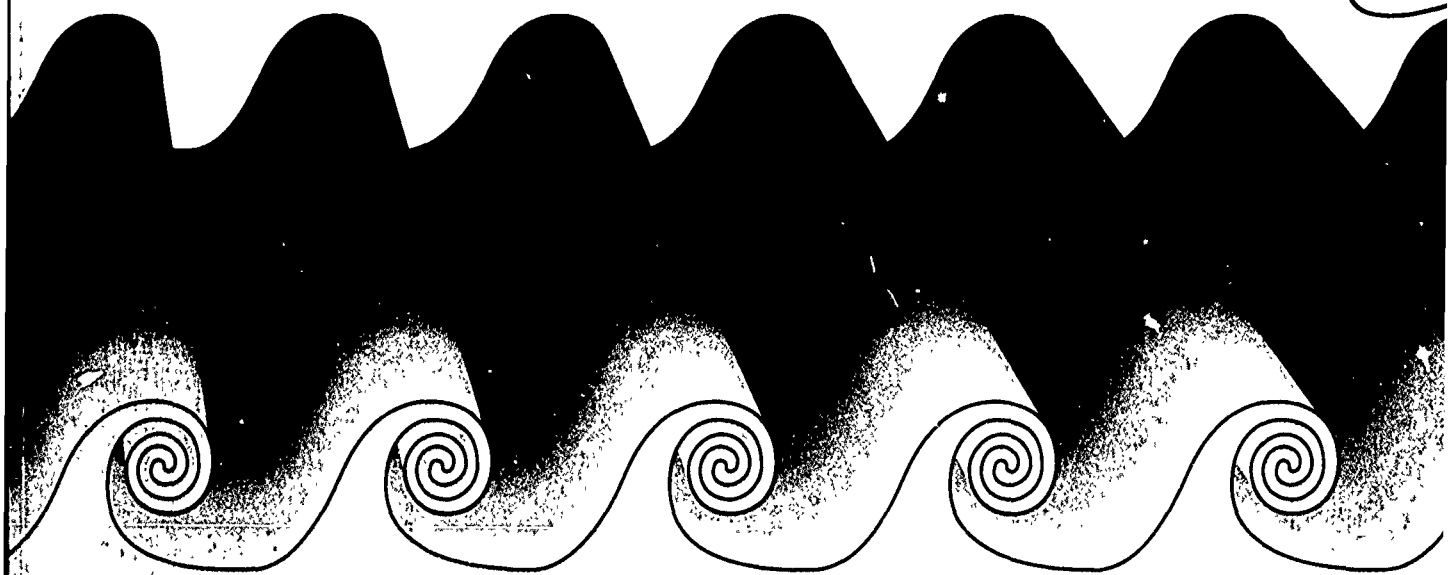
AD-A245 650

'Aha Huliko'a



DYNAMICS OF OCEANIC INTERNAL GRAVITY WAVES

DTIC
ELECTE
FEB 0 4 1992
S D
D

92-02786

This document has been approved for public release and sale; its distribution is unlimited.

Proceedings
Hawaiian Winter Workshop
University of Hawaii at Manoa
January 13 - 18, 1991

'Aha Huliko'a



DYNAMICS OF OCEANIC INTERNAL GRAVITY WAVES

PROCEEDINGS
'Aha Huliko'a
Hawaiian Winter Workshop
University of Hawaii at Manoa
January 15 - 18, 1991

PETER MÜLLER
DIANE HENDERSON
editors

Sponsored by the U.S. Office of Naval Research,
the School of Ocean and Earth Science and Technology,
and the Department of Oceanography, University of Hawaii

Cover, title page by Brooks Bays, SOEST Publication Services

FOREWORD

The sixth 'Aha Huliko'a† Hawaiian Winter Workshop was held from January 15 to 18, 1991 at the East-West Center in Honolulu, Hawaii. The topic was the "Dynamics of oceanic internal gravity waves."

The workshop marked a deep penetration into the subject of internal waves and mixing. Progress in theory and observations have advanced the field to the point where it becomes feasible to construct a global model to predict the internal-wave field and diapycnal mixing. The participants were tasked to assess the physical basis for such a model. Their lectures are published in these proceedings as submitted in camera ready form by the authors. The order of the papers loosely follows the agenda of the workshop, covering observations of internal-wave induced mixing, the theory and modeling of specific dynamical processes, and a discussion of major issues such as tidal versus wind forcing or interior versus boundary mixing. Also included is a summary of the meeting, which appeared in Eos, Transactions of the American Geophysical Union.

The workshop, made possible by a grant from the U.S. Office of Naval Research, was hosted by the Department of Oceanography of the newly established School of Ocean and Earth Science and Technology of the University of Hawaii. The excellent facilities of the East-West Center and the capable staff directed by James McMahon contributed greatly to the success of the meeting. The local organization and logistical arrangements were expertly and cheerfully handled by Crystal Miles. This proceedings volume came into existence through the creative and dedicated research of the scientists who gathered in Hawaii and provided the articles that follow and with the help of Brooks Bays, May Izumi, and Paula Yoshioka.

Peter Müller
Diane Henderson

Department of Oceanography
School of Ocean and Earth Science and Technology
1000 Pope Road
University of Hawaii
Honolulu, HI 96822

†'Aha Huliko'a is a Hawaiian phrase meaning an assembly that seeks into the depth of a matter.

DTIC COPY INSPECTED

Accesion For
NTIS CRA&I
DTIC TAB
U ... ...ed
J ...
By
D ...

This work is related to Department of the Navy Grant N-00014-90-J-4010 issued by the Office of Naval Research. The U.S. Government has a royalty-free license throughout the world in all copyrightable material herein.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION School of Ocean and Earth Science and Technology 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code) Department of Oceanography, Univ. of Hawaii 1000 Pope Road Honolulu, HI 96822	7b. ADDRESS (City, State, and ZIP Code) Department of the Navy 800 N. Quincy Street Arlington, Virginia 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-90-J-4010
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. 4222922-02 / TASK NO / WORK UNIT ACCESSION N

11. TITLE (Include Security Classification)
Dynamics of Oceanic Internal Gravity Waves

12. PERSONAL AUTHOR(S)
Müller, Peter and Henderson, Diane (eds.)

13a. TYPE OF REPORT	13b TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Workshop Proceedings	FROM 7/1/90 TO 12/31/9	November 1991	520

16. SUPPLEMENTARY NOTATION
Proceedings, 'Aha Huliko'a, Hawaiian Winter Workshop, January 1991, Honolulu, Hawaii

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
internal gravity waves, inertial oscillations, internal tides, finestructure, vortical motion, nonlinear interaction, mixing

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

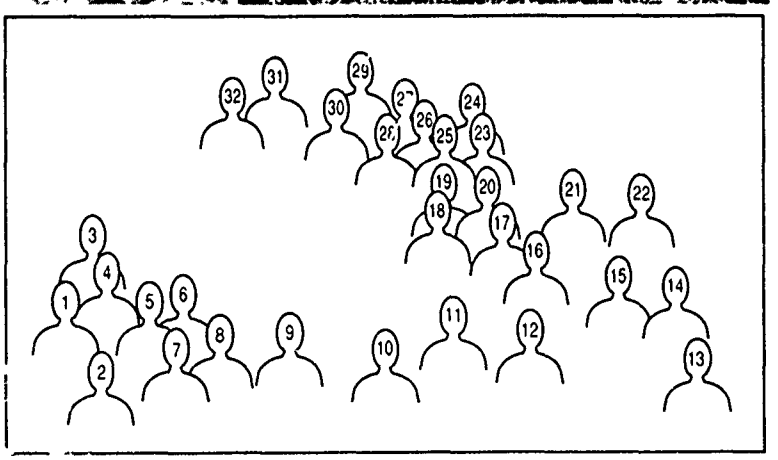
These proceedings contain the lectures given at the sixth 'Aha Huliko'a Hawaiian Winter Workshop on "Dynamics of oceanic internal gravity waves" and a summary of the workshop, which appeared in Eos, Transactions, American Geophysical Union. The lectures and the summary cover the major aspects of internal wave dynamics and internal wave induced mixing.

92 2 07 027

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT XX UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
SECURITY CLASSIFICATION OF THIS PAGE





PARTICIPANTS

1. Peter Müller
2. Crystal Miles
3. Dave Ramsden
4. Chris Garrett
5. Alex Warn-Varnas
6. Ken Watson
7. Mel Briscoe
8. Greg Holloway
9. Mike Gregg
10. Tom Dillon
11. Trevor McDougall
12. Henry Abarbanel
13. Rob Pinkel
14. Dave Fritts
15. Colin Hines
16. Tom Sanford
17. Murray Levine
18. Terry Ewart
19. Frank Henyey
20. Eric Kunze
21. Dave Rubenstein
22. Ren-Chieh Lien
23. Kraig Winters
24. Steve Anderson
25. Jim Riley
26. Laurie Padman
27. Gary Geernaert
28. Dave Siegel
29. Eric D'Asaro
30. Eric Hirst
31. Denis Gilbert
32. Claude Frankignoul

Charlie Eriksen

Naihuai Xu

Table of Contents

Foreword iii

Participants v

Internal Wave Shear and Dissipation
Michael C. Gregg, D.P. Winkel, and Thomas B. Sanford 1

Internal Wave Dissipation in a Non-Garrett-Munk Ocean
L. Padman, T.M. Dillon, H.W. Wijesekera, Murray D. Levine, C.A. Paulson, and Robert Pinkel 31

Testing the Critical Reflection Hypothesis
Denis Gilbert 53

Observations of Near-Inertial Internal Waves and Mixing in the Seasonal Thermocline
Charles C. Erikson 71

On the Statistics of Fine Scale Strain in the Thermocline
Robert Pinkel and Steven Anderson 89

Spatial Structure of Thermocline and Abyssal Internal Waves
Thomas B. Sanford 109

Estimates of Small-Scale Horizontal Divergence and Relative Vorticity in the Ocean
Ren-Chieh Lien and Peter Müller 143

Measurements of Ertel Vorticity Finestructure in the Eastern North Atlantic
Eric Kunze and Thomas B. Sanford 157

Symmetry Preserving Mode Truncations of Inviscid Geophysical Fluid Dynamical Equations
Henry D.I. Abarbanel and Ali Rouhi 179

Weak Wave and Vortex Interactions
James J. Riley, Marie-Pascale G. Lelong, and Donald N. Slinn 193

Internal Wave-Wave Resonance Theory: Fundamentals and Limitations
Eric Hirst 211

Choosing Variables for Internal Wave Dynamics
Frank S. Henyey 227

Is Scattering or Reflection More Effective in Causing Boundary Mixing?
Naihuai Xu and Peter Müller 237

On the Exchange of Energy Between Surface and Internal Wave Fields
Kenneth M. Watson 251

The Saturation of Middle-Atmosphere Gravity Waves
Colin O. Hines 261

Diagnosing Diapycnal Mixing
Kraig B. Winters and Eric A. D'Asaro 279

Energy Transfers Across an Internal Wave/Vortical Mode Spectrum
Dave Ramsden and Greg Holloway 295

Large-Eddy Simulation of Internal Wave Motions
David A. Siegel 315

Numerical Modeling of the Large-Scale Dynamics of Internal Waves
David Rubenstein 341

Parameterizing Mixing in Inverse Models
Trevor McDougall 355

Oceanic Model Testing
Claude Frankignoul 387

Acoustic Implications of a New Model
Terry Ewart and Stephen Reynolds 399

Inferences of Gravity Wave Processes From Atmospheric Spectra
David C. Fritts 417

Paradigm Lost?
Chris Garrett 433

A Strategy for Investigating and Modeling Internal Wave Sources and Sinks
Eric A. D'Asaro 451

Observing Oceanic Internal Waves: What have we learned? What can we learn?
Murray D. Levine 467

Overview of Ocean Background Conditions and Modeling of Internal Waves
Alex Warn-Varnas 481

Internal Gravity Waves and Mixing
Peter Müller, Eric A. D'Asaro, and Greg Holloway 499

INTERNAL WAVE SHEAR AND DISSIPATION

M. C. Gregg, D. P. Winkel, and T. B. Sanford

Applied Physics Laboratory and School of Oceanography, College of Ocean and Fishery Sciences, University of Washington, Seattle, Washington, 98105-6698

ABSTRACT

With the Multi-Scale Profiler (MSP), we are able to obtain shear spectra extending from vertical scales of hundreds of meters to the viscous cutoff of small-scale turbulence. Comparing shear spectra from five sites reveals varied spectral shapes and amplitudes over 0.01–0.1 cpm (cycles per meter). The amplitudes converge, however, near 0.14 cpm, just beyond the start of the rolloff of the internal wave range of the spectrum. Cross-spectra of the u (east) and v (north) velocity components reveal significant coherence squared, phases of $\approx \pm 90°$, and corresponding asymmetries in clockwise and anticlockwise rotary spectra at the beginning of the rolloff in most spectra. We interpret these signatures of near-inertial motions as evidence that the rolloff is caused by critical layer interactions of waves having vertical scales of about 10 m being advected by larger-scale waves. Similar signatures occur irregularly throughout the rolloff range, usually 0.1–1 cpm, consistent with critical-layer interactions continuing with increasing wavenumber until the waves break down into turbulence. For waves close to the Garrett and Munk spectrum (GM76), this breakdown occurs at wavenumbers larger than 1 cpm. Owing to the variety of spectral shapes at low wavenumbers, variances of kinetic energy and shear do not change in the same proportion when spectra depart from GM76. Spectra from low latitudes in the central Pacific fail to exhibit the inverse dependence on the Coriolis parameter, f, predicted by Munk (1981). Their shape also differs substantially from GM76. Some have lower kinetic energies than mid-latitude spectra but retain comparable shear variances. In addition, all of the low-latitude spectra roll off much more steeply than do the mid-latitude spectra. The steep rolloff forms a weak spectral gap separating internal waves and turbulence.

INTRODUCTION

For several decades, most turbulence in the ocean's interior has been attributed to breaking internal waves, even though the mechanisms of breaking remain obscure. Possibilities include shear instability, advective overturning, strain, and critical layers formed by large waves advecting smaller waves. The probability of breaking is estimated using the Garrett and Munk (1975) model spectrum, $\Phi(\vec{k}, \omega)$, of the wavenumber and frequency content of the internal wave field. Aside from under the arctic ice cap (Levine et al., 1985), observations find internal wave intensities at or above GM, but rarely below, e.g., Smart (1988). Consequently, GM appears to describe the background state, with average forcing in equilibrium with average dissipation by breaking.

We are attempting to observe both equilibrium and nonequilibrium internal wave states and to determine how vertical wavenumber spectra depart from GM, what clues these departures offer about internal wave dynamics, and how they alter $\langle\varepsilon\rangle$, the average turbulent dissipation rate. Here we examine vertical shear spectra obtained with the Multi-Scale Profiler (MSP). The spectra extend from 0.01 cpm (cycles per meter) to the viscous cutoff, usually near 10 cpm. They were obtained at five sites, four of which were used by Gregg (1989) in comparing moments of 10-m-shear with $\langle\varepsilon\rangle$.

In examining the shear spectra, we ask

- How do spectral shapes at low wavenumbers vary with changes in spectral amplitude? Are the variations systematic, and do kinetic energy and shear variances change proportionately?
- Do the spectra change systematically with latitude?
- How do spectra from the thermohaline staircase observed during CSALT compare with spectra at a similar latitude in a normal profile?
- Do the spectra contain signatures of critical layer interactions, proposed by Holloway (1980) as the mechanism causing the rolloff?
- Does the wavenumber at which internal waves roll off vary with spectral amplitude?
- Is Nasmyth's universal spectrum for turbulence in stratified fluids an adequate description of the dissipation range?

BACKGROUND

Spectra of vertical strain (Gregg et al., 1973; Gregg, 1977a) and of vertical shear (Gargett et al., 1981; Sherman and Pinkel, 1991) change from being nearly flat, k_3^0, at low wavenumbers to approximately k_3^{-1} at wavenumbers higher than about 0.1 cpm. We define k_{k-1} as the wavenumber where the rolloff occurs. Munk (1981) incorporates the rolloff into the GM model, but also inadvertently changes the relationship between kinetic energy and shear variance. Consequently, we use the intermediate model, known as GM76, with a modification for the rolloff, as given by Gregg and Kunze (1991).

Holloway (1980) points out that internal waves with short vertical wavelengths are strongly advected by larger, and hence faster, internal waves. Away from inertial and buoyancy frequencies, i.e., $f \ll \omega \ll N$, the horizontal phase speed of internal waves is

$$c_h \approx \frac{N}{\beta} \quad [\mathrm{m\ s^{-1}}] \tag{1}$$

where β is the vertical wavenumber in radians per meter. For $\beta = 0.063$ m^{-1} ($k_3 = 0.01$ cpm), $c_h \approx 0.09$ m s^{-1}, compared with an rms velocity for the entire wave field of $u_{\mathrm{rms}} = 0.07$ m s^{-1}, obtained by integrating GM76 for $N = 0.0056$ s^{-1} (3 cycles per

hour). With increasing wavenumber, $c_h \rightarrow u_{\text{rms}}$, equaling it at $\beta = 0.08$ m^{-1} ($k_3 =$ 0.013 cpm). Holloway convincingly argues that wave/wave interactions cannot be weak for $c_h \leq u_{\text{rms}}$. Why then does the shear spectrum remain flat to 0.1 cpm instead of rolling off at $k_3 = 0.01$–0.02 cpm?

In discussing the rolloff, Munk (1981) cites Ericksen (1978), who computes the gradient Richardson number, $Ri \equiv N^2/(\Delta U/\Delta z)^2$, from moored sensors having $\Delta z = 7$ m. A scatter plot shows few values smaller than 1/4, the necessary condition for shear instability. Defining the Richardson function

$$Ri(k_3) \equiv \frac{\langle N^2 \rangle}{\int_0^{k_3} \Phi_{\text{SHEAR}}(\xi)\, d\xi} \tag{2}$$

Munk concludes that the internal wave field is characterized by $Ri(k_{k-1}) = \text{O}(1)$ and argues that the rolloff should shift with changes in energy as

$$E \times k_{k-1} = \text{constant} \tag{3}$$

where E is the dimensionless energy density in GM76. This criterion should apply for shear instability, advective overturning, and instability due to lateral strain. Sherman and Pinkel (1991) point out that when $E = E_{\text{GM76}}$, $Ri(0.1) \approx 0.5$ and $Ri(k_3) = 1$ occurs well into the rolloff.

Gregg (1977b) interprets the rolloff region of scalar spectra as composed of decaying internal waves and irreversible density finestructure, i.e., density structure produced by mixing. Gargett et al. (1981) interpret the rolloff as the buoyancy subrange of turbulence and take its upper bound as the buoyancy wavenumber

$$k_{\text{B}} \equiv \left(\frac{\langle N^2 \rangle^{3/2}}{\langle \varepsilon \rangle} \right)^{1/2} \quad \left[\text{m}^{-1} \right] \tag{4}$$

They identify the corresponding buoyancy length scale, $L_{\text{B}} = 2\pi k_{\text{B}}^{-1}$, as the vertical scale of the largest overturns. To explore whether the rolloff results from buoyancy-modified turbulence, they assume that the turbulence is controlled only by $\langle N^2 \rangle$ and $\langle \varepsilon \rangle$ and nondimensionalize their spectra using k_{B} and

$$\Phi_{\text{B}} \equiv \frac{\langle N^2 \rangle}{k_{\text{B}}} \quad \left[\frac{\text{s}^{-2}}{\text{m}^{-1}} \right] \tag{5}$$

In the rolloff region, the scaling reduces amplitude variations among their spectra from a factor of 4 to a factor of 2. The slope of the rolloff is similar to predictions by some theories of the buoyancy subrange (Shur, 1962; Lumley, 1964), but Gargett et al. cannot determine whether the agreement is fortuitous or results from the rolloff actually being a buoyancy subrange.

We believe that the agreement is fortuitous because Gargett et al. are inconsistent in taking L_{B} as the scale of the largest overturns and in identifying the rolloff region with the

buoyancy subrange. Their identification of L_B with the largest overturns has been confirmed by Dillon (1982) and subsequent investigators, who find that root-mean-square (rms) overturning scales are about $1.25(\langle\varepsilon\rangle / \langle N^2\rangle^{3/2})^{1/2}$ or $0.2L_B$ in the notation of Gargett et al. (1981). (Gargett et al. and Dillon differ by 2π in defining the buoyancy scale.) Because the largest overturns must be several times larger than Dillon's rms scale, L_B is a good upper bound. This, however, means that the rolloff region cannot contain three-dimensional turbulence, which is assumed by Shur (1962) and Lumley (1964) when they consider how stratification flattens the overturns to produce anisotropic structures.

In the atmosphere, internal wave shear spectra also roll off as k_3^{-1} at high wavenumbers, but at low wavenumbers the spectra often rise with increasing wavenumber (Smith et al., 1987). This shape so concentrates velocity and shear variance near the rolloff that the waves are sometimes modeled as a single Fourier component. Stratospheric waves originate in the troposphere and grow rapidly with altitude in response to the exponential decrease in air density. As a consequence, the rolloff shifts to lower wavenumbers with increasing altitude, and the shift is attributed to saturation of the wave field. Fritts (1984, 1989) reviews the extensive literature, which includes much debate about the mechanisms saturating the wave field. Dewan and Good (1986) argue that both shear instability and convective overturning produce spectra of the form

$$\Phi_{SAT}(k_3) = \frac{b\langle N^2\rangle}{\beta} \quad \left[\frac{s^{-2}}{m^{-1}}\right] \tag{6}$$

where b is a constant. They obtain $b = 4$ for shear instability when each wavenumber component saturates individually and $b = 1$ for convective instability, leading to the conclusion that waves reach convective instability before shear instability. Dewan and Good also note that turbulent layers can be much thinner than the scale at which the spectrum saturates. The spectrum rolls off because waves at that scale lose energy. However, the turbulent overturns extracting the energy may be only a few percent of the scale of the wave. Smith et al. (1987) conclude that the instability criteria used by Dewan and Good are too large and obtain $b = 1/2$ for a single Fourier component and $b = 1/6$ for a spectrum of superposed waves.

Hines (1991a) holds that the saturated spectrum in the atmosphere results from the same mechanism as proposed by Holloway for the ocean, i.e., by Doppler shifting and critical layer interactions, rather than by shear instability or convective overturning. Furthermore, he develops a spectrum for this mechanism that asymptotically approaches (6) (Hines, 1991b).

Discussions of internal waves in the ocean have concentrated on the extensive observations in the upper ocean at mid-latitude. The limited observations available from the deep ocean (Sanford, 1991) and from low latitudes (Wunsch and Webb, 1979; Eriksen, 1980) appear different. To incorporate the low-latitude observations, Munk (1981) modifies GM76 by replacing the dimensionless energy density, E, with E'/f, where $E' = 8.8 \times 10^{-7}$ and f is the Coriolis parameter. This changes the model from constant energy in the wave field to constant spectral density, thereby accommodating Wunsch and

Webb's report that " ...on the equator the frequency spectrum does not change shape but the vertical and horizontal coherences are much reduced compared to midlatitude values, suggestive of higher modes." With this change, the shear spectrum becomes

$$\Phi_{\mathrm{SHEAR}}(k_3) = \frac{3E'b^3N_0^2}{2j_*\pi f}\frac{\beta^2}{(1+\beta/\beta_*)^2} \quad \left[\frac{\mathrm{s}^{-2}}{\mathrm{m}^{-1}}\right] \tag{7}$$

In their theoretical models of wave/wave interactions, McComas and Müller (1981) and Henyey et al. (1986) give the energy flux through the spectrum as proportional to f. Consequently, $\langle\varepsilon\rangle$ is also proportional to f. Both Müller and Henyey (personal communications, 1990) allow that the f scaling is dictated mostly by the need for correct dimensions rather than by serious consideration of the effect of f on wave/wave interactions. Nevertheless, both predict that for constant shear variance $\langle\varepsilon\rangle \to 0$ approaching the equator. Coupled with Munk's (1981) prediction (7), one would expect to find the anomalous situation of increasing shear variances and decreasing dissipation rates.

OBSERVATIONS AND DATA ANALYSIS

Encompassing a six-fold variation in f, the observations come from one cruise in the Atlantic and two in the Pacific (Table 1). The Atlantic data were taken in October 1985 during CSALT and span the thermohaline staircase lying east of Barbados at pressures of 3–6.5 MPa (Gregg and Sanford, 1987). Retaining much of the steppiness in the density field, the CSALT velocity profiles are not primarily signatures of internal waves but are included for comparison with the other profiles, which are dominated by internal waves in regions with little or no mean shear (Fig. 1). PATCHEX was conducted a year later on the western side of the California Current, and the profiles appear typical of the open ocean, except for ubiquitous salt-stabilized temperature inversions (Gregg and Sanford, 1988). After the primary PATCHEX observations, we worked for two days in a coastal jet off Crescent City, California, obtaining the profiles referred to as PATCHEX north. Both sets of PATCHEX profiles extend to 10 MPa. The following spring, while transiting to and from the equator for TROPIC HEAT 2, we took three test drops to 5 MPa at 6°N and five drops to 10 MPa at 11°N.

Table 1: MSP data sets.

Cruise	Dates	Position	$10^5 f/\mathrm{s}^{-1}$	# Profiles	$p_{\max}$/MPa
CSALT	Nov 23–24, 1985	12°N, 56.5°W	3.02	5	7
PATCHEX	Oct 17–24, 1986	34°N, 127°W	8.13	28	10
PATCHEX north	Oct 26–27, 1986	42°N, 126°W	9.73	5	10
TROPIC HEAT 2	Mar 28–29, 1987	6.0°N, 143.5°W	1.52	3	5
TROPIC HEAT 2	Apr 18–19, 1987	11.5°N, 134.8°W	2.90	5	10

Because GM76 describes the average behavior of random wave fields, we tested the PATCHEX and PATCHEX north shear profiles for randomness by first-differencing u (east) and v (north) with $\Delta z = 10$ m. After WKB scaling, $\Delta u/\Delta z$ and $\Delta v/\Delta z$ have

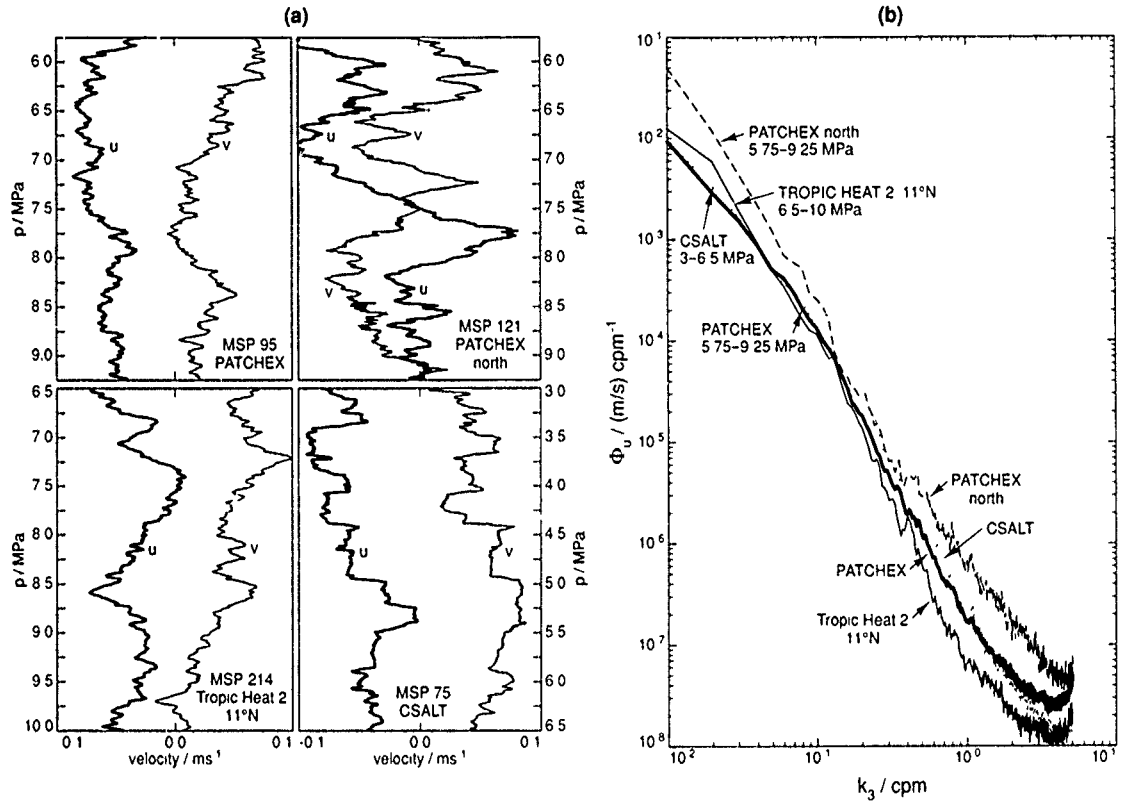


Figure 1: Typical velocity profiles (a) and average velocity spectra (b). PATCHEX north has the largest velocity fluctuations and the highest spectral levels at low and at high wavenumbers. The CSALT profiles show some of the steppiness of the density field, but their average spectrum is similar to that of the PATCHEX profiles. The spectrum for TROPIC HEAT 2 at 11°N is dominated by noise for $k_3 > 3$ cpm. All spectra are slightly aliased just before the Nyquist wavenumber at 5 cpm.

normal probability densities with zero means and are independent of each other (Gregg et al., 1991). In addition, we used the run test (Bendat and Piersol, 1971) and the cumulative periodogram (Jenkins and Watts, 1969) to determine that these shear profiles have no significant departures from vertical randomness. These data sets, therefore, satisfy the criteria for random internal wave fields assumed by Garrett and Munk (1975). The other data sets are too small for adequate testing, but we find no visual suggestions of deterministic features.

To avoid the near-surface internal wave duct, we analyze only data below the shallow thermocline and, when the records are sufficiently long, take two sets of spectra, referred to as shallow and deep. Depending on the record, the sections are 2.5–3.5 MPa thick. Considering all sections, the average buoyancy frequency has a small range, $\langle N^2 \rangle^{1/2} = 0.0028$–$0.0042$ s^{-1}.

The MSP senses velocity with two pairs of electrodes, two airfoil probes, and a two-axis N. Brown acoustic current meter (ACM). The acoustic channels are recorded at 62.5 Hz, corresponding to $\Delta z = 4$ mm at typical fall rates of 0.25 m s^{-1}. To convert the

acoustic data to u and v, we adapt previous models (Evans et al., 1979; Hayes et al., 1984) for the MSP configuration. Corrected velocities are low-passed with a 2-kPa Bartlett filter and digitized at 1 kPa. At low wavenumbers, these data are compared with velocities derived from the motionally induced electromagnetic signals, which are much less sensitive to vehicle motion. At high wavenumbers, we compute velocity from the airfoil probes, which sense fluctuations for 1–2 cpm $< k_3 \leq 100$ cpm.

Data are conditioned for spectral analysis by subtracting the mean, first-differencing, and applying the Hann filter. For the ACM records, u and v are put in real and complex arrays, and Singleton's (1969) algorithm is used to take Fourier transforms of pieces containing 1000 points (corresponding to 1 MPa). This yields two autospectra, $\Phi_u(k_3)$ and $\Phi_v(k_3)$, and the co- and quadspectra, $\Phi^c_{uv}(k_3)$ and $\Phi^q_{uv}(k_3)$, all of which are corrected for first-differencing, for the loss in variance to the Hann window, and for the Bartlett window. Successive pieces are overlapped by 50%, resulting in 4–6 pieces for sections of 2.5–3.5 MPa. Wavenumbers are given in cycles per meter as $k_3 = (0.01/\Delta p)$ cpm, where Δp is the pressure interval in megapascals.

For one profile, the degrees of freedom per spectral estimate is

$$\nu_1 = 2L^2/(1.056L - 0.056) \tag{8}$$

where L is the number of pieces overlapped by 50% (D. Percival, personal communication, 1991). After averaging R records, the total number of degrees of freedom is $\nu = R \times \nu_1$. With $L = 6$ and $R = 28$, PATCHEX has $\nu = 321$. At the other extreme, TROPIC HEAT 2 at 6°N has $L = 4$ and $R = 3$, yielding $\nu = 23$. Defining $\hat{\Phi}(k_3)$ as a spectral estimate and $\Phi(k_3)$ as the true spectrum, the 95% confidence limits are

$$\left[\frac{\nu\hat{\Phi}(k_3)}{\chi^2_{\nu;0.025}} \leq \Phi(k_3) \leq \frac{\nu\hat{\Phi}(k_3)}{\chi^2_{\nu;0.975}}\right] \tag{9}$$

where $\chi^2_{\nu;0.025}$ is the chi-square distribution evaluated for ν degrees of freedom and the 0.025 percentile. These limits are plotted as shading around the spectra.

Adding the two autospectra forms the spectrum of total velocity,

$$\Phi_{\mathrm{VEL}}(k_3) = \Phi_u(k_3) + \Phi_v(k_3) \quad [(\mathrm{m/s})^2\ \mathrm{cpm}^{-1}] \tag{10}$$

which is twice the spectrum of specific kinetic energy. Near the Nyquist wavenumber, 5 cpm, all spectra are slightly aliased, in spite of the Bartlett window. Some have negligible noise, e.g., PATCHEX north; others, e.g., TROPIC HEAT 2 at 11°N, become noisy near 3 cpm. Taking the noise level as 1.8×10^{-8} m^2 s^{-1} and the noise bandwidth as 2 cpm gives an rms noise of 0.16 mm s^{-1}.

The velocity spectra are used to compute the shear spectra

$$\Phi_{\mathrm{SHEAR}}(k_3) = (2\pi k_3)^2\ \Phi_{\mathrm{VEL}}(k_3) \quad [\mathrm{s}^{-2}\mathrm{cpm}^{-1}] \tag{11}$$

which are cut off before 2 cpm to avoid noise contamination (Fig. 2).

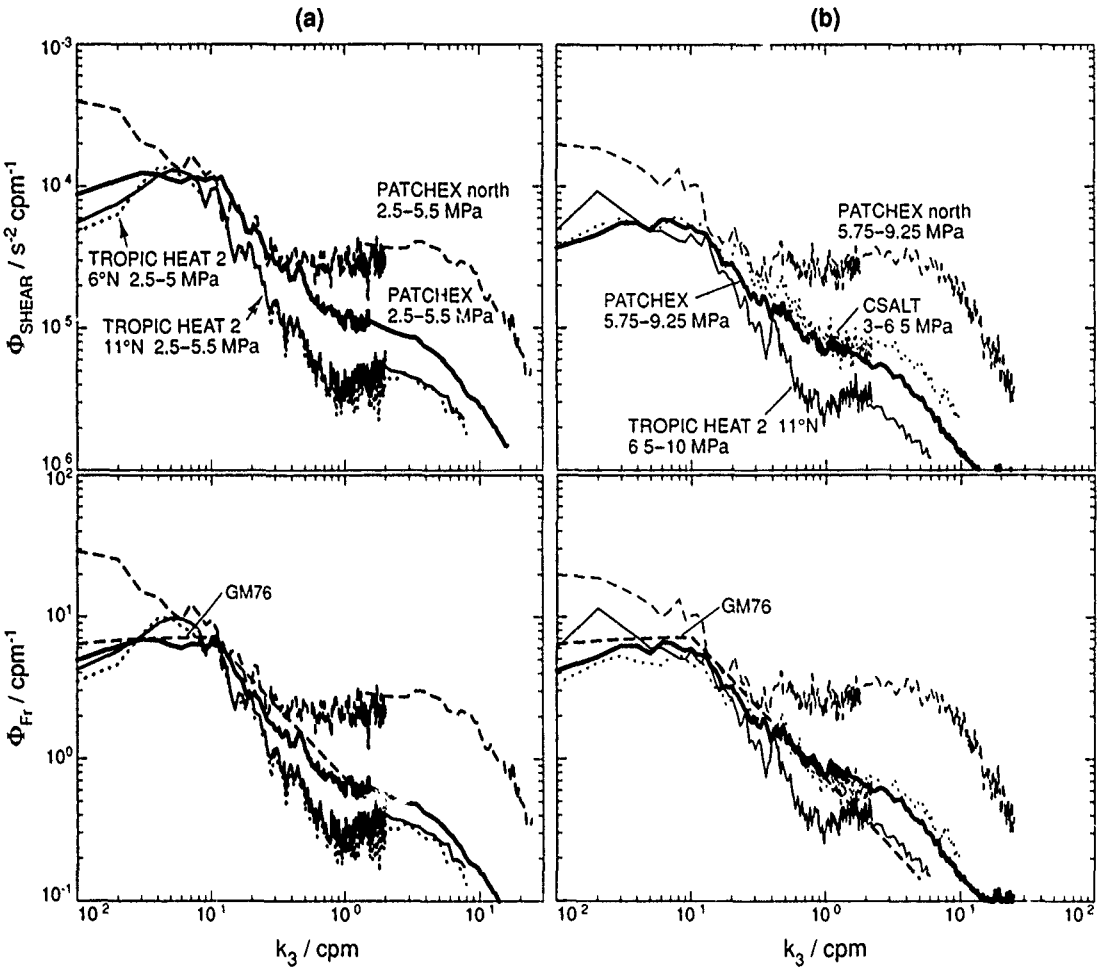


Figure 2: Average shear and Froude spectra for the shallow (a) and deep (b) sections. GM76 is shown on the Froude spectra as a heavy dashed line. Although the Froude spectra differ by a factor of 5 at 0.01 cpm and by 10–20 at high wavenumbers, they have nearly the same amplitude at 0.14 cpm.

To compare shear spectra from records having different stratification, we present them as Froude spectra

$$\Phi_{\mathrm{Fr}}(k_3) \equiv \left\langle N^2 \right\rangle^{-1} \Phi_{\mathrm{SHEAR}}(k_3) \quad [\mathrm{cpm}^{-1}] \tag{12}$$

For reference, we use the Froude spectrum for GM76,

$$\Phi_{\mathrm{Fr}}^{\mathrm{GM76}}(k_3) = \frac{3Eb^3N_0^2}{2j_*\pi} \frac{\beta^2}{(1+\beta/\beta_*)^2} \quad \left[\frac{1}{\mathrm{m}^{-1}}\right] \tag{13}$$

and the saturated atmospheric spectrum with $b = 1/2$,

$$\Phi_{\mathrm{Fr}}^{\mathrm{SAT}}(k_3) = \frac{1}{2\beta} \quad \left[\frac{1}{\mathrm{m}^{-1}}\right] \tag{14}$$

Co-spectra and quadspectra are used for coherence-squared

$$coh^2(k_3) \equiv \frac{\Phi^c_{uv}(k_3)^2 + \Phi^q_{uv}(k_3)^2}{\Phi_u(k_3)\,\Phi_v(k_3)} \tag{15}$$

and phase

$$\phi(k_3) \equiv -\frac{180^\circ}{\pi} \arctan\left(\frac{-\Phi^q_{uv}(k_3)}{\Phi^c_{uv}(k_3)}\right) \tag{16}$$

The sign convention gives $\phi > 0$ when v leads u with increasing pressure. Confidence limits (95%) are computed using Percival (1991).

Because the ACM data are limited to $k_3 \leq 3$–4 cpm, the vertical spectrum is completed using data from the airfoil probes. The two airfoil probes on the MSP are aligned with their sensitive axes in the same direction, which aids in removing spectra contaminated by plankton impacts. Consequently, we analyze the spectrum of one channel, compare with the spectrum of other channel to detect noise spikes, assume horizontal isotropy, and multiply by 2 to obtain $\Phi_{VEL}(k_3)$. For the deep data, spectra are taken over intervals of 0.07–0.08 MPa, again with 50% overlap. For the shallow data, we use a new robust procedure (Hess et al., 1991) and take spectra over 0.02 MPa intervals. Comparisons of the two procedures reveal no significant differences in average spectra. As the airfoils are recorded at 125 Hz, their spectra have many degrees of freedom, e.g., $\nu = 5,253$ for PATCHEX. Consequently, some 95% confidence limits are too narrow to be seen on plots. We form the composite spectra by overlaying spectra from the ACM and airfoil probes, without adjusting amplitudes. Variance-preserving plots show that the dissipation ranges are moderately well resolved, except for the deep spectrum at 11°N, which has the weakest signals (Fig. 3).

At high wavenumbers, the observed spectra are compared with the Nasmyth universal turbulence spectrum (Oakey, 1982). It depends only on the kinematic viscosity and $\langle\varepsilon\rangle$, which we obtain by averaging the ε's we compute routinely over 5 or 10 kPa intervals. Owing to its nonlinear dependence on $\langle\varepsilon\rangle$, the Nasmyth spectrum of $\langle\varepsilon\rangle$ does not equal the average of the individual Nasmyth spectra; numerical simulations show that the two average spectra can diverge significantly, depending on the distribution of ε (H. Seim, personal communication, 1991). Therefore, although the Nasmyth spectrum is a useful approximate reference, we should not expect exact agreement.

AMPLITUDES AND SHAPES

From Figure 2 and from the individual spectra in Figures 4–7, we draw the following conclusions:

- Deep and shallow spectra from the same data set are usually more alike than are spectra from different sets. The shallow spectra from 6°N and 11°N are the exception; they are nearly identical even though they were taken 1000 km and 3 weeks apart.

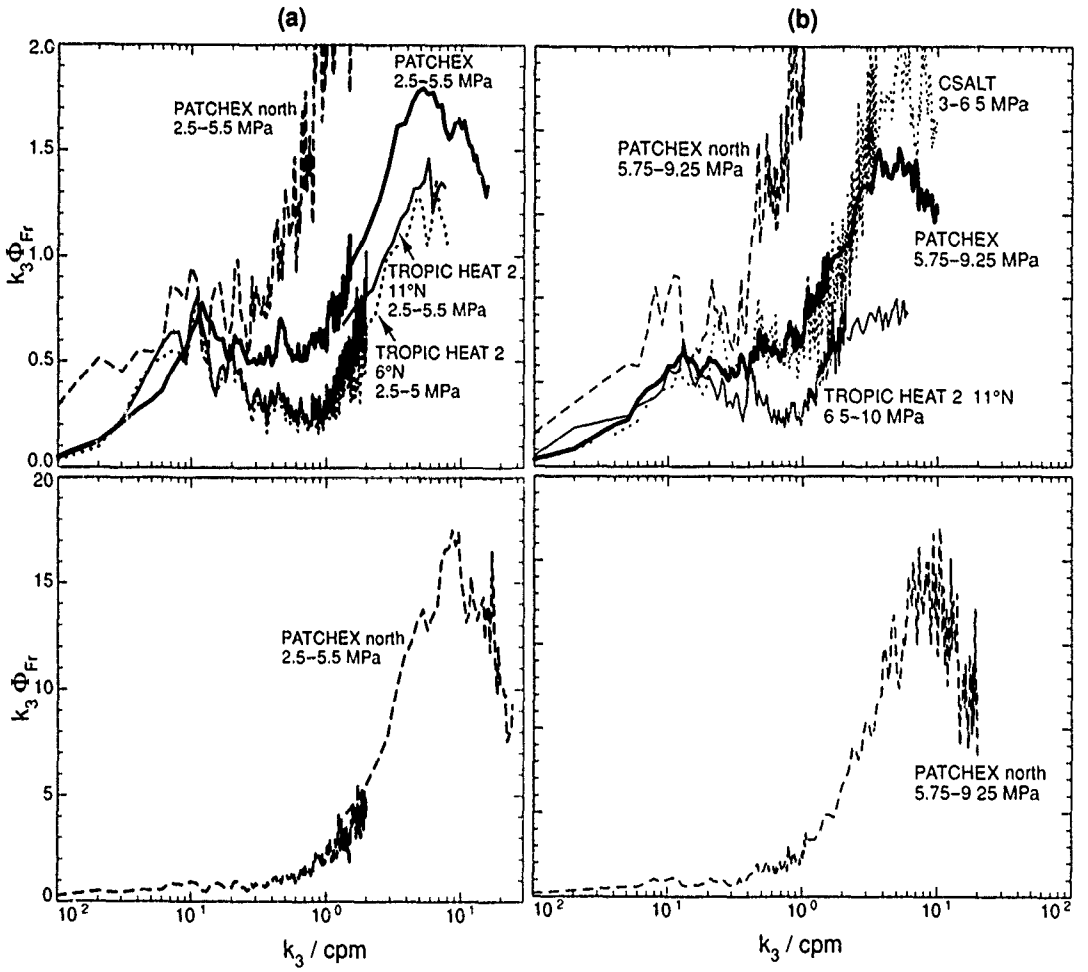


Figure 3: Variance-preserving plots of the spectra in Figure 2.

- All spectra have nearly the same amplitude at 0.14 cpm: $(0.65\text{–}1.0) \times \Phi_{\mathrm{Fr}}^{\mathrm{GM76}}(1.4)$. By contrast, their amplitudes vary by factors of 5 at 0.01 cpm and 20 in the dissipation range.

- Between 0.01 and 0.1 cpm the spectra exhibit little uniformity in amplitude or shape. Average slopes span -0.55 to +0.75, but in most cases, power laws fail to represent the spectra accurately. Adjustment of the slope to meet GM76 near 0.14 cpm is the only consistent pattern; spectra above GM76 at 0.01 cpm slope downward toward this common point, and those starting below GM76 slope upward.

- The k_3^{-1} rolloff for $k_3 > 0.1$ cpm is the most uniform feature for CSALT, PATCHEX, and PATCHEX north. The rolloff terminates at high wavenumbers where it intersects the Nasmyth spectrum.

- Froude spectra often have local maxima at or near $k_{k^{-1}}$. These are most evident in variance-preserving plots (Fig. 8).

- The three TROPIC HEAT 2 spectra peak near 0.08 cpm and roll off as $k_3^{-1.4}$ for $k_3 > 0.1$ cpm. The steep rolloff produces a weak spectral gap between internal waves and dissipation.
- Turbulence is so weak in most records that the dissipation range follows the Nasmyth spectrum only in the viscous cutoff. PATCHEX north is the exception, with a distinct inertial subrange, extending over half a decade in wavenumbers.

Because most of the spectra change slope near 0.1 cpm, we estimate average slopes using linear fits over 0.01–0.1 cpm and 0.1–1 cpm (Table 2), except where local maxima call for different limits. Most fits over 0.01–0.1 cpm are nominal, as few spectra are truly linear.

Table 2: Average slopes and standard deviations of shear spectra. Owing to particular features, some fits cover restricted intervals: (a) 0.1–0.4 cpm, (b) 0.1–0.37 cpm, (c) 0.01–0.045 cpm, (d) 0.01–0.06 cpm, (e) 0.02–0.1 cpm.

Cruise	p/MPa	0.01–0.1 cpm	0.1–1 cpm
PATCHEX	2.50–5.50	$+0.07 \pm 0.04$	-1.01 ± 0.02
PATCHEX	5.75–9.25	$+0.16 \pm 0.05$	-0.89 ± 0.02
PATCHEX north	2.50–5.50	-0.54 ± 0.04	-0.96 ± 0.08[a]
PATCHEX north	5.75–9.25	-0.33 ± 0.06	-1.03 ± 0.09[b]
TROPIC HEAT 2 6°N	2.50–5.00	$+0.75 \pm 0.13$[c]	-1.43 ± 0.03
TROPIC HEAT 2 11°N	2.50–5.50	$+0.50 \pm 0.04$[d]	-1.38 ± 0.03
TROPIC HEAT 2 11°N	6.50–10.0	-0.55 ± 0.04[e]	-1.37 ± 0.03
CSALT	3.00–6.50	$+0.15 \pm 0.05$	-0.89 ± 0.03

Closest to GM76 throughout the full range of wavenumbers (Fig. 4), the PATCHEX spectra are 0.8 and 0.6 times GM76 at 0.01 cpm. From there, they slope upward as $k_3^{+0.07}$ and $k_3^{+0.16}$. The shallow spectrum makes a sharp transition at the rolloff and initially decreases more steeply than k_3^{-1}, dropping close to the saturated spectrum from the atmosphere which has an amplitude about 50% smaller than GM76 when $b = 1/2$. The deep spectrum rolls off more gradually and stays closer to k_3^{-1}. Near 1 cpm, both PATCHEX spectra cease to roll off as steeply and lie close to the Nasmyth spectrum until 5–8 cpm, where they fail to roll off as steeply as the reference. Such weak turbulence is unlikely to produce an inertial subrange; the viscous rolloff occurs at 2–3 cpm, only an octave past the end of the internal wave part of the spectrum.

The PATCHEX north spectra are 3–4 times GM76 at 0.01 cpm and slope downward to the common point at 0.14 cpm (Fig. 5). By 0.02 cpm, the shallow spectrum is close to an extension of the k_3^{-1} portion of GM76, and it comes closer with increasing wavenumber. The deep spectrum is initially nearly flat and then curves downward to approach the k_3^{-1} extension near 0.08 cpm. For $k_3 > 0.1$ cpm, both PATCHEX north spectra are nearly identical to GM76 until 0.3–0.4 cpm, where they make sharp transitions to the $k_3^{1/3}$ inertial subrange of the turbulent spectrum. (Note that both the ACM and the

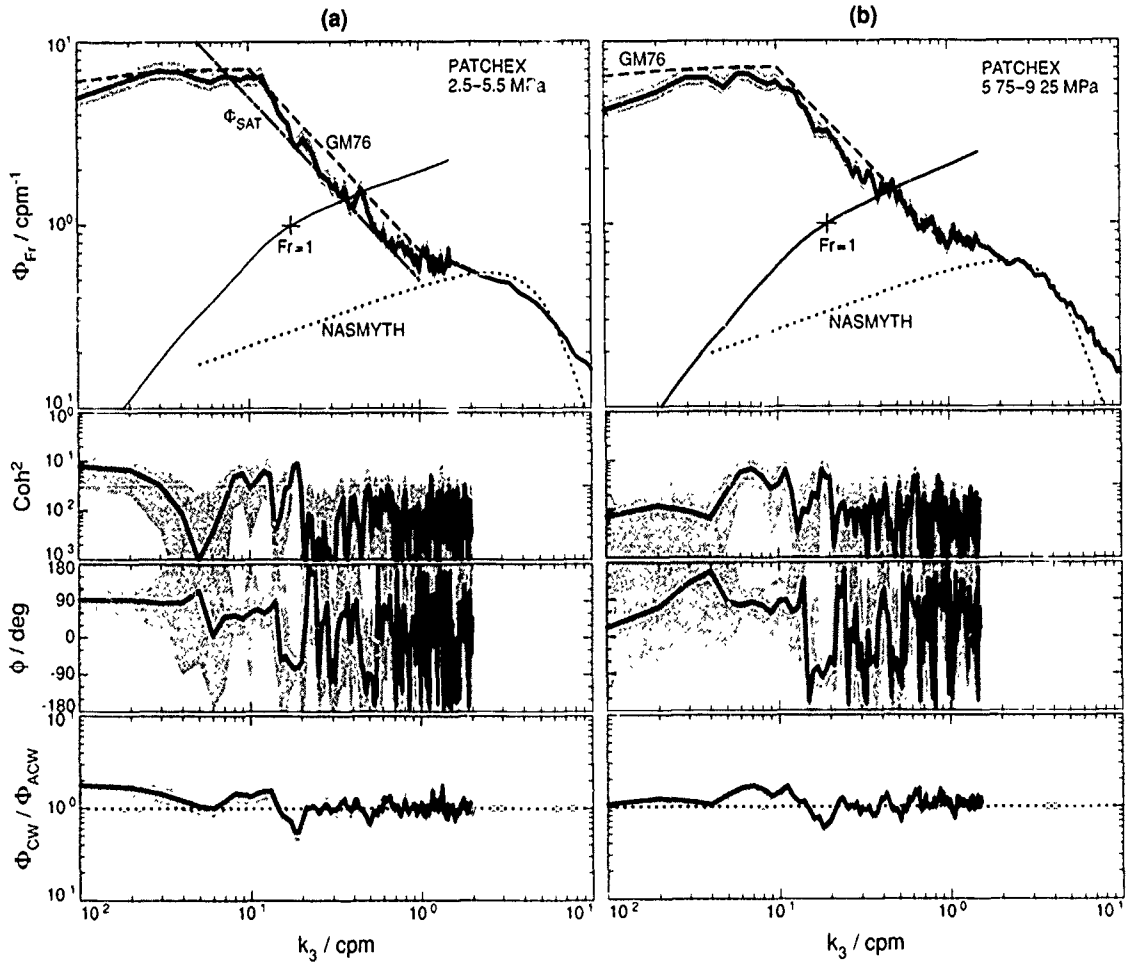


Figure 4: Froude spectra, coherence-squared, phase, and the ratio of clockwise (CW) to anticlockwise (ACW) velocities for PATCHEX. Shading shows 95% confidence limits. Also shown in the top panel are GM76, the saturated internal wave spectrum applied to atmospheric observations ($b = 1/2$), Nasmyth's universal turbulence spectrum in stratified fluids, and the Froude function. Data from the acoustic current meter is used for $k_3 < 1.5$ cpm.

airfoil probes detect the inertial subrange.) Beyond the viscous cutoff, these spectra continue to decrease (Fig. 3), but less steeply than the Nasmyth spectrum. The total range from where the spectrum initially follows $k_3^{+1/3}$ to the peak of the dissipation spectrum is a factor of 6–8, less than a decade. Either the close agreement with the inertial subrange is fortuitous, or it indicates that locally universal turbulence can occur at wavenumbers only 2–3 times larger than the energy-containing overturns.

The shallow spectra for TROPIC HEAT 2 are surprisingly similar, even to their small irregularities (Figs. 2 and 6). From amplitudes 0.6 times GM76 at 0.01 cpm, they slope upward to broad maxima near 0.08 cpm. Their most distinctive feature, however, is the steepness of the rolloff, $k_3^{-1.4}$. The combination of the rise to a peak near 0.08 cpm and the rapid rolloff at high wavenumbers resembles the narrow-band structures observed in

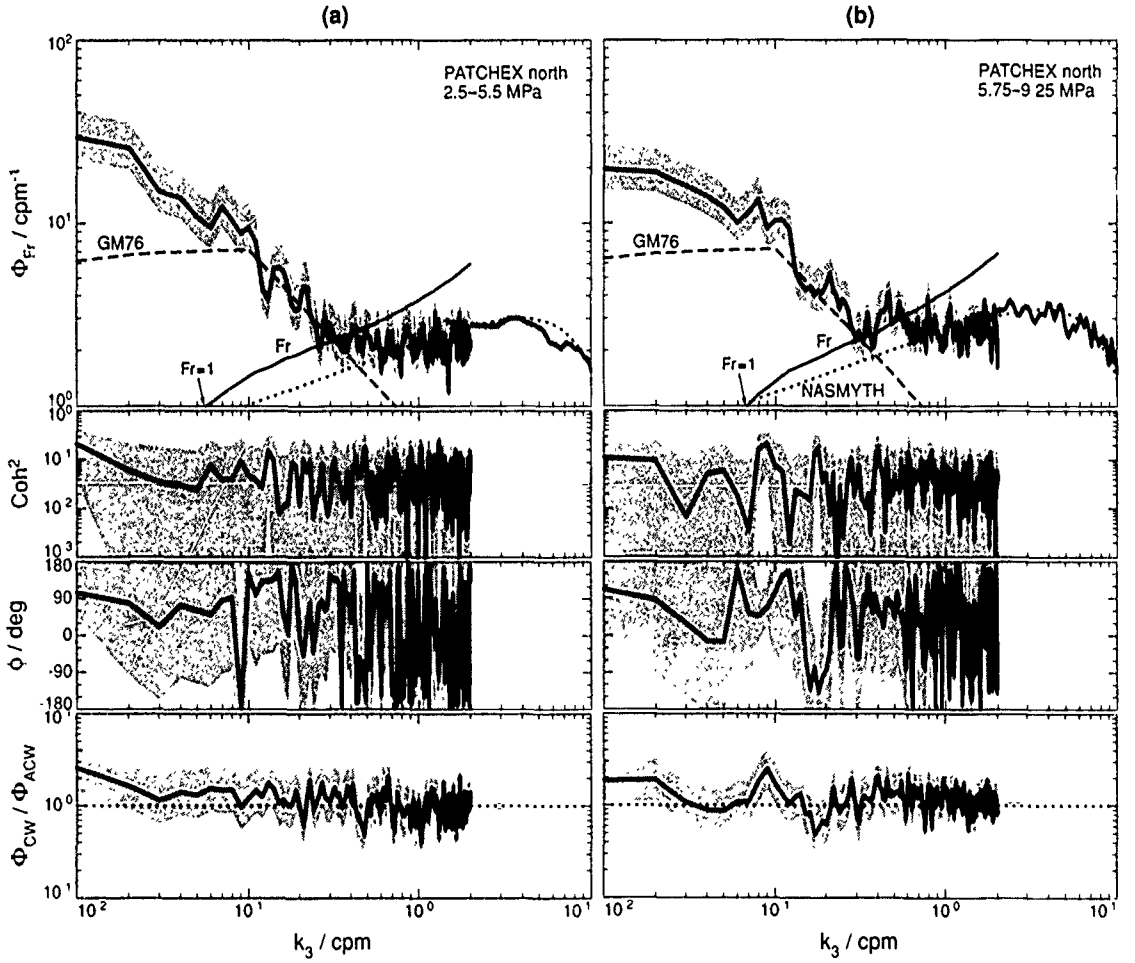


Figure 5: Spectra for PATCHEX north, in the same format as the previous figure.

the stratosphere more than the broad-band fields observed in the mid-latitude ocean. By 1 cpm, the spectra are about 1/3 of GM76. As a result of this steep rolloff, their turbulent spectra are more distinctive than for PATCHEX, even though $\langle \varepsilon \rangle$ is smaller.

The deep spectrum at 11°N starts at GM76, rises to almost twice GM76 at 0.02 cpm, and then slopes gradually downward until just past 0.1 cpm, where it too rolls off rapidly into a weak, but distinct, dissipation range (Fig. 7).

The CSALT spectrum (Fig. 7) is close to the PATCHEX spectra and to GM76, which is unexpected in view of the markedly different appearances of the profiles (Fig. 1). At 0.01 cpm, CSALT is about half of GM76, slightly lower than the PATCHEX spectra. It slopes upward to 0.03 cpm, where it flattens until it rolls off near 0.1 cpm. Unlike PATCHEX, however, this spectrum remains below GM76 when it is nearly flat and consequently rolls off from a lower amplitude than do the other spectra. Initially, the rolloff is less steep than k_3^{-1}, but for $k_3 > 2$ cpm the rolloff does not differ significantly from the model.

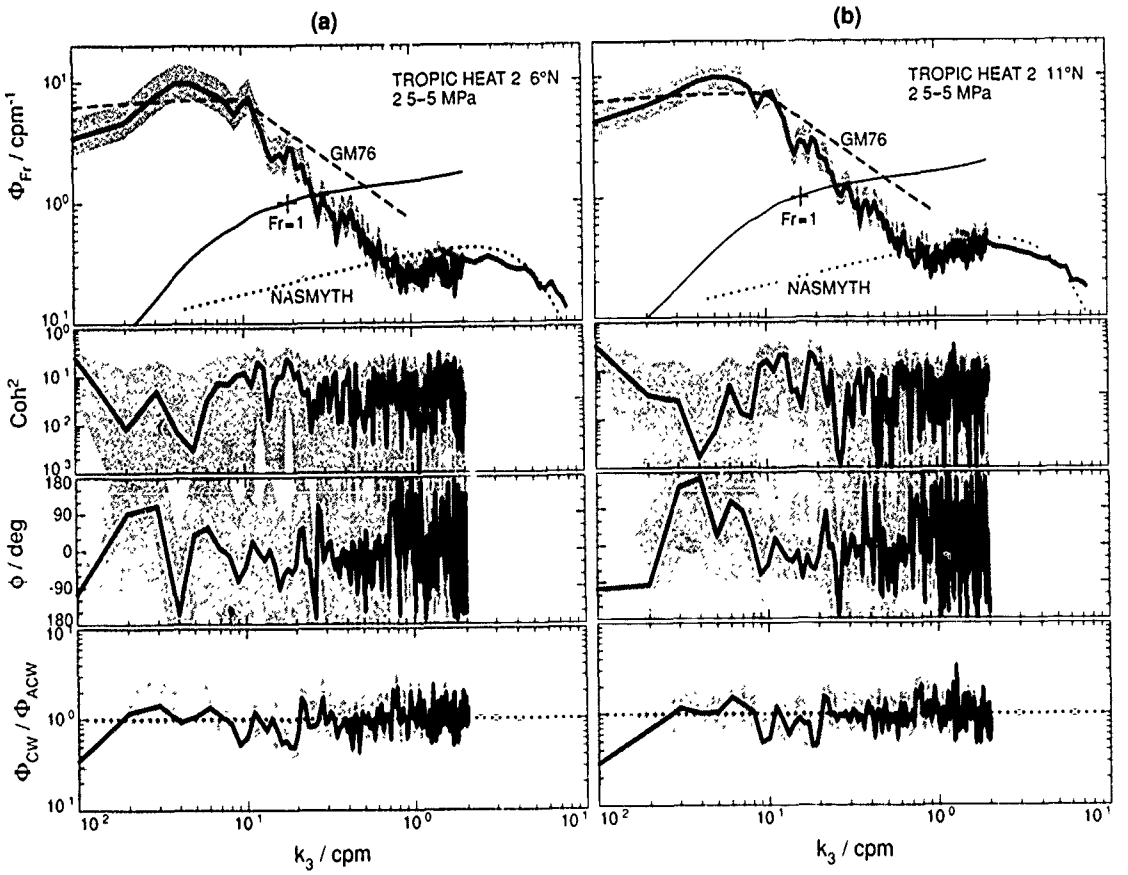


Figure 6: Shallow Froude spectra at 6°N and 11°N during TROPIC HEAT 2. They are nearly identical, including small fluctuations, even though they were taken 1000 km apart at an interval of three weeks.

For this collection of spectra, variance-preserving plots of the internal wave range (Fig. 8) differ considerably, as expected from the differences in the log-log plots. The deep PATCHEX spectrum is closest to GM76 and has the expected concentration of variance in the rolloff (a slope of k_3^{-1} on a log-log plot is flat on a variance-preserving format). The shallow PATCHEX spectrum has a distinct peak at the beginning of the rolloff, and both PATCHEX north spectra have similar peaks at slightly lower wavenumbers, corresponding to the start of their rolloffs. Steep rolloffs of the TROPIC HEAT 2 spectra result in variance contributions decreasing between 0.1 and 1 cpm, constituting the weak spectral gap mentioned previously. Finally, for CSALT the variance is more concentrated at high wavenumbers than for any other spectrum, consistent with concentration of much of the shear across density steps several meters thick (Gregg and Sanford, 1987).

VARIANCES AND SIGNIFICANT WAVENUMBERS

Owing to the variety of their shapes, differences in the spectra cannot be described adequately by E, the dimensionless energy density. In particular, variances of velocity and of shear do not always vary in the same proportion when spectra depart from GM76.

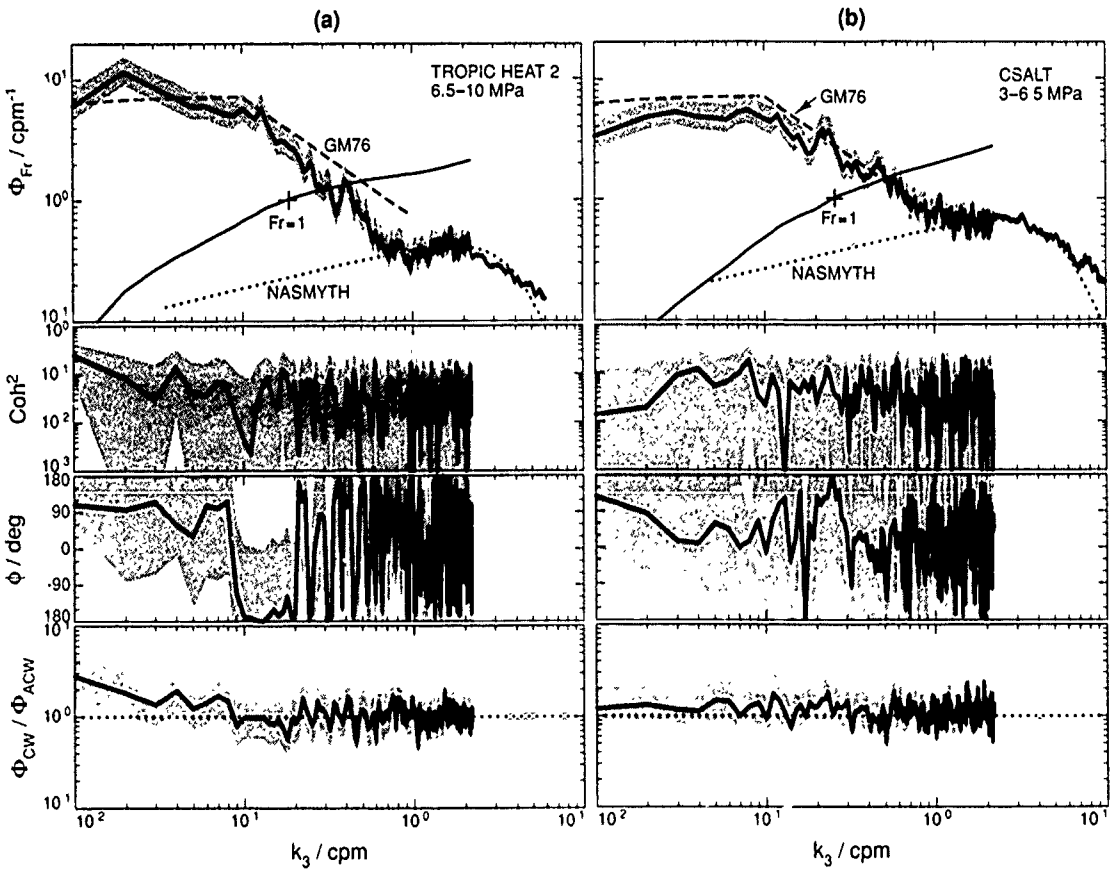


Figure 7: Froude spectra for the deep TROPIC HEAT 2 spectrum and for CSALT.

Because the velocity spectra are 'red,' velocity variances are dominated by the lowest spectral estimate and receive little contribution past the first 4 or 5 estimates. Consequently, they depend on the length of the data window and have relatively few degrees of freedom. As a compromise between length and degrees of freedom, for the longer records we compute $\Phi_{\mathrm{VEL}}(k_3)$ from 2.4 MPa pieces, overlapped by 50%. GM76 is digitized at the same wavenumbers and integrated in the same manner to yield

$$\frac{KE}{KE_{\mathrm{GM}}} \equiv \frac{\int_{0.004}^{k_{\mathrm{IW}}} \Phi_{\mathrm{VEL}}(k_3)\, dk_3}{\int_{0.004}^{k_{\mathrm{IW}}} \Phi_{\mathrm{GM76}}(k_3)\, dk_3} \tag{17}$$

where k_{IW}, the upper limit of the internal wave regime, makes little difference to the integral, and the factors of 1/2 relating kinetic energy and velocity variances cancel. As shown in Table 3, kinetic energy variances are (0.6–4.2)× GM76, consistent with previous observations that internal waves in the open ocean do not fall far below GM76. The contrast between PATCHEX and PATCHEX north appears to be between a typical 'quiet' site in open ocean and an energetic region near strong forcing, presumably the coastal jet several hundred meters above the depths of these spectra.

Contrary to expectations, the low-latitude kinetic energies are less than or only modestly larger than GM76. If the site at 6°N had the same E' as PATCHEX, from

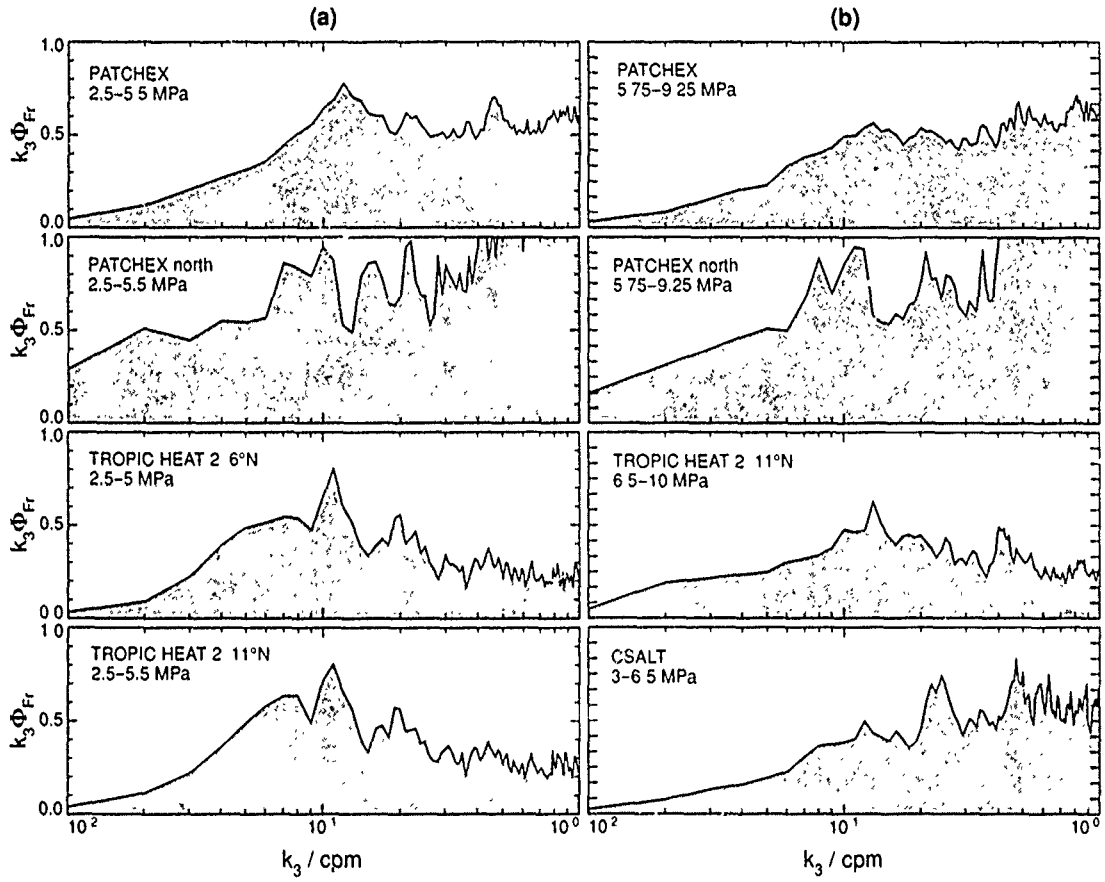


Figure 8: Variance-preserving plots of the internal wave spectra. The deep PATCHEX spectrum is closest to GM76 and has uniform contributions to the variance in the rolloff between 0.1 and 1 cpm. Both PATCHEX north spectra rise rapidly near 0.4 cpm, where turbulence becomes the dominant signal. The TH2 spectra have broad maxima centered at 0.08 cpm; hence a weak spectral gap separates internal waves from turbulence. Several spectra have local maxima near 0.1 cpm, indicating concentrations of variance at the beginning of the rolloff.

integrating (7), we would expect the kinetic energy at 6° to be $f_{34°}/f_{6°} = 5.4$ times higher than PATCHEX. Instead, we find the kinetic energy at 6°N to be 0.83 times PATCHEX. With only three profiles, this could be a statistical anomaly, but finding the same spectrum and level at 11°N suggests that this condition prevailed over a large area, at least in the spring of 1987. Moreover, we would have expected the level at 11°N to be half of that at 6°N. To the contrary, we find little difference. We conclude, therefore, that (7) may not be an accurate representation of the latitudinal structure of the internal wave field.

Analogous to the Richardson function used by Munk (1981), we quantify the shear variance with the Froude function

$$Fr(k_3) \equiv \int_{k_L}^{k_3} \Phi_{\rm Fr}(\xi)\, d\xi \qquad (18)$$

Table 3: Variances and evaluations of proposed universal constants. $\frac{KE}{KE_{GM}}$ is the ratio of kinetic energy in the observed spectrum to that in GM76 digitized and integrated numerically over the same wavenumbers. Spectra with 3.5 MPa data windows were integrated from bounds of 0.004 cpm, and the others from 0.01 cpm. Values of k_{k-1} are taken from Table 4.

Cruise	p MPa	$\frac{KE}{KE_{GM}}$	$\left(\frac{KE}{KE_{GM}}\right) \times k_{Fr=1}$ cpm	$\left(\frac{KE}{KE_{GM}}\right) \times k_{k-1}$ cpm
PATCHEX	2.50–5.50	0.86	0.15	0.10
PATCHEX	5.75–9.25	0.64	0.13	0.06–0.07
PATCHEX north	2.50–5.50	3.86	0.23	0.15–0.42
PATCHEX north	5.75-9.25	4.23	0.30	0.17–0.55
TROPIC HEAT 2 6°N	2.50–5.0	0.71	0.13	0.06–0.09
TROPIC HEAT 2 11°N	2.50–5.50	0.81	0.13	0.06–0.11
TROPIC HEAT 2 11°N	6.50–10.0	1.38	0.35	0.12–0.19
CSALT	3.00–6.50	0.71	0.18	0.06–0.09

where the lower limit, k_L, makes little difference because the principal contributions come from about a decade of wavenumbers in the middle or high end of the internal wave band, resulting in well-resolved shear variances and many degrees of freedom. Overlaying the Froude functions (Fig. 9) reveals PATCHEX north standing well above the others. The others vary, but much less than the differences between any of them and PATCHEX north. For example, at 0.01 cpm the deep spectrum from 11°N is largest of those in the 'pack,' but at 2 cpm it is less than PATCHEX. Froude functions are also plotted on Figures 4 to 7, and in Table 4 they are evaluated at several significant wavenumbers.

We choose k_{IW} as the wavenumber where an increase in slope marks the start of the dissipation range. PATCHEX and CSALT spectra have only minor inflections marking k_{IW}, but, owing to the more rapid internal wave rolloff, the TROPIC HEAT 2 spectra exhibit well-defined changes. For all cases, $k_{IW} < k_B$, often by about a factor of 2. For PATCHEX north, k_{IW} occurs at 0.3 and 0.4 cpm, where the k_3^{-1} rolloff ends and the slope changes to $k_3^{+1/3}$. We also see that k_B occurs at 0.7 and 0.6 cpm, the beginning of the inertial subrange. We have yet to investigate overturning scales thoroughly in the scalar records, but, as the inertial subrange implies isotropic turbulence, this narrow difference should be the transition from overturns affected by stratification to those not affected. If so, 0.3–0.7 cpm is much too narrow to be considered a buoyancy subrange.

Evaluating the Froude function at k_{IW}, the shear variance that can be assigned to internal waves has a relatively small range, corresponding to $Fr(k_{IW}) = 1.5$–2.5. Thus, in all cases the dissipation range begins well before $Fr = 4$, the necessary condition for shear instability. This does not imply that individual turbulent events do not have $Fr = 4$, only that it is not met on average throughout the profile. When we use the airfoil spectra to extend the Froude function through the dissipation range, we find that some spectra do not reach $Fr = 4$ even by the end of the dissipation range.

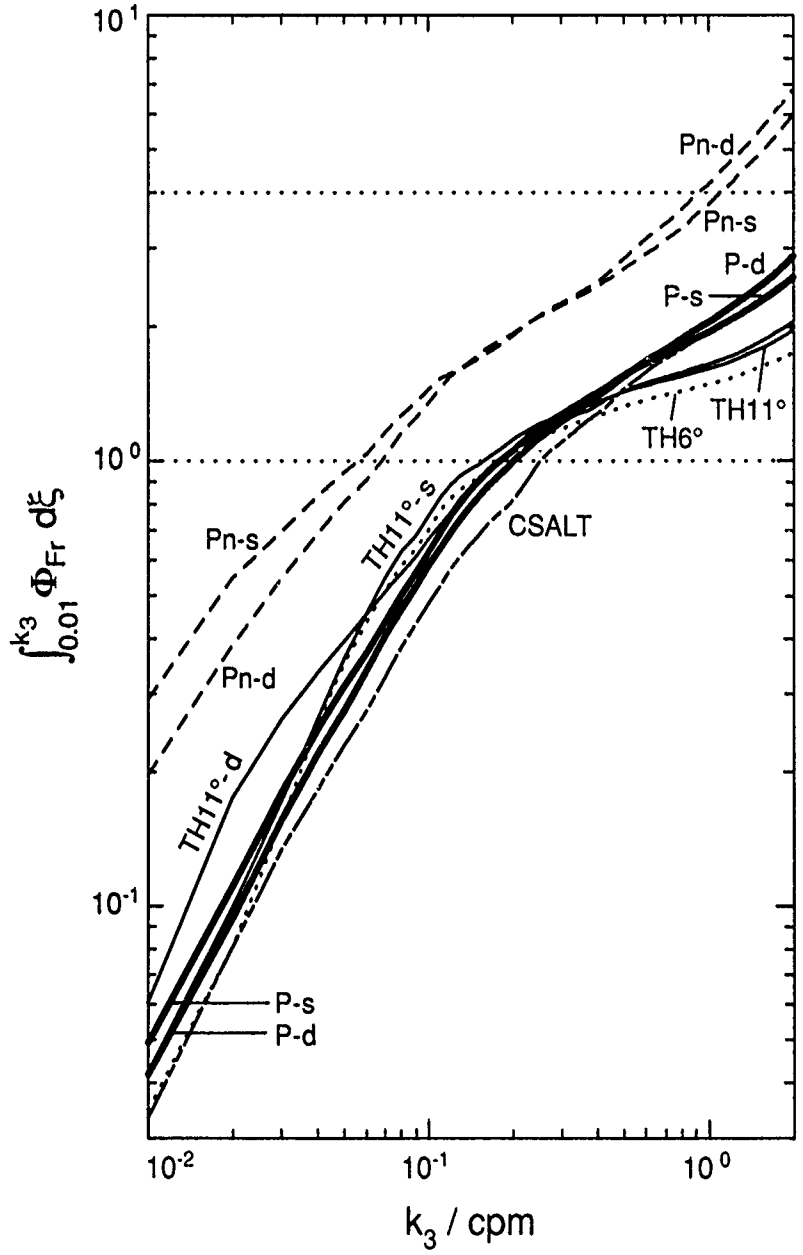


Figure 9: Comparison of all Froude functions shows two classes: PATCHEX north and the others.

The beginning of the rolloff, which we estimate visually as $k_{k^{-1}}$, is not sharply defined for most spectra, but occurs over about an octave of wavenumbers. Considering all records, the limits are 0.04 cpm $\leq k_{k^{-1}} \leq$ 0.14 cpm. Munk (1981) and Gargett et al. (1981) argue that the rolloff begins where the Froude function equals 1, i.e., $k_{\mathrm{Fr}=1} = k_{k^{-1}}$. However, from Figures 4–7 and Table 4 we see that, except for PATCHEX north, $k_{\mathrm{Fr}=1} = (1.5\text{–}3) \times k_{k^{-1}}$. As noted by Sherman and Pinkel (1991), this is also obtained by

integrating GM76, and it agrees with their observations for 1.5–1.8 MPa during PATCHEX. Again, PATCHEX north is the exception, having $k_{k-1} = k_{\mathrm{Fr}=1}$. Why does the rolloff occur well before $k_{\mathrm{Fr}=1}$ for the other spectra? Does the mechanism of instability differ? For example, the analysis of instability in the atmosphere shows that convective overturning should precede shear instability. Aspect ratios of oceanic internal waves generally preclude convective overturning, but were conditions unusual for PATCHEX north?

Table 4: Significant wavenumbers and, in the last two columns, evaluations of the Froude function at the low and high wavenumber ends of the rolloff. Because the location of the rolloff, k_{k-1}, is often indistinct, the estimates are made visually on intersection of the 95% confidence limits with the rolloff or its extension to lower wavenumber. Lower bound for the Froude function, $Fr(k_3)$, is 0.01 cpm.

Cruise	p MPa	k_{k-1} cpm	$k_{\mathrm{Fr}=1}$ cpm	k_{IW} cpm	k_{B} cpm	$Fr(k_{k-1})$	$Fr(k_{\mathrm{IW}})$
PATCHEX	2.50–5.50	0.12	0.18	1.0	2.3	0.75	1.9
PATCHEX	5.75–9.25	0.09–0.11	0.20	1.0	2.0	0.52–0.64	2.0
PATCHEX north	2.50–5.50	0.04–0.11	0.06	0.3	0.7	0.83–1.53	2.2
PATCHEX north	5.75–9.25	0.04–0.13	0.07	0.4	0.6	0.69–1.60	2.5
TROPIC HEAT 2 6°N	2.50–5.00	0.08–0.13	0.19	0.9	2.7	0.58–0.86	1.5
TROPIC HEAT 2 11°N	2.50–5.50	0.08–0.13	0.16	0.8	1.8	0.63–0.93	1.6
TROPIC HEAT 2 11°N	6.50–10.0	0.09–0.14	0.25	0.8	2.6	0.61–0.87	1.6
CSALT	3.0–6.5	0.08–0.12	0.25	1.2	2.0	0.38–0.57	2.1

Because we cannot obtain the total kinetic energy from these profiles, we test (3) as

$$\frac{KE}{KE_{\mathrm{GM}}} \times k_{k-1} = \text{constant} \tag{19}$$

Closely related is

$$\frac{KE}{KE_{\mathrm{GM}}} \times k_{\mathrm{Fr}=1} = \text{constant} \tag{20}$$

found by Duda and Cox (1989) to be a good representation of their profiles. In both cases, the measure of the cutoff, k_{k-1} or $k_{\mathrm{Fr}=1}$, varies inversely with the energy level. In neither case do (19) or (20) accurately describe our observations (Table 3). Instead, there appear to be several groups having approximately the same products. One group, containing PATCHEX, shallow TROPIC HEAT 2, and CSALT, has $(KE/KE_{\mathrm{GM}}) \times k_{\mathrm{Fr}=1} =$ 0.13–0.18 and $(KE/KE_{\mathrm{GM}}) \times k_{k-1} = 0.06$–0.11, compared with 0.23–0.30 and 0.15–0.55 for the PATCHEX north spectra, which constitute a second group. The deep spectrum from 11°N is a group by itself, with (19) close to PATCHEX north and (20) intermediate between the first two groups. In any event, neither $k_{\mathrm{Fr}=1}$ nor k_{k-1} decreases sufficiently to compensate the increase in kinetic energy for PATCHEX north relative to PATCHEX. Therefore, (19) and (20) fail to describe even our mid-latitude data.

VERTICAL SYMMETRY, COHERENCE, AND PHASE

Although GM76 assumes vertical symmetry, observations often reveal asymmetry. Most kinetic energy in the internal wave field resides at near-inertial frequencies and is polarized, with significant coherence between u and v. In the northern hemisphere, near-inertial motions with downward group velocity have clockwise rotation (CW, with $\phi = +90°$) and those with upward group velocity have anticlockwise rotation (ACW, with $\phi = -90°$). Rotary decomposition of velocity profiles often reveals $\Phi_{CW}/\Phi_{ACW} > 1$, corresponding to an excess of downward propagating energy (Leaman and Sanford, 1975).

At low wavenumbers, the shallow PATCHEX record has small, but significant, coh^2, $\phi \approx +90°$, and $\Phi_{CW}/\Phi_{ACW} > 1$ for $k = 0.01$–0.03 cpm (Fig. 4). (Because confidence limits for ϕ depend inversely on coh^2, ϕ is unreliable when coh^2 is insignificant.) At 0.01 cpm, the flux asymmetry is large, upward exceeding downward by nearly 2:1. The excess decreases with increasing wavenumber, and does not differ significantly from 1 for 0.04–0.07 cpm. Over nearly the same interval, coh^2 also decreases and is not significantly different from 0. Coherence-squared is again significant between 0.08–0.13 cpm and 0.16–0.2 cpm. In the first instance, $\phi \approx +60°$ (the confidence limits include $+90°$), and $\Phi_{CW}/\Phi_{ACW} > 1$, all consistent with downward propagation. In the second instance, $\phi \approx -90°$ and $\Phi_{CW}/\Phi_{ACW} < 1$, indicating upward propagation. Furthermore, replotting the spectra with a linear wavenumber axis reveals similar patterns throughout 0.1–1 cpm (Fig. 10). In these cases, however, the significance is much less.

At low wavenumbers, the deep PATCHEX record does not have significant coh^2, and Φ_{CW}/Φ_{ACW} is not significantly different from 1 (Fig. 4). The vertical symmetry of this record may be a consequence of its distance from surface forcing. Lack of coherence at low wavenumbers, however, does not appear to affect intermediate scales. Coherence-squared is significant for 0.05–1.2 cpm and 0.18–2.0 cpm, with the same pattern of phase and flux asymmetry as found in the shallow spectrum; i.e., the first coherent band has $\phi \approx +90°$ and $\Phi_{CW}/\Phi_{ACW} > 1$, the second has $\phi \approx -90°$ and $\Phi_{CW}/\Phi_{ACW} < 1$. Furthermore, using a linear wavenumber axis shows additional coherent patterns throughout the rolloff (Fig. 10).

We are not aware of previous reports of near-inertial signatures at high wavenumbers. To the contrary, after decomposing a time series of velocity profiles into mean, near-inertial, and high-frequency internal waves, Sanford (1991) finds that "At all wavelengths shorter than about 100 m, the higher-frequency waves are more energetic." Occurrence of the first coherent band at the beginning of the rolloff is a signature one would expect if the rolloff results from strong Doppler shifting creating critical layers for waves of this scale. The small magnitudes of coh^2 demonstrate that somewhat less than 10% of the variance at the rolloff scale is involved, which is plausible. Bands of coherent u and v structures throughout the rolloff are also consistent with wave/wave interactions, moving energy to higher wavenumbers until the waves break down into turbulence at $k_3 > 1$ cpm. Existence of coherent structures is also inconsistent with the rolloff being a buoyant subrange of three-dimensional turbulence.

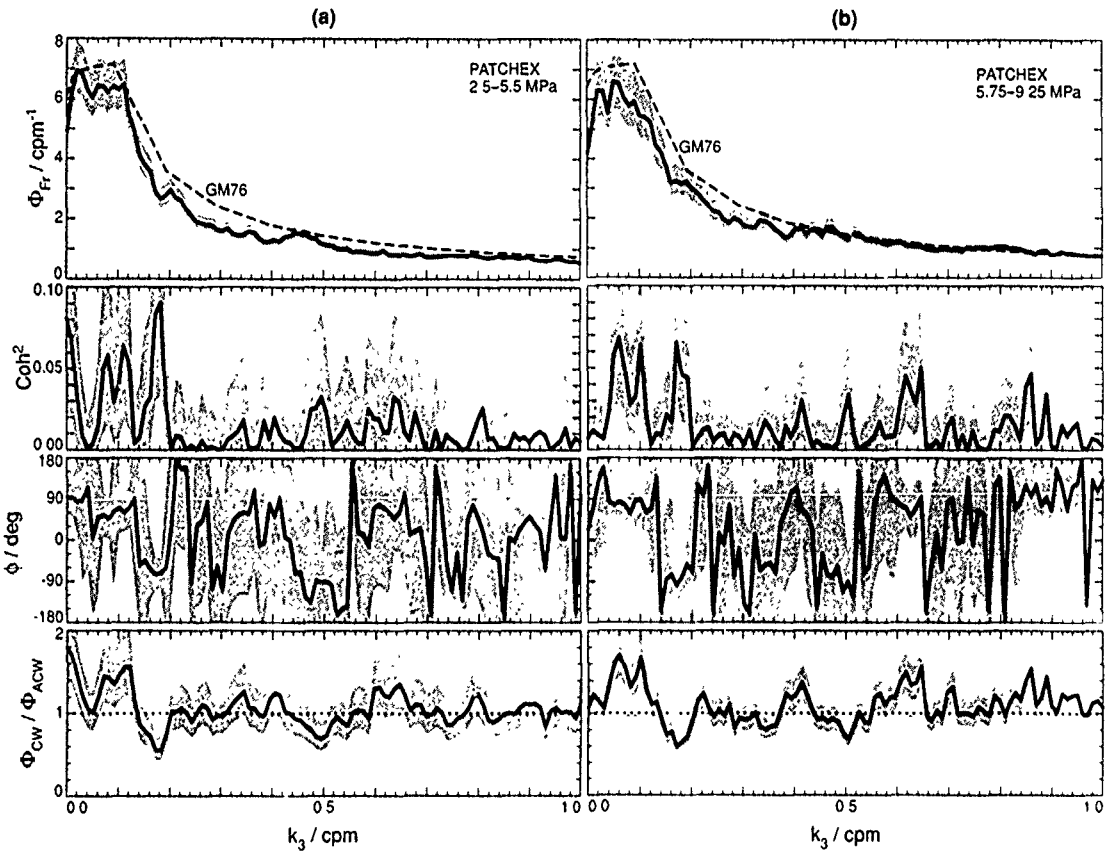


Figure 10: The PATCHEX spectra on a linear wavenumber axis. Although not as well defined as the two patterns at the beginning of the rolloff, the region from 0.1–1 cpm has numerous sections with significant coh^2, $\phi \approx \pm 90°$, and matching Φ_{CW}/Φ_{ACW} greater than or less than 1.

Similar coherent patterns occur in some of the other spectra, but, owing to fewer degrees of freedom, their significance is much less. At 0.01 cpm, the PATCHEX north spectra have $coh^2 = 0.1$–0.2, $\phi \approx +90°$, and $\Phi_{CW}/\Phi_{ACW} \geq 2$ (Fig. 5). Thus, even for the deeper record, energy at low wavenumbers is strongly dominated by downward propagating waves, and these signatures drop off with increasing wavenumber. For the shallow record, only one wavenumber near the rolloff has significant coh^2, but for the deep record the band is wider, 0.07–0.1 cpm, and coincides with the significant local spectral maximum which is very prominent in the variance-preserving plot (Fig. 8).

The shallow TROPIC HEAT 2 spectra are quite different, being dominated by upward-propagating energy: $coh^2 \approx 0.3$, $\phi \approx -90°$, and $\Phi_{CW}/\Phi_{ACW} \approx 0.3$ (Figs. 6 and 7). Curiously, coh^2 is equally large for the deep spectrum at 11°N , but $\phi \approx +90°$ and $\Phi_{CW}/\Phi_{ACW} \approx 3$. Thus, at low wavenumbers, the shallow records are strongly dominated by upward-propagating energy and the deep record is strongly dominated by downward-propagating energy, a situation we cannot explain.

The shallow spectrum at 11°N has two additional bands with significant coh^2: 0.09–0.13 cpm and 0.18–0.2 cpm (Fig. 6). For both, $\phi \approx -90°$ and $\Phi_{CW}/\Phi_{ACW} < 1$, demonstrating upward propagation. Perhaps upward propagation in the second of these bands results from the strong dominance of upward propagation at 0.01 cpm. There is a suggestion of the same pattern at 6°N, but the degrees of freedom are too small for it to be significant.

The deep spectrum at 11°N has only one spectral estimate with significant coherence near the rolloff (Fig. 7). The phase has very wide confidence limits that include −90°.

Finally, CSALT has no patterns suggesting critical layer interactions (Fig. 7).

SUMMARY AND DISCUSSION

Returning to the questions posed in the introduction, for these observations we find

- Spectral shapes change with amplitude principally at wavenumbers less than 0.1 cpm; near 0.14 cpm all of the Froude spectra have nearly the same amplitude. Consequently, the only systematic structure to variations for $k_3 = 0.01$–0.1 cpm is sloping downward when the amplitude exceeds GM76 and sloping upward when the amplitude is smaller than GM76. Owing to the irregular shapes, kinetic energy and shear variances do not change in the same proportion.

- Spectra from low latitudes in the central Pacific differ markedly from those at mid-latitude, but the differences are not those predicted by Munk's (1981) addition of latitude dependence to GM76. Munk predicts kinetic energies 3 and 5 times GM76 at 11°N and 6°N, respectively; whereas we observe 0.8 and 1.4 times GM76 in the shallow and deep spectra at 11°N, and 0.7 times GM76 in the shallow spectrum at 6°N. Both shallow spectra have band-limited shear variances concentrated near 0.08 cpm and roll off as $k_3^{-1.4}$, rather than as k_3^{-1} at mid-latitude. Although not as peaked as the shallow spectra, the deep spectrum at 11°N rolls off as steeply and also has a spectral gap separating internal waves and turbulence.

- The CSALT spectra are similar to PATCHEX and to GM76, the principal difference being somewhat lower levels near 0.01 cpm. In view of the different appearance of the CSALT velocity profiles, the similarity of the spectra to PATCHEX and to GM76 is unexpected. It is not clear, however, what spectral shape is typical at that latitude. The CSALT spectra, taken at 12°N, differ markedly from those observed at 11.5°N during TROPIC HEAT 2. More spectra at low latitudes are needed to define the norm, assuming one exists.

- In most of the spectra, bands of near-inertial energy (as indicated by significant coherence-squared between u and v with phases of ±90°) span the rolloff, k_{k-1}, and recur irregularly throughout the rolloff region, usually 0.1–1 cpm. These signatures are unexpected, as high-frequency waves generally dominate wavenumbers greater than 0.01 cpm (Sanford, 1991), and we interpret them as evidence of critical layer

interactions causing the rolloff. Preliminary examination of overturning scales in these profiles shows vertical scales consistent with k_B, i.e., tens of centimeters or less except for PATCHEX north which has overturns of about 1 m. Thus, in both cases, the rolloff cannot contain the three-dimensional turbulence needed for a buoyancy subrange.

- The PATCHEX north spectrum rolls off at a lower wavenumber than does the PATCHEX spectrum, as expected from the larger velocity and shear variances of PATCHEX north. The decrease in k_{k-1}, however, is only about half the decrease predicted by Munk's (1981) $E \times k_{k-1} = \text{constant}$. For the other records, k_{k-1} shows little systematic shift with changes in kinetic energy but all of them roll off where $Fr = 0.5$–0.8.

- When evaluated with $\langle \varepsilon \rangle$, Nasmyth's spectrum adequately approximates the observed spectra in the dissipation range, particularly for the PATCHEX north records which are strongly turbulent. Our spectra, however, roll off more gradually beyond the viscous cutoff, and we do not know whether this results from the averaging procedure or is more fundamental.

Considering the range of spectral amplitudes, the variety of spectral shapes for 0.01–0.1 cpm is not surprising. Energy enters at low wavenumbers, and the spectrum is likely to contain signatures of the generation mechanisms, much like energy-containing scales of turbulence. (This is especially true for these data, which were collected at relatively shallow depths and may have sampled forced rather than freely propagating waves.) The observed departures from white shear spectra, i.e., from k_3^0, imply some degree of coherent structure in the low-wavenumber shear field. This is a general consequence of a white spectrum and a unit impulse autocorrelation function being a Fourier transform pair, and is the basis for the cumulative periodogram test for correlatedness (Jenkins and Watts, 1969). As generation mechanisms are identified, it may be possible to infer mechanisms from spectral signatures and to relate changes in kinetic energy to changes in shear. At present, however, we cannot identify the mechanisms. Nor can we accurately predict variations in shear from those in kinetic energy.

During eight days of sampling, internal waves did not change appreciably during PATCHEX. The other observations were collected much more rapidly and thus represent only snapshots. However, the similarity of the shallow spectra from 6°N and 11°N makes it plausible to assume that these also represent a steady state. Furthermore, the long time scale of the thermohaline staircase east of Barbados is strong evidence that CSALT also sampled a steady state.

In spite of significant differences in low-wavenumber shear and in the steepness of their rolloffs, PATCHEX and TROPIC HEAT 2 have nearly the same average dissipation rates, as does CSALT. These rates are minimal, slightly above 10^{-10} W kg^{-1}. Minimal dissipation rates reflect the tendency of all thermodynamic systems to evolve to states of minimum entropy production consistent with the constraints imposed upon them (De Groot and Mazur, 1962). By contrast, $\langle \varepsilon \rangle$ is not minimal for PATCHEX north.

We conclude that the most likely interpretation of the spectra and their dissipation rates is

1. PATCHEX and TROPIC HEAT 2 sampled steady states or nearly steady states of freely propagating waves close to equilibrium; i.e., the observations typify long-term averages at those sites, $\langle\varepsilon\rangle$ represents the average energy flux through the shear spectra, and the spectral shapes are adjusted to maintain that energy flux.
2. The contrast in spectral shapes between PATCHEX and TROPIC HEAT 2 results from differences in large-scale forcing of the internal wave field and in latitudinal changes in the rate at which wave/wave interactions transfer energy.
3. Either PATCHEX north was in steady state with active forcing or it was decaying from previous forcing.

In view of the limited sampling, we obviously cannot prove these statements; we regard them as hypotheses to be tested by further observations and perhaps by numerical calculations. In Figure 9, we have already noted how the Froude function for PATCHEX north stands above the others. Plotting ratios of the observed Froude functions to the Froude function for GM76 shows how these functions change with wavenumber (Fig. 11). Except for PATCHEX north, the Froude ratios converge to about 0.8, regardless of whether they start above or below 1. All of these records have low dissipation rates, as shown by the variance-preserving dissipation plots in the right-hand panels. The ratios for PATCHEX north drop with increasing wavenumber, but terminate at k_{IW}, the end of the internal wave region, before reaching 0.8. They correspond to the only high dissipation spectra. We conclude that $Fr(k_{\mathrm{IW}})$ is likely the major factor establishing $\langle\varepsilon\rangle$.

Most sampling during TROPIC HEAT 2 was concentrated on the equator, but we also profiled at 1°N and 2°N. At low wavenumbers, these profiles have Froude ratios that fill in the space in Figure 11 between PATCHEX north and the deep record at 11°N. Unlike PATCHEX north, however, the Froude ratios for the 2°N–0°N spectra drop steeply with increasing wavenumber and terminate near 0.8. They also have minimal dissipation rates and overlay the low values on the variance-preserving spectra. Owing to the strong mean shears within 2° of the equator, wave/wave interactions are likely to be more complicated than for the records we have presented. Nevertheless, the shapes of their spectra are consistent with the hypothesis that $\langle\varepsilon\rangle$ is controlled by $Fr(k_{\mathrm{IW}})$.

The rapid rolloff in the low-latitude spectra is consistent with reduced wave/wave flux feeding energy to small scales. But, if that flux is linearly proportional to f, why isn't the rolloff steeper at 6°N than at 11°N. And, why are the dissipation rates at these stations nearly the same as at PATCHEX? Because velocity and shear variances for $k_3 \leq k_{k-1}$ are comparable to PATCHEX, we expect $\langle\varepsilon\rangle$ at 11°N to be 1/3 of PATCHEX and $\langle\varepsilon\rangle$ at 6°N to be 1/5 of PATCHEX.

The location of the rolloff, k_{k-1}, shows the same two groups of profiles that we obtain from the Froude functions, namely PATCHEX north, with $Fr(k_{k-1}) \approx 1$, and the others,

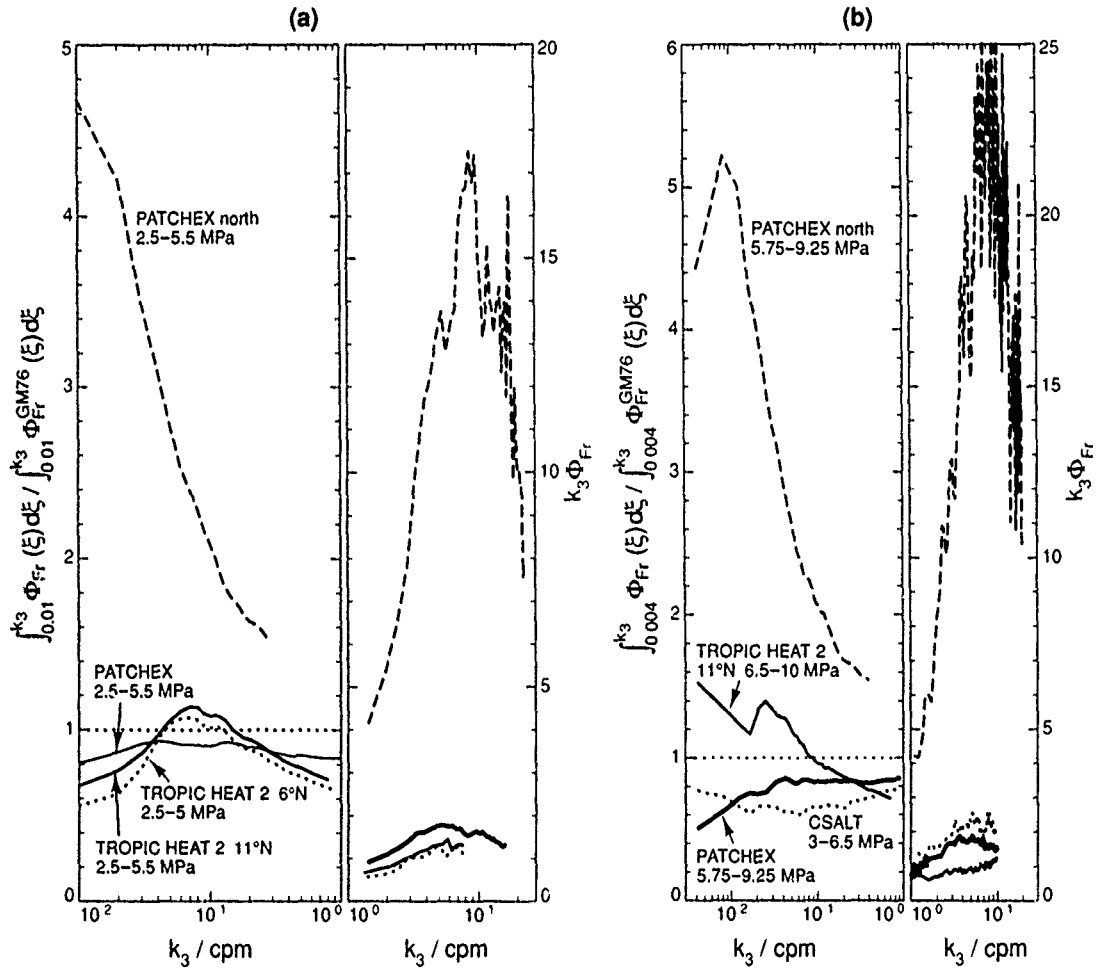


Figure 11: Ratios of observed Froude functions to the Froude function and variance-preserving dissipation spectra for shallow spectra (left) and for the deep spectra plus CSALT (right). Froude plots are cut off at k_{IW}. In spite of larger differences in amplitude and shape among the spectra, only PATCHEX north stands out.

with $Fr(k_{k^{-1}}) \approx 0.5$. Although we do not consider the rolloff a buoyancy subrange, we applied the buoyancy scaling used by Gargett et al. (1981) to PATCHEX and PATCHEX north, which have similar values of f. The results are dramatic (Fig. 12), achieving a better collapse than obtained by Gargett et al. In addition, the collapse extends across the internal wave range and is not limited to the rolloff. If this collapse is not fortuitous, the scaling must be an internal wave scaling rather than a buoyancy scaling. As in the models of wave/wave interactions by McComas and Müller (1981) and Henyey et al. (1986), for a wave field in equilibrium $\langle\varepsilon\rangle$ is also the rate at which energy flows through the spectrum. In which case, the striking collapse of PATCHEX and PATCHEX north indicates that the shape of the PATCHEX north spectrum is close to that required to maintain a uniform energy flux.

Occurrence of the rolloff at 0.1 cpm, instead of near 0.013 cpm as predicted by Holloway (1980), indicates that $c_h \approx u_{RMS}$ is not the criterion for Doppler shifts leading to

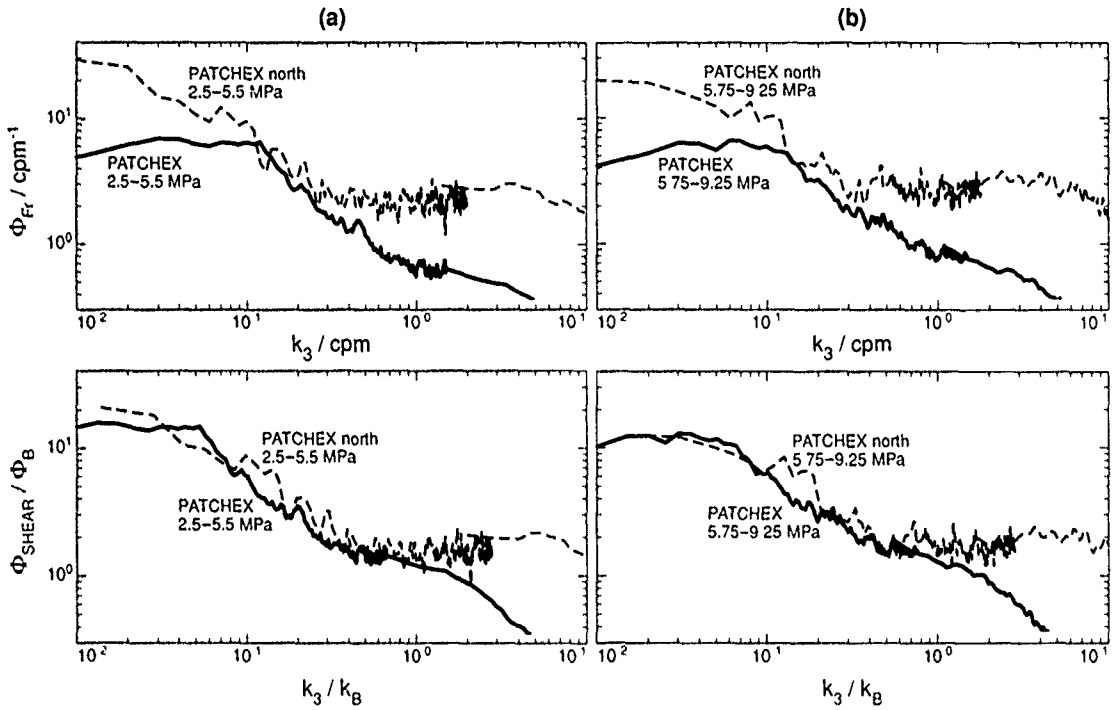


Figure 12: Froude and buoyancy scalings applied to the shallow (left) and deep (right) spectra from PATCHEX and PATCHEX north. The buoyancy scaling gives a much better collapse throughout the internal wave range, $k_3/k_B < 1$ in our interpretation, but not in the turbulent range, $k_3/k_B > 1$.

critical layers. We suggest two reasons for the apparent discrepancy in the scale of the rolloff: 1) The probability distribution for velocity differences produced by large-scale waves is skewed. Consequently, even though $ch \approx u_{RMS}$ at 0.013 cpm, half or more of the distribution at that scale has smaller velocity differences. 2) Critical layers also require alignment in direction between the 'test' wave and the field of the large-scale waves. For a horizontally isotropic wave field, this further reduces the likelihood of a critical layer when $ch \approx u_{RMS}$.

When u and v are normally distributed, $u^2 + v^2$ has a chi-square distribution (when suitably normalized), which is skewed. Consequently, the most probable velocity difference experienced by a wave is less than u_{RMS}. In Figure 13, we show probability density functions of $\log_{10}[(\Delta u)^2 + (\Delta v)^2]$ for a range of Δz. For PATCHEX, when $\Delta z = 50$ m, corresponding to $c_h \approx u_{RMS}$, half of the velocity differences are less than c_h^2, and the probability of matching amplitude and direction must be relatively low. For $\Delta z = 10$ m, however, only a few percent of the velocity differences are less than c_h^2, and the probability of matching amplitude and direction must be much larger. The probability densities for PATCHEX north are not as well defined, but show few velocity differences less than c_h^2 for $\Delta z = 50$–25 m, consistent with the rolloff beginning at lower wavenumbers.

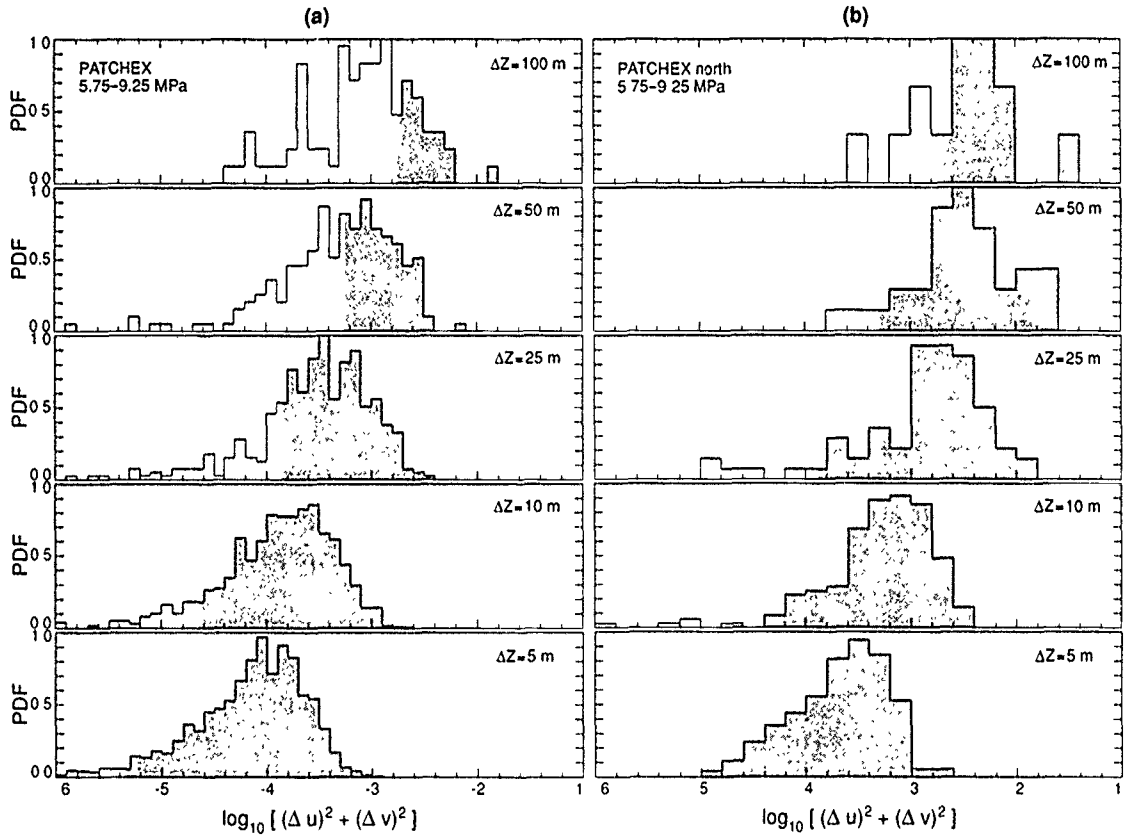


Figure 13: Probability density functions of $(\Delta V)^2$ for different vertical lengths, Δz. Shading shows where $\Delta V \geq c_h$, the horizontal phase speed of internal waves having the same vertical scale. The histograms are fully shaded only when $\Delta z \leq 10$ m.

Much of this discussion is speculation in an attempt to formulate the hypotheses needed to move from internal wave kinematics to dynamics. These records demonstrate that, until we better understand dynamics, we need to continue fully resolving the vertical shear spectrum. Further observations are needed to determine the full range of shapes and amplitudes for shear spectra, but, even more, to reveal how these characteristics evolve during strong forcing.

Acknowledgments. Collection of the MSP data was funded by the Office of Naval Research and the National Science Foundation. The Office of Naval Research funded this analysis through the *Mixing to Mesoscale* University Research Initiative at the University of Washington (Contract N00014-86-K-0690). Don Percival and Harvey Seim provided useful discussions and comments. Contribution 1891 of the School of Oceanography.

REFERENCES

Bendat, J. S., and A. G. Piersol, 1971: *Random Data: Analysis and Measurement Procedures.* Wiley-Interscience, New York, 407 pp.

De Groot, S. R., and P. Mazur, 1962. *Non-Equilibrium Thermodynamics*. North-Holland Pub. Co., Amsterdam, 510 pp.

Dewan, E. M., and R. E. Good, 1986: Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere. *J. Geophys. Res.*, **91**, 2742–2748.

Dillon, T. M., 1982: Vertical overturns: A comparison of Thorpe and Ozmidov length scales. *J. Geophys. Res.*, **87**, 9601–9613.

Duda, T. F., and C. S. Cox, 1989: Vertical wave number spectra of velocity and shear at small internal wave scales. *J. Geophys. Res.*, **94**, 939–950.

Eriksen, C. C., 1978: Measurements and models of fine structure: Internal gravity waves, and wave breaking in the deep ocean. *J. Geophys. Res.*, **83**, 2989–3009.

Eriksen, C. C., 1980: Evidence for a continuous spectrum of equatorial waves in the Indian Ocean. *J. Geophys. Res.*, **85**, 3285–3303.

Evans, D. L., H. T. Rossby, M. Mork, and T. Gytre, 1979: Yvette—a free-fall shear profiler. *Deep-Sea Res.*, **26**, 703–718.

Fritts, D. C., 1984: Gravity wave saturation in the middle atmosphere: A review of theory and observation. *Rev. Geophys. Space Phys.*, **22**, 275–308.

Fritts, D. C., 1989: A review of gravity wave saturation processes, effects, and variability in the middle atmosphere. *PAGEOPH.*, **130**, 343–371.

Gargett, A. E., P. J. Hendricks, T. B. Sanford, T. R. Osborn, and A. J. Williams III, 1981: A composite spectrum of vertical shear in the upper ocean. *J. Phys. Oceanogr.*, **11**, 1258–1271.

Garrett, C. J. R., and W. H. Munk, 1975: Space-time scales of internal waves: A progress report. *J. Geophys. Res.*, **80**, 291–297.

Gregg, M. C., 1977a: A comparison of finestructure spectra from the main thermocline. *J. Phys. Oceanogr.*, **7**, 33–40.

Gregg, M. C., 1977b: Variations in the intensity of small-scale mixing in the main thermocline. *J. Phys. Oceanogr.*, **7**, 436–454.

Gregg, M. C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.*, **94**, 9686–9698.

Gregg, M. C., and E. Kunze, 1991: Shear and strain in Santa Monica Basin. *J. Geophys. Res.*, in press.

Gregg, M. C. and T. B. Sanford, 1987: Shear and turbulence in thermohaline staircases. *Deep-Sea Res.*, **34**, 1689–1696.

Gregg, M. C., and T. B. Sanford, 1988: The dependence of turbulent dissipation on stratification in a diffusively stable thermocline. *J. Geophys. Res.*, **93**, 12,381–12,392.

Gregg, M. C., C. S. Cox, and P. W. Hacker, 1973: Vertical microstructure measurements in the central North Pacific. *J. Phys. Oceanogr.*, **3**, 458–469.

Gregg, M. C., H. Seim, and D. B. Percival, 1991: Statistics of shear and turbulent dissipation in profiles. *J. Geophys. Res.*, submitted.

Hayes, S. P., H. B. Milburn, and E. F. Ford, 1984: TOPS: a free-fall velocity and CTD profiler. *J. Atmos. Oceanic Technol.*, **1**, 220–236.

Henyey, F.S., J. Wright, and S. M. Flatté, 1986: Energy and action flow through the internal wave field: An eikonal approach. *J. Geophys. Res.*, **91**, 8487–8495.

Hess, W. C., D. B. Percival, and M. C. Gregg, 1991: Robust estimation of the rate of turbulent dissipation. (in preparation)

Hines, C. O., 1991a: The saturation of gravity waves in the middle atmosphere. Part I: Critique of linear-instability theory. *J. Atmos. Sci.*, in press.

Hines, C. O., 1991b: The saturation of gravity waves in the middle atmosphere. Part II: Development of Doppler-spread theory. *J. Atmos. Sci.*, in press.

Holloway, G., 1980: Oceanic internal waves are not weak waves. *J. Phys. Oceanogr.*, **10**, 906–914.

Jenkins, G. M. and D. G. Watts, 1969: *Spectral Analysis and Its Applications.* Holden-Day, San Francisco, 525 pp.

Leaman, K. D., and T. B. Sanford, 1975: Vertical energy propagation of inertial waves: A vector spectral analysis of velocity profiles. *J. Geophys. Res.*, **80**, 1975–1978.

Levine, M. D., C. A. Paulson, and J. H. Morison, 1985: Internal waves in the Arctic Ocean: Comparison with lower-latitude observations. *J. Phys. Oceanogr.*, **15**, 800–809.

Lumley, J. L., 1964: The spectrum of nearly inertial turbulence in a stably stratified fluid. *J. Atmos. Sci.*, **21**, 99–102.

McComas, C. H., and P. Müller, 1981: The dynamic balance of internal waves. *J. Phys. Oceanogr.*, **11**, 970–986.

Munk, W. H., 1981: Internal waves and small-scale processes. In *Evolution of Physical Oceanography, Scientific Surveys in Honor of Henry Stommel*, edited by B. A. Warren and C. Wunsch, 264–291, MIT Press, Cambridge, Mass.

Oakey, N. S., 1982: Determination of the rate of dissipation of turbulent energy from simultaneous temperature and velocity shear microstructure measurements. *J. Phys. Oceanogr.*, **12**, 256–271.

Percival, D., 1991: A note on confidence intervals for the magnitude squared coherence. *J. Time Ser. Anal.*, submitted.

Sanford, T. B., 1991: Spatial structure of thermocline and abyssal internal waves. This volume.

Sherman, J. T., and R. Pinkel, 1991: Estimates of the vertical wavenumber-frequency spectra of vertical shear and strain. *J. Phys. Oceanogr.*, **21**, 292–303.

Shur, G. N., 1962: Experimental studies of energy spectrum of atmospheric turbulence. *Tr. Tsent. Aerolog. Observ.*, **43**, 79–90.

Singleton, R. C., 1969: An algorithm for computing the mixed radix fast Fourier transform. *IEEE Trans. Audio and Electroacoustics*, **AU-17**, 93–103.

Smart, J. H., 1988: Comparison of modelled and observed dependence of shear on stratification in the upper ocean. *Dyn. Atmos. Oceans*, **12**, 127–142.

Smith, S. A., D. C. Fritts, and T. E. VanZandt, 1987: Evidence for a saturated spectrum of atmospheric gravity waves. *J. Atmos. Sci.*, **44**, 1404–1410.

Wunsch, C. and S. Webb, 1979: The climatology of deep ocean internal waves. *J. Phys. Oceanogr.*, **9**, 235–243.

INTERNAL WAVE DISSIPATION IN A NON-GARRETT-MUNK OCEAN

L. Padman[1], T.M. Dillon[1], H.W. Wijesekera[1], M.D. Levine[1], C.A. Paulson[1], and R. Pinkel[2]

[1]College of Oceanography, Oregon State University, Corvallis, OR 97331-5503.
[2]Scripps Inst. of Oceanography, La Jolla, CA 92093.

ABSTRACT

Recent work by M.C. Gregg has suggested that the dissipation rate of turbulent kinetic energy in the oceanic thermocline may be closely related to the internal wave shear variance measured at 10 m vertical scales. If this relationship holds, it significantly decreases the experimental effort required to estimate eddy viscosities and diffusivities. However, as Gregg has noted, the diffusivities associated with his model in a Garrett-Munk (GM) canonical wave field are quite small, and it therefore appears that mixing driven by internal waves may only be important where the mean energy level is substantially above GM. In this paper we explore the relationship between internal waves and microscale dissipation from a region of energetic mixing, near the Yermak Plateau in the Arctic Ocean. The dissipation rate is an order of magnitude above that predicted by Gregg's model, and we shall discuss several aspects of the local internal wave field which may contribute to this result. We also suggest one way in which Gregg's model might be modified to account for this new data set.

INTRODUCTION

The recent paper by Gregg (1989) (hereinafter G89), which suggested a close relationship between internal wave (IW) shear and dissipation rate, ϵ, has prompted a number of studies aimed either at (a) applying his results to other appropriate data sets, or (b) searching for locations in which the scaling model does not work. As Gregg himself has indicated (pers. comm., 1990), anomalous environments may be regions of large diapycnal transport of both momentum and scalars, perhaps dominating the basin-averaged diapycnal fluxes. It is therefore necessary to

determine why these regions differ from the cases studied in G89, and whether a predictive model, based on observed IW field properties, can be found for these regions. G89 noted that the diffusivities implied by the model applied to a GM wave field at the canonical energy density level are small, and may therefore not be relevant to the dynamics of the mid-ocean thermocline, except perhaps on time scales of decades or longer. Apparently then, the places to look for vertical mixing rates which are sufficiently large to be dynamically important are regions where the IW energy density is substantially higher than the canonical value. However, as the present analysis will show, regions where the energy density is not necessarily abnormally large, but the wave field deviates in other ways from the GM assumptions, can also be regions of large diapycnal transports.

One of the aims of Gregg's study was to find ways to reduce the field effort involved in estimating ϵ. One possibility, which he explored, was to use Expendable Current Profilers (XCPs) to determine the shear at 10 m vertical scales. For the data sets which he considered, shear variance was a good indicator of ϵ in some averaged sense. However, this method still requires that the experimentalist be present to launch the XCPs, and the questions of sampling bias in ϵ estimates due to small sample size (Baker and Gibson, 1987) still remain. In an ideal world, one would obtain estimates through all types of events which might contribute to ϵ within the time and space intervals of interest. For example, if ϵ could be linked to some basic variables which are obtainable from long-term moorings (*e.g.* IW energy density, and buoyancy frequency, N), then ϵ could be "monitored" throughout the life of the mooring. Unfortunately, as we shall show below, a model based simply on energy and N doesn't work well in some environments where mixing is important. However, there may be higher order IW statistics which could be measured, and used to improve models. The requirements for models are not as strict as they might seem at first, since even a factor of two error in a model prediction might be much less than the errors involved in extrapolating the results of a short-duration microstructure or XCP program.

The data set which we shall use to explore this topic was obtained from the Oceanography ("O") Camp, deployed on the pack ice in March-April 1989 during the Coordinated Eastern Arctic Experiment (CEAREX). A background paper which fully describes the program and the data is in press (Padman and Dillon, 1991; hereinafter PD91), however a short review is provided in the following section of the current paper. We then review some of the anomalous characteristics of the O Camp IW climate, followed by a discussion of the implications for predictive modelling of ϵ in this environment. There are two seemingly dichotomous

aims to this study: first, we wish to show how the observed wave field is inconsistent with the assumptions of existing wave/dissipation models; and second, we wish to indicate that some modelling success can be retained despite these inconsistencies. Hopefully, however, the two approaches are actually complementary, even if the empirical modeling is primarily a short-term effort awaiting the construction of more versatile IW dissipation models.

THE CEAREX PROGRAM

CEAREX was a multi-investigator Arctic oceanographic experiment conducted from several platforms, including the O Camp from which the data discussed in this paper were obtained. Apart from direct microstructure measurements, data from O Camp include Acoustic Doppler Current Profiler (ADCP) measurements in the upper several hundred meters. and densely sampled CTD data, as well as water depth and meteorological information. A schematic of the various programs at O Camp is shown in Fig. 1. The drift track (Fig. 2) was determined from ap-

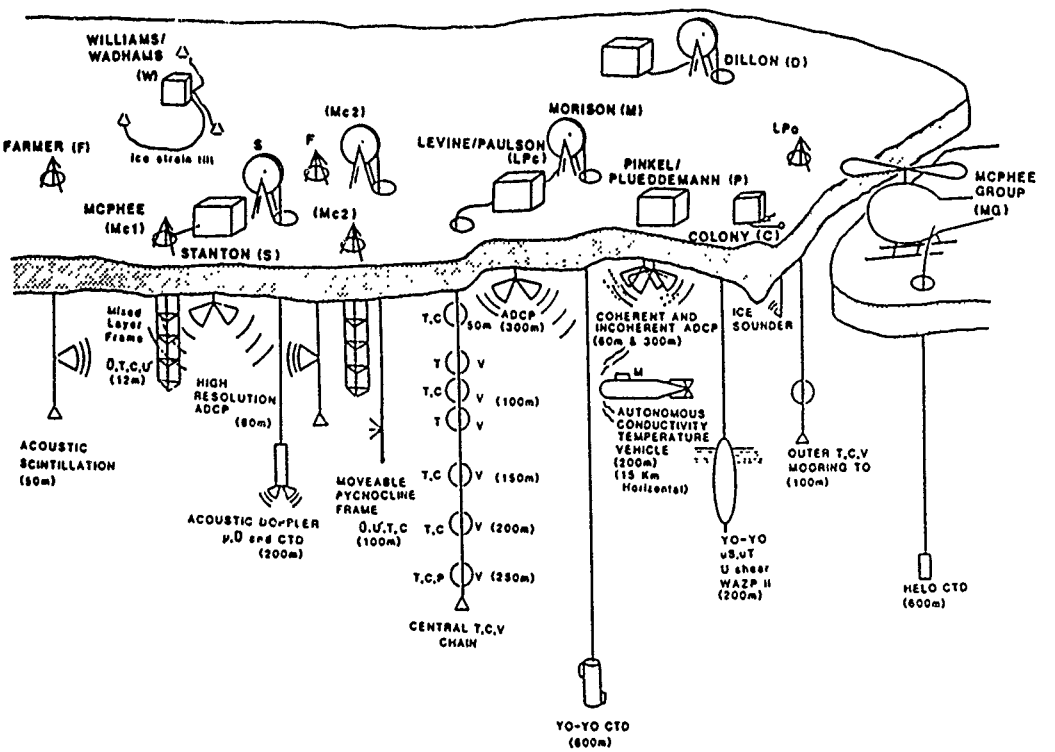


Fig. 1. Schematic of CEAREX O Camp, showing the principal sampling programs.

proximately hourly satellite fixes, then smoothed with the complex demodulation algorithm described by McPhee (1988). This algorithm assumes that the ice motion consists solely of tidal and inertial oscillations, and low-frequency drift.

Approximately 1500 microstructure profiles were made with the Rapid-Sampling Vertical Profiler (RSVP) (Caldwell, *et al.*, 1985; Padman and Dillon, 1987) from March 31 to April 25, 1989. Data were obtained between the surface and a typical maximum depth of 340 m, and the cycling time between profiles was usually 15–20 minutes. The RSVP (Fig. 3) is a tethered, freefall profiler about 1.3 m long, equipped with sensors for measuring pressure, P, orthogonal microscale velocity shears, $u_z(= \partial u/\partial z)$ and $v_z(= \partial v/\partial z)$, temperature, T, and conductivity, C. The average fall rate is about 0.7 m s^{-1}, constrained by a drag element consisting of an annular brush near the rear of the probe during descent, but able to slide towards the RSVP's nose for improved retrieval dynamics. A profile to 340 m takes about 8 minutes. The raw data sampling rate during CEAREX was 256 Hz for all channels. Microscale temperature was measured with a Thermometrics FP07 thermistor (T_1) projecting forward from the probe nose assembly. Gregg and Meagher (1980) measured a 3-dB attenuation point for similar thermistors to be approximately 15 Hz. Conductivity was measured with a Neil Brown Instru-

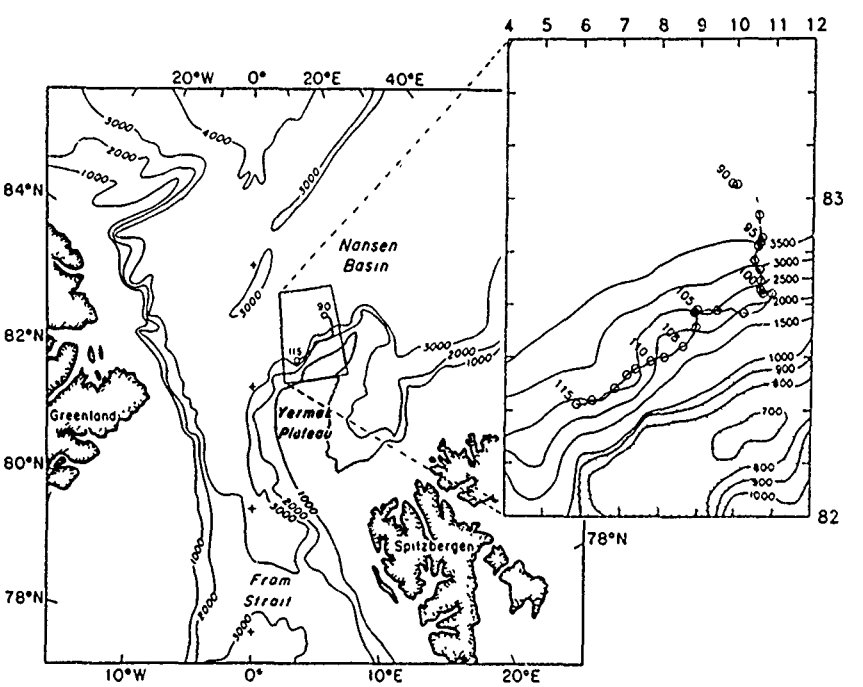


Fig. 2. Drift track of the CEAREX O Camp. Depth contours are in meters. Symbols are given once daily, and are marked in day-of-year (1989).

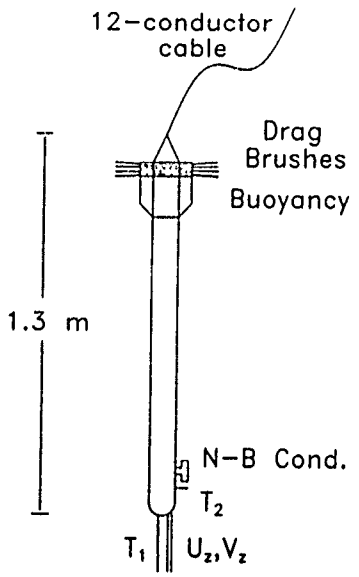


Fig. 3. Schematic of the Rapid-Sampling Vertical Profiler. Sensors are thermistors (T_1 and T_2), conductivity cell (*N-B Cond.*), pressure transducer (P), and airfoil shear probes (U_z and V_z).

ment Systems (NBIS) conductivity cell ($N - B\ Cond.$) mounted on the side of the probe, 0.15 m above the probe tip. For salinity determination a second FP07 thermistor (T_2) was mounted adjacent to the conductivity sensor. The NBIS cell has a response length of O(0.1) m (Gregg *et al.*, 1982), so that at the nominal fall rate the time constant of the conductivity cell is 0.14 s. Post-analysis of the conductivity data indicated a small, time-varying calibration offset voltage: conductivities have therefore been corrected by comparison with CTD data, provided by J. Morison. Vertically averaged salinity, S, was determined from similarly filtered T and C: the effective resolution for S and density, σ_t, is about 0.2 m. Least significant bit (*lsb*) resolutions of the raw 16-bit records are about 1.5×10^{-4}°C in temperature, and 1.5×10^{-5} S m^{-1} in conductivity. Typical *rms* noises, based on measurements deep within non-turbulent surface mixed layers, are comparable to the *lsb* resolutions.

Velocity microstructure was measured with two orthogonally mounted airfoil shear sensors (U_z, V_z) on the RSVP's nose. These probes (Osborn and Crawford, 1980) have a spatial resolution of about 0.02 m, sufficient to resolve most of the "universal", or Kolmogorov, shear spectrum for typical oceanic dissipation rates. Estimates of ϵ were made for approximately 1.4 m depth intervals (2 s of data)

by integrating the velocity shear spectra in the wavenumber range of 2 to 20 cpm. Assuming isotropy of velocity fluctuations in this wavenumber band,

$$\epsilon = \frac{15}{2}\nu\left\langle\frac{u_z^2 + v_z^2}{2}\right\rangle \tag{1}$$

where ν is the kinematic viscosity of seawater, about 1.8×10^{-6} m^2 s^{-1} at these low temperatures, and $\langle\ \rangle$ denotes vertical averaging. The noise level based on measured microscale shears in the quietest regions appears to be about 10^{-9} W kg^{-1}, substantially higher than during the Arctic Internal Wave Experiment, AIWEX, (Padman and Dillon, 1987). We believe that this is due to a change in the dynamics of the RSVP, which in CEAREX was an air–filled, pressure–sealed case compared with an oil–filled instrument in AIWEX. The probable result of the latter change, which was compensated for by a reduction in the number and size of drag and buoyancy elements near the instrument's tail, is an increase in motion near the instrumented nose. However, ϵ in energetic mixing patches was 10^3 times greater than the noise level, so that the mixing processes which contribute most to the time–averaged evolution of the large–scale hydrographic fields were well resolved. The noise level was lowest in regions where the thermal gradients were smallest, notably in a warm, almost isothermal Atlantic Layer slab which was sampled at the end of the experiment. This suggests that the shear probes' thermal response, discussed by Osborn and Crawford (1980), may also contribute to setting the noise level on ϵ measurements.

An *upper* limit on fully resolved shear spectra is determined by the finite response length scale of the airfoil shear probes. The universal shear spectrum contains energy at all scales larger than the Kolmogorov microscale,

$$l_k = 2\pi(\nu^3/\epsilon)^{1/4} \tag{2}$$

although the spectral peak occurs at wavelengths of about $10l_k$. For $\epsilon = 1 \times 10^{-6}$ m^{2}s^{-3}, l_k is about 0.01 m, and the peak spectral density lies at wavelengths of about 0.1 m. Some correction can be made for the frequency response of the shear probes, however in all the data discussed in this paper the correction to ϵ was less than 20%, which is smaller than the potential calibration errors on the probes themselves. Therefore, in this paper, no spectral corrections have been applied.

The complete transect of ϵ averaged over 4 hours and 10 vertical meters is shown in Fig. 4. Time, t, is given throughout this paper in decimal day-of-year (UTC), where t=1.0 is 00:00 h on January 1,1989. High dissipation rates near the sur-

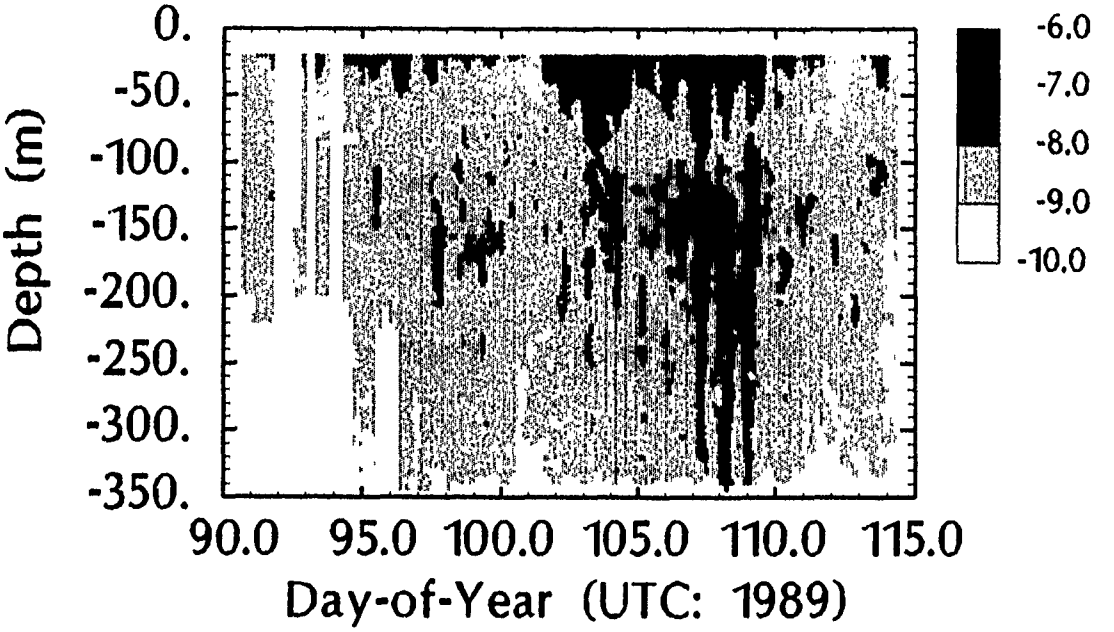


Fig. 4. Transect of dissipation rate, $\log_{10}(\epsilon)$, throughout the O Camp drift. Data have been averaged over 4 hours and 10 vertical meters.

face between t=101 and 110 are related to high surface stresses resulting from enhanced diurnal tidal currents which are a well-documented phenomenon in this region (Hunkins, 1986; PD91). Of more interest to the present paper are the large, quasi-diurnal bursts of high dissipation in the pycnocline (below about 100 m), which also occur during this period of strong diurnal tides. We shall discuss the possible causes of these bursts after reviewing the IW field which is assumed to initiate the mixing.

THE INTERNAL WAVE FIELD

The O Camp phase of CEAREX was exceptionally well set up for the study of the IW field (see Fig. 1). Simultaneous measurements were made of the hydrographic and current variability throughout the upper several hundred meters of the water column, encompassing almost all of the density stratification. A horizontal array provided information on the propagation direction of specific wave packets, while unique measurements of ice tilt and strain related to the passage of wave packets were also collected (P. Czipott, pers. comm., 1990). The ice camp drifted southward from the deep water of the Nansen Basin, up the plateau

slope, then approximately followed the 2000 m isobath towards the southwest (Fig. 2). Based on water depth, and the low frequency modulation of both ϵ and the apparently dynamically significant diurnal tide (PD91), the drift track is divided into four sections. Fig. 5 shows this division, with plots of water depth, the diurnal-band major axis current speed, and ϵ averaged over the pycnocline, $\langle \epsilon_{pyc} \rangle$. As this figure shows, $\langle \epsilon_{pyc} \rangle$ is substantially higher in Period 3 than elsewhere, although it is dominated by a few large events. Period 1 exhibits relatively low turbulence levels. However, as we shall show below, the time-averaged rate in this period is still much larger than the G89 model would suggest. D'Asaro and Morison (1991) provide an independent view of the spatial variability of mixing rates relative to topography in the eastern Arctic: their XCP measurements suggest that the wave shear is significantly higher over topography, but that the decay scale of the wave field is hundreds of kilometers. On that basis, even the deep Nansen Basin data in the present study (Period 1) lie within the radiation field associated with the Yermak Plateau.

The mean profiles of buoyancy frequency, N, dissipation rate, ϵ, and wave field energy density, E_{meas}, for the entire experiment are shown in Fig. 6. These profiles were made by first mapping the velocity field from the ADCPs to approximate isopycnal coordinates in order to reduce the vertical "smearing" by the large-amplitude diurnal oscillations. The resultant along-isopycnal averages are then plotted as a function of mean isopycnal depth. This "semi-Lagrangian" approach will be used throughout this paper. The energy density, E_{meas}, was calculated as

$$E_{meas} = \frac{1}{2} \left[\langle \eta^2 \rangle N^2 + \langle U^2 \rangle + \langle V^2 \rangle \right] \tag{3}$$

where

$$\langle \eta^2 \rangle = \int_f^{\omega_{nyq}} \Phi_\eta(\omega) d\omega \tag{4}$$

and

$$\langle U^2,\ V^2 \rangle = \int_f^N \left[\Phi_u(\omega),\ \Phi_v(\omega) \right] d\omega \tag{5}$$

$\Phi_\eta(\omega)$ is the power spectral density of RSVP isopycnal displacements, $\Phi_u(\omega)$, $\Phi_v(\omega)$ are the power spectral densities of orthogonal velocity components, and f and ω_{nyq} are the local inertial frequency ($\approx 1.45 \times 10^{-4}$ s^{-1}) and the Nyquist frequency ($\omega_{nyq} \approx 1$ cycle per hour for RSVP data) respectively. The vertical velocity variance is negligible compared with the horizontal components, and is therefore ignored in determining E_{meas}.

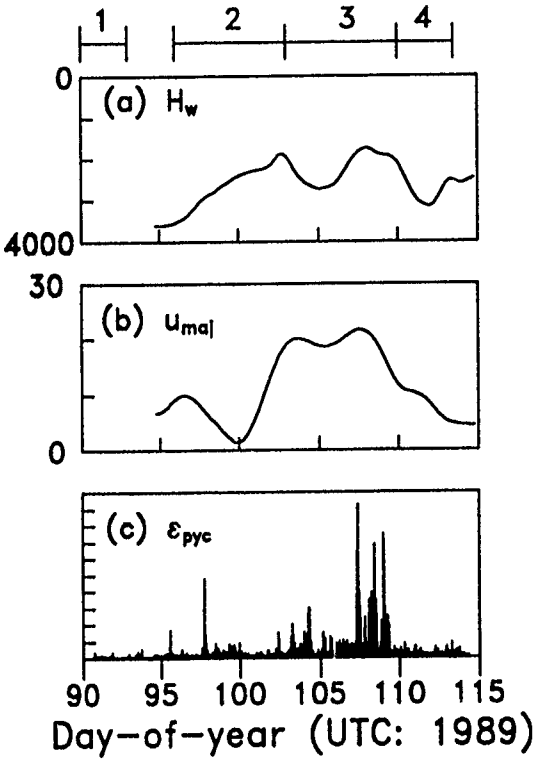


Fig. 5. (a) water depth (m), (b) diurnal-band major axis current (cm s^{-1}), and (c) dissipation rate (m^2s^{-3}) averaged through the pycnocline. Horizontal bars indicate the division of the experiment into 4 periods.

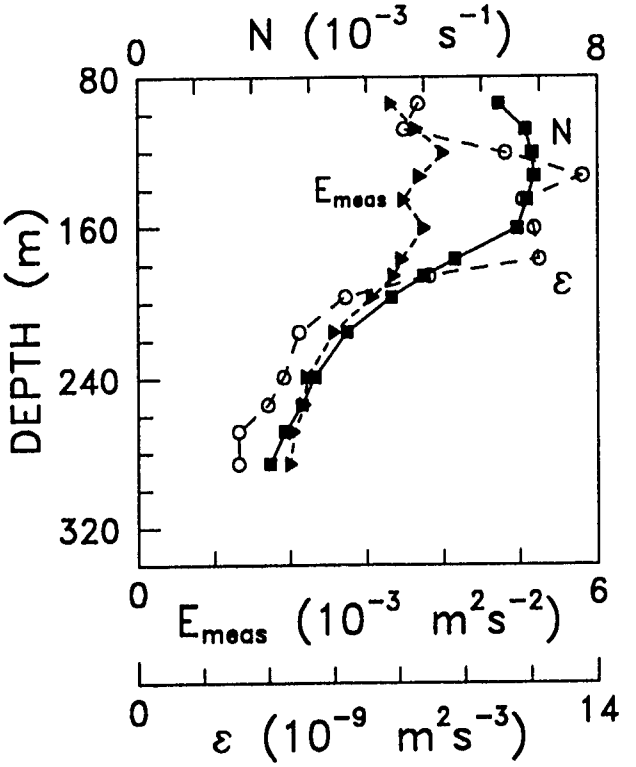


Fig. 6. Mean profiles of buoyancy frequency (N), dissipation rate (ϵ), and internal wave energy density (E_{meas}), for the period $96 < t < 114$.

The highest dissipation rates and energy densities occur in the upper, strongly stratified pycnocline, with a decay to lower values in the less strongly stratified region associated with the Atlantic Water core near 250 m. Based on these profiles, we consider the pycnocline to consist of 3 regions; the strongly stratified upper pycnocline, $100 < z < 170$ m; a transition zone, $170 < z < 220$ m; and the lower pycnocline, $220 < z < 270$ m. The wave field experienced by each region in each period varies widely, with commensurate changes in the mean dissipation rates.

Table 1 presents some basic characteristics of the IW field for each of the time and depth ranges. Buoyancy frequency changes by a factor of about two from the upper (U) to lower (L) pycnocline, however the mean profile of $N(z)$ changes very little from one period to the next. Except for the first period, however, the

energy density scaled by the N-scaled GM level varies significantly with depth, indicating that WKB scaling of the wave field is not valid. The measured dissipation rate varies from about 1 $\times 10^{-9}$ to 20 $\times 10^{-9}$m^{2}s^{-3}, however, values of less than 2 $\times 10^{-9}$m^{2}s^{-3} may be seriously biased by noise (see the description of the RSVP, above). The *rms* internal wave band strain, $\lambda^{IW}_{10,rms}$, is a measure of the IW-induced variability of N, and is discussed in more detail below.

Table 1: Mean properties of the internal wave field and dissipation rates for the four time periods shown in Fig. 5, and for mean isopycnal depth ranges of 100–170 m (U), 170–220 m (T), and 220–270 m (L). Tabulated parameters are: buoyancy frequency, $\langle N \rangle$ ($\times 10^{-3}$ s^{-1}); mean energy density, $\langle E_{meas} \rangle$ ($\times 10^{-3}$ m^{2}s^{-2}); mean energy density scaled by N-scaled Garrett-Munk energy density, E_o; *rms* internal-wave-band strain evaluated over 10-m mean isopycnal separation, $\lambda^{IW}_{10,rms}$; equivalent number of vertical modes, j_*; measured mean dissipation rate, ϵ_{meas} ($\times 10^{-9}$m^{2}s^{-3}); and decay time scale, $\tau = \langle E_{meas} \rangle / \epsilon_{meas}$ (days).

Period	Depth range	$\langle N \rangle$	$\langle E_{meas} \rangle$	$\langle E_{meas} \rangle / E_o$	$\lambda^{IW}_{10,rms}$	j_*	ϵ_{meas}	τ
1	U	6.7	1.12	0.30	–	6	2.0	6
	T	4.8	0.66	0.25	–	6	1.2	6
	L	2.9	0.41	0.26	–	6	1.6	3
2	U	6.7	2.23	0.60	0.37	7	7	4
	T	4.8	2.66	1.01	0.40	7	4	8
	L	2.9	2.24	1.40	0.48	7	2.0	13
3	U	6.7	5.79	1.56	0.43	4	20	3
	T	4.8	4.09	1.55	0.51	4	16	3
	L	2.9	3.95	2.48	0.53	4	9	5
4	U	6.7	4.16	1.13	0.36	5	6	8
	T	4.8	3.33	1.26	0.39	5	4	10
	L	2.9	3.54	2.22	0.32	5	2.0	20
GM		5.2	2.9	1.0	0.27	3	–	–

A particularly interesting aspect of data from the Arctic Ocean is the close proximity of the semi-diurnal tidal lines to the inertial frequency, f. For example, at 83° N, the frequency of the M_2 tidal line is about $0.97f$. Therefore, while the M_2 tide is formally a subinertial oscillation, the addition of relatively little negative

relative vorticity, $-|\zeta|$, to the planetary vorticity can result in the M_2 baroclinic tide becoming a free wave (*i.e.* $\omega > f - |\zeta/2|$). This effect is discussed in more detail in D'Asaro and Morison (1991). Perhaps more importantly, it becomes difficult or impossible to determine the free wave contribution from a near-inertial spectral peak, especially when measured from a non-stationary ice camp.

Fig. 7 shows frequency spectra for several relevant variables during the ice drift. These spectra are taken along an isopycnal with a mean depth near 150 m, near the ϵ peak, for the period $97.0 < t < 113.0$. The velocity spectrum is dominated by diurnal energy, with a secondary peak, as expected, near the inertial frequency. The cartesian coordinate shear spectrum shows that the diurnal tide is almost entirely barotropic, so that the near-inertial peak dominates. When the *magnitude* of the shear vector is analyzed (Fig. 7d) the dominant peak is again near-inertial, consistent with a shear vector that rotates with time rather than showing appreciable magnitude changes. Other spectra show the principal frequencies of variability in N, strain, and ϵ, and will be discussed further below. However, for the moment it is worth noting that the dominant time scales for variability of ϵ are diurnal and approximately 6 hours, the latter *not* corresponding to significant peaks in any of the obvious forcing functions.

With this degree of wave field variability in mind, we now consider higher order statistics of the displacement field, to see if there might be some obvious reason for the failure of the G89 model in this environment. One possible candidate is the effective vertical wavenumber bandwidth, β_*, which appears in the McComas and Müller (1981) and Henyey *et al* (1986) models (Eqs. 9 and 10, below). In practice, because β_* is expected to be a function of N, we consider instead the variability in the "equivalent number of vertical modes", $j_* = bN_o\beta_*/\pi N$, where N_o is a canonical buoyancy frequency ($0.0052\ \mathrm{s}^{-1}$), and b is the scale depth of the thermocline (1300 m). A typical mid-latitude value is $j_* = 3$.

Desaubies and Smith (1982) used an analytical approximation (Desaubies and Gregg, 1981) to the GM spectrum to show that shear variance was simply related to kinetic energy density by

$$\langle U_z^2 \rangle = \beta_* \beta_c \langle U^2 \rangle \tag{6}$$

where β_c is the cutoff wavenumber in the IW field, which Gargett *et al* (1981) suggested was constant at about $0.6\ \mathrm{m}^{-1}$. Desaubies and Smith argued that the statistics of mixing in a GM IW field are dependent only on the *rms* strain, λ_{rms}, where

$$\lambda_{rms}^2 = \langle (\partial\eta/\partial z)^2 \rangle \tag{7}$$

and η is an isopycnal displacement from its mean position. Desaubies and Smith speculated that the (at the time) apparent constancy of λ_{rms} in the open ocean

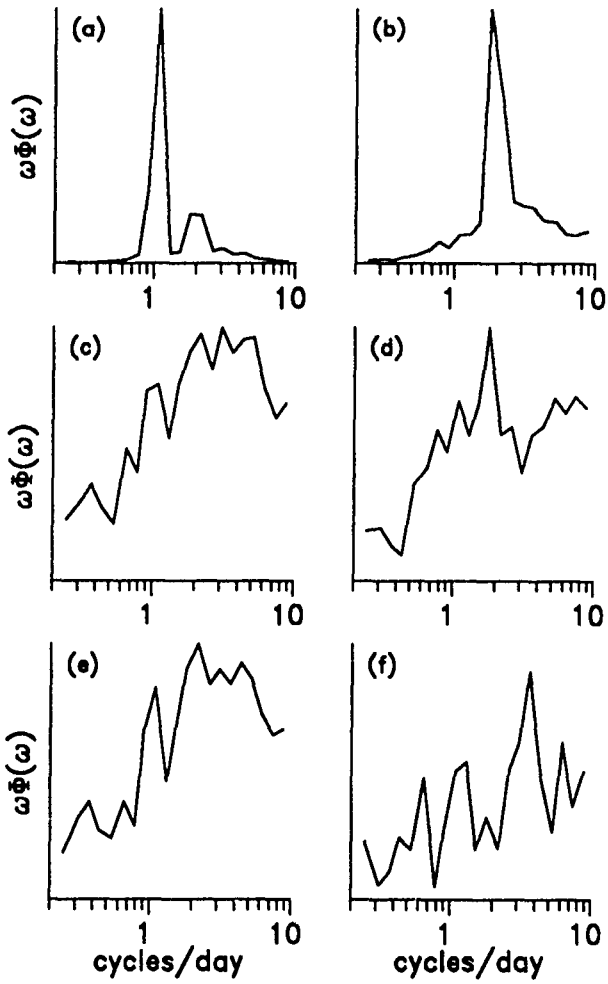


Fig. 7. Area-preserving frequency spectra of (a) currents, (b) shear components, (c) buoyancy frequency, (d) shear *magnitude*, (e) strain, and (f) pycnocline-averaged dissipation rate.

indicated that the ϵ was related to the saturation of the IW field at a level where breaking instabilities rapidly drained energy from the waves. Their numerical simulations indicated that ϵ was extremely sensitive to λ_{rms}, as well as being strongly influenced by the addition of near-inertial shear to the GM level. The overall result was a model which, while producing plausible diffusivities, was too sensitive to be usefully applied to field measurements. Nevertheless, the concept that strain might be an important parameter to monitor is valuable, and we shall show below that the CEAREX data set exhibits strain rates which are much larger than are typical in the mid-ocean thermocline.

As with other analyses in this paper, strain is measured in isopycnal-following coordinates. We have traced 100 isopycnals, having a mean separation of Δz = 1.7 m, and spanning the range of 100 to 270 m mean isopycnal depth, through the RSVP data set between t=96.0 and t=114.0. The original profiles were pre-filtered to 0.7 m resolution in order to minimize the impact of salinity spiking. A local value of strain is estimated as

$$\lambda_\delta = \frac{\eta(\overline{z_i}) - \eta(\overline{z_i} + \delta)}{\delta} \tag{8}$$

where δ is the mean separation of isopycnal pairs between which strain is being estimated. For consistency with G89 we will generally use λ_{10}. This choice of δ = 10 m also reduces the noise associated with interpolation from the equi-spaced (in depth) original profiles. The time series for finite-differenced strain between each isopycnal pair is then filtered to remove variability at frequencies below $0.8f$, most of which is due to compression of isopycnals as the barotropic diurnal tidal currents flow across the plateau slope. The *rms* strain is then simply the square root of the variance of λ_δ evaluated throughout the time and depth ranges of interest. The value of λ_{rms}, based on a 10 m mean isopycnal separation and after filtering, is denoted $\lambda^{IW}_{10,rms}$. Values of $\lambda^{IW}_{10,rms}$ are included in Table 1 for the three depth ranges and the latter three time periods (there were insufficient profiles in the first period to allow reliable isopycnal tracking). However, note that the true *rms* strain (Eq. 7) is very sensitive to the cutoff wavenumber of the wave field: the choice of β_c=0.63 m^{-1} is based on Gargett *et al* (1981), rather than being substantiated by the present data.

The values of $\lambda^{IW}_{10,rms}$ in Table 1 are all much larger than the 10 m *rms* strain expected for a GM ocean with j_* = 3. The dissipation rate which the Desaubies and Smith model therefore predicts is orders of magnitude higher than in a canonical ocean. There is a trend in periods 2 and 3 towards higher strain rates with decreasing N. Desaubies and Smith (1982) noted that, with a WKB-scaled wave field, $\lambda_{\delta,rms}$ should be independent of depth (and N). As we found with the N-dependence of energy density, this result suggests that WKB scaling is inappropriate to the present environment.

Another aspect of the wave field which deviates markedly from the GM76 assumptions is the presence of intermittent, but large-amplitude, wave packets. Fig. 8 shows the passage of one such packet as measured from a thermistor moored at 150 m. As suggested by PD91, there is some evidence that the mixing rates are correlated with these wave packets, which are in turn weakly correlated with the observed amplitude of the diurnal currents. One hypothesis is therefore that these waves are generated somewhere over the Plateau by the cross-shelf flow of the diurnal tide, then propagate towards deep water. While the generation mechanism itself is unclear, this scenario is supported by D'Asaro and

Morison's (1991) estimates of the decay scale of shear variance away from topography, and our own observations of the spatial dependence of ϵ from O Camp. Furthermore, both ice tilt measurements (P. Czipott, pers. comm., 1990) and estimates of phase propagation from the horizontal array of T, C and current sensors at O Camp indicate a cross-slope phase propagation towards deeper water. Coherent wave packets such as this violate the GM76 assumptions both of random phase between different waves, and isotropy. It is not clear how to take such packets into account, other than to re-run the HWF eikonal model with such waves included, however there is a strong possibility that such anomalous waves contribute substantially to the observed mean mixing rates.

The anisotropic nature of these near-N waves is also apparent in plots of the vertical coherence of semi-Lagrangian alongslope and cross-slope currents as a function of frequency (Fig. 9). While there is no anisotropy apparent in energy density when smoothed over a number of days, the cross-slope current is much more vertically coherent than the alongslope current. Our view is that this relatively high coherence arises from the passage of these near-N wave packets, even though the contribution of the packets to the time-averaged near-N energy density is small.

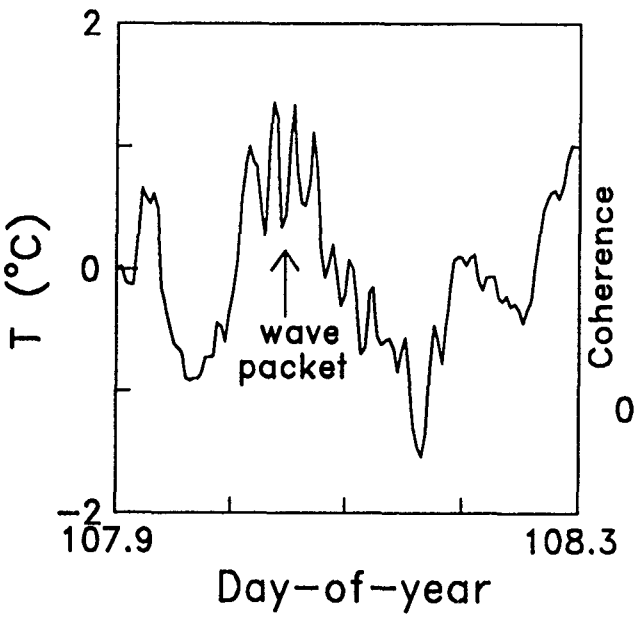


Fig. 8. Temperature at 150 m during the passage of a high-frequency wave packet past O Camp.

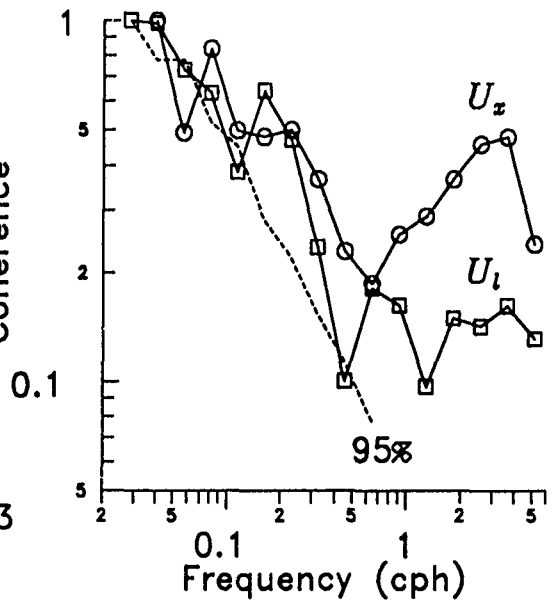


Fig. 9. Coherence, over 20 vertical meters, for cross-slope (U_x) and alongslope (U_l) currents in the pycnocline during the most energetic mixing period $103<t<110$. Dashed line indicates the 95% confidence level.

Finally, we are left not fully understanding the cause of the 6-hour peak in the area-preserving spectrum of ϵ (Fig. 7f). We have seen (Figs. 7c and 7d) that no such peaks exist in the density gradient or shear magnitude spectra. A cross-spectrum of shear *magnitude* and N shows no coherence at this frequency. However, a small but significant coherence is found near 4 cycles/day in the cross-spectrum of cross-slope shear with N, and similarly with along-slope shear and N. We postulate that the high frequency wave packets are in some way related to the 6-hour waves, and that the enhanced mixing at this periodicity is related to low Richardson number events which occur on time scales which are too short to be resolved by the present CTD data. Fig. 10 shows the band-passed (1/7 to 1/5 hour) isopycnal displacements plotted as a function of mean isopycnal depth, for the 5-day period when mixing was most energetic. There is evidence of both downward energy propagation (upward phase), particularly in the first 2 days, and upward energy propagation for $107 < t < 109.5$. The highest dissipation rates (Fig. 4) occur during the period of upward energy flux in this frequency band.

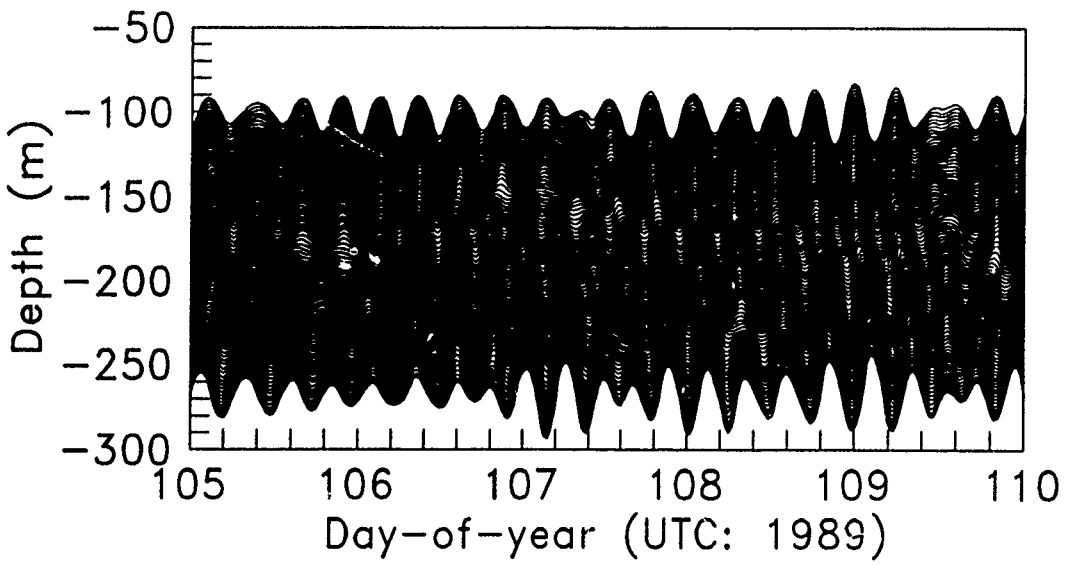


Fig. 10. Band-passed ($1/7 < \omega < 1/5$ cph) density field for the period $105 < t < 110$. For presentation, the mean density at each depth has been added to the band-passed signal.

Carrying this scenario further, we suspect that the 6-hour waves might be generated at the seabed as a higher harmonic of the cross-slope diurnal tidal current, propagate upwards into the main pycnocline, then create the near-N wave packets as a response to the rapidly increasing shear as the vertical wavelength is compressed by increasing N. This would be consistent with the highest values of ϵ being found *below* the depth of maximum N (see Fig. 6), since significant energy would presumably be lost from the upward-propagating waves in this region.

The proposed mechanism (above) is clearly highly speculative, and it is probable that it will remain so in the foreseeable future. The available bathymetry in this region is totally inadequate for an assessment of the bottom slope in any detail, while the buoyancy frequency profile of the deep ocean in this region is also inadequately sampled. It is therefore not possible to construct a robust model of wave reflection and propagation in this region. Nevertheless, we believe that all the data point towards a vertical anisotropy in the wave fluxes, at least within certain frequency bands, which again is a violation of the GM tenets.

DISSIPATION RATE MODELS

We have shown in the previous Section that the wave field during the CEAREX O Camp drift differs in many respects from the canonical GM wave field. It would not be surprising, therefore, if models based on the GM parameterization failed to adequately describe ϵ in such anomalous wave fields. Nevertheless, as was suggested in the Introduction, an approximate model might still prove valuable by overcoming the necessarily limited sampling domain of most microstructure and XCP surveys. We therefore consider some basic statistical properties which might allow existing models to be extended to cope with this new data set. We emphasize, however, that this effort does not imply increased deterministic knowledge of the processes involved in energy transfer through anomalous wave spectra. The aim is simply to look for a refined empirical model which might be useful while the wave field dynamics continue to be explored.

A complete discussion of existing models of ϵ related to properties of the IW field is given in Wijesekera *et al.* (1991). We simply provide the relevant equations against which the CEAREX data will be compared.

The model of McComas and Müller (1981) (hereinafter MM), which is based on weak resonant interactions between internal waves, is:

$$\epsilon_{MM} = \left(\frac{27\pi}{32\sqrt{10}} + 1\right) \pi^2 b^2 N^2 f j_*^2 E_{GM}^2 , \tag{9}$$

where E_{GM} $(= E/b^2 N N_o)$ is the GM dimensionless energy.

The model of Henyey, Wright and Flatté (1986) (hereinafter HWF) is based on an eikonal approach in which the nonlinear interactions are dominated by scale-separated interactions, with no effect on the large-scale background field. HWF predict a mean value of ϵ in a GM ocean of

$$\epsilon_{HWF} = \left[\frac{1.67}{\pi}\right] f j_*^2 b^2 N^2 E_{GM}^2 \, cosh^{-1}\left[\frac{N}{f}\right] \tag{10}$$

Note that the forms of ϵ_{MM} and ϵ_{HWF} are similar: as G89 noted, $\epsilon_{MM} \approx 7\epsilon_{HWF}$ for reasonable values of f and N.

Finally, the G89 model, expressed in terms of 10 m shear, S_{10}, is

$$\epsilon_{G89} = 7 \times 10^{-10} \frac{\langle N^2 \rangle}{N_0^2} \frac{\langle S_{10}^4 \rangle}{\langle S_{GM}^4 \rangle} \tag{11}$$

where $\langle S_{GM}^2 \rangle$ is the variance of shear at vertical scales greater than 10 m in the GM canonical, N-scaled ocean. If, as G89 assumed, the 10 m shear variance is simply related to the energy density by

$$\frac{S_{10}^2}{S_{GM}^2} = \frac{E_{IW}}{E_{GM}}, \tag{12}$$

where $E_{IW} = E_{meas}/b^2 N N_o$, then $\epsilon_{G89} \approx \epsilon_{HWF}$ for the mid-latitude data which he considered, provided GM canonical values of j_* (=3) and b (=1300 m), are used in calculating ϵ_{HWF}. G89 also investigated the f-dependence in the other models, but the test was inconclusive. The evaluation of S_{10}^2 from the O Camp ADCP data is difficult: Wijesekera *et al.* (1991) attempt to quantify the shear variance which is not resolved as a result of the vertical averaging scales for currents and finite-differencing to obtain shear; however the result is very sensitive to the choice of cutoff wavenumber, β_c, and the slope of the shear spectrum at the unresolved wavenumbers. For comparison with the measured dissipation rates, we therefore introduce an "energy-dependent" Gregg model, ϵ_E, based on (11) and (12):

$$\epsilon_E = 7 \times 10^{-10} \frac{\langle N^2 \rangle}{N_0^2} \frac{E_{IW}^2}{E_{GM}^2} \tag{13}$$

The essential difference between Gregg's scaling model and the MM and HWF models is that Gregg's model requires measurements of the velocity shear, while the MM and HWF models require estimates of the wave field energy density, E_{meas}, and the vertical wavenumber bandwidth, expressed as an equivalent number of vertical modes, j_*. Since energy density is relatively easily obtained from current meter moorings, the problem is therefore in estimating j_*. Gregg's assumption that j_*=3 is clearly not universally true: previous work in the Canada Basin (Levine, 1990; D'Asaro, pers. comm., 1990) indicate that $j_* \approx 30 - 60$ in that region. The Canada Basin is a site where E_{meas} is only about 10% of GM, but the shear variance associated with the greater number of equivalent modes is much closer to mid-latitude levels. We estimate j_* in the present data by calculating the correlation, r, between pairs of isopycnals with a mean separation of about 10 m. Then, for $\beta_c \gg \beta_*$, $\beta_* \approx -\ln(r)/\delta$: δ = 10 m (from Desaubies and Smith (1982), Eq. 16). The value of j_* is then given by $bN_o\beta_*/\pi N$. The

measured vertical coherences indicate that $j_* \approx 6$ during CEAREX, although it actually varies from about 4 to 7 between periods (Table 1). The MM and HWF models therefore predict values of ϵ which are about 4 times greater than for $j_* = 3$. The f-scaling in those two models also implies dissipation rates which are a factor of two higher in the Arctic than at mid-latitudes.

It is not known yet why j_* in the Arctic is larger than in the deep, mid-latitude oceans. Nevertheless, if an experiment is somehow able to estimate j_*, then predictions of ϵ could be made based on both E and j_*. Table 2 shows how the three models compare with the measured averages for the same time and potential density ranges discussed above. In the mean, the energy-based, modified Gregg model underestimates ϵ_{meas} by about a factor of 6, although a factor of two improvement occurs if the prediction is f-scaled. MM overpredicts ϵ_{meas} by a factor of 3, while HWF underpredicts by a factor of 2.3. However, there is a large scatter in predictive ability for each model between different time and depth ranges.

Table 2: Measured mean dissipation rate, ϵ_{meas} ($\times 10^{-9} \mathrm{m^2 s^{-3}}$); and model predictions scaled by ϵ_{meas} for the time and depth ranges shown in Table 1.

Period	Depth range	ϵ_{meas}	ϵ^*_{MM}	ϵ^*_{HWF}	ϵ^*_E	ϵ^*_λ
1	U	2.0	0.55	0.07	0.06	–
	T	1.2	0.33	0.04	0.03	–
	L	1.6	0.09	0.01	0.01	–
2	U	7	2.0	0.29	0.06	0.21
	T	4	5.0	0.73	0.15	0.73
	L	2.0	7.0	1.0	0.21	2.1
3	U	20	1.55	0.22	0.14	0.90
	T	16	0.94	0.13	0.09	1.13
	L	9	1.56	0.22	0.14	2.11
4	U	6	4.17	0.60	0.25	0.77
	T	4	4.0	0.58	0.24	1.03
	L	2.0	9.0	1.3	0.53	1.05
Mean	–	–	3.02	0.43	0.16	1.02

The final column in Table 2 is an empirical prediction of dissipation rate, based on the assumption that the higher strain rates are indicative of higher shear and/or greater probability of instabilities occurring in the wave field. We assume that the strain-based predicted dissipation rate, ϵ_λ, is given by

$$\epsilon_\lambda = 7 \times 10^{-10} \frac{\langle N^2 \rangle}{N_0^2} \left(\frac{\lambda^{IW}_{10,rms}}{\lambda^{DS}_{10,rms}} \right)^4 \tag{14}$$

i.e. functionally similar to the G89 version (Eq. 11), with *rms* strain replacing shear. The 10-m strain in the GM canonical wave field, following Desaubies and Smith (1982) is $\lambda^{DS}_{10,rms} \approx 0.27$. The mean predictive ability of this empirical model is exceptionally good. We emphasize that there is little theoretical justification for this model, which is a hybrid of Gregg's scaling arguments and Desaubies and Smith's strain-based numerical simulations. In a GM wave field, shear is predominantly near-inertial, while the peak in the strain spectrum is shifted towards higher frequencies. In the present data set, the internal wave continuum, $\omega \gg f$, is amplified more than the near-inertial band near the plateau, so that the ratio of *rms* strain to *rms* shear is increased. Our results suggest that variability of N in the present data set will therefore be more important to the occurrence of the low Richardson number events (which we believe lead to turbulence) than in the mid-ocean thermocline, where variability of shear is most significant. Since both the G89 and Desaubies and Smith models assume a fixed relationship between *rms* shear and strain, it is impossible to determine whether strain should be the correct variable to track.

With the above caveats in mind, however, the apparent success of strain-based scaling of dissipation (assuming that the data sets analyzed by G89 all have canonical strain rates), suggests that dissipation rates can be estimated with some success from measurements of IW energy density and strain. A mooring with adequate vertical and temporal resolution of the density field might therefore be used to obtain dissipation rates on sufficiently large vertical and temporal scales, say 50 m and several days. These requirements would be most easily met by sufficiently rapid automatic yo-yo CTD devices with resolutions of about 1 m and a sampling rate greater than about 1 cycle/hour. The vertical span of the profile would need to encompass the upper and lower bounds of isopycnal excursions, which have *rms* vertical motions of $O(10)$ m in a GM ocean.

SUMMARY

Our studies of simultaneous measurements of IW properties and microscale dissipation rates in a region where many of the GM assumptions do not hold (although $E_{IW}/E_{GM} = O(1)$), leads us to the following conclusions.

(1) Large values of ϵ are found where the wave field is not well modeled by a Garrett-Munk spectrum, even where the energy density is comparable to the canonical values.

(2) Some modelling success might still be achievable in these regions, provided the vertical wavenumber bandwidth, β_*, can be estimated (perhaps from vertical and/or horizontal coherences in either the velocity or hydrographic fields). Automated yo-yo CTD profilers and a vertical current meter array appears to be the most likely technique for acquiring the required data from moorings.

We believe that it is still too early to claim even a "first-order" understanding of the processes by which energy is transferred through the IW spectrum to the dissipation scales. There is, however, some prospect that sufficiently robust empirical models can be developed to allow predictions of ϵ from long-term moorings, at least to within the accuracy obtainable by extrapolating the estimates from short-duration microstructure programs.

ACKNOWLEDGMENTS

We wish to express our appreciation for the efforts and hospitality of Peter Müller and Crystal Miles, who organized the 'Aha Huliko'a workshop. Discussions with several participants, notably Eric Kunze, Mike Gregg, and Greg Holloway, have been valuable to this study. This work has been sponsored by ONR, contracts N00014-87-K-0009 and N00014-90-J-1042 (LP, TM, HW, ML, CP), and N00014-90-J-1099 (RP).

REFERENCES

Baker, M.A. and C.H. Gibson, 1987: Sampling turbulence in the stratified ocean: statistical consequences of strong intermittency. *J. Phys. Oceanogr.*, **17**, 1817–1836.

Caldwell, D.R., T.M. Dillon and J.N. Moum, 1985: The Rapid Sampling Vertical Profiler. *J. Atmos. Oceanic Technol.*, **2**, 615–625.

D'Asaro, E.A., and J.H. Morison, 1991: Internal waves and mixing in the Arctic Ocean, *Deep-Sea Res.*, in press.

Desaubies, Y.J.F. and M.C. Gregg, 1981: Reversible and irreversible finestructure. *J. Phys. Oceanogr.*, **11**, 541–556.

Desaubies, Y., and W.K. Smith, 1982: Statistics of Richardson number and instability in oceanic internal waves. *J. Phys. Oceanogr.*, **12**, 1245–1259.

Gargett, A.E., P.J. Hendricks, T.B. Sanford, T.R. Osborn, and A.J. Williams III, 1981: A composite spectrum of vertical shear in the upper ocean. *J. Phys. Oceanogr.*, **11**, 1258–1271.

Gregg, M.C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.*, **94**, 9686–9698.

Gregg, M.C., and T.B. Meagher, 1980: The dynamic response of glass rod thermistors. *J. Geophys. Res.*, **85**, 2779–2786.

Gregg, M.C., J.C. Shedvin, W.C. Hess and T.B. Meagher, 1982: Dynamic response calibration of the Neil Brown conductivity cell. *J. Phys. Oceanogr.*, **12**, 720–742.

Henyey, F.S., J. Wright, and S.M. Flatté, 1986: Energy and action flow through the internal wave field: an eikonal approach. *J. Geophys. Res.*, **91**, 8487,-8495.

Hunkins, K., 1986: Anomalous diurnal tidal currents on the Yermak Plateau. *J. Marine Res.*, **44**, 51–69.

Levine, M.D., 1990: Internal waves under Arctic ice pack during the Arctic Internal Wave Experiment: the coherence structure. *J. Geophys. Res.*, **95**, 7347,-7357.

McComas, C.H., and P. Müller, 1981: The dynamic balance of internal waves. *J. Phys. Oceanogr.*, **11**, 970–986.

McPhee, M.G., 1988: Analysis and prediction of short-term ice drift. *J. Offshore Mechanics and Arctic Engineering*, **110**, 94–100.

Osborn, T.R. and W.R. Crawford, 1980: An airfoil probe for measuring turbulent velocity fluctuations in water. In: *Air-Sea Interaction: Instruments and Methods*, Plenum Press, New York, 801 pp.

Padman, L., and T.M. Dillon, 1987: Vertical heat fluxes through the Beaufort sea thermohaline staircase, *J. Geophys. Res.*, **92**, 10799–10806.

Padman, L., and T.M. Dillon, 1991: Turbulent mixing near the Yermak Plateau during the Coordinated Eastern Arctic Experiment. *J. Geophys. Res.*, in press.

Wijesekera, H.W., L. Padman, T.M. Dillon, M.D. Levine, C.A. Paulson, and R. Pinkel, 1991: A comparison of internal wave dissipation models with CEAREX observations. (in preparation).

TESTING THE CRITICAL REFLECTION HYPOTHESIS

Denis Gilbert
Institute of Ocean Sciences, P.O. Box 6000, Sidney, B.C., Canada, V8L 4B2.

ABSTRACT

According to linear internal wave theory, the reflection of internal waves off a bottom of uniform slope, in a uniformly rotating, uniformly stratified fluid, should lead to energy enhancement and a cross-isobath alignment of motions at the frequency for which the wave ray slope equals the bottom slope. Current meter data from the near-equatorial Atlantic and from the continental rise and slope off Nova Scotia are used to test this hypothesis.

1 INTRODUCTION

The reflection of internal waves off a plane rigid surface differs markedly from the reflection of electromagnetic or acoustic waves. In Optics or Acoustics, the incident and reflected wave rays make the same angle with respect to the normal to the reflecting surface, whereas for internal waves, the incident and reflected wave rays make the same angle with respect to the vertical. The unusual nature of the law of reflection for internal waves is a direct consequence of their dispersion relation which states that, for waves of a given frequency, energy must propagate at a given angle with respect to the vertical. As linear theory requires frequency to be conserved upon reflection, the angles which the incident and reflected wave rays make with respect to the vertical must be the same.

Close to the critical frequency ω_c for which the wave ray slope equals the bottom slope, simple arguments (Phillips 1977, p.227) show that, upon reflection, the wavenumber, energy density, and shear associated with the incident waves are greatly amplified, so that shear instability and energy dissipation are more likely. Phillips (1963) first pointed this out for inertial waves incident on a bottom of constant slope in a rotating, homogeneous, inviscid fluid, and later generalised the theory to internal gravity waves in Phillips (1966). In both cases he assumed that the direction of energy propagation of the incident waves was normal to the isobaths.

For arbitrary angle of incidence with respect to the isobaths, Sandstrom (1966) pointed out that whereas the component of the wavenumber vector normal to the isobaths is generally amplified upon reflection, the component of the wavenumber vector parallel to the isobaths remains unchanged. Consequently, the major axis of the current ellipses should become more closely aligned with the cross-isobath direction upon reflection. Sandstrom (1966, p.78) was the first to report observational evidence of energy enhancement at ω_c, using thermistor data from the Bermuda slope (Haurwitz, Stommel and Munk, 1959) [1]. He also performed laboratory experiments which clearly

[1] Wunsch (1972) also used that data and drew attention to the energy enhancement at ω_c.

demonstrated the amplification of parcel motions that results from internal wave reflection off a sloping bottom (figure 21, p.71), and published that work in Sandstrom (1969).

The most recent surge of interest for the problem of internal wave reflection off sloping bottoms is largely due to the work of Eriksen (1982), who presented observational evidence of energy and shear enhancement near the critical frequency at a few mooring sites, and also provided evidence for a cross-isobath alignment of motions near ω_c in a few of these cases. In a later paper, Eriksen (1985) explored the implications for ocean mixing of internal wave reflection off sloping bottoms. He argued that internal wave breaking at sloping boundaries may cause diapycnal mixing of global oceanic significance, possibly accounting for Munk's (1966) canonical value of $K_v \approx 10^{-4}\mathrm{m}^2\mathrm{s}^{-1}$ in the abyssal ocean. This possibility was further examined by Garrett and Gilbert (1988).

In section 2 of this paper, we give a brief summary of the linear theory of internal wave reflection off a bottom of constant slope, in a uniformly rotating, uniformly stratified fluid. In section 3, we examine current meter data from the Western Boundary Sill Experiment to try to verify whether a near-bottom enhancement of motions with 3–4 day periods can be attributed to internal wave reflection, as Eriksen (1982) suggested. In section 4, current meter data from the continental rise and slope off Nova Scotia are used in order to look for evidence of energy enhancement and/or cross-isobath alignment of motions near ω_c. A summary and general discussion follow in section 5.

2 THEORY

The dispersion relation for internal waves is given by (Gill, 1982)

$$\omega^2 = N^2 \cos^2\theta + f^2 \sin^2\theta \tag{1}$$

$$\text{or} \qquad \tan\theta = \left(\frac{\mathrm{N}^2 - \omega^2}{\omega^2 - \mathrm{f}^2}\right)^{1/2}, \tag{2}$$

where N is the buoyancy frequency, f is the inertial frequency, and θ is the angle which the wavenumber vector $\mathbf{k} = (k, \ell, m)$ makes with respect to the horizontal. Since the direction of energy propagation is perpendicular to the direction of phase propagation ($\mathbf{c_g} \perp \mathbf{k}$) for internal waves, the frequency at which the wave ray slope matches a bottom slope of $\tan\alpha$ can be obtained by substituting $\theta = \pi/2 - \alpha$ into (1), yielding

$$\omega_c^2 = N^2 \sin^2\alpha + f^2 \cos^2\alpha. \tag{3}$$

It can be shown (Eriksen, 1982, Gilbert, 1990) that for constant values of f, N and $\tan\alpha$, the wavenumber amplification is given by

$$\frac{m_r}{m_i} = \frac{a^2 + 2a\cos\phi_i}{a^2 - 1} + 1, \qquad a = \tan\alpha \tan\theta, \tag{4}$$

where a is a frequency-dependent parameter (2), ϕ_i is the angle which the horizontal component of $\mathbf{k_i}$ makes with the onslope direction, and m_i, m_r denote the vertical components of the incident and reflected wavenumber vectors respectively. By requiring that $\mathbf{u} \cdot \hat{\mathbf{n}} = 0$ at the bottom, where $\hat{\mathbf{n}}$ is a unit vector normal to the bottom, it can also be shown that (Eriksen, 1982)

$$\frac{E_r}{E_i} = \left(\frac{m_r}{m_i}\right)^2, \tag{5}$$

$$\text{and} \qquad \phi_r = \sin^{-1}\left(\left|\frac{m_i}{m_r}\right| \sin\phi_i\right), \tag{6}$$

where E_i, E_r are the energy density of the incident and reflected waves respectively, and ϕ_i, ϕ_r are the angles which the horizontal components of $\mathbf{k_i}$, $\mathbf{k_r}$ make with the onslope direction. There is a singularity in (4) when $a \to 1$ ($\omega \to \omega_c$), so that linear inviscid theory (5) predicts we should observe enhanced energy density near ω_c above sloping bottoms. It also predicts (6) a cross-isobath alignment of motions near ω_c ($|m_i/m_r| \to 0$ and hence $\phi_r \to 0$ or π when $a \to 1$).

3 DATA FROM THE WESTERN BOUNDARY SILL EXPERIMENT

To measure the flow of Antarctic Bottom Water entering the Northwest Atlantic Basin, the buoy group of the Woods Hole Oceanographic Institution deployed two moorings between the Ceara Rise and the mid-Atlantic ridge from December 9, 1977 to December 5, 1978 (Whitehead and Worthington, 1982, hereafter WW82). Mooring 636 was deployed at 4°2.5′N, 39°40.6′W, and mooring 637 was deployed at 4°1.3′N, 39°19.0′W. Both moorings had current meters at 10m, 50m, 100m and 200m above the bottom. The local depth was 4456m at mooring 636, and 4304m at mooring 637 (see Fig. 1).

Eriksen's (1982) most convincing evidence of energy enhancement near ω_c came from mooring 636 of this experiment. This is shown here on Fig. 2, where we see that the near-bottom energy enhancement ranges roughly from 0.005 cph to 0.05 cph (periods between 20 and 200 hours), with a peak in energy density centered at 0.0117 cph (85.5 hour or 3.56 day period). The inertial frequency at that location is 0.00587 cph (170.3 hour period), so that the energy peak occurs well within the internal waveband at $\omega \sim 2f$.

Following Eriksen (1982) in taking $N \approx 0.8$ cph, and assuming that the bottom slope at mooring 636 is 0.015, as estimated from the East-West transect shown on Fig. 1, we obtain $\omega_c = (N^2 \sin^2\alpha + f^2 \cos^2\alpha)^{1/2} \approx 0.0133$ cph. This estimate of the critical frequency compares well with the energy enhancement shown on Fig. 2.

However, Eriksen (1982, p. 533) pointed out that while the energetic motions with $\omega \approx 0.012$ cph at mooring 636 could be due to critical internal wave reflection, they violate one of the basic predictions of specular reflection theory (6), namely that the major axis of the current ellipses should be oriented normal to the isobaths. Instead the current ellipses are roughly parallel to the inferred North-South orientation of the isobaths

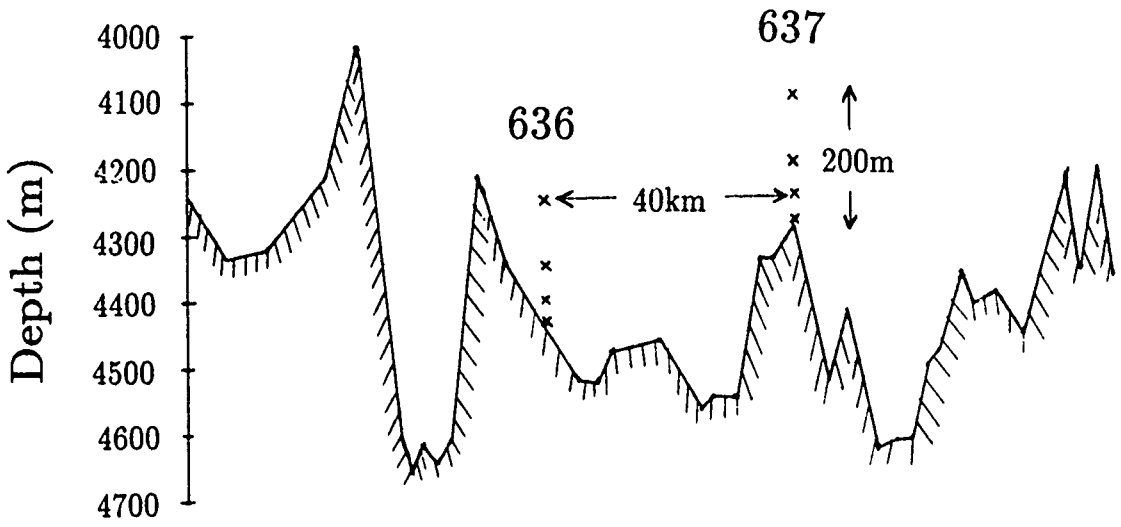


Figure 1: Topography in the vicinity of moorings 636 and 637 of the Western Boundary Sill Experiment, based upon the East-West transect at $4°N$ shown on figure 4 of Whitehead and Worthington (1982).

(see Eriksen's figure 9). This is not a trivial discrepancy, as the mechanism leading to energy enhancement near the critical frequency involves a 'squeezing' of the reflected wavebeam (relative to the width of the incident wavebeam) in the vertical plane normal to the isobaths (Phillips, 1977, figure 5.13). It is hard to envisage how internal wave reflection could lead to enhanced motions in a vertical plane roughly parallel to the isobaths. Further examination of the current meter data from the Western Boundary Sill Experiment therefore seems warranted.

Comparison with mooring 637

Figure 3 shows the kinetic energy spectra 50m[2], 100m and 200m above the bottom at mooring 637. A first obvious difference in the kinetic energy spectra is that an inertial peak ($f = 5.87 \times 10^{-3}$ cph) appears to be present at mooring 637, whereas it could not be discerned at mooring 636 (Fig. 2). A second difference is that while mooring 637 shows some near-bottom kinetic energy enhancement at $\omega = 0.0117$ cph, the energy levels are not nearly as elevated as at mooring 636; comparing the 50m records from each mooring, we find that the spectral level at mooring 637 is about 5 times lower than at mooring 636. The range of frequencies over which we observe near-bottom kinetic energy enhancement is also narrower at mooring 637.

Nonetheless, since moorings 636 and 637 both display near-bottom energy enhancement at $\omega = 0.0117$ cph, it seems worthwhile to find out whether motions at the two moorings are

[2]Data from the current meter 10m above the bottom were judged to be of lesser quality by NODC (National Oceanographic Data Center) and so were not sent to us.

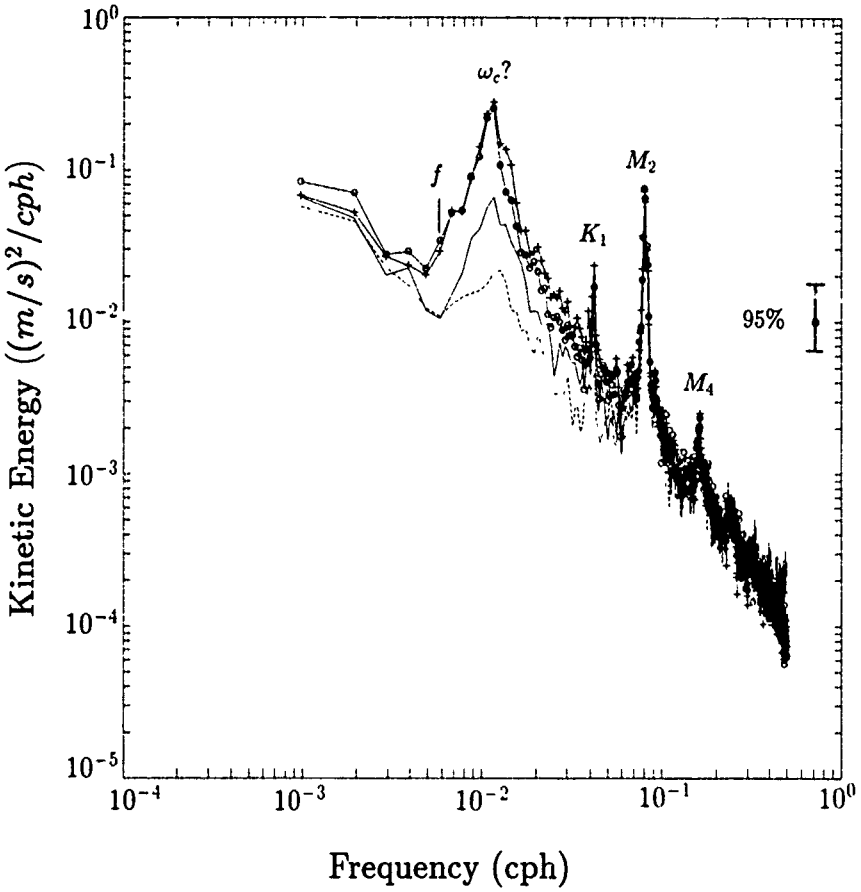


Figure 2: Kinetic energy spectra at 10m (+), 50m (o), 100m (—) and 200m (- - -) above the bottom at mooring 636 ($\nu = 30$). The inertial frequency is $f = 5.87 \times 10^{-3}$ cph.

correlated at that frequency. Figure 4 shows cross-spectra between the current meter 200m above the bottom at mooring 636, and the current meter 50m above the bottom at mooring 637 (4256m and 4254m deep respectively).

For the clockwise velocity signals, there is a strong coherence peak at the semidiurnal tide (not shown here), a few weak peaks near the diurnal tide, and another peak spreading across four adjacent frequency bins from 0.0117 cph to 0.0146 cph, where the signals are not significantly out of phase. The latter coherence peak coincides perfectly with the range of frequencies with the most kinetic energy within the internal waveband 50m above the bottom at mooring 637 (Fig. 3). The temperature signals are also correlated in the 0.0137–0.0146 cph band, where they are not significantly out of phase.

Thus Fig. 4 shows that for $\omega \approx 0.012$ cph, the motions at mooring 636 are coherent and not significantly out of phase with those at mooring 637, some 40 km away. Significant coherences within the internal waveband over horizontal distances of the order of a few tens of kilometers are rare but not unheard of in the deep ocean. For example, Fu (1981)

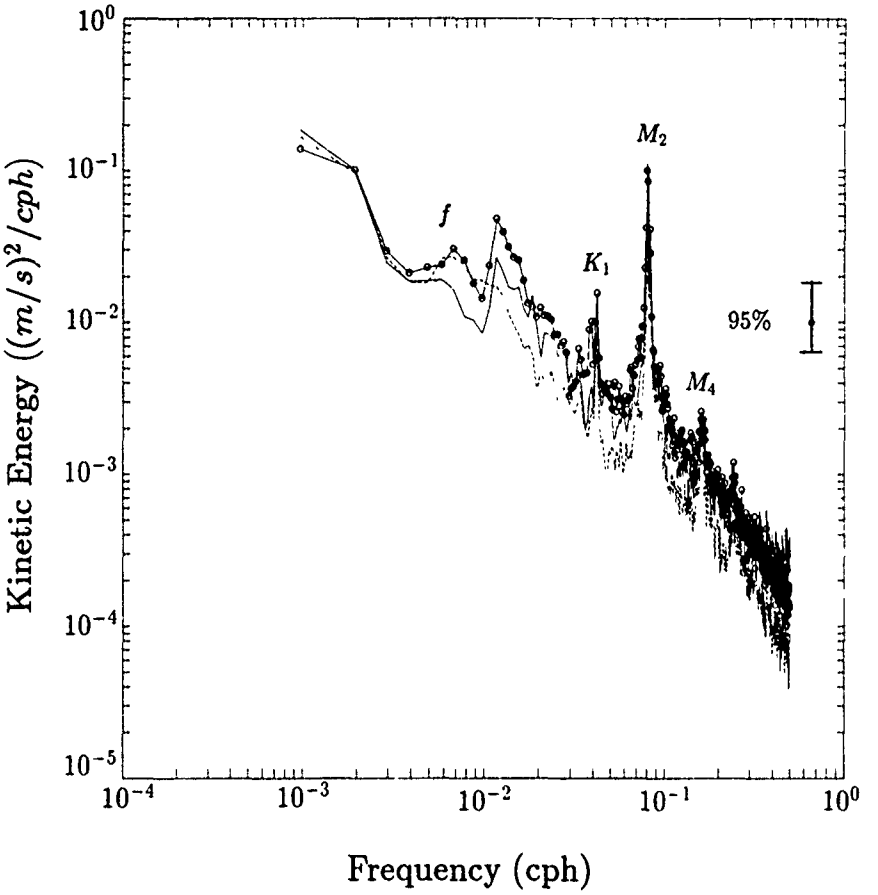


Figure 3: Kinetic energy spectra 50m (o), 100m (—) and 200m (- - -) above the bottom at mooring 637 ($\nu = 30$).

obtained significant coherences over horizontal separations up to 70 km at mid-latitudes for the inertial band. For the mid-latitude internal wave model proposed by Garrett and Munk (1972), the horizontal distance at which the coherence drops to 0.5 is given by

$$\Delta X_{0.5} = \frac{58 \text{ m} \cdot 3 \text{ cph}}{(\omega^2 - f^2)^{1/2}}, \tag{7}$$

where ω and f are expressed in cph. For $\omega = 1.17 \times 10^{-2}$ cph and $f = 5.87 \times 10^{-3}$ cph, this yields $\Delta X_{0.5} = 17$ km, so that the coherence should be less than 0.5 (or the squared coherence should be less than 0.25) at a separation of 40 km. Wunsch and Webb (1979) showed that at low latitudes, currents are generally coherent over smaller spatial scales than would be predicted by (7), so that the observed coherence is rather surprising. It becomes even more puzzling when we consider the fact that based on the critical reflection hypothesis, the observed field of motions, if dominated by the reflected wavefield, should

be unusually rich in motions with small horizontal and vertical scales [3] near the critical frequency (4). Cross-spectra between various pairs of current meters at mooring 636 suggest that the dominant vertical scales of motion at $\omega \approx 0.12$ cph are also large (Gilbert, 1990).

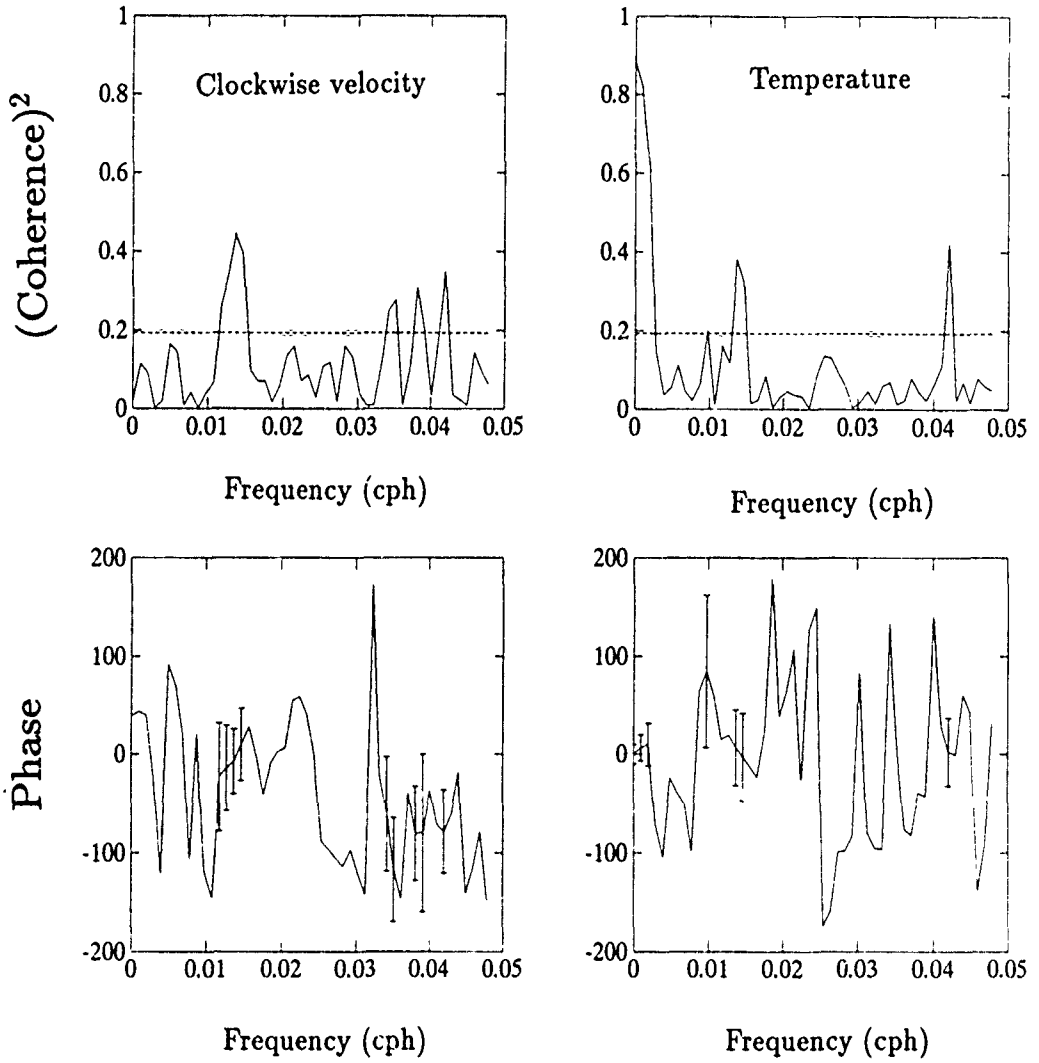


Figure 4: Cross-spectra between the current meter 200m above the bottom at mooring 636, and the current meter 50m above the bottom at mooring 637, both of which were at the same depth. The cross-spectrum for clockwise velocity is shown on the left, and that for temperature is shown on the right. A positive phase means that the signal at mooring 637 leads the signal at mooring 636. The dashed line represents the 95% significance level for zero true coherence on the upper plots, and the 95% confidence intervals for phase are shown on the lower plots ($\nu = 30$).

[3]The horizontal and vertical scales of motion should be reduced by the same factor upon reflection, because conservation of frequency implies that $\tan\theta_r = \tan\theta_i$, which in turn implies that $m_r/m_i = \kappa_r/\kappa_i$, where m and κ are the vertical and horizontal components of the wavenumber vector respectively

Rotary spectra and the ratio P_{++}/P_{--}

Using the clockwise (u_-) and anticlockwise (u_+) rotary velocity components

$$u_- = (u - iv)/\sqrt{2}, \tag{8}$$
$$u_+ = (u + iv)/\sqrt{2} \tag{9}$$

as in Müller et al. (1978), we computed the clockwise and anticlockwise autospectra P_{--} and P_{++} at each current meter. For linear, freely-propagating internal waves, the ratio of anticlockwise to clockwise kinetic energy should be given by (Fofonoff, 1969)

$$\frac{P_{++}}{P_{--}} = \left(\frac{\omega - f}{\omega + f}\right)^2. \tag{10}$$

On Fig. 5, we plotted P_{++}/P_{--} as a function of frequency at 10m, 50m, 100m and 200m above the bottom at mooring 636. Interestingly enough, we find large deviations from (10) in the range of frequencies for which near-bottom kinetic energy enhancement is observed. More interesting still, we find that in the neighbourhood of the energy peak ($\omega \approx 0.012$ cph), the largest departure from (10) occurs at 10m above the bottom, followed by the departures at 50m, 100m and 200m in that order. There thus seems to be a one to one correspondance between kinetic energy enhancement and the degree of departure from (10) near $\omega \approx 0.012$ cph. Eriksen (1982, p.533) hinted at that when he mentioned that "the current ellipses at $\omega \sim 2f$ are more narrow than would be expected from linear internal waves, even if waves were unidirectional."

Trapped waves with $\omega > f$?

Thompson and Luyten (1976) have provided evidence for the existence of Rhines' (1970) bottom-trapped waves at frequencies lower than the inertial frequency f. In the present context however, what is interesting about those waves is that the maximum allowed frequency, $N \sin\alpha \approx 0.8$ cph $\times$ 0.015 $\approx$ 0.012 cph, is larger than f, so that bottom-trapped motions could in principle occur in the frequency band normally reserved to freely-propagating internal waves ($f < \omega < N$). The possibility of bottom-trapped waves with $\omega > f$ has not received much attention in the literature, but Huthnance (1989) points out that "continuity of mode forms near $\omega = f$ (Huthnance, 1978) suggests that trapped waves approaching $\omega = f$ from lower frequencies should continue as nearly-trapped waves for ω exceeding f. The bottom trapped waves of Rhines (1970) in realistic contexts are obvious candidates." Huthnance and Baines (1982) called for a more detailed investigation of the phenomenon, pointing out that we do not know much about the possible radiational energy losses of those waves when $N \sin\alpha > f$.

The bottom-trapped motions described by Rhines (1970) are rectilinear, implying the ratio of anticlockwise to clockwise kinetic energy

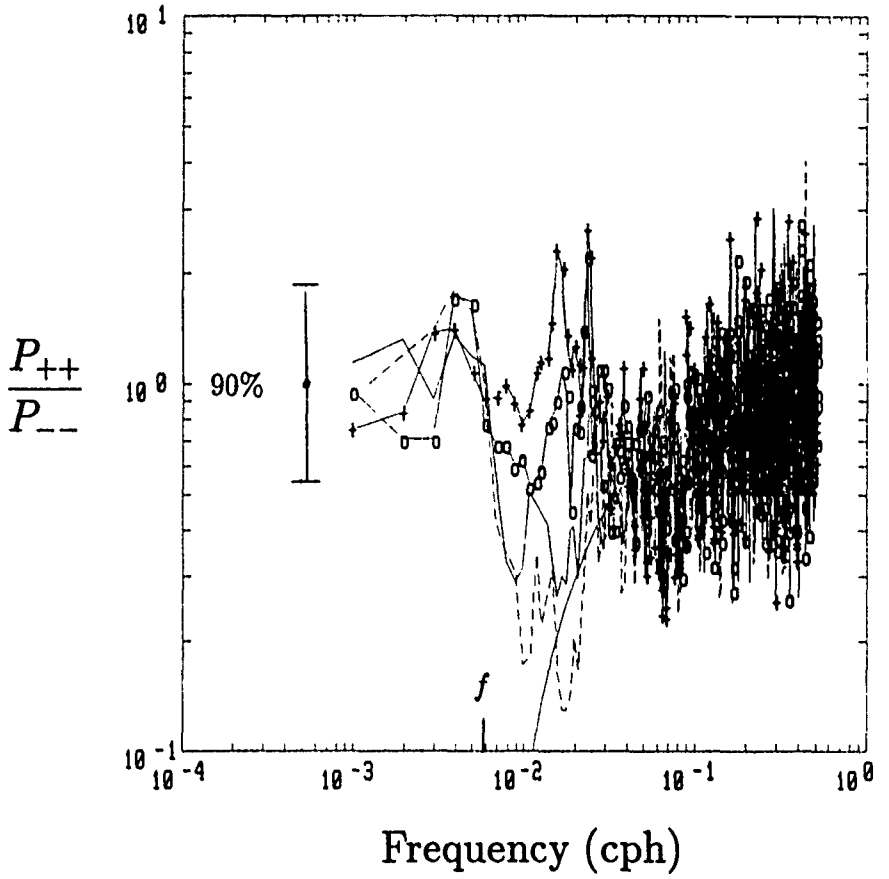


Figure 5: Ratio P_{++}/P_{--} at 10m (+), 50m (o), 100m (—) and 200m (- - -) above the bottom at mooring 636 ($\nu = 30$). The theoretical ratio (10) for linear, freely propagating internal waves, is also shown for the purpose of comparison. The 90% confidence intervals assume a Fisher's F distribution for the ratio of two independent chi-squared distributions (Jenkins and Watts, 1968, p.85), and so may not be appropriate.

$$\frac{P_{++}}{P_{--}} = 1. \tag{11}$$

Looking at Fig. 5, we find that at 10m and 50m above the bottom, P_{++}/P_{--} is closer to 1 than to the internal wave line for $\omega \leq 0.025$ cph. This result prompted Gilbert (1990) to try to verify whether the orientation ϕ of these motions changes with frequency according to

$$\phi = \cos^{-1}\left(\frac{\omega}{N \sin\alpha}\right), \tag{12}$$

as in Rhines' theory [4], but he found no significant changes of ellipse orientation within that range of frequencies. Thus whereas Eriksen's (1982) interpretation of the 3-4 day motions at mooring 636 as critically reflected internal waves fails to explain why the

[4]The notation used here differs from that of Rhines, ϕ is measured counterclockwise from the onslope direction to ensure consistency with the earlier definition of ϕ given in section 2.

current ellipses near ω_c are not perpendicular to the inferred North–South orientation of the isobaths, an interpretation of those motions in terms of Rhines' internal edge waves (Gilbert, 1990) fails to explain the lack of turning of the current ellipses with frequency.

4 DATA FROM THE CONTINENTAL RISE AND SLOPE OFF NOVA SCOTIA

The Bedford Institute of Oceanography deployed several moorings on the Scotian Rise and Slope during the 1970's and 1980's. Those moorings were deployed for the purpose of studying sub-inertial motions (e.g. Louis et al., 1982, Smith and Petrie, 1982), not for the purpose of studying internal wave reflection off sloping bottoms. Consequently, the vertical spacing between current meters is large at most moorings; there is usually only one current meter in the first 100m above the bottom, which does not allow us to study the vertical structure of some of the small-scale processes known to occur above sloping topography (e.g. Thorpe et al., 1990). Nevertheless, it should be possible to look for evidence of near-bottom energy enhancement at ω_c i. a manner analogous to Eriksen (1982).

Tables 1 and 2 provide a convenient summary of the current meter data used here, and Fig. 6 shows the location of the moorings. More detailed information about the data can be found in Lively (1979a, 1979b and 1984).

A list of criteria

Linear internal wave reflection theory predicts we should observe energy enhancement (5) at the critical frequency ω_c above sloping topography. For the data set under consideration here, we could try to verify this by looking for

1a) a spectral peak above the background energy level at ω_c

1b) near-boundary energy enhancement at ω_c for instruments at different heights above the bottom, but on the same mooring

1c) near-boundary energy enhancement at ω_c for instruments at the same depth, but on different moorings

1d) a spectral level at ω_c larger than that given by the GM79 model spectrum.

The other major prediction of linear internal wave reflection theory is that current ellipses should tend to orient themselves normal to the isobaths upon reflection (6). Hence we should also look for

2) a pronounced anisotropy at ω_c, the major axis of current ellipses being oriented in the cross-isobath direction.

Criterion 1a was used by Sandstrom (1966) and Eriksen (1982) at mooring sites where the critical frequency was well separated from the energetic tidal and inertial frequencies.

Such a separation is necessary if we wish to unambiguously interpret a spectral peak at ω_c as the signature of internal wave reflection. Unfortunately, ω_c is close to f or M_2 at nearly all the mooring sites considered here (see table 3), so that criterion 1a was only used at mooring S2.

Table 1: Mooring locations on the Scotian Rise and Slope. The letter S stands for the Shelf Break experiment, and R stands for the Risex experiment.

Mooring	Latitude	Longitude	Bottom depth	Current meter depths
S1	42°48.6′N	63°30.0′W	250m	20,50,100,150,230m
S2	42°46.8′N	64°00.0′W	250m	20,50,230m
S3	42°45.0′N	63°30.0′W	710m	230,500,690m
S4	42°40.4′N	63°30.0′W	1010m	50,100,150,500,690,990m
S5	42°30.5′N	63°30.0′W	1550m	50,150,1530m
S6	43°00.5′N	63°30.0′W	170m	20,50,100,150m
S7	42°41.7′N	64°00.0′W	710m	230,690m
S8	42°01.0′N	63°30.0′W	2550m	70,1500,2530m
R1	41°20.0′N	63°58.0′W	3600m	200,500,1000,2900,3500m
R2	41°27.0′N	63°30.0′W	3600m	3500m
R3	41°00.0′N	65°00.8′W	3600m	3500m
R4	41°38.6′N	64°17.7′W	3000m	200,1000,2900m
R5	41°53.2′N	64°31.7′W	2500m	200,500,1000,2400m

Table 2: Mooring deployment periods on the Scotian Rise and Slope. Periods A to G are from the Shelf Break Experiment, and periods H to J are from the Risex Experiment.

Mooring deployment period	Deployment date (dd/mm/yy)	Recovery date (dd/mm/yy)	Length of deployment (days)	Comments
A	13/12/75	06/04/76	115	Mooring S1 only.
B	06/04/76	05/07/76	90	S1 and S4 only.
C	05/07/76	17/10/76	104	
D	17/10/76	16/12/76	60	
E	16/12/76	02/04/77	107	
F	02/04/77	09/07/77	98	
G	09/07/77	04/01/78	179	Data from S1, S2,S4,S6 only.
H	04/11/80	03/03/81	119	Poor data return, except for R5.
I	03/03/81	06/05/81	64	Poor data return.
J	06/05/81	24/10/81	171	Good data return, except for R5.

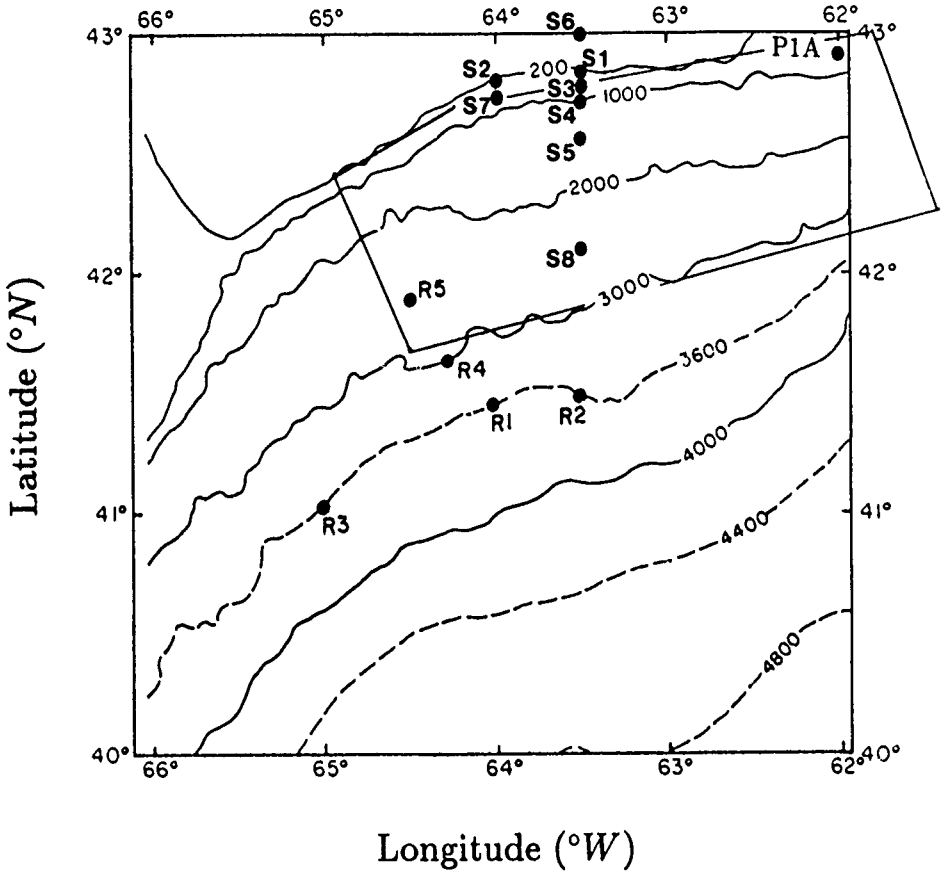


Figure 6: Map of the Scotian Rise and Slope showing the location of moorings used in this paper. The letter S refers to moorings from the 1975–78 Shelf Break experiment, R refers to moorings from the 1980–81 Risex Experiment, and mooring P1A is Petrie's (1975) mooring 1A. The hydrographic data used to compute the Brunt-Vaisala frequency $N(z)$ comes from subarea 33 of Drinkwater and Trites (1987), whose perimeter is shown here as a thin line.

Criterion 1b was also used by Eriksen (1982) at a few moorings. It can be used at most of the moorings considered here, as the majority of them have two or more current meters in the vertical. However, we should bear in mind that this criterion fails to take into account the slantwise propagation of internal wave energy, as pointed out by Gilbert and Garrett (1989), and so may be ambiguous. In all cases where we have used criterion 1b, the kinetic energy spectra from instruments at different depths were normalised with respect to a common value of N, taking into account WKB scaling (Briscoe, 1975, Gill 1982, p.300).

Criterion 1c was used by Thorpe (1987a, figure 11) at a site where the critical frequency was close to the inertial frequency ($\omega_c \approx f$), a situation where this criterion should be most useful due to nearly horizontal wave rays. It was used in a few cases for the present data set.

Criterion 1d does not on its own constitute evidence of critical internal wave reflection, but it could be used to support successful tests based on other criteria, especially when $\omega_c > M_2$, where there is some degree of universality in the spectral levels (Wunsch, 1976). Fu (1981) has shown that such universality does not exist near f in the deep ocean, so that criterion 1d was not used at moorings where $\omega_c \sim f$. It was not used at moorings where $\omega_c \sim M_2$ either, since internal tides were excluded from the internal wave model of Garrett and Munk (1972). Application of criterion 1d was thus limited to mooring S2 for the present data set.

Criterion 2 was successfully used by Eriksen (1982) for mooring sites on Muir Seamount and the New England continental slope, but failed at mooring 636 of the Western Boundary Sill Experiment. Application of criterion 2 was limited here to those moorings where ω_c was appreciably greater than f (near-inertial motions are quasi-circular and so the orientation of their 'major' axis tends to be less stable).

Hundreds of autospectra were computed from the data set described in tables 1 and 2, some of which are shown in Gilbert (1990). A summary of his results is given in table 3.

Table 3: Results of tests aimed at determining whether critical internal wave reflection was observed at moorings from the Shelf Break and Risex Experiments. In column 1 the critical frequency (in cph) at each of the mooring sites is written in parentheses. For the purpose of comparison, note that $f \sim 0.056$ cph, and $M_2 = 0.081$ cph. In columns 2 and 3, the capital letters refer to the mooring periods listed in table 2, and the numbers in parentheses refer to the criteria described in this section.

Mooring	Successful tests (95% significant)	Unsuccessful tests (not significant)	Comments
S1(.086)		D(2),E(2),G(2)	
S2(.132)	D(1d,2),E(1d,2), G(1d,2)	D(1a),E(1a), G(1a)	
S3(.076)	E(2),F(2)	D(1b,2),E(1b,1c), F(1b,1c)	
S4(.060)	C(1b),E(1b)	F(1b)	Large bump at $\approx 1000m$
S5(.058)	C(1b,1c),F(1b)	D(1b,1c),F(1c)	
S6(.059)		F(1c)	
S7(.076)	C(1b,2),E(2), F(1b,2)	D(1b,2),E(1b)	
S8(.057)		F(1b)	sig. at 80% level
R1(.055)		J(1b)	K.E.(3500m) > K.E.(2900m)
R2(.056)			K.E.($\omega = f$) $\sim$ R1 (3500m)
R3(.055)			K.E.($\omega = f$) $\sim$ R1 (3500m)
R4(.056)	J(1c)	J(1b)	
R5(.057)		H(1b)	

5 SUMMARY AND DISCUSSION

The WBSE data

In section 3 of this paper, we reexamined data from mooring 636 of the Western Boundary Sill Experiment to try to verify whether the energetic 3–4 day motions seen at that mooring are due to critical internal wave reflection, as postulated by Eriksen (1982). We pointed out, as Eriksen did, that those motions are parallel to the inferred orientation of the isobaths, not normal to them, thus violating one of the predictions of specular internal wave reflection theory.

We also pointed out that the 3–4 day motions are coherent over a horizontal scale of at least 40 km (Fig. 4). Such a large coherence scale appears to be inconsistent with the transfer of energy to smaller scales that should occur due to internal wave reflection. Gilbert (1990) also showed that the 3–4 day motions at mooring 636 are coherent over a vertical scale of at least 200m, and the measured phase leads or lags are generally not significantly different from zero over that distance, implying large vertical scales of motion.

Another puzzling feature of the 3–4 day motions at mooring 636 is that they do not satisfy the consistency relation (10) for freely-propagating internal waves (Fig. 5). At 10m and 50m above the bottom, those motions are essentially rectilinear, apparently more consistent with the prediction of Rhines (1970) for internal edge waves (11). Agreement with Rhines' theory is quite limited however, as the orientation of those quasi-rectilinear motions does not change with frequency according to (12) (Gilbert, 1990).

Perhaps this illustrates the need for a careful investigation of the manner in which the consistency relations for internal waves should change as we approach a sloping bottom. Let us consider here the inviscid problem of a wave which reflects off a vertical wall with $\phi_i \neq 0$. The onslope velocity component of the incident and reflected waves must cancel each other at the wall in order to satisfy the boundary condition of no normal flow, but the alongslope velocity component of the incident and reflected waves are phase-locked at the wall and add up constructively. Consequently, in the immediate vicinity of a vertical wall, motions should be rectilinear, and we should have $P_{++}/P_{--} = 1$ instead of $P_{++}/P_{--} = [(\omega - f)/(\omega + f)]^2$ (Barry Ruddick, personal communication). We need to verify whether this result would still hold above sloping topography, as it may help explain the variation of P_{++}/P_{--} with height above the bottom [5] observed at mooring 636 for example (Fig. 5). Preliminary results suggest that a slope 2 to 3 times steeper than 0.015 might be consistent with Fig. 5, but that would leave the energy peak at $\omega \approx 0.012$ cph shown on Fig. 2 unexplained.

Our lack of success in trying to interpret the 3–4 day motions at mooring 636 is largely due to our poor knowledge of the topography. Figure 1 is only based on an East–West

[5] Mooers (1973) suggests that departure from the ratio (10) may be a good indicator of the validity of the low spatial coherence hypothesis of Fofonoff (1969)

transect, and so cannot tell us what is the orientation and magnitude of ∇h. The most detailed bathymetric map available for the area, produced by Moody et al. (1979) and shown on figure 2 of WW82, does not help us solve that problem either due to its lack of resolution.

Proximity to the equator ($f \sim 10^{-5} s^{-1}$ at 4°N) and the possibility of equatorially trapped waves (Moore and Philander, 1977, Eriksen, 1980) may also complicate the analysis, a fact that was overlooked by Gilbert (1990) and probably deserves closer attention. Among other things, the traditional approximation of neglecting the horizontal component of the Earth's rotation vector Ω in the equations of motion may have to be reexamined near the equator.

The Scotian Slope data

In section 4, historical current meter data from the Scotian Rise and Slope were used in order to look for evidence of critical internal wave reflection. To that end, a set of criteria was proposed and discussed, and then applied to the data. Energy enhancement and/or cross-isobath alignment of motions near ω_c was found to be significant at the 95% level in some cases (see table 3), but was generally not very pronounced. This could be due to the overall concavity of the Scotian Rise and Slope (Gilbert and Garrett, 1989), but other factors may also affect the likelihood of observing energy enhancement at the critical frequency.

Gilbert (1990) suggests that when $N \sin\alpha \ll f$ at a given mooring site, so that $\omega_c \sim f$ (3), the orientation of the isobaths could be one such factor. The linear reflection laws of Eriksen (1982), valid on an f-plane, predict that maximum wavenumber amplification upon reflection should occur for onslope incident energy propagation (i.e. $\phi_i = 0$). However on a β-plane, due to the turning latitude effect, near-inertial motions are more likely to have $\phi_i \approx \pm\pi/2$ when the onslope direction is poleward (e.g. Kroll, 1975). This is roughly the case on the Scotian Slope, possibly explaining the overall absence of pronounced energy enhancement at $\omega_c \sim f$ [6] (note that $\omega_c \approx f$ at 9 of the 13 moorings in table 3). A more quantitative investigation of this phenomenon could probably be carried out using the wave functions of Munk and Phillips (1968).

Finally we would like to draw the attention of the reader to the non-linear, specular reflection theory of Thorpe (1987b), which raises the possibility of singularities for wave ray slopes different from the bottom slope. When a train of finite amplitude internal waves travelling in a vertical plane normal to the slope gets reflected, resonance between the incident and reflected waves is possible at second order for a bottom slope less than 0.15 and a wave ray slope less than 0.58. This condition is met almost everywhere on the Scotian Rise and Slope for example. Gilbert (1990) offered an explanation for the unusual

[6] Fu (1981) found that the largest energy peaks at $\omega = f$ in the abyssal ocean occured in the vicinity of the Mid-Atlantic Ridge, which runs roughly North-South, so that the onslope direction is zonal and near-inertial waves are more likely to have $\phi_i \approx 0$.

size of the M_4 peak seen at Petrie's (1975) mooring 1A (see Fig. 6) in terms of that theory, and also explained some aspects of the observations at moorings S3 and S7 using Thorpe's theory. Unlike the first order resonance, higher order resonances should be unaffected by the boundary concavity criterion of Gilbert and Garrett (1989).

Acknowledgments

The useful advice and patience of Chris Garrett during the few years spent at Dalhousie University are gratefully acknowledged. Financial support for this work was provided by the U.S. Office of Naval Research.

REFERENCES

Briscoe, M.G. 1975. Preliminary results from the trimoored Internal Wave Experiment (IWEX). *J.Geophys.Res.*, **80**: 3872–3884.

Drinkwater, K.F. and R.W. Trites, 1987. Monthly means of temperature and salinity in the Scotian Shelf region., *Canadian technical report of fisheries and aquatic sciences No.1539*, iv + 101p.

Eriksen, C.C. 1980. Evidence for a continuous spectrum of equatorial waves in the Indian Ocean. *J. Geophys. Res.*, **85**: 3285–3303.

Eriksen, C.C. 1982. Observations of internal wave reflection off sloping bottoms. *J. Geophys. Res.*, **87**: 525–538.

Eriksen, C.C. 1985. Implications of ocean bottom reflection for internal wave spectra and mixing. *J. Phys. Oceanogr.*, **15**: 1145–1156.

Fofonoff, N.P. 1969. Spectral characteristics of internal waves in the ocean. *Deep-Sea Res.*, **16**: 58–71.

Fu, L.L. 1981. Observations and models of inertial waves in the deep ocean. *Rev. Geophys. Space Phys.*, **19**: 141–170.

Garrett, C.J.R. and W.H. Munk 1972. Space-time scales of internal waves. *Geophys. Fluid Dyn.*, **2**: 225–264.

Garrett, C.J.R. and D. Gilbert, 1988. Estimates of vertical mixing by internal waves reflected off a sloping bottom. In *Small-scale turbulence and mixing in the ocean.*, J.C.J. Nihoul and B.M. Jamart (editors), Elsevier, Amsterdam, pp. 405–423.

Gilbert, D. 1990. Theory and observations of internal wave reflection off sloping topography. *Ph.D. Thesis*, Dalhousie University, xix + 183 pp.

Gilbert, D. and C.J.R. Garrett 1989. Implications for ocean mixing of internal wave scattering off irregular topography. *J. Phys. Oceanogr.*, **19**: 1716–1729.

Gill, A.E., 1982. *Atmosphere-Ocean Dynamics.* Academic Press, New York, 662 pp.

Haurwitz, B., H. Stommel and W.H. Munk, 1959. *On the thermal unrest in the ocean. In* The atmosphere and sea in motion, Rossby Memorial Volume, B. Bolin (editor), Rockefeller Univ. Press, pp. 74–94.

Huthnance, J.M. 1978. On coastal trapped waves: analysis and numerical calculation by inverse iteration. *J.Phys.Oceanogr.*, **8**: 74–92.

Huthnance, J.M. 1989. Internal tides and waves near the continental shelf edge. *Geophys. Astrophys. Fluid Dynamics*, **48**: 81–106.

Huthnance, J.M. and P.G. Baines 1982. Tidal currents in the Northwest African upwelling region. *Deep-Sea Res.*, **29**: 285–306.

Jenkins, G.M. and D.G. Watts, 1968. *Spectral analysis and its applications.*, Holden-Day, Oakland, California, 525 pp.

Kroll, J. 1975. The propagation of wind-generated inertial oscillations from the surface into the deep ocean. *J.Mar.Res.*, **33**: 15–51.

Lively, R.R. 1979a. Current meter and meteorological observations on the Scotian Shelf: December 1975 to January 1978. Volume 1: December 1975 to December 1976. Bedford Institute of Oceanography data series report BI–D–79–1, 280 pp.

Lively, R.R. 1979b. Current meter and meteorological observations on the Scotian Shelf: December 1975 to January 1978. Volume 2: December 1976 to January 1978. Bedford Institute of Oceanography data series report BI–D–79–1, 368 pp.

Lively, R.R. 1984. Current meter observations on the Scotian Rise for November 1980 to October 1981. *Canadian technical report of hydrography and ocean sciences no. 50,* 156 pp.

Louis, J.P., B.D. Petrie and P.C. Smith 1982. Observations of topographic Rossby waves on the continental margin off Nova Scotia. *J.Phys.Oceanogr.*, **12**: 47–55.

Moody, R., D.E. Hayes and S. Connary 1979. Bathymetry of the continental margin of Brazil, *Map number 832,* Am. Assoc. Petrol. Geol., Tulsa, Oklahoma.

Mooers, C.N.K. 1973. A technique for the cross spectrum analysis of pairs of complex-valued time series, with emphasis on properties of polarized components and rotational invariants. *Deep-Sea Res.*, **20**: 1129–1141.

Moore, D.W. and S.G.H. Philander 1977. Modeling of the tropical oceanic circulation, in *The Sea,* vol. 6, pp. 319–361, Interscience, New York.

Müller, P., D.J. Olbers and J. Willebrand 1978. The Iwex spectrum *J.Geophys.Res.*, **83**: 479–500.

Munk, W.H. 1966. Abyssal recipes. *Deep-Sea Res.*, **13**: 207–230.

Munk, W.H. and N. Phillips 1968. Coherence and band structure of inertial motion in the sea. *Rev.Geophys.Space Phys.*, 6: 447–472.

Petrie, B. 1975. M2 surface and internal tides on the Scotian shelf and slope. *J. Mar. Res.*, **33**: 303–323.

Phillips, O.M. 1963. Energy transfer in rotating fluids by reflection of inertial waves. *Phys. Fluids*, **6**: 513–520.

Phillips, O.M., 1966. *The Dynamics of the Upper Ocean.* Cambridge University Press, New York, 261pp.

Phillips, O.M., 1977. *The Dynamics of the Upper Ocean. 2nd ed.,* Cambridge University Press, New York, 336 pp.

Rhines, P.B. 1970. Edge-, Bottom-, and Rossby waves in a rotating stratified fluid. *Geophys. Fluid Dyn.*, 1: 273–302.

Sandstrom, H., 1966. The importance of topography in generation and propagation of internal waves. *Ph.D. thesis,* University of California at San Diego, xi + 105 pp.

Sandstrom, H. 1969. Effect of topography on propagation of waves in stratified fluids. *Deep-Sea Res.*, **16**: 405–410.

Smith, P.C. and B.D. Petrie 1982. Low-frequency circulation at the edge of the Scotian Shelf. *J. Phys. Oceanogr.*, **12**: 28–46.

Thompson, R.O.R.Y. and J.R. Luyten 1976. Evidence for bottom-trapped topographic Rossby waves from single moorings. *Deep-Sea Res.*, **23**: 629–635.

Thorpe, S.A. 1987a. Current and temperature variability on the continental slope. *Phil. Trans. R. Soc. London, Series A*, **323**: 471–517.

Thorpe, S.A. 1987b. On the reflection of a train of finite amplitude internal gravity waves from a uniform slope. *J. Fluid Mech.*, **178**: 279–302.

Thorpe, S.A., P. Hall and M. White 1990. The variability of mixing at the continental slope. *Phil. Trans. R. Soc. Lond. A*, **331**: 183–194.

Whitehead, J.A. and L.V. Worthington 1982. The flux and mixing rates of Antarctic Bottom Water within the North Atlantic. *J. Geophys. Res.*, **87**: 7903–7924.

Wunsch. C. 1972. The spectrum from two years to two minutes of temperature fluctuations in the main thermocline at Bermuda. *Deep-Sea Res.*, **19**: 577–593..

Wunsch, C. 1976. Geographical variability of the internal wavefield: A search for sources and sinks. *J. Phys. Oceanogr.*, **6**: 471–485.

Wunsch, C. and S. Webb 1979. The climatology of deep ocean internal waves. *J. Phys. Oceanogr.*, **9**: 235–243.

OBSERVATIONS OF NEAR-INERTIAL INTERNAL WAVES AND MIXING IN THE SEASONAL THERMOCLINE

Charles C. Eriksen
School of Oceanography WB-10
University of Washington
Seattle, WA 98195

ABSTRACT

Wind forced near-inertial internal waves reduce gradient Richardson number sufficiently to induce microscale mixing within the seasonal pycnocline in eastern North Pacific observations. The net buoyancy change inferred from a moored time series of upper ocean observations implies turbulent dissipation rates sufficient to damp near-inertial kinetic energy in the mixed layer in several days.

INTRODUCTION

Wind stress is well recognized as a source of both internal gravity waves and mixing in the upper ocean. Swiftly moving wind systems generate internal waves whose phase speeds match their translation speeds. Since most storms translate at speeds at the high end of the range of possible internal wave phase speeds, storm-generated internal waves tend to be only slightly superinertial in frequency (Kundu and Thomson, 1985; D'Asaro, 1989). While the details of how stress applied by the wind is transmitted through the upper ocean are far from clear, both observations and theoretical considerations point to nearly linear variation of stress through a surface mixed layer such that stress effectively vanishes only slightly deeper than this layer. Because there is a relatively sharp transition between acceleration in the mixed layer and the stratified region beneath it, strong shears are found at the base of the mixed layer. This shear, in turn, is responsible for continued deepening of a mixed layer, as recognized in model parameterizations (Pollard et al., 1973; Price et al., 1986). As this shear is dominated by variance at slightly superinertial frequencies, motions in the seasonal pycnocline take the form of propagating near-inertial internal gravity waves. Observations in the OCEAN STORMS program discussed here indicate that these wind-forced internal waves are sufficiently energetic to account for mixing implied by temporal evolution of upper ocean density structure.

UPPER OCEAN CURRENT AND DENSITY STRUCTURE

The observations discussed here come from a single 10-month Profiling Current Meter (PCM) record collected at 47° 35′ N, 139° 23′ W in the eastern North Pacific. The record begins August 22, 1987 and ends June 12, 1988, spanning the complete fall and winter cooling periods and the beginning of spring warming. The PCM profiled every 4 hours, averaging currents, temperature, and electrical conductivity into 5 m thick bins from 195 to 35 m depth. The subsurface PCM mooring was set 15 m deeper than in other deployments in anticipation of the severe sea state expected (and encountered) in the OCEAN STORMS site. The upward profile transit routinely took 20 to 25 minutes to complete, as the instrument used an electric pump system to control its buoyancy, hence its movement along the upper section of the mooring. The unique feature of PCMs is that they collect simultaneous current and density profiles at fixed locations in the upper ocean over extended periods. The instruments are capable of collecting roughly 2000 profiles, depending on the depth to which profiles are made, the strength of ambient currents and water temperature. Details of the design and operation of the engineering prototype PCM are given in Eriksen et al. (1982).

Upper ocean density structure at the OCEAN STORMS site is dominated by two pycnoclines; the upper seasonal one is dominated by temperature stratification while the deeper permanent one is dominated by salinity stratification. During the course of the winter, the shallow seasonal pycnocline is not merely eroded by storm-induced mixing, it itself is deepened as the mixed layer deepens, as indicated in the density record contoured in Figure 1. From August through October, the PCM does not profile high enough in the water column to encounter the mixed layer. Over the course of November and December, the mixed layer both deepens and becomes more dense. The seasonal pycnocline maintains a thickness of about 20 m during this period whereafter it merges with the permanent pycnocline centered at roughly 115 m depth. Although the stratification remains weak at depths shallower than about 90 m from January through March, this part of the water column is not strictly uniformly well mixed. Restratification begins in April and by late May a new seasonal pycnocline is evident.

The current structure which accompanies these changes in stratification is dominated by near-inertial motions (inertial period $2\pi/f$=16.254 hr) concentrated in the surface mixed layer but occasionally penetrating both the seasonal and permanent pycnoclines. The stick vector plot of currents in Figure 2 reveals an enormous burst of kinetic energy starting in early October at the shallowest depth profiled and appearing later deeper, suggestive of a wave packet propagating. High current largely is confined to the mixed layer as it deepens in November and December, but more bursts of high current in the pycnocline occur in January and less intensely through April and even May.

Near-inertial frequency motions dominate the spectrum of current for all depths and are particularly evident in the October storm event. An expanded view of density and current

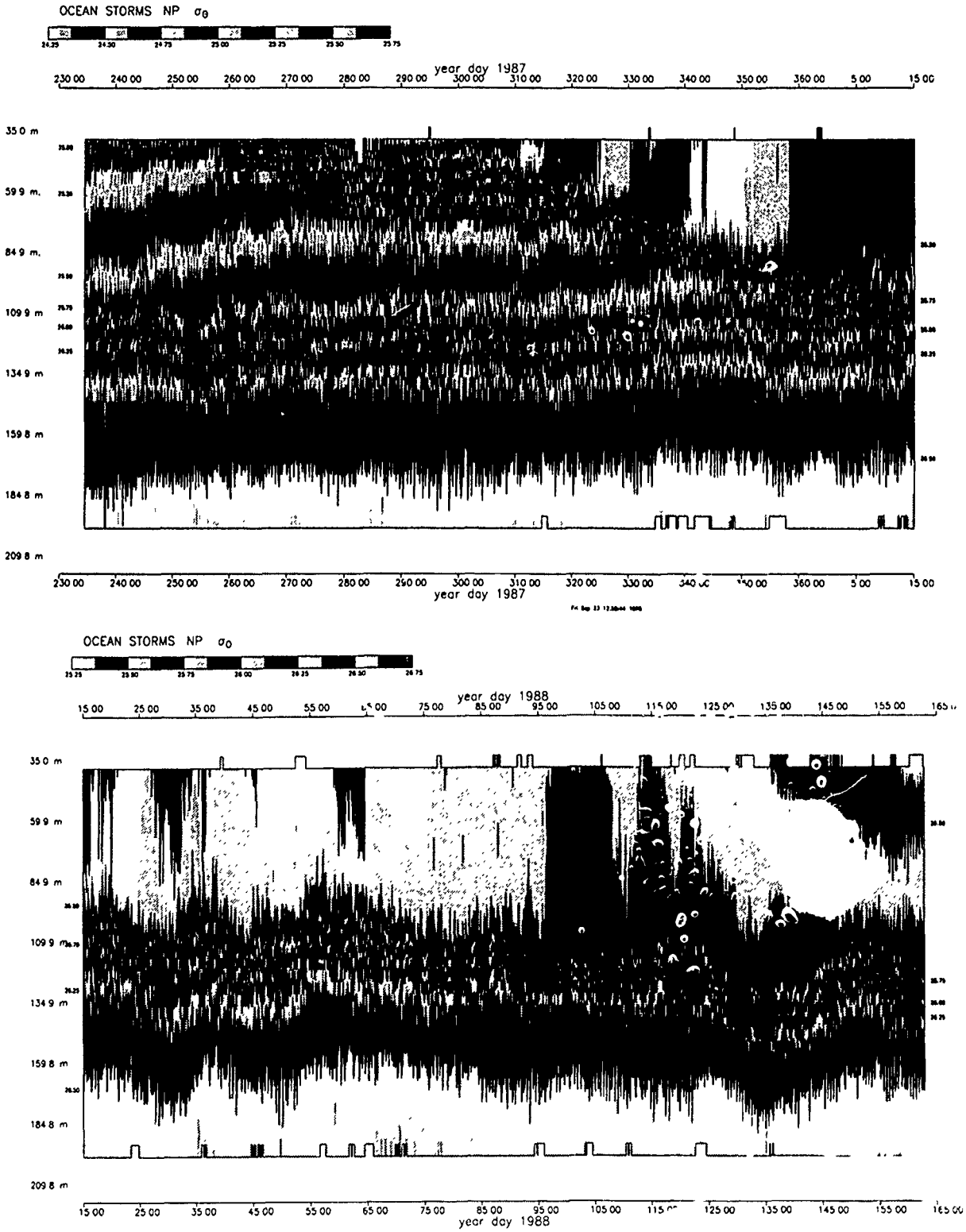


Figure 1. Potential density $\sigma_{\theta 3}$ in 5-m depth bins every 4 hr contoured with a gray scale which repeats every 0.25 kg/m. The two frames start at 18 August 1987 (year day 230) and 15 January 1988 (year day 15), respectively, and depths are given at the left of each.

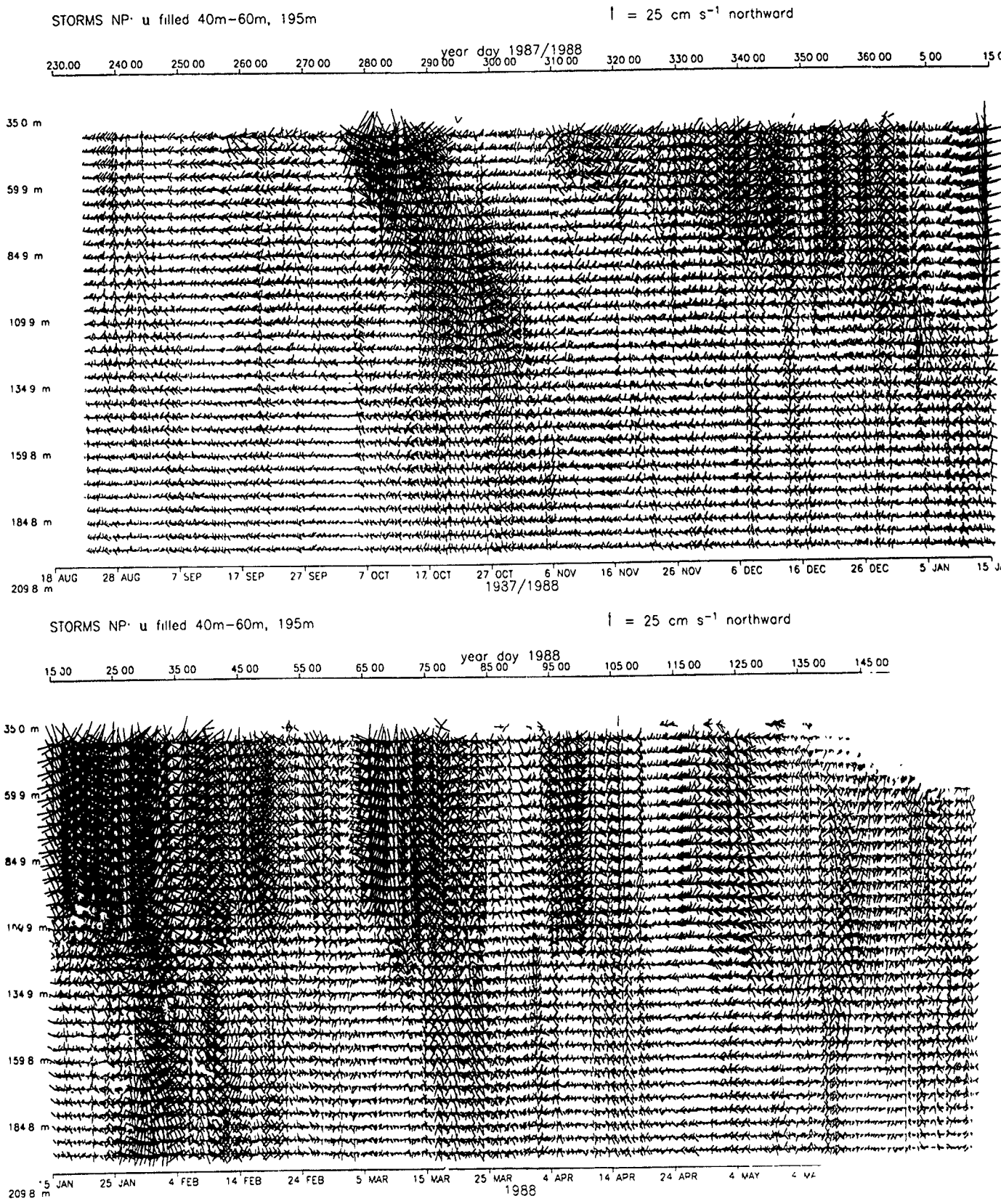


Figure 2. Current time series for each depth bin plotted as stick vectors where the scale is given at the upper right of each frame. The frames are the same as in Fig. 1.

is given in Figures 3 and 4 covering the period September 12 to November 1, 1987. Of particular interest is the jump in stratification structure which takes place on October 4 just beneath the seasonal pycnocline at depths from 55 to 75 m (Figure 3, left panel). This jump from weak to stronger stratification coincides with the appearance of strong near-inertial rotary currents (Figure 4, left panel). Density surfaces oscillate vertically with amplitudes as much as 10 m peak to peak with a near-inertial period in the week following the onset of the current burst. Interpreting these as due to vertical motion is consistent with regarding the horizontal currents as due to near-inertial internal waves with horizontal wavelengths of a few hundred km. A less prominent transition in density structure and near-inertial current activity can be found on September 14 at 40 to 50 m depth. Horizontal advection is inadequate to explain these density changes (as will be discussed below). The implication is that water within the pycnocline undergoes incomplete mixing in response to internal wave activity, that is, mixing which is sufficient to reduce but not eliminate stratification locally.

This mixing is accomplished through elevation of shear and accompanying reduction of Richardson number. Reduction of Richardson number is associated with shear instability, an idea confirmed by microstructure observations of turbulent mixing (Peters et al.,1988). Vertical shear increases to amplitudes as high as 0.02 s^{-1} in near-inertial waves in the October 4 storm (Figure 5). Elevated vertical shear is also apparent in the top three depth bins for which a 10 m centered first difference can be evaluated (45, 50, and 55 m) in the September 14 storm. Shear also tends to be somewhat higher within the permanent pycnocline, a feature presumed due to amplification of horizontal currents as internal waves propagate through a region of stronger stratification (e.g., WKB scaling). Gradient Richardson number Ri is the square of the ratio of buoyancy frequency N to vertical shear $\mathbf{u}_z$, hence a plot of a vector whose components are N^2 and $\mathbf{u}_z^2$ can reveal variations in both the stability of flow and the potential and kinetic energy associated with stratification and shear. Such a plot is given in Figure 6, where the slope of each stick vector relative to the time axis is the instantaneous inverse Richardson number Ri^{-1}. During both the October 4 and September 14 storms Ri reduces to unity or less.

ESTIMATES OF MIXING RATES

The substantial changes in upper ocean stratification evident in this OCEAN STORMS PCM record strongly suggest that mixing is important in effecting these changes. Compared to other regions of the ocean, mesoscale eddy activity in the eastern North Pacific is relatively weak. Nevertheless, advection still makes significant contributions to density changes observed at a fixed location. In order to infer quantitative estimates of mixing activity from a sequence of profiles at a fixed location, estimates of horizontal and vertical advection of density must be added to the observed rate of density change so that their sum can be considered the convergence of microscale turbulent density flux. Consider the equations for density and horizontal momentum conservation:

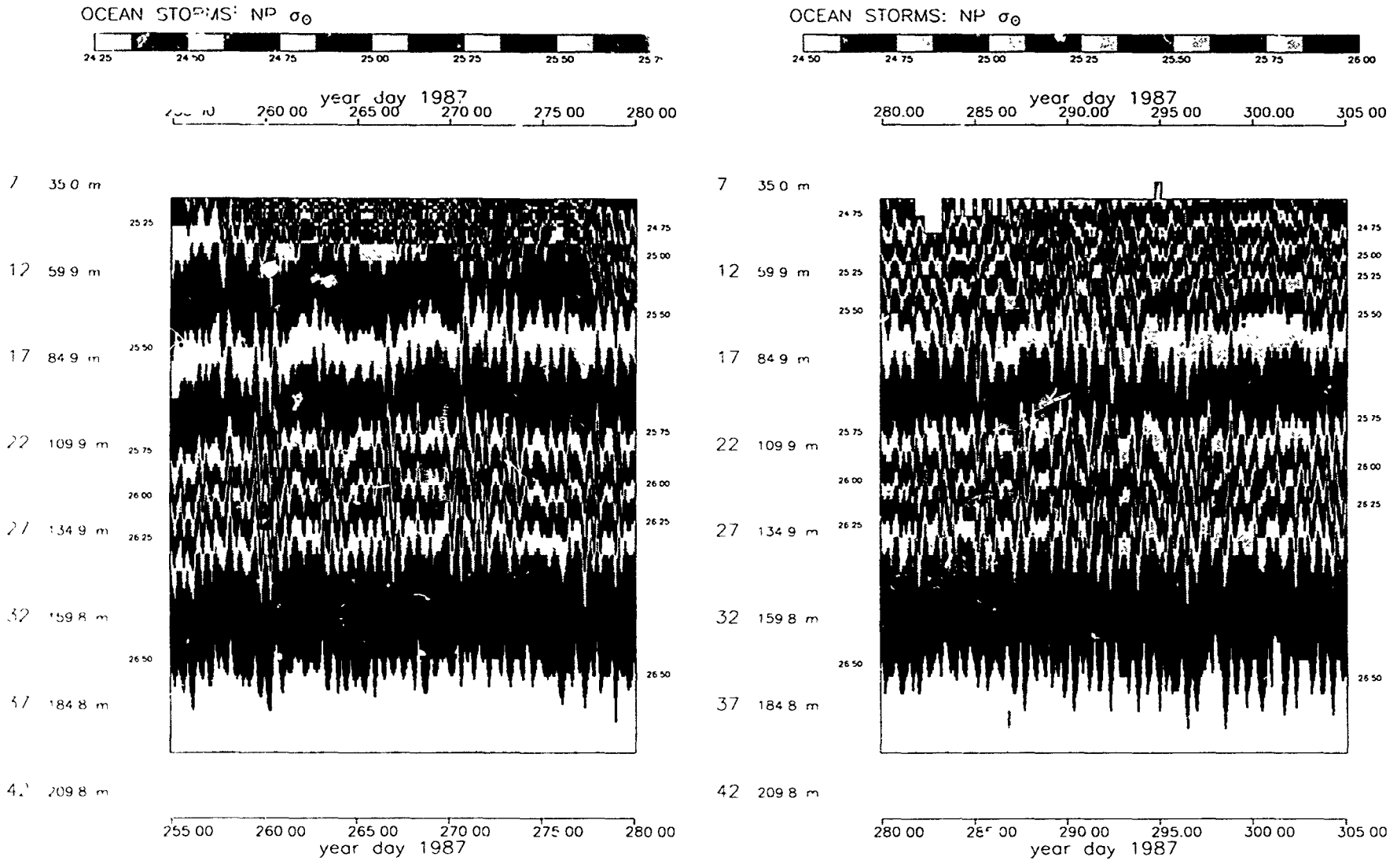


Figure 3. Expanded time scale contour plot of potential density σ_θ for the period 12 September to 1 November 1987, otherwise as in Fig. 1. September 14 and October 4 are year days 257 and 277, respectively.

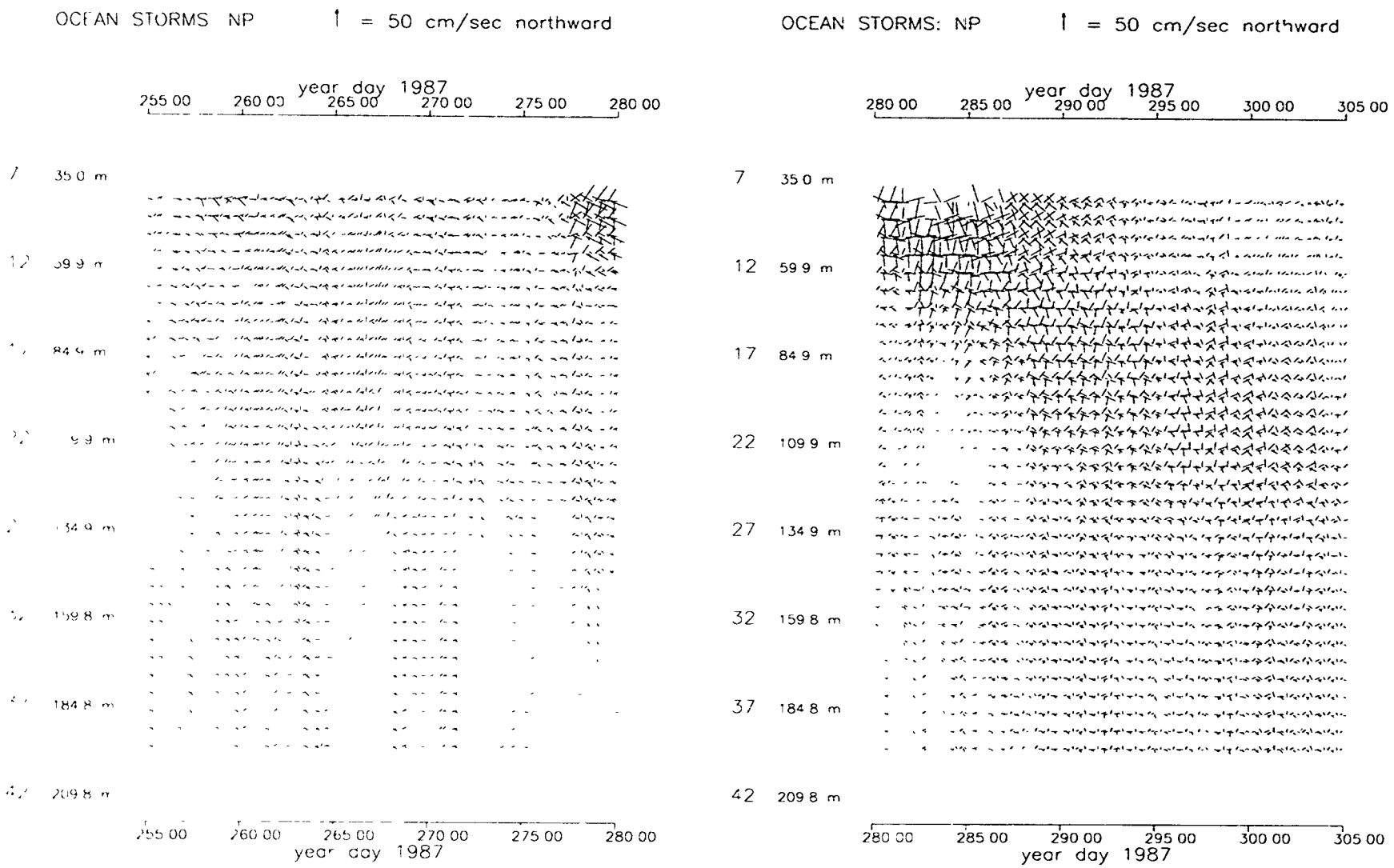


Figure 4. Expanded time scale current time series for the period shown in Fig. 3, otherwise as in Fig. 2.

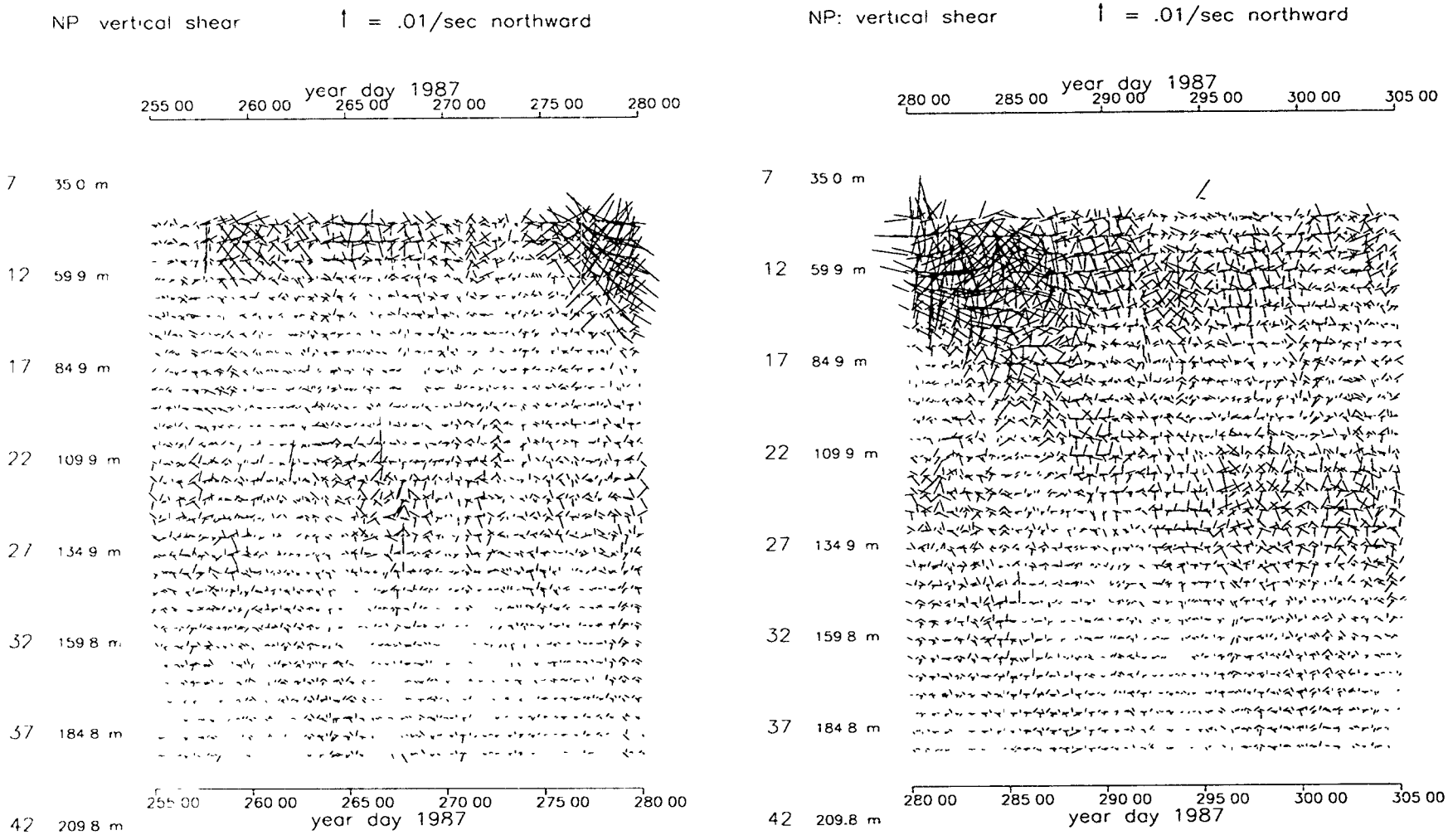


Figure 5. Vertical shear time series calculated as a center first difference over 10 m and plotted as stick vectors for the period shown in Fig. 3.

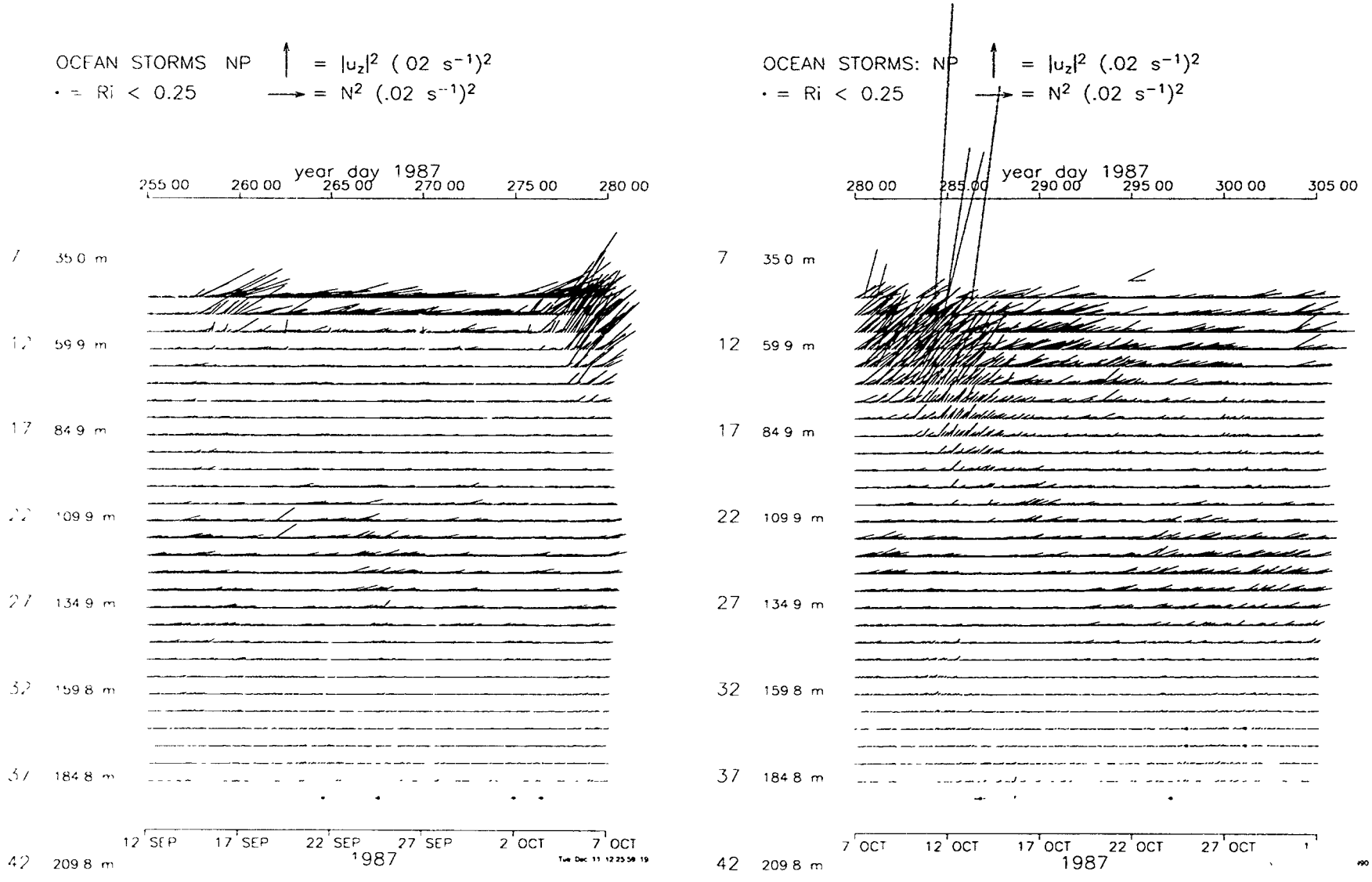


Figure 6. Vector time series of a quantity whose components are shear squared and buoyancy frequency squared plotted as in Figs. 4 and 5. The slope of each vector above the time axis is inverse Richardson number.

$$\rho_t + \mathbf{u}\cdot\nabla\rho + w\rho_z = -\langle w'\rho'\rangle_z \qquad (1)$$

$$\mathbf{u}_t + \mathbf{u}\cdot\nabla\mathbf{u} + w\mathbf{u}_z + 2\Omega\times\mathbf{u} + \rho_0^{-1}\nabla p = -\langle \mathbf{u}'w'\rangle_z \qquad (2)$$

where the primed quantities represent turbulent microscale variations and brackets the averages of their products. The first equation states that density changes following a particle are due solely to the vertical convergence of vertical turbulent transport of density. The second states that horizontal acceleration of a fluid parcel in a rotating frame of reference is balanced by a pressure gradient force and vertical convergence of vertical turbulent transport of horizontal momentum. Turbulence is considered as any process producing non-vanishing correlation between flow components at time and space scales too small to be resolved by the observations. It is also assumed that only vertical transports of density and momentum by turbulence are important. At low frequencies the momentum balance is geostrophic provided turbulent stresses $\tau=-\rho_0\langle \mathbf{u}'w'\rangle$ can be ignored. This simplification is poorest within the mixed layer and is of questionable validity within the seasonal pycnocline. Nevertheless, we make it here in order to obtain estimates of horizontal density gradients through the thermal wind equations. Then (1) may be written as

$$\rho_t - f\rho_0 g^{-1}(uv_z - vu_z) + w\rho_z = -\langle w'\rho'\rangle_z \qquad (3)$$

where $f = 2\Omega\cdot\sin(\text{latitude})$ is the inertial frequency. Since we are without direct measurements of vertical velocity, we must assume that turbulent mass fluxes $\langle w'\rho'\rangle$ effectively vanish at some depth beneath the mixed layer so that w can be estimated from (3) for the restricted depth range for which density is assumed to change purely advectively. Since w must also effectively vanish at the sea surface because changes in sea level are very slight, w can be interpolated for the mixed layer and the depth range within the pycnocline for which mixing is presumed of potential importance. The linear inviscid modes of oscillation of a rotating stratified fluid over a flat bottom are separable into vertical and horizontal structures. The equation governing vertical structure indicates that w is linear within a mixed layer, giving another constraint on the shape of w(z). Matching deep w behavior inferred from the non-mixing version of (3) with a linear portion in the mixed layer which vanishes at the sea surface makes possible an estimate of vertical density advection throughout the region where turbulent transport of density is important. Hence, through the assumptions stated, the PCM record of current and density profiles is sufficient to estimate the substantial derivative of density (all three terms on the left side of (3)). Vertical integration of density change following a fluid parcel then gives an estimate of turbulent mass flux $\langle w'\rho'\rangle$, that is, the quantity whose convergence represents mixing.

Although mixing events in the seasonal pycnocline take place on time scales identical to those of storms, even the relatively weak mesoscale eddy environment of the OCEAN STORMS site requires quantities to be smoothed temporally to yield stable estimates of

advective contributions. A filtered density record is given in Figure 7 where fluctuations with periods shorter than 20 days have been suppressed. Unfortunately, the number of weights necessary for the filter shortens the record at each end, so much so that the low-passed record begins at the onset of the October 4 storm. Contours initially within the seasonal pycnocline descend over time before eventually outcropping, indicative of mixing within the pycnocline. The second half of the record (lower frame, Figure 7) reveals little mixing activity and a tendency for the pycnocline to rise and fall due to mesoscale fluctuations. The time derivative of low-pass filtered density shows both the late fall season increase in density within the mixed layer and accompanying decrease beneath it (Figure 8). The horizontal advective contribution to density, assuming vertical shear is purely geostrophic, is comparable in size to local temporal change (Figure 9). This term (the second in equation 3) is positive within the mixed layer from late October through mid-January, indicating that lighter water is being imported laterally into the region. This tendency is nearly cancelled within the pycnocline by upwelling during this period, as evidenced by negative values of the vertical advection term $w\rho_z$ (Figure 10). It may be noted that any error in estimating horizontal advection will necessarily tend to be cancelled by an error in the opposing sense in implied vertical advection as (3) suggests.

The vertical buoyancy flux $-g\rho_0^{-1}\langle w'\rho'\rangle$ implied by the sum of the terms on the left side of equation (3) is given in Figure 11. While uncertainties in estimating advection make the second half of the record somewhat suspect, there is clearly a large upward buoyancy flux within the mixed layer during the fall cooling period. This flux averages roughly 10^{-7} W/kg at the surface. This buoyancy flux is presumed carried by convective activity within the mixed layer. Buoyancy flux changes sign at the base of the mixed layer so that buoyancy is transported downward within the seasonal pycnocline and deposited there. This downward flux reaches values as high as 10^{-8} W/kg. If turbulent mass flux is expressed as being diffusive, following the definition $\langle w'\rho'\rangle = -K\rho_z$, then the implied diffusivity K within the seasonal pycnocline falls in the range $1\text{-}5\times10^{-4}$ m^2/s. These are reasonable values for what is expected in a highly smoothed description of mixing.

DISCUSSION AND CONCLUSIONS

The implication of the downward buoyancy flux within the seasonal pycnocline is that turbulent dissipation must be high enough to decay the observed near-inertial internal wave activity substantially. The "dissipation method" (see Gregg (1987), Osborne (1980)) declares that flux Richardson numbers in stratified turbulence are limited to values of about 0.2. Measurements in strongly sheared equatorial currents indicate the ratio of buoyancy flux to turbulent kinetic energy dissipation to be about 0.1 (Peters et al., 1988). Hence the dissipation associated with turbulence is 5 to 10 times the buoyancy flux being effected. If the buoyancy flux of turbulence in the seasonal pycnocline is $O(10^{-8}$ W/kg), then the dissipation rate of kinetic energy must be $O(10^{-7}$ W/kg). This is the rate at which energy is presumed removed from near-inertial motions in the OCEAN STORMS observations. A typical kinetic energy content of mixed layer near-inertial

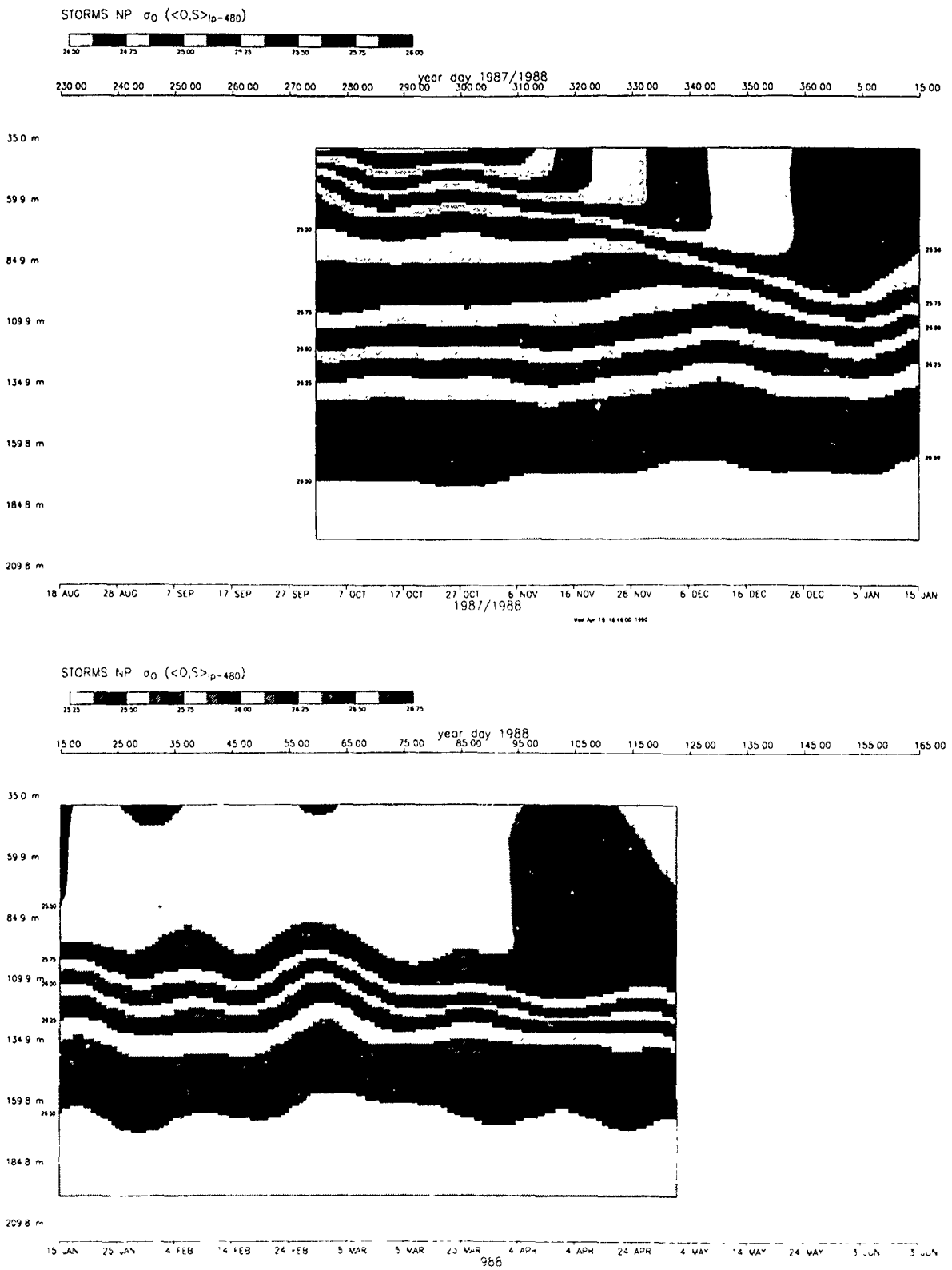


Figure 7. Low-pass filtered potential density σ_θ contoured as in Fig. 1. The start and end of the record has been deleted to accommodate the filter.

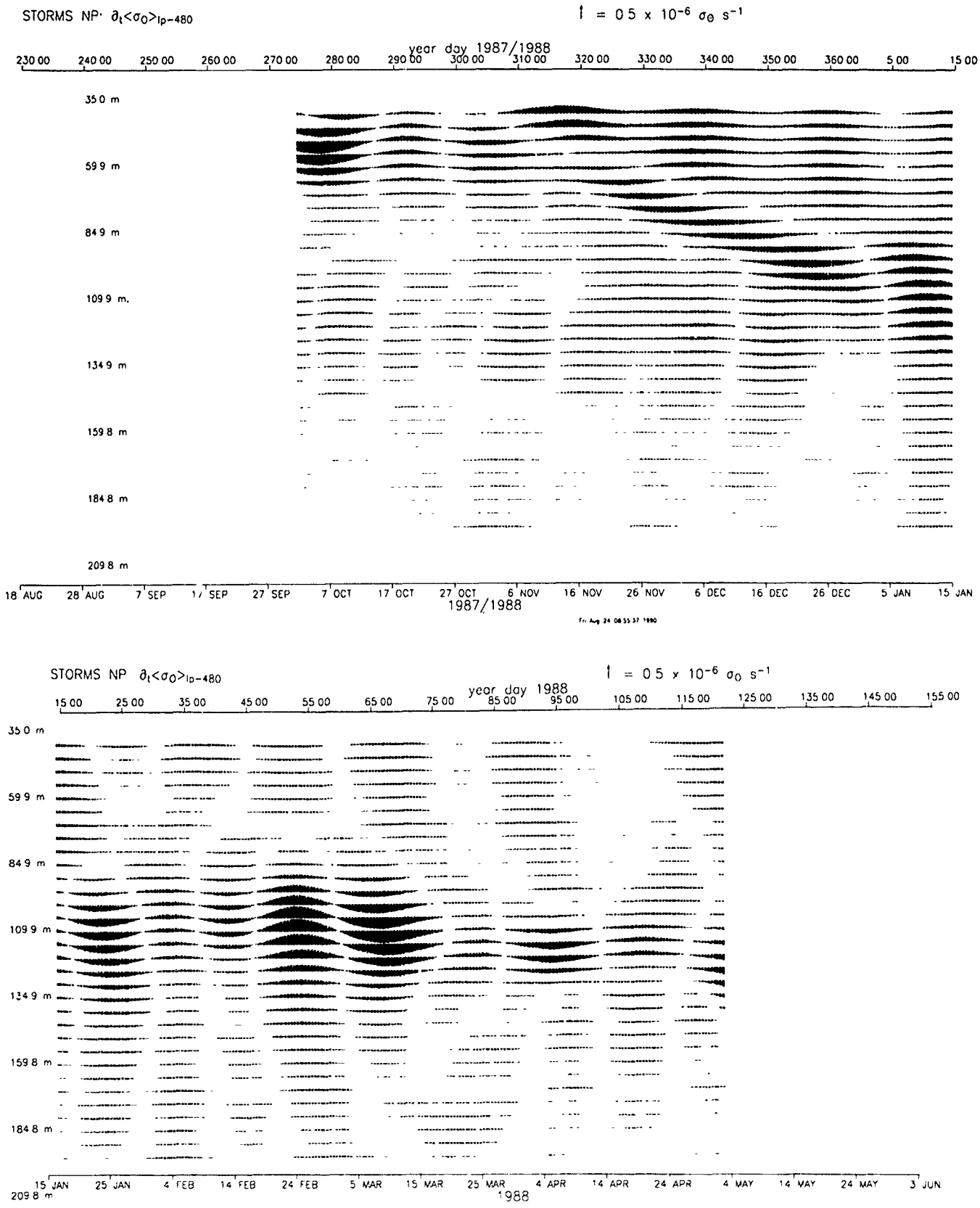


Figure 8. Time rate of change of low-pass filtered potential density where the scale in the upper right of each frame is 0.5 x 10^{-9} kg/m^3/s.

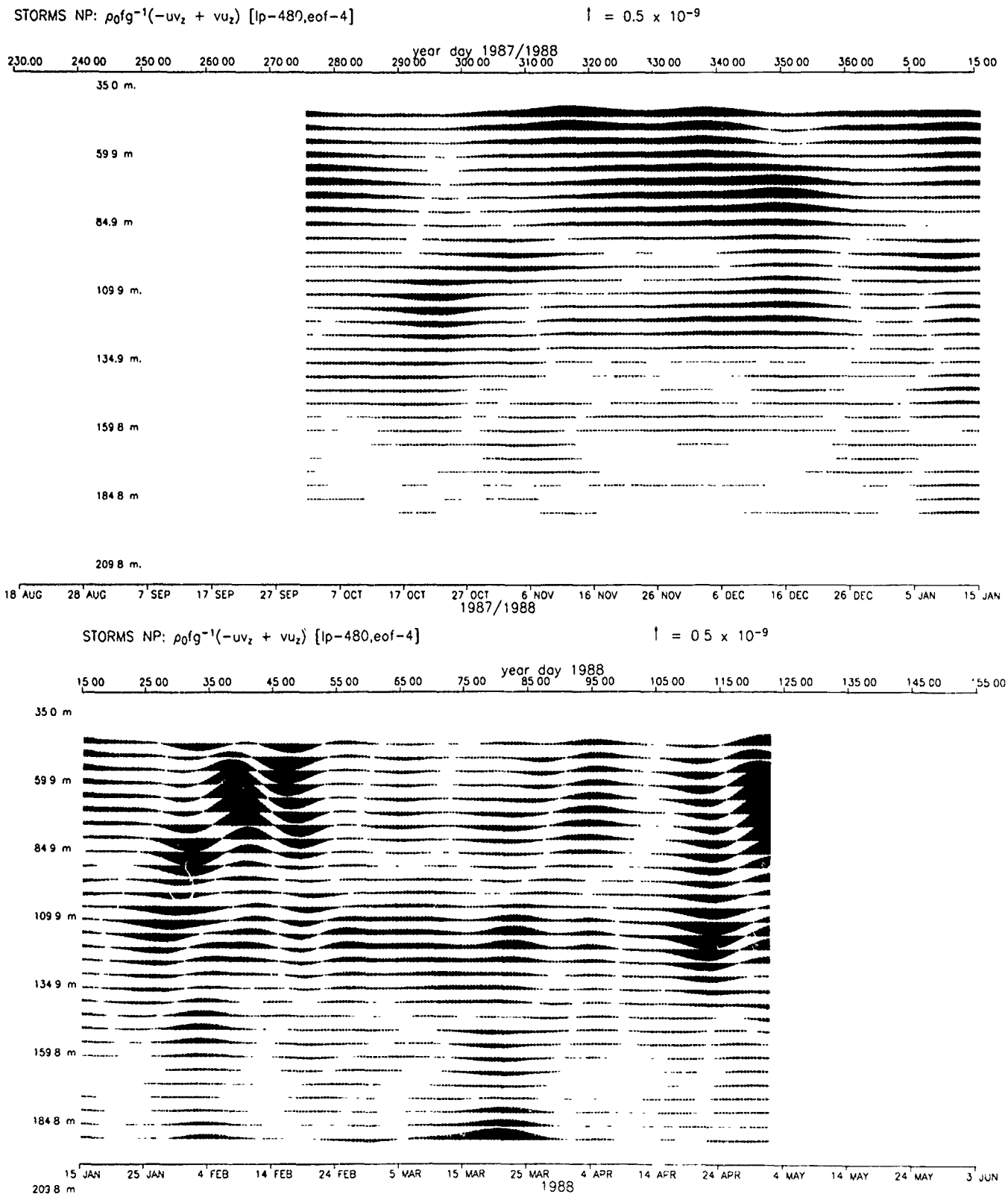


Figure 9. Horizontal advection of potential density calculated assuming vertical shear is geostrophic and plotted with the same scale as Fig. 8.

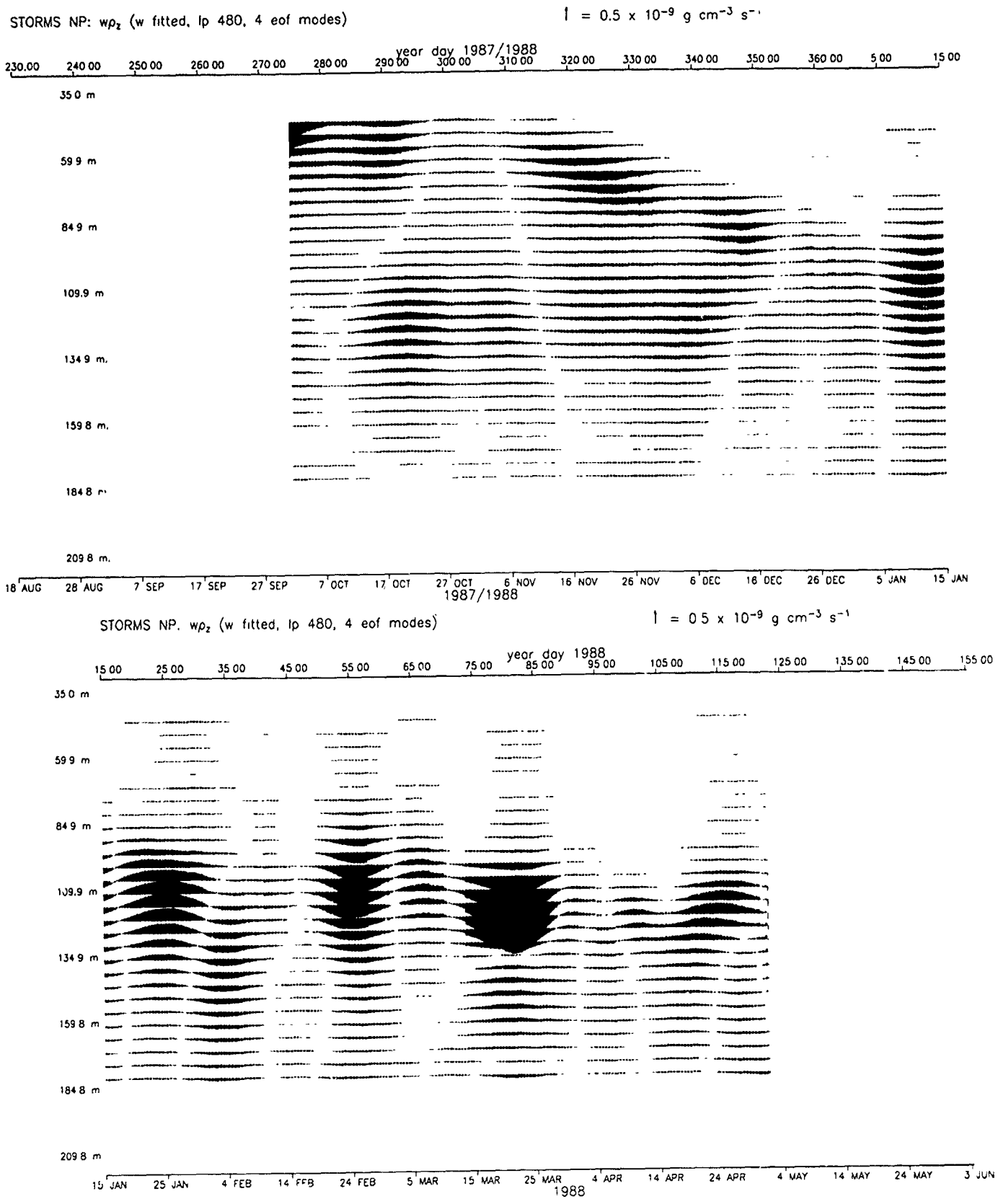


Figure 10. Vertical advection of potential density plotted as in Figs. 8 and 9.

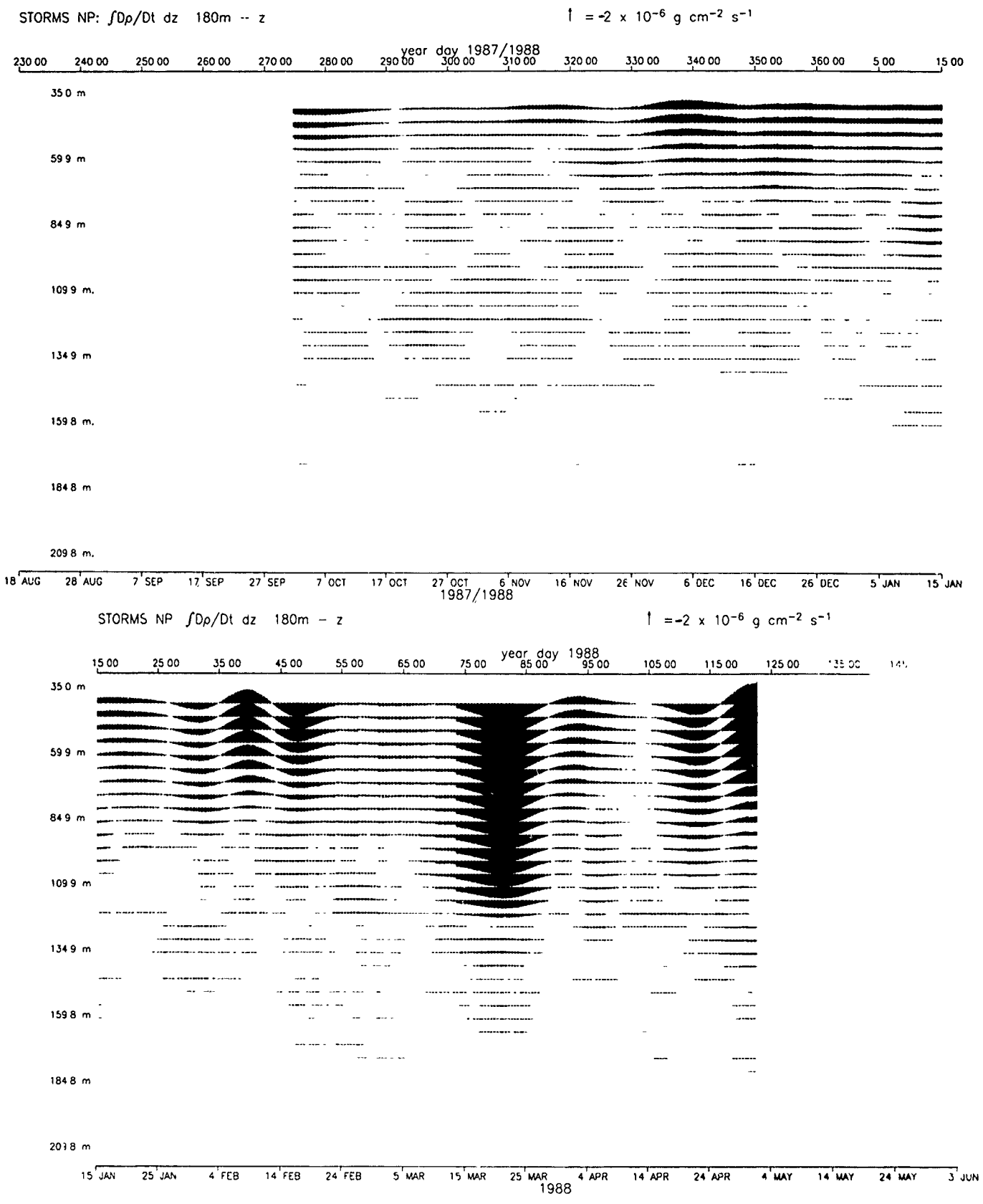


Figure 11. Vertical mass flux inferred from the vertical integral of the substantial derivative of density plotted with a scale of -2 x 10^{-5} kg/m^2/s (equivalent to a buoyancy flux scale of 2 x 10^{-5} W/kg, i.e., an upward stick corresponds to upward buoyancy flux).

motions is the layer depth h times half the mean square current current. A decay time for mixed layer near inertial motions can then be estimated as the ratio of this to the rake of kinetic energy removal by turbulent dissipation within the pycnocline:

$$t_{decay} = hh'^{-1} \langle u^2/2 \rangle / \varepsilon$$

where h′ is the depth range over which dissipation at rate ε is taking place within the pycnocline. This decay time works out to be about one week in the OCEAN STORMS observations.

The decay of near-inertial motions in the mixed layer is presumably governed by two processes: wave propagation and frictional damping. Calculations of D'Asaro (1991) demonstrate that both linear and nonlinear model simulations of the October 4 storm leave near inertial energy in the mixed layer much longer than the OCEAN STORMS observations indicate. His best simulation of the character of the descending internal wave beam include frictional damping. Our calculations of the amount of mixing that takes place as the mixed layer and seasonal pycnocline descend together imply that enough turbulent dissipation of kinetic energy must take place at the same time to extract a considerable fraction of the energy in the internal wave field. This energy is lost ultimately to dissipation at molecular scales. The picture that emerges is that a substantial portion of the near-inertial internal wave energy that would result from a given storm wind stress pattern is lost to friction through generation of shear at the base of the mixed layer in the top 20 m or so of the pycnocline. Actual wave amplitudes are reduced accordingly and a significant fraction of the work done by wind stress goes into raising potential energy of the water column through mixing by shear instability within the upper pycnocline.

REFERENCES

D'Asaro, E. A., 1989: The decay of wind-forced mixed layer inertial oscillations due to the β effect.. J. Geophys. Res., *94*, 2945-2056.

D'Asaro, E. A., 1991: The need for better internal wave models. this volume

Eriksen, C. C., J. M. Dahlen, and J. T. Shillingford, Jr., 1982: An upper ocean moored current and density profiler applied to winter conditions near Bermuda. J. Geophys. Res., *87*, 7879-7902.

Gregg, M. C., 1987: Diapycnal mixing in the thermocline: A review. J. Geophys. Res., *92*, 5249-5286.

Kundu, P. K., and R. E. Thomson, 1985: Inertial oscillations due to a moving front. J. Phys. Oceanogr., *15*, 1076-1084.

Pollard, R. T., P. B. Rhines, and R. O. R. Y. Thompson, 1973: The deepening of the wind-mixed layer. Geophys. Fluid Dyn., *3*, 381-404.

Price, J. F., R. A. Weller, and R. Pinkel, 1986: Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling and wind mixing. J. Geophys. Res., *91*, 8411-8427.

Osborne, T. R., 1980: Estimates of the local rate of vertical diffusion from dissipation measurements. J. Phys. Oceanogr., *10*, 83-89.

Peters, H., M. C. Gregg, and J. M. Toole, 1988: On the parameterization of equatorial turbulence. J. Geophys. Res., *93*, 1199-1218.

ON THE STATISTICS OF FINE SCALE STRAIN IN THE THERMOCLINE

Robert Pinkel
Marine Physical Laboratory, 0213
Scripps Institution of Oceanography
University of California, San Diego
La Jolla, California 92093-0213

Steven Anderson
Marine Physical Laboratory, 0213
Scripps Institution of Oceanography
University of California, San Diego
La Jolla, California 92093-0213

INTRODUCTION

The interrelation between the large vertical-scale motion field, the fine scale (1-100 m) field, and the micro-scale is poorly understood. It is thought that motions become significantly non-Gaussian at vertical scales shorter than about 50 m (Holloway 1983). However, "non-Gaussian" is an extremely general description of a flow field. In this work we examine the fine-scale statistics of the vertical strain field in the sea. Strain is here defined as the vertical gradient of isopycnal vertical displacement. The objective is to identify a specific class of probability density functions (pdfs) which characterize the fine scale field during its transition from highly skewed (micro-scale) to Gaussian (large scale) behavior.

Theoretical studies of non-linear processes often assume quasi-Gaussian statistics. The non-linear condition is approximated through a perturbation expansion about a zeroth-order Gaussian state. If a joint-normal form is assumed for the pdf of vertical displacement of isopycnal pairs, $\eta(\rho_i),\eta(\rho_j)$, it is easily shown that the pdf of isopycnal separation Δz_{ij} is also Gaussian (Desaubies and Gregg 1981, henceforth DG81). There is always a finite probability that $\Delta z_{ij}(t)$ will vanish, resulting in singular values for vertical gradients of passive scalars $\frac{\partial \theta}{\partial z} = \frac{\theta(\rho_i) - \theta(\rho_j)}{\Delta z_{ij}(t)}$. From a mathematical viewpoint, a Gaussian zeroth order state is an awkward starting point for the description of the fine scale field.

Knowledge of the pdfs of isopycnal separations enables the statistical modeling of a number of phenomena of physical interest in addition to vertical gradient fluctuations. Measurements of variance can be used to infer skewness, kurtosis, etc., once a form for the separation pdf is established.

This work complements an introductory paper, Pinkel et al 1991, (Henceforth P91). A more complete discussion of statistical matters is presented in Pinkel and Anderson 1991 (PA91).

Isopycnal displacement data obtained in the 1986 experiment PATCHEX are used in this study. The data are derived from a series of nine thousand CTD profiles from the sea surface to 560 m, obtained over an 18.75 day interval. Three types of isopycnal separation statistics are accumulated. Probability density functions of isopycnals separation are formed at varying mean separations $\overline{\Delta z}$. These statistics are gathered in both fixed depth (Eulerian) and fixed density (semi-Lagrangian) reference frames. In addition, discrete probability functions are formed, describing the probability of occurrence of varying numbers of isopycnals in fixed vertical intervals.

The discrete probability distributions formed in the fixed depth intervals are found to be very nearly Poisson for vertical bins, H, of order 3 m and greater. The corresponding Eulerian and Lagrangian pdfs of isopycnal separation are very nearly gamma pdfs, as would be predicted from elementary Poisson theory.

This surprising finding enables the simple modeling of strain and gradient statistics in both reference frames. A single parameter, λ, specifies the entire description at all vertical scales $\overline{\Delta z} > 3$ m.

The measurements and data are described next. Statistical results are then presented, followed by discussion of these findings.

MEASUREMENTS

The data considered here are a set of 9,000 CTD profiles from the surface to 560 m. These were obtained during October 1986 from the Research Platform FLIP. FLIP was located at 34°N, 127°W, approximately 500 km west of Point Conception, CA. Position was maintained to within 300 m by a two point moor. Water depth at the site is 4 km.

The CTD's used are Seabird Instruments model SBE-9s. Two such instruments are profiled. The upper unit is cycled from the surface to 320 m. The lower system covers the depth range 250-560 m. Profiles are repeated at 3 min intervals. The drop rate of the sensors is approximately 3.5 m/s. It is not necessary to pump water through the conductivity cell to achieve adequate spatial resolution at this drop rate.

Following a time response correction to the temperature sensor, the vertical resolution of both the temperature and conductivity sensors is limited to 2 m by a low pass filter (Sherman 1989). Density profiles are then produced. A set of 560 isopycnals, of mean separation 1 m, is then followed for the duration of the data experiments (Fig. 1). The experimental approach is discussed in greater detail in P91.

The three hour record presented in Figure 1 represents a small portion of the 18.75 day data set. In it one sees a general trend toward decreasing isopycnal depth, associated with the baroclinic tide. Superimposed on this trend are higher frequency (1-2 cph) internal waves. These are extremely coherent with depth. Against this large scale background, the fine scale straining of the density field is seen. Isopycnals converge to form "sheets" of high vertical gradient and diverge, forming low gradient "layers". The typical time-scale for the fine scale variation appears to be from one-half to several hours, in this short record.

Protagonists in the present study are:

the isopycnal separation $\Delta z_{ij} \equiv z(\rho_i,t) - z(\rho_j,t)$

the normalized separation $\gamma_{ij}(t) = \Delta z_{ij}(t)/\overline{\Delta z_{ij}}$

and the finite difference strain $\hat{\gamma}_{ij} = \gamma_{ij}(t) - 1$.

The finite difference strain can be thought of as an approximation to the actual strain, $\partial\eta(\rho,t)/\partial z$, where $\eta \equiv z(\rho,t) - \overline{z}(\rho)$. Alternatively, separation variance statistics can be considered as precise estimates of the structure function

$$F(\overline{\Delta z}) = 1/\pi \int_{-\infty}^{\infty} (1 - \cos k\overline{\Delta z})S(k)dk = \langle \gamma^2 | (\overline{\Delta z}) \rangle - 1 \quad \text{(Tennekes and Lumley, 1972).} \qquad 1)$$

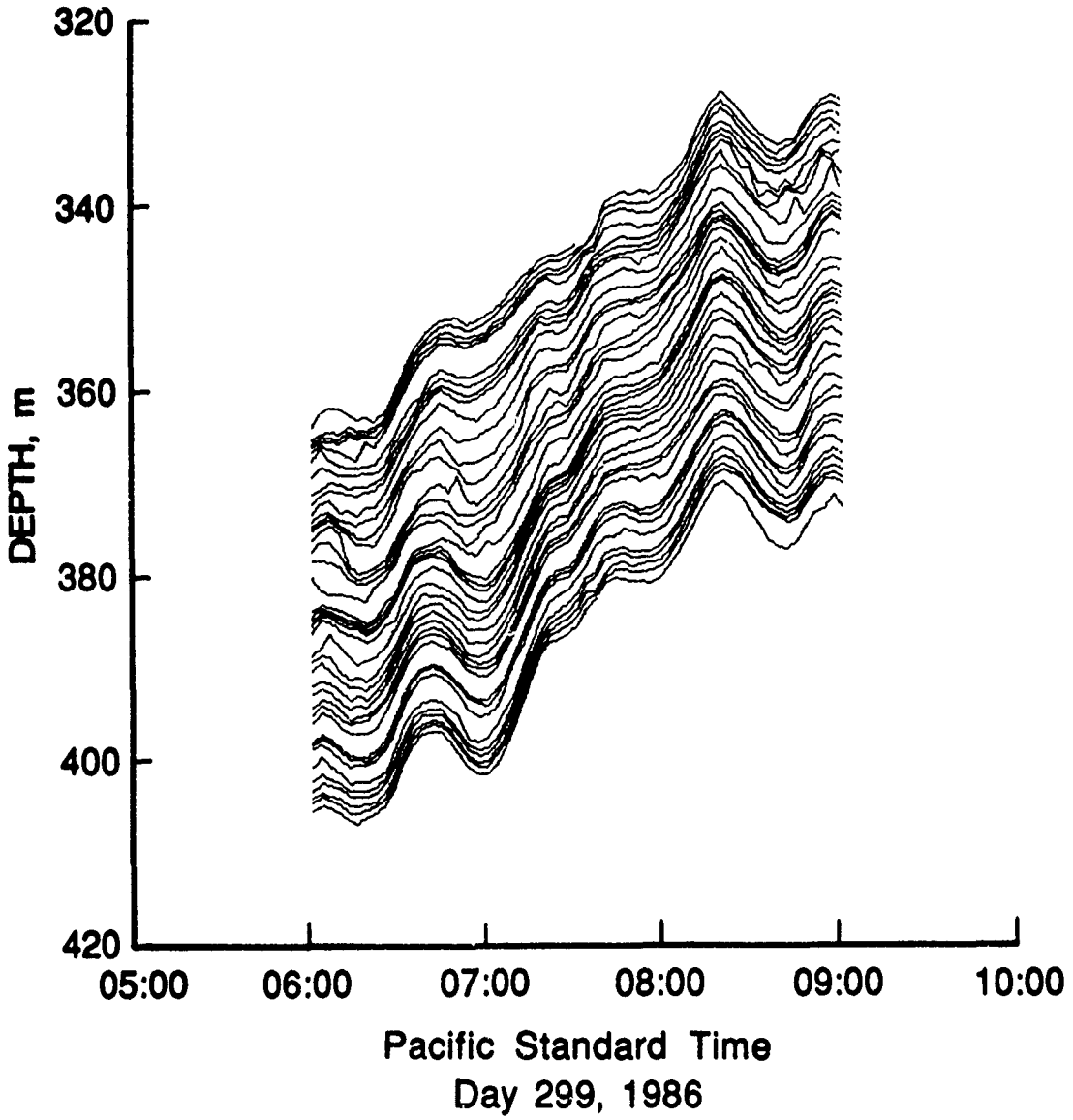


Fig. 1. An example of isopycnal depth fluctuations as seen in the PATCHEX data set. The statistics of isopycnal separation are the focus of the present study.

Here, S(k) is the vertical wavenumber spectrum of strain.

Obviously, no technique can produce quality estimates of the spectrum at high wavenumber in the absence of accurate measurements at small vertical scale Δz. Evenly spaced estimates of F at intervals of Δz = n meters, n = 1, N, result in a wavenumber spectral estimate with Nyquist wavenumber of .5 cpm, and wavenumber resolution of N^{-1} m (McKean, 1974).

Three aspects of the measurements, sensor noise (ε), resolution, and statistical precision impact the present discussion. Sensor noise results in error in the estimates of density profiles. Noise has both correlated

and uncorrelated aspects. At separations greater than a few meters, the noise which influences the estimate of one isopycnal is independent of that influencing another. Strain variance estimates are biased in the presence of noise.

$$\langle\gamma^2\rangle = \frac{\langle\Delta z^2\rangle}{\overline{\Delta z}^2} + \frac{2\langle\epsilon^2\rangle}{\overline{\Delta z}^2} \qquad 2)$$

The error in the estimate of the depth of a given isopycnal is approximately .25 m rms in the PATCHEX data set. This corresponds to a strain variance bias of .12 at $\overline{\Delta z}$ = 1 m, decreasing to .005 at $\overline{\Delta z}$ = 5 m. Thus the bias can be as large as 20 % of the total signal at $\overline{\Delta z}$ = 1 m (assuming the errors associated with two closely spaced isopycnals are indeed uncorrelated). Since the variance bias decreases as $\overline{\Delta z}^{-2}$ while the strain signal decreases more nearly as $\overline{\Delta z}^{-1}$, (Fig. 7) random error quickly becomes insignificant.

Noise has an effect on the pdfs of separation. The observed pdfs are a convolution of the true pdf of separation with the pdf, presumably Gaussian, of the noise. Thus, observed pdfs will be broader than the true, with their variance increased by the variance of the noise. This effect will be significant at small separation $\overline{\Delta z}$.

Resolution is a concern when considering closely spaced isopycnals. One would expect to observe fewer than the actual number of instances of close isopycnal spacing (small γ) in pdfs of separation. Conversely, fewer observations of "many isopycnals found in a fixed size bin", are expected in the discrete probability functions presented below. The finite resolution of the CTD is particularly damaging to the present study given that the spatial autocorrelation of strain transitions from positive (...if two adjacent isopycnals are squeezed into a sheet, it is likely that the immediate neighboring isopycnals are being drawn into the sheet...) to negative (...if five isopycnals are being drawn into a sheet, it is likely that there will be an absence of isopycnals, a layer, five m away) at about the resolution scale of the CTD, 2 m.

At sufficiently small scales the strain correlation is high and positive. A principal conjecture of P91 is that strain statistics approach log-normal form in this region. If the log-normal regime indeed exists, it occurs at scales unresolvable by the present CTD. Further exploration of the log-normal issue awaits improved instrumentation.

Statistical precision is often a concern when trying to compare pdfs estimated from data with classical functional forms. In this work, estimates of pdfs are formed in four 100 m depth ranges, 100-200 through 400-500 m. Thus, 9×10^5 measurements (9,000 profiles * 100 isopycnals) are available for each 100 m estimate. At issue is the fraction of these points that is statistically independent. This question is discussed by Briscoe (1977) in his investigation of the Gaussianity of the horizontal velocity and vertical displacement of the IWEX data set. From numerical simulations Briscoe finds that the effective decorrelation time for displacement is of order 1/2 day. A nineteen day displacement time series would consist of 38 independent samples, corresponding to 76 degrees of freedom.

The situation is more complicated for the strain field, given the significance of non-linear distortion at small scale. P91 demonstrate that the characteristic lifetime of "layer" events ($\gamma > 1$) is shorter than "sheet" events ($\gamma < 1$). The number of independent estimates of $\gamma < 1$ events per unit time is less than that of $\gamma > 1$ events. Monte Carlo simulations of the strain field will not model this effect appropriately unless the bi-spectrum of the field is properly specified.

In PA91 an attempt to estimate the statistical precision of the strain, as a function of γ itself, is presented. The variability of independent estimates of the strain pdf, from depth to depth at fixed γ, is used to determine the effective number of degrees of freedom. When large mean separations, $\Delta z \geq 3$ m, are considered, the observations are consistent with an 100 degree of freedom process, more or less independent of γ. At very small mean separations, variations in statistical stability with γ is clearly seen. The effective number of degrees of freedom varies from 80 to 100 as γ increases from .2 to 2.

THE PDF OF STRAIN AND DISPLACEMENT

Joint pdfs of strain and displacement have been formed using the PATCHEX isopycnal displacement time series. The pdfs are formed for isopycnal pairs of mean separation $\Delta z = 1 - 50$ m. For each mean separation, the pdfs are binned into 100 displacements ($\pm$ 50 m) by 100 strains ($\gamma = 0$ to 5) in four depths zones ($\bar{z}$ = 100 - 200 through 400 - 500 m).

Two sets of joint pdfs are produced. Lagrangian pdfs are formed by tracking the evolution of specific isopycnals pairs (ρ_i,ρ_j) through the 19 day data set. Corresponding Eulerian pdfs are formed by tracking the separation between that pair of isopycnals, separated in the mean by Δz, which is spanning a specific fixed depth, z_0, at each instant of time. The Eulerian study is repeated for fixed depths of 100-500 m, at 1 m increments. The resulting pdfs are averaged into 100-200 through 400-500 m bins, in correspondence with the averaging used in the Langrangian study.

A representative joint pdf of strain and displacement is presented in Fig. 2. The pdf is formed in a Lagrangian frame, using isopycnals of mean separation 3 m. Although not clearly apparent in the figure, it can be shown (PA91) that the data are consistent with the assumption that the joint pdf is separable, and that the pdf of average displacement $\eta_{ij} = (\eta(\rho_i) + \eta(\rho_j))/2$ is Gaussian. The pdf of strain is clearly non-Gaussian, with observations of close isopycnal spacing ($\gamma < 1$) more frequent than those of large separation ($\gamma > 1$).

THE PROBABILITY DENSITY FUNCTION OF STRAIN

The joint pdfs previously discussed can be integrated with respect to displacement to produce univariate pdfs of strain. Strain pdf are presented in Figs. 3 and 4 for the semi-Lagrangian and Eulerian frames respectively, for mean separations of Δz of 1-10 m. Sample pdfs have been formed for mean separation as great as 50 m. At scales greater than 10 m these appear very nearly Gaussian. Nevertheless, skewness and kurtosis estimates are significant to separations of order 30 m.

The observed pdfs have been fit to a variety of classical forms, including Rayleigh, Weibull, log normal and gamma. Significant discrepancies are subjectively apparent in all comparisons, with notable exception of the gamma pdf, which fits very well (Figs. 3, 4, light curves). The gamma pdf has the form

$$G(x) = \frac{\beta^{\alpha+1} x^{\alpha} e^{-\beta x}}{\Gamma(\alpha)} \qquad 3)$$

with mean $\langle x \rangle = \alpha/\beta$ and variance $\langle x^2 \rangle - \langle x \rangle^2 = \alpha/\beta^2$ (eg. Papoulis 1984)

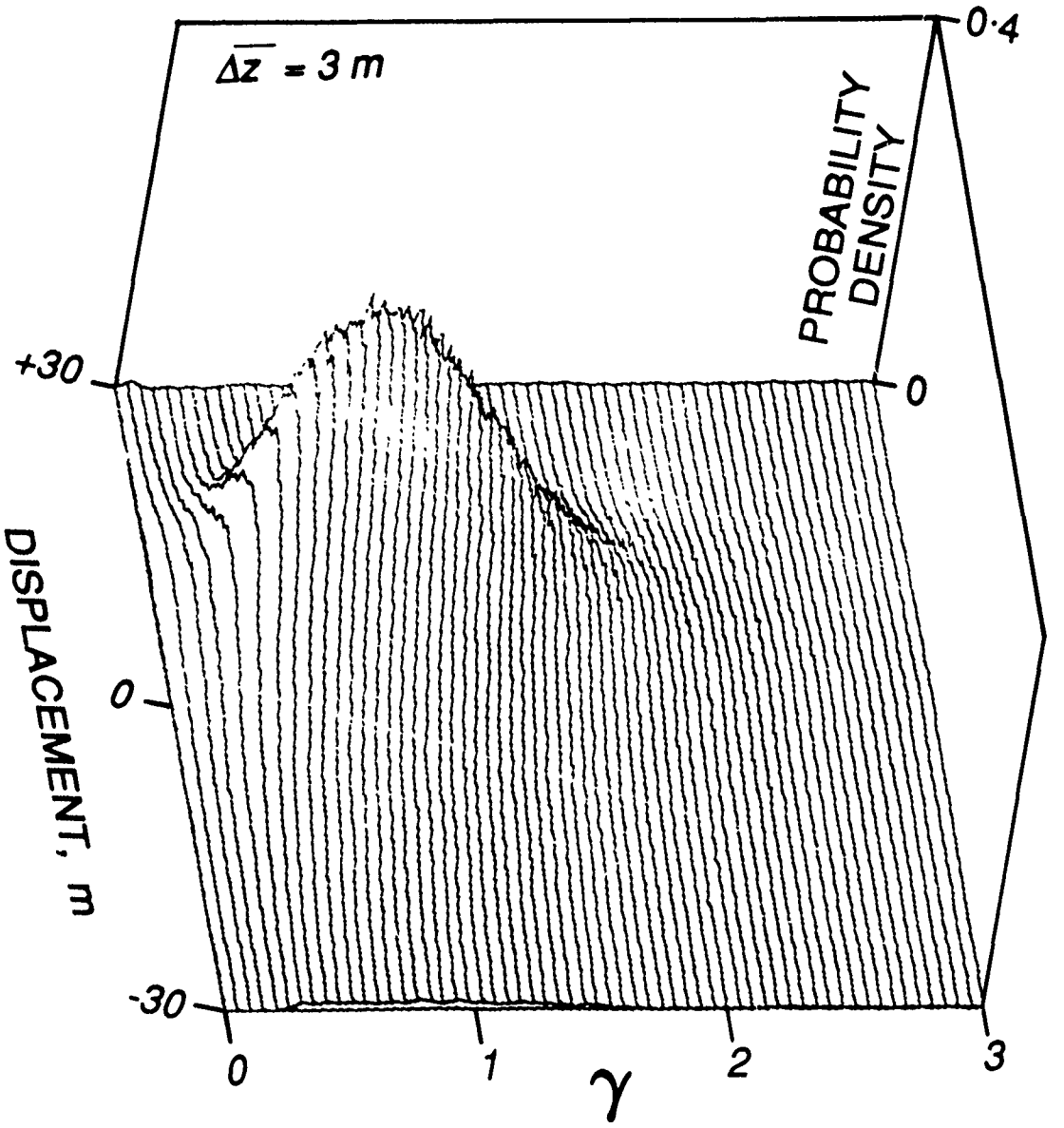


Fig. 2. An example of a joint probability denotes function of strain and displacement. This estimate was formed in a semi-Lagrangian frame, considering isopycnal pairs separated in the mean by 3 m. The PATCHEX data are consistent with the hypothesis that strain and displacement are independent quantities, with displacement obeying Gaussian statistics.

The semi- Lagrangian data are constrained to have $\langle\gamma\rangle = 1$, $\langle\Delta z\rangle = \overline{\Delta z}$, by initial choice of isopycnals. Hence $\alpha = \beta\overline{\Delta z}$. The fits presented in Fig. 3 are thus one parameter fits, with sample variance matched to the model variance. The Eulerian observations are not constrained to unity mean. The fits are thus two parameter fits. The observed mean and variance are used to set model pdf parameters in Fig. 4.

The model gamma pdf is seen to fit the observations well in the 200-300 m depth range, except at separations less than 4 m. The fits are comparable in the other depth ranges, with the exception of the 300-400

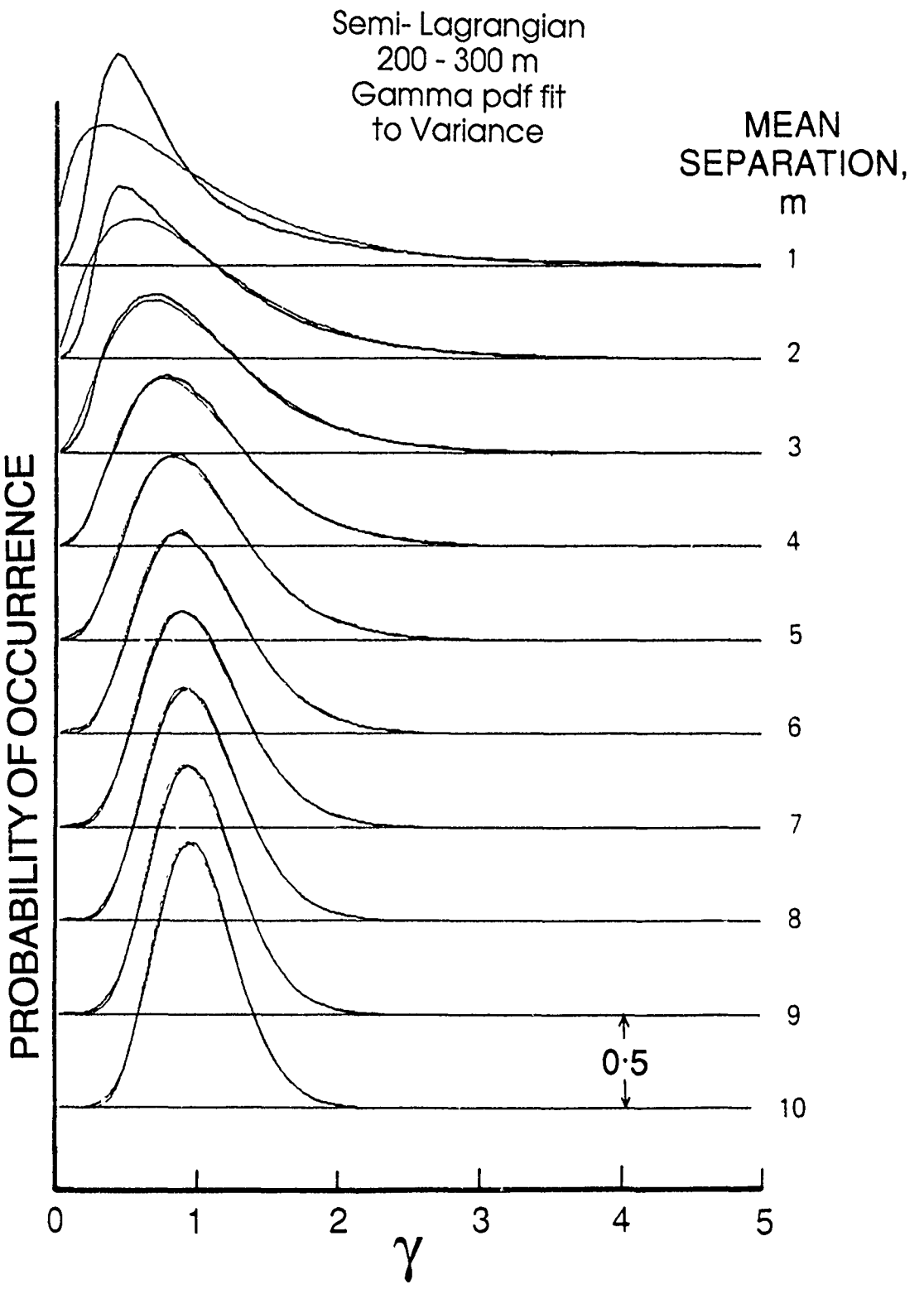


Fig. 3. Probability density functions of normalized separations, γ, formed in a semi-Lagrangian frame, for mean isopycnal separations 1-10 m. Light lines give model gamma pdfs, constrained to have unity mean and the observed variance. Data from 200-300 m depth zone are presented.

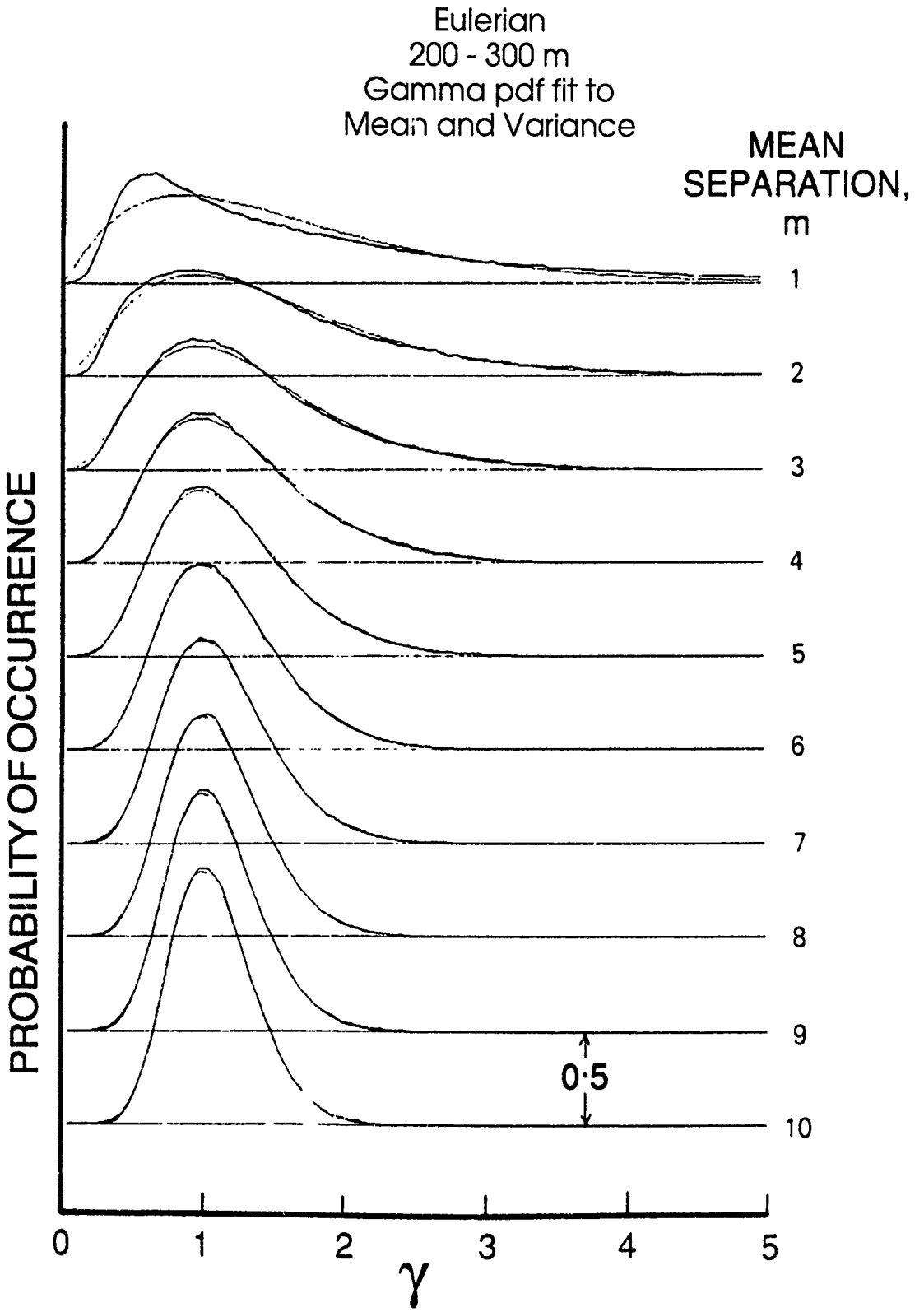


Fig. 4. Probability density functions of normalized separation, γ, as in Fig. 3 except formed in an Eulerian frame. Light lines give model gamma pdfs, constrained to have mean and variance identical to the observations. Data from 200-300 m depth zone are presented.

m interval, where the Lagrangian pdfs appear distorted at small γ over a range of $\overline{\Delta z} = 3 - 7$ m. This issue is addressed in PA91. The fits could be improved by employing a least squares fitting procedure. However, optimizing the fit is not the point of the present exercise.

Gamma pdfs are associated with the classical theory of Poisson processes. They describe the statistics of separation between the occurrence of Poisson "events" (Papoulis, 1984). If, indeed, a simple Poisson statistics describe the non-Gaussian behavior of the fine scale field, the problem of modeling the motion field in this regime can be significantly advanced.

POISSON STATISTICS

Poisson statistics describe the occurrence of discrete "events". If the probability of occurrence of these events is uniform and the occurrence of one event in no way influences the occurrence of any other events, Poisson statistics will apply (Papoulis, 1984). The Poisson probability function gives the probability that the number of events which occur in a dimensional interval of length H will equal any specified value, k.

$$P(n = k \mid H) = \frac{(\lambda H)^k e^{-\lambda H}}{k!} \qquad 4)$$

The Poisson probability function has the interesting property that the mean number of "events" occurring in an interval H, λH, is equal to the variance in the number of events.

It is not clear exactly what constitutes a "Poisson event" in the context of the fine scale variability in the thermocline. We have tracked a set of isopycnal surfaces with mean spacing arbitrarily chosen to be 1 m. A Poisson-like investigation can be conducted, tracking the number of isopycnals which are found to occur in fixed depth bins of varying size H. This is done for the four 100-m depth ranges used in the previous studies. The results are presented for the 200-300 m range in Fig. 5. Not surprisingly, the mean number of isopycnals found in a depth bin of thickness H is H, given our initial choice of isopycnals mean separation. The oceanic signal is seen in the variation of the higher moments with H. To facilitate comparison with the classical Poisson distribution, which is constrained to have mean equal to variance, the observed probability functions are re-scaled in terms of equivalent "Poisson events", $\hat{n} = \lambda n$. Here $\hat{n}$ is the number of "events" alleged to occur in the fixed depth bin, n is the observed number of isopycnals, and λ is a scale factor chosen such that the mean number of "events" is equal to the event variance. The fitting of model Poisson probability functions (Fig. 5, light line) to the data is accomplished by setting the model mean equal to the rescaled observed mean, λH. The fit is impressive, with significant discrepancies apparent only in 1 and 2 bin sizes.

THE THERMOCLINE AS A POISSON PROCESS

The excellent fits of the Gamma and Poisson pdfs to the observations at vertical scales greater than 3 m encourages the adoption of Poisson statistics as a zeroth order model for the thermocline. In this section we review relevant aspects of Poisson theory. Subsequently we discuss observed departures from the predictions of the Poisson model, and why these departures are to be expected in a real fluid.

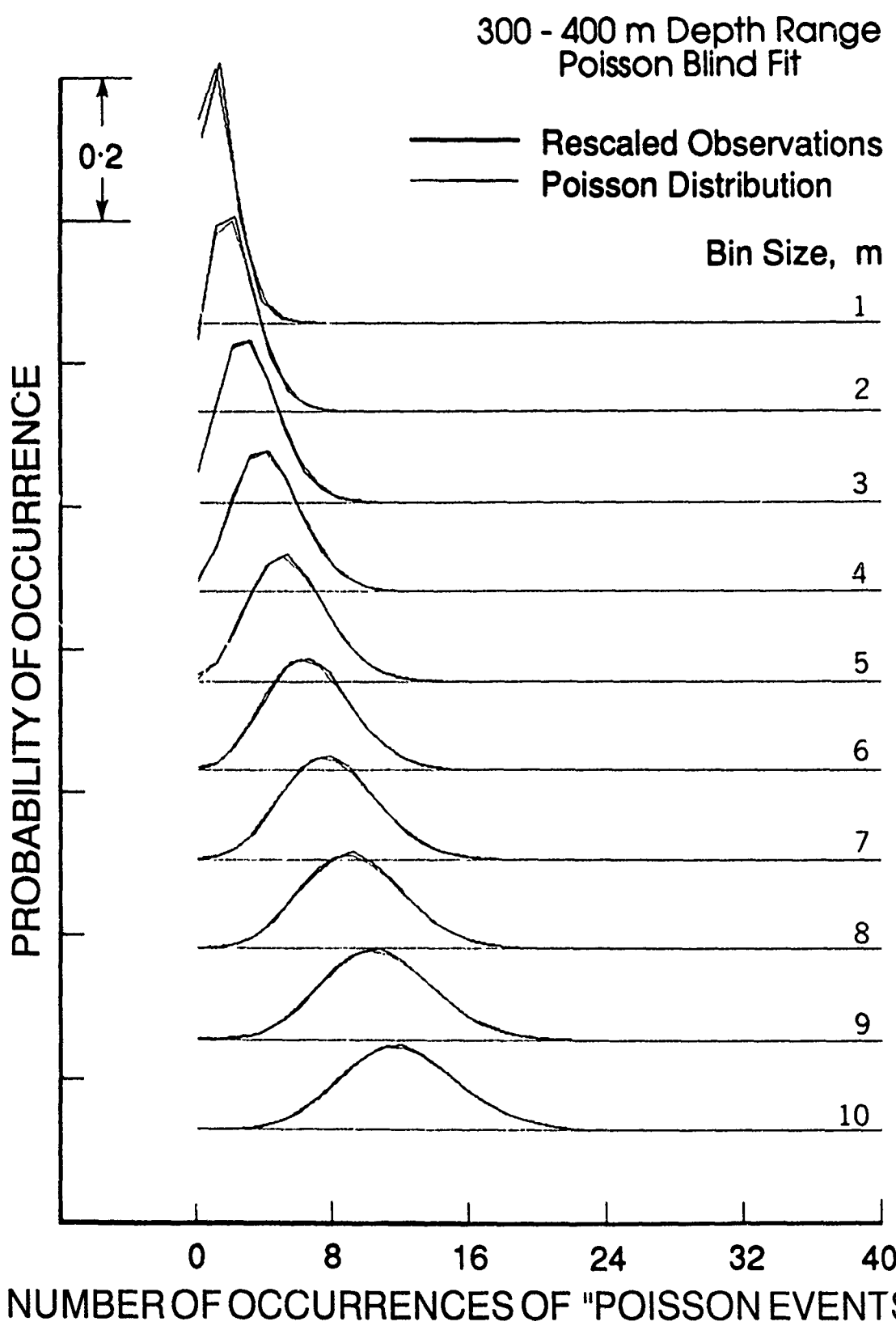


Fig. 5. Probability function of the number of isopycnals found in fixed depth bins of size H = 1-10 m. The observed number of isopycnals is re-scaled by the factor λ, such that the sample mean equals the variance. The light curve gives a theoretical Poisson distribution. Data from the depth range 300-400 m are presented here.

Papoulis (1984) notes that the statistics of separation of Poisson events take two distinct forms. If two adjacent events are considered at random, the gamma pdf

$$G(\Delta z \mid 1, \lambda) = \lambda e^{-\lambda \Delta z}/\Gamma(1) \qquad 5)$$

describes the separation statistics. However if the two adjacent random points are constrained to bracket some fixed point, z_0, the pdf of separation is given by

$$G(\Delta z \mid 2, \lambda) = \frac{\lambda^2 \Delta z e^{-\lambda \Delta z}}{\Gamma(2)}. \qquad 6)$$

The added constraint of requiring the points to bracket a fixed point alters the statistics.

We identify (6) as the appropriate pdf for isopycnal separation in an Eulerian frame, at the fixed depth z_0. Expression (5) is the appropriate semi-Lagrangian separation pdf, for a mean separation $\overline{\Delta z} = \lambda^{-1}$. When considering the pdf of distance between an arbitrary random point and one removed by (n – 1) intermediate points, the order of the appropriate gamma pdf is simply increased e.g.,

$$G(\Delta z \mid n, \lambda) = \frac{\lambda^n \Delta z^{n-1} e^{\lambda \Lambda z}}{\Gamma(n)}. \qquad 7)$$

The continuous observations presented here are consistent with the interpretation that Poisson "events"" occur every λ^{-1} meters. Thus the discrete parameter n can be replaced by its continuous equivalent $\lambda^* = \lambda \overline{\Delta z}$.

$$G^L(\Delta z \mid \overline{\Delta z}, \lambda) = \lambda \cdot \frac{(\lambda \Delta z)^{\lambda^*} e^{-\lambda \Delta z}}{\Gamma(\lambda^*)} \qquad 8)$$

$$G^E(\Delta z \mid \overline{\Delta z}, \lambda) = \frac{\lambda(\lambda \Delta z)^{\lambda^* + 1}}{\Gamma(\lambda^* + 1)} e^{-\lambda \Delta z} = \gamma\, G^L(\Delta z \mid \overline{\Delta z}, \lambda) \qquad 9)$$

The skewness seen in the semi-Lagrangian frame reflects the relatively passive advection of the density field by the velocity field, on the fine scale. Isopycnals which find themselves close together experience nearly identical advecting velocity fields. Hence they remain together for a relatively long time. The Eulerian pdf is less skewed than its semi-Lagrangian counterpart, for a given mean separation $\overline{\Delta z}$. This reflects the fact that when isopycnal pairs are traveling closer together than average, they are less likely to span a specific reference depth than when they are farther apart. Not surprisingly, the chance of spanning a given point while simultaneously having separation Δz increases linearly with Δz. This result is reflected in equations 5,6 and 8,9, and is demonstrated for arbitrary strain pdfs in PA91, under the assumption that strain and displacement are independent quantities.

The immense power in the Poisson model results from the fact that the single dimensional parameter λ (m^{-1}) describes the variability of strain not just at a particular vertical scale $\overline{\Delta z}$, but at all scales.

Thus, mean isopycnal separation is given by

$$<\Delta z_L> = \frac{\lambda^*}{\lambda} \equiv \overline{\Delta z} \qquad 10)$$

$$<\Delta z_E> = \frac{\lambda^* + 1}{\lambda} = \overline{\Delta z} + \lambda^{-1}$$

Separation variance is:

$$<\Delta z^2>_L - \overline{\Delta z}^2 = <\Delta\eta^2>_L = \lambda^* / \lambda = \overline{\Delta z} / \lambda$$

$$<\Delta z^2>_E - <\Delta z>_E^2 = <\Delta\eta^2>_E = \lambda^* + \frac{1}{\lambda^2} = \overline{\Delta z} / \lambda + 1 / \lambda^2 \qquad 11)$$

The growth of variance with mean separation is linear. If this pattern were seen in the data down to arbitrarily small separation, it would correspond to a white vertical wavenumber spectrum of strain,

$$S_\gamma(k) = \lambda^{-1} m \qquad k > 0$$

$$S_\gamma(k) = 0 \qquad k = 0 \qquad 12)$$

The corresponding isopycnal vertical displacement spectrum is of k^{-2} form.

THE DOMAIN OF VALIDITY OF THE POISSON MODEL

One anticipates departures from the Poisson model at both large and small vertical scales. This can be seen by considering the variance of isopycnal separation, Eq. 11, which is predicted to grow linearly without bound as mean separation increases. This behavior is a consequence of the fact that the model k^{-2} vertical wavenumber spectrum of displacement extends to arbitrarily low wavenumber. In a finite depth ocean, the governing physics will surely change as the scale of the waveguide thickness is approached, resulting in a departure from Poisson behavior.

At small vertical scales, of order λ^{-1}, the model itself becomes inconsistent. The Eulerian separation variance fails to vanish at $\overline{\Delta z} = 0$. This corresponds to a limit of infinite strain variance in the Eulerian frame as the vertical scale tends to zero. The Poisson model Cox number presented in Table 1 also becomes singular at small scales.

There is a simple interpretation for these a-physical aspects of the Poisson model. The lack of predictability is central to th concept of a Poisson process. The occurrence of a specific Poisson "event" in no way influences the occurrence of subsequent events. However, at sufficiently small scales in the ocean one expects the strain field to be correlated. The oceanic strain spectrum cannot be white. It presumably is band limited, with a bandwidth which is the inverse of this correlation scale. To the extent that a "sheet and layer" model of the strain field is valid, one expects the autocorrelation function of strain to assume negative as well as positive values. (If a large number of isopycnals is found concentrated in a given region, a "sheet", it is likely that there will be a relative absence of isopycnals, a "layer" nearby.) A purely Poisson model is not appropriate at small vertical scales, where the strain field is correlated.

Discrepancies between the observations and the idealized Poisson model can be seen in plots of observed variance vs. mean separation. These are presented in Figs. 6 and 7 for both the Eulerian and semi-Langrangian studies. Data are presented to vertical separations of 50 m.

Focusing attention on the semi-Lagrangian separation variance, Fig. 7 c,d, disagreement with the Poisson model is found at both smaller and larger scale. At small scales, the variance first increases more rapidly than $\overline{\Delta z}^{+1}$ ($\overline{\Delta z} < 4$ m), then more gradually ($\overline{\Delta z} > 4$ m). This corresponds to the positive and negative regions of strain correlation, as sensed by the low resolution CTDs used in the experiment.

At large separations, the variance increase never does approach the $\overline{\Delta z}^{+1}$ Poisson form, except perhaps in the 100-200 m depth range. Simple simulations suggest that this discrepancy is related to the depth variability of the wavefield. Isopycnal displacement variance changes significantly with depth, as a result of the changing profile of Vaisala frequency. WKB theory suggests that at depths removed from the sea surface or sea floor,

$$\langle \eta^2 \rangle(z) = \langle \eta^2 \rangle_0 \frac{N_0}{N(z)} \qquad 13)$$

(Garrett and Munk 1972).

While this classical expression is only approximately correct for the PATCHEX data, it could be used to rescale (WKB stretch) the displacement field to produce a data set more nearly homogeneous in depth. This has not been done here, in the interest of presenting the basic results in a model independent format.

A related check on the applicability of the Poisson model is to examine the variability in the observed estimate of λ, the Poisson "constant", as a function of bin size H or mean separation $\overline{\Delta z}$. This is presented in Fig. 8, for the fixed depth Poisson study and the semi-Langrangian isopycnal separation study. At 10 m scales λ has a minimum value of approximately 1.1 m^{-1}. Slight increases in this constant are apparent at both small and larger scales. Again, the finite correlation of the strain field is thought to cause the small scale departure from ideal Poisson behavior. The non-homogeneity of the wavefield with depth is a suggested source of the discrepancy at large scale.

THE MODELING OF VERTICAL GRADIENTS

If one abandons the assumption of Gaussianity, it simplifies the modeling of vertical gradients in the thermocline. There is no need to linearize about small departures from the mean gradient. Recalling the expression for the instantaneous gradient of a passive scalar,

$$\frac{\partial \theta}{\partial z} = \frac{\Delta \theta}{\overline{\Delta z}} \cdot \gamma^{-1}(t),$$

interest is focused on the statistics of $\mu(t) \equiv \gamma^{-1}(t)$. If $p(\gamma)$ is an arbitrary pdf of γ ($p(\gamma) \equiv 0$ for $\gamma \leq 0$), then $g(\mu) = p(\mu^{-1})/\mu^2$ gives the corresponding pdf of the scalar gradient.

Using this approach, one can derive general expressions for quantities of interest.

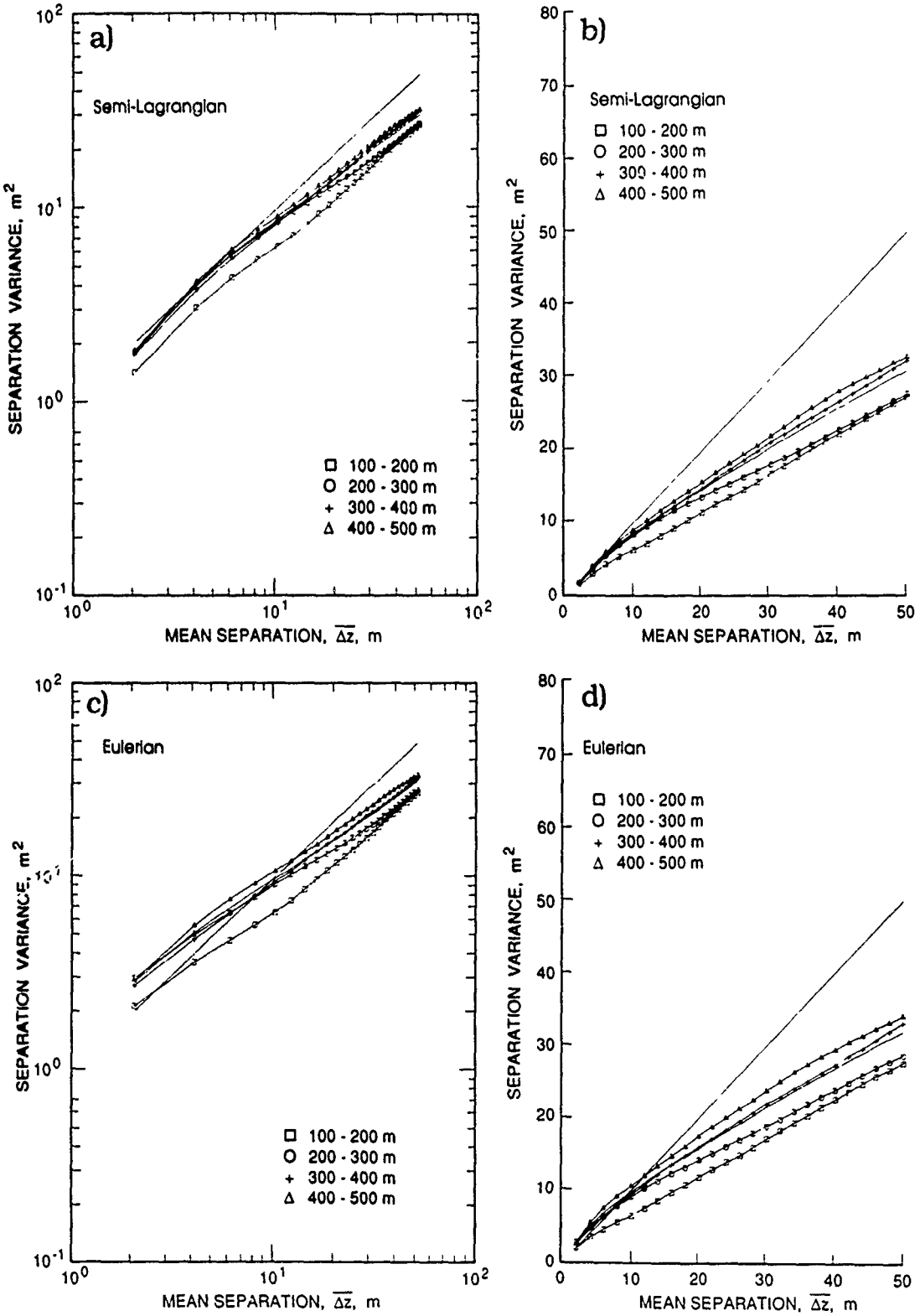


Fig. 6. Isopycnal separation variance as observed in semi-Lagrangian (a,b) and Eulerian (c,d) frames. A Poisson model would indicate a linear increase of variance with range.

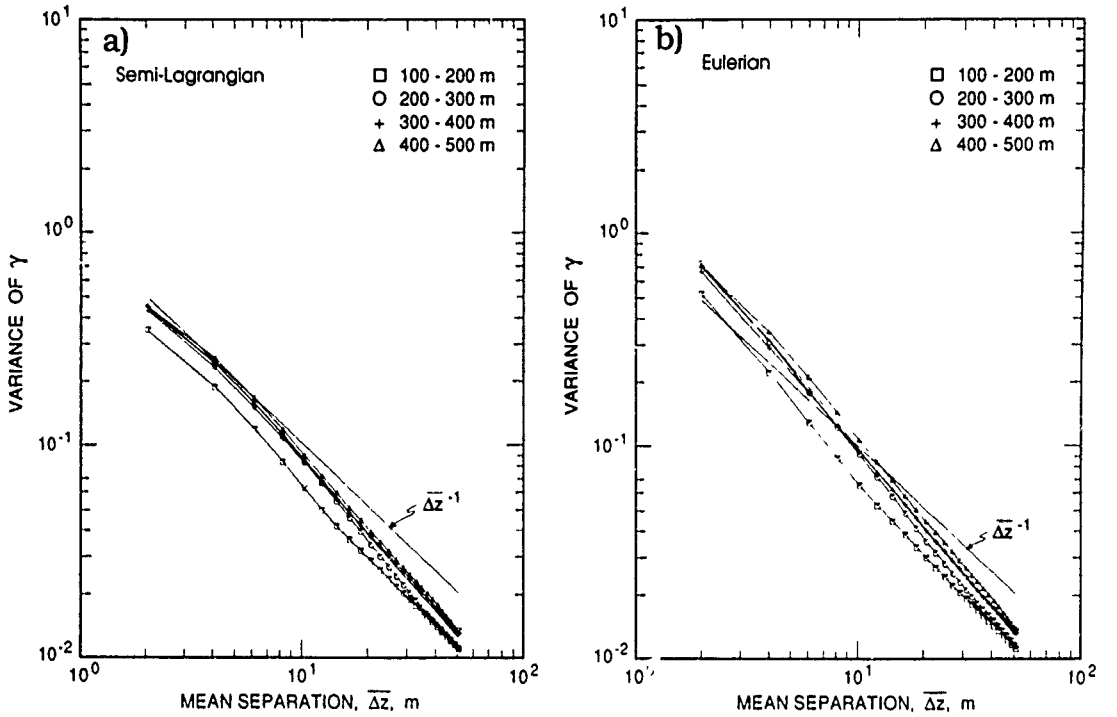


Fig 7. Normalized separation variance,1 $\langle\gamma^2\rangle$ - 1, as observed in semi-Lagrangian (a) and Eulerian (b) frames.

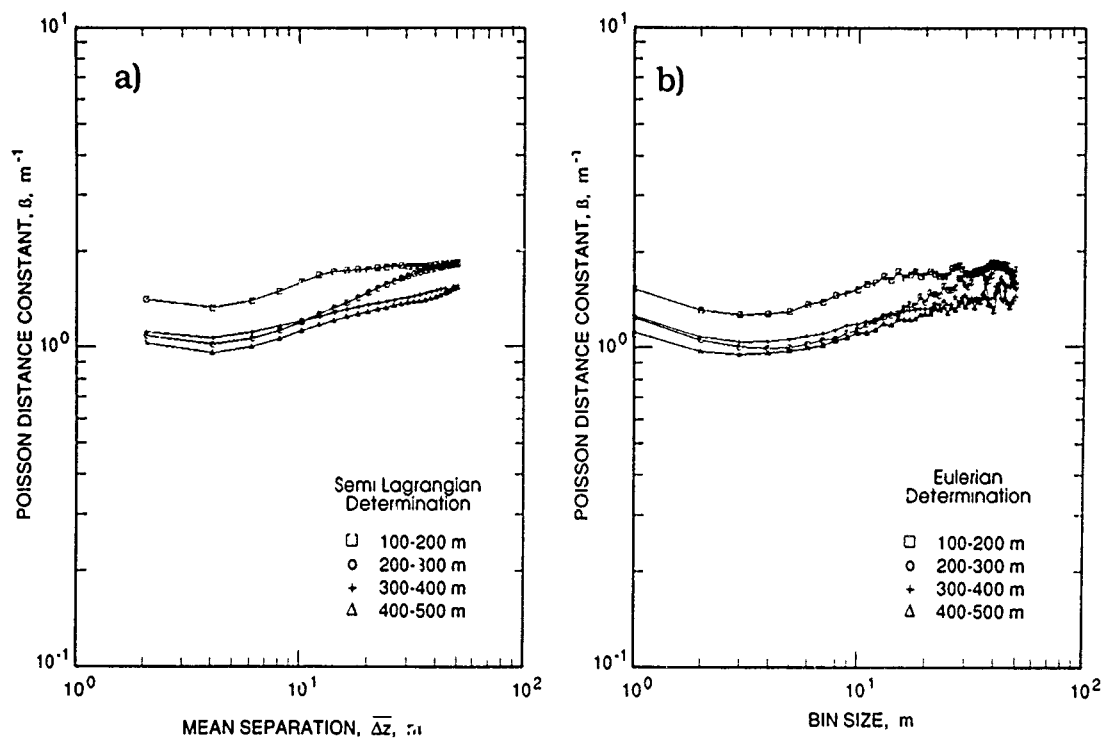


Fig. 8. a. The Poisson scale factor λ as determined by the mean to variance ratio of the semi-Lagrangian pdfs. b. The scale factor as determined by the mean to variance ratio of the number of isopycnals in fixed depth bins of varying size, H.

The Eulerian mean gradient is given by:

$$<\frac{\Delta\theta}{\Delta z}>_E = \frac{\Delta\theta}{\overline{\Delta z}}<\mu>_E = \frac{\Delta\theta}{\overline{\Delta z}}\int_0^\infty \mu g^E(\mu\,|\,\overline{\Delta z})d\mu$$

$$= \frac{\Delta\theta}{\Delta z}\int_0^\infty \gamma^{-1}p^E(\gamma\,|\,\overline{\Delta z})d\gamma$$

$$= \frac{\Delta\theta}{\Delta\overline{z}}\int_0^\infty p^L(\gamma\,|\,\overline{\Delta z})d\gamma = \frac{\Delta\theta}{\overline{\Delta z}} \qquad 14)$$

Here, we make use of the fact that $p^E(\gamma) = \gamma p^L(\gamma)$ is a general result, applying not only to the Poisson process (PA91). Also, the Eulerian Cox fine scale number

$$C_E \equiv <(\frac{\Delta\theta}{\Delta z})^2>_E / <\frac{\Delta\theta}{\Delta z}>_E^2 = \qquad 15)$$

$$\int_0^\infty \gamma^{-2}p^E(\gamma\,|\,\overline{\Delta z})d\gamma = \int_0^\infty \gamma^{-1}p^L(\gamma\,|\,\overline{\Delta z})\,d\gamma = <\mu>_L$$

In a semi-Langrangian frame, the mean gradient is seen to be

$$<\frac{\Delta\theta}{\Delta z}>_L = \frac{\Delta\theta}{\overline{\Delta z}}\int_0^\infty \gamma^{-1}p^L(\gamma\,|\,\overline{\Delta z})d\gamma \qquad 16)$$

Similarly,

$$C_L \equiv <(\frac{\Delta\theta}{\Delta z})^2>_L / \frac{\Delta\theta}{\Delta z}>_L^2 = \int_0^\infty \gamma^{-2}p^L(\gamma\,|\,\overline{\Delta z})d\gamma / (\int_0^\infty \gamma^{-1}p^L(\gamma\,|\,\overline{\Delta z})d\gamma)^2 \qquad 17)$$

These results are independent of the specific form of the separation pdf $p(\gamma\,|\,\overline{\Delta z})$.

Here we have demonstrated that the pdf of separation are very nearly gamma pdfs at vertical scales greater than 3 m. A summary of the corresponding model pdfs for strain and vertical gradients, as well as expressions for the Cox numbers are presented in Table 1. Note that, over the range of validity of this model, the fine scale Cox number is of the order two or less, far below values typically encountered at microstructure scales.

TABLE 1. Poisson Model Results

$\lambda = 1.1$ $\quad \lambda^* = \overline{\Delta z}\,\lambda$ $\quad \gamma = \Delta z/\overline{\Delta z}$ $\quad \mu = 1/\lambda$

	Semi Lagrangian	Eulerian
Strain, γ	$p^L(\gamma \mid \overline{\Delta z}) = \frac{\lambda^* (\lambda^*\gamma)^{\lambda^* - 1} e^{-\lambda^*\gamma}}{\Gamma(\lambda^*)}$	$p^E(\gamma \mid \overline{\Delta z}) = \frac{(\lambda^*\gamma)^{\lambda^*} e^{-\lambda^*\gamma}}{\Gamma(\lambda^*)}$
Vertical Gradients, μ	$g^L(\mu \mid \overline{\Delta z}) = \frac{(\lambda^*/\mu)^{\lambda^*} e^{-\lambda^*/\mu}}{\mu\Gamma(\lambda^*)}$	$g^E(\mu \mid \overline{\Delta z}) = \frac{(\lambda^*/\mu)^{\lambda^*+1} e^{-\lambda^*/\mu}}{\mu\Gamma(\lambda^*+1)}$
Cox Number	$C_L = (\lambda^* - 1)/(\lambda^* - 2)$	$C_E = \lambda^*/(\lambda^* - 1)$

The Poisson model pdfs can be tested against existing observations of fine scale gradients. The observations of Gregg 1977, DG81 are presented in Fig. 9. Using a single value $\Delta\theta/\overline{\Delta z}$ for Gregg's local mean gradient, and the parameter $\lambda_0 = 1.1\ m^{-1}$ set by our Eastern Pacific 1986 observations, it is seen that Gregg's central Pacific observations, taken nearly a decade earlier, are well fit by the model. Significant discrepancies are seen only at small differencing intervals Δz. Here, separation statistics deviate from the Poisson-gamma model, perhaps approaching a log-normal form. In addition, negative temperature gradients are sometimes seen in Gregg's observations. These intrusive/overturn events are outside the scope of the present model.

SUMMARY

A series of 9000 CTD profiles from the surface to 560 m has been used to study the statistics of the fine scale strain field in the thermocline. A set of isopycnals, of 1 m mean spacing, is tracked for the 18.75 day duration of the observations. Three statistical studies are performed. Isopycnal separation statistics are formed in both isopycnal following and fixed-depth reference frames. A related investigation simply counts the number of isopycnals found in fixed depth intervals of varying scale. At vertical scales greater than 2-3 m, the statistics of the isopycnal counting study are Poisson. The corresponding isopycnal separation statistics are described by gamma pdfs. The relationship between these three types of information is very nearly as predicted by classical Poisson theory. (e.g., Papoulis 1984).

In this study, the instantaneous separation between isopycnals is simply the sum of the separations of the intervening isopycnals. For example:

$$z_{220} - z_{200} = (z_{210} - z_{200}) + (z_{220} - z_{210})$$
$$= (z_{201} - z_{200}) + (z_{202} - z_{201}) + \ldots + (z_{219} - z_{218}) + (z_{220} - z_{219}) \quad \text{etc.}$$

Here z_{220} indicates the instantaneous depth of the isopycnal whose mean depth is 220 m. It is easily shown that the semi Lagrangian gamma pdf associated with 20 m (say) mean separation can be given as the convolution of the gamma pdf of 10 m separation with itself, or the convolution of the pdf of 5 m separation, repeated 4 times, etc. This key mathematical property of the gamma pdf is only applicable to the present problem if the various "sub-separations" which are added together to form the larger separation are themselves statistically independent. Statistical independence implies that the separation (strain)

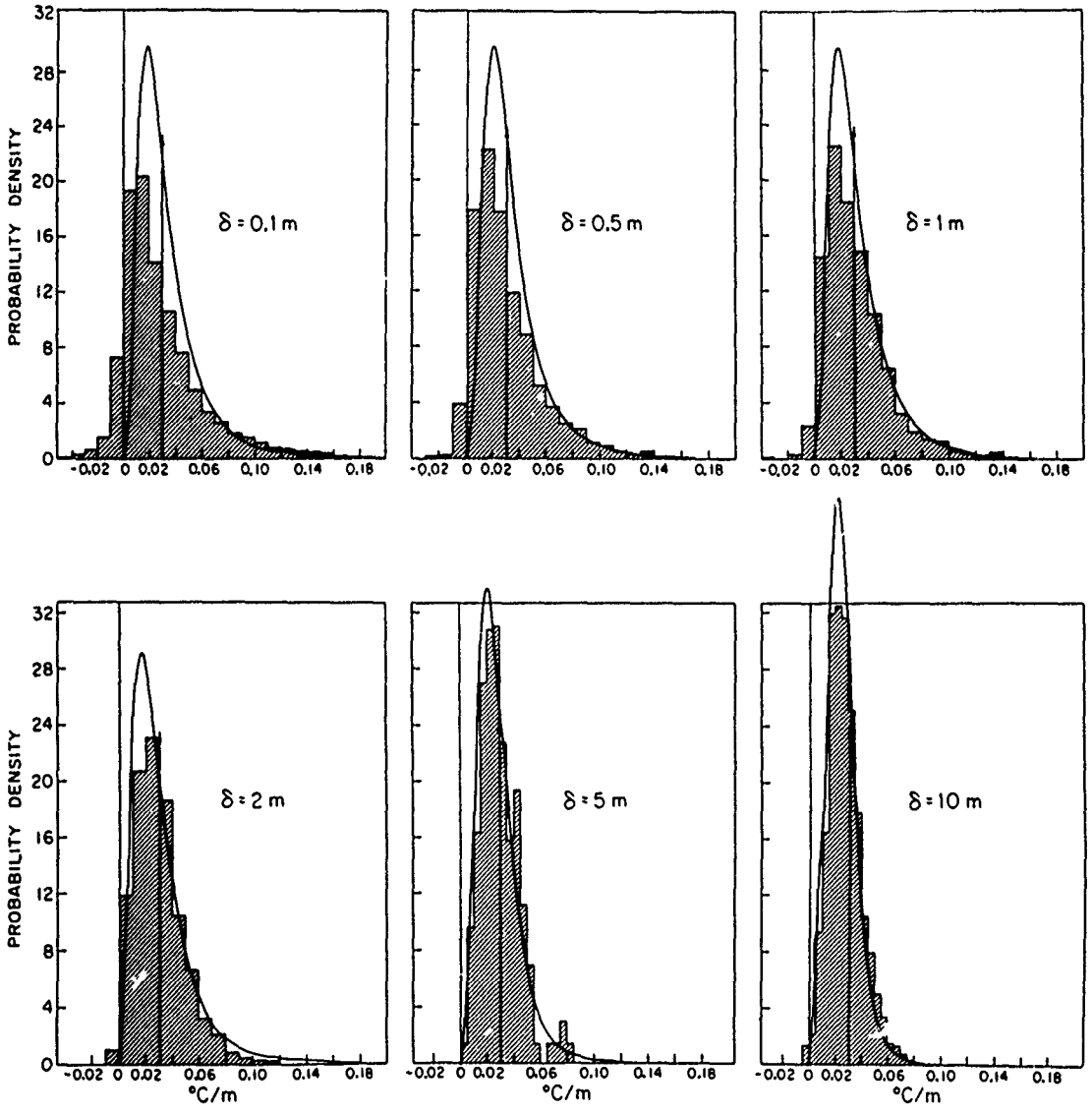


Fig. 9. Histograms of temperature gradient obtained at a variety of fixed separations $\Delta\bar{z}$ by Gregg in the Central Pacific in 1977 from Desaubies and Gregg (1981). The solid curves give the predictions of the Poisson model using the single constant $\lambda = 1.1$ m from the East Pacific 1986 PATCHEX data set.

statistics are uncorrelated. At sufficient small scale where the strain field is correlated we expect the Poisson model to deviate from observation.

A single constant λ (~1.1 m^{-1} here) should describe the mean, variance, skewness, etc. at all vertical scales where the model is applicable. The associated vertical wavenumber spectrum of strain should be white, with spectral level λ^{-1}. In fact, modest variation in λ is seen in the observations. A realistic description of (the second moment of) the strain field requires specification of high and low wavenumber bounds on the spectrum. These bounds are necessary to insure finite variance of the strain field and to acknowledge the finite depth of the ocean. Additional physical principles, beyond the scope of the Pois-

son model, need to be invoked to determine these bounds. An investigation of the non-Poisson aspects of the present data, with a focus on the small-scale/high wavenumber region of the spectrum, will be presented in a subsequent work.

In spite of the weak but apparent discrepancy between the observations and the Poisson model at small vertical scale, the Poisson-gamma pdfs represent a powerful starting point for a description of the fine scale fields in the sea.

REFERENCES

Briscoe, M.G., 1977: Gaussianity of internal waves. *J. Geophys. Res.*, 82, 2117-2126.

Desaubies, Y.J.F., and M.C. Gregg, 1981: Reversible and irreversible fine structure. J. *Geophys. Res.*, 11, 541-556.

Garrett, C.J.R., and W.H. Munk, 1972: Space-time scales of internal waves, *Geophys. Fluid Dyn.*, 3, 225-264.

Gregg, M.C., 1977: A comparison of fine structure spectra from the main thermocline. *J. Phys. Oceanogr.*, 7, 33-40.

Gregg, M.C., 1989: Scaling turbulent dissipation in the thermocline. J. *Geophys. Res.*, 94, 9686-9698.

Holloway, G. , 1983: A conjecture relating oceanic internal waves and small scale processes. *Atmos.-Ocean*, 21, 107-122.

McKean, R.S., 1974: *Probability, Random Variable and Statistic Processes.* McGraw-Hill, 576 pp.

Pinkel, R., and S. Anderson, 1991: Toward a statistical description of fine scale strain in the thermocline. *J. Phys Oceanogr.*, in press.

Pinkel, R., J.T. Sherman, J.A. Smith, and S. Anderson, 1991: Strain Observations of the Vertical Gradient of Isopycnal Vertical Displacement. *J. Phys. Oceanogr.*, 21, 527-540.

Sherman, J.T., 1989: Observations of Fine Scale Shear and Strain. PhD. Thesis, University of California Press.

Tennekes, H., and J.L. Lumley, 19782: A First Course in Turbulence. M.I.T. Press, Cambridge, MA, 300 pp.

SPATIAL STRUCTURE OF THERMOCLINE AND ABYSSAL INTERNAL WAVES

Thomas B. Sanford

Applied Physics Laboratory, College of Ocean and Fishery Sciences, University of Washington, Seattle, WA 98105-6698

ABSTRACT

One hundred velocity profiles over nearly the entire water column are analyzed in terms of the vertical and horizontal structure of the steady and internal-wave velocity contributions. The temporal decomposition into steady, inertial, and high-frequency internal-wave contributions is obtained from several brief time series at one location. The spatial structure of the velocity field is obtained from simultaneous profiles at separations ranging from 15 m to 15 km. WKB normalization of wave amplitude and stretching of vertical wavenumber are performed on the inertial and internal wave components. A reference Brunt-Väisälä frequency (N_0) of 3 cph is used. Spectra are computed on the WKB scaled profiles over more than 90% of the water column and separately over the upper and lower halves. The total internal wave field exhibits equipartition between the east and north velocity components, a decrease in energy density at the lowest wavenumber, and an overall dependence of KE on vertical wavenumber (m) as $m^{-2.5}$. Most of the internal wave energy is contributed by inertial motions, and of this most occurs at (scaled) wavelengths of 100–500 m with a peak at mode 3 (500-m wavelength). The inertial period contributions reveal a power law of m^{-3}, while the high-frequency internal waves ($\omega > 2f$) follow a slope of m^{-2}. There is strong vertical polarization (CW > ACW) of the near-inertial contribution but little or none for the higher-frequency motions. The scaled water column is divided in half for upper-half/lower-half comparisons. There is more scaled inertial KE in the lower half of the water column. The high-frequency internal waves have comparable total energy, but the upper-half spectrum shows a deficiency (or the lower half shows an excess) at vertical wavelengths shorter than 100 stretched meters. The upper half is where the time-mean shear is largest and critical layer processes might be playing a significant role. On the other hand, the lower half may exhibit the influences of bottom scattering of internal waves which tend to "whiten" the spectrum. A few simultaneous, but spatially separated, velocity profiles are differenced, and the KE of the resulting profiles have a zero-correlation scale of about 15–20 km for the total internal wave field. The difference profiles consist primarily of inertial waves. Thus, the deep water inertial waves exhibit wavelengths of 60–80 km (four times the zero-correlation distance) or less in contrast to much longer scales reported for the SML and uppermost thermocline.

1. INTRODUCTION

The least measured aspect of oceanic internal waves is the spatial structure. This is especially true in the abyssal oceans. The paucity of velocity observations there reflects more the difficulty in making appropriate measurements than a lack of interest in or recognition of the importance of the spatial structure. Regardless of the cause, ignorance of the horizontal variability of the velocity field is profound. This is an unfortunate circumstance, because there are

numerous reasons to study the spatial structure of internal wave velocities. The spatial characteristics must be observed to improve our understanding of

- wave dispersion
- effects of advection and Doppler contamination
- WKB scaling and power laws
- wave propagation (momentum, heat, and kinetic energy fluxes)
- wave nonlinearities, such as wave-wave interactions
- wave mixing and instabilities
- interactions of waves and mean currents
- spatial inhomogeneities and patchiness.

Most of the previous spatial studies have been in the upper ocean, and of these most are of temperature. For example, Hayes (1975), Katz (1975), and Stegen *et al.* (1975) present observations of vertical profile series, horizontal tows, and profile sections, respectively. The pioneering observations in IWEX (Müller *et al.*, 1978) and GATE (Käse and Olbers, 1979) provided important thermocline measurements over a spatial array of moored instruments. Considerable measurements of upper ocean velocity and temperature have been obtained with XCPs (Kunze and Sanford, 1984; D'Asaro, 1984; Kunze and Sanford, 1986). The spatial distribution of inertial waves forced by moving hurricanes is discussed by Sanford *et al.* (1987). Acoustic Doppler profiling with enhanced depth capabilities has been conducted by Pinkel with great success in the tropical Pacific thermocline. Basically, these various observations do not reach deeper than about 1000 m, the base of the permanent thermocline.

The goals of this work are to describe the dominant vertical and horizontal scales of internal wave variability as revealed in profiles of horizontal velocity. An attempt is made to separate the contribution of inertial- from higher-frequency (i.e., $\omega > 2f$) internal waves and contrast thermocline and abyssal motions. Special attention is given to the observation and analysis of subthermocline internal waves.

2. THE OBSERVATIONS

The observations were collected in the Sargasso Sea of the western North Atlantic Ocean, in a region with a relatively flat bottom, at least 400 km from appreciable topographic features (except for drops 306–310, which are not included in the spectra). The observations were taken in 1973 and 1975 during the MODE-I (MODE Group, 1978) and FAME (Sanford and Hogg, 1977) programs. The profile locations are presented in Fig. 1, with additional information contained in Table 1. All the data were taken in the summer and fall in generally light winds and low sea states.

A free-fall velocity profiler operating on the principles of geomagnetic induction (Sanford *et al.*, 1978) was used to profile the entire water column. The device is called an EMVP, for Electro-Magnetic Velocity Profiler. In the water depth involved for these data (5200–5500 m), the EMVP required about 1.5 hours for a down or up traverse of the water column. When operating in a time-series mode, occupying one site repeatedly, the EMVP was deployed about once every 6 hours. A set of seven time series of profiles was analyzed. One series was obtained in shallower water (4500 m) than the other six. The remaining group of six time-series profiles, consisting of 50 drops, or 100 profiles (down and up traverses are used separately), are WKB scaled and spectrally analyzed. A second measurement configuration, intended to reveal horizon-

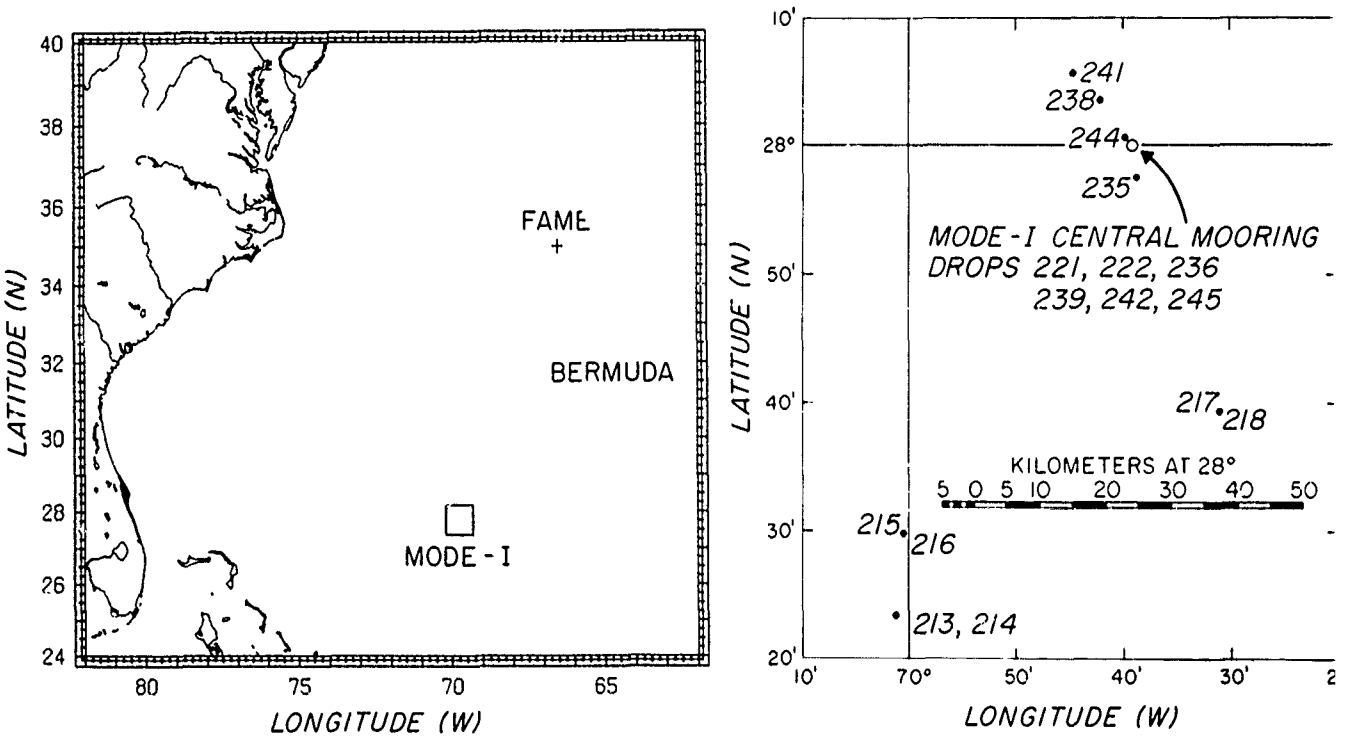


Fig. 1. Locations of profiles used in this study. All data were taken in the western North Atlantic in the Sargasso Sea. The smaller-scale figure shows locations of simultaneous, separated profiles around the MODE-I central mooring.

tal gradients, was deployed in which two EMVPs were preset to fall simultaneously at a known horizontal separation. Over a dozen such horizontally spaced pairs were obtained, but only four were made at horizontal separations of 1 km or greater.

The data quality is good, with errors not exceeding 1 cm/s rms. This level is consistent with the analysis of Sanford *et al.* (1978) and the measured rms differences between closely spaced profiles. The vertical resolution on the original profiles is about 8 dbars, which becomes 10 dbars after interpolation to a uniformly spaced series.

Because of intermittent failures of the internal digital recorder and the application of data quality criteria, gaps exist in some of the profile series. The recorder-produced gaps seldom exceed 10% on the better recorder, but sometimes 50% of the data is missing on the recorder aboard the second EMVP. A quality screening is applied to the data, such that if the noise in the EM data exceeds the equivalent of about 1 cm/s rms, no velocity determination is made. This condition occurred frequently above 100 m depth because of bubble formation in the electrode arms. When the gaps are short and not too frequent, they are filled by interpolation. When the gaps are more extensive, the data are not used or, if part of a spatially separated pair, only measured points are used.

The horizontal spacing between two simultaneously dropped profilers was determined acoustically using synchronized acoustic transmissions from the two profilers. Travel times for

Table 1. Profiles used in analysis.

Group	Type	Location
202-206	Time Series	27° 58′ N 69° 38′ W
213-214	Spaced Drops	27° 23′ N 70° 01′ W
215-216	Spaced Drops	27° 29′ N 70° 00′ W
217-218	Spaced Drops	27° 39′ N 69° 31′ W
219-228	Time Series	28° 00′ N 69° 39′ W
221-222	Spaced Drops	28° 00′ N 69° 39′ W
230-245	Time Series	28° 00′ N 69° 39′ W
235-236	Spaced Drops	28° 00′ N 69° 39′ W
238-239	Spaced Drops	28° 00′ N 69° 39′ W
241-242	Spaced Drops	28° 00′ N 69° 39′ W
244-245	Spaced Drops	28° 00′ N 69° 39′ W
251-260	Time Series	35° 00′ N 66° 30′ W
262-274	Time Series	35° 00′ N 66° 30′ W
306-310	Time Series	32° 58′ N 64° 23′ W
320-324	Time Series	35° 00′ N 66° 30′ W

direct and bottom-reflected acoustic paths were used to determine probe separations. Also, at short separations the distance could be measured visually; at the longest separations (> 4 km) the position of the surface release for the distant EMVP was calculated by interpolation between deployment and recovery positions as determined from LORAN. The displacement of the probe during a profile is negligible compared with that due to the surface current action before and after a profile. For separations of 100 m or less, it was possible to confirm visually that the profilers returned to the surface with approximately the same horizontal separation as at launch.

3. DATA ANALYSIS

Most of the kinetic energy of these velocity profiles resides in the low-frequency, quasi-geostrophic motions, with internal waves contributing only about 20% of the total horizontal kinetic energy. Two methods are used to separate the internal waves from the background. In the first method, a 1-day or longer series of profiles at a location is fitted to time-mean and inertial-period components by a rotary, least-square decomposition method described by Sanford (1975). The time-mean profile is subtracted from the individual profiles; the residue profiles are principally composed of internal waves. When both the time-mean and inertial profiles are subtracted from the original profiles, the residue profiles are of higher-frequency internal waves.

The second scheme involves the subtraction of two simultaneous but horizontally separated profiles. The residue profile is taken to be composed of internal waves, since the difference in low-frequency structure over 10 km is assumed to be negligible. This assumption is based on the

measured correlation length of about 100 km for the MODE eddy (MODE Group, 1978). Both methods of internal wave discrimination are barely adequate, yet both permit useful analysis of internal waves in the presence of strong planetary motion.

The third method for removing the time-mean or planetary motion is to fit individual profiles to vertical eigenmodes for linear, flat-bottom, and zero-mean-flow Rossby waves. It was thought that the modal fit would filter out the internal waves. The eigenmodes are solutions to the equation (Pedlosky, 1979)

$$\left[\frac{f^2}{N^2}W_z\right]_z + \lambda W = 0, \tag{1}$$

where (W) is the eigenfunction, the subscript denotes differentiation with respect to depth z, and N and f are the Brunt-Väisälä and Coriolis frequencies. The separation constant (λ) is $-[(-\beta k/\omega) + k^2 + l^2]$, where β is the horizontal (meridional) gradient of f, ω is the wave frequency, and k and l are the east and north horizontal wavenumbers. The boundary conditions at the sea surface and bottom are

$$W = 0 \text{ at } z = 0 \text{ and } -H\ . \tag{2}$$

The hypothesis is that an energetic mode will remain constant from fit to fit, being relatively unaffected by the internal waves. However, it is found that the decompositions show that the modal estimates, even for the most energetic first mode, fluctuate greatly over the brief series. The cause of this behavior is large, low-mode inertial period components that contaminate the eigenmode computation. Eigenmode fitting is very useful, however, when applied to the time-mean profiles derived from the rotary, least-square decomposition.

The time series profiles are decomposed into time-mean, inertial, and high-frequency internal wave components. Since the subinertial motion is so energetic, it is better to remove most of it from the series at the first step. The difference between the original profiles and the arithmetic mean profile is called the internal wave field. The next stage is to subject the total internal wave field to the rotary, least-square demodulation at a frequency equal to $1.05\,f$, a dominant frequency often observed in moored current meter records (Webster, 1968). From this procedure are obtained two velocity profiles: the low-frequency (which is small since the arithmetic mean has been removed) and the inertial. The small, low-frequency component is added to the arithmetic mean to yield the best estimate of the subinertial (time-mean) profile. The inertial profile consists essentially of speed and direction values referenced to some arbitrary, prescribed time t_o. To compute the inertial profile at a given time t, it is necessary to add to the computed direction an amount given by $1.05f\ (t - t_o)$. Thus the higher-frequency internal wave field is obtained by subtracting the time-mean and the appropriately time-advanced inertial profiles from each of the original profiles.

The independence of the inertial and higher frequency profiles (i.e., lack of correlation) is demonstrated in Appendix A.

Both down and up profiles have been used. No vertical interpolation has been performed; if more than 2/3 of the data at a given 10 dbar level is missing, no fit is performed at that level. That is, if 10 profiles were used (say 202–206, down and up), then 7 or more of the 10 velocity values must be available at that depth level for an acceptable fit to be made.

The fitting for steady and inertial components for the series 202–206 is shown in Fig. 2. This series provides a clear example of the compact description often possible for subinertial profiles. About 95% of the baroclinic energy is contained in the steady profile. The relative smoothness of the steady profile demonstrates the dominance of the steady component and the ability of a few vertical profiles to obtain an adequate determination of the mean flow. Sanford (1975) showed how a pair of velocity profiles taken half an inertial period apart nicely reveals the steady and inertial contributions.

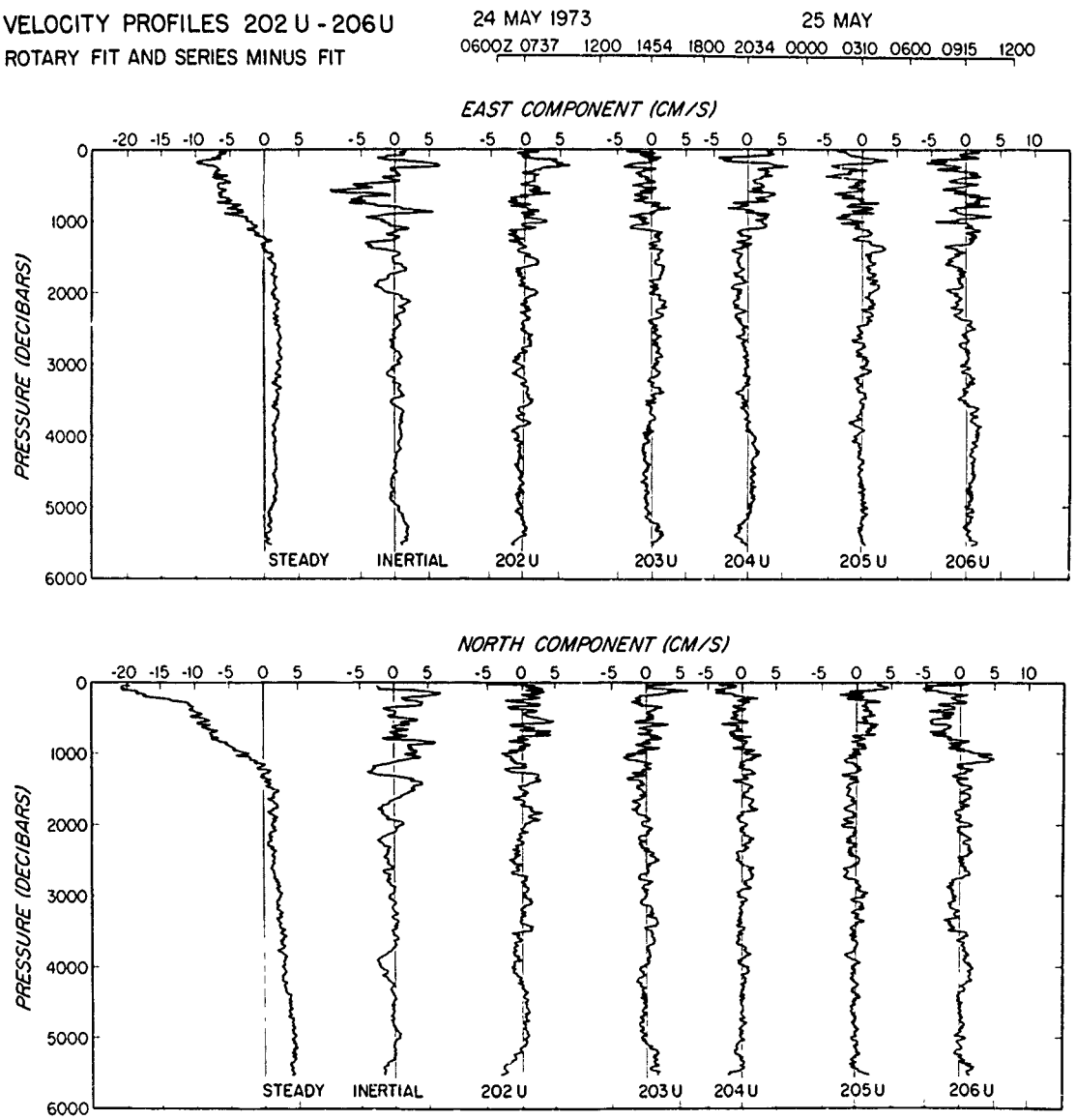


Fig. 2. Estimates at each 10-dbar level of the steady, inertial, and residual profiles (result of removing steady and inertial components from the original profiles). The inertial component was computed at a period of 25.6 hours (1.05 f) and is shown for the reference time of 0000 24 May 1973. The drop time of the individual profiles in this series is shown in the upper scale.

The modal composition of the 202–206 steady profile is shown in Fig. 3. Only the first mode is shown in actual magnitude; the higher modes are shown at a magnification of 5. The first mode is the dominant component, with the third adding some near-surface flow.

The time-mean profiles are presented in Fig. 4, and the inertial-period profiles are shown in Fig. 5 for the seven-profile series. The data sets are denoted as 202–206, 219–228, 230–245, 251–260, 262–274, 306–310, and 320–324, where the numbers are the first and last drop numbers used in each series. Note that drops 306–310 are in shallower water, which is still very deep but significantly less than at the other sites.

Two characteristics of these data should be mentioned. First, the time-mean profiles have considerable shear in the upper 1000 m; 320–324 also has higher mode, deeper shears. In general, the vertical shear is much less below 1000–1500 m than in the upper ocean. Second, the inertial fits are not totally independent of each other, and the energy and vertical scales are clearly larger in the FAME data set compared with MODE-I profiles. The two experiments were conducted at different latitudes (28°N vs 35°N), at different seasons (summer vs fall), and under different weather conditions. Profiles obtained in the POLYMODE experiment around 30°N, just north of the MODE-1 site, exhibit low levels of inertial motions. The latitudinal dependence of inertial motions has been investigated by Fu (1981), who found that the inertial peak grows and sharpens with higher latitude. With a small sample, it is not possible to determine if there are seasonal, weather, or latitudinal effects, as opposed to the expected variability due to sampling.

The modal fits for six time series (i.e., Fig. 4 without 262–274) are shown in Fig. 6. The first mode, which is plotted at one-fifth amplitude, is clearly dominant. Only the right-most profiles exhibit more robust second or third modes. The large second mode in 306–310 is probably the result of a large, thermocline eddy at this site, seen more clearly in Fig. 4.

A WKB scaling has been applied to both the temporal and the spatial series. After the subinertial profile is removed, the resulting internal-wave profiles are scaled in amplitude and the independent variable—pressure or depth. Removal of the subinertial flow is achieved by either the rotary decomposition of a time series or the subtraction of spatially separated drops. The WKB procedure (Leaman and Sanford, 1975) uses the local Brunt-Väisälä frequency, $N(z)$, to normalize wave amplitude by dividing by $[N(z)/N_o]^{1/2}$ and to stretch (shrink) the vertical scale in proportion to $N(z)/N_o$. A value of 3 cph is taken for N_o throughout this paper, although different N profiles were used for each data set since the observations were taken at different locations and seasons. An example of the Brunt-Väisälä profile and the actual stretched coordinate transformation diagram for the MODE data are shown in Figs. 7a and 7b. The WKB stretching and normalization are discussed further in Appendix B.

The inertial, high-frequency, and raw profiles minus the time-mean profiles (i.e., the total internal wave band) are WKB stretched and normalized, denoted by s and n, according to the procedure of Leaman and Sanford (1975). The resulting profiles are about 1800 sdbars long (Fig. 7) with velocity values every 10 sdbars [pressure in decibars (dbars) and depth in meters (m) or stretched meters (sm) are used interchangeably in this paper]. Short gaps in the original data are closed by linear interpolation. For spectral computations, 160 data points are used, which excludes about 200 m at the top and bottom of the water column.

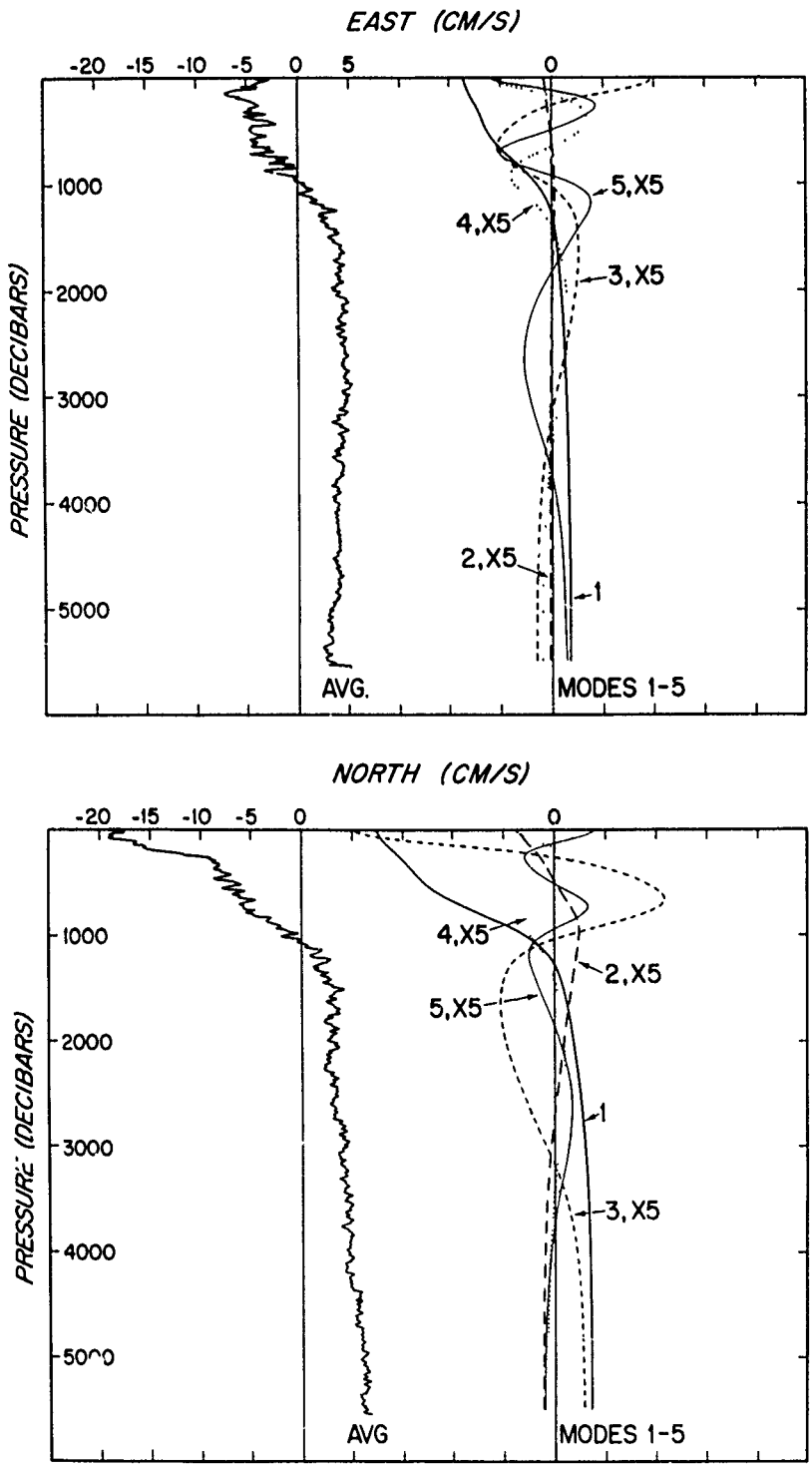


Fig. 3. The first five, flat-bottom, normal-mode amplitudes for the steady velocity profile of Fig. 2. Mode 1 is dominant in both velocity components, while mode 3 is needed to account for the shear above 300 dbars. Modes above the first are plotted at five times their values.

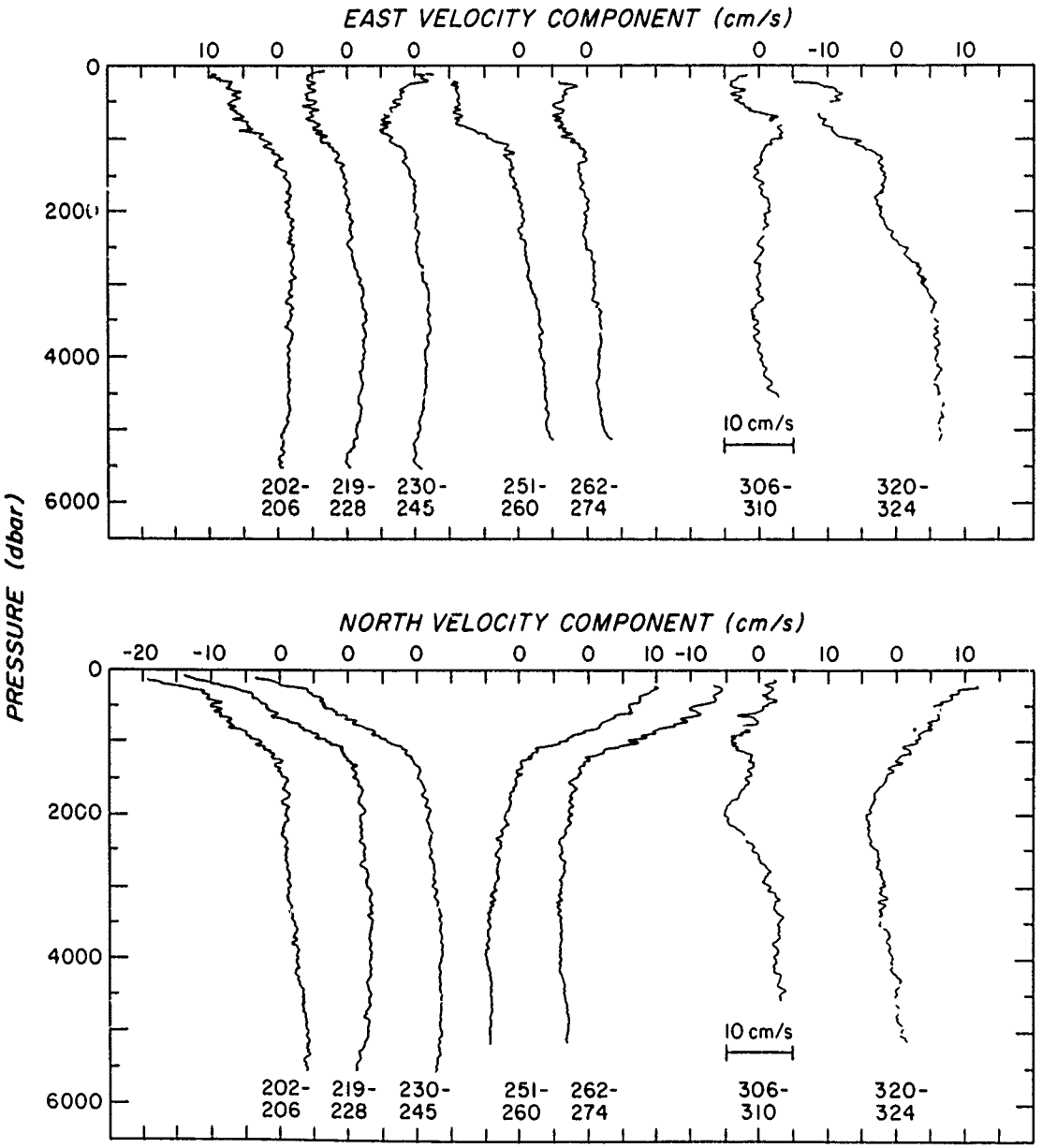


Fig. 4. Steady velocity profiles for seven time series. The numbers under each profile correspond to the first and last drop used in each series.

4. THE VERTICAL STRUCTURE OF INTERNAL WAVES

The internal wave profiles were subjected to Fourier analysis. Although velocity values from nearly the whole water column are used, the vertical wavelengths do not correspond to modes. First, neglect of the upper 200 m biases the computations; second, only the even modes correspond to spectral estimates (the odd modes are not individually resolved). This situation occurs because the modal eigenfunctions are similar to $\cos n\pi Z/H$ whereas the Fourier represen-

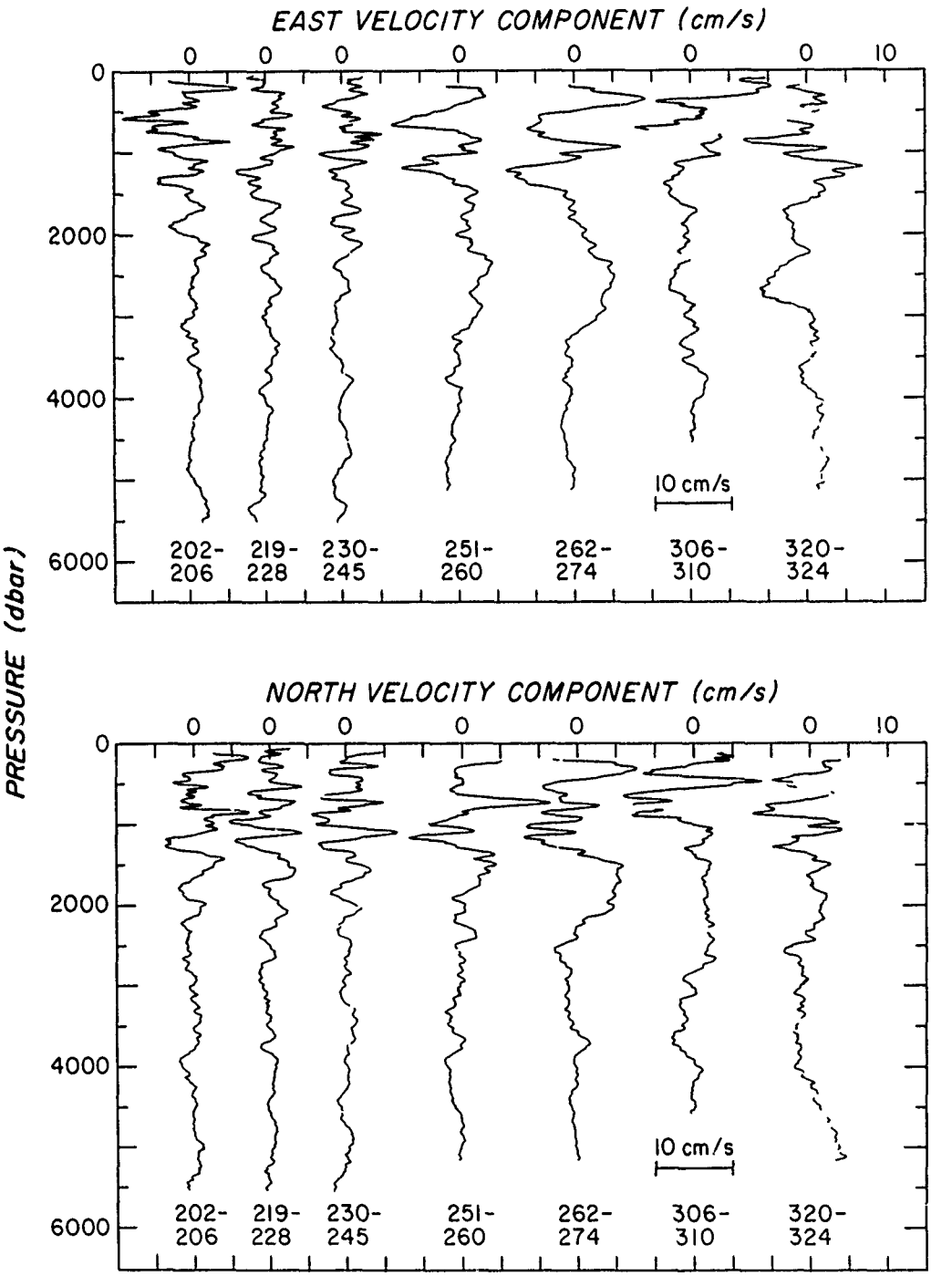


Fig. 5. Inertial profiles for the same seven time series used for Fig. 4.

tation uses $\sin 2n\pi Z/H$ and $\cos 2n\pi Z/H$. No band averaging is done on the spectra; ensemble averaging is performed over the six inertial determinations and 100 higher-frequency internal wave profiles. It is clear from the visual similarity of adjacent inertial fits that these data are not independent realizations. Because there are three separate groups of profiles, it is assumed that

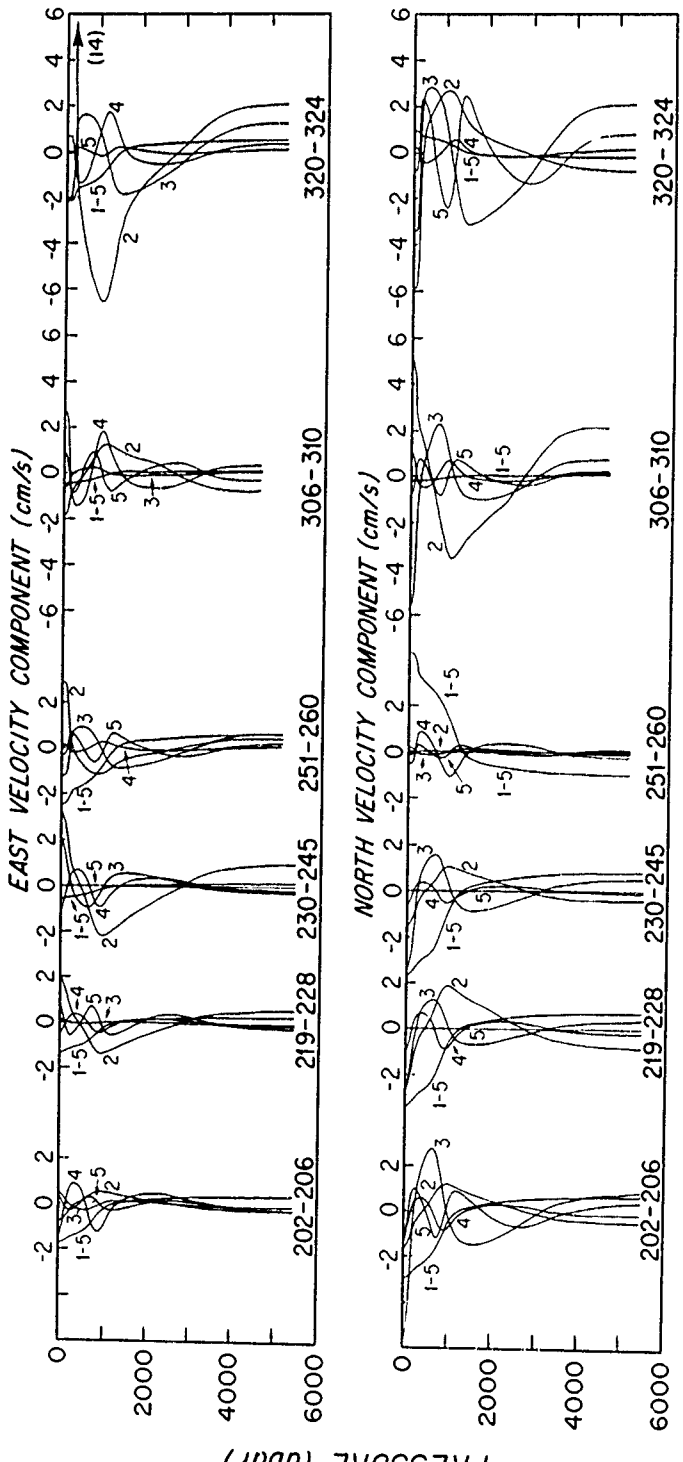


Fig. 6. Modal fits for six profile time series. Fits are made separately for the east and north velocity components. The contributions to the "steady" profile are shown in correct amplitudes for modes 2–5. Mode 1, the dominant one, is shown at 20% of its computed amplitude.

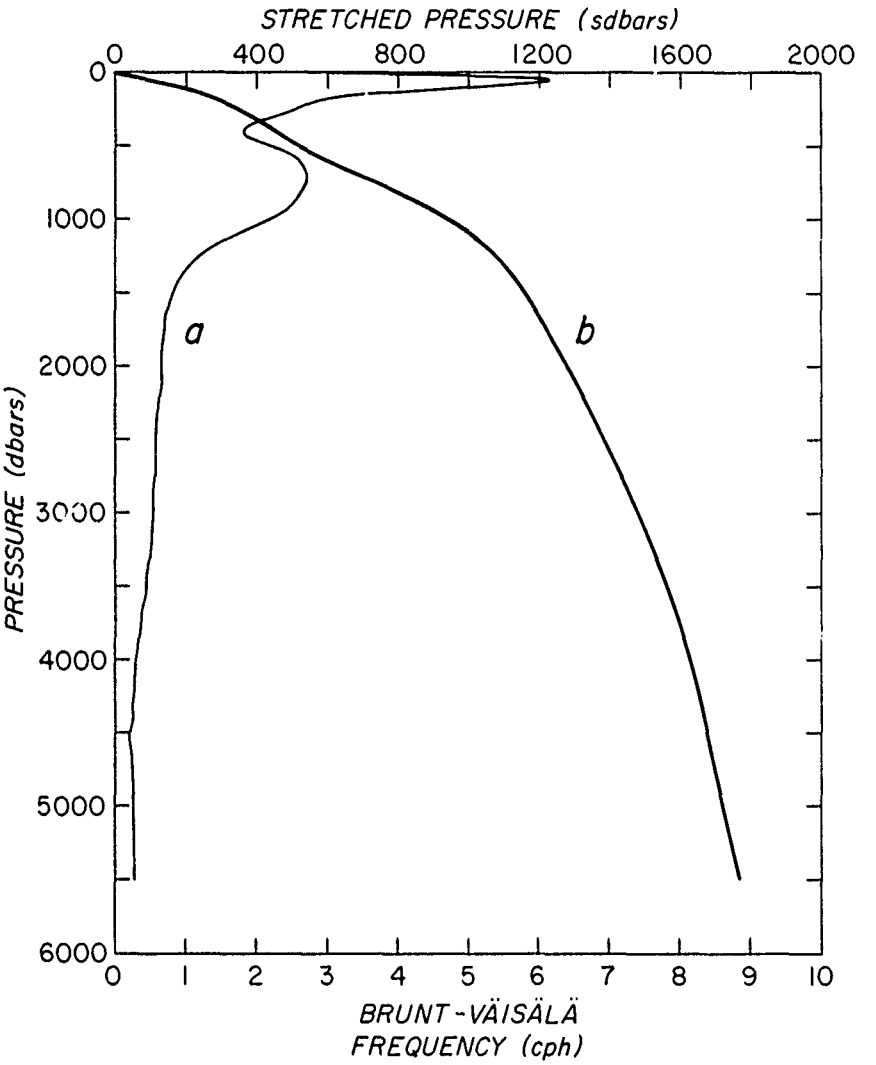


Fig. 7. Brunt-Väisälä frequency profile (line a) at the MODE-I mooring site and relationship of stretched pressure to measured pressure (line b).

these are independent and that the true confidence intervals are at least not larger than those based on two to three realizations. For the high-frequency profiles, down and up profiles are not independent in the deep water, but adjacent profiles taken every 4–7 hours are independent, except insofar as these waves are influenced by common environmental factors such as mean vertical shear or weather. For purposes of statistical inference, it is assumed that the set consists of 50 independent profiles, allowing for some common features on down and subsequent up profiles in deep water.

There are no significant differences between the energy in the east and north components. The separate contributions to the kinetic energy spectrum are shown in Fig. 8 for the original data minus the time-mean profiles. Equipartition between the velocity components is expected (Garrett and Munk, 1972) and observed (IWEX; Müller *et al.*, 1978). These results do not rule out nonequipartition over short portions of the water column or near sources and topography (Johnson and Sanford, 1980; Gordon, 1978).

Only at the lowest wavenumbers is the spectrum consistent with the top-hat dependence of the first GM model (Garrett and Munk, 1972). The spectral slope tends to be flat over estimates 1, 2, and 3 (1600, 800, and 533 sm) followed by a slope of –2 until at a wavelength of about 50 sm and smaller; the high wavenumber slope is about –2.5. The corresponding shear spectrum has a slope between zero and –0.5. The slope is probably not significantly nonwhite and basically agrees with the proposed shear spectrum of Gargett *et al.* (1981).

There is a clear difference in the spectral slope between the inertial and higher-frequency waves (Fig. 9). At vertical wavelengths of about 100 m and longer, the inertial motions are

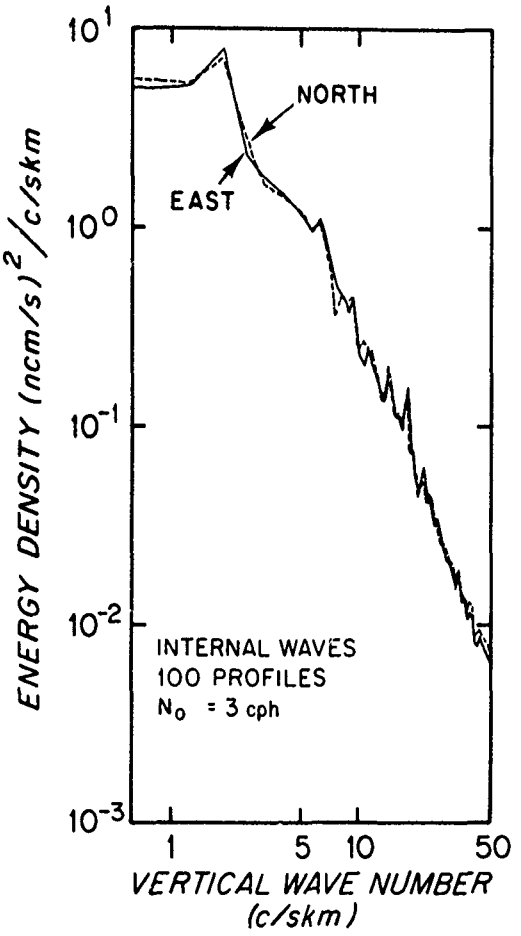


Fig. 8. Energy density spectra in stretched and normalized form for all 100 internal wave profiles for east and north velocity components. The spectrum of kinetic energy would be one-half the sum of the east and north contributions.

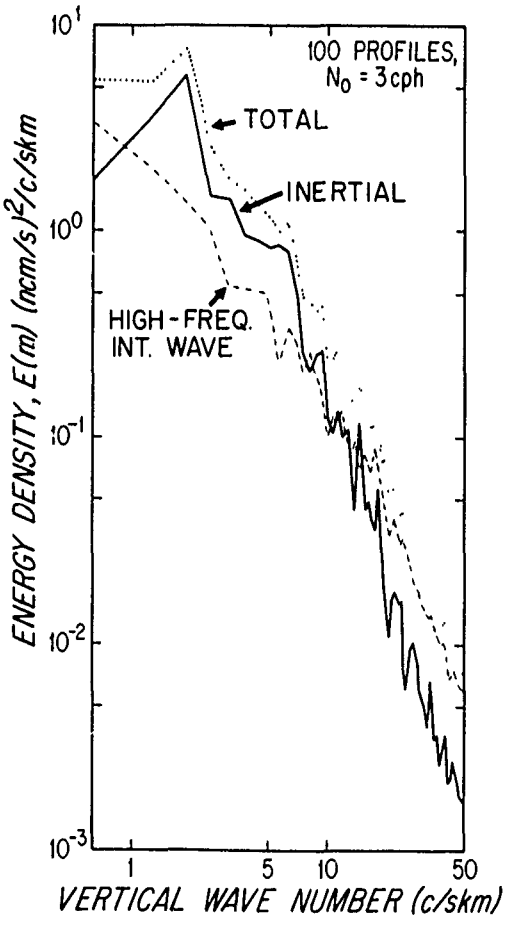


Fig. 9. Kinetic energy density spectra for total internal wave field, the inertial wave contribution, and the high-frequency internal waves. The ordinate is kinetic energy and is one-half the sum of the east and north velocity components.

dominant, except for the longest vertical scale, at which they are (or may be) inferior. At all wavelengths shorter than about 100 m, the higher-frequency waves are more energetic. Over the high-wavenumber region (say 5 c/skm and larger), the spectra slopes are about –3 for the inertial and –2 for the HF waves. In both cases the slope depends on the interval over which the fit is made. For example, around 20 c/skm (50 m), there is a slight bump (energy excess) to both spectra which, if heavily weighted in the fit, would force higher slopes to be obtained. Excluding this bump, at the highest wavenumbers (> 20 c/skm) the spectral slopes approach –2.5 for the inertial component and –2.0 for the high-frequency waves.

Other spectral slope determinations have largely resulted from temperature observations. Hayes (1975) reports a slope of –2.5 based on temperature profiles from a CTD; Cairns and Williams (1976) derive a slope of –2 from a drifting, vertically moving temperature sensor, as does Gregg (1977) from his free-fall temperature profiler. The temperature field is relatively unenergetic near f because the motions have little vertical velocity but maximum horizontal kinetic energy. Thus temperature spectra are lacking the lower frequency (and lower wavenumber), energetic inertial signatures. Also, it may be that vertical detrending of temperature profiles (to reduce the contribution from the time-mean temperature field) removes much of the longer vertical wavelength components.

The conventional and rotary spectra (Leaman and Sanford, 1975) of these data are presented in Fig. 10 in the energy-preserving form. The low wavenumber energy cutoff is clearly revealed in panel A beyond a wavelength of about 500 sm. The peak at 2 c/skm is very pronounced, showing that 500 m is a dominant vertical wavelength. There is a secondary peak around 5–6 c/skm, near the third harmonic of the main peak. The inertial energy is principally contained between wavelengths of 150 and 500 sm. At both the lowest and highest wavenumbers, the high-frequency internal waves are more energetic than the inertial motions.

A similar result was reported by Müller *et al.* (1978) who found during the IWEX experiment that there was little energy in the first three inertial modes (estimates 1 and 2 here), and that most of the energy at these vertical wavelengths was contributed by higher-frequency internal waves. Perkins (1970) found that the third vertical mode was the dominant one in the Mediterranean. The global wave functions of Munk and Phillips (1968) and Fu (1981) have turning latitudes which decrease with vertical mode number. Near-inertial waves will be reflected (attenuated) in a manner inversely related to vertical mode number as they propagate northward. Thus, the observed near-inertial wave field would be deficient of the lowest modes unless these modes were locally forced.

Panels B and C of Fig. 10 show the rotary decompositions of the inertial and high-frequency components. As has been reported by many investigators (e.g., Leaman, 1976; Kundu, 1976), the clockwise rotating (with respect to depth, looking downward) contributions are more energetic. The ascendency of the CW over the ACW is clear in these data, but was not observed in IWEX data (Müller *et al.*, 1978) from a mooring in the MODE-1 area based on calculations of the vertical component of velocity from temperature measurements. It is likely that the IWEX observations contain excessive noise from temperature finestructure.

These are the first abyssal vertical spectra of frequency-separated contributions; previous profile results were based on the whole frequency spectrum (Leaman, 1976), used inertial decomposition over only a short vertical interval and applied no WKB adjustments (Rossby and

Sanford, 1976), or were of thermocline motions (Pinkel, 1981, 1985).

The peak near 2 c/skm in the inertial period energy stands out in the CW component and is missing in the ACW. The CW contribution exceeds the ACW at all vertical wavelengths down to about 100 m. The ACW part is more regular in its wavenumber behavior, while the CW may have some harmonic structure as previously mentioned. It is evident that the ACW inertial component is similar in energy and spectral form to that of the higher-frequency internal waves.

Clearly, CW polarized waves dominate over ACW ones, and this demonstrates that most of the energy is associated with a downward group velocity (CW). The direction of energy flux, on the other hand, is not easily determined from these data. There is insufficient frequency (or, equivalently, spatial) information to calculate the group velocities associated with the upward- and downward-going waves. Near the inertial frequency, the vertical group velocity is

$$Cg = -\frac{(\omega^2 - f^2)}{f\,m}. \tag{3}$$

The convention is that $m > 0$ corresponds to CW polarized waves and $m < 0$ to ACW waves. The inertial decomposition is not selective enough in frequency to allow ω to be determined separately for the CW and ACW contributions. In fact, it would take a 100-day-long time series

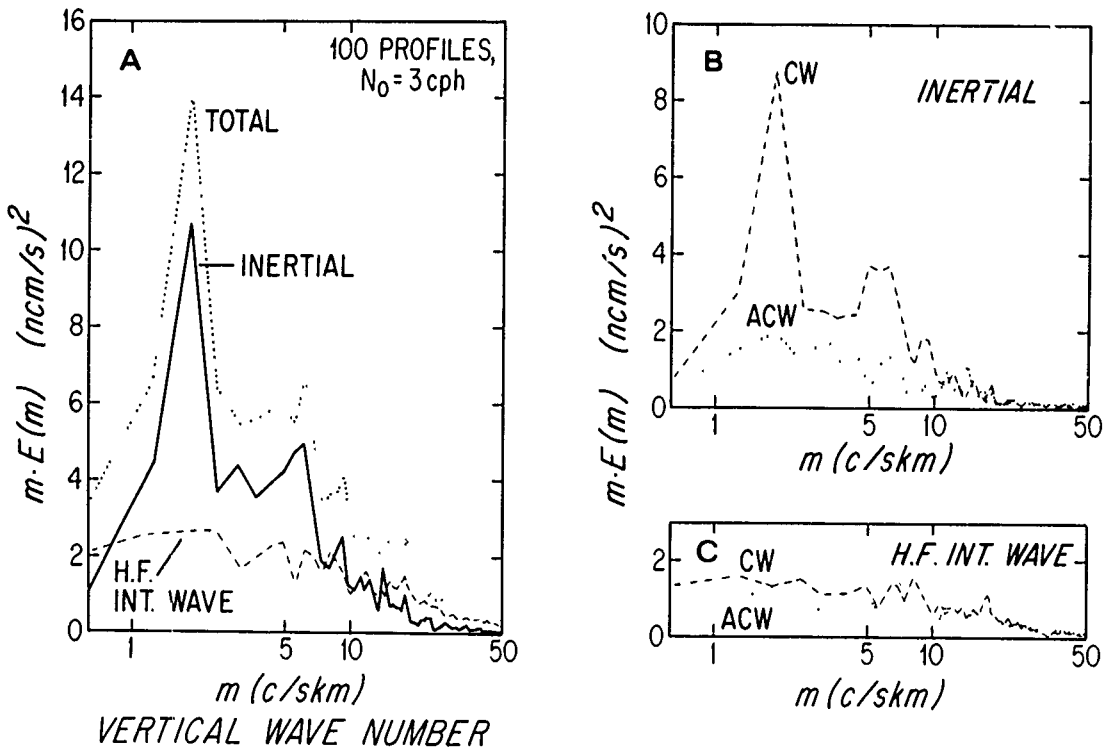


Fig. 10. A: Spectra of Fig. 9 expressed in energy preserving form. B: Inertial component separated into CW and ACW depth-polarized components. C. High-frequency internal waves expressed as CW and ACW spectra.

to get the resolution to about 1% of f. Spatial information, however, is readily used to resolve small differences in Cg, which can be written for $\omega^2 << N^2$ as

$$Cg \doteq -\frac{N^2 k^2}{f\, m^3}, \tag{4}$$

where N is the Brunt-Väisälä frequency and k is the modulus of the horizontal wavenumber associated with the energy at the vertical wavenumber m.

In panel C of Fig. 10, the CW and ACW spectra are shown for the high-frequency internal waves. The energies are about equal, demonstrating the vertical symmetry of these waves. Models generally hypothesize vertical symmetry (e.g., Garrett and Munk, 1975), and wave interactions predict equipartition (McComas and Bretherton, 1977).

A check of the WKB scaling results from comparing the upper half of the stretched water column (≈200–1000 dbars of real ocean) with the lower half (1000–5000 dbars). Such a comparison is shown in Fig. 11. Although the energy levels are not greatly different, demonstrating the usefulness of energy scaling, the vertical wavenumber spectra are systematically and increasingly different at the high-wavenumber end. Several explanations for this discrepancy occur. One is that the WKB scaling amplifies noise in the lower half relative to the upper half. To a first approximation, the measurement noise is essentially the same throughout the water column, so one might expect that the WKB scaling should boost the lower-half noise relative to that in the upper ocean. In Appendix B it is shown that this effect is compensated for by the adaptive filtering, the use of more data in the deep water. Thus the discrepancy probably results from more complex, non-WKB effects or structural changes in the ocean.

The upper-half versus lower-half spectra are presented in Figs. 12 and 13 for the inertial and high-frequency waves. The behavior of the two classes of motion is different; the deeper inertial motions seem more energetic by a factor of about 2 at all vertical wavenumbers, whereas the upper HF waves are more energetic until beyond 10 c/skm, where the spectrum drops rapidly. A high wavenumber cutoff in this range has been reported by Pinkel (1984, 1985) in his upper ocean profiles. The cutoff is not seen in the lower half of the ocean. Another possibility is that the lower half exhibits the results of energy flux from low wavenumbers to high wavenumbers as might be expected from the "whitening" effect of bottom scattering (Eriksen, 1985; Rubenstein, 1988).

The nearly constant ratio between upper- and lower-half inertial-energy densities suggests a wavenumber-independent factor. It is necessary to work with more complete WKB expressions (Phillips, 1977) for horizontal kinetic energy E and vertical wavenumbers m:

$$E \propto \frac{(1 + f^2/\omega^2)\,(N^2 - f^2)}{(\omega^2 - f^2)^{3/2}\,(N^2 - \omega^2)^{1/2}}, \tag{5}$$

and

$$m \propto \frac{(N^2 - \omega^2)^{1/2}\, k}{(\omega^2 - f^2)^{1/2}}. \tag{6}$$

For $\omega \sim f$ it is clear that E normalized by N and divided by the vertical wavenumber resolution (to yield spectral energy density) is

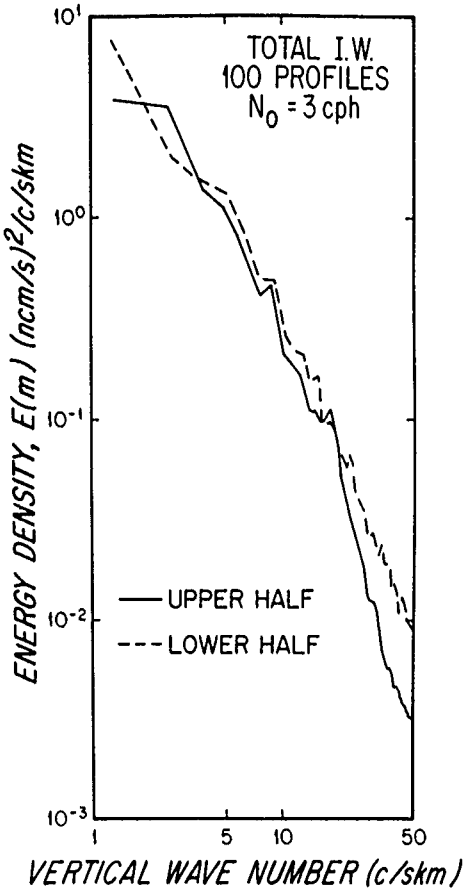


Fig. 11. Kinetic energy density spectra of total internal wave field separated into upper and lower halves of the WKB-stretched water column.

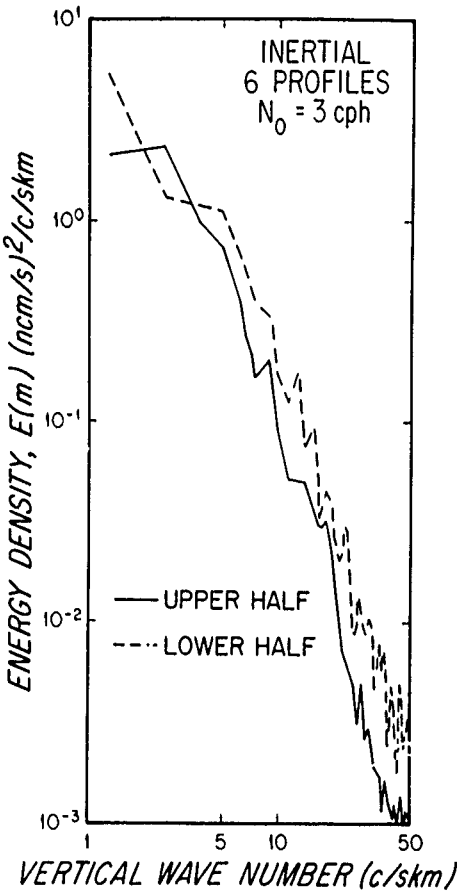


Fig. 12. Kinetic energy density spectra of inertial wave field separated into upper and lower halves of the WKB-stretched water column.

$$\frac{E^*}{\delta m^*} = \frac{E/N}{\delta m/N} = \text{normalized energy density} \propto \frac{1}{(\omega^2 - f^2)}\,, \tag{7}$$

with

$$m^* = m/N \propto \frac{1}{(\omega^2 - f^2)^{1/2}}\,. \tag{8}$$

Possibly the discrepancy in Fig. 12 results from incorrect N profiles or real departures from WKB scaling. As discussed in Appendix B, the normalized and stretched energy density is independent of N such that it does not change in value if an incorrect N profile is used. The wavenumber scale does change in this case, resulting in a lateral shift of the spectrum. It is not possible for N to be wrong by a factor of 1.5–2, but it does seem likely that the kinematic changes predicted by the WKB expressions are not in equilibrium. That is, the downward-propagating inertial waves, the existence of which is demonstrated by the rotary spectra, must

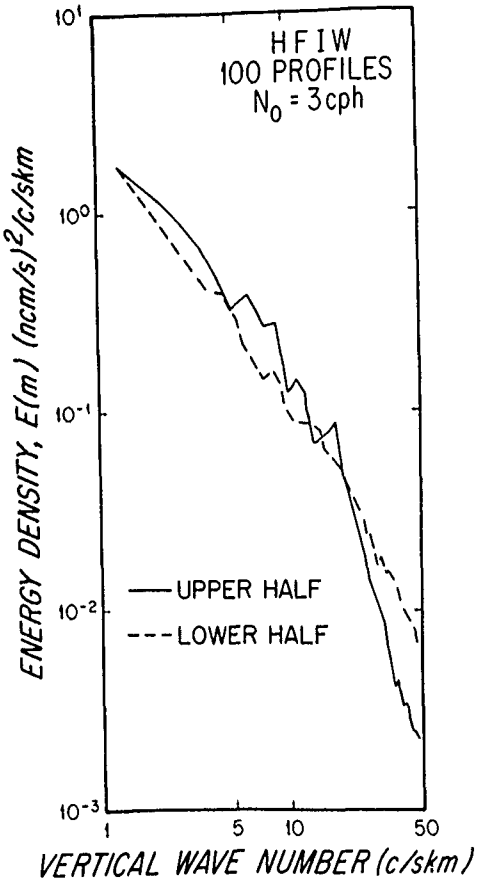


Fig. 13. Kinetic energy density spectra of high-frequency internal waves separated into upper and lower halves of the WKB-stretched water column. Data gaps result when one or the other profiler experienced noise or loss of data because of tape recorder malfunctions. Table 2 presents the WKB-normalized KE for each spaced pair.

continually undergo structural changes as they go deeper into the lower N region. Thus, there might be a tendency for the wave structure to reflect that of a wave in equilibrium with a slightly higher-than-local value of N. Hence, in the deep water, it is reasonable to expect energy that should be at a certain WKB-assigned wavenumber to appear at a *higher* one, since N is essentially monotonically decreasing below the thermocline.

Another possibility is that, if $\omega^2 - f^2$ were smaller in the lower half compared with above, the observed spectral shift in Fig. 12 would occur. That is, as a wave propagates into the deep water, $\omega^2 - f^2$ decreases (by processes to be discussed later), and the WKB-scaled energy (using only N) increases and appears at a larger stretched wavenumber. A spectral value will move along the line of slope +2 in the log $E^*/\delta m^*$ vs log m^* plane as ω changes near f. This will occur at all wavenumbers, resulting in a constant offset on the log/log plots of Fig. 12.

A possible mechanism for an $\omega \rightarrow f$ shift in the deep water is the kinematic effects of mean vertical (and, in general, horizontal) shears. In each of the data sets, there was a significant time-mean vertical shear of about 20–40 cm/s over the upper 1000 m. The deep-water shears were much weaker, although if scaled according to Rossby-wave modes (a similar scaling to that used here for internal waves) the shears are comparable to those higher in the water column. To investigate the behavior of near-inertial waves in the presence of vertical shear, it is appropriate to replace ω by the intrinsic frequency, $\omega - \mathbf{k} \cdot \mathbf{U}(z)$, where $\mathbf{k}$ is the horizontal wavenumber and $\mathbf{U}$ is

the time-mean (of frequency << ω) horizontal velocity. As near-inertial waves propagate downward from the sea surface, some will experience a frequency shift such that $\omega - \mathbf{k} \cdot \mathbf{U}$ approaches f, and the energy density as represented in Eq. (7) greatly increases. Other waves experience the opposite effect and blend into the higher-frequency components. The vertical wavenumber is correspondingly increased, which means that the deep water spectrum is translated along a line of slope = 2, resulting in high energies at every wavenumber. Once below the main shear zone, the refraction ceases, but the waves retain the enhanced energy.

Continuing with this line of speculation, the high-frequency waves also experience shear refraction. The situation here is that, as $N^2 - \omega^2$ becomes small, modifications occur such that horizontal KE density/N becomes larger, while m/N becomes smaller. Explicitly, as $\omega \rightarrow N$ in a region of uniform N, the expressions are

$$\frac{E^*}{\delta m^*} \propto (N^2 - \omega^2)^{-1} \tag{9}$$

and

$$N^2 - \omega^2 \propto N^4 k^{-2} m^{*2} . \tag{10}$$

Hence,

$$\frac{E^*}{\delta m^*} \propto N^{-4} k^2 m^{*-2} . \tag{11}$$

Thus a point on the spectrum will move along a line of slope equal to –2. For small changes, the spectral distortions would be hard to see, since the data have a slope of about –2.

A second consequence of shear refraction is the possibility of critical layer processes. Internal waves of a given frequency have smaller horizontal phase velocities as the vertical wavenumber is increased. Away from N and f, the horizontal phase velocity can be approximated as

$$c = N_o/m^* . \tag{12}$$

Critical layer absorption can occur as the waves propagate through an ambient velocity shear. Let δU be the velocity difference between the level of wave generation and the point at which $\delta U = c$ where critical layer processes occur. For a given δU over the domain of interest, waves of vertical wavenumber greater than m_c^* (where $m_c^* = N_o/\delta U$) will experience refraction and critical layer processes. The energy discrepancies seen in Fig. 13 may result from shear refraction and critical layer absorption. A two-part internal wave field is hypothesized by Munk (1981). At low vertical wavenumbers ($m^* < m_c^*$), the energetic, *intrinsic* waves occur, while at smaller scales ($m^* > m_c^*$), there are *compliant* waves which are greatly modified by interactions with the intrinsic field.

The change in slope in Fig. 13 occurs at a wavelength of 50 sm. The corresponding velocity difference, δU, is 5 cm/s, a value not frequently observed over 50 sm in the time-mean flow but one common to the inertial motions. The shear over 50 sm is taken to be caused by an inertial wave with a vertical wavelength of 200–300 sm.

5. THE HORIZONTAL STRUCTURE OF INTERNAL WAVES

Certainly the most unexplored aspect of internal waves is their horizontal structure. To a limited extent, IWEX (Briscoe, 1975; Müller *et al.*, 1978) described the horizontal wavenumber dependence of internal waves. Much more extensive observations have been reported by Katz (1975) using towed sensors, Stegen *et al.* (1975) from XBT patterns, and Pinkel (1975) using a spaced array of temperature profilers and, later, horizontally directed Doppler sonar. There are a few, but widely divergent, estimates of the horizontal wavelengths of near-inertial waves based on moored current meters (Webster, 1968; Fu, 1981; Pollard, 1980). More recent observations with XCPs by D'Asaro (1984) have revealed spatial coherence scales of internal waves in the upper ocean. The data set presented here is a small but unique one demonstrating a clear pattern to the horizontal dependence of internal waves.

The original intention was to compare these observations with the expectations of Garrett and Munk (1975) expressed in their dropped horizontal coherence (DHC). The pursuit of DHC calculations was suspended because the data set was too small to yield reliable estimates. Moreover, interpolation was necessary across the data gaps before WKB vertical stretching was applied. Thus it was decided that the vertical wavenumber and horizontal dependence of DHC could not be adequately tested with the present data.

Rather than the DHC calculations, the mean square differences between the horizontally spaced profiles are computed. This quantity is identical to a structure function of zero depth lag. The KE estimates are based only over depth intervals in which *both* profiles are good; there are gaps in the difference profiles wherever either profile was missing data. Measurements from the surface to 5550 dbars, every 10 dbars, are used for 556 possible velocity differences. The velocity differences are normalized by $[N_o/N(z)]^{1/2}$, the WKB amplitude scaling.

A set of profile differences, before WKB scaling, is shown in Fig. 14. There is a clear progression toward larger ΔX. Also, the dominant vertical scale increases with ΔX, consistent with the dispersion relationship between vertical and horizontal wavelengths for inertial-period internal waves.

An attempt has been made to demonstrate the dominance of the inertial energy at the ΔX = 12.5 km separation. The inertial fit was computed for the MODE-I center time-series drops 236 through 245. This fit was then rotated in time to the time of the 241/242 drop pair. The rotated fit and the difference profiles are presented in Fig. 15. There is substantial agreement throughout the water column, although the correspondence is clearly better below the thermocline (1000 m). The correspondence may be improved by rotating the inertial fit slightly, but the present example serves to demonstrate the presence of inertial waves in the velocity differences. Incidentally, the superior agreement below 1000 m may add credibility to the argument in the previous section about $\omega^2 - f^2$ being smaller (and, possibly, more narrow band) in the deep water than in the upper ocean.

Table 2 presents the normalized variances of the difference profiles. At the smaller horizontal separation, the normalized kinetic energy is 2.5 ncm^2/s^2. Since the depth-averaged N is 1 cph, the average kinetic energy of the difference profile is about 0.8 cm^2/s^2. There is a steady trend toward higher kinetic energies at larger horizontal separations, culminating in about an order-of-magnitude-larger value at 12.5 km. The mean square differences between the two EM sensors (for redundancy, there is a pair of EM sensors on each profiler) are on average less than

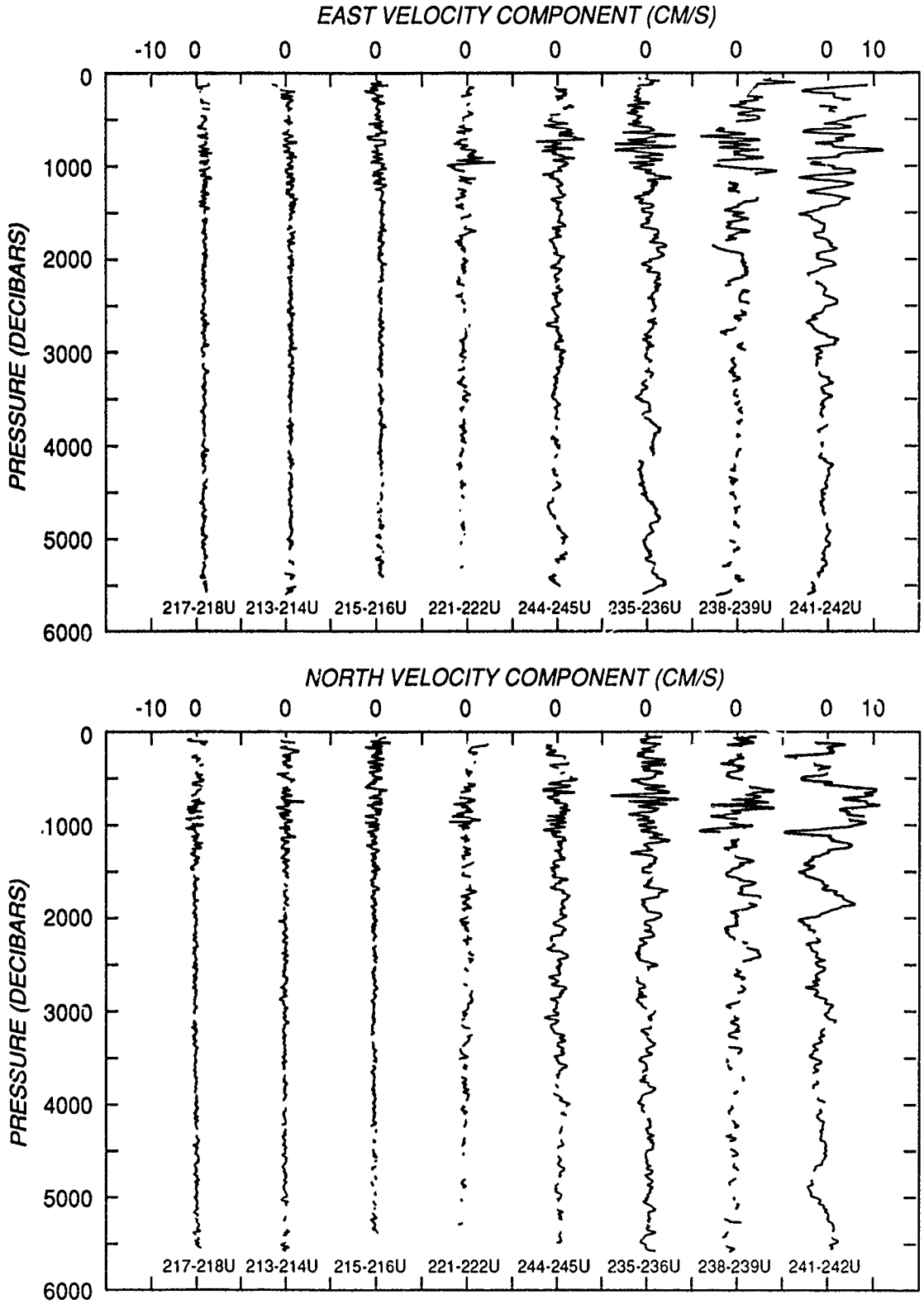


Fig. 14. Profiles of the differences in velocity components for pairs of simultaneous drops. The numbers beneath each profile represent the drop numbers of the pair, and the letters denote use of either the down (D) or up (U) portion of the profiles.

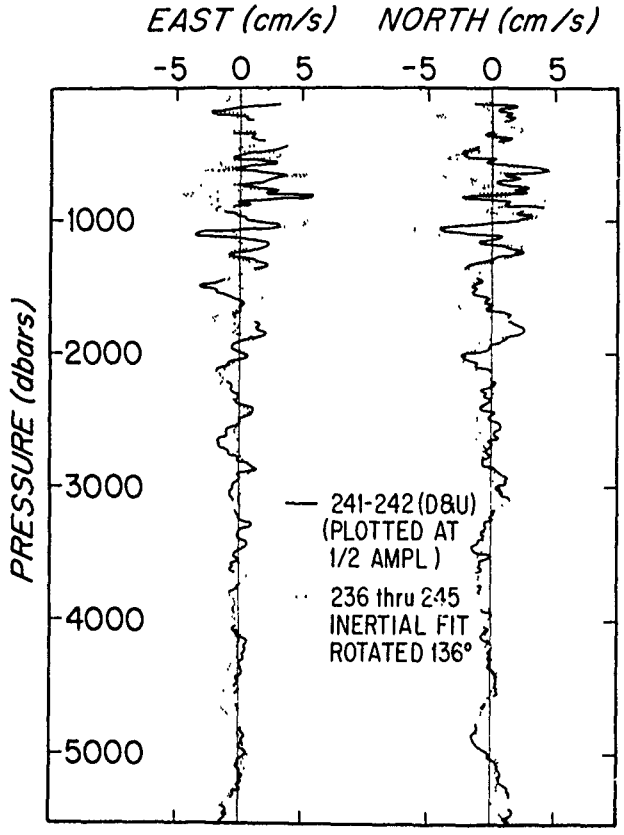


Fig. 15. Difference profile for drops 241 and 242 (down and up portions averaged) plotted at half amplitude (a value chosen to increase visual comparison) versus the inertial fit of the time series at the central mooring (including drop 242) rotated 136° forward in time from the reference time used for display of fit in Fig. 5.

0.5 ncm^2/s^2, which suggests that much of the interprofiler variance is due to apparent velocity differences induced by depth (pressure) errors acting on a shear field. That is, it is thought that most of the variance in profile differences at small ΔX originates from systematic pressure errors between the profilers. A pressure offset shifts one profile vertically with respect to the other, resulting in apparent velocity differences when small vertical-scale shear is present.

The background internal-wave energy has been computed over the MODE center profiles minus a time-mean profile. The time-mean profile was constructed from all MODE center profiles from drop 235 through drop 245, the profiles constituting the largest horizontal separations and largest energy differences. The residue profiles are considered to be of internal waves and have an average normalized energy of 15.4 ncm^2/s^2. Thus the difference profiles at $\Delta X = 6.7$ km contain a kinetic energy level about equal to that in the internal-wave background field.

Table 2. A summary of WKBJ normalized horizontal kinetic energy of difference profiles. The average N over the water column is 1 cph, and the value of N_o is 3 cph.

Profiles	Separation (km)	% good of 556 pts	ΔKE (ncm/s)
217-218D	0.015	96	1.8
			2.1
217-218U	0.015	89	2.3
213-214D	0.05	89	2.4
			2.4
213-214U	0.05	87	2.3
215-216D	0.1	42	2.7
			2.5
215-216U	0.1	83	2.3
221-222D	0.4	37	2.6
			2.9
221-222U	0.4	55	3.1
244-245D	1.5	60	4.7
			5.1
244-245U	1.5	82	5.5
235-236D	4.8	76	11.1
			11.1
235-236U	4.8	94	11.0
238-239D	6.7	47	14.6
			14.3
238-239U	6.7	64	14.0
341-242D	12.5	86	23.2
			23.6
341-242U	12.5	93	23.9

The results shown in Table 2 are plotted in Fig. 16. The solid line is a fit to the data with an intercept of 2.5 ncm^2/s^2, which is taken to be the noise level common to measurements at all separations. The dashed line has the same slope as the solid line but with the additive noise removed.

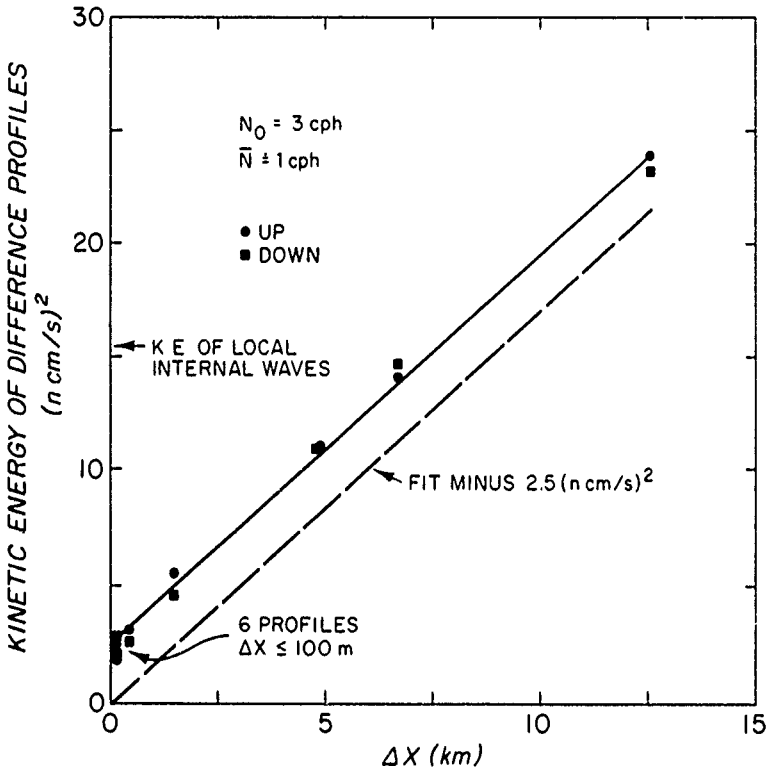


Fig. 16. Kinetic energy of the difference profiles for simultaneous pairs of Fig. 15 in normalized energy units. Both down and up portions of profile pairs are shown. If the energy of difference pairs at a small separation is uncorrelated noise, the fit can be reduced by this noise (about 2.5 ncm/s), and the result is shown as a dashed line. One way to view these data is to recognize that the profiles have zero correlation when the kinetic energy of the difference profiles exceeds the average kinetic energy of the local internal wave field (15 (ncm/s)^2).

The cross-structure function can be expressed in terms of correlation functions and the internal-wave energy in each profile going into the pair. Let u_1, v_1 represent profile 1 of a pair and u_2, v_2 represent the other profile. Then

$$KE \text{ of } \Delta\underline{v} = \frac{1}{2}\left[\overline{(u_1-u_2)^2} + \overline{(v_1-v_2)^2}\right], \tag{13}$$

where the overbar represents a vertical average over all (good) data, which for the following is assumed to be the whole water column. The cross correlation function is

$$R_{u_1u_2} = \frac{u_1u_2}{\left(\overline{u_1^2}\cdot\overline{u_2^2}\right)^{1/2}}, \tag{14}$$

with a similar expression for $R_{v_1v_2}$. Assuming the internal wave statistics are spatially uniform, then $\overline{u_1^2} = \overline{u_2^2} \equiv \overline{u^2}$, $\overline{v_1^2} = \overline{v_2^2} \equiv \overline{v^2}$, and u and v are isotropic ($R_{u_1u_2} = R_{v_1v_2} \equiv R$); then

$$\tfrac{1}{2}\,\overline{|\Delta v|^2} = 2(1-R)\cdot\tfrac{1}{2}\overline{v^2}. \tag{15}$$

The KE of difference profiles should be zero at $\Delta X = 0$, increasing to four times the average KE of v as R goes toward –1 (only for single wave) and ultimately approaching two times the average KE of v as R goes to zero for large ΔX. From Fig. 16 it is evident that $\frac{1}{2} |\Delta v|^2 = v^2$ (i.e., $R = 0$) for $\Delta X \sim 15$–20 km. The separations are not large enough to reveal the asymptotic behavior of the difference KE.

The MODE-I eddy does not contribute significantly to the velocity differences presented in Fig. 16. The absence of low-mode eddy structure in the differences of Fig. 15 demonstrates the absence of mesoscale eddy contamination. The MODE-I eddy had a zero correlation scale of 100 km (MODE Group, 1978) and a correlation function of the form

$$R = \cos\frac{\pi\Delta X}{200\ \text{km}} \quad \text{for } \Delta X \ll 100\ \text{km}\,. \tag{16}$$

The eddy field had a local KE of 25 $(\text{cm/s})^2$, or in normalized units was 75 ncm^2/s^2 since $N_o / \bar{N} \doteq 3$. Thus the eddy interference is estimated to be about 1 ncm^2/s^2 at $\Delta X = 6.7$ km and 3 ncm^2/s^2 at 12.5 km. It may be that the mean square differences at $\Delta X > 15$ km would be strongly influenced by mesoscale contamination. On the other hand, the drops at large ΔX were aligned parallel with the flow and, hence, normal to the first-order velocity gradient. Moreover, little of the MODE-I eddy structure, primarily first mode, is seen in the velocity differences of Fig. 14.

6. CONCLUSIONS

The present study describes the spatial structure of thermocline and abyssal internal waves. Velocity profiles usually span only the upper ocean and are seldom deeper than the base of the main thermocline. WKB normalization and stretching are applied to the velocity components and vertical coordinate, respectively. The WKB-scaled profiles exhibit much more vertical homogeneity and can be used to reveal differences between upper-ocean and abyssal internal-wave characteristics.

The principal conclusions of the analyses of these velocity profiles are

1. East and north contributions to total internal wave kinetic energy are equal as is consistent with statistical isotropy in the horizontal.
2. Bulk of the internal wave energy is contributed by near-inertial motions.
3. Frequency-wavenumber dependence of internal wave spectra is nonseparable: near-inertial motions exhibit a power law $\propto m^{-3}$, whereas higher frequency waves exhibit a power law $\propto m^{-2}$.
4. Near-inertial wave motion is clockwise polarized with depth and is most energetic at 500-m wavelength with lesser energy contributed by the band 100–300 m.
5. High-frequency internal waves exhibit no departures from isotropy, either in the vertical (i.e., polarization) or in the horizontal coordinates.
6. Upper-ocean internal waves exhibit a spectral cutoff in high-vertical wavenumber energy compared with the lower half of the water column, suggesting action by dissipative effects in the upper ocean and/or scattering effects ("whitening") in the abyssal ocean.

7. A horizontal, spatial correlation scale of 15–20 km existed for the MODE-I internal wave field.

Even the crude frequency separation possible with these short time series reveals the distinctly different spectral slopes for the inertial waves compared with the high-frequency internal waves. Spectral stability is achieved with ensemble averaging, rather than band averaging. As a result, the spectra retain the lowest wavenumber components, including essentially one wavelength over the entire water column.

The near-inertial motions predominate at vertical wavelengths longer than 100–200 m, at which point the high-frequency internal waves become more energetic. These differences may be expressions of wave–wave processes. McComas and Müller (1981a) find that, under conditions of energy or action fluxes governed by the processes of induced diffusion and dissipation, a spectral slope of –2 at all frequencies is not an equilibrium solution in the presence of flux and dissipation. In contrast, power law slopes of –2.5 at inertial and –2.0 at high frequencies are near-equilibrium solutions. The steeper slope for near-inertial motions supports the view that the wave field may be in equilibrium.

These results demonstrate that a separable frequency and vertical wavenumber description of these data is not appropriate. Garrett and Munk (1972, 1975) suggested a spectral model in which the same vertical wavenumber dependence was hypothesized at all frequencies. Clearly, such a model is not supported by these observations. The maverick is the inertial motions; it seems likely that the higher-frequency waves do follow a frequency-independent vertical structure. The inertial contribution, on the other hand, is more distinct owing to spotty generation and its isolation or uniqueness among the internal wave continuum. It seems less interactive and more independent than the rest of the internal waves (McComas and Bretherton, 1977; Olbers, 1983; McComas and Müller, 1981a,b).

At the longest vertical wavelengths, approximately one or two cycles per (normalized) water depth, the CW and ACW energy components of the inertial motions decrease markedly, becoming about equal at the lowest wavenumber. The tendency toward modes or symmetric energy composition is to be expected because of surface and bottom reflections; the decrease in energy for wavelengths greater than 500 sm is not easily understood. Pollard (1970) predicts that mode 1 is most easily forced by surface stress because of its large vertical group velocity. Perhaps the time-mean shear acts to destroy, or dismember, the lowest modes.

The separation of the total internal waves into upper and lower halves revealed mostly a divergence of the spectra of the upper ocean relative to the lower part for wavelengths less than 100 m. The inertial component does not exhibit this behavior; the main characteristic of this component is that the lower half is more energetic at all wavenumbers. The source of the divergence of the spectra is the high-frequency internal waves. At vertical wavelengths shorter than 50 m, the slope of the upper-ocean spectra changes from –2 to –4. One possible cause is the interaction of the internal waves with the time-mean thermocline shear. Bottom scattering should move energy from longer vertical wavelengths to shorter. Hence, the upper half/lower half differences may result from the whitening effects of bottom scattering (Rubenstein, 1988).

The zero-correlation distance of 15–20 km is considerably smaller than previously reported. Earlier measurements have largely been near the sea surface and have observed much longer horizontal scales. For example, Pollard (1980) reported horizontal wavelengths of several hundred

meters in the SML. D'Asaro (1984) observed inertial-wave coherence scales that seemed to decrease with depth, consistent with the present, much deeper results. The increase in variance between profiles obtained up to 12.5 km apart, essentially a structure function, is useful in estimating errors in comparing measurements taken at separated sites, such as profiler and moored current meter comparisons.

Velocity profiles over the entire water column are useful for studies of internal waves. The observed energy levels are lower in the deep ocean than in the thermocline but not after WKB scaling. Deep internal waves exhibit characteristics that should be more nearly in equilibrium since they are more distant from surface sources and thermocline interactions. The spatial structure of internal waves is a sensitive measure of wave frequency through the dispersion relation. Wave period determinations can be obtained from the spatial structure of the dominant internal wave with better resolution than a long time-series of moored observations. It is likely that models for internal wave characteristics would be better tested on subthermocline motions. Abyssal velocity observations are needed to evaluate proposals that abyssal mixing rates and diffusivity are larger than in the upper ocean. Methods are available for more extensive subthermocline observations in support of modern internal wave studies.

APPENDIX A

In Section 3 it was asserted that the inertial fit produces a profile that is uncorrelated with the residue or higher-frequency profiles. To examine this concept further, consider a simple, two-component model for the internal wave field:

$$u + iv = A\, e^{ift} + B\, e^{i\omega t} . \tag{A1}$$

The inertial decomposition is the solution to

$$\frac{1}{T}\int_0^T (A\, e^{ift} + B\, e^{i\omega t} - C\, e^{ift})^2 dt = \text{minimum} , \tag{A2}$$

where the summation over the finite number of profiles has been replaced with an integration over the duration T. The solution is

$$C(z) = A(z) + B(z) I , \tag{A3}$$

where

$$I = \frac{1}{T}\int_0^T e^{i(\omega - f)t} dt . \tag{A4}$$

The high-frequency internal waves are the total internal wave field minus the inertial fit:

$$u + iv - C\, e^{ift} = B\, (e^{i\omega t} - I\, e^{ift}) . \tag{A5}$$

The energy spectra with respect to vertical wavenumber averaged over all profiles are as follows, where the symbols ^ and * represent a Fourier transform with respect to depth and complex conjugate, respectively.

For the total wave field:

$$\frac{1}{T}\int_o^T \left[\hat{A}^2+\hat{B}^2+\hat{A}\hat{B}^*\ e^{-1(\omega-f)t}+\hat{A}^*\hat{B}\ e^{i(\omega-f)t}\right]dt \tag{A6}$$

$$=\hat{A}^2+\hat{B}^2+\hat{A}\hat{B}^*I^*+\hat{A}^*\hat{B}I\ .$$

For the inertial component:

$$\frac{1}{T}\int_0^T \left[\hat{A}^2+\hat{B}^2+\hat{A}\hat{B}^*\ e^{-i(\omega-f)t}+\hat{A}^*\hat{B}\ e^{i(\omega-f)t}\right]dt \tag{A7}$$

$$=\hat{A}^2+\hat{B}^2I^2+\hat{A}\hat{B}^*I^*+\hat{A}^*\hat{B}I\ .$$

For the higher-frequency component:

$$=\frac{1}{T}\int_0^T \hat{B}^2\left[1-I^*\ e^{i(\omega-f)t}-I\ e^{i(\omega-f)t}+I^2\right]dt \tag{A8}$$

$$=\hat{B}^2(1+I^2-I^*I-II^*) \tag{A9}$$

$$=\hat{B}^2(1-I^2)\ .$$

The sum of the inertial and higher-frequency components is

$$=\hat{A}^2+\hat{B}^2+\hat{A}\hat{B}^*I^*+\hat{A}^*\hat{B}I\ . \tag{A10}$$

Hence, the total energy equals the sum of the parts. In general, for a continuum of internal waves, the spectrum of the total internal wave energy should equal the sum of the inertial and high-frequency waves as here computed.

APPENDIX B

The WKB scaling and normalization (s & n) involves three steps:

1. Normalization based on a chosen reference Brunt-Väisälä frequency:

$$U_1^*(z)=(N_o/N)^{1/2}\,U(z). \tag{B1}$$

2. Adaptive filtering in which original observations are averaged over an interval determined by the local value of N/N_o:

$$U^*(z)=\frac{N(z)}{N_o\Delta z}\int_{\zeta_1}^{\zeta_2} U_1^*(\zeta)d\zeta\ , \tag{B2}$$

where

Δz = the data spacing in the original profile (assumed to be uniform).

$\zeta_1 = z-N_o\Delta z/2\,N(z)$

$\zeta_2 = z+N_o\Delta z/2\,N(z)$.

3. Stretching introduces a new depth scale, z^*, that is related to the measured depth scale:

$$z^* = \frac{1}{N_o}\int_o^z N(\zeta)\,d\zeta\,. \tag{B3}$$

Schematically the procedure is shown in Fig. B1.

The horizontal kinetic energy in the s & n series is

$$\frac{1}{2}\left[\overline{U^{*2}}+\overline{V^{*2}}\right] = \frac{1}{2H^*}\int_{-H^*}^{o}\left\{\left[U^*(z^*)\right]^2+\left[V^*(z^*)\right]^2\right\}dz^*\,. \tag{B4}$$

For now, ignore the effect of the adaptive filter and note that

$$dz^* = \frac{N(z)}{N_o}\,dz\,, \tag{B5}$$

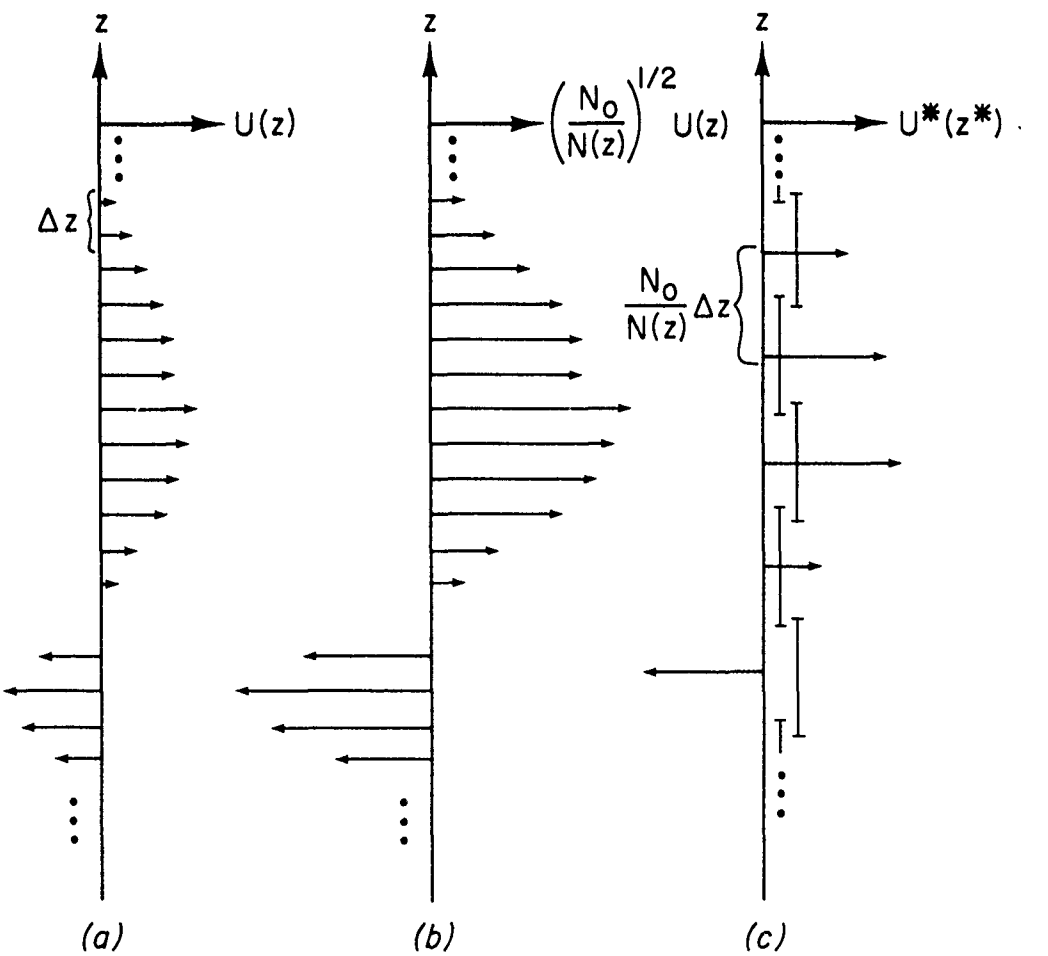


Fig. B1. (a) Example of a profile of velocity component $U(z)$. (b) The same profile after scaling by $[N_0/N(z)]^{1/2}$. (c) The vertical sampling resulting from WKB stretching of the vertical coordinate.

$$\bar{N} = \frac{1}{H}\int_{z}^{o} N(\zeta)\, d\zeta \tag{B6}$$

and

$$H^* = \int_{-H}^{o} \frac{N(\zeta)}{N_o} d\zeta = \frac{H\bar{N}}{N_o} . \tag{B7}$$

So

$$\frac{1}{H^*} \int_{-H^*}^{o} [U^*(z^*)]^2 dz^* = \frac{N_o}{H\bar{N}} \int_{-H}^{o} U^2(z) dz \tag{B8}$$

or

$$\overline{U^{*2}} = \frac{N_o}{\bar{N}}\, \overline{U^2} , \tag{B9}$$

when no adaptive filtering is used.

In the case in which the $U^*(z^*)$ series is split in two, the upper half $\bar{N} \sim 2$ cph and the lower half $\bar{N} \sim 0.7$ cph. Thus the KE is $N_o/\bar{N} \sim 3$ times larger in the stretched and normalized series than in the original. If the noise, U_N, were independent of depth, then

$$\overline{U_N^2}_{upper} = 1.5\, \overline{U_N^2}$$

and

$$\overline{U_N^2}_{lower} = 4.5\, \overline{U_N^2} .$$

Hence, an amplification of noise is expected in the lower half relative to the upper. This is what is seen.

However, include now the adaptive filter which averages over measurements with a box-car width proportional to $N_o/N(z)$. The noise variance should decrease in proportion to $N(z)/N_o$ (i.e., inversely to the number of data points taken into the average). The noise contribution is now $[N(z)/N_o]\, U_N^2$, and the stretched and normalized energy is

$$\overline{U_N^{*2}} = \overline{\frac{N_o}{N}\left[\frac{N}{N_o}\, U_N^2\right]} \sim \overline{U_N^2} . \tag{B10}$$

Therefore, no amplification of noise is expected.

According to Sanford *et al.* (1978) the difference between simultaneous velocity profiles by two EMVP instruments is less than 0.7 cm/s rms. The noise variance is then ~0.5 $(\text{cm/s})^2$ If it is considered as spectrally white and due to equal contributions from each instrument, then

$$E(m) = \frac{0.5\ (\text{cm/s})^2}{50\ \text{cpkm}} \cdot 1/2 = 5 \cdot 10^{-3}\ (\text{cm/s})^2/\text{cpkm} .$$

This variance should be independent of depth and comparable to that seen on the upper half versus the lower half spectra of Fig. 11. That is, the difference between the curves of Fig. 11 could be explained by larger noise in the deep water.

This analysis suggests that the differences between the upper and lower half may be real but they are close to the noise level, although there is no expectation for more noise in the lower half than in the upper.

With the WKB scaling and normalization, the horizontal kinetic energy *density* estimates are independent of N. The KE density is $|\hat{U}^*(m^*)|^2/\delta m^*$ where $\hat{}$ represents the Fourier transform and δm^* is the wavenumber resolution, i.e., $2\pi/H^*$. Since $U^* = (N_o/N)^{1/2}U$ and $\delta m^* = N_o \delta m/N$, and under the WKB assumption variations in N are small compared with those of U at any m, then $\hat{U}^{*2}/\delta m^* = \hat{U}^2/\delta m$. Therefore, spectral density estimates are invariant to N but are shifted by WKB scaling and normalization with respect to m. For example, suppose the wrong value of N was used for some profiles. The values of energy density will not be incorrect, but the wavenumbers will be. Hence, if there was a strong single wave in all velocity profiles, the use of a wrong N profile would result in correct energy density estimates assigned to the wrong vertical wavenumber.

ACKNOWLEDGMENTS

This work was possible only through the strenuous efforts of a number of people. Bob Drever, John Dunlap, Art Bartlett, and Ed Denton developed and operated the profilers with occasional but significant help from others. The Office of Naval Research supported this effort for many years, including instrument development, data collection, data analysis, manuscript preparation, and meeting attendance. I regret that it has taken so long to prepare these results and sincerely appreciate the support and patience of several ONR program managers. A number of scientists have encouraged the publication or discussed aspects of these data. Chris Garrett, Ann Gargett, Greg Holloway, and Eric D'Asaro have contributed thoughts on the subject of abyssal internal waves and the present results. I wish more of their wisdom could have been understood and incorporated in this manuscript.

REFERENCES

Briscoe, M. G., 1975: Preliminary results from the tri-moored internal wave experiment (IWEX). *J. Geophys. Res.*, **80**, 3872–3884.

Cairns, J. L., and G. O. Williams, 1976: Internal wave observations from a midwater float, Pt. 2. *J. Geophys. Res.*, **81**, 1943–1950.

D'Asaro, E. A., 1984: Wind forced internal waves in the North Pacific and Sargasso Sea, *J. Phys. Oceanogr*, **14**, 781–794.

Eriksen, C. C., 1985: Implications of ocean bottom reflection for internal wave spectra and mixing. *J. Phys. Oceanogr.*, **15**, 1145–1156.

Fu, L-L., 1981: Observations and models of inertial waves in the deep ocean. *Rev. Geophys. Space Phys.*, **6**, 447–472.

Gargett, A. E., P. J. Hendricks, T. B. Sanford, T. R. Osborn and A. J. Williams III, 1981: A composite spectrum of vertical shear in the upper ocean. *J. Phys. Oceanogr.*, **11**, 1258–1271.

Garrett, C. J. R., and W. H. Munk, 1972: Space-time scales of internal waves. *Geophys. Fluid Dyn.*, **2**, 225–264.

Garrett, C. J. R., and W. H. Munk, 1975: Space-time scales of internal waves: A progress report. *J. Geophys. Res.*, **80**, 291–307.

Gordon, L., 1978: Internal wave climate near the coast of northwest Africa during JOINT-I. *Deep-Sea Res.*, **25**, 625–643.

Gregg, M. C., 1977: Variations in the intensity of small-scale mixing in the main thermocline. *J. Phys. Oceanogr.*, **7**, 436–454.

Hayes, S. P., 1975: Preliminary measurements of the time lagged coherence of vertical temperature profiles. *J. Geophys. Res.*, **80**, 307–311.

Johnson, C. L., and T. B. Sanford, 1980: Anomalous behavior of internal gravity waves near Bermuda. *J. Phys. Oceanogr.*, **10**(12), 2021–2034.

Käse, R., and D. Olbers, 1979: Wind-driven inertial waves observed during Phase III of GATE. *Deep-Sea Res.*, **26**, 191–216.

Katz, E. J., 1975: Tow spectra from MODE. *J. Geophys. Res.*, **80**, 1163–1167.

Kundu, P. K., 1976: An analysis of inertial oscillations observed near Oregon coast. *J. Phys. Oceanogr.*, **6**, 879–893.

Kunze, E., and T. B. Sanford, 1984: Observations of near-inertial waves in a front. *J. Phys. Oceanogr.*, **14**, 566–581.

Kunze, E., and T. B. Sanford, 1986: Near-inertial wave interactions with mean flow and bottom topography near Caryn Seamount. *J. Phys. Oceanogr.*, **16**, 109–120.

Leaman, K. D., 1976: Observations on the vertical polarization and energy flux of near-inertial waves. *J. Phys. Oceanogr.*, **6**, 894–908.

Leaman, K. D., and T. B. Sanford, 1975: Vertical energy propagation of inertial waves: A vector spectral analysis of velocity profiles. *J. Geophys. Res.*, **80**, 1975–1978.

McComas, C. H. III, and F. P. Bretherton, 1977: Resonant interaction of oceanic internal waves. *J. Geophys. Res.*, **82**, 1397–1412.

McComas, C. H. III, and P. Müller, 1981a: Time scales of resonant interactions among oceanic internal waves. *J. Phys. Oceanogr.*, **11**, 139–147.

McComas, C. H. III, and P. Müller, 1981b: The dynamic balance of internal waves. *J. Phys. Oceanogr.*, **11**, 970–986.

MODE Group, 1978: The Mid-Ocean Dynamics Experiment. *Deep-Sea Res.*, **25**, 859–910.

Müller, P., D. Olbers and J. Willibrand, 1978: The IWEX spectrum, *J. Geophys. Res.*, **83**, 479–500.

Munk, W., 1981: Internal waves and small-scale processes. In: *Evolution of Physical Oceanography*, B. Warren and C. Wunsch Eds., MIT Press, Cambridge, MA.

Munk, W., and N. Phillips, 1968: Coherence and band structure of inertial motions in the sea. *Rev. Geophys.*, **6**, 447–472.

Olbers, D., 1983: Models of the oceanic internal wave field, *Rev. Geophys. Space Phys.*, **21**, 1567–1606.

Pedlosky, J., 1979: *Geophysical Fluid Mechanics*. Springer-Verlag, New York, 624 pp.

Perkins, H., 1970: Inertial oscillations in the Mediterranean. Ph.D. thesis, Mass. Inst. of Technol./Woods Hole Oceanogr. Inst., Woods Hole, MA.

Phillips, O. M., 1977: *The Dynamics of the Upper Ocean*, 2nd edition. Cambridge University Press, London, 336 pp.

Pinkel, R., 1975: Upper ocean internal wave observations from FLIP. *J. Geophys. Res.*, **80**, 3892–3910.

Pinkel, R., 1981: On the use of Doppler sonar for internal wave measurements. *Deep-Sea Res.*, **28**, 269–289.

Pinkel, R., 1984: Doppler sonar observations of internal waves: The wave-number frequency spectrum. *J. Phys. Oceanogr.* **14**, 1249-1270.

Pinkel, R., 1985: A wavenumber–frequency spectrum of upper ocean shear. *J. Phys. Oceanogr.*, **15**, 1453–1469.

Pollard, R. T., 1970: On the generation by winds of inertial waves in the ocean. *Deep-Sea Res.*, **17**, 795–812.

Pollard, R. T., 1980: Properties of near-surface inertial oscillations. *J. Phys. Oceanogr.*, **10**, 385–398.

Rossby, H. T., and T. B. Sanford, 1976: A study of velocity profiles through the main thermocline. *J. Phys. Oceanogr.*, **6**, 766–774.

Rubenstein, D., 1988: Scattering of inertial waves by rough bathymetry. *J. Phys. Oceanogr.*, **18**, 5–18.

Sanford, T. B., 1975: Observations of the vertical structure of internal waves. *J. Geophys. Res.*, **80**, 3861–3871.

Sanford, T. B., and N. G. Hogg, 1977: The North Atlantic fine and microstructure cruise *KNORR 52* and *EASTWARD 75-12*. Tech. Rep. 77-11, Woods Hole Oceanogr. Inst., Woods Hole, MA.

Sanford, T. B., R. G. Drever and J. H. Dunlap, 1978: A velocity profiler based on the principles of geomagnetic induction. *Deep-Sea Res.*, **25**, 183–210.

Sanford, T. B., P. G. Black, J. R. Haustein, J. W. Feeney, G. Z. Forristall and J. F. Price, 1987: Ocean response to a hurricane. Part I: Observations. *J. Phys. Oceanogr.*, **11**, 2065–2083.

Stegen, G. R., K. Bryan, J. L. Held and F. Ostapoff, 1975: Dropped horizontal coherence based on temperature profiles in the upper thermocline. *J. Geophys. Res.*, **80**, 3841–3847.

Webster, F., 1968: Observations of inertial-period motions in the deep sea. *Rev. Geophys.*, **6**, 473–490.

ESTIMATES OF SMALL–SCALE HORIZONTAL DIVERGENCE AND RELATIVE VORTICITY IN THE OCEAN

Ren-Chieh Lien and Peter Müller

Department of Oceanography
School of Ocean and Earth Science and Technology
University of Hawaii at Manoa
Honolulu, Hawaii 96822

ABSTRACT

An attempt to estimate horizontal divergence and the vertical component of relative vorticity at small scales is made using measurements from Internal Wave Experiment. These two quantities are very useful to decompose the small–scale motion into the vortical and gravity modes. Fluctuations of horizontal divergence and relative vorticity estimated using a three–point array are found to be attenuated and mutually contaminated at small horizontal scales. Assuming a horizontal isotropy condition, these sampling errors can be represented as two array response functions which act as horizontal wavenumber spectral windows imposing on wavenumber–frequency spectra of horizontal divergence and relative vorticity. Examining effects of array response functions for the GM–76 spectrum suggests that both the attenuation and contamination effects must be considered in obtaining the relative vorticity spectrum, whereas only the attenuation effect is important for the horizontal divergence spectrum since the contamination from relative vorticity is negligible.

INTRODUCTION

Small-scale motions in the ocean consist of both gravity waves and vortical motion. The vortical motion is distinguished from gravity waves by carrying the perturbation potential vorticity in the system. Therefore, it is likely that small-scale vortical motion plays an important role in the enstrophy (variance of potential vorticity) cascade in the ocean.

A normal mode decomposition scheme was proposed by Müller (1984) to project small-scale oceanic motion into the gravity and vortical modes. It can be achieved most conveniently using fields of horizontal divergence, the vertical component of relative vorticity, and vortex stretching (Lien, 1990). Attempts to estimate these three fields at small scales in the ocean have been made using measurements from Internal Wave Experiment (IWEX; Briscoe, 1975) by Müller et al. (1988) and by Lien (1990). Estimates

of "area-averaged" horizontal divergence and relative vorticity were obtained using spatially discrete velocity measurements.

It was pointed out by Prater (1989) and Kunze et al. (1990) that the three-point array of IWEX will detect "relative vorticity" even in a pure horizontal divergence flow field (free of relative vorticity). This paper examines potential errors of estimates of horizontal divergence and relative vorticity. The estimation of horizontal divergence and relative vorticity at small scales using IWEX measurements is described in the following section. Next, their spectral analysis, associated array response functions, and their effects for the GM-76 spectrum model are discussed. Conclusions are summarized in the last section.

ESTIMATES OF HORIZONTAL DIVERGENCE (HD) AND RELATIVE VORTICITY (RV)

An attempt to separate small-scale oceanic fluctuations into the gravity and vortical modes can be achieved conveniently using fields of HD, RV, and VS at small scales. To estimate these fields requires oceanic measurements of horizontal velocity and temperature with a sufficient spatial resolution. Oceanic measurements from IWEX seem to be suitable for such calculations.

The IWEX was conducted in late 1973 over a 42-day period. A trimooring array was designed on which 20 current meters (17 VACM and 3 EG&G 850) and temperature sensors were deployed in the main thermocline of the Sargasso Sea (27°44′ N, 69°51′ W). Horizontal velocity components, temperature, and temperature difference over a vertical distance of 1.74 m were measured. Horizontal spacing between sensors ranged from 1.4 m to 1600 m and vertical spacing from 2.1 m to 1447 m. Sampling interval was 225 s, except at the lowest level (2050 m depth) which was sampled every 900 s. The trimooring array was a nearly perfect tetrahedron (roughly 6 km on a side) with the apex at the top of the main thermocline at 604 m depth and the deepest current meter and temperature sensor at a depth of 2050 m. A schematic diagram of IWEX is shown in Figure 1. The mooring was very stable during the entire experiment. Pressure records showed ±0.2 m displacement at the apex and about ±6 m at 3000 m. A detailed description of IWEX was given by Tarbell et al. (1976). The IWEX measurements provide an opportunity to estimate spatial gradients in the time and space scales of small-scale motions. Measurements from 15 current meters and temperature sensors are used at five horizontal planes where measurements are available at all three legs. Characteristics of the five IWEX levels are described in Table 1.

Area-averaged horizontal divergence $\overline{HD}$ and relative vorticity $\overline{RV}$ can be obtained from velocity measurements on a horizontal plane using Stokes' and Gauss' theorems as

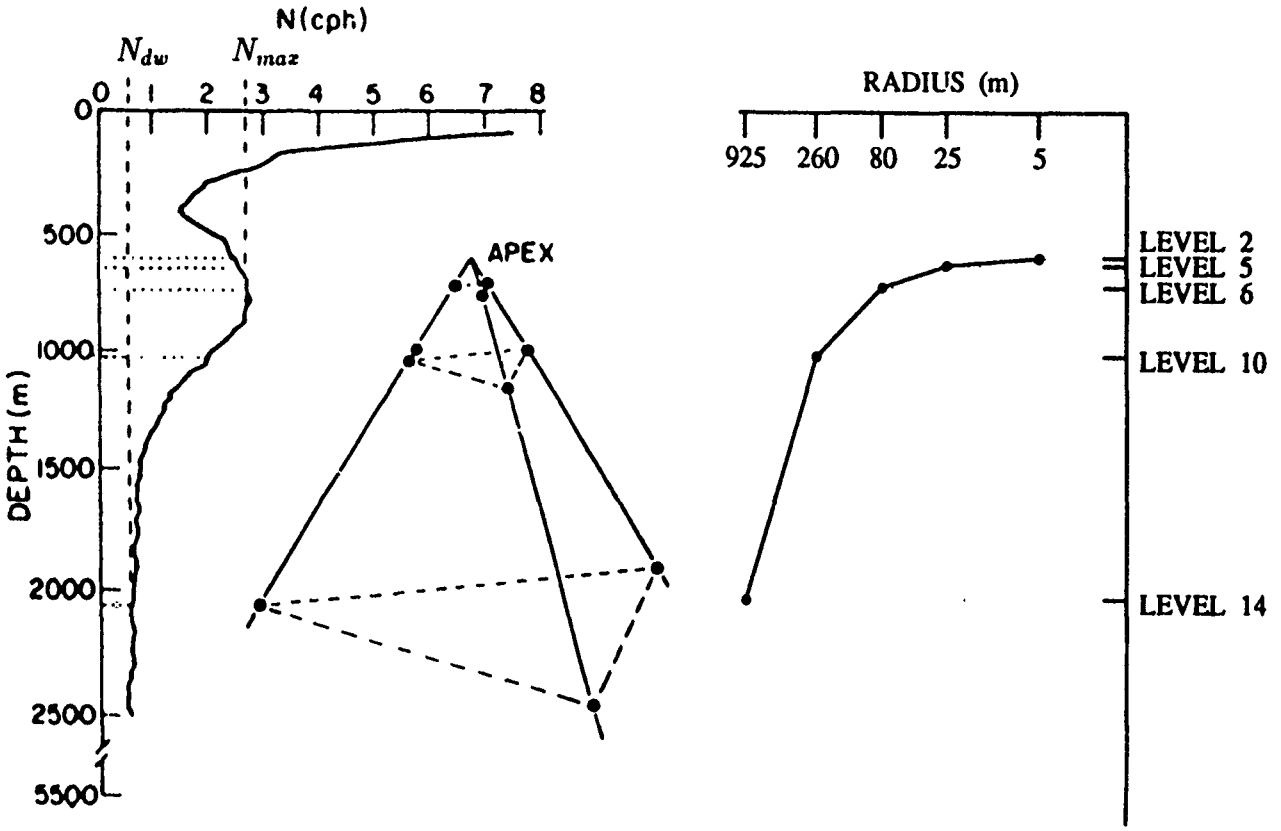


Figure 1. Schematic view of the geometry of the IWEX array and profiles of the Brunt-Väisälä frequency $N(z)$ and horizontal radius $R(z)$. Points indicate current meter positions. There are ten more current meters near the apex which are not shown. The levels that contain three current meters are indicated. The maximum Brunt-Väisälä frequency in the main thermocline is $N_{max} = 2.76$ cph. In the deep water column below 2050 m, N is almost constant, $N_{dw} = 0.36$ cph

$$\overline{HD} = \frac{1}{A} \oint \underline{u} \cdot d\underline{n}, \tag{1}$$

$$\overline{RV} = \frac{1}{A} \oint \underline{u} \cdot d\underline{t}. \tag{2}$$

Here $\underline{t}$ and $\underline{n}$ are the tangential and the normal unit vectors along the circumference of an area A, and $\underline{u}$ is the horizontal velocity vector.

Assuming N velocity sensors ($N = 3$ for IWEX trimooring array) evenly located on a circle on a horizontal plane (Figure 2), $\overline{HD}$ and $\overline{RV}$ are estimated by approximating the circle integration by a discrete sum of the radial and normal velocity components along the circle, i.e.,

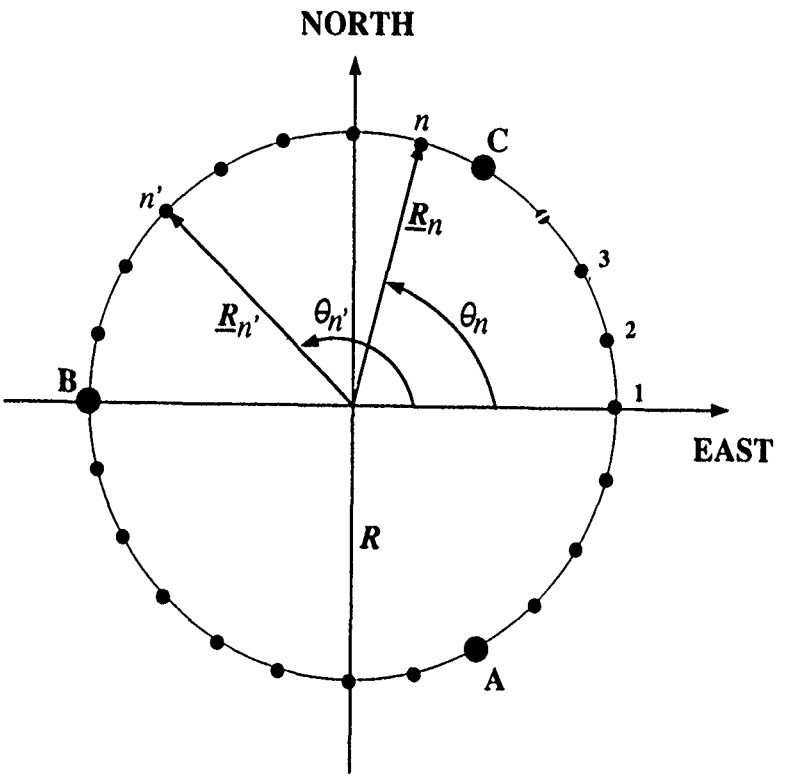


Figure 2. Configuration of N current sensors evenly located on a circle at a horizontal plane. Three larger dots, indexed by A, B, and C, denote positions of current sensors on the IWEX trimooring array. The radius of the circle is R, θ is the orientation of the current sensor counterclockwise from the east, and $\underline{R}$ is the corresponding position vector.

Table 1. Characteristics of five IWEX levels with three current meters

Level	Depth (m)	Radius (m)	N (cph)	Number of sampling
2	606	4.9	2.54	1800
5	640	25.4	2.60	12000
6	731	80.3	2.76	12000
10	1023	260.0	2.05	4800
14	2050	925.0	0.66	3900

$$\overline{HD} = \frac{1}{\pi R^2} \sum_{n=1}^{N} u_r^{(n)} \frac{2\pi R}{N}, \tag{3}$$

$$\overline{RV} = \frac{1}{\pi R^2} \sum_{n=1}^{N} u_t^{(n)} \frac{2\pi R}{N}, \tag{4}$$

where R is the radius of the circle, and $u_r^{(n)}$ and $u_t^{(n)}$ are the radial and normal velocity components of the n^{th} sensor defined as

$$u_r^{(n)} = u_r(R, \theta_n) = u(R, \theta_n) cos(\theta_n) + v(R, \theta_n) sin(\theta_n), \tag{5}$$

$$u_t^{(n)} = u_t(R, \theta_n) = -u(R, \theta_n) sin(\theta_n) + v(R, \theta_n) cos(\theta_n). \tag{6}$$

Here, θ_n is the angle counterclockwise from the east of the n^{th} current sensor. For the IWEX trimooring array, area-averaged horizontal divergence $\overline{HD}$ and relative vorticity $\overline{RV}$ are estimated as

$$\overline{HD} = \frac{2}{3R}\left(\frac{u_A + u_C}{2} - u_B\right) + \frac{1}{\sqrt{3}R}(v_C - v_A), \tag{7}$$

$$\overline{RV} = \frac{2}{3R}\left(\frac{v_A + v_C}{2} - v_B\right) - \frac{1}{\sqrt{3}R}(v_C - u_A). \tag{8}$$

Alternatively, $\overline{HD}$ and $\overline{RV}$ can be obtained using velocity gradients estimated from the least squares fitting of horizontal velocity measurements. These two approaches are found to be equivalent. Time series of $\overline{HD}$ and $\overline{RV}$ at five different depths of IWEX are obtained. Applying the run test and the goodness test, the time series of $\overline{HD}$ and $\overline{RV}$ are found to be stationary and Gaussian distributed (Lien, 1990).

SPECTRAL ANALYSIS OF $\overline{HD}$ and $\overline{RV}$

Time series of $\overline{HD}$ and $\overline{RV}$ are first divided into segments with a length of 1024 (2^{10}) data points. Successive segments are 50% overlapped. Each segment is subjected to a Hanning window and is fast Fourier transformed. The one-sided spectrum, averaged over all segments, is further averaged over adjacent frequency components resulting in 40 frequency points spaced about equally on a logarithmic scale.

Measurement errors on estimates of $\overline{HD}$ and $\overline{RV}$ have been discussed by Müller et al. (1988). Variance due to the instrumental noise can be estimated by separating velocity measurements at level 2 (R = 4.9 m) into coherent signal and incoherent noise components. At level 2, observed frequency spectra of $\overline{HD}$ and $\overline{RV}$ are dominated by the incoherent noise component and will be excluded from our analysis. At other levels, variance due to incoherent noise is removed from the frequency spectral estimates of $\overline{HD}$ and $\overline{RV}$.

Frequency spectra of $\overline{HD}$ and $\overline{RV}$ at levels 5, 6, 10, and 14 are displayed in Figures 3 and 4. Frequency spectra $S_{\overline{HD}}(\omega)$ and $S_{\overline{RV}}(\omega)$ are of the same order at each level and decrease systematically with increasing radius or depth.

Array Response Function

If $\overline{HD}$ and $\overline{RV}$ are obtained from N velocity measurements located on a circle with a radius R, their frequency spectral estimates $S_{\overline{HD}}(\omega; R, N)$ and $S_{\overline{RV}}(\omega; R, N)$ do not represent exactly area-averaged frequency spectra of horizontal divergence and relative vorticity.

Area-averaged horizontal divergence $\overline{HD}$ can be expressed in terms of Fourier transforms as

$$\overline{HD} = \frac{2}{NR}\sum_{n=1}^{N}\int_0^\infty d^2\underline{k}\left[u(\underline{k})e^{i\underline{k}\cdot\underline{R}_n}cos(\theta_n) + v(\underline{k})e^{i\underline{k}\cdot\underline{R}_n}sin(\theta_n)\right]. \tag{9}$$

Assuming statistical homogeneity and stationarity conditions, frequency spectral estimates $S_{\overline{HD}}(\omega; R, N)$ can be written as

$$\begin{aligned} S_{\overline{HD}}(\omega; R, N) = \; & \frac{4}{N^2R^2}\sum_{n=1}^{N}\sum_{n'=1}^{N}\int_0^\infty d\underline{k}^2 e^{i\underline{k}\cdot(\underline{R}_n-\underline{R}_{n'})} \\ & \{P_{uu}(\underline{k},\omega)cos(\theta_n)cos(\theta_{n'}) + P_{vv}(\underline{k},\omega)sin(\theta_n)sin(\theta_{n'}) \\ & + P_{uv}(\underline{k},\omega)cos(\theta_n)sin(\theta_{n'}) + P_{vu}(\underline{k},\omega)sin(\theta_n)cos(\theta_{n'})\}. \end{aligned} \tag{10}$$

Here, $P_{uu}(\underline{k},\omega)$, $P_{vv}(\underline{k},\omega)$, $P_{vu}(\underline{k},\omega)$, and $P_{uv}(\underline{k},\omega)$ are wavenumber-frequency autospectra of u and v, and wavenumber-frequency co-spectra between u and v. Assuming a horizontally isotropic flow field, velocity frequency spectra at a given space lag $(\underline{R}_n - \underline{R}_{n'})$ can be expressed as

$$\begin{aligned} P_{uu}(\omega; \underline{R}_n - \underline{R}_{n'}) = \; & \int_0^\infty d\alpha \left\{[J_0(\alpha r) + cos(\theta_n + \theta_{n'})J_2(\alpha r)]\frac{S_{HD}(\alpha,\omega)}{2\alpha^2}\right. \\ & \left. - [J_0(\alpha r) - cos(\theta_n + \theta_{n'})J_2(\alpha r)]\frac{S_{RV}(\alpha,\omega)}{2\alpha^2}\right\}, \end{aligned} \tag{11}$$

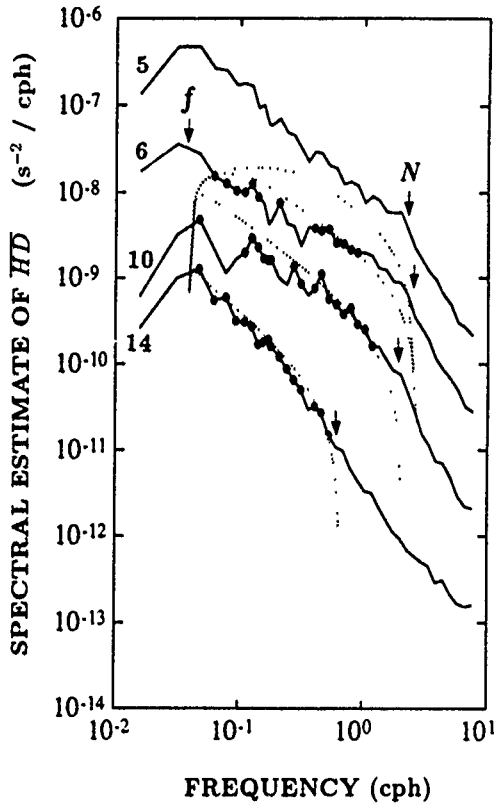


Figure 3. Observed frequency spectra of horizontal divergence at four IWEX levels (solid lines) and comparison with the GM-76 spectrum model (dotted lines). The attenuation and contamination array response functions have been applied to the GM model. Solid circles show agreement between observed spectral estimates and the GM model within 95% confidence interval.

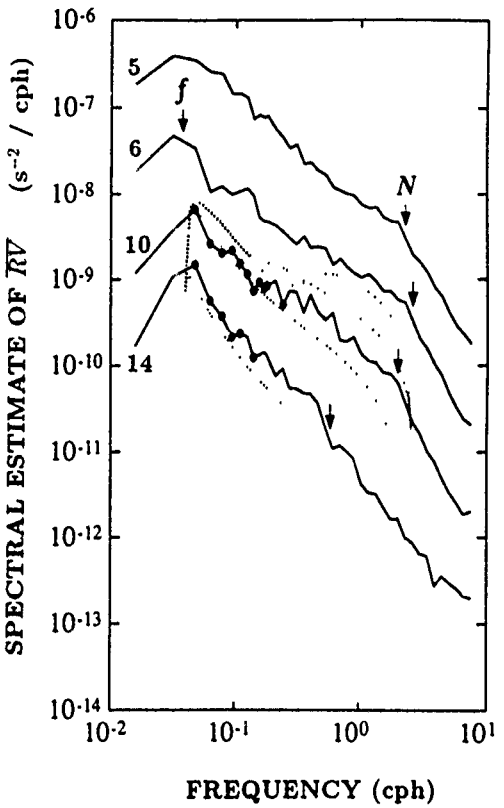


Figure 4. Observed frequency spectra of relative vorticity at four IWEX levels (solid lines) and comparison with the GM-76 spectrum model (dotted lines). The attenuation and contamination array response functions have been applied to the GM model. Solid circles show agreement between observed spectral estimates and the GM model within 95% confidence interval.

$$P_{vv}(\omega;\underline{R}_n-\underline{R}_{n'}) = \int_0^\infty d\alpha \left\{ [J_0(\alpha r) - cos(\theta_n+\theta_{n'})J_2(\alpha r)] \frac{S_{HD}(\alpha,\omega)}{2\alpha^2} + [J_0(\alpha r) + cos(\theta_n+\theta_{n'})J_2(\alpha r)] \frac{S_{RV}(\alpha,\omega)}{2\alpha^2} \right\}, \tag{12}$$

$$P_{uv}(\omega;\underline{R}_n-\underline{R}_{n'}) = P_{vu}(\omega;\underline{R}_n-\underline{R}_{n'}) = \int_0^\infty d\alpha J_2(\alpha r) sin(\theta_n+\theta_{n'}) \frac{S_{HD}(\alpha,\omega)-S_{RV}(\alpha,\omega)}{2\alpha^2}, \tag{13}$$

where $r = 2\alpha R sin(\frac{\theta_n - \theta_{n'}}{2})$, and α is the magnitude of horizontal wavenumber (after Batchelor, 1953). J_0 and J_2 are Bessel functions of zeroth and second orders.

Therefore, the frequency spectral estimate $S_{\overline{HD}}(\omega; R, N)$ can be expressed in terms of real wavenumber-frequency spectra $S_{HD}(\alpha,\omega)$ and $S_{RV}(\alpha,\omega)$ as

$$S_{\overline{HD}}(\omega; R, N) = \int_0^\infty d\alpha \left\{ S_{HD}(\alpha,\omega) F(\alpha R, N) + S_{RV}(\alpha,\omega) G(\alpha R, N) \right\}. \tag{14}$$

Similarly, the frequency spectral estimate of area-averaged relative vorticity can be expressed as

$$S_{\overline{RV}}(\omega; R, N) = \int_0^\infty d\alpha \left\{ S_{RV}(\alpha,\omega) F(\alpha R, N) + S_{HD}(\alpha,\omega) G(\alpha R, N) \right\}, \tag{15}$$

where

$$F(\alpha R, N) = \frac{2}{N^2 \alpha^2 R^2} \sum_{n=1}^{N} \sum_{n'=1}^{N} \left[cos(\theta_n - \theta_{n'}) J_0 \left(2\alpha R sin(\frac{\theta_n - \theta_{n'}}{2}) \right) + J_2 \left(2\alpha R sin(\frac{\theta_n - \theta_{n'}}{2}) \right) \right], \tag{16}$$

and

$$G(\alpha R, N) = \frac{2}{N^2 \alpha^2 R^2} \sum_{n=1}^{N} \sum_{n'=1}^{N} \left[cos(\theta_n - \theta_{n'}) J_0 \left(2\alpha R sin(\frac{\theta_n - \theta_{n'}}{2}) \right) - J_2 \left(2\alpha R sin(\frac{\theta_n - \theta_{n'}}{2}) \right) \right]. \tag{17}$$

For the IWEX configuration ($N = 3$), two array response functions have the form of

$$F(\alpha R) = \frac{2}{3\alpha^2 R^2} \left[1 - J_0(\sqrt{3}\alpha R) + 2J_2(\sqrt{3}\alpha R) \right], \tag{18}$$

$$G(\alpha R) = \frac{2}{3\alpha^2 R^2} \left[1 - J_0(\sqrt{3}\alpha R) - 2J_2(\sqrt{3}\alpha R) \right]. \tag{19}$$

Array response functions $F(\alpha R, N)$ and $G(\alpha R, N)$ are displayed in Figures 5 and 6 for cases of three, nine, and an infinite number of current sensors ($N = 3, 9$, and ∞) along a circle on a horizontal plane. The array response function F behaves as a lowpass wavenumber filter representing the effect of the finite size of the circle. Fluctuations of horizontal divergence and relative vorticity at scales smaller than the size of the circle are attenuated in our estimates. Therefore, F is termed the attenuation array response function. Note that increasing the number of current sensors on the circle cannot eliminate the attenuation problem. The array response function G behaves as a bandpass wavenumber filter. Fluctuations of horizontal divergence and relative vorticity are mutually contaminated in our estimates at horizontal scale comparable to the sensor separation since the tangential and normal components of horizontal velocity at scale smaller than the sensor separation is not detectable. The peak of the contamination array response function moves to higher wavenumber (smaller horizontal scale) and its magnitude is reduced with the increasing number of current sensors on the circle. In the ideal case of an infinite number of current sensors on the circle, the contamination array response function vanishes.

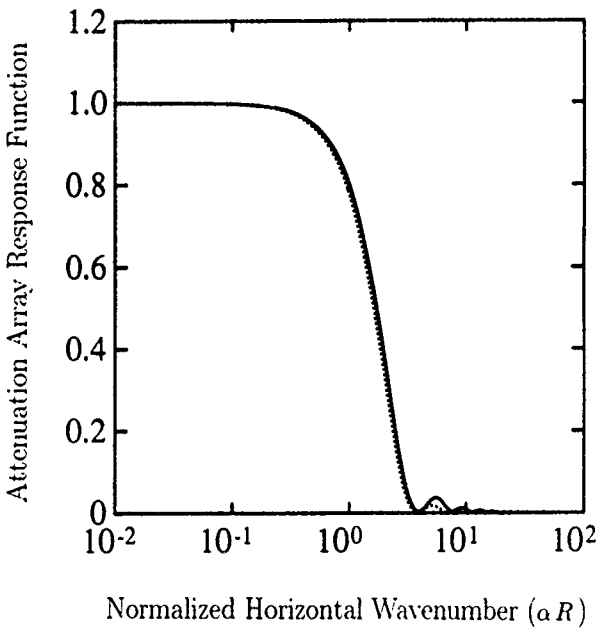


Figure 5. Attenuation array response function for cases of 3 (solid line), 9 (dotted line), and an infinite number of current sensors on a horizontal plane. For the case of an infinite number of current sensors, $F(\alpha R) = 4J_1^2(\alpha R)/(\alpha R)^2$, and it is not distinguishable from the case of 9 current sensors.

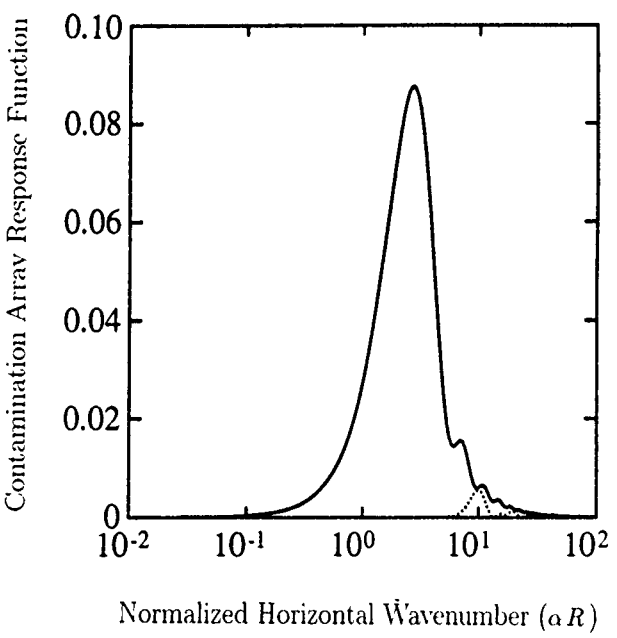


Figure 6. Contamination array response function for cases of 3 (solid line), 9 (dotted line), and an infinite numbter of current sensors, the contamination array response function vanishes.

Effects of Array Response Functions for GM–76 Spectrum

Using the polarization relation of internal waves, frequency spectra $S_{\overline{HD}}(\omega; R, N = 3)$ and $S_{\overline{RV}}(\omega; R, N = 3)$ of the GM–76 spectrum model (Cairns and Williams, 1976) are obtained. Effects of array response functions (see equations 18 and 19) are evaluated employing the wavenumber structure of the GM model in which a vertical cutoff wavenumber of 0.1 cpm is assigned.

Characteristic wavenumbers of the attenuation and contamination array response functions are estimated (Table 2) which vary for different array sizes and sensor separations at four IWEX levels. Note that the characteristic wavenumbers are the property of array response functions, and do not depend on the GM model. Applying the vertical cutoff wavenumber prescribed in the GM model and the dispersion relation of linear internal waves, frequency regimes (Table 2) are identified in which the GM spectrum is affected by array response functions.

The effects of array response functions for the GM model are illustrated in Figures 7 and 8 in which horizontal divergence and relative vorticity frequency spectra of the GM model at IWEX level 10 are estimated. In the absence of array response functions, the horizontal divergence and relative vorticity frequency spectra of the GM model have zero and –2 spectral slopes, respectively. By incorporating the attenuation array response functions, equivalent to area-averaging horizontal divergence and relative vorticity on a circle, frequency spectral slopes of horizontal divergence and relative vorticity become steeper. Since high frequency waves are associated with high wavenumbers whose fluctuations are attenuated, the attenuation is more apparent in the high frequency regime. Including the contamination from relative vorticity, the GM frequency spectrum of horizontal divergence is not significantly modified since the spectral ratio of relative vorticity to horizontal divergence is f^2/ω^2 for linear internal waves. On the contrary, the GM frequency spectrum of relative vorticity is markedly modified because of the contamination from horizontal divergence (Figure 8). Agreements between observed frequency spectra of horizontal divergence and relative vorticity with the corresponding GM spectra are greatly improved by including attenuation and contamination array response functions in the model. Both array response functions must be considered in the calculation of relative vorticity frequency spectrum of the GM model, whereas only the attenuation array response function is important to estimate the horizontal divergence spectrum of the GM model. The comparison of observed frequency spectra of horizontal divergence and relative vorticity with the corresponding GM model including both array response functions at all four levels is presented in Figures 3 and 4. Note that the GM model predicts $\overline{HD}$ much better than $\overline{RV}$. This discrepancy might be due to the existence of vortical motion, which has RV but no HD, or due to the inadequacy of the GM model. This issue is more fully discussed by Lien and Müller (1991).

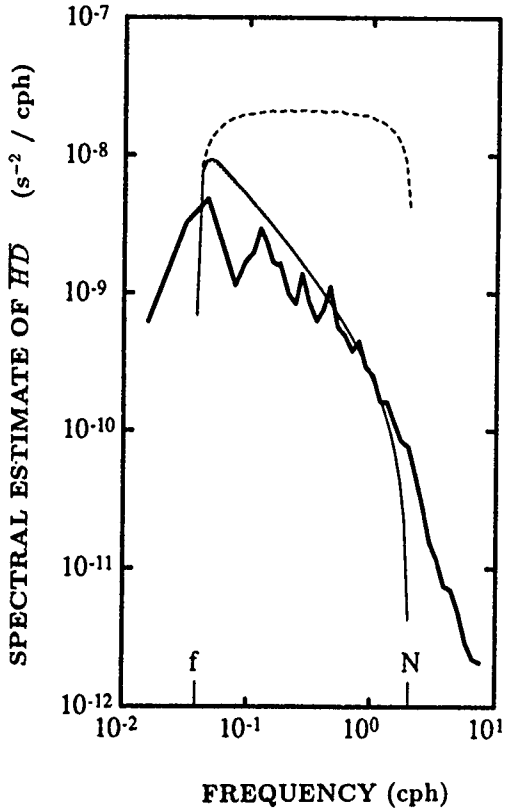


Figure 7. Effects of array response functions on the GM-76 spectrum. The thick solid line denotes the frequency spectral estimate $S_{\overline{HD}}(\omega)$ at IWEX level 10. The GM horizontal divergence spectrum in the absence of attenuation and contamination effects is shown by the dashed line. The dotted line is the horizontal divergence spectrum of the GM model including the attenuation effect, and the thin solid line includes both the attenuation and contamination effects. The dotted and the thin solid lines are not distinguishable.

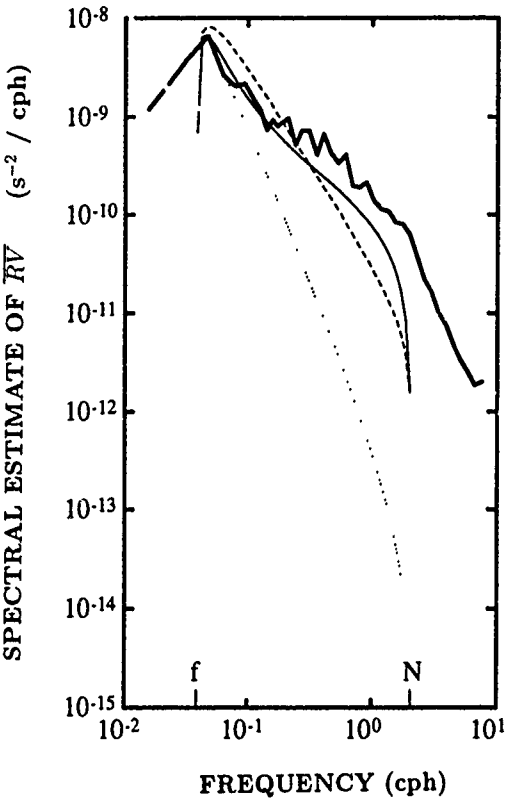


Figure 8. Effects of array response functions on the GM-76 spectrum. The thick solid line denotes the frequency spectral estimate $S_{\overline{RV}}(\omega)$ at IWEX level 10. The GM relative vorticity spectrum in the absence of attenuation and contamination effects is shown by the dashed line. The dotted line is the relative vorticity spectrum of the GM model including the attenuation effect, and the thin solid line includes both the attenuation and contamination effects.

SUMMARY AND DISCUSSION

The estimation of horizontal divergence and relative vorticity is motivated by the need to conveniently project small-scale motion into the gravity and vortical modes using the normal mode decomposition. This paper addresses the practical problems of estimating horizontal divergence and relative vorticity using discrete sampling in space of horizontal velocity measurements.

Table 2. Characteristic horizontal wavenumbers of array response functions and the corresponding frequencies obtained by assuming the dispersion relation of linear internal waves and the vertical cutoff wavenumber β_c of 0.1 cpm. The attenuation array response function drops to 0.90 at α_t, and decreases with increasing horizontal wavenumber by a -2 power law. Frequency spectra of $\overline{HD}$ and $\overline{RV}$ are affected by the attenuation array response function at frequencies greater than $\omega_t(=\sqrt{[N^2\alpha_t^2 + f^2\beta_c^2]/[\alpha_t^2 + \beta_c^2]})$. At lower frequencies, there is no attenuation effect. The contamination array response function behaves as a bandpass filter with its peak at α_c with an amplitude about 0.1, and decreases at a -2 power law away from α_c. Frequency spectra of $\overline{HD}$ and $\overline{RV}$ do not have significant effects of contamination array response function at frequencies smaller than $\omega_c,(=\sqrt{[N^2\alpha_c^2 + f^2\beta_c^2]/[\alpha_c^2 + \beta_c^2]})$.

Level	α_t (cpm)	ω_t (cph)	α_c (cpm)	ω_c (cph)
5	6.3×10^{-3}	0.12	1.7×10^{-2}	0.44
6	2.0×10^{-3}	0.05	5.4×10^{-3}	0.15
10	6.0×10^{-4}	0.04	1.7×10^{-3}	0.05
14	2.0×10^{-4}	0.04	5.0×10^{-4}	0.04

There are two types of sampling errors in the estimation of horizontal divergence and relative vorticity. Fluctuations of horizontal divergence and relative vorticity are attenuated and mutually contaminated at smaller horizontal scales corresponding to the array size and the sensor separation. Decontamination requires horizontal wavenumber structures of horizontal divergence and relative vorticity, which are not yet available. Moreover, the contamination problem causes the normal mode decomposition to be infeasible using IWEX measurements (Lien, 1990).

Effects of array response functions were examined using the GM–76 spectrum model. It was found that both the attenuation and contamination error should be considered in the calculation of the relative vorticity spectrum. Since the variance of relative vorticity is much smaller than that of horizontal divergence for linear internal waves (except on the inertial frequency), only the attenuation effect needs to be considered in obtaining the horizontal divergence spectrum, and the contamination from the relative vorticity is negligible.

ACKNOWLEDGMENTS:

We would like to thank Dr. Eric Kunze, Mark Prater and Niklas Schneider for stimulating discussions and constructive ideas. We are also grateful to Diane Henderson and Twyla Thomas for their editorial assistance. This research is supported by the Office of Naval Research.

REFERENCES:

Batchelor, G. K., 1953: The theory of homogeneous turbulence, Cambridge University Press, 197 pp.

Briscoe, M. G., 1975: Preliminary results from the trimoored internal wave experiment (IWEX), J. Geophys. Res., 80, 27, 3872–3884.

Cairns, J. L., and G. D. Williams, 1976: Internal wave observations from a midwater float, Part II, J. Geophys. Res., 81, 1943–1950.

Kunze, E., A. J. Williams, III and M. G. Briscoe, 1990: Interpreting shear and strain finestructure from a neutrally-buoyant float, J. Geophys. Res., 95, 18111–18125.

Lien, R.-C., 1990: Normal mode decomposition of small-scale oceanic motions, Ph.D. Dissertation, University of Hawaii at Manoa. 128 pp.

Lien, R.-C. and P. Müller, 1991: Normal mode decomposition of small-scale oceanic motions, submitted to J. Phys. Oceanogr.

Müller, P., 1984: Small scale vortical motions. Internal Gravity Waves and Small-Scale Turbulence, Proc., 'Aha Huliko'a Hawaiian Winter Workshop, Hawaii Institute of Geophys., Spec. Pub., P. Müller and R. Pujalet, Eds.

Müller, P., R. C. Lien, and R. Williams, 1988: Estimates of potential vorticity at small scales in the ocean, J. Phys. Oceanogr., 18, 401–416.

Prater, M. D., 1989: Simulations of flow past a triad of current meters, Unpublished manuscript.

Tarbell, S. M., M. G. Briscoe, and D. Chausse, 1976: A compilation of moored current data and associated oceanographic observations, 1973 Internal Wave Experiment (IWEX), Tech. Rep., 75–68, Vol. IX, Woods Hole Oceanogr. Inst., Woods Hole, MA.

MEASUREMENTS OF ERTEL VORTICITY FINESTRUCTURE IN THE EASTERN NORTH ATLANTIC

Eric Kunze
School of Oceanography, University of Washington, WB-10, Seattle, WA 98195

Thomas B. Sanford
Applied Physics Laboratory, HN-10, University of Washington, Seattle, WA 98105

ABSTRACT

Two velocity and temperature profile surveys collected near Ampere Seamount in the eastern North Atlantic reveal ertel vorticity finestructure on horizontal scales of 3-7 km and vertical wavelengths of 40-400 m (Burger numbers of 0.05-30). Just as on basin scales, this finescale ertel vorticity signal is dominated by stretching. The dynamic signal, as characterized by the relative vorticity and horizontal convergence, is consistent with the GM internal wave model and is dominated by near-inertial waves on the resolved scales. The dominance of stretching indicates that ertel vorticity finestructure is associated with passive density finestructure of very low aspect ratio. This suggests that it was not formed recently by flow separation at the seamount but might be an artifact of subduction of surface mixed layers or injection of benthic boundary layers into the pycnocline.

INTRODUCTION

At the large wavelengths that dominate horizontal velocity and vertical displacement in the ocean interior, fluctuations are largely linear, allowing identification of subinertial frequency flows as quasigeostrophic and superinertial motions as internal gravity waves. But at high wavenumbers, advective nonlinearity acts to Doppler shift and alias variance across frequency space, making identification of dynamics from Eulerian frequency impossible; many intrinsic frequencies will contribute to the same Eulerian frequency. Given that the aspect ratios of the dominant finescale motions are $\leq f/N \simeq 0.01$ (Marmorino *et al.*, 1987; Gregg *et al.*, 1986, Itsweire *et al.*, 1989), vertical advection by low-mode internal waves should be the main perpetrator of advective nonlinearity.

Using current-meter and temperature time-series from the IWEX trimooring array, Briscoe (1977) and Müller *et al.* (1978) found finestructure in the internal-wave frequency band with amplitudes of 2 cm s^{-1} and vertical scales of 2 m that contained too much velocity relative to temperature variance to be explained by linear internal wave dynamics. Excess finescale velocity in the continuum frequency band was also reported by Eriksen (1978). Kunze *et al.* (1990) confirmed this result and found excess strain in the near-inertial band. For linear internal waves, the shear-to-strain ratio $V_z/(N\zeta_z) = 1$ at intermediate intrinsic frequencies, $f << \omega_o << N$. As $\omega_o \rightarrow f$, this ratio approaches infinity while as $\omega_o \rightarrow N$, it approaches zero. Müller *et al.* (1978) also found that the vertical coherence scale for velocity was smaller than for temperature. Two sets of dynamics have

been proposed to explain the velocity finestructure: stratified two-dimensional turbulence (vortical mode) (Holloway, 1983; Müller, 1984) and near-inertial waves (Kunze *et al.*, 1990). Near-inertial waves dominate variance at larger scales (e.g., Garrett and Munk, 1979). Kunze *et al.* demonstrated that the finestructure properties were consistent with a finescale internal wave spectrum dominated by near-inertial waves. Resolving the dynamics of the finescale has important implications for stirring and mixing of water properties.

Stratified two-dimensional turbulence is the stratified, nonlinear analog of two-dimensional quasigeostrophic turbulence (Riley *et al.*, 1981; Lilly, 1983; Müller, 1984). Such motion would have subinertial intrinsic frequencies. Müller (1984) referred to stratified two-dimensional turbulence as 'vortical mode' because ertel vorticity anomalies can be associated with it while internal gravity waves can have no ertel vorticity fluctuations associated with them. Following a water parcel, ertel vorticity can only be modified by irreversible processes (Ertel, 1942; Pedlosky, 1978; Haynes and McIntyre, 1986). For example, Lelong and Riley (1991) have shown that wave/wave and wave/vortex interactions do not modify ertel vorticity. Thus, while internal waves can carry energy into the ocean interior, ertel vorticity anomalies are tied to a water parcel. Vortices have been observed in laboratory stratified-wake experiments (Lin and Pao, 1979) and as a decay product in numerical simulations of dissipative turbulence (Riley *et al.*, 1981; Staquet and Riley, 1989; Herring and Metais, 1989). They also appear in numerical simulations of 'forced internal waves' if the forcing projects on ertel vorticity anomalies (Metais and Herring, 1989; Holloway and Ramsden, 1990). Metais and Herring found that a numerically simulated field of random internal waves without ertel vorticity anomalies did not create vortical mode. Because turbulence and mixing are weak and confined to small vertical scales in the ocean interior (Moum and Osborn, 1986; Gregg, 1987; Gregg, 1989; Yamazaki *et al.*, 1990), we anticipate that ertel vorticity anomalies produced there to be likewise weak and confined to small vertical scales where molecular viscosity will rapidly eliminate any dynamic signal. However, considerable finestructure may be generated in regions of strong forcing, such as at boundaries, and survive little affected as it is carried into the ocean interior by subduction. The question of what fractions of finescale variance in the ocean are internal waves and stratified two-dimensional turbulence can only be answered with measurements.

A first attempt to estimate ertel vorticity on the finescale was undertaken by Müller *et al.* (1988). Using the IWEX array velocity measurements, they estimated the relative vorticity directly and stretching from the time-integral of horizontal convergence. The vorticity and convergence spectra were identical. Kunze *et al.* (1990) questioned their claim of having found an ertel vorticity signal on the grounds that (i) like quasigeostrophic flow, vortical mode should have little or no horizontal convergence so their 'stretching' signal was internal wave strain, (ii) near-inertial internal waves have vorticity as well as convergence, and (iii) in a blue spectra, a three-point array will detect 'vorticity' even in a pure convergence field and vice versa. Lien (1990) illustrated that this can explain the large horizontal scales from IWEX. But at radii of 25 m, the IWEX-aliased GM vorticity lies below GM convergence unlike the measurements. On these scales, GM isopycnal slopes tilt vertical shear into vertical vorticity sufficiently to explain the measurements. Thus, the IWEX measurements of velocity finestructure are explicable as internal waves, and Müller *et al.*'s (1988) conclusion that they identified ertel vorticity finestructure is suspect.

Using measurements from a neutrally buoyant float, Kunze *et al.* (1990) showed that, given typical finescale aspect ratios $\leq f/N \simeq 0.01$ (Marmorino *et al.*, 1987; Gregg *et al.*, 1986, Itsweire *et al.*, 1989), the observed finescale shear-to-strain ratio of 2.2 was too high to be explained by stratified two-dimensional turbulence. However, it was consistent with a finescale internal wave spectrum dominated by near-inertial waves. Indeed, the observed shear-to-strain ratio is only slightly above the GM value of 1.7 (Munk, 1981). The semi-Lagrangian strain spectra of Sherman and Pinkel (1991) supports the interpretation that most shear variance at high wavenumber is Doppler-smeared near-inertial motions rather than vortical mode. In contrast, D'Asaro and Morehead (1991) found that the incoherent signal at 40-m vertical wavelength in velocity profile surveys had a potential-to-horizontal kinetic energy ratio (*PE/HKE*) and Burger number ($N^2k_H^2/f^2k_z^2$) consistent with vortical mode and not internal wave dynamics (Fig. 1). These measurements were collected under the ice in the Beaufort Sea. The incoherent (vortical) signal contained as much energy as the high-frequency internal waves in the weak internal wave regime under the ice. None of these measurements were able to estimate ertel vorticity.

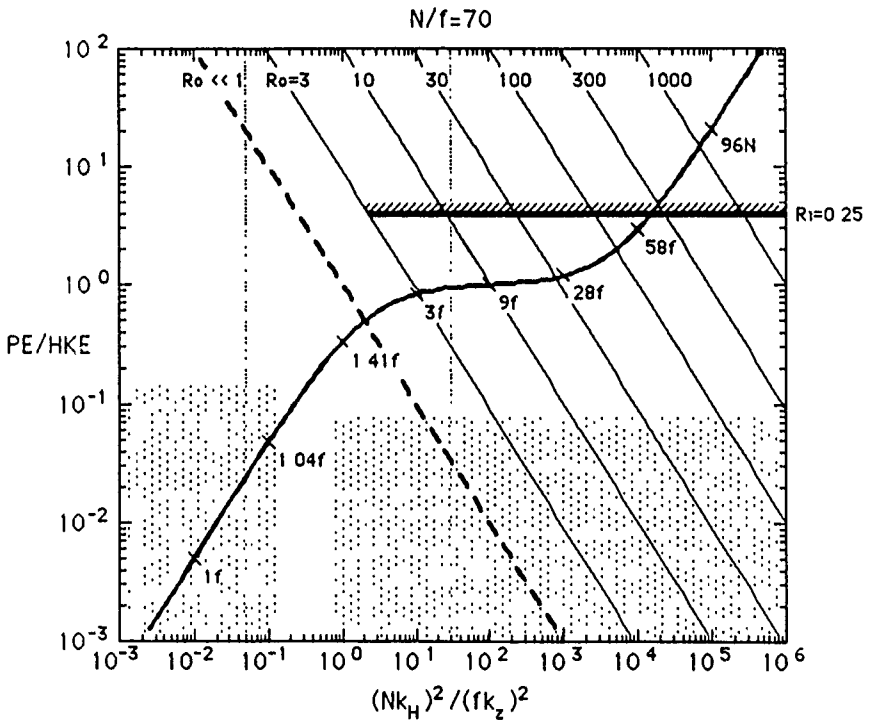


Figure 1. A dynamic diagram of the ratio of potential-to-horizontal kinetic energy *PE/HKE* versus Burger number $(N^2k_H^2)/(f^2k_z^2)$ from D'Asaro and Morehead (1991). The *S*-curve is the consistency relation (marked with intrinsic frequencies) for internal waves with $N/f = 70$ [$PE/HKE = B/(B+2)$]. The thick-dashed diagonal corresponds to geostrophy [$PE/HKE = 1/B$] and the thick solid diagonals to high Rossby number vortical mode [$PE/HKE = Ro^2/B = 1/Ri$]. The two blocks of stippling correspond to the coherent signal (left), which is in the near-inertial wave part of the diagram, and the incoherent signal (right), which is in the geostrophic/vortical regime, from D'Asaro and Morehead's XCP surveys in the Beaufort Sea. The two vertical dotted lines bound the Burger numbers resolved by the Ampere Seamount surveys.

In September 1988, an expedition was made to Ampere Seamount at 35° 03.5' N, 12° 52.2' W in the eastern North Atlantic (Fig. 2) to look for evidence of ertel vorticity finestructure which would be a signature of stratified two-dimensional turbulence. The Coriolis frequency at this latitude is 8.3×10^{-5} s $^{-1}$ It was thought that the likelihood of finding this signal would be greater in the vicinity of a seamount where flow separation might lead to shed eddies in the wake. However, as no finescale ertel vorticity measurements have been made in the ocean, this was not crucial. Eddy-shedding did not appear to be playing a role in the data we collected.

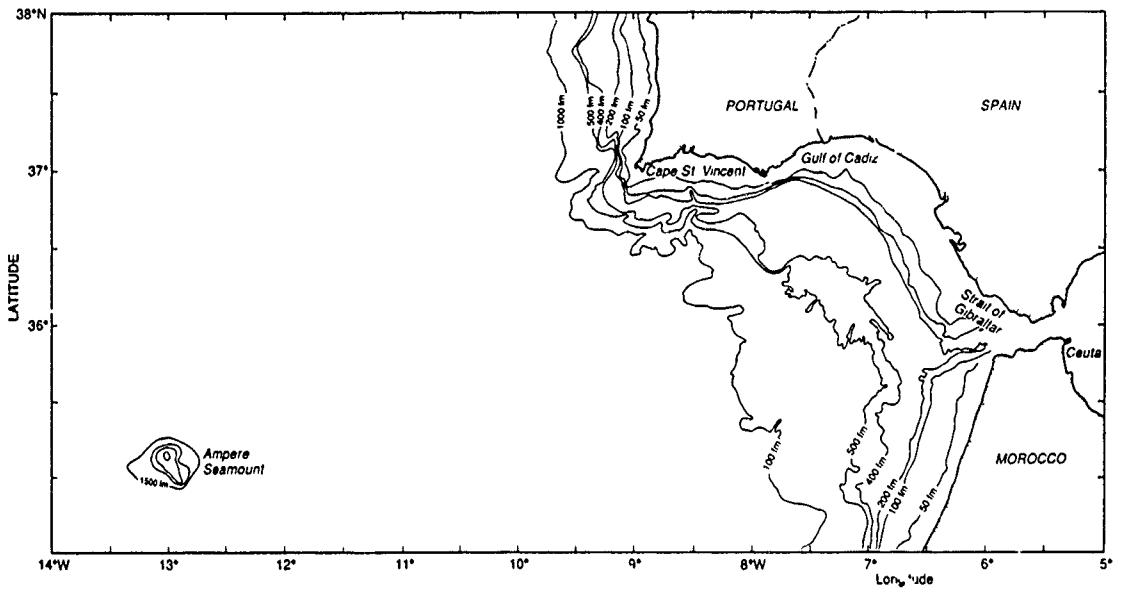


Figure 2. Bathymetry in the eastern North Atlantic showing the location of Ampere Seamount at 35° 04'N, 12° 52'W outside the mouth of the Gulf of Cadiz The summit comes within 53 m of the surface.

MEASUREMENTS

We first undertook an XBT box survey 60 km on a side around the seamount (Fig. 3) to look for mesoscale impinging geostrophic flow. No signal was apparent. Four drifters were then deployed within 10 km of the summit to examine the near-field mean flow. Three of these drifters were recovered and redeployed on the eastern flank for a total of seven drifter tracks. The drifters were drogued to 100-200 m using a high-drag line in that depth range and a low-drag line to the surface (Drever and Kennelly, 1991). Radar reflectors were used to locate the drifters. The ship came alongside each drifter to determine its range and bearing to a radar beacon moored on the summit (Fig. 3). Range is accurate to ±200 m and bearing to ±2°. Positioning based on Loran C was good to one kilometer in x and y. The drifters were tracked from 6-10 September 1988. On the northern flank of the seamount and to the east in an interior boundary layer, drifters 1 and 2 moved persistently toward the east with average speeds of 5-7 cm s^{-1} (Fig. 4). Tidal fluctuations of ±5 cm s^{-1} were also primarily in the u velocity component. On the eastern flank, motion was

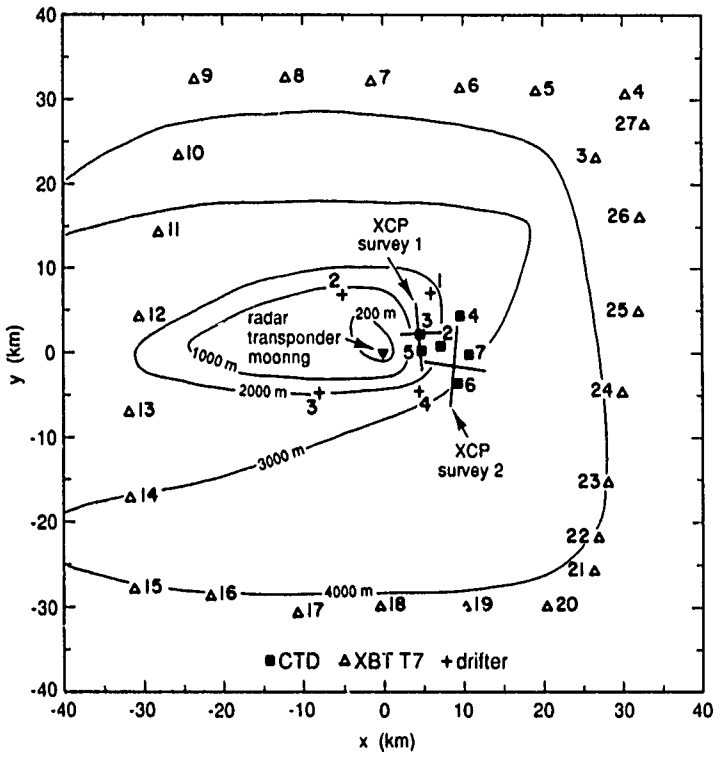


Figure 3. Sampling in the neighborhood of Ampere Seamount including an XBT box survey (Δ) around the seamount, drifter deployments (+) tagged to 100-200 m depth, six CTD casts (❑) and two cross-shaped XCP surveys on the eastern flank. The XCP surveys contained 27 and 28 profiles at spacings of 0.3-1 km. A radar transponder navigation mooring was located on the summit (∇).

weak and random for the first three days, then accelerated to the south during 9 September, reaching speeds of ~10 cm s^{-1}. Drifter 3, deployed southwest of the summit, moved southward at ~13 cm s^{-1}. However, on recovery it was discovered to have lost its drogue so its track is not representative of the flow at 100-200 m depth. The last drifter was recovered in the middle of 9 September.

Based on the lateral shear between drifters on the northern and eastern flanks of Ampere Seamount during the first three days of the drifter deployment (Fig. 5), which correspond to vorticities ~0.5f, it was thought that flow separation was most likely to occur in the northeast corner of the summit. Eddies generated by this mechanism would be found east and north of the summit so two cross-shaped finescale expendable current profiler (XCP) surveys were conducted near the northeast flank on the afternoons of 8 and 9 September (Fig. 3). Six CTD casts to 2000-m depth were made here between the XCP surveys to determine the local T,σ–relation. The XCP surveys contained 28 and 27 probes with 12 to 14 XCPs in each leg and drop spacings of 0.3-1 km. Survey 1 was effectively 4 km in diameter and survey 2 was 7 km across. Each survey was completed in three hours in an effort to minimize biasing of the spatial gradients by temporal variability. Because these horizontal scales are dominated by near-inertial waves, this appears to have been effective. Vertical wavenumber spectra of profile sums and differences near the intersection of the legs suggest that temporal variability has little impact on wavelengths greater than 100 m.

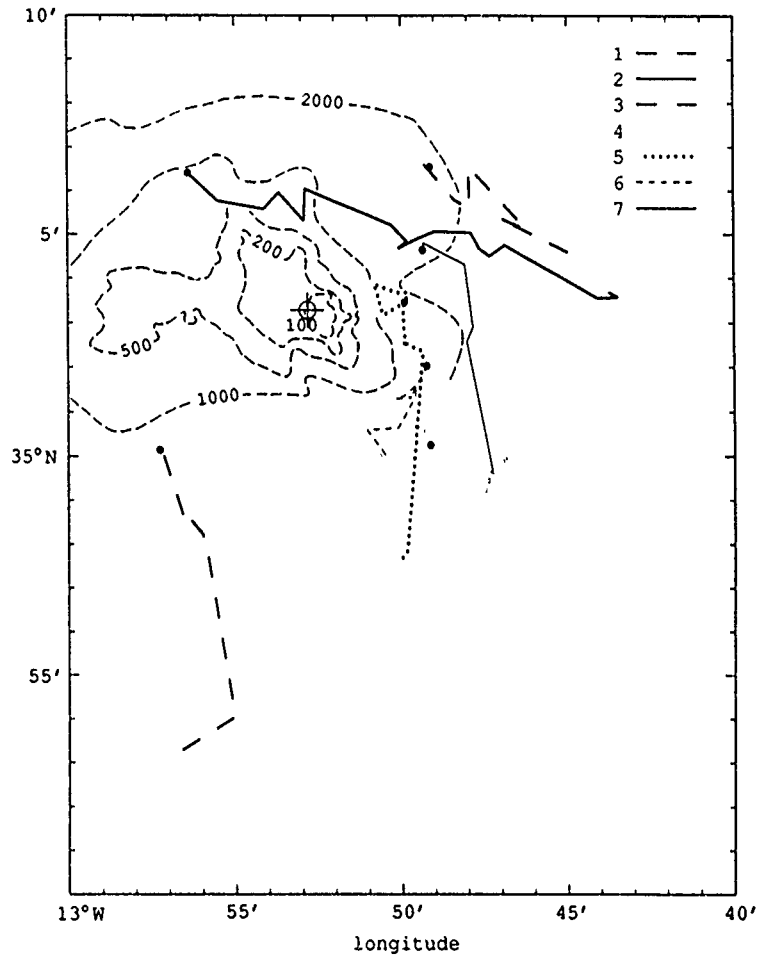


Figure 4. The seven drifter tracks span 6 to 10 September superimposed on Ampere Seamount bathymetry. Drifters on the northern flank of Ampere moved eastward at 5-7 cm s^{-1}, extending off the flank in an interior boundary layer to the east. Those on the eastern flank moved little for the first three days, then moved southward at speeds exceeding 10 cm s^{-1}. The drifter to the southwest moved south at 13 cm s^{-1} but it lost its drogue.

Expendable current profilers electromagnetically measure the horizontal velocity (u , v) relative to a depth-independent constant (e.g., measures the baroclinic flow) and measure temperature T with an XBT thermistor. Data values are recorded every 0.3 m from the surface to ~1600-m depth. The upper 50 m is generally contaminated by temporally aliased velocities associated with the surface swell. In the pycnocline, the oceanic signal typically falls below the instrumental noise level for wavelengths smaller than 10 m. Therefore, standard pre-analysis XCP processing smooths the raw data with a 6-m triangular window every 3 m. The velocities have rms errors of ±0.4 cm s^{-1} and temperature errors of ±0.15 °C. We will be using these data to compute vertical and horizontal gradients. The velocity error corresponds to shear errors of $\pm 10^{-4}$ s^{-1} ~ $0.03N$ over 40 m (vertical) and $\pm 10^{-6}$ s^{-1} ~ $0.013f$ over 4 km (horizontal). The temperature error corresponds to $\pm 4 \times 10^{-3}$ °C m^{-1} $\equiv 4 \times 10^{-6}$ s^{-2} ~ $0.3N^2$ over 40 m and $\pm 4 \times 10^{-5}$ °C m^{-1} $\equiv$ 4×10^{-8} s^{-2} ~$0.15fN$ over 4 km given $\partial b/\partial T \sim 10^{-3}$ m s^{-2} $°C^{-1}$. Thus, with comparison to f and N, the velocity contributes smaller errors to an ertel vorticity signal (see below) than temperature. Assuming resolved vertical scales of 40-400 m and horizontal scales of 3-7 km , the range of resolved Burger number is 0.05-30 (between the dotted vertical lines in Fig. 1). For internal wave dynamics, this range includes frequencies up to $3f$. For vortical dynamics, Rossby numbers up to ten are in principle measurable.

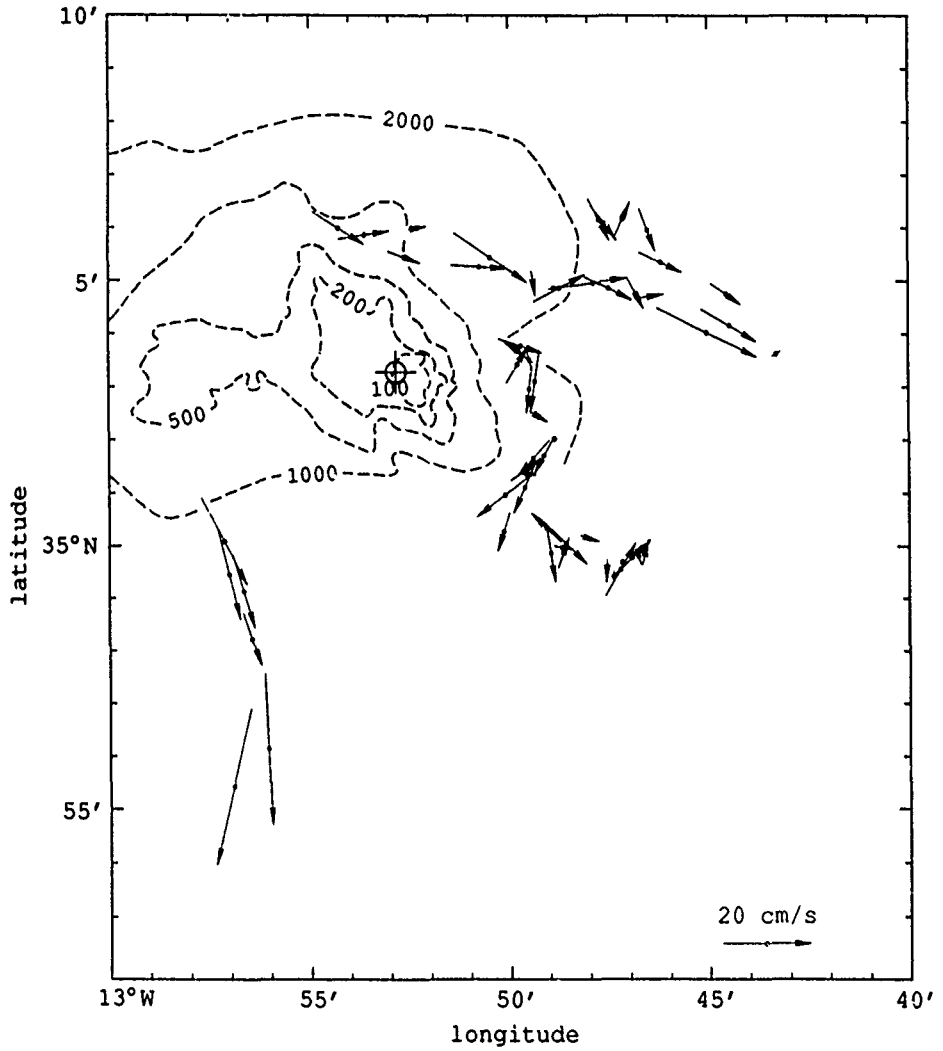


Figure 5. Drifter velocities around Ampere Seamount for the period prior to the acceleration to the south on the eastern flank.

ERTEL VORTICITY

Review

Following a water parcel, the ertel vorticity cannot be modified except by irreversible processes: external forcing, dissipative turbulence and mixing (Ertel, 1942; Pedlosky, 1978). Thus, no ertel vorticity fluctuations are associated with nondissipative internal waves. For linear internal waves, the linear stretching and relative vorticity contributions to ertel vorticity are equal and opposite. At high wavenumbers where nonlinearity becomes important, nonlinear contributions of the ertel vorticity may contribute, but no total ertel vorticity anomalies arise in the absence of internal wave dissipation (Lelong and Riley, 1991).

The ertel vorticity will be defined here as the dot product of the total vorticity ($2\Omega + \nabla \times V$) and the gradient of the bucyancy $\nabla(\bar{B} + b)$

$$\Pi = (2\Omega + \nabla \times V)\cdot\nabla(\bar{B} + b) \tag{1}$$

where Earth's rotation, $2\Omega = (0, 2\Omega\cos(lat°), 2\Omega\sin(lat°)) = (0, f\cot(lat°), f)$, and the buoyancy $\bar{B} + b = -g\delta\rho/\rho_0$ has been split into a largescale background $\bar{B}(z)$ varying only with depth and smaller scale anomalies $b(x,y,z,t)$. The splitting of the vertical buoyancy gradient is not necessary but is common practice and a convenient way of separating large and small scale s. Neglecting the meridional component of Earth's rotation and the vertical velocity w, Eq. (1) can be expanded into a background, linear and nonlinear part

$$\Pi = f\bar{N}^2 + fb_z + (v_x - u_y)\ \bar{N}^2 + (v_x - u_y)b_z - b_x v_z + b_y u_z \tag{2}$$

(i) (ii) (iii) (iv) (v) (vi)

where subscripts denote derivatives and $\bar{N}^2 = \partial\bar{B}/\partial z$ is the background buoyancy frequency. The background ertel vorticity (i) involves only the smoothed stratification $\bar{B}(z)$ and the planetary vorticity f. The linear ertel vorticity anomaly is made up of buoyancy-gradient anomalies (ii) and relative vorticity (iii). Nonlinear ertel vorticity anomalies arise from coupling of buoyancy-gradient anomalies and relative vorticity (iv) and the twisting terms (v) and (vi).

Data

Estimating the different components of (2) requires determination of both horizontal and vertical gradients of u, v and b. This is achievable with cross-shaped XCP surveys provided that the surveys are coherent so that (i) horizontal gradients can be estimated and (ii) temperature can be used as a proxy for buoyancy.

To demonstrate that the survey was coherent, Fig. 6 displays the eastwest section of east velocity u from survey 2. Fluctuations of ±10 cm s^{-1} are seen on scales of tens to hundreds of meters with vertical structure down to scales of 50 m coherent across much of the transect. Vertical least-squares cubic fits to the profiles (dashed lines) reveal little variance associated with vertical wavelengths greater than 1000 m. As wavelengths ~1000 m can be affected by the barotropic offset when profiles are not of identical length and our main interest here is in the finescale ($\lesssim$ 100 m), the cubic fits were removed before subsequent analysis. Because the profiles are horizontally coherent, we are justified in estimating the horizontal gradients $\partial/\partial x$ from the eastwest legs and $\partial/\partial y$ from the northsouth legs using least-squares fits. These are then combined to determine dynamically relevant variables. Figure 7a displays the relative vorticity $(v_x - u_y)/f$ for survey 1 as a function of depth z and the horizontal scale over which the fit was performed Δr. As the fitting scale Δr increases, so does the number of profiles going into the fit (number at bottom of each profile) while the standard deviation (stippling), as determined using a bootstrap technique, decreases. At Δr=4 km , the relative vorticity has significant fluctuations on scales of tens to

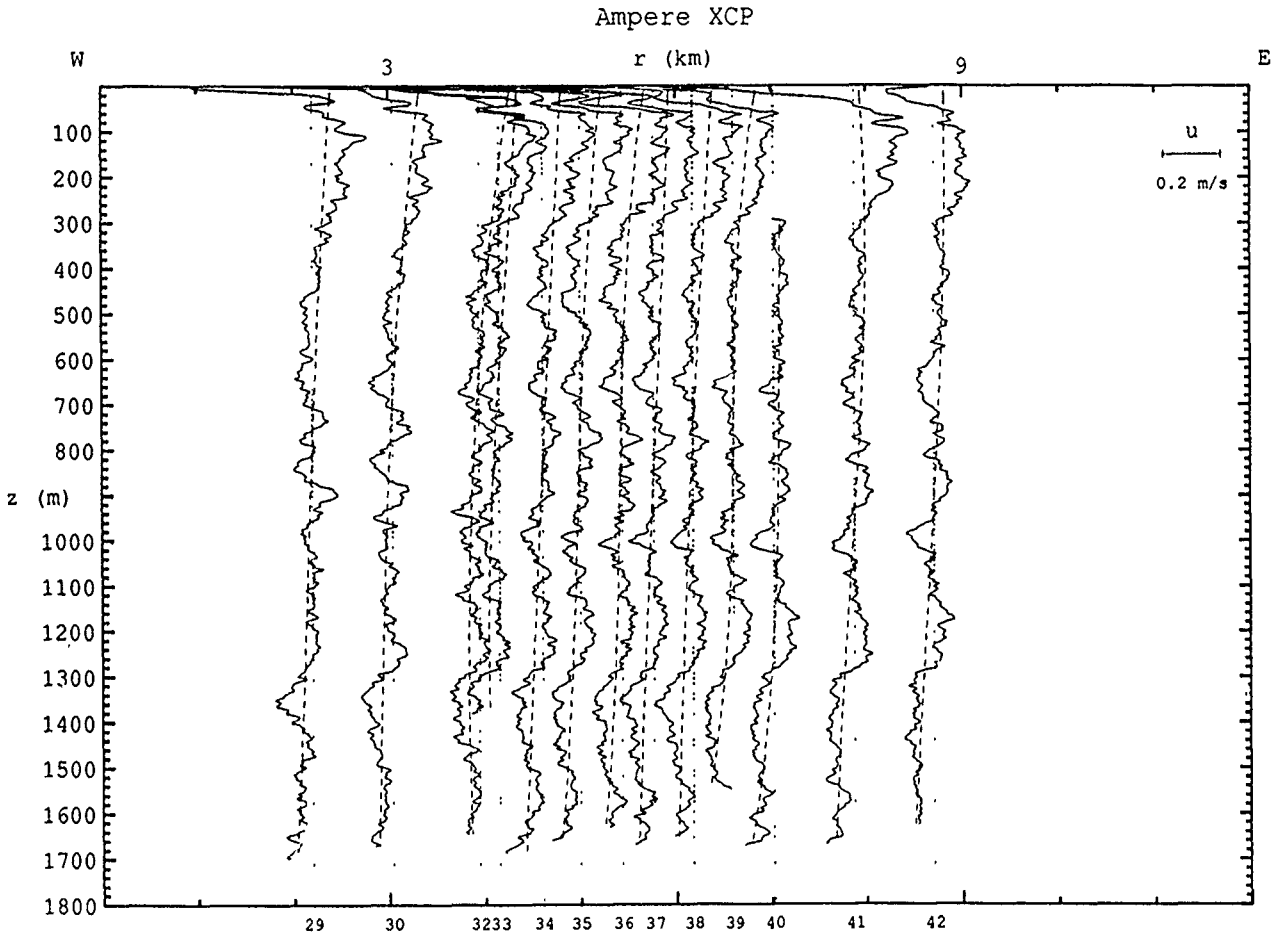


Figure 6. A sample section of u velocity profiles from the east-west leg of XCP survey 2. The dashed lines indicate cubic fits to the profiles. Fluctuations coherent over much of the transect are evident down to scales of tens of meters.

hundreds of meters with rms values of 0.2-0.3f, an order of magnitude above instrument noise. Unresolved oceanic variance is ~0.1f. The horizontal convergence $(u_x + v_y)/f$ for survey 1 (Fig. 7b) has similar characteristics. Inertially rotating the velocity profiles to a common time before carrying out the fits did not alter the details of the vorticity and convergence profiles appreciably.

Figure 8 displays the T,σ–relation from the CTD casts collected near the seamount. The relation is tight for temperatures $T < 10$ °C (depths greater than 1250 m) and for 12 °C $< T <$ 16 °C (150-600 m depth) but not in the salt-stratified isothermal layer between 600 and 1200 m, nor in the upper 150 m where water-mass gradients are evident. In this paper we will confine our attention to 150-600 m depth and use $b(T) = T \cdot \partial b/\partial T$ based on least-squares fits to the CTD data to relate buoyancy to temperature. For 150-600 m depth, $\partial b/\partial T \simeq 10^{-3}$ m s^{-2} °C^{-1} and the buoyancy frequency $\bar{N} = 3.3\text{-}3.8 \times 10^{-3}$ s^{-1}.

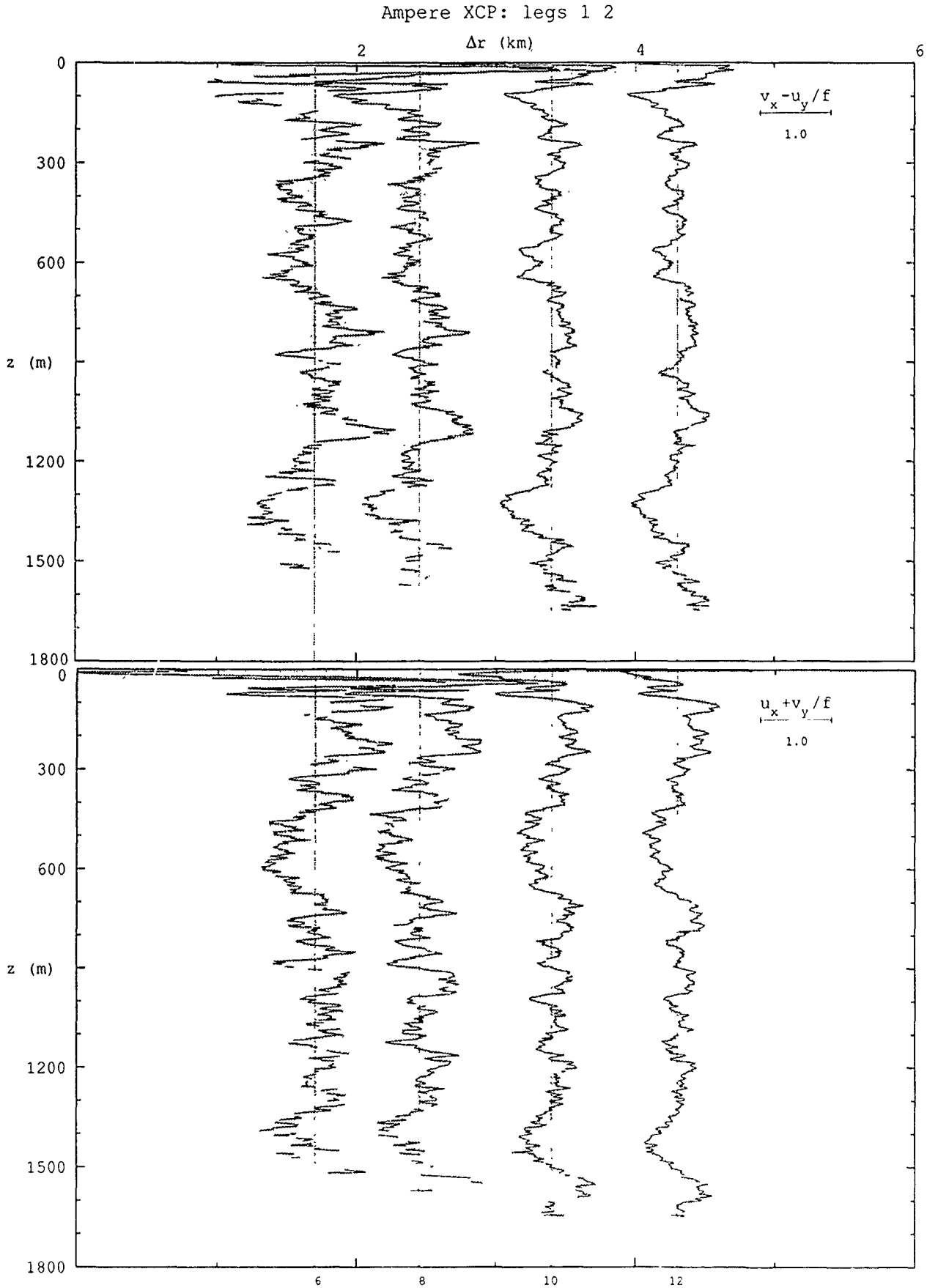

Ampere XCP: legs 1 2
Δr (km)
2
4
6
0
300
600
z (m)
1200
1500
1800
$v_x - u_y / f$
1.0
$u_x + v_y / f$
1.0
6
8
10
12

Figure 7a. (Facing page, top) Survey 1 relative vorticity $(v_x - u_y)/f$ from least-squares horizontal fits as function of depth z and the fitting scale Δr. As the fitting scale increases, the number of drops in the fit increases (numbers at bottom of each profile) and the standard deviation (stippling) diminishes. For a fitting scale $\Delta r = 4$ km , significant fluctuations occur on scales greater than 50 m with rms values of 0.2-0.3f.

Figure 7b. (Facing page, bottom) As in Fig. 7a but for horizontal convergence $(u_x + v_y)/f$. Magnitudes and scales are comparable to those of relative vorticity.

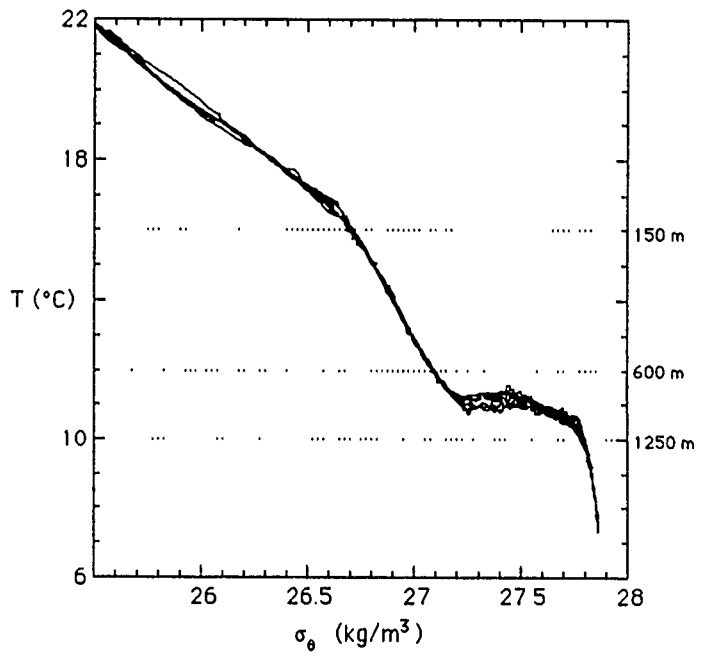


Figure 8. Temperature T versus density σ_t from the six CTDs collected east of Ampere. For temperatures less than 10 °C (depths greater than 1250 m) and between 12-16 °C (150-600 m), the T,σ-relation is tight, allowing temperature to be used as a proxy for buoyancy (density). A salt-stratified isothermal layer occupies the depth range 650-1200 m and water-mass gradients are evident in the upper 150 m.

Spectra

The sum of the five linear and nonlinear terms of the ertel vorticity (2) reveals that there is ertel vorticity finestructure present between 150- and 600-m depth. The nonlinear terms contribute little to the signal, that is, on the scales of the survey, the ertel vorticity is linear. The stretching term dominates. These results are most clearly seen in vertical wavenumber spectra of the ertel vorticity terms (Fig. 9). The spectra are presented in both log-log and variance-preserving forms. Figure 9a displays the ertel vorticity component spectra for survey 1 and a fitting scale Δr=4 km and Fig. 9b for survey 2 and Δr=7 km. The stretching fb_z (thick solid) clearly dominates the variance, peaking at 50-100 m wavelength. This scale is at least in part due to the horizontal averaging over the fitting scale which smooths over higher wavenumbers. The nonlinear terms are at least an order of magnitude smaller at all wavenumbers. The other linear term, the vorticity $\zeta \bar{N}^2$, is an order of magnitude smaller than the stretching at 50-100 m wavelength but comparable in magnitude at lower wavenumbers. It exceeds stretching in the lowest wavenumber band in the second survey but this is a consequence of the cubic-fit separation of background $\bar{N}^2$ and anomaly b_z which

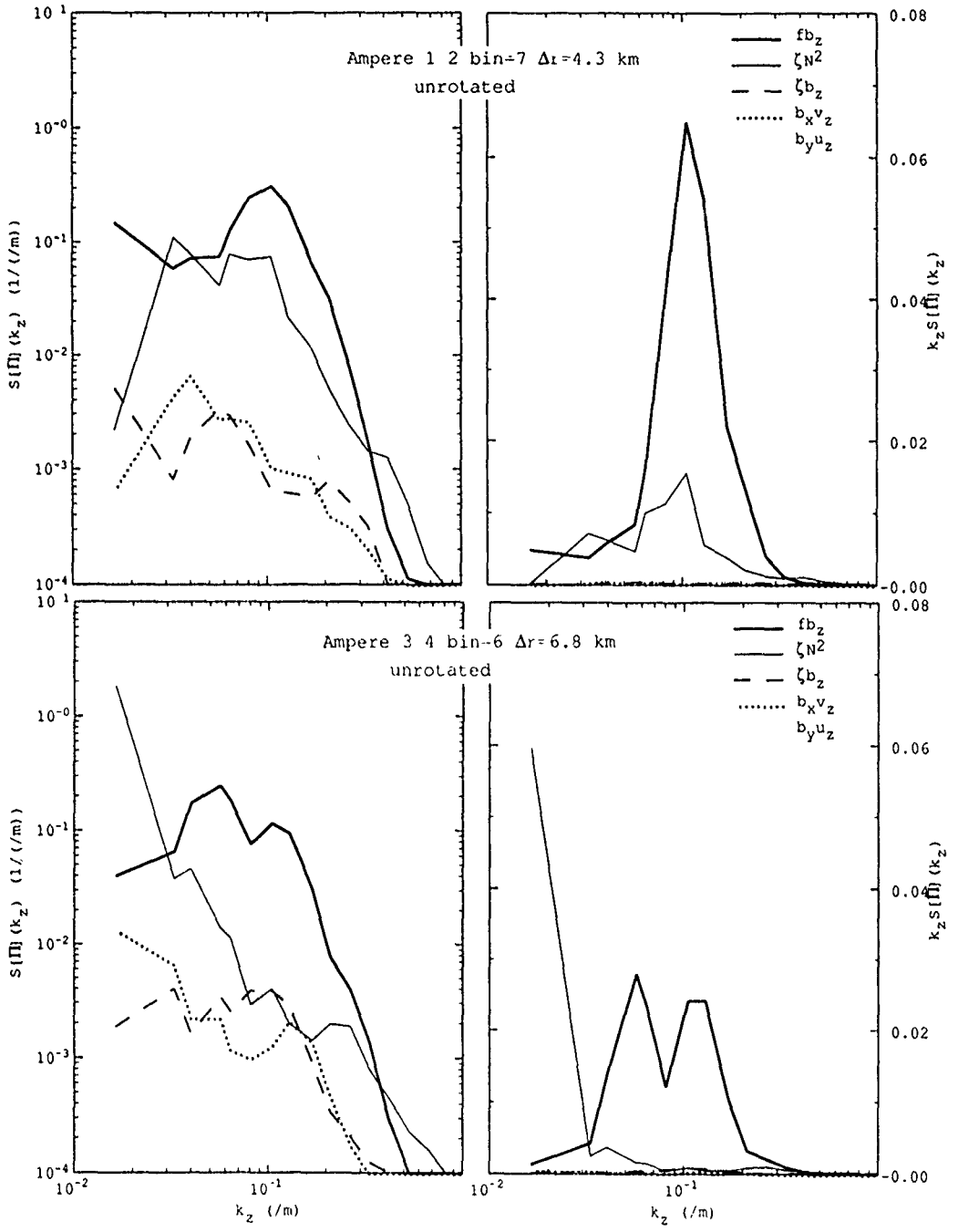


Figure 9. (a, top) Vertical wavenumber spectra from survey 1 with horizontal fitting scale $\Delta r = 4$ km for the linear and nonlinear terms in the ertel vorticity (2) including stretching fb_z, vorticity $(v_x - u_y)\,\bar{N}^2$, $(v_x - u_y)b_z$, and twisting $b_x v_z$ and $-b_y u_z$ terms. The spectra are presented in both log-log (left) and variance-preserving (right) forms. The stretching peak at 60-m wavelength contains an order of magnitude more variance than the other terms. At lower wavenumbers ($\lambda_z > 100$ m), vorticity and stretching become comparable. At higher wavenumbers ($\lambda_z < 20$ m) vorticity slightly exceeds stretching due to the 12-m scale over which vertical gradients were computed. (b, bottom) As in Fig. 9a but for survey 2 with 7-km fitting. The stretching excess peak ranges from $\lambda_z = 30$-160 m. Vorticity exceeds stretching at the lowest resolved wavenumber as a consequence of the cubic fit used to separate $\bar{N}^2$ and fb_z.

assigns all low wavenumber variance to the background. At high wavenumbers, vorticity exceeds stretching because of the 12-m scale over which first-differencing to estimate b_z was performed.

Comparing the measured stretching (strain) and vorticity spectra with the GM model (Figs. 10 and 11) reveals that, while the stretching greatly exceeds the GM model in the 50-100 m wavelength band, the vorticity spectra are comparable to GM (for this comparison, the high horizontal wavenumbers that the array does not resolve have been eliminated from the GM model with a simple boxcar lowpass filter with a cutoff $k_H = \pi/\Delta r$). In survey 1, the vorticity has a weak peak at 50-100 m but is generally similar in shape and level to GM. In survey 2, the vorticity lies on the GM model except at the largest resolved vertical wavelength (λ_z = 384 m). Thus, while stretching greatly exceeds vorticity indicating the presence of vortical mode, the dynamic (velocity) signal differs little from the GM model. This suggests that the velocity signal could be predominantly due to internal waves. This can be tested by comparing the measured vorticity and convergence fields with the filtered GM model (Fig. 12). For survey 1 and Δr= 4 km , the observations share with the GM model the tendency for (i) convergence to exceed vorticity at low vertical wavenumbers, (ii) convergence and vorticity to be nearly identical at high vertical wavenumbers, and (iii) convergence to peak at slightly lower wavenumbers than vorticity. The same cannot be said of survey 2 where vorticity slightly exceeds convergence at most wavenumbers. This is *not* consistent with internal-wave dynamics for which $(\nabla_H \cdot V) \geq (\nabla_H \times V)$. We caution that the difference between the two spectra is small and may not be significant. If this is the case, the survey 2 spectra suggest a larger relative contribution from near-inertial waves [for which $(\nabla_H \cdot V) \simeq (\nabla_H \times V)$] than the GM model and survey 1.

For survey 1, for which the convergence exceeds the vorticity at most wavenumbers, consistent with internal-wave dynamics, the ratio of these two quantities can be used to estimate the intrinsic frequency. For $\omega_O << N$, assuming a single dominant wave, the intrinsic frequency depends on the ratio of the convergence ∇ to the vorticity ζ

$$\omega_O(k_z) = f \frac{\nabla(k_z)}{\zeta(k_z)} \quad . \tag{3}$$

Because of the similar treatment of vorticity and convergence, this estimate is more reliable than other intercomparisons that could be made. The intrinsic frequency is displayed as a function of vertical wavenumber in Fig. 13 for Δr=4 km. Where it falls below ω_O, the model assumptions fail. For the variance-containing lower wavenumbers, the intrinsic frequencies lie between 1.2f and 3f.

SUMMARY AND CONCLUSIONS

Measurements were made in the vicinity of a seamount to look for ertel vorticity finestructure. Since the ertel vorticity of a fluid parcel can only be modified by irreversible processes, no ertel vorticity fluctuations are associated with internal gravity waves. Thus, finding ertel vorticity finestructure would be evidence of non-internal wave dynamics on the finescale, that is, stratified two-dimensional turbulence aka vortical mode (Riley *et al*, 1981; Lilly, 1983; Müller, 1984). We

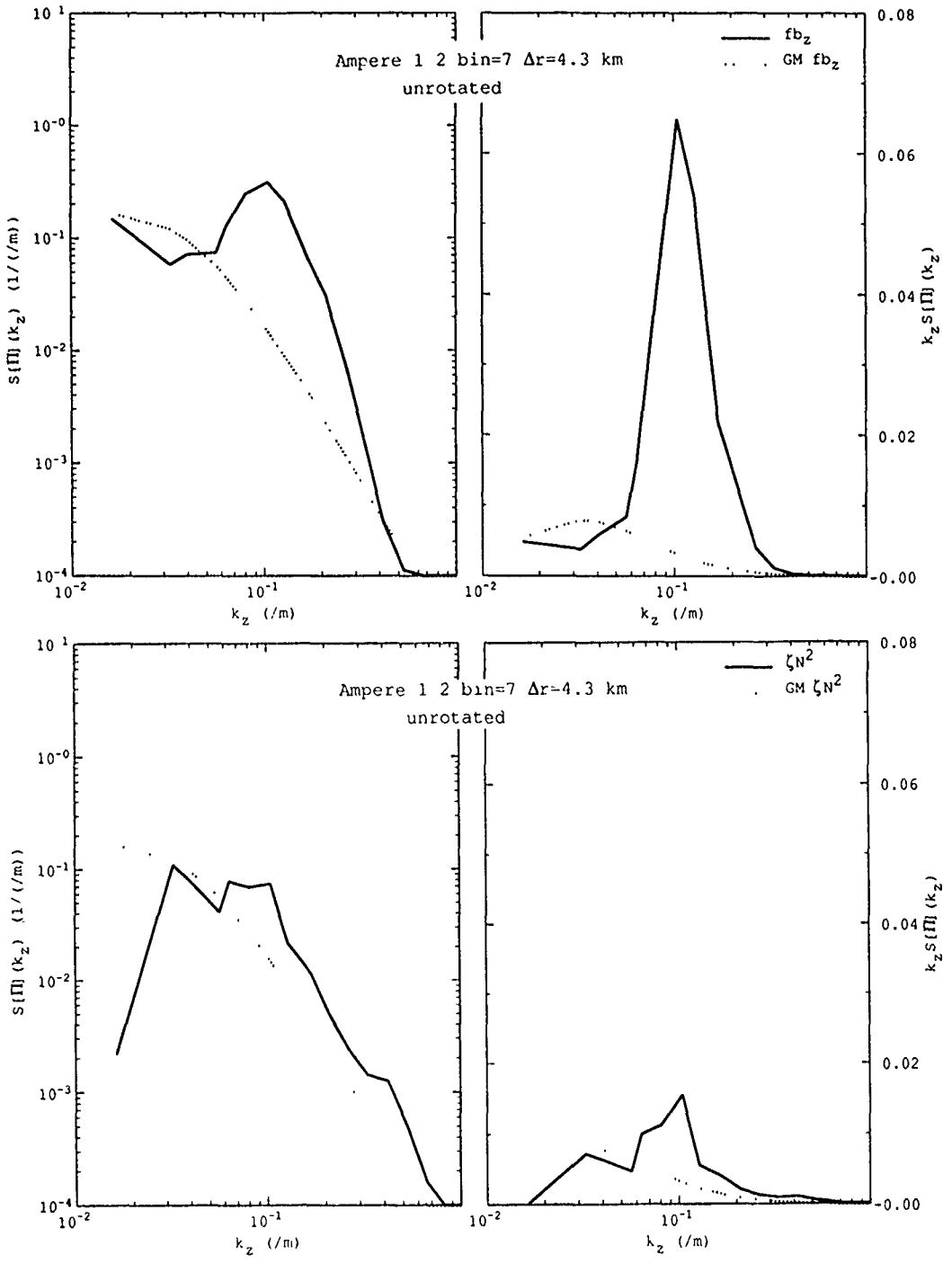


Figure 10. (a, top). Comparison of the vertical wavenumber spectrum for 4- km averaged stretching fb_z from survey 1 with that from the GM internal wave model filtered to remove unresolved high horizontal wavenumbers ($k_H > \pi/\Delta r$). The observed stretching is higher than GM for $6\times10^{-2}\text{m}^{-1}<k_z<2.5\times10^{-1}\ \text{m}^{-1}$. (b, bottom) Comparison of the vertical wavenumber spectrum for the 4-km fit vorticity $(v_x - u_y)\,\bar{N}^2$ from survey 1 with the GM internal wave model. Levels are comparable.

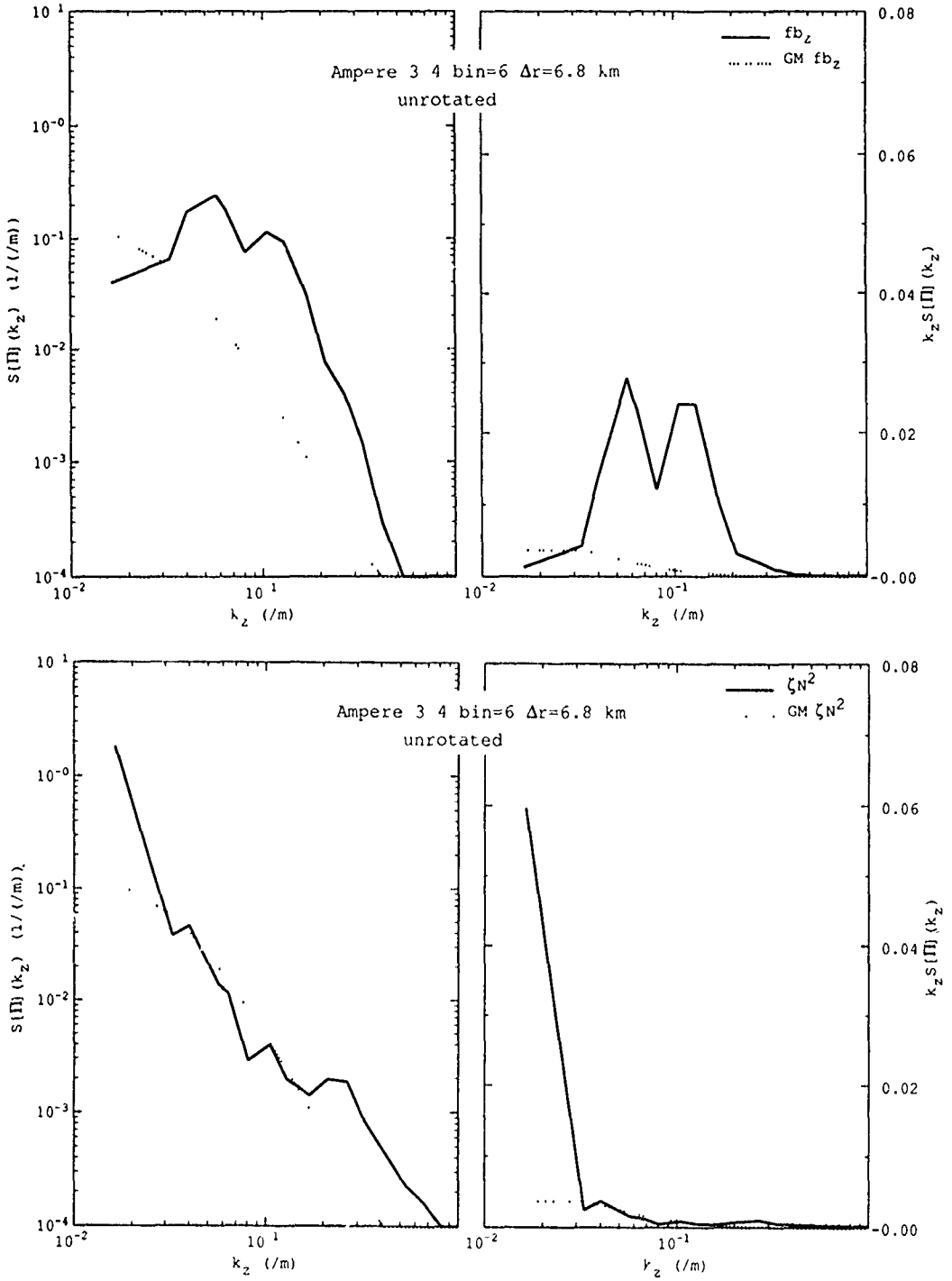
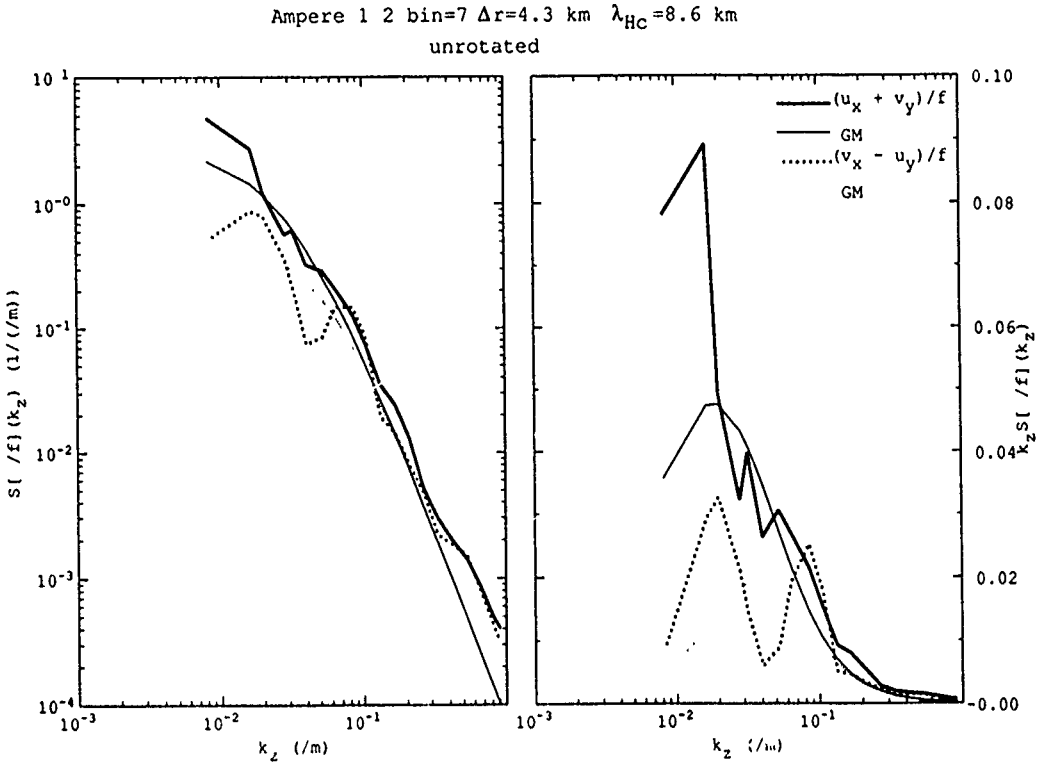


Figure 11. (a, top). As in Fig. 10a but for a 7-km fitting scale on survey 2. Again, observed stretching exceeds GM. (b, bottom) As in Fig. 10b but for a 7-km fitting scale on survey 2. Observed vorticity exceeds GM at the largest resolved wavelength (λ_z = 384 m).


Ampere 1 2 bin=7 Δr=4.3 km λHc=8.6 km
unrotated
(ux + vy)/f
GM
(vx - uy)/f
GM
S(/f)(kz) (1/(/m))
kzS(/f)(kz)
kz (/m)
kz (/m)

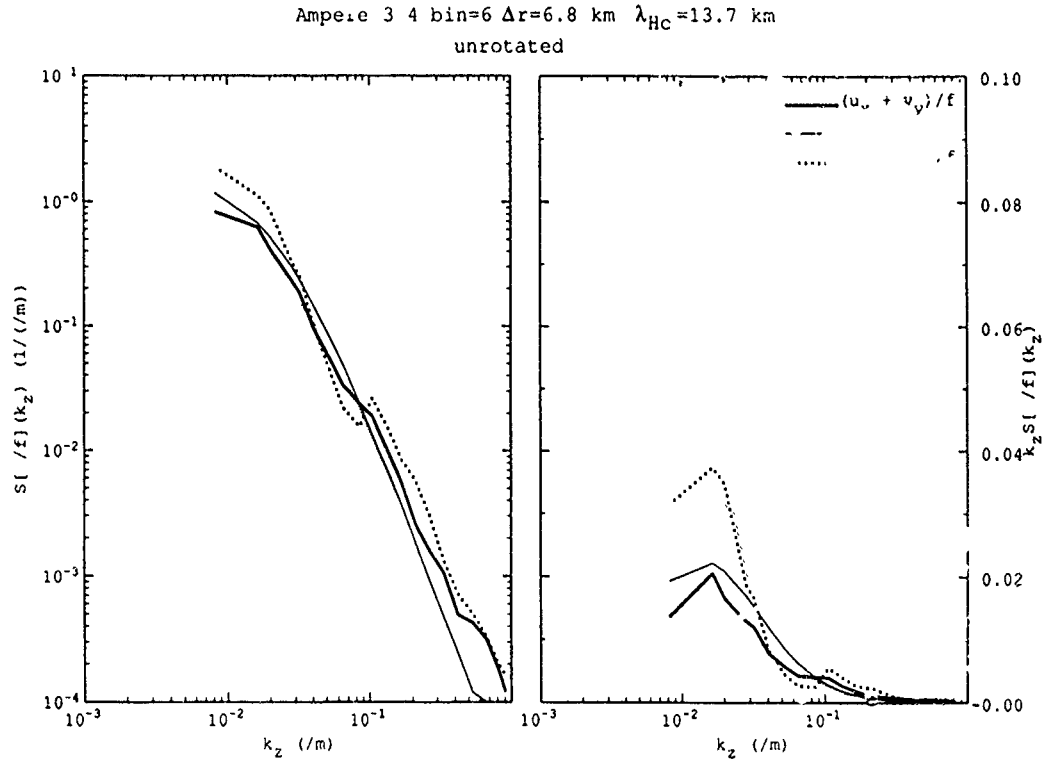

Ampere 3 4 bin=6 Δr=6.8 km λHc=13.7 km
unrotated
(ux + vy)/f
S(/f)(kz) (1/(/m))
kzS(/f)(kz)
kz (/m)
kz (/m)

Figure 12a. (facing page, top). Comparison of the observed vertical wavenumber spectra for horizontal convergence and relative vorticity from survey 1 (thick curves) with the GM model integrated to horizontal wavenumbers of $\pi/8.6$ km $^{-1}$ (thin curves). This upper integration bound is consistent with the fitting scale of the horizontal gradients which filters out higher wavenumbers. Like the GM model, the observed convergence exceeds vorticity at low wavenumbers and is comparable at high wavenumbers. Also, the vorticity peaks at slightly higher wavenumbers than convergence in both measurements and model. This suggests that the dynamic (velocity) signal is dominated by internal waves.

Figure 12b. (facing page, bottom) As in Fig. 12a but for survey 2 and the GM model integrated to horizontal wavenumbers of $\pi/13.7$ km $^{-1}$. In this case, the observed vorticity slightly exceeds convergence at most wavenumbers, inconsistent with linear internal wave dynamics. The difference is small and may not be significant.

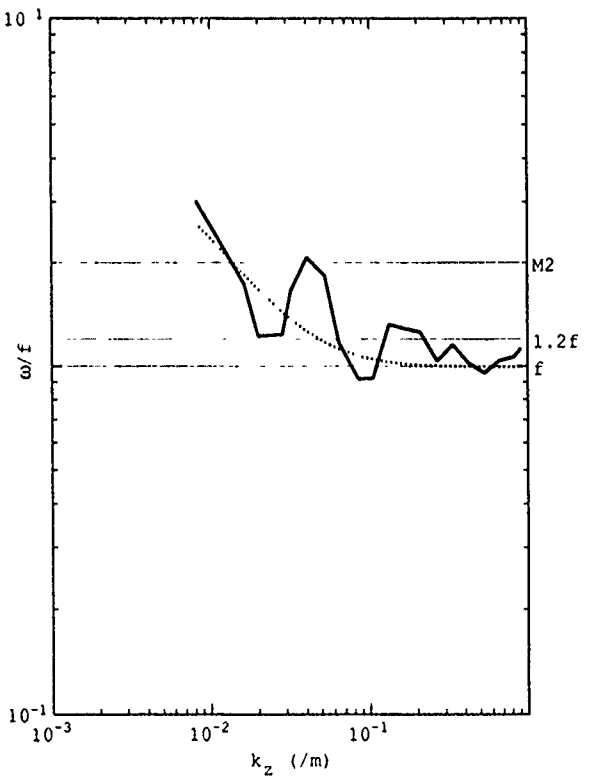


Figure 13. Intrinsic frequency versus vertical wavenumber as inferred from the ratio of horizontal convergence to relative vorticity (3) in survey 1 with Δr=4.3 km (thick solid) and the ratio of convergence to vorticity for the filtered GM model (dotted). This assumes a single internal wave dominates. Values below f are inconsistent with (3). At variance-containing wavenumbers, values lie between 1.2-2.5f, consistent with the filtered GM model.

caution that the reverse is not true. An absence of ertel vorticity finestructure does not preclude the existence of stratified two-dimensional turbulence because vortices can have compensating contributions from the different terms (2) making up ertel vorticity (McWilliams, personal communication, 1991).

The velocity and temperature profile surveys described here reveal ertel vorticity finestructure on vertical wavelengths of 40-400 m and horizontal scales of 3-7 km near a seamount. As on basin scales, these scales are dominated by stretching fb_z. Thus, away from strong eddies, stretching can

be used to estimate ertel vorticity down to scales of a few kilometers. The implied high potential to horizontal kinetic energy ratio and low aspect ratios (see Fig. 1) are reminiscent of the density finestructure, aka pancake eddies or blini, discussed in the internal-wave literature in the early 70's (Phillips, 1971; Garrett and Munk, 1971; McKean, 1974; Eriksen, 1978; Levine and Irish, 1981). The stretching spectra peaks at ~60-m wavelength.

The dynamic (velocity) signal contains roughly equal amounts of relative vorticity $(v_x - u_y)/f$ and horizontal convergence $(u_x + v_y)/f$. For comparison, high-frequency internal waves have greater convergence than vorticity, near-inertial waves comparable convergence and vorticity, and stratified two-dimensional turbulence greater vorticity than convergence. Thus, the velocity signal is most consistent with near-inertial internal waves. This is borne out by the close agreement between the observed and GM vertical wavenumber spectra for these quantities at least in survey 1. With the exception of the $b_y u_z$ in survey 1 (Fig. 9a), the nonlinear vorticity and twisting contributions to the ertel vorticity, $(v_x - u_y)b_z$, $b_x v_z$ and $b_y u_z$, differ little from zero.

The velocity finescale being dominated by near-inertial waves is consistent with the findings of Kunze *et al.* (1990) and Sherman and Pinkel (1991). The dominance of the stretching term fb_z in the ertel vorticity indicates that the observed anomalies are very nearly passive and geostrophic. This suggests that they were not recently formed by flow separation around Ampere Seamount since, in the near field, a wake of shed eddies should have a strong dynamic signal; flow separation at topography has been proposed as a generation mechanism for Arctic eddies (D'Asaro, 1988) and Meddies (Prater and Sanford, 1990) which have strong dynamic signals. Alternative explanations include:

- a signature of irreversible processes in the pycnocline.
- an artifact of subduction of surface mixed-layers or injection of benthic boundary layers from the flanks of Ampere Seamount (Armi, 1978).

As discussed in the introduction, the 60-m wavelengths of the finestructure are too large to be due to pycnocline turbulence given typical measured microstructure dissipation levels (Moum and Osborn, 1986; Gregg, 1987; Gregg, 1989; Yamazaki *et al.*, 1990). They are, however, similar to typical winter mixed-layer depths. Following subduction, vertically stacked mixed layers would slowly collapse and lose their dynamic signature (Gill, 1981). Dissipative turbulence and mixing concentrated at the base of the mixed layer would smear any pronounced sheet-and-layer structure, blending them into the background stratification. Alternatively, Nabatov and Ozmidov (1988) reported evidence of actively mixing layers in the pycnocline a few kilometers from Ampere Seamount. These layers appeared to have been generated by tidal advection of benthic boundary layers off the summit. While tidal advection would limit the excursion of the layers to within a few kilometers of the seamount, geostrophic flow like that we observe could carry them further. Emery (personal communication, 1991) and Roden (personal communication, 1991) have also found evidence of detached benthic boundary layers at the depth of the summit in the vicinity of Fieberling Guyot in the North Pacific. While our measurements indicate stretching ertel vorticity finestructure throughout the pycnocline, we cannot exclude this mechanism yet. We infer that the observed finestructure is not related to flow separation but is due to injection of mixed layers from either the

surface or benthic boundary. Additional analysis is needed to distinguish between these two boundary sources.

Acknowledgments: We would like to thank John Dunlap for his assistance in modifying Eric D'Asaro's XCP data acquisition program to acquire XBT data, Maureen Kennelly for handling the organization and logistics, Tom Lehman for the CTD work, and Captain Paul Howland and the crew of OCEANUS for their able assistance. This work was supported by ONR grant N00014-90-J-1100.

REFERENCES

Armi, L, 1978: Some evidence of boundary mixing in the deep ocean. *J. Geophys. Res.*, **83**, 1971-1979.

Briscoe, M.G., 1977: On current finestructure and moored current-meter measurements of internal waves. *Deep-Sea Res.*, **24**, 1121-1131.

D'Asaro, E.A., 1988: Generation of submesoscale vortices: A new mechanism. *J. Geophys. Res.*, **93**, 6685-6693.

D'Asaro, E.A., and M.D. Morehead, 1991: Internal waves and velocity finestructure in the Arctic Ocean. *J. Geophys. Res.*, **96**, submitted.

Drever, R.G., and M.A. Kennelly, 1991: Design and use of a simple, low-cost drifter. in preparation.

Eriksen, C.C., 1978: Measurements and models of finestructure, internal gravity waves, and wave breaking in the deep ocean. *J. Geophys. Res.*, **83**, 2989-3009.

Ertel, H, 1942: Ein neuer hydrodynamischer Wirbelsatz. *Meteorol. Z.*, **59**, 277-281.

Garrett, C.J.R., and W. Munk, 1971: Internal wave spectra in the presence of finestructure. *J. Phys. Oceanogr.*, **1**, 196-202.

Garrett, C.J.R., and W. Munk, 1979: Internal waves in the ocean. *Ann. Rev. Fluid Mech.*, **11**, 339-369.

Gill, A.E., 1981: Homogeneous intrusions in a rotating stratified fluid. *J. Fluid Mech.*, **103**, 275-295.

Gregg, M.C., 1987: Diapycnal mixing in the thermocline: A review. *J. Geophys Res.*, **92**, 5249-5286.

Gregg, M.C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.*, **94**, 9686-9698.

Gregg, M.C., E.A. D'Asaro, T.J. Shay and N. Larson, 1986: Observations of persistent mixing and near-inertial internal waves. *J. Phys. Oceanogr.*, **16**, 856-885.

Haynes, P.H., and M.E. McIntyre, 1986: On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.*, **44**, 828-841.

Herring, J.A., and O. Metais, 1989: Numerical experiments in forced stably-stratified turbulence. *J. Fluid Mech.*, **202**, 97-115.

Holloway, G., 1983: A conjecture relating oceanic internal waves and small-scale processes. *Atmos.-Ocean*, **31**, 107-122.

Holloway, G., and D. Ramsden, 1989: Energy transfers by internal gravity wave interactions/stably stratified turbulence, presented at the *7th Conference on Atmospheric and Oceanic Waves and Stability*, Am. Meteorol. Soc., San Francisco, CA.

Itsweire, E.C., T.R. Osborn and T.P. Stanton, 1989: Horizontal distribution and characteristics of shear layers in the seasonal thermocline. *J. Phys. Oceanogr.*, **19**, 301-320.

Kunze, E., M.G. Briscoe, A.J. Williams III, 1990: Interpreting shear and strain finestructure from a neutrally-buoyant float. *J. Geophys. Res.*, **95**, 18,111-18,125.

Lelong, M.-P., and J.J. Riley, 1991: Internal wave-vortical mode interactions in strongly stratified flows. *J. Fluid Mech.*, submitted.

Levine, M.D., and J.D. Irish, 1981: A statistical description of temperature finestructure in the presence of internal waves. *J. Phys. Oceanogr.*, **11**, 676-691.

Lien, R.-C., 1990: Normal mode decomposition of small-scale oceanic motions. Ph.D. thesis, U of Hawaii, 128 pp.

Lilly, D.K., 1983: Stratified turbulence and the mesoscale variability of the atmosphere. *J. Atmos. Sci.*, **40**, 749-760.

Lin, J.-T., and Y.-H. Pao, 1979: Wakes in stratified fluids. *Ann. Rev. Fluid Mech.*, **11**, 317-338.

Marmorino, G.O., L.J. Rosenblum and C.L. Trump, 1987: Finescale temperature variability: The influence of near-inertial waves. *J. Geophys. Res.*, **92**, 13,049-13,062.

McKean, R.S., 1974: Interpretation of internal wave measurements in the presence of finestructure. *J. Phys. Oceanogr.*, **4**, 200-213.

McWilliams, J.C., 1989: Statistical properties of decaying geostrophic turbulence. *J. Fluid Mech.*, **198**, 199-230.

Metais, O., and J.R. Herring, 1989: Numerical simulations of freely evolving turbulence in stably stratified fluids. *J. Fluid Mech.*, **202**, 117-148.

Moum, J.N., and T.R. Osborn, 1986: Mixing in the main thermocline. *J. Phys. Oceanogr.*, **16**, 1250-1259.

Müller, P., 1984: Smallscale vortical motion, in *Internal Gravity Waves and Smallscale Turbulence, Proc. 'Aha Huliko'a Hawaiian Winter Workshop*, P. Müller and R. Pujalet, Eds., Hawaiian Inst. Geophys., Honolulu, 249-262.

Müller, P., D.J. Olbers and J. Willebrand, 1978: The IWEX spectrum. *J. Geophys. Res.*, **83**, 479-500.

Müller, P., R.-C. Lien and R. Williams, 1988: Estimates of potential vorticity at small scales in the ocean. *J. Phys. Oceanogr.*, **18**, 401-416.

Nabatov, V.N., and R.V. Ozmidov, 1988: Study of turbulence above seamounts in the Atlantic Ocean. *Oceanology*, **28**, 161-166.

Pedlosky, J., 1979: *Geophysical Fluid Dynamics*. Springer-Verlag, 624 pp.

Phillips, O.M., 1971: On spectra measured in an undulating layered medium. *J. Phys. Oceanogr.*, **1**, 1-6.

Prater, M.D., and T.B. Sanford, 1990: Generation of Meddies off Cape St. Vincent, Portugal. *EOS Transac., Amer. Geophys. Union*, **71**(43), 1416.

Riley, J.J., R.W. Metcalfe and M.A. Weissman, 1981: Direct numerical simulations of homogeneous turbulence in density-stratified fluids. in *Nonlinear Properties of Internal Waves*, 76, B.J. West, Ed., Amer. Inst. Phys., New York, 79-112.

Sherman, J.T., and R. Pinkel, 1991: Estimates of the vertical wavenumber-frequency spectra of vertical shear and strain. *J. Phys. Oceanogr.*, **21**, in press.

Staquet, C., and J.J. Riley, 1989: A numerical study of a stably-stratified mixing layer. *Turbulent Shear Flows*, **6**, Selected Pap. 6th Intl., Symp. 6th, 381-397.

Yamazaki, H., R.G. Lueck and T. Osborn, 1990: A comparison of turbulence data from a submarine and a vertical profiler. *J. Phys. Oceanogr.*, **20**, 1778-1786.

Symmetry Preserving Mode Truncations of Inviscid Geophysical Fluid Dynamical Equations

Henry D.I. Abarbanel[1]
Marine Physical Laboratory, Scripps Institution of Oceanography,
and
Department of Physics, University of California, San Diego, Mail Code 0402, La Jolla, CA 92093-0402

and

Ali Rouhi
Institute for Nonlinear Science, University of California, San Diego, Mail Code 0402, La Jolla, CA 92093-0402

Abstract

We investigate the role of potential vorticity in nearly two dimensional flows of importance in geophysical fluid dynamics. Potential vorticity conservation arises from particle interchange symmetry in the Lagrangian formulation of fluid dynamics and is associated with an infinite dimensional symmetry group. In truncating the number of degrees of freedom of these fluid flows, as one does when making numerical integrations of the theory, it is not possible to keep the full infinite dimensional symmetry group. We show, in the context of the shallow water equations, how to modify the symmetry algebra and construct a Hamiltonian for the fluid which preserves the maximum symmetry consistent with the finite number of retained degrees of freedom and which becomes the original fluid as the number of degrees of freedom increases to infinity. The construction is done in planar geometry without rotation, but it also goes through for f or β plane settings, for flows on a sphere (rotating or not) and for stratified fluids. The latter application includes both internal and surface gravity waves.

1 Introduction

This is a talk about methods of truncating or restricting the number of degrees of freedom in equations of motion of relevance to geophysical fluid dynamics while preserving the symmetries leading to conservation laws respected by those evolution equations. In particular, the symmetry which will concern us here is that of particle relabeling in

[1]Institute for Nonlinear Science

Lagrangian coordinates [Eckart 1960] which leads to conservation of potential vorticity in either Lagrangian or Eulerian formulations of the theory. The results provide a consistent mode truncation of the full continuum theory which preserves the invariance of as many of the conserved quantities of the continuum theory as is consistent with the number of retained degrees of freedom. Further, as the number of degrees of freedom goes to infinity, the original continuum theory is recovered and the full set of conserved quantities is recovered as well. This provides the possibility of reducing the number of degrees of freedom of a continuum geophysical fluid dynamics flow to a finite number, the only situation which can be treated numerically, and still preserving the maximum possible symmetry of the underlying theory.

The methods we present here are Hamiltonian, and the fluid dynamics is cast in Lagrangian realization. The advantage of this is that the underlying Lagrangian theory is canonical in the classical mechanics sense and the symmetries of the theory are manifest and easy to deal with. The correspondir g Eulerian theory is non-canonical and the symmetries are hidden or "mysterious". The reason for this disguise of the symmetries is that the Eulerian theory is "reduced" from the Lagrangian formulation by considering the flow only on hypersurfaces in phase space where the conserved quantities are constant. The manifestation of these symmetries in terms of conserved quantities seems unnatural in Eulerian formulation while appearing quite natural in Lagrangian formulation.

An outline of this talk is as follows:

- (1) Lagrangian formulation of the Shallow Water Equations and the Internal Wave Equations
- (2) Invariance under particle interchange and potential vorticity conservation.
- (3) Truncating the Fourier modes and SU(N) symmetry
  - Algebra of Symmetry Generators and Dynamical Variables
  - Conserved Quantities
- SU(N) symmetric Hamiltonian, H_N; $N \to \infty$ leads to usual equations

In the SU(N) symmetric truncated theory there are $\approx N$ conserved quantities. As $N \to \infty$, we recover the continuum theory and an infinite number of conserved quantities.

This talk focuses on the shallow water equations [Pedlosky 1979] because it is for these that we have concrete results at this time. The shallow water equations are also formulated on an f-plane, that is Cartesian or flat geometry, in this talk. We know how to extend the results to flows on the surface of a sphere, but the algebra is difficult and will be reported elsewhere. In progress is work on extending these results to quasi-two dimensional geophysical flows including internal waves on a plane (f or β) and surface gravity waves. The reader will see that our methods are generally applicable to flows with a conserved potential vorticity. If there is driving and damping also present in the physical setting, then we can regard the work here as establishing a finite set of coordinates for such dynamics. When the driving or damping is not significant, then in the coordinates we present the required conservation laws are respected automatically. In that sense they provide a rational choice of truncated modes for all numerical work on geophysical problems where quasi-two dimensionality is a feature.

Our motivation for concentrating on potential vorticity modes is two fold:

- The work of Müller and co-workers [Müller 1988a, Müller 1988b] has provided evidence for the geophysical importance of potential vorticity carrying motions even at small scales.
- Conserved quantities are always important for constraining the allowed physical motions of a system and for checking numerical integrations of those equations of motion.

2 Lagrangian Fluids

In the description of fluids by the Lagrangian method [Abarbanel 1987] we are required to give the position of a fluid particle $\mathbf{y}(\mathbf{r},t)$ and its canonical momentum $\mathbf{p}(\mathbf{r},t)$ for each particle label $\mathbf{r}$, which is a two or three dimensional continuum of labels for particles. The evolution equations of these variables follows from an Action Principle which is really just Hamilton's Principle. This states that the action S in d-dimensions:

$$S(\mathbf{y},\mathbf{p}) = \int_{t_1}^{t_2} dt \int d^d r\, \rho_0(\mathbf{r})\{\frac{1}{2}\frac{\partial \mathbf{y}(\mathbf{r},t)}{\partial t}\cdot\frac{\partial \mathbf{y}(\mathbf{r},t)}{\partial t} - \epsilon(\rho,s)\} \tag{1}$$

is stationary under changes of $\mathbf{y}(\mathbf{r},t)$ near the orbit of the system. Here the internal energy density $\epsilon(\rho,s)$ is a thermodynamic quantity from which the pressure is derived. It is a

function of the density and the specific entropy. In this expression for the action all partial derivatives with respect to time are with $\mathbf{r}$ held fixed. The density $\rho(\mathbf{y})$ is

$$\rho(\mathbf{y}(\mathbf{r},t)) = \rho_0(\mathbf{r})\frac{\partial(\mathbf{r})}{\partial(\mathbf{y}(\mathbf{r},t))}, \tag{2}$$

and $s = s(\mathbf{r})$ is the entropy per unit volume. Varying S with respect to $\mathbf{y}(\mathbf{r},t)$ with $\mathbf{r}$ and t fixed leads to the equations of motion

$$\begin{aligned}
\rho(\mathbf{y})\frac{\partial^2\mathbf{y}(\mathbf{r},t)}{\partial t^2} &= -\nabla p(\mathbf{y},t), \\
\frac{\partial\rho(\mathbf{y},t)}{\partial t} &= -\rho(\mathbf{y})\nabla_{\mathbf{y}}\cdot(\frac{\partial\mathbf{y}(\mathbf{r},t)}{\partial t}), \\
\frac{\partial s}{\partial t} &= 0, \\
p &= \rho^2\epsilon_\rho.
\end{aligned} \tag{3}$$

To reach the Eulerian formulation of fluid dynamics we identify a fixed point in space $\mathbf{x}$ with the location $\mathbf{y}(\mathbf{r},t)$ of a particular fluid particle at time t. This defines a particular label $\mathbf{R}(\mathbf{x},t)$ which identifies the fluid particle which arrives at $\mathbf{x}$ at the appointed time, so

$$\mathbf{x} = \mathbf{y}(\mathbf{R}(\mathbf{x},t),t), \tag{4}$$

and the Eulerian velocity $\mathbf{u}_E(\mathbf{x},t)$ is defined as

$$\mathbf{u}_E(\mathbf{x},t) = \frac{\partial\mathbf{y}(\mathbf{r},t)}{\partial t}|_{\mathbf{r}=\mathbf{R}(\mathbf{x},t)}. \tag{5}$$

The Lagrangian derivative at fixed label $\mathbf{r}$ becomes

$$\frac{\partial}{\partial t}|_{\mathbf{r}} = \frac{\partial}{\partial t}|_{\mathbf{x}} + \mathbf{u}_E(\mathbf{x},t)\cdot\nabla_{\mathbf{x}}. \tag{6}$$

The Eulerian formulation at fixed spatial points $\mathbf{x}$ is a **reduced** description of the fluid theory [Marsden 1984] since it describes flows restricted to surfaces in the fluid state space which have constant values of the conserved potential vorticity. Lagrangian fluid dynamics describes the evolution of *six* fields: the canonical coordinates $\mathbf{y}(\mathbf{r},t)$ and their canonical momenta $\mathbf{p}(\mathbf{r},t)$. Eulerian fluid dynamics describes the evolution of *five* fields: $\mathbf{u}_E(\mathbf{x},t), \rho(\mathbf{x},t)$, and the specific entropy $s(\mathbf{x},t)$. This reduction in number can be traced to the restriction of the flows to motion on constant potential vorticity surfaces, and that brings us to potential vorticity and its interpretation.

In this talk we consider the shallow water equations as our paradigm for a nearly two dimensional fluid with a conserved quantity. We wish to truncate to a finite number the continuum degrees of freedom of the fluid and to do so in a fashion which preserves a subset of the symmetry leading to potential vorticity conservation The truncated theory must become the correct continuum theory as the number of modes goes to infinity.
The shallow water equations result from stationarity of the action

$$S = \int_{t_1}^{t_2} dt \int d^2r\, h_0(\mathbf{r}) \{ \frac{1}{2} \frac{\partial \mathbf{y}(\mathbf{r},t)}{\partial t} \cdot \frac{\partial \mathbf{y}(\mathbf{r},t)}{\partial t} - \frac{g}{2} \frac{1}{J} \}, \tag{7}$$

with $J = \frac{\partial(\mathbf{y}(\mathbf{r},t))}{\partial(\mathbf{r})}$. One can absorb the initial "height" $h_0(\mathbf{r})$ into the definition of the labels $\int d^2r h_0(\mathbf{r}) \rightarrow \int d^2r$ without any loss of generality, and we do that to simplify our formulae. The canonical momentum is defined in the usual way as the derivative of the Lagrangian $S = \int_{t_1}^{t_2} dt \int d^2r L[\mathbf{y}(\mathbf{r},t), \partial_t \mathbf{y}(\mathbf{r},t)]$ with respect to $\partial_t \mathbf{y}(\mathbf{r},t)$, so $\mathbf{p}(\mathbf{r},t) = \partial_t \mathbf{y}(\mathbf{r},t)$. The shallow water Hamiltonian is then

$$H(\mathbf{y},\mathbf{p}) = \frac{1}{2} \int d^2r\, [|\mathbf{p}(\mathbf{r},t)|^2 + \frac{g}{J}], \tag{8}$$

and the equations of motion follow from the Poisson bracket relation

$$\frac{\partial \bullet}{\partial t} = \{\bullet, H(\mathbf{y},\mathbf{p})\}, \tag{9}$$

using the fundamental Poisson bracket

$$\{y_a(\mathbf{r},t), p_b(\mathbf{r}',t)\} = \delta_{ab} \delta^2(\mathbf{r}-\mathbf{r}'). \tag{10}$$

Under particle interchanges which preserve the density (or $h_0(\mathbf{r})$ here) the action is invariant. This is formally expressed by requiring that $\delta_{\mathbf{r}} S = 0$ with $\mathbf{y}(\mathbf{r},t)$ and density held fixed, and was pointed out first by Eckart in 1960 [Eckart 1960]. The conserved quantity which results from this symmetry of the action is the potential vorticity

$$q(\mathbf{r},t) = \epsilon_{ab} \frac{\partial \mathbf{y}_a(\mathbf{r},t)}{\partial r_a} \frac{\partial \mathbf{p}_a(\mathbf{r},t)}{\partial r_b}, \tag{11}$$

and

$$\frac{\partial q(\mathbf{r},t)}{\partial t} = 0. \tag{12}$$

Translating this into Eulerian variables using the prescription given above results in

$$q_E(\mathbf{x},t) = \frac{\hat{z}\cdot\nabla_{\mathbf{X}}\times\mathbf{u}_E(\mathbf{x},t)}{h_E(\mathbf{x},t)}, \tag{13}$$

and

$$(\frac{\partial}{\partial t} + \mathbf{u}_E\cdot\nabla_{\mathbf{X}})q_E(\mathbf{x},t) = 0. \tag{14}$$

$h_E(\mathbf{x},t)$ is the usual Eulerian fluid thickness in shallow water theory and comes from the Lagrangian quantity $h(\mathbf{r},t) = \frac{1}{J}$.

These conservation laws lead to the statements that for arbitrary functions G

$$\frac{\partial}{\partial t}\int d^2r\, G(q(\mathbf{r},t)) = 0, \tag{15}$$

and

$$\frac{\partial}{\partial t}\int d^2x\, h_E(\mathbf{x},t)G(q_E(\mathbf{x},t)) = 0. \tag{16}$$

These are an infinite number of conserved quantities associated with the local particle interchange symmetry. Next we examine the algebra associated with this symmetry noting it is $q(\mathbf{r},t)$ which is the infinitesimal generator of the symmetry.

Before delving into the algebra let us make the connection with internal wave dynamics. For internal waves the flows are three dimensional and the Hamiltonian is

$$H(\mathbf{y},\mathbf{p}) = \int d^3r\,\{\frac{|\mathbf{p}-\rho_0\hat{\mathbf{R}}|^2}{2\rho_0} + \rho_0 g y_3(\mathbf{r},t) + \rho_0\epsilon(\rho)\}, \tag{17}$$

where $\hat{\mathbf{R}}$ is the rotational potential whose curl is $\hat{z}f(\mathbf{r})$, and the initial density $\rho_0(r_3)$ is taken to depend on the vertical coordinate only. The quantity conserved under particle interchange for this theory is

$$q(\mathbf{r},t) = \sum_{\alpha=1}^{3}\{\frac{\partial y_\alpha(\mathbf{r},t)}{\partial r_1}\frac{\partial}{\partial r_2}[\frac{\partial y_\alpha(\mathbf{r},t)}{\partial t} + \hat{\mathbf{R}}(\mathbf{y})_\alpha] - (r_1\leftrightarrow r_2)\}, \tag{18}$$

or in Eulerian variables

$$q_E(\mathbf{x},t) = (\hat{z}f + \nabla\times\mathbf{u}_E)\cdot\nabla\rho. \tag{19}$$

3 Algebra of Particle Interchange Symmetry

To exhibit the algebra associated with the particle interchange symmetry of the shallow water equations, it is easier to go from configuration space r to Fourier

space [Abarbanel 1991]. For this we place the configuration space in a box of size $L \times L$ and define Fourier transforms via

$$\begin{aligned} f(\mathbf{r}) &= \sum_{\mathbf{n}=-\infty}^{+\infty} F(\mathbf{n}) \exp[i\kappa \mathbf{n} \cdot \mathbf{r}], \\ F(\mathbf{n}) &= \frac{1}{L^2} \int d^2 r f(\mathbf{r}) \exp[-i\kappa \mathbf{n} \cdot \mathbf{r}]. \end{aligned} \tag{20}$$

Here the vector $\mathbf{n} = [n_1, n_2]$ with n_i are integers $n_i = 0, \pm 1, \pm 2, , \ldots, \pm\infty$, and $\kappa = \frac{2\pi}{L}$. With this Fourier transform pair we define $\mathbf{Y}(\mathbf{n}, t)$ and $\mathbf{P}(\mathbf{n}, t)$ as

$$\begin{aligned} y_\alpha(\mathbf{r}, t) &= \frac{1}{L} \sum_{\mathbf{n}=-\infty}^{+\infty} Y_\alpha(\mathbf{n}, t) \exp[i\kappa \mathbf{n} \cdot \mathbf{r}], \\ p_\alpha(\mathbf{r}, t) &= \frac{1}{L} \sum_{\mathbf{n}=-\infty}^{+\infty} P_\alpha(\mathbf{n}, t) \exp[i\kappa \mathbf{n} \cdot \mathbf{r}], \end{aligned} \tag{21}$$

with the normalization chosen so the fundamental Poisson bracket becomes

$$\{Y_\alpha(\mathbf{n}, t), P_\beta(\mathbf{m}, t)\} = \delta_{\alpha\beta} \delta_{0, \mathbf{m}+\mathbf{n}}. \tag{22}$$

The Fourier components of the potential vorticity are taken as

$$q(\mathbf{r}, t) = \frac{(2\pi)^2}{L^4} \sum_{\mathbf{n}} Q(\mathbf{n}, t) \exp[i\kappa \mathbf{n} \cdot \mathbf{r}], \tag{23}$$

which leads to

$$Q(\mathbf{n}, t) = \sum_{\mathbf{m}, \mathbf{m}'} \mathbf{m}' \times \mathbf{m} \, P_\alpha(\mathbf{m}) Q_\alpha(\mathbf{m}') \delta_{\mathbf{n}, \mathbf{m}+\mathbf{m}'}, \tag{24}$$

where $\mathbf{m}' \times \mathbf{m} = m_1' m_2 - m_2' m_1$ is the z component of the cross product among vectors. With these definitions of Fourier components we can easily evaluate the Poisson brackets of the $Q(\mathbf{n})$ which are the generators of the local particle interchange symmetry with the $\mathbf{Y}(\mathbf{n})$, the $\mathbf{P}(\mathbf{n})$, and themselves. This leads to

$$\begin{aligned} \{Q(\mathbf{n}), Y_\alpha(\mathbf{m})\} &= \mathbf{n} \times \mathbf{m} \, Y_\alpha(\mathbf{m} + \mathbf{n}), \\ \{Q(\mathbf{n}), P_\alpha(\mathbf{m})\} &= \mathbf{n} \times \mathbf{m} \, P_\alpha(\mathbf{m} + \mathbf{n}), \\ \{Q(\mathbf{n}), Q(\mathbf{m})\} &= \mathbf{n} \times \mathbf{m} \, Q(\mathbf{m} + \mathbf{n}), \end{aligned} \tag{25}$$

so the algebra of the $Q(\mathbf{n})$ closes, as it must if we have a symmetry, and the Fourier components of the $\mathbf{y}$ and the $\mathbf{p}$ transform under the algebra as "vectors". The factors of

$\mathbf{m} \times \mathbf{n}$ are the *structure constants* of the group of particle interchange symmetry.

Perhaps a more familiar example of this kind of algebra is that of three dimensional angular momentum in classical mechanics. The angular momentum $\mathbf{L} = \mathbf{q} \times \mathbf{p}$ or $L_a = \epsilon_{abc} q_b p_c$; $a, b = 1, 2, 3$ has the following Poisson brackets with the coordinates $\mathbf{q}$, the momenta $\mathbf{p}$, and $\mathbf{L}$ which follow from the fundamental bracket $\{q_a, p_b\} = \delta_{ab}$:

$$\begin{aligned} \{L_a, q_b\} &= \epsilon_{abc} q_c, \\ \{L_a, p_b\} &= \epsilon_{abc} p_c, \\ \{L_a, L_b\} &= \epsilon_{abc} L_c. \end{aligned} \tag{26}$$

Any quantity which satisfies $\{L_a, v_b\} = \epsilon_{abc} v_c$ is a vector under three dimensional rotations which are generated by $\mathbf{L}$. The dot product $\mathbf{v} \cdot \mathbf{v}$ is unchanged under rotations since $\{L_a, \mathbf{v} \cdot \mathbf{v}\} = 0$, and $\mathbf{L} \cdot \mathbf{L}$ is the invariant of the algebra of the rotation group. Rotational invariance of the dynamics of a system is guaranteed by having an Hamiltonian $H(\mathbf{p}, \mathbf{q})$ which satisfies

$$\{L_a, H(\mathbf{p}, \mathbf{q})\} = 0. \tag{27}$$

This also leads to the conservation (under evolution in time under $H(\mathbf{p}, \mathbf{q})$) of $L^2 = \mathbf{L} \cdot \mathbf{L}$ and any function of L^2.

A critical aspect of the angular momentum algebra which we must establish for our particle interchange algebra is that the Poisson brackets satisfy the Jacobi identity

$$\{L_a, \{L_b, L_c\}\} + \{L_b, \{L_c, L_a\}\} + \{L_c, \{L_a, L_b\}\} = 0, \tag{28}$$

for this guarantees that a combination of rotations is also a rotation and that under evolution through a **finite** time under $H(\mathbf{p}, \mathbf{q})$ angular momentum is conserved.

Now we return to the shallow water equations. The final ingredient we require for constructing the truncated Hamiltonian for shallow water flow is the Fourier decomposition of the Jacobian and the transformation properties of these Fourier coefficients under particle interchange. This decomposition is easily established to be

$$\begin{aligned} J &= \frac{\partial(\mathbf{y}(\mathbf{r}, t))}{\partial(\mathbf{r})} \\ &= \frac{(2\pi)^2}{L^4} \sum_{\mathbf{n}} \rho(\mathbf{n}, t) \exp[i\kappa \mathbf{n} \cdot \mathbf{r}], \\ \rho(\mathbf{n}) &= \frac{1}{2} \sum_{\mathbf{m}, \mathbf{m}'} (\mathbf{m}' \times \mathbf{m}) \, \mathbf{Y}(\mathbf{m}) \times \mathbf{Y}(\mathbf{m}') \, \delta_{\mathbf{n}, \mathbf{m}+\mathbf{m}'}, \end{aligned} \tag{29}$$

from which

$$\{Q(\mathbf{n}), \rho(\mathbf{m})\} = \mathbf{n} \times \mathbf{m}\, \rho(\mathbf{m}+\mathbf{n}) \tag{30}$$

follows. So the Fourier components of the Jacobian are also vectors under the transformations generated by the Fourier components of potential vorticity.

Just as with the three dimensional angular momentum example above, the Jacobi identity among the $Q(\mathbf{n})$ is critical in guaranteeing that finite particle interchange transformations such as are generated by finite time evolution under the shallow water Hamiltonian lead to potential vorticity conservation.

Before displaying our truncated shallow water theory we recall how the potential vorticity $Q(\mathbf{n}, t)$ is conserved in the case with an infinite number of Fourier components. For this we need to compute the Poisson bracket of $Q(\mathbf{n}, t)$ with the Hamiltonian

$$H = \frac{1}{2} \sum_{\mathbf{n}=-\infty}^{+\infty} P_\alpha(\mathbf{n}) P_\alpha(-\mathbf{n}) + \frac{g}{2} \int d^2 r \, \frac{1}{J}. \tag{31}$$

The Poisson bracket with the first term in II is up to a factor of 2

$$\sum_{\mathbf{m},\mathbf{m}'} (\mathbf{m}' \times \mathbf{m})\, \mathbf{P}(\mathbf{m}) \cdot \mathbf{P}(\mathbf{m}') \delta_{\mathbf{n},\mathbf{m}+\mathbf{m}'}, \tag{32}$$

which vanishes because of symmetry in the $\mathbf{m}, \mathbf{m}'$ sum. The Poisson bracket with the second term is (up to a constant factor)

$$\frac{1}{\kappa^2} \int d^2 r \frac{\partial(\exp[-i\kappa \mathbf{n} \cdot \mathbf{r}], J^{-1})}{\partial(\mathbf{r})}, \tag{33}$$

which vanishes by integration by parts. In a mode truncated theory the first part of this will remain: the kinetic energy will still Poisson commute with potential vorticity, but integration by parts will be absent since we will no longer have a continuum theory in label space.

4 Truncating the Number of Modes; a New Potential Vorticity Algebra

Now we restrict the number of Fourier modes allowed to the variables $\mathbf{Y}(\mathbf{n}, t)$ and $\mathbf{P}(\mathbf{n}, t)$ by keeping the Fourier sums in the bounds $-M \leq n_i \leq M$ for $i = 1, 2$. We now have N^2 degrees of freedom where $N = 2M + 1$. The fundamental Poisson bracket among the $\mathbf{Y}(\mathbf{n})$

and the $\mathbf{P}(\mathbf{n})$ is unchanged except the rule is to keep all Fourier indices within $[-M, M]$, so when $\mathbf{n}+\mathbf{m}$ appears it is to be so restricted. The problem comes when we go to the Poisson brackets of the potential vorticity $Q(\mathbf{n})$ with the coordinates or the canonical momenta or the Fourier components of the Jacobian or with itself. In this we encounter the cross product $\mathbf{m} \times \mathbf{n}$ which is the structure constant for the group action of $Q(\mathbf{n})$ in the Euclidian space of Fourier indices. By our truncation of modes we have changed the space of Fourier modes from the plane to that of a two dimensional torus; this is because we have introduced an effective periodicity in Fourier labels. To match this and preserve the Jacobi identities we replace $\mathbf{m} \times \mathbf{n}$ by

$$\mathbf{n} \times \mathbf{m} \longrightarrow \frac{1}{\kappa_N} \sin[\kappa_N(\mathbf{n} \times \mathbf{m})], \tag{34}$$

where $\kappa_N = \frac{2\pi}{N}$. Clearly as $N \to \infty$ this reduces back to the Euclidian space version $\mathbf{n} \times \mathbf{m}$. For finite N, which is our concern here, we have an effective periodicity in Fourier space now respected by the new structure constants. What is truly remarkable, however, is that this simple replacement of $\mathbf{m} \times \mathbf{n}$ also respects the Jacobi identity so a group structure is retained [Hoppe 1989].

With these new structure constants we can write the Fourier decomposition of the potential vorticity

$$Q_N(\mathbf{n}) = \frac{1}{\kappa_N} \sum_{\mathbf{m}=-M}^{M} \sin[\kappa_N(\mathbf{n} \times \mathbf{m})]\mathbf{P}(\mathbf{m}) \cdot \mathbf{Y}(\mathbf{n}-\mathbf{m}), \tag{35}$$

and for the Jacobian Fourier components, we write

$$\rho_N(\mathbf{n}) = \frac{1}{\kappa_N} \sum_{\mathbf{m}=-M}^{M} \sin[\kappa_N(\mathbf{n} \times \mathbf{m})]Y_1(\mathbf{m})Y_2(\mathbf{n}-\mathbf{m}). \tag{36}$$

The Poisson brackets of this new $Q_N(\mathbf{n})$ with any of $\mathbf{Y}(\mathbf{n}), \mathbf{P}(\mathbf{n}), \rho_N(\mathbf{n})$ or $Q(\mathbf{n})$ takes the form

$$\{Q_N(\mathbf{n}), f(\mathbf{m})\} = \frac{\sin[\kappa_N(\mathbf{m} \times \mathbf{n})]}{\kappa_N} f(\mathbf{m}+\mathbf{n}), \tag{37}$$

with $f(\mathbf{m})$ any component of the canonical coordinates or canonical momentum or $\rho_N(\mathbf{m})$ or $Q_N(\mathbf{m})$. This set of Poisson brackets now defines a finite algebra of particle interchange transformations generated by the $Q_N(\mathbf{n})$. It also defines anything which transform as $f(\mathbf{m})$ here as a vector under this new transformation group. The group structure is guaranteed by satisfying the Jacobi identity, the demonstration of which is a tedious task left to the

dedicated reader. Our job now is to establish a Hamiltonian $H_N(\mathbf{Y},\mathbf{P})$ in these truncated variables which Poisson commutes with $Q_N(\mathbf{n})$ and becomes just the shallow water Hamiltonian as $N \to \infty$. The easiest method is to seek **invariants** of the finite particle interchange algebra (it happens to be SU(N)) and construct H_N out of them.

4.1 Invariants of the Truncated Algebra

The idea is to use the transformation properties of **vectors** $f(\mathbf{m})$ under the $Q_N(\mathbf{n})$ algebra

$$\{Q_N(\mathbf{n}), f(\mathbf{m})\} = \frac{\sin[\kappa_N(\mathbf{m}\times\mathbf{n})]}{\kappa_N} f(\mathbf{m}+\mathbf{n}), \tag{38}$$

to form "dot products" $C_p(f)$ such that

$$\{Q_N(\mathbf{n}), C_p(f)\} = 0. \tag{39}$$

The $C_p(f)$ made out of powers of $f(\mathbf{m})$ are

$$\begin{aligned}
C_2(f) &= \sum_{\mathbf{m},\mathbf{m}'=-M}^{M} f(\mathbf{m})f(\mathbf{m}')\delta_{0,\mathbf{m}+\mathbf{m}'} \\
C_3(f) &= \sum_{\mathbf{n}_1,\mathbf{n}_2,\mathbf{n}_3} f(\mathbf{n}_1)f(\mathbf{n}_2)f(\mathbf{n}_3)\delta_{0,\mathbf{n}_1+\mathbf{n}_2+\mathbf{n}_3}\exp[i\kappa_N(\mathbf{n}_1\times\mathbf{n}_2+\mathbf{n}_1\times\mathbf{n}_3+\mathbf{n}_2\times\mathbf{n}_3)] \\
&\vdots \\
C_{L+1}(f) &= \sum_{\mathbf{n}_1\ldots\mathbf{n}_L}\prod_{\alpha<\beta}\exp[i\kappa_N(\mathbf{n}_\alpha\times\mathbf{n}_\beta]\, f(\mathbf{n}_1)f(\mathbf{n}_2)\ldots f(\mathbf{n}_L)f(-(\mathbf{n}_1+\mathbf{n}_2+\ldots\mathbf{n}_L)).
\end{aligned} \tag{40}$$

So these are generalized "powers of vectors".

The kinetic energy term in the truncated Hamiltonian

$$KE_N = \frac{1}{2}\sum_{\mathbf{n}=-M}^{M}\mathbf{P}(\mathbf{n})\cdot\mathbf{P}(-\mathbf{n}), \tag{41}$$

is just $C_2(\mathbf{P})$ up to a constant. The term involving $\frac{1}{J}$ requires some thought. The idea is to express $\frac{1}{J}$ as a power series around some finite value J_0 and then truncate the sum with N terms. Then we replace each of the integrals of $\frac{J}{J_0}$ by $C_p(\rho_N)$ up to constants. The natural value of J_0 is unity since for small displacements $\mathbf{Y}(\mathbf{r},t) = \mathbf{r}$ + small terms, and for $\mathbf{Y} = \mathbf{r}, J = 1$. For general J_0 we write

$$\begin{aligned}\frac{1}{J} &= \frac{1}{J_0-(J_0-J)} \\ &= \frac{1}{J_0}\sum_{k=0}^{\infty}(1-\frac{J}{J_0})^k \\ &\approx \frac{1}{J_0}\sum_{k=0}^{N-1}(1-\frac{J}{J_0})^k \\ &= \frac{1}{J}(1-(1-\frac{J}{J_0})^N), \end{aligned} \tag{42}$$

which is very nearly $\frac{1}{J}$ when $0 < J < 2J_)$ and N is large.

In the expression for the potential energy in H_N involving $\int d^2r\frac{1}{J}$ we make this replacement for J^{-1} and specifically set

$$\int d^2r J^{p+1} \longrightarrow \frac{(2\pi)^{2p+2}}{L^{2(2p+1)}}C_{p+1}(\rho_N), \tag{43}$$

so our truncated Hamiltonian is

$$H_N = \frac{1}{2}\sum_{\mathbf{n}=-M}^{M}\mathbf{P}(\mathbf{n})\cdot\mathbf{P}(-\mathbf{n}) + \frac{g}{2J_0}\sum_{k=1}^{N}C_k^N\frac{(2\pi)^{2k-2}}{L^{2(2k-2)}}C_{k-1}(\rho_N). \tag{44}$$

This Hamiltonian, by construction, has zero Poisson bracket with $Q_N(\mathbf{n})$. Further $\{C_p(Q_N), H_N\} = 0$ as well.

This constitutes our mode truncated shallow water Hamiltonian and is an explicitly SU(N) symmetric approximation to the continuum shallow water theory from which we started. As the number of modes goes to infinity, the continuum theory is recovered in all its details. For finite N, the symmetry constraints of particle interchange are respected as accurately as possible.

5 Conclusions

In this talk we have presented insight into the origins of potential vorticity conservation and in doing so have investigated the algebra of infinitesimal operations associated with the particle interchange symmetry responsible for that conservation law. The generators of local infinitesimal particle interchanges are the potential vorticity at a point, and in the continuum theory their Poisson bracket algebra is infinite dimensional.

We then showed how to truncate the modes of the shallow water theory, expressed in Fourier space of its Lagrangian representation, and to alter the symmetry algebra so it

remains a symmetry algebra of the finite degree of freedom theory. In the planar geometry where we worked, this replacement was straightforward.

In the future we shall address several questions:

- the application and numerical investigation of this kind of truncation to inviscid two dimensional incompressible flow. This simplest of all theories of fluid flow has only one Eulerian dynamical field which can be taken to be the vorticity out of the plane of flow, and the algebra of this variable in Eulerian representation parallels that discussed here for the potential vorticity.
- the extension of the construction presented here to two dimensional flow on a sphere (rotating, if you like).
- the extension of these ideas to planar and spherical stratified flow for the study of internal waves and surface waves.
- the numerical investigation of these symmetric finite degree of freedom systems to understand the role played by the symmetry constraints.
- investigation of the "statistical mechanics" of these symmetric Hamilton systems and of the paths to chaos in the systems.

Another avenue of substantial interest is to understand the Eulerian version of our Lagrangian formulations of these symmetric theories. This is both for general interest and since the numerical investigation of the symmetry preserving mode truncated theories may well be easier in Eulerian variables.

Finally, since damping and driving are physical ingredients of any real observations of the ocean, we expect that these inviscid or Hamiltonian discussions will serve as means for identifying variables in which to investigate both the inviscid and the dissipative physical settings. The advantage of the variables thus suggested is that when length scales and time scales are large enough that viscosity is unimportant, all conservation laws one would want to be respected are respected.

Acknowledgments

This work has been supported in part by a Contract with the Office of Naval Research, Small Scale Physical Oceanography Program, Code 1122SS.

References

Abarbanel, H.D.I. and A. Rouhi, 1991: Symmetric Truncations of the Shallow Water Equations. UCSD/INLS Preprint, February, 1991.

Abarbanel, H.D.I. and D.D. Holm, 1987: Nonlinear Stability Analysis of Inviscid Flows in Three Dimensions: Incompressible Fluids and Barotropic Fluids. *Phys. Fluids* 30, 3369-3382.

Eckart, C., 1960: Variational Principles in Hydrodynamics. *Phys. Fluids* 3, 421-427.

Hoppe, J.: Diffeomorphism Groups, Quantization, and $SU(\infty)$. Int. *J. Mod. Phys. A*, 5235-5248.

Marsden, J., T. Ratiu, and A. Weinstein, 1984: Reduction and Hamiltonian Structures on Duals of Semidirect Product Lie Algebras. *Contemporary Mathematics* 28, 55-100.

Müller, P., 1988: Vortical Motions, in *Small-Scale Turbulence and Mixing in the Ocean*, Ed. J.C.J. Nihoul and B. M. Jamart, Elsevier, Amsterdam.

Müller, P., R-C. Lien, and R. Williams, 1988: Estimates of Potential Vorticity at Small Scales in the Ocean. *J. Phys. Oceanogr.* 18, 401-416.

Pedlosky, J., 1979: *Geophysical Fluid Dynamics*. Springer, New York.

WEAK WAVE AND VORTEX INTERACTIONS

James J. Riley, Marie-Pascale G. Lelong†, and Donald N. Slinn

University of Washington
Seattle, Washington 98195

ABSTRACT

We report on theoretical analysis and direct numerical simulations of internal wave and vortical mode interactions in stably stratified fluids. Utilizing multiple-time-scale perturbation analysis, a resonant interaction is found between a vortical mode and two internal waves in which the vortical mode plays the role of a catalyst. This interaction could cause significant modification of the internal wave field. Vortical mode self-interactions are found to be strongly nonlinear, and can be a significant source of internal waves. The numerical simulations indicate that the theoretical analysis is valid for the small parameter (Froude number) of order one or less. Furthermore, in each case computed in this range of Froude number, the vortical mode exhibited strong instabilities, transferring energy to larger horizontal scales.

1. INTRODUCTION

The traditional view of oceanographic flows at higher frequencies (between the local Coriolis and buoyancy frequencies) and smaller scales (between tens of meters up to about tens of kilometers) is that these flows mainly consist of inertial gravity waves. Laboratory visualizations have indicated, however, that quasi-horizontal meandering motions often exist superimposed upon an internal wave field. Such flows have been observed, e. g., in the later stages of decay of a turbulent wake (Lin and Pao, 1979), in the later stages of grid turbulence (Liu, 1980), and in the long-time behavior of a short duration jet (van Heijst and Flór, 1989) when such experiments are carried out in a stably stratified fluid.

In an attempt to explain the laboratory results and also some numerical simulation data, Riley et al. (1981) have suggested that the presence of quasi-horizontal structures, which have been termed vortical modes (Müller et al.,

†Present address: National Center for Atmospheric Research, Boulder, CO 80302

1986), is due to the domination of the flow by stable density stratification. Introducing a Froude number defined by $F = u'/NL$, where u' is an rms velocity, N is the buoyancy frequency, and L is an integral scale (characterizing the larger-scale structure), they argued that the vortical modes appear when F becomes small, i. e., when the influence of stratification becomes strong. Furthermore, they offered scaling arguments to explain why the flow should consist of both vortical modes and internal waves when F becomes small, and gave a mathematical decomposition of the velocity field into wave and vortical mode components. Subsequently Lilly (1983) extended these ideas to large-scale geophysical flows (with rotation), in which case the small parameter is a Rossby number. Lilly suggested that vortical modes might explain recent atmospheric spectral data (Gage, 1979) at intermediate scales. Müller (1988; see also Müller et al., 1988) has proposed that vortical modes explain the vertical fine-structure observed in the ocean internal wave field.

At the present time, little is known about the properties of vortical modes at the internal wave scales, for example, the vortical mode energy levels, spectral distributions, dynamical interactions, sources of energy, and sinks of energy. The goal of the work presented here is to elucidate the dynamic interactions of the vortical modes, and in particular: (i) to identify the principal interactions which affect vortical modes; (ii) to determine how weakly nonlinear theory must be modified to take into account these interactions; and (iii) to test the resulting theoretical predictions by comparison with results from direct numerical simulations of the fundamental equations of motion.

There are two key points upon which this theory rests. The first is that a small parameter exists. This is an assumption implicit in any wave theory, and implies that, in some sense, the nonlinearities are sufficiently weak that the mathematical concept of a wave is useful. This small parameter can be taken to be, for example, the ratio of the wave period to an advective time. For higher frequency motions this ratio is a Froude number, as utilized by Riley et al. (1981), whereas for lower frequency motions it is a Rossby number (Lilly, 1983).

The second point upon which the theory rests is the use of Ertel's potential vorticity in the decomposition of the flow fields (Müller et al., 1986). We define the internal wave field to be that part of the flow which does not contribute to the potential vorticity, and the vortical mode field to be that part of the flow which accounts for all of the potential vorticity. A principal rationale for using potential vorticity in the decomposition is that, as a dynamic quantity conserved following the motion in the absence of molecular diffusion, it cannot be associated with wave propagation. This decomposition reduces to that obtained when the governing equations are linearized (Müller et al., 1986), and Staquet and Riley (1989) have shown that this decomposition can be usefully extended to weakly nonlinear flows. It should be noted that, with this definition, the vortical mode

is the same as what is termed 'stratified turbulence' at intermediate atmospheric scales (Lilly, 1983), and the 'slow quasi-manifold' or the solution to the 'balanced equations' (McIntyre and Norton, 1991) at planetary scales. Furthermore Lilly (1983) has demonstrated that, at intermediate scales, the vortical mode satisfies the (nonlinear) geostrophic turbulence equations to lowest order.

In an effort to understand the vortical mode dynamics and their interaction with the internal wave field, we have recently examined the nonlinear interaction of simple monochromatic internal waves and vortical modes, in particular addressing wave-wave, wave-vortical mode, and vortical mode-vortical mode interactions (Lelong and Riley, 1991). In the next section, we briefly review the theory of Lelong and Riley and discuss their principal results. In the third section, results of direct numerical simulations of wave-vortical mode interactions are presented and compared with theory. In addition, simulations are presented for initial conditions consisting only of vortical modes. The behavior of the resulting flows is examined, and the results compared to the scaling analysis suggested by the theory. In the final section, the results are summarized and discussed.

2. RESULTS FROM PERTURBATION ANALYSIS

As mentioned in the previous section, vortical modes are visually observed in laboratory experiments when the Froude number based upon the energy containing range, i. e., $F = u'/NL$, is small. This Froude number can be interpreted as the ratio of two time scales: N^{-1}, the buoyancy period, and L/u', an advective time scale. In the experiments, the buoyancy period has thus become small compared to the advection time. F being small is also the usual requirement for linear or weakly nonlinear internal wave theory to be valid.

This suggests the use of multiple time scale perturbation analysis to analyze the dynamics of the internal waves and vortical modes. Lelong and Riley (1991) have carried out such an analysis, and begin by writing the velocity field $\vec{u}(\vec{x},t)$ in the form, as suggested by Lilly (1983):

$$\vec{u} = \nabla \wedge \psi \vec{i}_z + \nabla_H \phi + w \vec{i}_z \,. \tag{1}$$

[Note that this analysis has also been extended to rotating flows (Lelong, 1989) by starting with the decomposition suggested by Müller et al. (1986), which is an extension of the above expression to include the effects of rotation.] Here the subscript H denotes the horizontal component. To lowest order the stream function ψ (and an associated density field) completely determine the vortical mode, while the vertical velocity w (or ϕ, where $w = -\int \nabla^2 \phi \, dz$) specifies the internal wave field. When these expressions are substituted into the Navier-Stokes equations subject to the Boussinesq approximation, an equation for ψ, namely the vertical

vorticity equation, is obtained while an equation for w results from manipulation of the vertical and horizontal momentum equations. In addition, the equation for the buoyancy is required.

Velocities are nondimensionalized by u', lengths by L, and time by N^{-1}. (It is assumed that N is uniform in space.) All nondimensionalized dependent variables are expanded in power series of F, and terms of like powers of F are equated. Furthermore, a 'fast' time of the order of the buoyancy period and a 'slow' time of the order of the advective time are now both treated as independent variables. To lowest order (on the short time scale N^{-1}), linear theory is recovered. On this time scale, the vortical modes are steady. The equations at the next order give an $\mathcal{O}(F^2)$ correction to these solutions unless a resonance exists, in which case the form of the lowest order solutions has to be reassessed in order for the series expansion to remain valid. The elimination of secular terms from the second order equations yields the evolution of the solutions on the slow, advective time scale. The results depend on the specific problem considered, i. e., on the initial conditions. (Note that when this analysis is extended to the case with rotation, the small parameter is a Rossby number, and the theory proceeds in an analogous manner.)

Wave-Wave Interactions

For this case, the initial conditions are taken to be two linear internal waves and a linear vortical mode. The internal waves are arbitrarily oriented, while the vortical mode is assumed to be harmonic in space and is given in terms of a stream function as in Equation (1) above. The results of the analysis are an extension of resonant wave interaction theory in the presence of a vortical mode, and also the application of resonance theory to triads which are out of the vertical plane. Wave resonant triads are unaffected by the presence of the vortical mode, unless there exists a wave-vortical mode resonance (see the next subsection below). Furthermore, resonant triad analysis is readily extended to wave triads not lying in a vertical plane. The only deviation from the vertical plane case is that the gravity vector is effectively reduced by the cosine of the angle of the triad-containing plane with respect to the vertical.

The initial wave fields, being solutions to the linear wave equations, possess no potential vorticity. Since potential vorticity would be expected to be conserved for these weak wave interactions, the potential vorticity should remain zero. From the decomposition of the flow field discussed in the Introduction, no vortical mode would be expected to be generated. One key result of this theory, however, is that, for interactions out of a vertical plane, linearized potential vorticity can be produced at higher orders. This can lead to the erroneous conclusion that wave-wave interactions can be a source of vortical mode energy (e. g., Dong and Yeh, 1988). Potential vorticity is, however, exactly conserved at each order by such interactions so that no vortical mode energy is generated by such an interaction.

Another important related point is that, although resonances out of the vertical plane are as likely to occur as in-plane resonances, it is not clear that any of the present theories [e. g., McComas and Bretherton (1977)] include such interactions.

Wave-Vortical Mode Interactions

In this case the initial conditions consist of one internal wave (with wave number vector $\vec{\kappa}_1$ and frequency ω) and one vortical mode (with wave number vector $\vec{\kappa}_2$). The analysis predicts that a resonance can occur with a second internal wave (with wave number vector $\vec{\kappa}_3$) of the same frequency if the wave numbers satisfy

$$\vec{\kappa}_1 \pm \vec{\kappa}_2 \pm \vec{\kappa}_3 = 0\,. \tag{2}$$

The wave number vectors of the internal waves lie on a vertical cone whose surface is at angle $\theta = \cos^{-1}\omega$. (See Figure 1.) Note that this interaction could act to redistribute energy broadly in wave number space.

The nonlinear amplitude equations for this case can be solved analytically, and it is found that the role of the vortical mode is catalytic, i. e., it is needed for the interaction to occur, but it does not actively participate in the energy exchange. (This mechanism is reminiscent of the elastic scattering interaction discovered by Phillips, 1968, although the dynamics for the present case are significantly different.) The two wave modes exchange energy harmonically on the slow time scale, the vertical velocity being given by

$$w = \cos(\Gamma F t)\sin(\vec{\kappa}_1 \cdot \vec{x} - \omega t) - \sin(\Gamma F t)\cos(\vec{\kappa}_3 \cdot \vec{x} - \omega t)\,. \tag{3}$$

Here $2\pi/\Gamma F$ is the interaction period, where Γ is given by

$$\Gamma = \frac{B\kappa_1\kappa_3 \sin^2\theta \sin\Delta\gamma}{2}\left\{\cos^2\theta\cos\Delta\gamma + \sin^2\theta\right\}. \tag{4}$$

The parameter B is the amplitude of the vortical mode and $\Delta\gamma$ is the angle between the horizontal components of the wave number vectors of the two waves. Note that this same resonance exists in the case with rotation (Lelong, 1989).

Vortical Mode Interactions

In this final case, initial conditions consisting of two vortical modes have been considered. It is found that all interactions are resonant because the resonance condition imposed on the frequencies is identically satisfied, the frequency of any vortical mode being zero regardless of its wavenumber. As a consequence, the vortical mode equations are fully nonlinear on the vortical mode time scale (L/u'). It is found that, to lowest order, the vortical mode velocity, pressure, and density satisfy:

$$\frac{\partial}{\partial t}\vec{u}_H + \vec{u}_H \cdot \nabla \vec{u}_H = -\nabla p + R^{-1}\nabla^2 \vec{u}_H \tag{5a}$$

$$\nabla \cdot \vec{u}_H = 0 \tag{5b}$$

$$0 = -\frac{\partial p}{\partial z} - \rho \tag{5c}$$

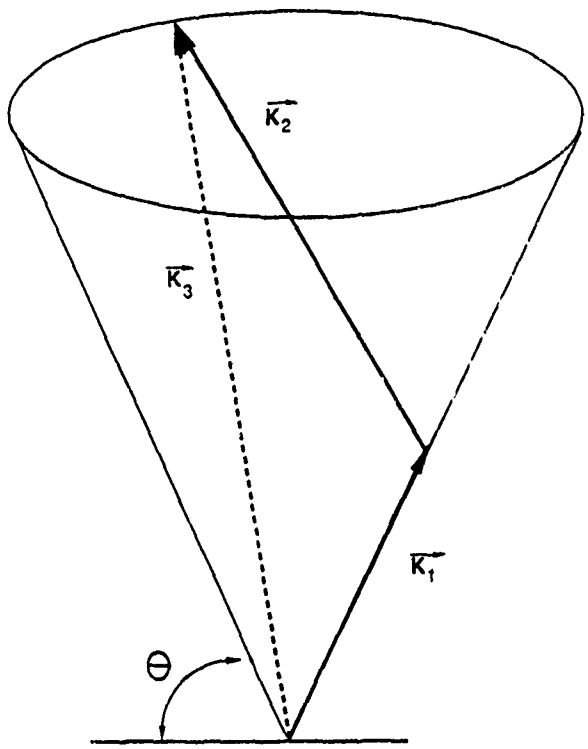


Figure 1. Wave/vortical mode resonant triad.

Here $R = u'L/\nu$ is the Reynolds number. Note that these equations describe horizontal motion in each horizontal plane, but with vertical variation retained. This result was suggested by Riley et al. (1981) using heuristic arguments. In addition to the conclusion that the vortical modes satisfy Equation (5), it was found that the vortical modes excite internal waves at higher order, with the amplitude of the internal waves scaling as F (or the energy as F^2). Note that, with rotation, the lowest order equations are equivalent to the (nonlinear) geostrophic turbulence equations (Lilly, 1983).

3. SIMULATION RESULTS

In this section we present results of direct numerical simulations and comparisons with theory for both wave-vortical mode interactions and vortical mode self-interactions. The simulations employ pseudo-spectral numerical methods with leap-frog time-stepping (smoothed every 25 time steps). The wave-vortical mode simulations were performed on $32 \times 32 \times 32$ point computational grids, while the vortical mode interactions used $64 \times 64 \times 64$ point grids. All calculations were performed on an Ardent Titan Mini-Supercomputer Server.

Wave-Vortical Mode Interactions

As discussed in the previous section, a resonance was found to exist between a vortical mode and two internal waves if the two waves are of the same frequency and the three wave number vectors form a triangle on the surface of a vertically-oriented cone. In order to examine these interactions we have performed direct numerical simulations for a number of different resonance conditions, and for a variety of different Froude numbers. The following case is typical of the results from these simulations, and represents the interaction of an internal wave with a horizontally shearing current. The latter is a special case of a vortical mode. The initial vertical velocity of the internal wave satisfies

$$w(\vec{x},0) = \cos(\vec{\kappa}_1 \cdot \vec{x}),$$

while the vortical mode is initially given by

$$\vec{u}(\vec{x},0) = \left(\kappa_{2_2} B \sin(\vec{\kappa}_2 \cdot x), -\kappa_{2_1} B \sin(\vec{\kappa}_2 \cdot x), 0\right).$$

We choose

$$\vec{\kappa}_1 = (2,0,2),$$

giving a wave propagating at 45° to the horizontal. The vortical mode wave number is

$$\vec{\kappa}_2 = (-2,2,0),$$

giving a horizontal current uniform in x_3, but which is sinusoidal in the horizontal with wave number magnitude $|\vec{\kappa}_2| = 2\sqrt{2}$. Based upon the resonance conditions [Equation (2)], the resonant wave should be found at

$$\vec{\kappa}_3 = (0,2,2).$$

The parameter B is taken to be 2.0, and the Froude number F to be 0.02. From the scaling analysis, this implies that the wave frequency ω in the simulations is $2\pi/F$, while the interaction frequency Γ is 2.0. The perturbation theory indicates that both the kinetic and the potential energy in the initial internal wave ($\vec{\kappa}_1$) should vary with time as $0.5\cos^2(2t)$, and that in the forced wave ($\vec{\kappa}_3$) as $0.5\sin^2(2t)$.

Figure 2 contains a plot of the potential energy at the forced wave number $\vec{\kappa}_3$ versus time taken from the simulations. Also plotted is the theoretical prediction. We see that the simulations follow the theory fairly closely for a time of about 2 to 2.5, at which point the potential energy in this wave mode begins to oscillate (near the buoyancy frequency), and the results begin to deviate significantly from the theory. More insight into this problem is given by examining the kinetic energy of each mode for this case, as shown in Figure 3. According to the theory, the kinetic energy in the vortical mode should remain uniform in time. Again we see that the simulation results follow the predicted oscillation in the wave kinetic energy up to a time of about 2 to 2.5, at which point the computed solutions rapidly diverge from the theoretical predictions. Furthermore, the kinetic energy in the vortical mode remains approximately constant up to this time, and it then begins to exhibit a large deviation from its initial value.

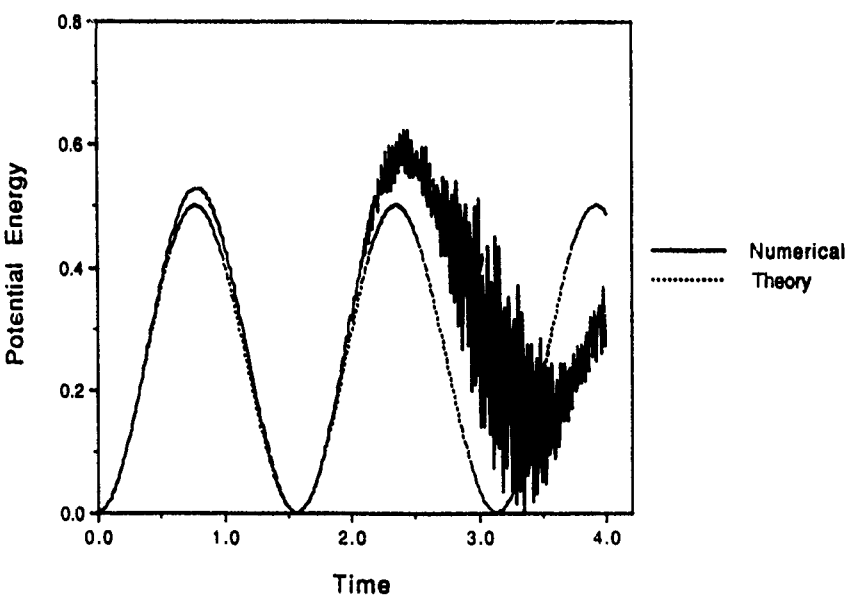


Figure 2. Potential energy in the forced wave $\vec{\kappa}_3$; $F_t = 0.02$.

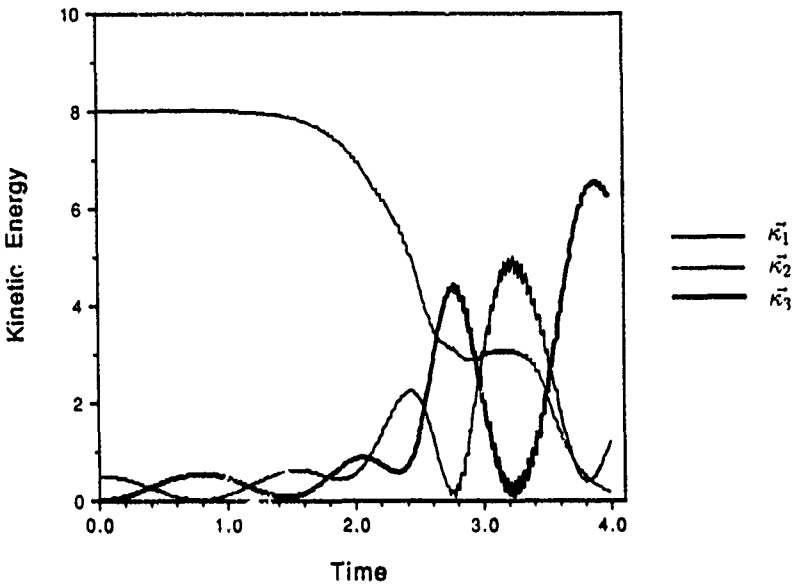


Figure 3. Kinetic energy in all three modes; $F_t = 0.02$.

Visual analysis of the results of this case determined that the horizontally shearing motion was subject to a shear instability which arose dramatically at about a time of 2. Furthermore this instability transferred energy mainly into the vortical mode components at wave numbers $\vec{\kappa}_1$ and $\vec{\kappa}_3$. Prior to this instability, however, the perturbation theory gave accurate predictions of the interactions.

Vortical Mode Interactions

In order to examine nonlinear vortical mode interactions we have considered initial conditions consisting of a sum of spatially harmonic vortical modes. The initial velocity field is given by

$$\vec{u}(\vec{x},0) = \cos x_3(\cos x_1 \sin x_2, -\sin x_1 \cos x_2, 0)\,, \tag{6}$$

while the initial perturbation density field is identically zero. For the nonstratified case this initial condition defines the Taylor-Green problem (Taylor and Green, 1937), with the velocity field oriented such that it is initially horizontal. This problem has received much attention in the literature (e. g., Orszag, 1971). The perturbation analysis predicts that as $F \rightarrow 0$, the flow field should satisfy Equation (5) to lowest order. The exact solution to these equations, satisfying the above initial conditions, is

$$\vec{u}(\vec{x},0) = e^{(-3R^{-1}t)} \cos x_3(\cos x_1 \sin x_2, -\sin x_1 \cos x_2, 0)\,. \tag{7}$$

The nonlinear terms in the momentum equation [Equation (5a)] are exactly balanced by the pressure gradient, so that the equations become linear. The streamline pattern remains independent of time as the velocity field decays due to viscosity. This solution is an extension of a well-known solution for the two-dimensional Taylor-Green problem (see, e. g., Staquet, 1985).

We performed a series of simulations for a number of different Froude numbers, using the above initial conditions for each case. For all the cases presented the initial Reynolds number R was fixed at 200. Figure 4 gives a plot of the volume integrated kinetic energy of the horizontal velocity (normalized by its initial value) for these different cases. The case with $F = \infty$ corresponds to the Taylor-Green problem, and agrees with previously published simulation results. We see that increasing the stratification decreases the decay rate, as might be expected.

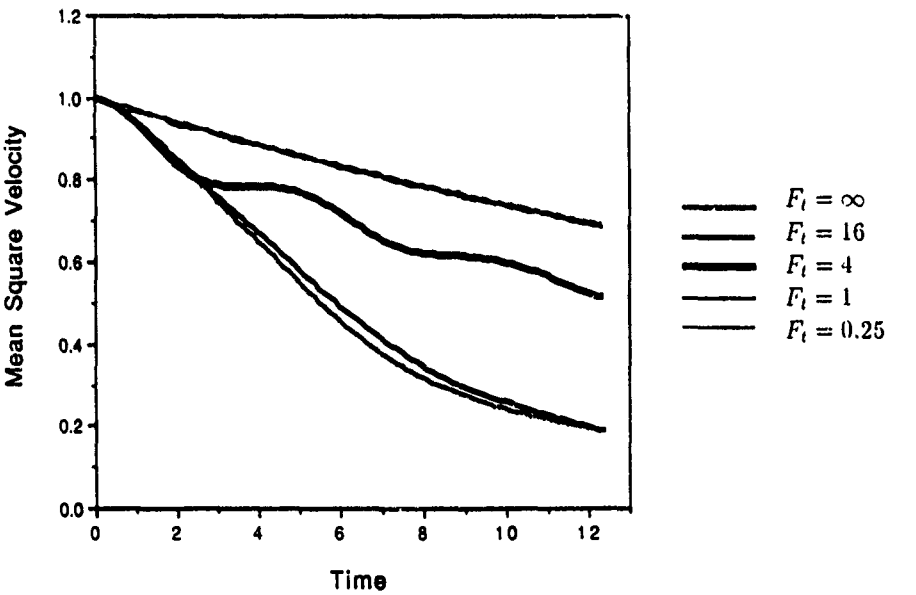


Figure 4. Taylor-Green problem: normalized kinetic energy of horizontal velocity for various Froude numbers.

Furthermore as F becomes small, roughly $F \leq 1.0$, the computed solutions coincide, as expected from the perturbation theory. Figure 5 gives a plot of the same data on an expanded scale, for the cases with stronger stratification. Also plotted is the result from the theoretical solution, Equation (7). We see that the solution to the predicted asymptotic equations agrees very well with the simulation results for $F \leq 1.0$.

Figure 6 contains plots of the wave energy for these same cases. The wave energy contains the potential energy plus the wave part of the kinetic energy, as defined by Equation (1). Note that the wave energy decreases significantly as F is

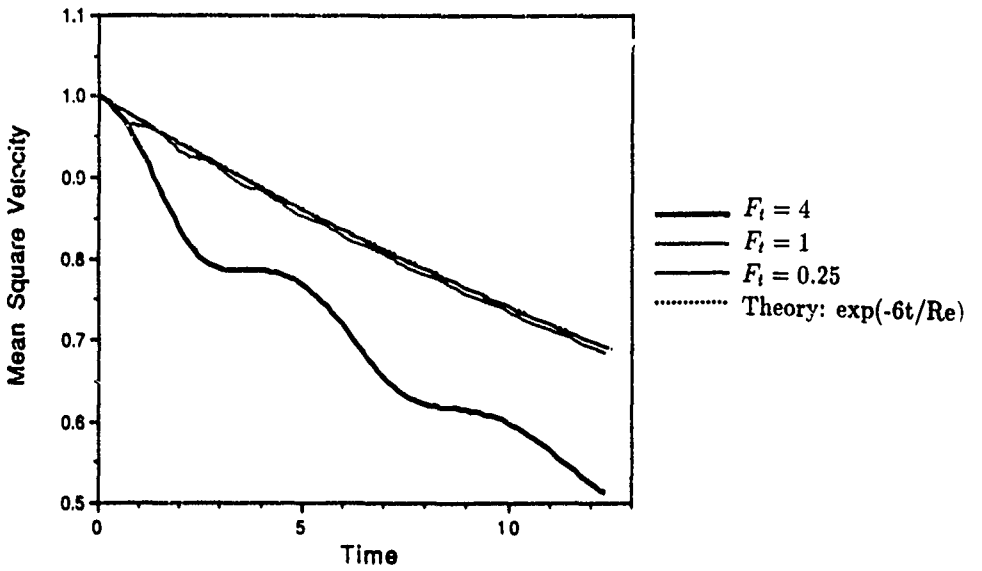


Figure 5. Taylor-Green problem: normalized kinetic energy of horizontal velocity for several Froude numbers: expanded scale.

decreased. When these results are replotted, scaled with F^{-2}, as suggested by the theory, then the wave energy collapses well for $F \leq 4.0$ (Figure 7). Therefore, for $F \leq 1.0$, the perturbation analysis is consistent with the results for both the horizontal kinetic energy, which consists mainly of the vortical mode, and the wave energy.

This present case is somewhat degenerate because, as F becomes small, the nonlinear and horizontal pressure gradient terms come into balance, leading to the simple viscous decay given by Equation (7). As mentioned, this is similar to the result for the two-dimensional Taylor-Green problem. It is well-known for the two-dimensional problem, however, that the solution is unstable to subharmonic perturbations. The length scales of the flow continually grow larger as energy is nonlinearly transferred to lower wave numbers. Therefore we also performed a series of simulations for different F using the above initial conditions, Equation (6), but with white noise added. Figure 8 gives a sequence of constant contours of ψ in a horizontal plane for the case $F = 1.0$. At this value of stratification these contours approximate streamlines in a horizontal plane. We see that the flow develops nonlinearly in time, as energy is continually transferred to larger scales, reminiscent of the two-dimensional problem. Figure 9 contains plots of ψ in two different horizontal planes at a later time. We see that the two layers have become decoupled, the flow in the two planes being very different. Finally, Figure 10 has plots of the total energy versus time for the different cases computed. Again the results converge for $F \leq 1.0$, consistent with the predictions of the perturbation theory.

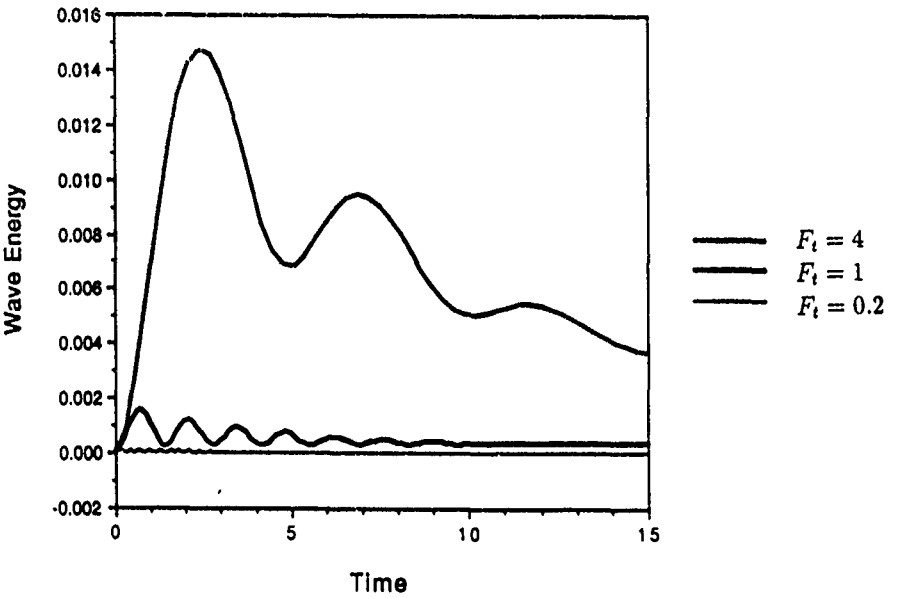


Figure 6. Taylor-Green problem with white noise: wave energy for three different Froude numbers.

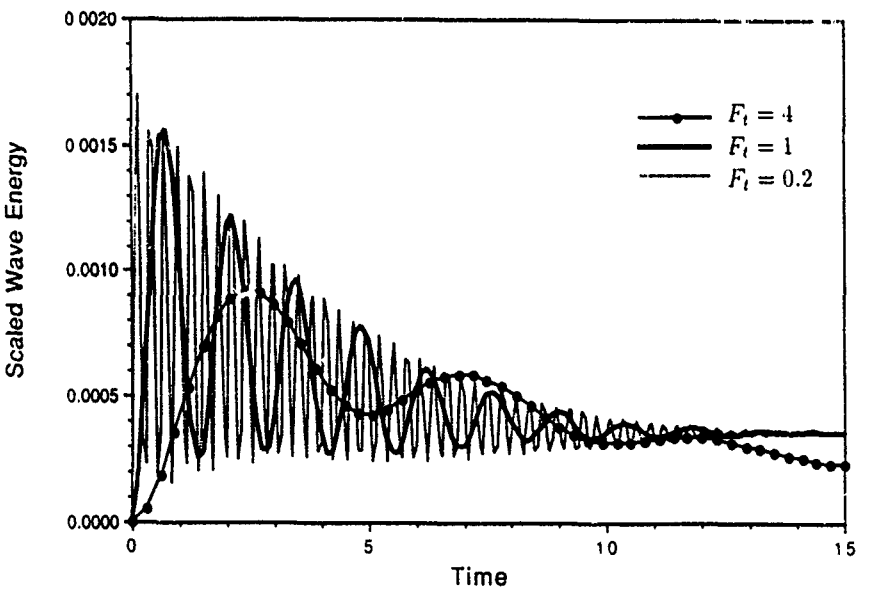


Figure 7. Wave energy scaled by $1/F_t^2$ for three different Froude numbers.

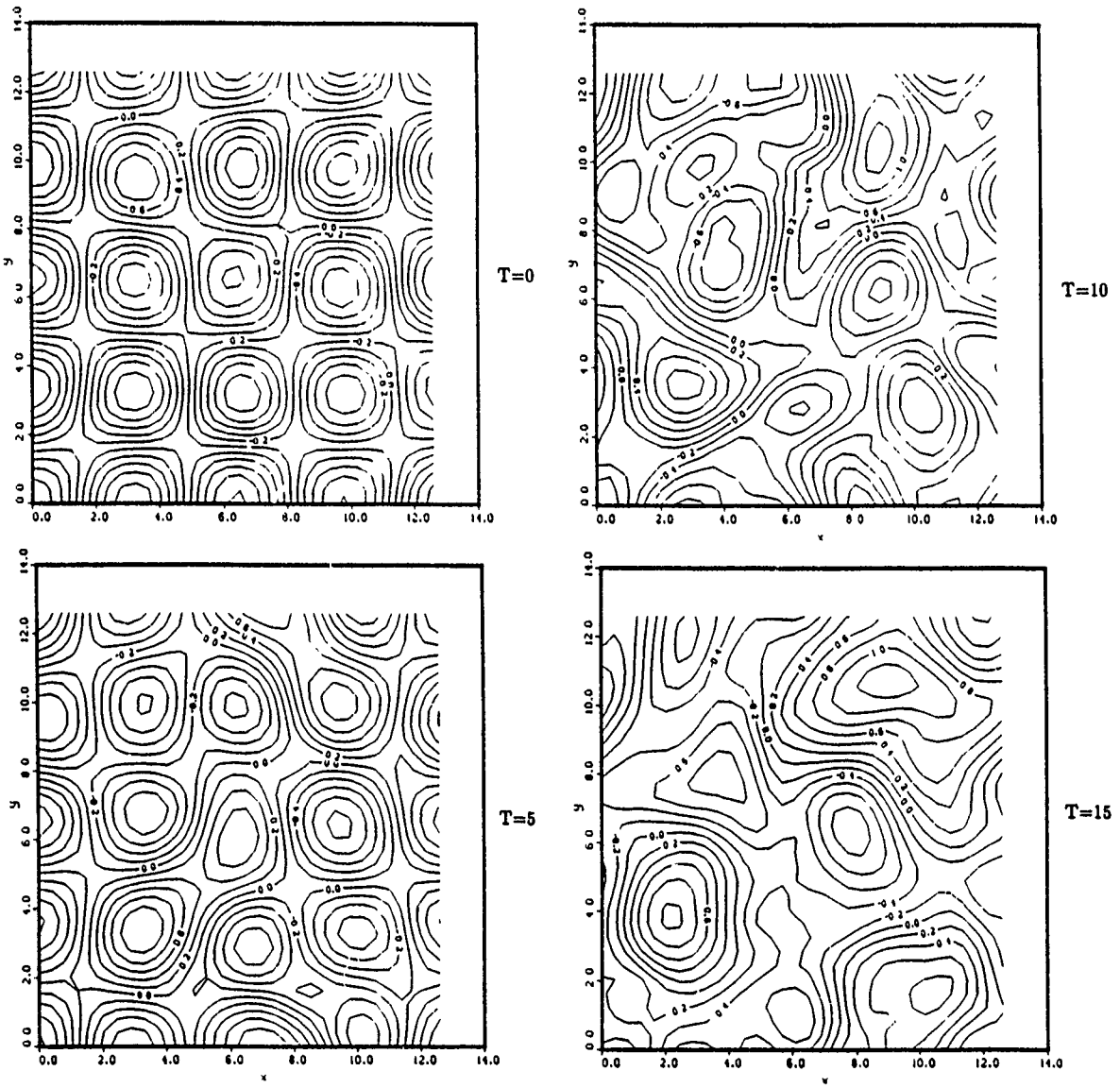


Figure 8. Sequence of constant contours of the stream function in a horizontal plane for $F_t = 0.2$ at T=0, T=5, T=10, and T=15.

4. CONCLUSIONS AND DISCUSSION

We have reported on theoretical analysis and direct numerical simulations of internal wave and vortical mode interactions in strongly stratified flows. The theoretical work utilizes multiple-scale expansions assuming the existence of a small parameter (e. g., the Froude number). Furthermore, a decomposition of the flow field into internal waves and vortical modes that is based upon Ertel's potential vorticity is employed. The objectives are: (i) to identify the principal interactions

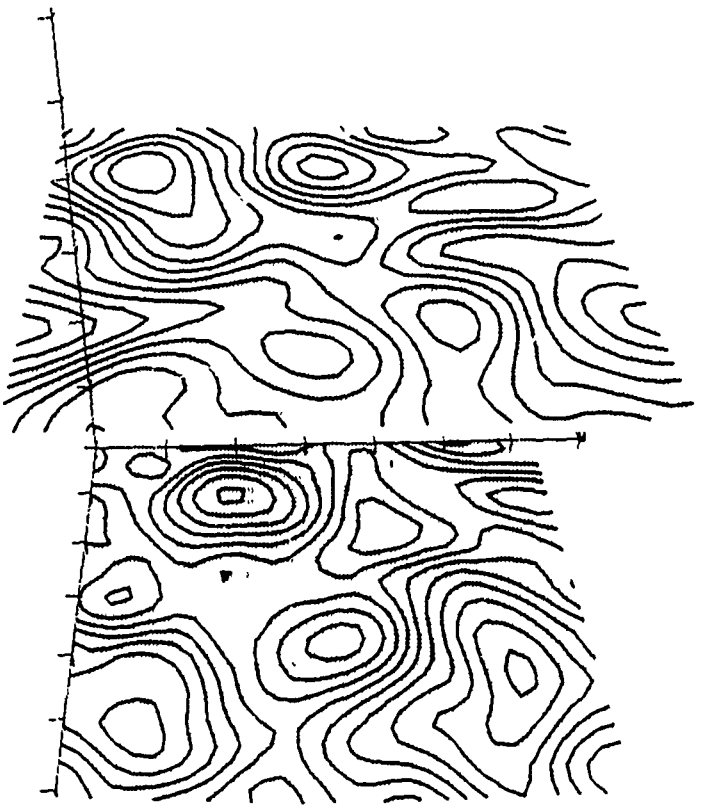


Figure 9. Constant contours of the stream function in two different horizontal planes for $F_t = 0.2$ at T=15.

which affect vortical modes ; (ii) to determine how weakly nonlinear theory must be modified to take into account these interactions; and (iii) to test the resulting theoretical predictions by comparisons with results of direct numerical simulations of the fundamental equations of motion.

Wave-wave interactions are considered theoretically, and previous results on wave resonance are reproduced. Resonances out of the vertical plane are also found, and an apparently erroneous conclusion regarding the excitation of vortical modes by internal wave interactions is explained. Another interaction examined is the resonance of two internal waves and a vortical mode. In this interaction the vortical mode plays the role of a catalyst, not exchanging energy with the waves, but being necessary for the interaction to occur. This interaction could have a significant effect on the development of an internal wave field. By comparing theoretical results with those from direct numerical simulations, we have found that the perturbation analysis predicts the interactions very well if the Froude number is small enough, approximately $F \leq 1.0$. From the numerical simulations we have also found, however, that the vortical modes considered were highly unstable, and ultimately experienced breakdown as the fluctuations in the flow increased. In all cases considered for this interaction the vortical mode consisted at least partially

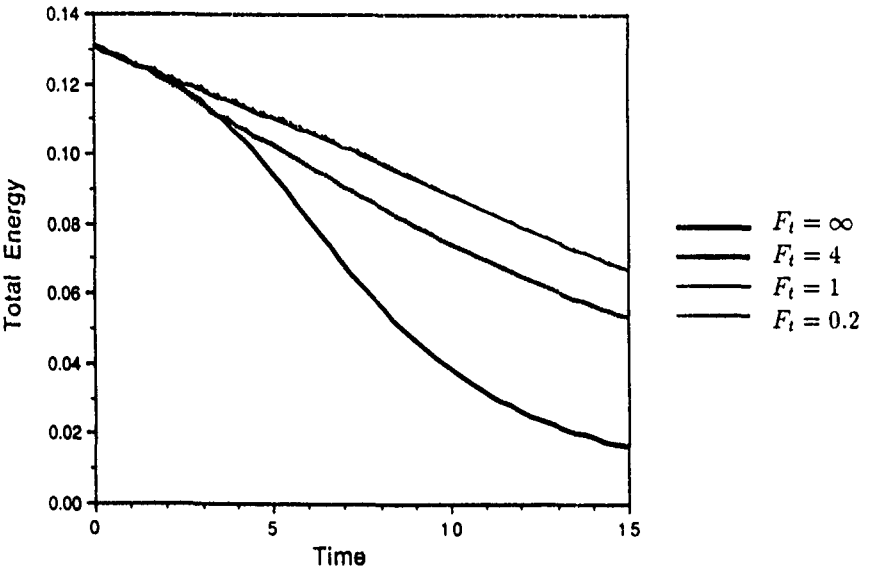


Figure 10. Taylor-Green problem: total energy for several Froude numbers.

of horizontally varying currents with multiple inflection points. These flows satisfy both Rayleigh's and Fjortoft's necessary conditions for instability (Drazin and Reid, 1981). Thus, it is not surprising that they were unstable. The instabilities appear to feed energy into the vortical modes at the same wave number as the internal waves. Once the unstable fluctuations grew to an appreciable amplitude, then the results of the simulations deviated strongly from the perturbation theory predictions.

Vortical mode self-interactions are also considered, and it is found that the lowest order governing equations are fully nonlinear. As a test case the problem of Taylor and Green is considered. An exact solution to the perturbation equations was found for this case, an extension of a well-known two-dimensional result. The results of the numerical simulations agreed well with this solution for approximately $F \leq 1.0$. Furthermore, the scaling of the potential energy predicted by the theory was also consistent with the simulation results when this condition was satisfied. When white noise was added to the initial Taylor-Green field, for small F the flow exhibited subharmonic instabilities similar to those observed for the two-dimensional Taylor-Green case. Again the perturbation equations as well as the predicted scaling of the potential energy were found to hold for approximately $F \leq 1.0$.

In both cases the simulations emphasize the importance of nonlinearity in the dynamics of the vortical modes. Furthermore, especially in the Taylor-Greene case, it is clear that upscale transfer of energy to larger horizontal scales occurs, a phenomena

suggested by Lilly (1983), and observed in the laboratory (Lin and Pao, 1979) and in other numerical simulations (Riley et al., 1981; Herring and Métais, 1989; Métais and Herring, 1989). It was also clear that the different horizontal layers tend to become uncorrelated, since the horizontal flow dynamics differ in each layer. How the layers remain weakly coupled, whether through viscous effects or shear instabilities, must surely depend strongly on the Reynolds number.

If vortical modes exist on geophysical scales then their spatial characteristics and dynamical properties become of interest. Clearly, if vortical modes somewhat similar to those considered here were prevalent, then the vortical mode field would be rapidly evolving in time and probably continually be subject to shear instabilities. However, if vortical modes existed as two-dimensionally (in the horizontal) stable flows, e.g., in stable rotation satisfying Rayleigh's circulation criterion (LeLong, 1989; McWilliams, 1985), then perhaps they would be more persistent dynamically.

REFERENCES

Dong, B., and K. C. Yeh. 1988. "Resonant and Nonresonant Wave-Wave Interactions in an Isothermal Atmosphere", *J. Geophys. Res.*, Vol. 93, No. D4, pp. 3729-3744.

Drazin, P. G., and W. H. Reid. 1981. *Hydrodynamic Stability.* Cambridge University Press.

Gage, K. S. 1979. "Evidence for a $k^{-5/3}$ Law Inertial Range in Mesoscale Two-Dimensional Turbulence", *J. Atmos. Sci.*, Vol. 36, pp. 1950-1954.

Herring, J. R., and O. Métais. 1989. "Numerical Experiments in Forced Stably-Stratified Turbulence", *J. Fluid Mech.*, Vol. 202, pp. 97-115.

Hopfinger, E. J. 1987. "Turbulence in Stratified Fluids: A Review", *J. Geophys. Res.*, Vol. 92, No. C5, pp. 5287-5304.

Lelong, M.-P. G. 1989. *Weakly-Nonlinear Internal Wave/Vortical Mode Interactions in Stably-Stratified Flows*, Ph. D. Thesis, University of Washington.

Lelong, M.-P. G., and J. J. Riley. 1991. "Internal Wave/Vortical Mode Interactions in a Stably-Stratified Flow", to be published in the *J. Fluid Mech.*

Lilly, D. K. 1983. "Stratified Turbulence and the Mesoscale Variability of the Atmosphere", *J. Atmos. Sci.*, Vol. 40, pp. 749-761.

Lin, J.-T., and Y.-H. Pao. 1979. "Wakes in Stratified Fluids", *Ann. Rev. Fluid Mech.*, Vol. 11, pp. 317-338.

Liu, H.-T. 1980. Private communication.

McComas, C. H., and F. P. Bretherton. 1977. "Resonant Interaction of Ocean Internal Waves", *J. Geophys. Res.*, Vol. 82, No. 9, pp. 1397-1412.

McIntyre, M. E., and W. A. Norton. 1991. "Potential-Vorticity Inversion on a Hemisphere", submitted to *J. Atmos. Sci.*

McWilliams, J. C. 1985. "Submesoscale Coherent Vortices in the Ocean", *Rev. Geophys.*, Vol. 23, pp. 165-182.

Métais, O., and J. R. Herring. 1989. "Numerical Simulations of Freely Evolving Turbulence in Stably-Stratified Fluids", *J. Fluid Mech.*, Vol. 202, pp. 117-148.

Müller, P. 1988. "Vortical Motions", in *Small-Scale Turbulence and Mixing in the Ocean*, ed. by J. C. J. Nihoul and B. M. Jamart, Elsevier, pp. 285-302.

Müller, P., G. Holloway, F. Henyey, and N. Pomphrey. 1986. "Nonlinear Interactions Among Internal Gravity Waves", *Rev. Geophys.*, Vol. 24, pp. 493-536.

Müller, P., R.-C. Lien, and R. Williams. 1988. "Estimates of Potential Vorticity at Small Scales in the Ocean", *J. Phys. Ocean.*, Vol. 18, No. 3, pp. 401-416.

Orszag, S. A. 1971. "Numerical Simulation of Incompressible Flows Within Simple Boundaries: Accuracy", *J. Fluid Mech.*, Vol. 49, Part 1, pp. 75-112.

Phillips, O. M. 1968. "The Interaction Trapping of Internal Gravity Waves", *J. Fluid Mech.*, Vol. 34, pp. 407-416.

Riley, J. J., R. W. Metcalfe, and M. A. Weissman. 1981. "Direct Numerical Simulations of Homogeneous Turbulence in Stably-Stratified Fluids", in *Proc. AIP Conf. on Nonlinear Properties of Internal Waves*, ed. by B. J. West, pp. 79-112.

Staquet, C. 1985. *Étude Numérique de la Turbulence Bidimensionnelle Homogène et Cisaillée.* Thèse Docteur de l'Institut National Polytechnique de Grenoble.

Staquet, C., and J. J. Riley. 1989. "On the Velocity Field Associated with Potential Vorticity", *Dyn. Atmos. Oceans*, Vol. 14, pp. 93-123.

Taylor, G. I., and A. E. Green. 1937. "Mechanisms of the Production of Small Eddies From Large Ones", *Proc. Roy. Soc.*, Vol. A 158, pp. 499-521.

van Heijst, G. J. F., and J. B. Flór. 1989. "Dipole Formation and Collisions in a Stratified Fluid", *Nature*, Vol. 340, No. 6230, pp. 212-215.

Internal Wave-Wave Resonance Theory: Fundamentals and Limitations

Eric Hirst
Applied Physics Laboratory
University of Washington
Mailstop HN 10
Seattle, Wa. 98115

ABSTRACT

First, two terms of a resonant transfer integral are derived for a simple one-dimensional analog to the equations of fluid motion. The purpose of this rather academic exercise is to more clearly illustrate the assumptions of the resonant interaction approximation. Second, it is shown that, in much of the region of phase space where nonlinear interactions are usually considered weak enough for the theory to be applicable, nonlinear transfers due to bottom scattering and other mechanisms cannot be ignored.

INTRODUCTION

Weakly nonlinear resonant interactions have been studied extensively as a model for predicting the temporal evolution of various broad-band wave spectra. Hasselmann (1962, 1966) derived a general wave-wave resonance theory, which I will refer to as the resonant interaction approximation (RIA). RIA has been used with some success on the surface wave interaction problem (WAMDI, 1988), but it is still unclear how applicable it is to internal waves. RIA-based internal wave spectral evolution equations were derived and evaluated by Olbers (1976), McComas and Bretherton (1977, henceforth MB), and Pomphrey, et al. (1980). Results of these calculations have yet to be verified for any realistic ocean internal wave field, however, and the validity of RIA for large portions of the internal wave spectrum has been called into question by Holloway (1980, 1982), Henyey et al. (1986), and Müller, et al. (1986).

The initial portion of this paper will attempt to highlight some of the basic points of an RIA derivation using a 1-D analog. The somewhat ad hoc derivation will follow Hasselmann (1966); a more careful derivation using the method of multiple time scales is given by Benney and Saffman (1966), who also begin with a one dimensional example for readability. It must be stressed that, since RIA reduces the dimension of any problem by one, the resonant "surface" for this example will just be a set of three points. The purpose of this exercise is strictly to clarify the mechanics of an RIA derivation, and not to offer a new way of solving one dimensional problems.

In the past, criticisms of RIA have focused on mapping out regions of phase space where RIA predicts interactions so strong that they violate the theory's assumptions of "weakness." There are regions of phase space, however, where RIA predicts interactions which are so weak that they will be overshadowed by other nonlinear mechanisms, most notably bottom scattering. The second portion of this paper attempts to map the regions in wavenumber and frequency space where resonant interaction theory either violates its own assumptions or predicts interactions so weak that they will be overshadowed by other mechanisms.

DERIVATION OF A SPECTRAL EVOLUTION EQUATION

Derivation of a resonant spectral evolution equation directly from the equations of fluid motion is, among other things, a monumental exercise in algebra. Fortunately, the critical steps may be highlighted by deriving a similar evolution equation for a one dimensional analog. The somewhat modified Korteveg-DeVries (KdV) equation,

$$\frac{\partial \phi}{\partial t} + \frac{\partial^3 \phi}{\partial x^3} + \varepsilon \phi \frac{\partial \phi}{\partial x} = 0, \tag{1}$$

has many of the same basic properties as the equations of fluid motion. It is first order in time, non-dissipative, dispersive, and contains a quadratic nonlinearity. The dependent variable $\phi(x,t)$ is assumed to be of order one, and ε is a small parameter, analogous to a Rossby number for low frequency internal waves.

The first step in our streamlined derivation will be to expand (1) in a Fourier series. If we assume a periodic solution in a domain $-L \le x \le L$, we may write

$$\phi(x,t) = \sum_n \hat{\phi}(n,t) e^{\frac{in\pi x}{L}}, \tag{2}$$

where the summation is over all integer mode numbers, and

$$\hat{\phi}(n,t) = \frac{1}{2L} \int_{-L}^{L} dx \, \phi(x,t) e^{\frac{-in\pi x}{L}}. \tag{3}$$

To handle the nonlinear term, we make use of the convolution rule,

$$\frac{1}{2L} \int_{-L}^{L} dx \, f(x,t) g(x,t) e^{\frac{-in\pi x}{L}} = \sum_{n_1} \sum_{n_2} \hat{f}(n_1,t) \hat{g}(n_2,t) \delta_{n, n_1+n_2}, \tag{4}$$

where δ_{n, n_1+n_2} is the Kronecker delta. The KdV equation may then be written

$$\frac{\partial \hat{\phi}}{\partial t} - ik^3 \hat{\phi} = \frac{-ik\varepsilon}{2} \sum_{n_1} \sum_{n_2} \delta_{n, n_1+n_2} \hat{\phi}_1 \hat{\phi}_2, \tag{5}$$

where $k = n\pi/L$ is the wavenumber, and $\hat{\phi}_j \equiv \hat{\phi}(n_j,t)$. Note that the factor $k/2 = (n_1+n_2)\pi/2L$ appears due to the Kronecker delta and the symmetry of the summation.

It would be possible to find a numerical scheme to evaluate (5) directly for a finite number of modes. The convolution sum could be handled by a fast Fourier transform algorithm, and a suitable procedure could be used to integrate in time. This technique is simply a pseudo-spectral

Fourier-Galerkin scheme, however, and has the same basic strengths and weaknesses as any of the competing numerical methods. To take advantage of the supposed weakness of the nonlinear term, we instead make a perturbation expansion about a first order linear solution.

Expanding our solution in powers of ε, we may write

$$\hat{\phi} = {}_1\hat{\phi} + {}_2\hat{\phi}\varepsilon + {}_3\hat{\phi}\varepsilon^2 + \cdots, \tag{6}$$

where ${}_j\hat{\phi} = O(1)$. The linear solution is then

$${}_1\hat{\phi} = A(n,t=0)e^{i\omega t}, \tag{7}$$

where $\omega \equiv k^3$ is the intrinsic frequency, and the time t is measured from the most recent time the spectrum was measured or calculated. The $A(n,0)$'s are chosen such that ${}_1\hat{\phi} = A = \hat{\phi}$ at $t = 0$. The second order solution will then be obtained from

$$\frac{\partial {}_2\hat{\phi}}{\partial t} - i\omega\, {}_2\hat{\phi} = \frac{-ik}{2}\sum_{n_1}\sum_{n_2}\delta_{k,k_1+k_2}A_1A_2e^{i(\omega_1+\omega_2)t}, \tag{8}$$

where the subscripts on amplitudes A_j and frequencies ω_j denote dependence on mode number $n_j = Lk_j/\pi$. (Unsubscripted forms will generally be dependent on mode number n.)

Equation (8) has the solution

$${}_2\hat{\phi} = \frac{-ik}{2}e^{i\omega t}\int_0^t d\tau \sum_{n_1}\sum_{n_2}\delta_{k,k_1+k_2}A_1A_2e^{i(\omega_1+\omega_2-\omega)\tau}. \tag{9}$$

This may be integrated in τ to obtain

$${}_2\hat{\phi} = \frac{-ik}{2}e^{i\omega t}\sum_{n_1}\sum_{n_2}\delta_{k,k_1+k_2}A_1A_2\Delta(\Omega,t), \tag{10}$$

where $\Omega \equiv \omega_1+\omega_2-\omega$, and

$$\Delta(\Omega,t) \equiv \frac{e^{i\Omega t}-1}{i\Omega}. \tag{11}$$

A similar but slightly more complicated expression may be obtained for ${}_3\phi$. While the third order quantities must be included in the full resonant transfer equation for both KdV and internal wave systems, they will be ignored in this discussion.

At this point, we should make some observations on just what "linear solution" we are expanding about. Recall that the purpose of RIA is to predict the temporal evolution of a given spectrum. Presumably at $t = 0$, we know the spectrum and the $A(N)$'s are chosen such that the first older solution ${}_1\phi$ is the complete solution ϕ. It is clear from (9) that ${}_2\phi = 0$ at $t = 0$; the same applies to higher order terms. But for larger times, the spectrum will evolve, and the $A(k)$'s will periodically have to be updated (and t reset to zero in my notation) for the series (6) to correctly approximate the true solution.

In an actual RIA calculation, phase information is sacrificed and $A(k)$ is not actually calculated. Instead we look for an equation governing the temporal evolution of an action density $a(k,t)$, defined as

$$\begin{aligned} a(k,t) &\equiv \frac{1}{\omega}<\hat{\phi}\hat{\phi}^*> \\ &= \frac{1}{\omega}<{}_1\hat{\phi}_1\hat{\phi}^*> \\ &+ \frac{\varepsilon}{\omega}\left\{<{}_1\hat{\phi}_2\hat{\phi}^*> + <{}_2\hat{\phi}_1\hat{\phi}^*>\right\} \\ &+ \frac{\varepsilon^2}{\omega}\left\{<{}_1\hat{\phi}_3\hat{\phi}^*> + <{}_2\hat{\phi}_2\hat{\phi}^*> + <{}_3\hat{\phi}_1\hat{\phi}^*>\right\} + \cdots, \end{aligned} \tag{12}$$

where the brackets represent ensemble averages and asterices indicate complex conjugates. Our choice of first order solution ${}_1\phi$ implies that the lowest (second) order action term $\omega^{-1}<{}_1\hat{\phi}_1\hat{\phi}^*>$ will be constant. If we can assume that all amplitudes $A(k)$ are uncorrelated and Gaussian, then the third order terms, with their associated triple correlations, vanish identically. It must be recognized that strong interactions in even an initially Gaussian wave field might set up correlations between interacting wave modes. In this case, the third order (cumulant) terms may not be neglected, and the theory will fail. If our Gaussian assumption is valid, we have

$$\frac{\partial a}{\partial t} = \frac{\partial}{\partial t}\frac{\varepsilon^2}{\omega}\left\{<{}_1\hat{\phi}_3\hat{\phi}^*> + <{}_2\hat{\phi}_2\hat{\phi}^*> + <{}_3\hat{\phi}_1\hat{\phi}^*>\right\} + \cdots. \tag{13}$$

To perform the averaging on the middle term, it is useful to first write equation (10) as a single sum without a Kronecker delta.

$${}_2\hat{\phi} = \frac{-ik}{2}e^{i\omega t}\sum_{n_1} A(n_1)A(n-n_1)\Delta_1, \tag{14}$$

where $\Delta_j \equiv \Delta(\omega(k_j)+\omega(k-k_j)-\omega,t)$. We then have

$$<{}_2\hat{\phi}_2\hat{\phi}^*> = \frac{k^2}{4}\sum_{n_1}\sum_{n_2}<A(n_1)A(n-n_1)A^*(n_2)A^*(n-n_2)>\Delta_1\Delta_2^*. \tag{15}$$

Again assuming that the A's are Gaussian random variables, we split up the fourth order moment into second order moments to obtain

$$<_2\hat{\phi}_2\hat{\phi}^*> = \frac{k^2}{4}\sum_{n_1}\sum_{n_2}\Big\{<A(n_1)A(n-n_1)><A^*(n_2)A^*(n-n_2)> + <A(n_1)A^*(n-n_2)><A(n-n_1)A^*(n_2)> + <A(n_1)A^*(n_2)><A(n-n_1)A^*(n-n_2)>\Big\}\Delta_1\Delta_2^*. \tag{16}$$

Because $<A(n_1)A(n_2)>$ vanishes for $n_1 \neq -n_2$ ($A^*(n)=A(-n)$), each of the second order moments in the first term vanishes for $n \neq 0$. Because of the k^2 factor in front of the integral, we may then drop this term.

The second term vanishes except when $n_1 + n_2 = n$. If we replace the summation variable n_2 with n_2+n, we find that the third term vanishes except when $n = n_1 - n_2$. We may then insert and make use of two new Kronecker deltas to write

$$<_2\hat{\phi}_2\hat{\phi}^*> = \frac{k^2}{4}\sum_{n_1}\sum_{n_2}\delta_{k,k_1+k_2}|\Delta(\omega_1+\omega_2-\omega,t)|^2 a_1 a_2 \omega_1\omega_2 + \delta_{k,k_1-k_2}|\Delta(\omega_1-\omega_2-\omega,t)|^2 a_1 a_2 \omega_1\omega_2. \tag{17}$$

Here we have made the substitution

$$a(n,t) \approx \frac{1}{\omega}<_1\hat{\phi}_1\hat{\phi}^*> = \frac{1}{\omega}<A(n)A^*(n)>, \tag{18}$$

valid for short amounts of time since the most recent update to the spectrum. The amount of time for which this is valid will be inversely proportional to the strength of the interactions.

The spectral evolution equation for $a(k)$ will be of the form

$$\frac{\partial a}{\partial t} \approx \frac{\varepsilon^2}{\omega}\frac{\partial}{\partial t}<_2\hat{\phi}_2\hat{\phi}^*>, \tag{19}$$

where again the (non-negligible) $<_1\hat{\phi}_3\hat{\phi}>$ terms have been dropped for simplicity. The time derivative acts only on the $|\Delta|^2$ functions on the right hand side of (17), which have the explicit form

$$|\Delta(\Omega,t)|^2 = \frac{2-2\cos(\Omega t)}{\Omega^2}. \tag{20}$$

The limiting behavior of these functions is essential to the validity of RIA theory. We find a short time limit of

$$\lim_{t\to 0}\left[|\Delta(\Omega,t)|^2\right] = t^2, \tag{21}$$

and a long time limit of

$$\lim_{t\to\infty}\left[|\Delta(\Omega,t)|^2\right] = 2\pi t\,\delta(\Omega), \tag{22}$$

where $\delta(\Omega)$ is now a Dirac delta. For intermediate values of t, we find a sinc-like function in Ω with a bandwidth $\approx 2\pi/t$, as shown in Figure1. The value of t chosen for this function must be of the same or lesser order than the amount of time our expansion (12) remains accurate and rapidly convergent, such that (18) remains valid.

RIA theory chooses to take the long time limit (22) as an approximation of $|\Delta|^2$. In doing so, it assumes that the first order solution holds steady long enough for the bandwidth of the function in Figure1 to become small in comparison with a characteristic bandwidth in the internal wave frequency spectrum. By "characteristic bandwidth," I mean a bandwidth over which amplitudes

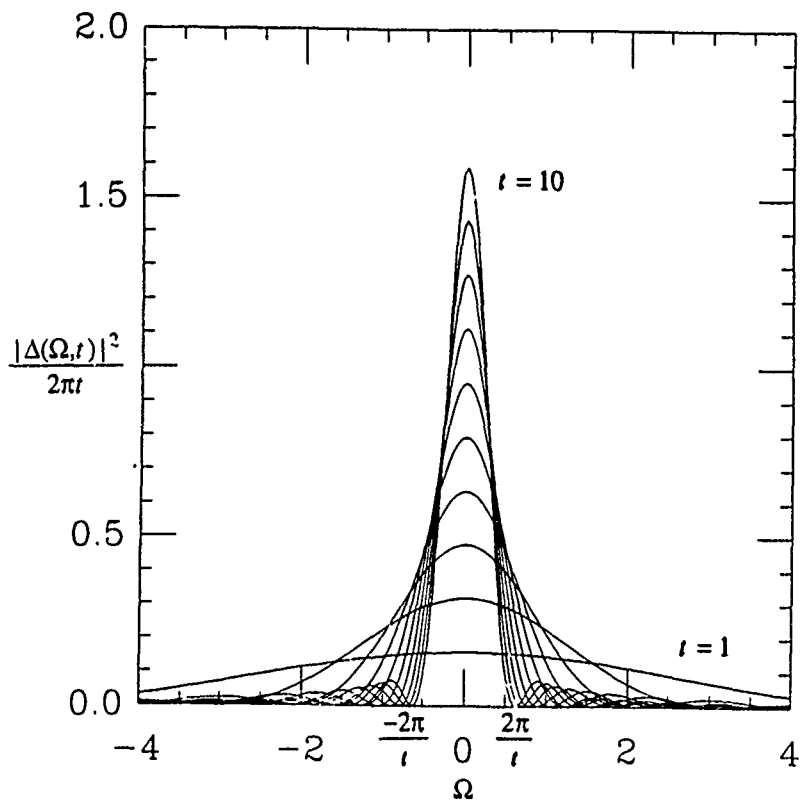


Figure 1. The function $|\Delta|^2/2\pi t$, approximated as a Dirac delta in RIA theory.

vary essentially linearly, such that the filter function depicted in figure 1 may be safely replaced by a Dirac delta. This bandwidth has traditionally been taken as the period of the wave n, yielding the validity criterion

$$\frac{2\pi\nu}{\omega} < 1, \tag{23}$$

where ν^{-1} is the interaction time over which (18) remains valid. It must be stressed that (23) is a necessary but not sufficient criterion for RIA's validity -- we must also be able to somehow justify dropping the cumulant term in (12). However, (23) is an easy test to perform, and Holloway's (1980) implementation of it is still the only quantitative attempt to establish a high wavenumber cutoff for RIA.

Applying (22), we obtain

$$\frac{\partial a}{\partial t} \approx \frac{\pi k^2 \varepsilon^2}{2\omega} \sum_{n_1} \sum_{n_2} \delta_{n,n_1+n_2} \delta(\omega_1+\omega_2-\omega) a_1 a_2 \omega_1 \omega_2 \tag{24}$$

$$+ \delta_{n,n_1-n_2} \delta(\omega_1-\omega_2-\omega) a_1 a_2 \omega_1 \omega_2,$$

plus similar forms for the $<_1\hat{\phi}_3\hat{\phi}>$ terms. In the limit of a continuous spectrum, we have

$$\frac{\partial a}{\partial t} \approx \frac{\pi k^2 \varepsilon^2}{2\omega} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dk_1 dk_2 \delta(k-k_1-k_2)\delta(\omega_1+\omega_2-\omega) a_1 a_2 \omega_1 \omega_2 \tag{25}$$

$$+ \delta(k-k_1+k_2)\delta(\omega_1-\omega_2-\omega) a_1 a_2 \omega_1 \omega_2.$$

Notice that RIA has reduced the dimension of the problem by one. Equation (25) is now a zero-dimensional integral(!); the corresponding equation for internal waves will be two-dimensional.

The transfer integral for internal waves takes the form (Müller, et al, 1986)

$$\frac{\partial a}{\partial t} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\mathbf{k}_1 d\mathbf{k}_2 T^+ \delta(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\delta(\omega_1+\omega_2-\omega)(a_1 a_2 - a a_1 - a a_2) \tag{26}$$

$$+ 2T^- \delta(\mathbf{k}-\mathbf{k}_1+\mathbf{k}_2)\delta(\omega_1-\omega_2-\omega)(a_1 a_2 + a a_1 - a a_2),$$

where a is the internal wave action density and T^+ and T^- are rather messy coupling coefficients depending on the wave vectors $\mathbf{k}$, $\mathbf{k}_1$, and $\mathbf{k}_2$. Resonant interactions, then, are confined to "triads" of three waves satisfying the conditions $\mathbf{k}=\mathbf{k}_1\pm\mathbf{k}_2$ and $\omega=\omega_1\pm\omega_2$.

The appeal of RIA theory is perhaps chiefly tied to the reduction in dimensions associated with the δ functions acting on the frequencies. However, the validity of this δ function is by no means assured, as suggested by the requirement (23). Additionally, the Gaussian assumption, which allowed us to drop the cumulants in (12), suggests another restriction on the strength of the interactions, namely that the "forcing" waves n_1 and n_2 must have rapid decorrelation times

compared to the interaction rates. This restriction is discussed briefly in Müller et al. (1986), in the context of their equation (70), but imposes a criterion which has been relatively difficult to quantify.

For very weak interactions, we may expect the perturbation expansion to be a useful model. However, it will be pointed out in the latter section of this paper that when resonant interactions become too weak, other nonlinear transfer mechanisms must also be considered.

VALIDITY AND APPLICABILITY OF RESONANCE THEORY

In the tradition of Holloway (1980), interaction times predicted by McComas (1977), and MB were compared with other readily available characteristic time scales. Approximate formulas for induced diffusion (ID) and parametric subharmonic instability (PSI) time scales, given by McComas and Müller (1981, henceforth MM), were applied to the GM76 (Cairns and Williams, 1976) model spectrum. The model spectrum and the resulting time scales are shown in Figures 2 and 3, respectively. Numerical values for the GM76 spectrum were inertial frequency $f = 7.0\times10^{-5}\text{s}^{-1}$, buoyancy frequency $N = 71f$, stratification scale $b = 1300\text{m}$, peak mode number $j^* = 3$, and dimensionless energy scale $E = 6.3\times10^{-5}$. The hydrostatic dispersion relation

$$\omega = \sqrt{f^2 + N^2 \frac{k_h^2}{k_z^2}}, \tag{27}$$

where k_h and k_z are the horizontal and vertical wave numbers, was used throughout.

Figure 3 is essentially an attempt to replicate figure 4 of McComas (1977) and figure 11 of MB, without solving the full transfer integral (26). It was generated using equations (11) and (22) of MM, with the ID equation (11) applying for frequencies $\omega > 2f$, and the PSI equation (22) applying for $\omega < 2f$. A value of $x = 1$, or $\omega' = 2.5f$, was chosen in MM's equation (22) to better fit the PSI time scales indicated by McComas (1977) and MB. (The transfer times in figure 8 of MM appear to be approximately ten times too high.)

The figures compare reasonably well, except at very low wave numbers, and at the region near $\omega = 2f$, where interaction rates pass through zero in the earlier plots. The discrepancies at low wave numbers do not substantially affect the results which will follow. Also, there is no reason to expect the actual *strength* of the interactions to become small near $\omega = 2f$. The rightward leaning "spikes" which appear in the earlier plots have therefore been truncated.

Mean Free Paths and Bottom Scattering

Horizontal and vertical "mean free paths," MFP_h and MFP_z of resonantly interacting internal waves, plotted in Figures 4 and 5, were then calculated by multiplying the resonant interaction times in Figure 3 with the horizontal and vertical group velocities. The mean free paths then correspond to the distances which waves may be expected to propagate before they are significantly altered by resonant interactions. If a wave propagates more than one ocean depth in the vertical (or one ocean width in the horizontal) before it has time to resonantly interact, it

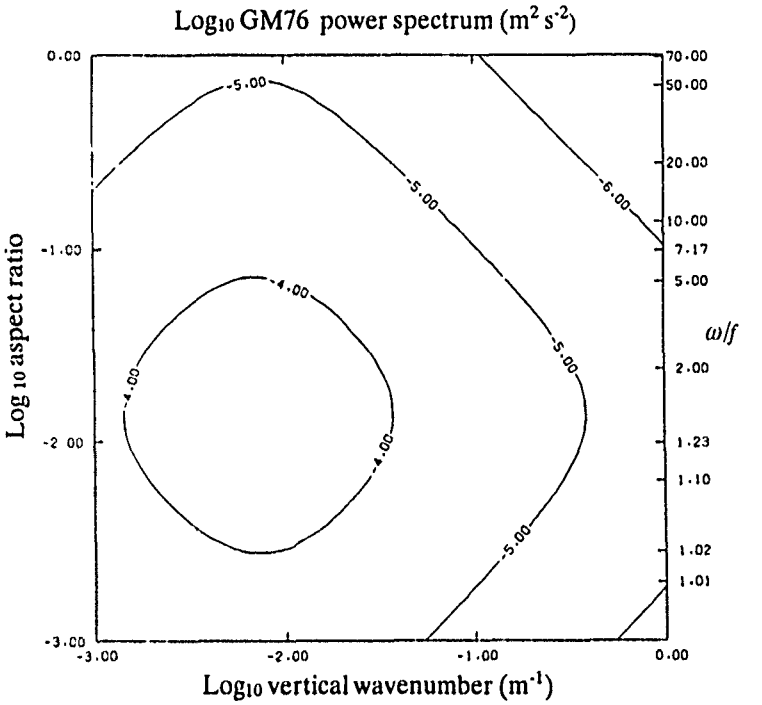


Figure 2. The GM76 model spectrum. The energy density is multiplied by the vertical wavenumber k_z and the aspect ratio α, such that the quantity plotted integrated over the area $d\log(k_z)d\log(\alpha)$ gives the energy in that area.

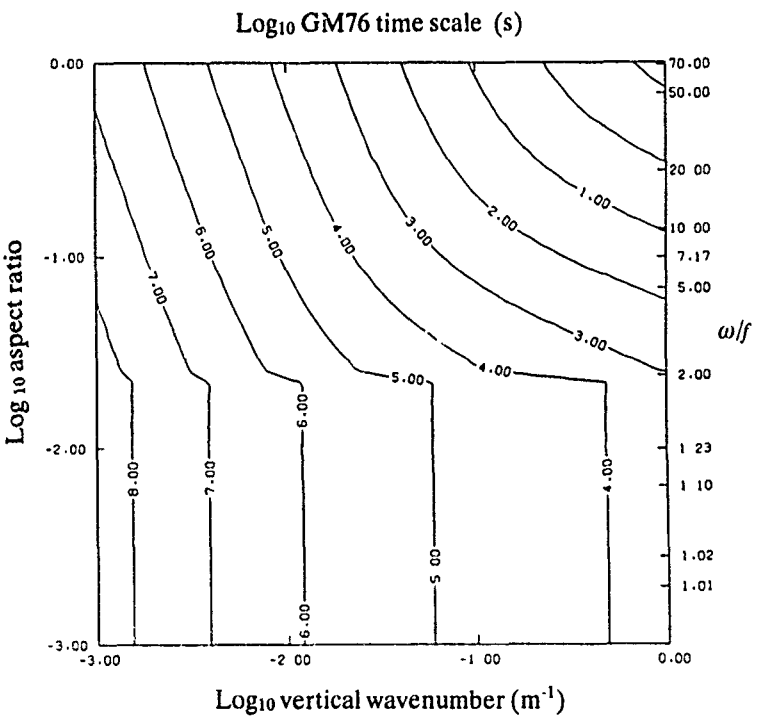


Figure 3. Interaction time scales predicted by RIA theory. The formulas of McComas and Müller (1981) are used to approximate values previously obtained by solving the full RIA transfer integral.

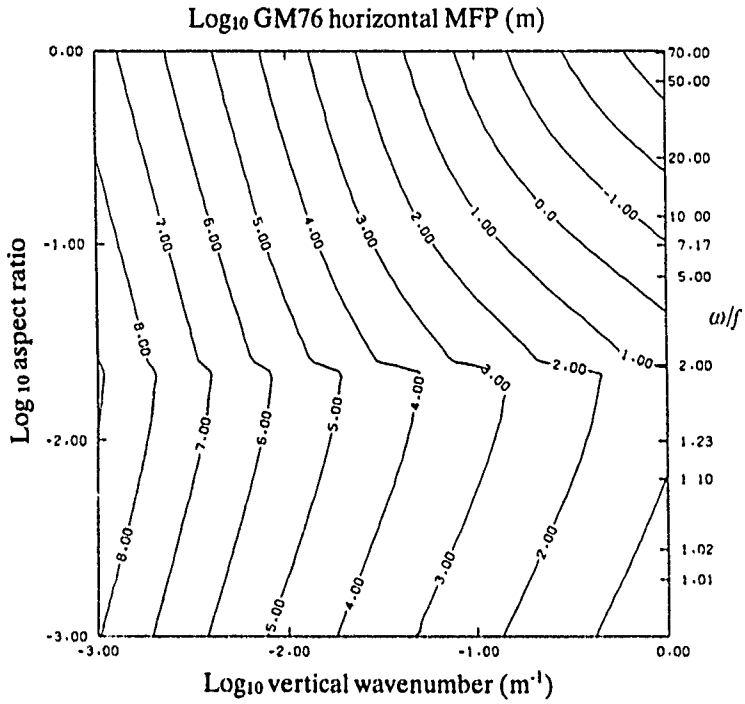


Figure 4. Horizontal "mean free paths" of resonantly interacting waves. In an infinite ocean, waves would be expected to propagate the distances shown before being consumed by resonant interactions. The quantity plotted is the product of the horizontal group velocity with the time scale of Figure 3.

cannot be well modeled without considering scattering effects. Figure 4 suggests that a substantial range of low wavenumber waves will be affected by bottom scattering.

Critical Layers

Two types of critical trapping may occur for low frequency waves, as schematically illustrated by Kunze and Müller (1989) in their figures 1A and 1B.

In the presence of variable baroclinic vorticity ζ, internal waves will experience an effective inertial frequency

$$f_{eff} = f + \frac{\zeta}{2}. \tag{28}$$

As a near-inertial wave propagates into a region in which f_{eff} exceeds the intrinsic frequency, its vertical group velocity will go to zero, sending its vertical wavenumber to infinity. This type of critical layer is discussed in detail in Kunze (1985). If we assume that f_{eff} varies linearly with depth as $|df_{eff}/dz| = 1.2f/3000\text{s}^{-1}\text{m}^{-1}$, where 3000m is the bottom depth, the change in f_{eff} which a wave may experience as it propagates may be estimated by

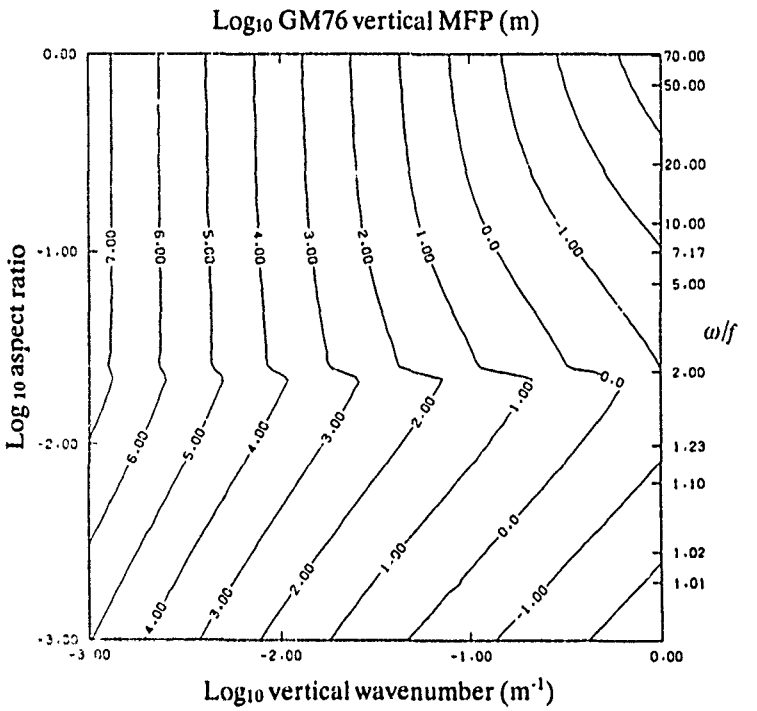


Figure 5. Vertical "mean free paths" of resonantly interacting waves. The quantity plotted is the product of the vertical group velocity with the time scale of Figure 3.

$$|\Delta f_{eff}| = \frac{1.2f}{3000\text{m}} \cdot \min(MFP_z, 3000\text{m}). \tag{29}$$

When $|\Delta f_{eff}| > \omega - f$, a wave may reach a critical layer of this type. An appropriate stall criterion will then be

$$\gamma > 1, \tag{30}$$

where $\gamma \equiv |\Delta f_{eff}|/(\omega - f)$ is plotted in Figure 6. It is seen that, for the parameters chosen, such critical layers may be experienced by waves with frequencies less than $1.2f$ and vertical wavelengths greater than approximately 300 meters.

A second type of critical layer may occur even when the effective Coriolis parameter is constant. A wave may have its intrinsic frequency shifted to f (or f_{eff}) by spatial variability in the geostrophic current. In a time independent mean flow U, the Eulerian frequency ω will be a constant along a ray path, while the intrinsic frequency will be given by $\omega_0 = \omega - \mathbf{k}\cdot\mathbf{U}$. Thus an appropriately oriented change of geostrophic current $\Delta U = \Delta U_{crit} \equiv (\omega - f)/k_h$ along a ray path will push the intrinsic frequency to f and cause the wave to stall. This type of critical layer is discussed in more detail by Olbers (1981).

ΔU_{crit} is plotted in Figure 7; Kunze and Müller (1989) suggest that critical layers are likely when

$$\Delta U_{crit} < 0.05\text{ms}^{-1}. \tag{31}$$

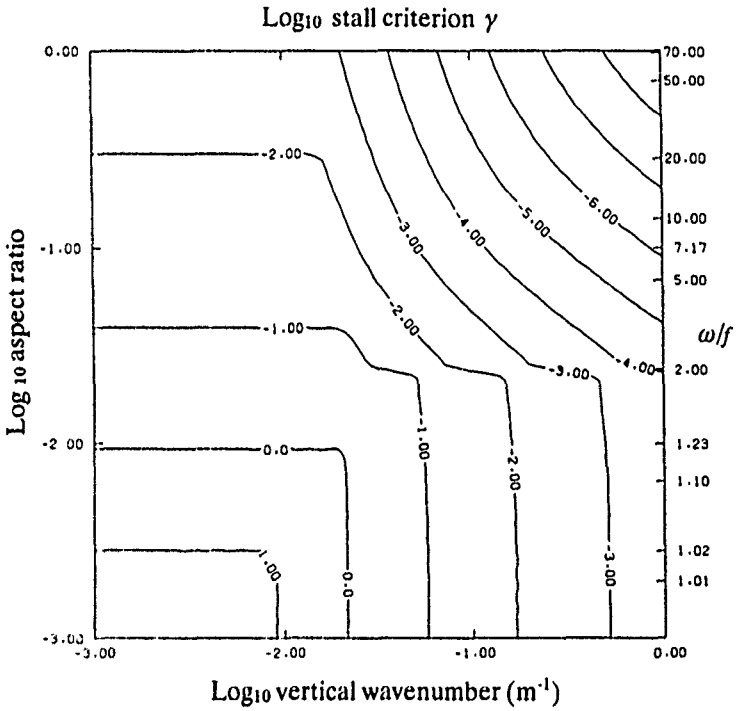


Figure 6. The stall criterion $\gamma = |\Delta f_{eff}|/(\omega - f)$. For $\gamma > 1$, the effective Coriolis frequency may increase to the wave frequency along a ray path, causing critical trapping.

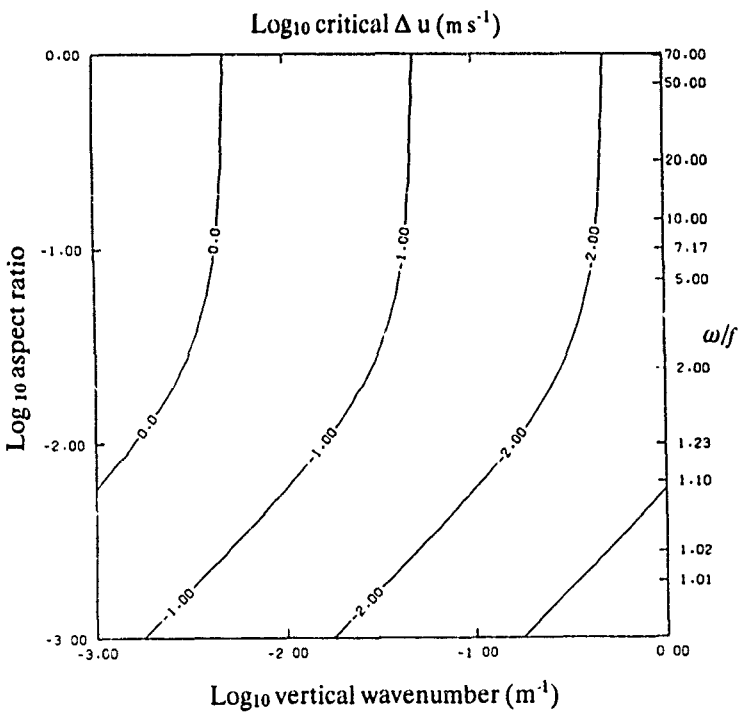


Figure 7. The change in the mean current along a ray path which, if appropriately oriented, will push a wave's intrinsic frequency below the inertial frequency, causing critical trapping.

The likelihood of a wave ever reaching such a critical layer will also depend strongly on that wave's mean free path, however. To account for this, we first define a Richardson number as

$$Ri \equiv \left[\frac{N \min(3000\text{m}, MFP_z)}{\Delta U} \right]^2. \tag{32}$$

Replacing ΔU with ΔU_{crit} gives a "critical Richardson number" of

$$Ri_{crit} \equiv \left[\frac{Nk_z \min(3000\text{m}, MFP_z)}{\omega - f} \right]^2, \tag{33}$$

which is plotted in Figure 8. We may supplement (31) by requiring that the shear be strong enough along a wave's mean free path for

$$Ri < Ri_{crit} \tag{34}$$

to hold. Since Kelvin-Helmholtz instability will occur when $Ri \leq 1/4$, we may conclude that the region in which this type of critical layer may occur will be bounded by the contours

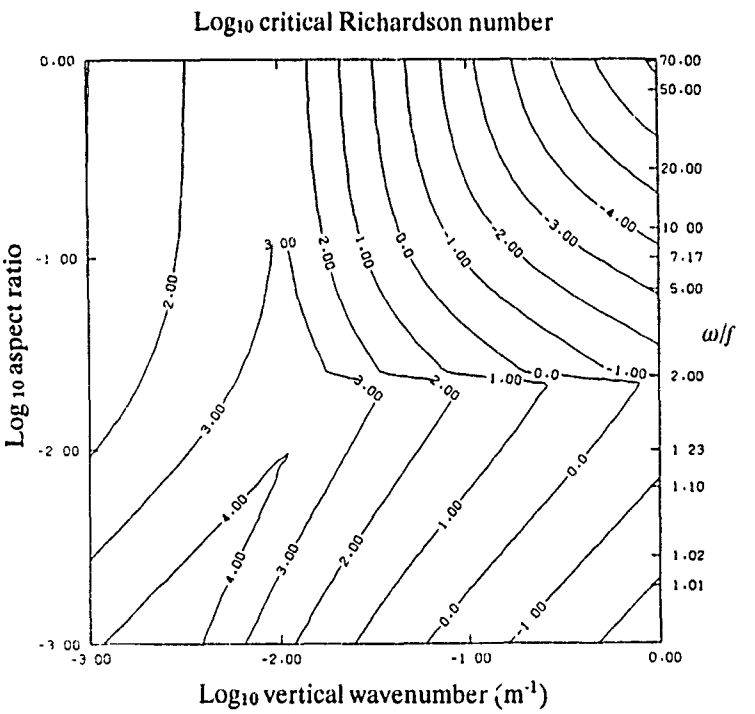


Figure 8. The critical Richardson number, on a scale of either the vertical mean free path or the ocean depth, required to cause trapping. If the actual Richardson number is greater than that shown, the shear will be too weak for this effect to occur in the vertical distance the wave can propagate before it is consumed by resonant interactions..

$\Delta U_{crit} = 0.05 ms^{-1}$ in figure 7 and $Ri_{crit} = 0.25$ in Figure 8.

It should be noted that the first phenomenon, in which $f_{eff} \rightarrow \omega$ along a ray path, is vertically anisotropic and may not occur among both upward and downward propagating waves simultaneously. The second phenomenon, in which ω_0 is shifted to f along a ray path, is strongly dependent on the orientation of the horizontal wave vector, which must be aligned with the mean flow. Finally, both of the critical layers described will no doubt be highly intermittent, and their importance in the overall spectral energy transfer is still unknown.

Validity regions

Figure 9 divides the internal wave spectrum into five sections, using information from Figures 5-9 and the various cutoff criteria listed in the text and in equations (23),(30), and (31). We first consider the three non-overlapping sections, divided by the solid lines. In the leftmost section, the vertical mean free path exceeds the canonical 3000 meter ocean depth, and we expect bottom reflection effects to be significant. This region also roughly corresponds to the region in which the horizontal mean free path is greater than 100 to 1000 kilometers, as indicated in Figure 4. In the rightmost section, which corresponds to the shaded portion of Holloway's (1980) figure 1a,

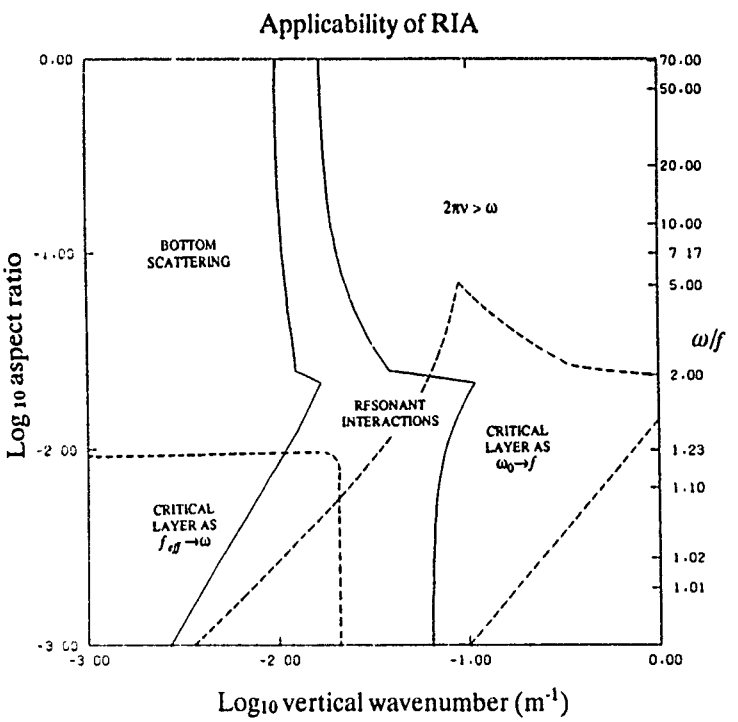


Figure 9. Map of RIA's potential applicability. In the rightmost solid region, the interaction time is less than one wave period, and we expect the perturbation expansion to fail. In the leftmost solid region, bottom reflections will be significant. In the center solid region, RIA may be useful. In the two dashed regions, the types of critical trapping discussed in the text may also be important.

the interaction time exceeds the wave period, and we expect the perturbation expansion to break down. (Again, this cutoff is a crude one, as equation (23) is not a sufficient condition for RIA's validity.)

In the center region, RIA may in fact prove useful. Several possible obstacles must still be recognized here, however. First, the Gaussian or random phase assumption must be tested to see if cumulants may in fact be ignored. Second, we see that much of this region is overlapped by the regions in which the two critical layer phenomena may (or may not) be significant, as indicated by the dashed curves.

The area inside the leftmost dashed region, based on the criterion (30), corresponds to waves which may be susceptible to the first type of critical layer discussed, in which $f_{eff} \rightarrow \omega$ along a ray path. The area inside the rightmost dashed region, bounded on the left by the curve $\Delta U_{crit} = 0.05 \text{ms}^{-1}$ and on the right and top by the curve $Ri_{crit} = 0.25$, corresponds to waves which may be susceptible to the second type of critical layer, in which $\omega_0 \rightarrow f$. Since the dashed regions overlap much of RIA's "applicability region," one might be tempted to abandon the RIA theory altogether at this point. However, it should be mentioned that RIA's parametric subharmonic instability mechanism yields an energy flux toward the near-inertial portion of the spectrum (cf. MB); we may (boldly and blindly) speculate that this flux is in perfect balance with critical layer dissipation.

SUMMARY

The one-dimensional analog has made resonance theory somewhat more accessible to all who might be interested. Also, it has been pointed out that there is a rather substantial region in phase space where bottom reflections may dominate the nonlinear transfer. Finally, two types of critical trapping mechanisms have been discussed for near-inertial waves. While the net contribution by these mechanisms to the overall nonlinear transfer is unknown, it has been shown that they may affect waves over large regions of phase space, including much of the region where resonant interactions may occur. Since RIA does not predict equilibrium for near-inertial waves, however, it has been suggested that critical layer interactions may in fact complement RIA theory in this region.

ACKNOWLEDGMENTS

I would like to thank Peter Müller for his indispensable guidance and suggestions. I would also like to thank Terry Ewart, Jim Riley, and Eric D'Asaro for their feedback and support, and Eric Kunze for taking time to discuss critical layers with me. This work was supported by discretionary funds at the Applied Physics Laboratory at the University of Washington.

REFERENCES

Benney, D. J., and P. G. Saffman, 1966: Nonlinear interactions of random waves in a dispersive medium, *Proc. Roy. Soc., Ser. A, 289*, 301-320

Cairns, J. L., and G. O. Williams, 1976: Internal wave observations from an internal float, 2, *J. Geophys. Res., 81*, 1943-1960.

Hasselmann, K., 1962: On the non-linear energy transfer in a gravity-wave spectrum, 1, General theory, *J. Fluid Mech., 12*, 481-500.

Hasselmann, K., 1966: Feynman diagrams and interaction rules of wave-wave scattering processes, *Rev. Geophys., 4*, 1-32.

Henyey, F.S., J. Wright, and S. M. Flatte, 1986: Energy and action flow through the internal wave field: an eikonal approach, *J. Geophys. Res., 91*, 8487-8495.

Holloway, G., 1980: Oceanic internal waves are not weak waves, *J. Phys. Oceanogr., 10*, 906-914.

Holloway, G., 1982: On the interaction time scales of oceanic internal waves, *J. Phys. Oceanogr., 13*, 293-296.

Kunze, E., 1985: Near-inertial wave propagation in geostrophic shear, *J. Phys. Oceanogr., 15*, 544-565.

Kunze, E., and P. Müller, 1989 The effect of internal waves on vertical geostrophic shear, in *Parameterization of Small-Scale Processes, Proceedings, 'Aha Huliko' a Hawaiian Winter Workshop*, edited by P. Müller and Diane Henderson, pp. 273-287, Hawaii Institute of Geophysics, Honolulu.

McComas, C. H., 1977: Equilibrium mechanisms within the oceanic internal wave field, *J. Phys. Oceanogr., 7*, 836-845.

McComas, C. H., and F. P. Bretherton, 1977: Resonant interaction of oceanic internal waves, *J. Geophys. Res., 83*, 1397-1412.

McComas, C. H., and P. Müller, 1981: Time scales of resonant interactions among oceanic internal waves, *J. Phys. Oceanogr., 11*, 139-147.

Müller, P., G. Holloway, F.S. Henyey, and N. Pomphrey, 1986: Nonlinear interactions among internal gravity waves, *Rev. Geophys., 24*, 493-536.

Olbers, D. J., 1976: Nonlinear energy transfer and the energy balance of the internal wave field in the deep ocean, *J. Fluid Mech., 74*, 375-399.

Olbers, D.J., 1981: The propagation of internal gravity waves in a geostrophic current, *J. Phys. Oceanogr., 11*, 1224-1233.

Pomphrey, N., J. D. Meiss and K. M. Watson, 1980: Description of nonlinear internal wave interactions using Langevin methods, *J. Geophys. Res., 85*, 1085-1094.

The WAMDI Group: Hasselmann, S., K. Hasselmann, E. Bauer, P. A. E. M. Janssen, G. J. Komen, L. Bertotti, P. Lionello, A. Guillaume, V. C. Cardone, J. A. Greenwood, M. Reistad, L. Zambresky, and J. A. Ewing, 1988: The WAM model -- a third generation ocean wave prediction model, *J. Phys. Oceanogr., 18*, 1775-1810.

CHOOSING VARIABLES FOR INTERNAL WAVE DYNAMICS

Frank S. Henyey

Arete Associates, P.O. Box 8050, La Jolla, CA 92038

ABSTRACT

The issue of which variables to use for internal wave dynamics is discussed. Advantages of Hamiltonian canonical variables are mentioned. The suggestion that vertically Lagrangian coordinates be used is reviewed. It would be useful to separate internal waves form the vortical mode, but no satisfactory separation has been found.

Ten years ago, a meeting was held at Scripps Institute of Oceanography (West, 1981) at which various speakers discussed the Hamiltonian formulation of internal wave dynamics, the use of a vertically Lagrangian coordinate system, and the mode of motion now termed the vortical mode. These topics comprise the subject of this paper. In these ten years, we have not completely solved the problem; we don't really know what variables are best to use.

For the case of surface waves, the situation is much better. We have found a set of variables for which the linear approximation is remarkably good (Creamer et al, 1989). The most important short-time nonlinear features of the wave field result not from nonlinear dynamics, but simply from a nonlinear change of variables. The Hamiltonian formulation and Lagrangian coordinates played a large role in finding these variables.

Hamiltonian variables (Henyey, 1983) have a number of advantages. A very important advantage is that conservation laws are very naturally handled in that framework. If one makes an approximation, one generally does not destroy conservation laws, only changes their form slightly. That is not true of approximating equations of motion. An example of a conservation law that has been misunderstood in non-Hamiltonian frameworks is conservation of action. A commonly studied system is the dynamics of infinitesimal waves interacting with a single periodic wave, imagined to be an exact solution of the equation of motion. For this problem, in the frame of reference moving with the phase velocity of the big wave, the flow is steady. As a result, the energy, frequency, and their ratio, the action are all exactly conserved. The key is that the energy is the value of the Hamiltonian function in that frame.

Another important conservation law is momentum conservation. In non-Hamiltonian frameworks a concept of "pseudomomentum" has arisen. This concept is entirely unneeded; there is no distinction between the terms that add up to the total momentum and the "pseudomomentum". The momentian density in terms of Hamiltonian canonical variables is

$$pv = \text{linear terms} - p_1 \nabla q_1 - p_2 \nabla q_2 - p_3 \nabla q_3 \tag{1}$$

Each pair of variables describe one of the linear modes; p_1, q_1, are sound, p_2, q_2, are internal waves, and p_3, q_3 are vortical mode. (Only one combination of p_3 and q_3 is physical. This leads to "gauge" freedom of choosing the other, as described by Henyey, 1983.)

If for example

$$q_2 = a\cos(\boldsymbol{k}\cdot\boldsymbol{x} - \omega\tau) \tag{2}$$

$$p_2 = b\sin(\boldsymbol{k}\cdot\boldsymbol{x} - \omega\tau) \tag{3}$$

then the average contribution to the momentum density is

$$\langle pv \rangle = abk/2 \tag{4}$$

The quadratic part of energy density turns out to be

$$E = ab\omega/2 \tag{5}$$

where ω is related to $\mathbf{k}$ by the linear dispersion relation, so the action density is

$$A = ab/2 \tag{6}$$

Another use of the Hamiltonian is in formal development. The Eikonal, or ray tracing method naturally follows from a Hamiltonian formulation (Henyey and Pomphrey, 1983). Various generalizations, such as Whitham's (Whitham, 1974) to narrow-band nonlinear waves, or the Gaussian beam method require the Hamiltonian or the closely related Lagrangian framework. Weak interaction transport theory is also most usefully derived from the Hamiltonian framework; the symmetry of the interaction coefficients which leads to detailed balance, entropy, and effective temperature is not true for arbitrary differential equations, but always occurs for Hamiltonian systems. The symmetry is obvious in a Hamiltonian derivation.

Closely related to the Hamiltonian formulation is the Lagrangian variational formulation. The canonical variational principle is that the action

$$S = \int \left(\sum P_j \dot{q}_j - H \right) \tag{7}$$

is stationary under independent variation of all p's and q's. Unfortunately, the stationary value is always a saddle point, never a minimum. In certain special cases, the problem can be constrained (such as by assuming steady state) to make the stationary point an extremum. Arnold's stability method, and other stability calculations are examples.

Finally, the Hamiltonian can be helpful for finding new variables. The theory of canonical transformations and the Hamilton-Jacobi equation can lead to better variables. In the surface wave case we found the variables by implementing the removal of nonresonant interactions, which always can be accomplished.

Variables can be a better choice to use if the linear approximation is more accurate. There are several ways of telling if the linearization might be better. One way is that the high frequency or wavenumber part of the spectrum might be smaller due to the absence of bound harmonics; the free waves are present in either case. Another way is that the Doppler shift might be smaller, so the dispersion relation is better satisfied. A theoretical way of determining that linearization is better is by the size of the nonlinear interaction terms. By these criteria, the replacement of density as a variable by vertical displacement that is referred to as removing "fine structure contamination," or in other words, reducing bound harmonics, is an improved choice of variables.

One possible improvement in the choice of variables is the use of the vertically Lagrangian coordinate system. This idea is not new; both Ripa and Milder in West (1981) mention it. More recently Odulo (1989) has written a paper advocating its use, but I do not understand his reasons for believing it to be superior. Sherman and Pinkel (1991) have analyzed internal wave data in terms of this coordinate system. Unfortunately, the results are not completely convincing. There is not good evidence that bound harmonics were reduced. They did find that the small waves were significantly reduced in frequency (Sherman used this for the cover picture of his thesis, see Figure 1). This presumably is associated with a reduction of the Doppler shift, but of course the dispersion relation cannot be evaluated in the absence of knowing the horizontal wavenumber.

What can be said theoretically? Not much. The potential energy becomes exactly quadratic, but nonlinearities associated with stratification have always been believed unimportant relative to the advective nonlinearity. The vertical advective terms are missing, but why shouldn't horizontal advection be equally important? In order to stimulate consideration of these variables, I present the equation of motion:

The new vertical coordinate is

$$z' = z - \zeta \tag{8}$$

where ζ is the vertical displacement. The advective time derivative is

$$d_t = \partial_t' + \boldsymbol{u} \cdot \nabla' \tag{9}$$

where

$$\nabla' = \boldsymbol{e}_x \partial_x' + \boldsymbol{e}_y \partial_y' \tag{10}$$

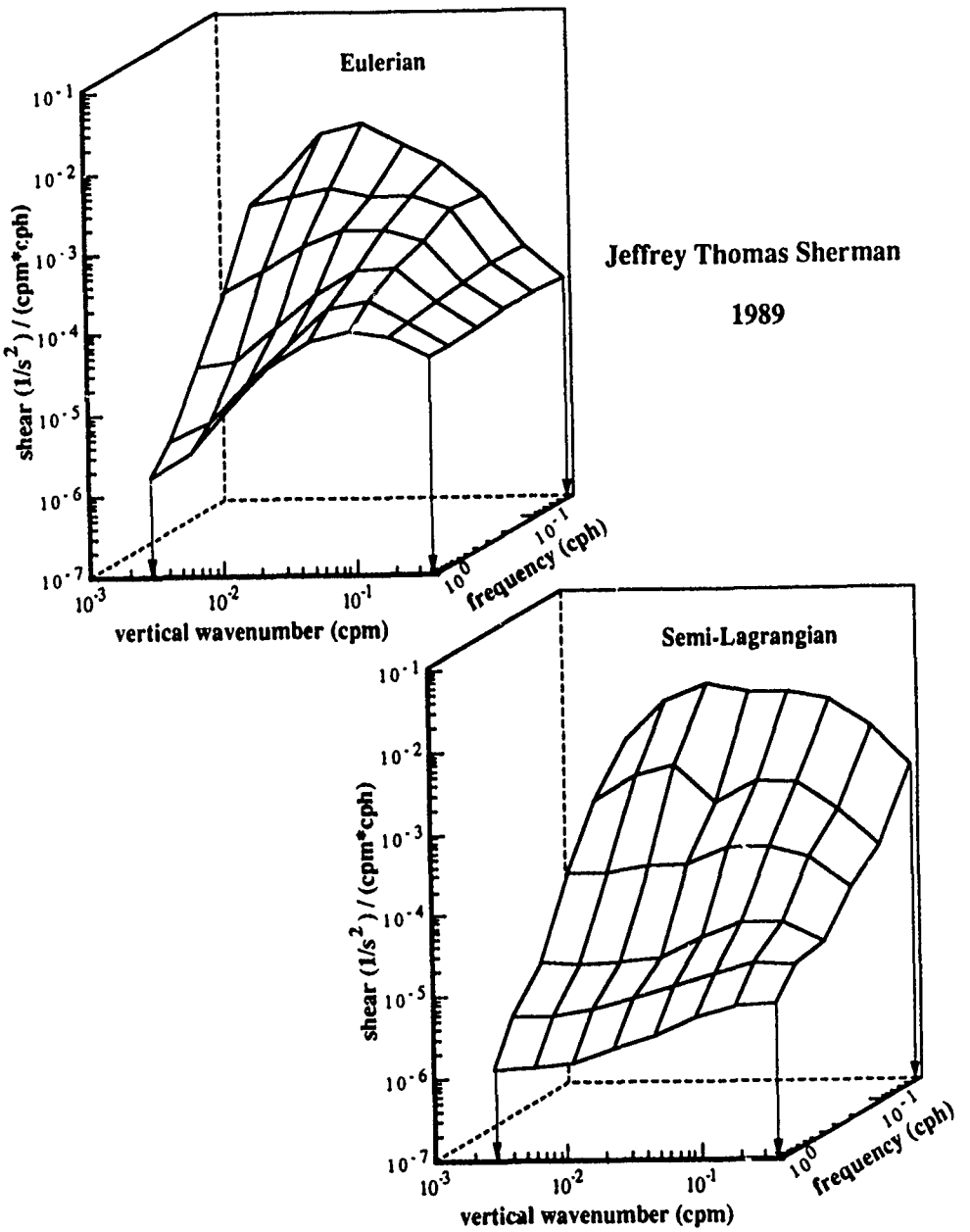


Figure 1. The cover of Sherman's Ph.D. thesis. This shows the change in the spectrum going from the Eulerian to the vertically Lagrangian coordinate system.

has only horizontal advection and **u** is the horizontal velocity. The Jacobian of the transformation is

$$S \equiv \frac{1}{1+\partial_z'\zeta} \tag{11}$$

The equations of motion are

$$d_t\zeta = w \tag{12}$$

$$d_t w = -s(N_{z'}^2\zeta + \partial_z' p) \tag{13}$$

$$d_t u = -\nabla'\zeta d_t w - f\times u - \nabla' p \tag{14}$$

$$s\partial_z' w + \nabla'\cdot u = s\nabla'\zeta\cdot\partial_z' u \tag{15}$$

The potential vorticity is

$$J = sf + (\hat{z} - s\nabla'\zeta)\cdot\nabla'\times v \tag{16}$$

Another reason to choose particular variables is to separate the types of motion; sound, internal waves, and vortical motion. This separation is reasonably easily carried out in the linear approximation by diagonalizing an operator or a Hamiltonian form, and has been done numerous times. The linear separation demonstrates an advantage of Hamiltonian variables.

Using the usual variables, the linear vortical mode is geostrophic. With Hamiltonian variables there is no distinction between the geostrophic and cyclostrophic cases. This is closer to reality, as there is little difference other than size between an eddy which rotates in a week and one that rotates in an hour. Nonlinear separations, in general, cannot be accomplished. Sound is generally separated by setting $\nabla\cdot v=0$ while retaining the value of N^2 rather than the actual stratification. This is adequate for most purposes, but not ideal.

The separation of internal waves from vortical mode is much less satisfactory. A pure internal wave field can be defined by the condition that the potential vorticity

$$J = (f+\omega)\cdot\nabla(z-\zeta) \tag{17}$$

is a constant (such as f). Even for a pure oceanic internal wave field of typical strength, the nonlinear terms in J are of comparable magnitude to the linear ones. It is clear that the linear separation of IWs from vortical mode is unsatisfactory. A step in the right direction has been

taken by Staquet and Riley (1989). They define that part of the velocity belonging to the vortical mode by a prescription which is equivalent to minimizing the energy subject to specified J and ζ fields. This prescription obviously does not separate ζ. One could say that ζ is purely internal wave, but a field with $J \neq 0$ and $\zeta = 0$ would radiate IWs in a time of order N^{-1}. A stationary eddy, which one would like to say is purely vortical, has nonzero ζ.

An alternative would be to say that J is one of the coordinates. It cannot be a canonical coordinate, as canonical fields have Poisson brackets that are delta function for canonical pairs, and vanish for a single field. Potential vorticity has the Poisson bracket (Abarbanel, private communication).

$$\{J(x),J(x')\} = \nabla J(x) \times \nabla \delta(x-x') \cdot \nabla(z-\zeta) \tag{18}$$

Even worse than not vanishing, this PB involves ζ which, (as discussed above) although not entirely IW, is mostly IW.

Having reviewed the choice of variables for internal wave dynamics, it is clear that the situation is not entirely satisfactory. For the surface wave problem, we are able to demonstrate variables and list their wonderful properties. For the IW problem, we find difficulties with all proposed schemes. This situation is indicative of the more general situation that internal wave dynamics is more difficult than that for other wave fields such as sound and surface waves. Progress has been made, however, and we are able at least to try out various possibilities.

REFERENCES

Creamer, D.B., F. Henyey, R. Schult, and J. Wright (1989): Improved linear representation of ocean surface waves. *J. Fluid Mech.* 205, 135-161.

Henyey, F. S. (1983): Hamiltonian description of stratified fluid dynamics. *Phys. Fluids* 26, 40-47.

Henyey, F.S. and N. Pomphrey (1983): *Dyn. Atmos. Oceans* 7, 189-219.

Milder, D.M. (1982): Hamiltonian dynamics of internal waves. *J. Fluid Mech.* 119, 269-282.

Odulo, A. (1989): Theory of nonlinear motions in a stratified fluid with a free surface. Unpublished.

Sherman, J. and R. Pinkel (1991): Estimates of the vertical wavenumber-frequency spectra of vertical shear and strain. *J. Phys. Oceanogr.* 21, 292-303.

Staquet, C. and J. Riley (1989): On the velocity field associated with potential vorticity. *Dyn. Atmos. Oceans* 14, 93-123.

West, B.J. (ed). 1981: Nonlinear Properties of Internal Waves, *AIP Conf. Proc.* 76 (American Inst. Phys. New York) see, in particular, the contributions of Pomphrey, Ripa, Abarbanal, and Milder (p. 345).

Whitham, G.B. (1974): Linear and nonlinear waves (John Wiley, New York).

SCALING OF INTERNAL WAVE MODEL PREDICTIONS FOR ε

Frank S. Henyey

Areté Associates, P.O. Box 8050, La Jolla, CA 92038

ABSTRACT

Both induced diffusion and the Henyey, Wright, and Flatté model predict the Gregg scaling law

$$\varepsilon \sim \left(\frac{S_{10}}{N}\right)^4 N^2 f \tag{1}$$

when the Garrett-Munk spectrum is used. This note discusses the scaling in a more general context. The more general scaling is needed:

1. near the equator, where f=0.

2. if the shear spectrum is not white in vertical wavenumber.

INTRODUCTION

Various model predictions of the energy flux through the internal wave field have been made. These model predictions involve not only the intrinsic physics put into the model, but also the assumed spectral form for the internal waves. Invariably, the Garrett-Munk (GM) spectrum is used. According to the models discussed here, GM is appropriate in steady-state conditions far from the equator. The question arises, what if these conditions are not met?

Gregg (1989) showed that a particular form describes data well. His form is a version of that predicted by both induced diffusion (ID) (McComas and Müller, 1981) and by the model of Henyey, Wright, and Flatté (1986) (HWF). These two models differ in their overall level, but have the same scaling form when applied to a Garrett-Munk spectral shape.

GENERALIZED SCALING

In order to obtain more general scaling expressions for these two models, we will assume that small-scale IWs obey a power law spectrum in vertical wavenumber, but not necessarily that of GM. We will denote the RMS shear on any vertical scale by S; the shear spectral level scales as S^2/k_v, where k_v is the vertical wavenumber describing that scale. GM has $S \sim k_v^{1/2}$.

Induced diffusion is described by a diffusivity in vertical wavenumber space. This diffusivity is given by (Müller et al, 1986, eq. 81).

$$D = \frac{k_h^2}{2} \int d(S^2)\, \delta(\Omega - V_g \cdot K) \tag{2}$$

Ω and K are the frequency and wavenumber of the large scale waves contributing dS^2 to the shear variance. V_g is the group velocity. For scaling purposes, the delta function behaves as the reciprocal of its argument, and we use the term $V_{gv}K_v$ as being indicative of the scaling. Thus, we obtain

$$D \sim \frac{k_h^2}{V_{gv}} \frac{S^2}{K_v}\bigg|_{max}. \tag{3}$$

The maximum wavenumber of the shearing waves is of order the wavenumber of the sheared wave. (For GM, S^2/K_v is constant, making this observation not necessary.) The group velocity is

$$V_{gv} = \frac{-N^2}{\omega} \frac{k_h^2}{k_v^3} \tag{4}$$

where we have assumed here and in what follows $\omega << N$; $k_h << k_v$.

Putting it all together, we have

$$\epsilon = D \partial_{k_v} E \tag{5}$$

where E is the energy spectral level in vertical wavenumber, and scales as

$$E \sim \frac{(S/k_v)^2}{k_v} = \frac{S^2}{k_v^3} \tag{6}$$

Therefore,

$$\epsilon \sim \frac{\omega S^4}{k_v^2 N^2} \tag{7}$$

Under steady state conditions, this form must be independent of k_v , leading to the GM $S \sim k_v^{1/2}$. More generally, for an evolving spectrum, the flux is only ε at the smallest scales. These scales are presumably determined by a Richardson number criteria S~N, so

$$\varepsilon \sim \omega N^2/k_c^2 \tag{8}$$

where k_c is the wavenumber at some fixed Richardson number.

The HWF model is based on deterministic evolution at small scales. The steady-state spectral form is not obtained analytically. Numerical results (Flatté, Henyey, and Wright, 1985) demonstrate $S \sim k_v^{1/2}$ in steady state for this model. Energy dissipation scales as

$$\varepsilon \sim \left|\frac{dk_v}{dt}\right| E \tag{9}$$

at some fixed Richardson number.

The change of vertical wavenumber is $|\, dk_v/dt \,| \sim Sk_h$, so

$$\varepsilon \sim Sk_h\left(\frac{S^2}{k_v^3}\right) \tag{10}$$

$$\sim\left(\frac{k_h}{k_v}\right)\frac{N^3}{k_c^2} \tag{11}$$

As long as ω-f is not typically much smaller than f, k_h/k_v scales as ω/N. Thus,

$$\varepsilon \sim \omega N^2/k_c^2, \tag{12}$$

exactly the same scaling as ID has. Except near the equator, typical values of ω are of order f. This is the origin of the f factor in the previously published scaling expressions. The models do not assume any process happens on time scales of 1/f. If there were no separation between horizontal and vertical scales, ω would scale as N, giving $\varepsilon \sim N^3/k_c^2$, which could be obtained by dimensional arguments. The reason the models predict otherwise is that the aspect ratio is a dimensionless parameter.

We now assume

$$\frac{S^2}{k_v} \sim A k_v^p \tag{13}$$

so p=0 for GM. Evaluating this at fixed Richardson number we get

$$N^2 \sim A k_c^{p+1} \tag{14}$$

Eliminating the unknown coefficient A, we get

$$\left(\frac{S}{N}\right)^2 \sim \left(\frac{k_v}{k_c}\right)^{p+1} \tag{15}$$

we solve for k_c and put it into ε, obtaining

$$\varepsilon \sim \frac{\omega N^2}{k_v^2}\left(\frac{S}{N}\right)^{\frac{4}{p+1}} \tag{16}$$

Following Gregg, we fix k_v at $2\pi/10$m, so we get

$$\varepsilon \sim \omega N^2 \left(\frac{S_{10}}{N}\right)^{\frac{4}{p+1}} \tag{17}$$

This agrees with Gregg for the GM values

$$\omega \sim f \tag{18}$$

$$p = 0. \tag{19}$$

and is the appropriate prediction in more general cases from both models.

REFERENCES

Flatté, S.M. F.S. Henyey, and J.A. Wright (1985): Eikonal calculations of short-wavelength internal wave spectra, *J. Geophys. Res.* 90, 7265-7272.

Gregg, M.C. (1989). Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.* 94, 9686-9698.

Henyey, F.S. J.A. Wright, and S.M. Flatté (1986): Energy and action flow through the internal wave field: an Eikonal approach. *J. Geophys. Res.* 91, 8487-8495.

McComas, C.H. and P. Müller (1981): The dynamic balance of internal waves. *J. Phys. Oceanogr.* 11, 970-986.

Müller, P., G. Holloway, F. Henyey, and N. Pomphrey (1986): Nonlinear interactions among internal gravity waves. *Rev. Geophys.* 24, 493-536.

IS SCATTERING OR REFLECTION MORE EFFECTIVE IN CAUSING BOUNDARY MIXING?

Naihuai Xu and Peter Müller

Department of Oceanography, University of Hawaii, 1000 Pope Road, Honolulu HI 96822

Abstract

Comparison is made of the redistributed energy flux resulting from scattering of two-dimensional internal waves off one-dimensional random topography with that resulting from the reflection of two-dimensional internal waves off a straight slope. Reflection redistributes much more energy flux than scattering does (2.86 $mW\,m^{-2}$ as opposed to 0.99 $mW\,m^{-2}$), but reflection redistributes less energy flux to high wavenumbers than does scattering (0.90 $mW\,m^{-2}$ as against 0.97 $mW\,m^{-2}$ to wavenumbers greater than 10^{-4} cpm). Scattering might thus be equally or more efficient than reflection in causing high shears and mixing near the bottom.

1. Introduction

Boundary mixing has been advocated (e.g., Ivey, 1987) as a process responsible for diapycnal mixing in the deep ocean. Diapycnal mixing is required to satisfy the global balances of mass and heat and to support observed poleward heat transport. Internal wave reflection at topography has been proposed (Eriksen, 1982, 1985; Garrett and Gilbert, 1988) to cause this boundary mixing. Recently, Xu (1991) and Müller and Xu (1991) have suggested scattering at random topography. Internal wave interaction with topography distorts the wave spectrum and redistributes the incoming energy flux within the wavenumber space. Waves scattered or reflected to high wavenumbers might break and cause boundary mixing.

The reflection process has been intensively studied by Eriksen (1982, 1985) and Garrett and Gilbert (1988). When the incident wave length is much shorter than the radius of curvature of topography, internal waves can be treated as encountering an infinite flat slope. The reflection process conserves the frequency and the tangential component of the wavenumber vector. Eriksen (1982, 1985) calculated the absolute difference between the incident and reflected energy flux which is comparable to the incident flux. This quantity does not tell how much energy flux could be used for generating shears, but can serve as the upper limit. In an attempt to assess quantitatively how much energy flux is available for mixing, Garrett and Gilbert (1988) assumed that waves with mode numbers greater than a critical mode

number undergo instability, break, and cause mixing. This critical mode number is determined by assuming that waves with mode numbers larger than the critical number cause the inverse Richardson number to become larger than one. Given typical parameters for the deep ocean, the energy flux available for mixing is about 0.25 mW m^{-2} (about 1.4% of incident flux), much smaller than absolute energy flux difference calculated by Eriksen (1985).

Scattering of internal waves at a random topography was theoretically analyzed by Müller and Olbers (1975) in the Bragg scattering or weak interaction limit. Olbers and Pomphrey (1981) estimated the redistributed energy flux in this limit and found it to be as small as $\sim 10^{-2} - 10^{-3}$ of the incident one. The energy flux available for mixing should be even smaller than this. However, the formula used to calculate the redistributed energy flux contains an algebraic error. Rubenstein (1988) treated a simplified problem: a two-dimensional wavefield interacting with one-dimensional topography. In a detailed study, he found significant energy flux transfer by assuming that the probability of scattering from an incident wavenumber to a scattered wavenumber is proportional to the ratio of the respective fluxes. Unfortunately, he interpreted this probability as a density in a vertical wavenumber space whereas it is a probability in horizontal wavenumber space. A systematic derivation and evaluation of the scattering integral is given in Xu (1991) and Müller and Xu (1991). Here we compare the scattering and reflection processes and concentrate on the question of which process is more efficient in causing boundary mixing, scattering or reflection.

We emphasize here the difference between the redistributed energy flux and the energy flux available for mixing. The redistributed flux represents the energy flux that is transferred to other wavenumbers, either lower or higher. When the energy flux is redistributed to lower wavenumbers, the field becomes more stable. The shear in the field decreases. This process does not favor mixing. On the other hand, if wave energy flux is transferred to higher wavenumbers, shear is increased and the flow field tends to be less stable and more prone to breaking. This process favors mixing. The energy flux available for mixing is the energy flux to high wavenumbers, which causes the shear to become larger than critical. This energy flux will drive the field to instability, and breaking of internal waves thus occurs.

2. Scattering process

We consider the simplified model of two-dimensional internal waves scattered off one-dimensional random topography or reflected off a straight slope.

Baines (1971) considered the two-dimensional internal wave interaction with an

arbitrary flat bump topography. The slope of the flat bump topography is required to be less than the slope of the group velocity. The radiation condition was formulated as a homogeneous integral equation. An incident wave impinging on this topography will generate a forward transmitted wave and a backward reflected wave. The solution to the system is obtained by solving a Fredholm integral equation of the second kind. If we further assume that the vertical wavelength of an incident wave is much smaller than the height of topography (the ratio of them is a small parameter ε_1), the Fredholm integral equation can be solved to the second order of ε_1 so that we can obtain the lowest order corrections of the scattered energy flux. The redistributed energy flux is then (Xu, 1991)

$$\begin{aligned} D^s(\omega,\alpha) &= \int_0^\infty d\alpha' \mu^2 V(\omega)[E(\omega,\alpha')\alpha - E(\omega,\alpha)\alpha'] \\ &\quad [S_1(\alpha+\alpha') + S_1(\alpha-\alpha')] \end{aligned} \tag{1}$$

where $E(\omega,\alpha)$ is the incident energy density spectrum, $S_1(\alpha)$ the topography height spectrum, $V(\omega)/\alpha$ the vertical group velocity, $\mu^2 = (N^2-\omega^2)/(\omega^2-f^2)$, α the modulus of horizontal wavenumber and ω the wave frequency. For each frequency, the energy flux is conserved, i.e., $\int D^s(\omega,\alpha)d\alpha = 0$. This can be seen from the antisymmetric character of

$$[E(\omega,\alpha')\alpha - E(\omega,\alpha)\alpha']. \tag{2}$$

The redistributed energy flux consists of two parts: the first part is the gain of energy flux at α from interaction of incident wavenumber α' with topographic wavenumber $\alpha \pm \alpha'$. More energy flux is gained at the higher wavenumbers than at the lower wavenumbers. This can be seen by looking at the first part of D^s denoted by $D^s_{11}(\omega,\alpha)$. Using k_1' to represent the horizontal wavenumber of an incident wave and k_1'' the topographic wave, then $D^s_{11}(\omega,\alpha)$ can be transformed as

$$\begin{aligned} D^s_{11}(\omega,\alpha) &= D^s_{11}(\omega,|k_1'+k_1''|) \\ &= \int_{-\infty}^\infty dk_1' \mu^2 V(\omega) E(\omega,k_1')|k_1'+k_1''|S_1(k_1'') \end{aligned} \tag{3}$$

where $\alpha = |k_1'+k_1''|$. The transfer function is proportional to α. The scattered wave energy flux will be larger for a high scattered wavenumber α than for low scattered wavenumber α. The second term of Eq. (1) acts like a friction to the incident energy flux. Waves that would be reflected to a certain wavenumber α are scattered to another wavenumber α'. Scattering always transfers energy flux to high wavenumbers. This result can be understood using statistical mechanics and the H-theorem (Müller and Xu, 1991).

Equation (1) can be derived by several approaches (see Xu, 1991). Rubenstein's hypothetical approach differs by a factor of μ and exaggerates the scattering

efficiency at low frequencies. This extra factor results from interpreting a probability density with respect to horizontal wavenumber as a density with respect to vertical wavenumber.

3. Reflection process

The two-dimensional internal wave reflection off a straight slope is considered here. Assume that the flat slope $z = sx$ lies in the $x - z$ plane where $s = \tan\varphi_o$, the inclination of the topography relative to the horizontal plane. The reflected wavenumber vector is denoted as $\vec{k} = (k\cos\theta, 0, k\sin\theta)$, $\alpha = |k\cos\theta|$. The inclination θ is determined by the frequency and radiation condition. The reflection law requires that the frequency and tangential component of wavenumber vector conserves in the reflection. The reflected energy spectrum in wavenumber space under the constraint of radiation and boundary conditions (no normal flow across the boundary) yields

$$E_r(\vec{k}) = E(\vec{k^i}) \tag{4}$$

where $\vec{k^i}$ is the incident wavenumber vector. It shows that in the case of reflection, the energy density spectrum at the reflected and incident wavenumber are the same. This conclusion is consistent with Eriksen's results (1982, 1985). Under the constraint of the radiation condition, incident waves exist in only certain permissible regions shown in Figure 1. For each frequency, there are two permissible incident wavenumbers, one of them is in the first quadrant, the other is in another quadrant. If the incident wavenumbers lie in the first quadrant of the $k_1 - k_3$ plane, the reflected horizontal wavenumbers are always greater than the corresponding incident ones, while if the incident wavenumbers are in any other quadrant, the reflected horizontal wavenumbers will be smaller than the corresponding incident ones. The energy flux at each frequency is therefore transferred to both high and low wavenumbers. The difference between the reflected and incident normal energy flux therefore is

$$D(\omega,\alpha) = \sum_{r=1}^{2}\left[E\left(\omega, \left|\frac{\alpha\cos(\theta_r - \varphi_o)}{\cos(\theta_r + \varphi_o)}\right|\right) - E(\omega,\alpha)\right]\frac{N^2 - f^2}{\omega\alpha}|\sin\theta_r\cos^2\theta_r\cos(\theta_r + \varphi_o)|, \tag{5}$$

where θ_1 and θ_2 correspond to the two inclinations of the permissible reflected wavenumbers. The redistributed energy flux vanishes when integrated over all horizontal wavenumbers. At the critical frequency defined by

$$\omega_c^2 = N^2\sin\varphi_o + f^2\cos^2\varphi_o, \tag{6}$$

one of the reflected wavenumbers ($\theta = \pi/2 + \varphi_o$) becomes infinity, the other one ($\theta = -\pi/2 - \varphi_o$) goes to zero. This implies that the energy flux is redistributed to

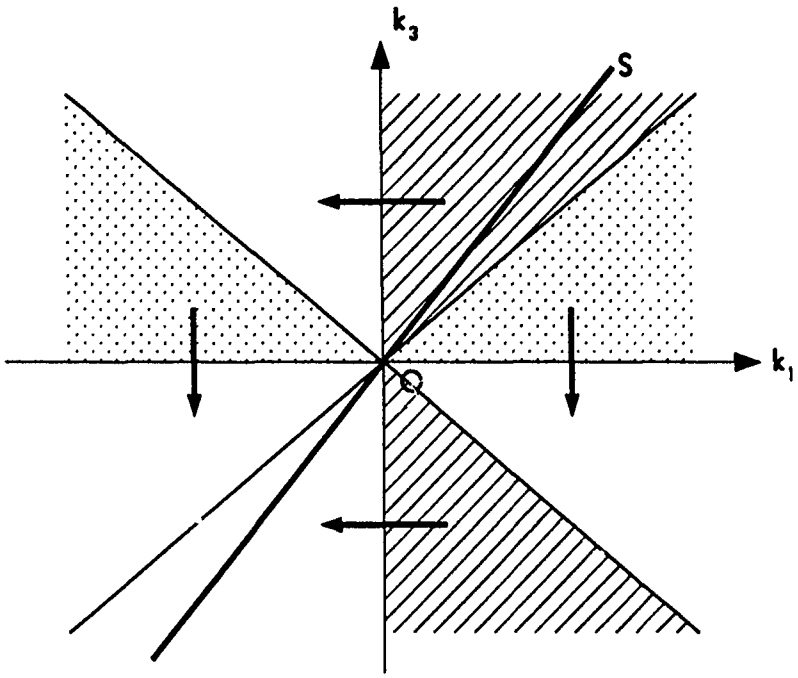


Figure 1: Regions of permissible incident and reflective waves in the horizontal-vertical wavenumber plane. The heavy solid straight line represents the bottom slope. The light solid lines represent the critical frequency cone. The cross-hatched and stippled regions are permissible incident wavenumbers. For the cross-hatched regions the reflection is subcritical, for the stippled region the reflection is supercritical. The incident wavenumbers reflect to regions are indicated by the arrows.

very high and low wavenumbers at the critical frequency.

4. Background Spectra

For analytical convenience, we choose the Garrett and Munk model GM76 (Desaubies, 1976)

$$E(\omega,\alpha) = b^2 N N_0 E_0 B(\omega)\frac{A(\alpha/\alpha_*)}{\alpha_*} \tag{7}$$

where

$$B(\omega) = \frac{2}{\pi}\frac{f}{\omega}(\omega^2 - f^2)^{-1/2} \tag{8}$$

$$A(\lambda) = \frac{2}{\pi}(1+\lambda^2)^{-1} \tag{9}$$

and $N_o = 5.2 \times 10^{-3} s^{-1}$, $E_o = 6 \times 10^{-5}$ and $b = 1.3 \times 10^3$ m. The bandwidth is given by

$$\alpha_*(\omega) = \left(\frac{\omega^2 - f^2}{N_o^2 - \omega^2} \right)^{1/2} \frac{\pi}{b} j_*, \tag{10}$$

here $j_* = 3$ is the frequency-independent equivalent mode number bandwidth. We set the high wavenumber cutoff

$$\alpha_{hc}(\omega) = \left(\frac{\omega^2 - f^2}{N^2 - \omega} \right)^{1/2} \beta_{hc}, \tag{11}$$

where $b_{uc} = 0.1$ cpm. Finite incident energy flux is achieved by assuming the lower wavenumber cutoff

$$\alpha_{lc}(\omega) = \lambda_l \alpha_* \tag{12}$$

where $\lambda_l \approx 0.154$ based on identical energy fluxes in both the continuous GM76 and the discrete GM81 (Munk, 1981) models.

The bottom spectrum is chosen as formulated by Bell (1975)

$$S_1(\alpha) = \tilde{F}_o \tilde{S}_1(k_1) \tag{13}$$

where $\tilde{F}_o = 0.5\pi F_o/\alpha_o = 2.0 \times 10^4 \pi^3$ cpm^{-2} is the variance and

$$\tilde{S}_1(k_1) = \begin{cases} \frac{\alpha \alpha_1}{(\alpha^2 + \alpha_1^2)^{3/2}} & \text{for } k_1 < \alpha_c^t \\ 0 & \text{for } k_1 > \alpha_c^t \end{cases} \tag{14}$$

describes the variation, the high wavenumber cutoff is α_c^t. The rms height of this spectrum is about 125 m.

In addition, we assume the parameters for typical deep ocean conditions are $f = 0.042$ cph (mid-latitude, 30°), $N = 0.4$ cph and $\varphi_o = 4°$ corresponding to $s = \tan \varphi_o = 0.07$.

5. Comparison of redistributed energy fluxes by scattering and reflection

With a specified interior internal wave spectrum and a topographic spectrum, we can evaluate the redistributed energy flux spectrum for each process.

Figure 2 shows the comparison of the redistributed energy fluxes $D(\omega, \alpha)$ as a function of α for frequency $\omega = 1.33f$ in a variance conserving representation. In the scattering process, energy flux is redistributed from low to high wavenumbers,

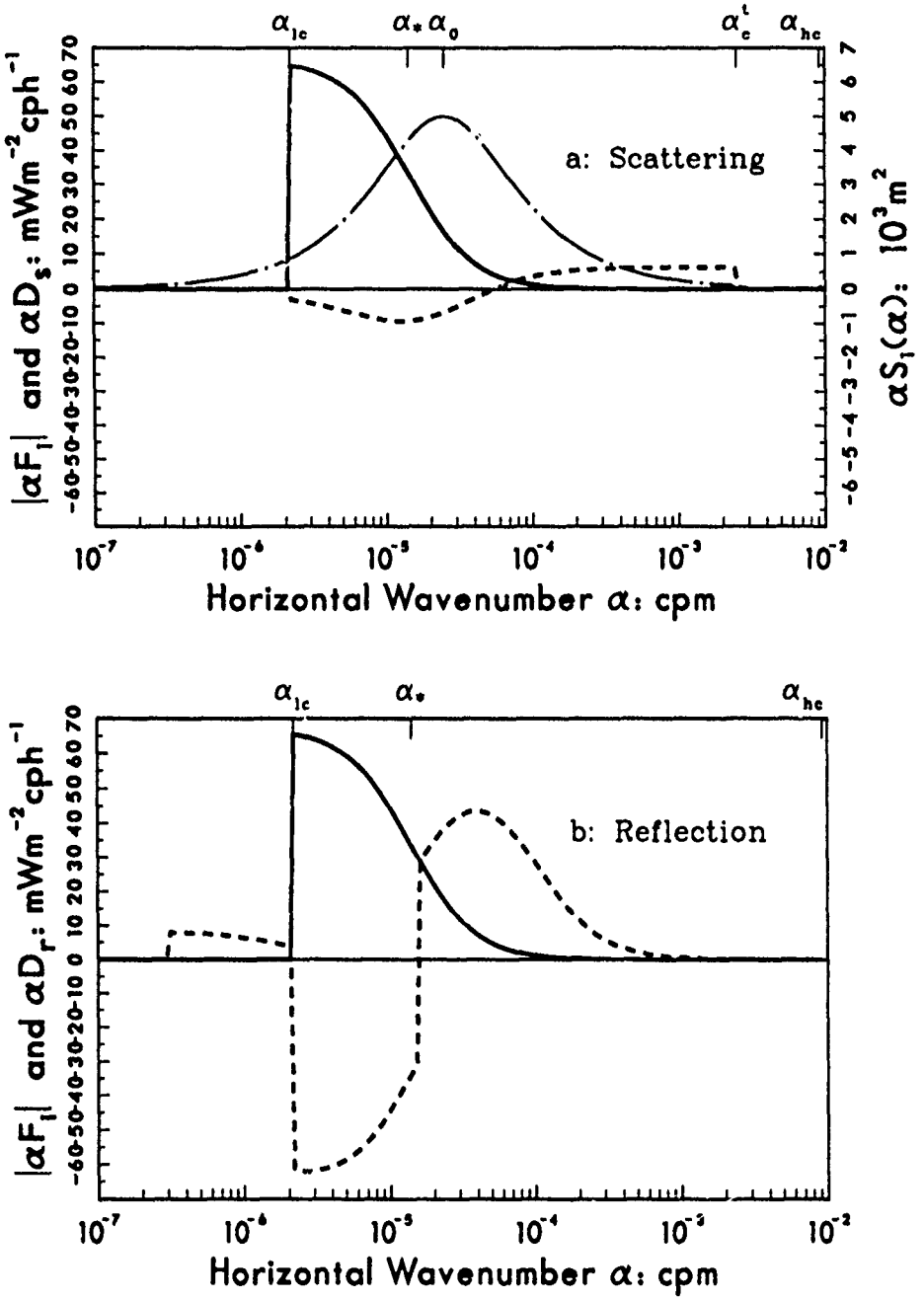


Figure 2: Comparison of redistributed energy fluxes $D(\omega,\alpha)$ by scattering or reflection. Incident energy flux $|F_i(\alpha,\omega)|$ (solid line) and redistributed energy flux $D(\alpha,\omega)$ (dashed line) as a function of horizontal wavenumber for frequency $\omega = 1.33f$. The representation is variance conserving. The wavenumbers α_{lc}, α_*, and α_{hc} are the low wavenumber cutoff, the bandwidth, and the high wavenumber cutoff of the incident internal wave spectrum, respectively. (a). For two-dimensional internal wave scattering model. $S_1(\alpha)$ (dash-dotted line) is the bottom spectrum. The wavenumber α_o and α_c^t are the bandwidth and high wavenumber cutoff of the topographic spectrum, respectively. (b). For two-dimensional internal wave reflection model. The bottom slope is $\gamma = 0.07$ and the critical frequency is $\omega_c = 1.2f$.

whereas in the reflection process energy flux is transferred from medium to both high and low wavenumbers. Because of no normal flow across the boundary, the area of negative lobes equals the area of positive lobes.

Integration with respect to wavenumber α yields the redistributed energy flux

$$D^{+}(\omega) = \frac{1}{2}\int d\alpha |D(\omega, \alpha)| \tag{15}$$

as shown in Figure 3. The total redistributed energy flux D^{+} is 3.89 mW m^{-2} out of 18.2 mW m^{-2} , the incident normal energy flux, for the reflection process, and 1.2 mW m^{-2} out of 17.6 mW m^{-2} for the scattering process. The reflection process redistributes 21% of the incident energy flux as against 6.8% for the scattering process. Reflection is much more efficient than scattering in redistributing energy flux. Scattering redistributes most of the energy flux near the inertial frequency, reflection around the critical frequency.

An important difference between the two processes is the frequency-integrated energy flux spectrum $D(\alpha)$ as shown in Figure 4. The reflection process

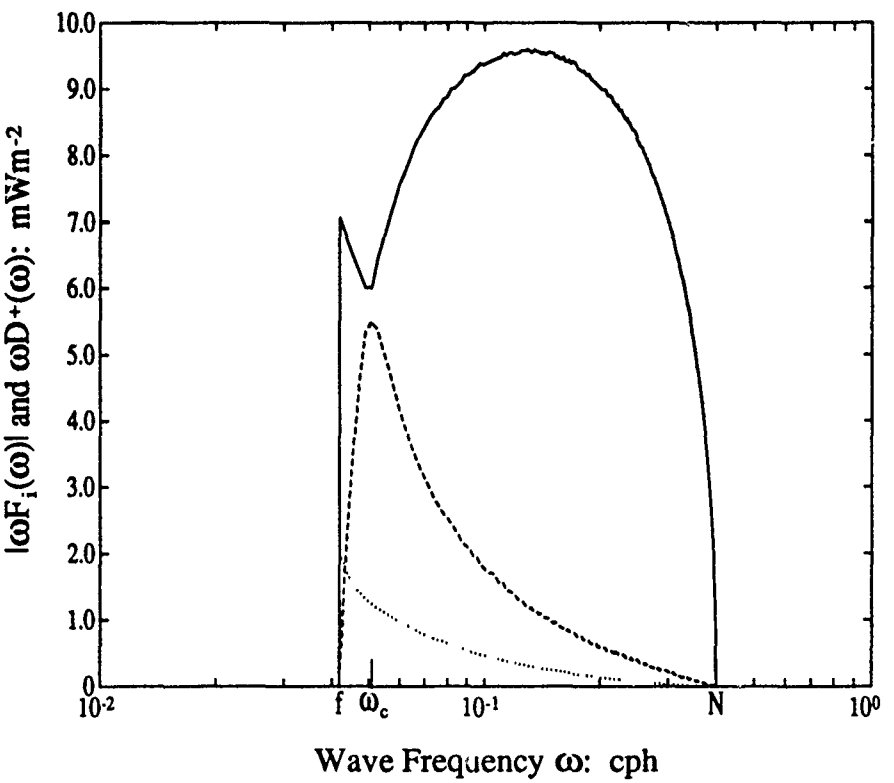


Figure 3: Comparison of two-dimensional scattering and reflection. Incident energy flux $|F_i(\omega)|$ (solid line), redistributed energy flux $D^{+}(\omega)$ by scattering (dashed line) and redistributed energy flux $D^{r}(\omega)$ by reflection (dotted line) as a function of frequency in a variance-conserving representation.

redistributes 2.86 mW m^{-2} of incoming energy flux (15.7% incident), scattering redistributes 0.99 mW m^{-2} (5.6% incident). These numbers are smaller than the integrated D^+ since contributions cancel by frequency integration. The scattering process redistributes energy flux to much higher wavenumbers. In the scattering process 0.97 mW m^{-2} out of 0.99 mW m^{-2} , the total redistributed energy flux is transferred to high wavenumbers above 10^{-4} cpm, whereas in the reflection process, 0.90 mW m^{-2} , which is less than 32% of the total redistributed energy flux, is moved to wavenumbers above 10^{-4} cpm. Scattering hence must be more efficient than critical reflection in increasing the shear and the inverse Richardson number near the bottom.

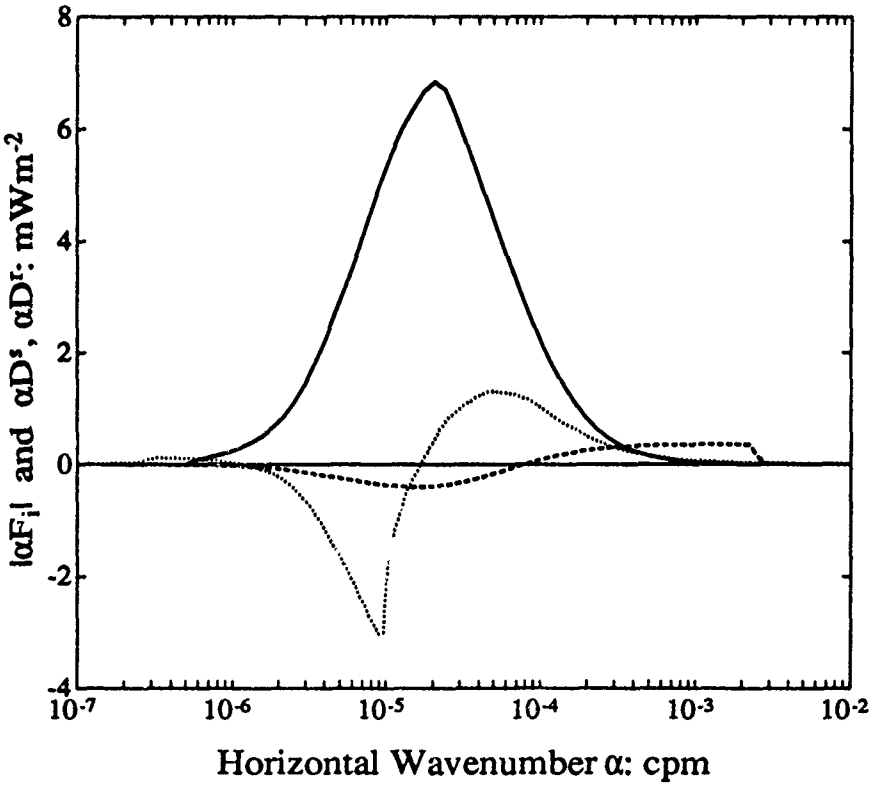


Figure 4: Comparison of two-dimensional scattering and reflection. Incident energy flux $|F_i(\alpha)|$ (solid line), energy flux $D^s(\alpha)$ redistributed by scattering (dashed line), and energy flux $D^r(\alpha)$ redistributed by reflection (dotted line) in a variance conserving representation.

6. Discussions

Here we discuss some possible restrictions:

Three-dimensional scattering model

Are the conclusions drawn from the comparison of scattering and reflection processes for the two-dimensional geometry still valid for three-dimensional model? Müller and Xu's (1991) and Xu's (1991) studies show that the two-dimensional model is a good representation of the three-dimensional internal wave interaction model in terms of the redistributed energy flux and other quantities. Using the perturbation method, Xu (1991) and Müller and Xu (1991) obtain a similar result to Eq. (16) for the redistributed energy flux

$$\begin{aligned} D^s(\omega,\vec{\alpha}) &= 2\int_{-\infty}^{\infty} d^2\vec{\alpha}_1 \frac{\mu^2}{\alpha^2\alpha_1^2} V(\omega)|\vec{\alpha}_1\cdot\vec{\alpha}_1 + i\frac{\vec{f}}{\omega}\cdot\vec{\alpha}\times\vec{\alpha}_1|^2 \\ &\quad [E(\omega,\vec{\alpha}_1)\alpha - E(\omega,\vec{\alpha}_1)\alpha_1]S_2(\vec{\alpha}-\vec{\alpha}_1), \end{aligned} \tag{16}$$

where $S_2(\vec{\alpha})$ is the two-dimensional random topography spectrum. $S_2(\alpha)$ is derived as

$$S_2(\alpha) = \frac{\tilde{F}_o \alpha_o \alpha}{(\alpha^2+\alpha_o^2)^{3/2}}, \tag{17}$$

a natural extension of one-dimensional bottom topography by the assumption of horizontal isotropy, and $E(\omega,\vec{\alpha})$ is specified as GM76 as before.

Numerical evaluation of Eq. (16) gives similar features as shown for the two-dimensional model in Figures 2, 3, and 4. Magnitudes differ only slightly by less than 15% of the corresponding two-dimensional model results. The total redistributed energy flux for the three-dimensional model is 1.14 $\mathrm{mW\,m^{-2}}$ as opposed to 0.99 $\mathrm{mW\,m^{-2}}$ for the two-dimensional case. The gain part of the redistributed energy flux is dominant at high horizontal wavenumbers and the lost part at lower horizontal wavenumbers. With this approximation, Eq. (16) can be integrated to yield

$$D(\omega,\alpha) \approx 2\frac{N^2-\omega^2}{\omega^2-f^2}\frac{1}{2}\frac{\omega^2+f^2}{\omega^2}V(\omega)[\alpha E(\omega)S_2(\alpha) - E(\omega,\alpha)C] \tag{18}$$

where $E(\omega) = \int d\alpha E(\omega,\alpha)$ and

$$C = \int d^2\alpha\alpha S_2(\alpha) = \tilde{F}_o\alpha_o\{\ln\frac{2\alpha_c^t}{\alpha_o} - 1\}. \tag{19}$$

This approximation turns out to be a good one; it is indistinguishable from the numerically evaluated solution plotted in Figure 2. The approximation (18) is

particularly useful in theoretical study to derive the quantities associated with the three-dimensional internal wave scattering model.

Limits of the scattering theory

For the establishment of both three-dimensional and two-dimensional internal wave scattering theory, an expansion is made in which two parameters are assumed small: ε_1 is the ratio of topography height to the incident vertical internal wavelength, ε_2 is the ratio of the slope of topography to the slope of wave rays, in the root mean square sense. The GM spectrum implies the variance of β

$$\langle \beta^2 \rangle \approx \beta_* b_{uc} = \left(\frac{2\pi}{1000 \text{ m}} \right)^2 \frac{N}{N_o}. \tag{20}$$

With the specified topographic spectrum S_2, we obtain

$$\varepsilon_1 = 2\pi \frac{125}{100} \left(\frac{N}{N_o} \right)^{1/2} \tag{21}$$

which is of order one for the deep ocean where $N/N_o \sim 10^{-1}$. Bell's spectrum implies a slope variance of

$$\gamma^2 = \int d\alpha S_2(\alpha) \alpha^2 = \tilde{F}_o \alpha_o \alpha_c^t \approx (0.2)^2, \tag{22}$$

hence

$$\varepsilon_2 = 0.2 \left(\frac{N^2 - \omega^2}{\omega^2 - f^2} \right)^{1/2} \tag{23}$$

which is larger than one for $\omega \leq 2f$. For typical ocean conditions, the expansion is only marginally correct and breaks down for near inertial oscillation. The transfer to high wavenumber is a general tendency not limited to weak interaction. It represents the approach to statistical equilibrium. Since the approach to equilibrium is generally faster the larger the nonlinearities are, our results can be expected to represent a lower bound.

Singularities in solutions:

Singularities might indicate possible breakdowns of the theory and therefore deserve careful study. In horizontal wavenumber space $\alpha \to 0$ and $\alpha \to \infty$ represent singularities causing the energy flux and energy density to become infinite. These singularities are overcome by assuming low and high wavenumber cutoffs. There still exist singularities in wave frequency. For the scattering process the energy density has a nonintegratable singularity but energy flux is finite. The reflected energy density spectrum at one of the critical angles $\theta = \pi/2 + \varphi_o$ (corresponding to the critical frequency ω_c) has a nonintegrable singularity. The reflected energy flux

spectrum is finite here and has the greatest contribution to the total reflected energy flux spectrum. This ensures that none of the frequencies become the singular point of scattered or reflected energy fluxes.

Energy flux available for mixing:

The essential quantity in causing boundary mixing is the redistributed energy flux available for mixing. For the typical parameters of ocean conditions specified in this paper, Garrett and Gilbert's (1988) results show that the energy flux available for mixing is about 0.25 $\mathrm{mW\,m^{-2}}$ and is due to the redistributed energy flux with wavenumbers beyond $\alpha_c \approx 2 \times 10^{-4}$ cpm for the three-dimensional internal wave reflection model. Using the Eq. (18) to estimate the corresponding energy flux available for mixing for the three-dimensional internal wave scattering model, we find about 0.80 $\mathrm{mW\,m^{-2}}$ of the redistributed energy flux will go to mixing.

7. Conclusions

Both scattering of internal waves off random topography (Müller and Olbers, 1975; Olbers and Pomphrey, 1981; Rubenstein 1988, Müller and Xu, 1991 and Xu, 1991) and reflection of internal waves off a straight slope (Eriksen, 1982 and 1985; Garrett and Gilbert, 1988) have been carefully studied. Here we assessed whether scattering or reflection is the more efficient process in causing boundary mixing?

In order to answer this question, we compared scattering and reflection together for the two-dimensional internal wave models. The scattering of three-dimensional internal waves off two-dimensional random topography is more realistic. The study shows that the simplified two-dimensional internal wave scattering model is a good approximation to this general three-dimensional scattering model. They show almost similar features and only differ slightly in magnitudes-by less than 15% in terms of redistributed energy flux. The conclusions are still valid for the general model of three-dimensional internal wave interactions. The scattering process is analyzed under the assumptions of (i) the height of topography is smaller than the vertical wavelength and (ii) the slope of topography is smaller than the wave slope, in the root mean square sense. For typical deep ocean conditions, these conditions are only marginally satisfied, especially the slope condition, which breaks down for near inertial waves. The reflection theory presented here is consistent with Eriksen's work. Comparison between scattering and reflection processes is made in detail. In the reflection process, energy flux is redistributed to both high and low wavenumbers, whereas in the scattering process scattered energy flux is transferred to high wavenumbers. Reflection redistributed much more normal energy flux than the scattering process does (2.86 $\mathrm{mW\,m^{-2}}$ as opposed to 0.99 $\mathrm{mW\,m^{-2}}$), but the

reflection process redistributes less energy flux to high wavenumbers, for example, to $\alpha > 10^{-4}$ cpm, than the scattering process (0.9 mW m^{-2} opposed to 0.97 mW m^{-2}). Since most of the redistributed energy flux goes to high wavenumbers, we could roughly estimate the redistributed energy flux available for mixing as 0.80 mW m^{-2} compared with 0.25 mW m^{-2} estimated for the reflection process by Garrett and Gilbert (1988). Scattering might thus be equally or more efficient than the reflection in causing shears and mixing near the bottom.

Acknowledgments

This work relates to Department of Navy Grant N00014-89-J-1315 issued by the Office of Naval Research. The United States Government has a royalty-free license throughout the world in all copyrightable material contained herein. This paper is School of Ocean and Earth Science and Technology contribution 2637.

References

Baines, P. G., 1971a: The reflexion of internal/inertial waves from bumpy surfaces, *J. Fluid Mech.*, 46, part 2, 273-291.

Bell, T. H., 1975: Statistical features of sea-floor topography, *Deep-Sea Res.*, 22, 883-892.

Desaubies, Y. J. F., 1976: Analytical representation of internal wave spectra, *J. Phys. Oceanogr.*, 6, 976-981.

Eriksen, C. C., 1982: Observations of internal wave reflection off sloping bottoms, *J. Geophys. Res.*, 87, 525-538.

Eriksen, C. C., 1985: Implications of ocean bottom reflection for internal wave spectra and mixing, *J. Phys. Oceanogr.*, 15, 1145-1156.

Garrett, C.J.R. and D. Gilbert, 1988: Estimates of vertical mixing by internal waves reflected off a sloping bottom, in *Small-scale turbulence and mixing in the ocean.* J. C. J. Nihoul and B. M. Jamart (editors), 405-424.

Ivey, G. N., 1987: The role of boundary mixing in the deep ocean, *J. Geophys. Res.*, Vol. 92, C11, 11,873-11,878.

Müller, P. and D. J. Olbers, 1975: On the dynamics of internal waves in the deep ocean, *J. Geophys. Res.*, Vol. 80, 3848-3860.

Müller, P. and N. Xu, 1991: Scattering of oceanic gravity waves off random bottom topography, submitted to *J. Phys. Ocean.*

Munk, W. H., 1981: Internal waves and small-scale processes, in *Evolution of Physical Oceanography Scientific Surveys in Honor of Henry Stommel*, edited by B. A. Warren and C. Wunsch, MIT Press, Cambridge, Mass., pp. 264-291.

Olbers, D. J. and N. Pomphrey, 1981: Disqualifying two candidates for the energy balance of oceanic internal waves, *J. Phys. Oceanogr.*, 11, 1423-1425.

Rubenstein, D., 1988: Scattering of inertial waves by rough bathymetry, *J. Phys. Oceanogr.*, 18, 5-18.

Xu, N., 1991: Interactions of internal waves with random bottom topography and a straight slope, M.S. thesis, Univ. of Hawaii.

On the Exchange of Energy Between Surface and Internal Wave Fields

Kenneth M. Watson
Marine Physical Laboratory, Scripps Institution of Oceanography
University of California, San Diego
San Diego, CA 92093-0701

1. Introduction

The proposal by Garrett and Munk (1972a) that there is a (more-or-less) steady internal wave (IW) spectrum has led to continuing search for the sources and sinks of IW energy. It has been generally thought that energy loss occurs principally by a cascade process to small spatial scales, where turbulent dissipation damps the fluid motion. Dissipation rates have been deduced by Garrett and Munk (1972b), Gargett et al. (1981), Osborn (1978), Lueck et al. (1983), Gregg et al. (1986) and Gregg (1989). These energy loss rates fall in the range 10^{-4} to 10^{-3} W m^{-2}, corresponding to decay times of 10 to 100 days, with 50 days often accepted as a nominal characteristic time.

Theoretical calculations of the energy loss rate have been given by McComas (1978), McComas and Müller (1981), Pomphrey et al. (1980), and by Flatte et al. (1985). These calculations predict dissipation rates in the 10^{-4} to 10^{-3} m^{-2} range.

Numerous mechanisms for producing IW's have been proposed. Some of these are illustrated in Fig 1, where the IW spectral domain is indicated in a frequency-vertical mode number diagram. Large scale flow over bottom topography (Bell, 1975), flows in the mixed layer (Bell, 1978, Kanthu, 1979), and mesoscale shears (Watson 1985) are among the mechanisms which appear to provide significant IW energy sources.

The transfer of energy between surface waves (SW's) and IW's has received considerable attention. In particular, models proposed by Watson et al. (WWC, 1976), Olbers and Herterich (OH, 1979), Dysthe and Das (DD, 1981), and Watson (1990) have predicted the rate at which IW's receive energy from SW's. These models have been based on weakly nonlinear interaction theory with coupling given by "resonant triads" of two surface and one internal wave modes. The energy exchange was significantly overestimated by WWC, who used a "locked phase" approximation and a numerical calculation using too few modes. The calculations of OH were limited to a specific mechanism ("spontaneous creation") and predicted an insignificant rate of energy exchange. To obtain an analytic prediction, DD studied what they called the "modulation mechanism" and assumed a pencil beam of SW's. They concluded that a significant energy flow did not occur for a realistic wind-wave field.

The IW/SW interaction was recently reviewed by Watson (1990). He noted that the formal expressions for the interaction obtained by WWC, OH, and DD were all equivalent, but that in none of these papers had the implications of the theory been

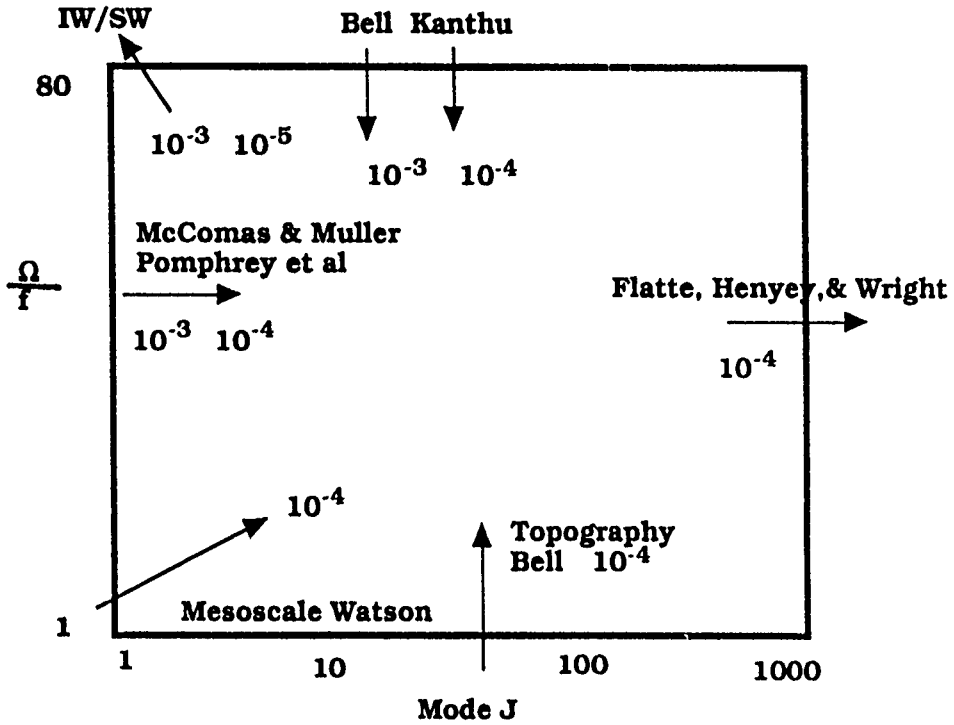


Fig. 1 Some suggested mechanisms of internal wave energy input and loss. The notation is that of Eqs. (1-4), f is the inertial frequency, and the units are W m^{-2}.

adequately developed. Watson investigated both the OH spontaneous creation mechanism and the DD modulation mechanisms[1]. He concluded that, except for winds greater than about 20 m s^{-1}, only the modulation mechanism is of practical significance[2]. Using this mechanism, energy transfer rates of about 10^{-4} Wm^{-2} were found, but with energy flow from the IW field to the SW field. This energy transfer rate was found to be significant only for the lowest vertical modes and for IW frequencies within a factor of ten of the peak Vaisala frequency.

In Watson (1990) an exponential profile for the Vaisala frequency was assumed. To obtain a more realistic estimate of the significance of IW/SW interactions, historical Vaisala frequency data and wind speed records have been collected and the theory applied to these. A preliminary account of the conclusions of this work is given here.

2. Linear Wave Models

The IW/SW interaction has been treated as a weakly nonlinear coupling of the two linear wave fields. A rectangular coordinate system with the z-axis directed

[1] Note that both these mechanisms were included in the formal analyses of WEC, OH, and DD.

[2] The small energy exchange rate of the spontaneous creation mechanism was observed by OH.

upward is assumed to have its origin in the local plane of the quiescent ocean surface.

The vertical component of IW fluid velocity is written as

$$w(x,z,t) = \sum_j \sum_K A_{jK}\, W_{jK}(z) \exp\left[i(K \bullet x - \Omega t)\right]. \tag{1}$$

Here A_{jK} specified the mode amplitude and W is a solution to the equation.

$$\left[\frac{d^2}{dz^2} + K^2\,(N^2/\Omega^2 - 1)\right] W_{jK} = 0\,, \tag{2}$$

$$W_{jK}\,(0) = 0\,.$$
$$W_{jk}\,(-b) = 0. \tag{3}$$

Here $z = -b$ is the assumed water depth and

$$\Omega = \Omega\,(j,K) \tag{4}$$

is the mode frequency determined from Eqs. (1), and (2), and (3). The fluid velocity at the surface $z = 0$ is $U(x,t)$. THE GM 76 IW spectrum is used when a spectrum is needed (see Gregg 1989, for the form adopted here).

The linear surface wave field is described by modes having wave number k and angular frequency

$$\omega_k = \sqrt{gk}\,. \tag{5}$$

The spectrum of SW energy/unit area at a location x is

$$E_s(x,k,t)\,. \tag{6}$$

The spectrum of ocean density is then

$$F_s(x,k,t) = E_s(x,k,t/\omega_k\,). \tag{7}$$

In the absence of SW modulation by the IW current U, we assume an "ambient" SW spectrum: $E_a(k)$, $F_a(k)$. The surface wave "modulation" M is defined as

$$M(x,k,t) = F_s/F_a\,. \tag{8}$$

For our applications, we have evaluated F_a from the models of Phillips (1985), Donelan et al. (1985), and Banner (1990). Our calculations are not sensitive to the minor differences among these models.

3. The Wave Interaction Model

The resonant triad condition for two SW modes having wavenumbers k and k' and one IW mode with wavenumber K is

$$k - k' = K \,.$$
$$\omega_k - \omega_k = \Omega\,(j,K) \,. \tag{9}$$

it is realistic to assume that

$$K << k \,.$$
$$\Omega << \omega_k \,. \tag{10}$$

These conditions let us write the second of Eqs. (9) in the form

$$c_g(k) \bullet K = \Omega \,. \tag{11}$$

where c_g is then SW group velocity. The equation

$$\left[c_g(k) + U\right] K = \Omega \tag{12}$$

is an obvious generalization of Eq. (11) and also represents the condition for a raypath turning point for SW's propagating through the current U.

The "standard" deviation of the SW/IW interaction as a weakly nonlinear resonant triad coupling begins with the fluid equations:

$$\rho \frac{dv}{dt} = -\nabla p + \rho g \,.$$
$$\frac{d\rho}{dt} = 0 \,. \tag{13}$$

Here ρ is the fluid density, $v = (u,w)$ the fluid velocity, p the pressure, and g the acceleration of gravity. The boundary conditions at the ocean surface are

$$\frac{d\varsigma}{dt} = w \text{ at } z = \varsigma\,(x,t), \text{ the ocean surface}$$

$$p(x,t) = \text{constant at } z = \varsigma \,. \tag{14}$$

To include triad interactions, Eqs. (13) and (14) are expanded to second order in wave amplitudes. Appropriate ensemble averaging and cumulant discard (see, for example, Olbers and Herterich, 1979) provide expressions for energy flow between SW and IW spectra.

The above procedure leads to the two mechanisms noted in the Introduction: "spontaneous creation" and the "modulation mechanism". A simple derivation of the energy exchange rate for the modulation mechanism was given in Watson (1990). We repeat this here. The conditions (9) let us use the transport equation:[3]

$$\left[\frac{d}{dt} + \dot{x} \bullet \nabla_x + k \bullet \nabla_k\right] F_k = 0 \,, \tag{15}$$

where

$$\dot{x} = \nabla_k H,k = -\nabla_x H,\ H = \omega_k + k \quad U\,. \tag{16}$$

On multiplying (15) by ω_k and integrating over k, we obtain

$$\dot{E}_s = \int \frac{d^2x}{A_0} \int d^2k k \bullet \dot{U}(x,t)\, F_s(x,k,t) = \dot{E}_i\,, \tag{17}$$

where the integration is taken over some area A_O large compared with the IW wavelength. Since U is oscillatory, we see that Eq. (17) depends on the value of Eq. (8).

The modulation of M may be evaluated in a straightforward manner from Eqs. (15) and (16), as was shown in Watson (1990). The rate of energy exchange between fields can then be obtained from Eq. (17).

When only a single IW mode is considered, we can define an e-folding rate using Eq. (17):

$$\dot{E}_i(K,j) = \nu(K,j)\, E_i(K,j)\,. \tag{18}$$

The sign of ν determines whether energy flows to or from the IW field. To obtain energy transfer for a finite IW spectral domain, an appropriate sum over K and j can be evaluated.

4. Energy Transfer Rates

To evaluate the significance of the energy transfer, seasonally averaged Vaisala profiles were obtained for approximately 25 locations in the North Pacific Ocean. Corresponding seasonally averaged wind profiles were also obtained for the same locations. An example of the Vaisala profiles is shown in Fig. 2.

The dependence of the energy transfer rate (W m^{-2}) for a summer profile is illustrated in Fig. 3. For wind speeds less than 17 m/s the transfer of energy is to the SW field. At higher wind speeds the IW field receives energy from the SW field. Since no seasonally averaged winds above 15 m/s were found, we conclude that the mean transfer of energy is from the IW field to the SW field. For summer profiles this is at a rate which seems to be significant for the net IW energy balance. For winter Vaisala profiles the rates are reduced by about an order of magnitude. This is illustrated in Figs. 4 and 5 where the rates of energy loss are shown as a function of horizontal wavelength and for modes 1 and 2.

A more detailed discussion of the implications of this process will be published elsewhere.

[3] A phenomenological model for surface wave damping is included in Eq. (15). This has little influence on the calculated transfer rates.

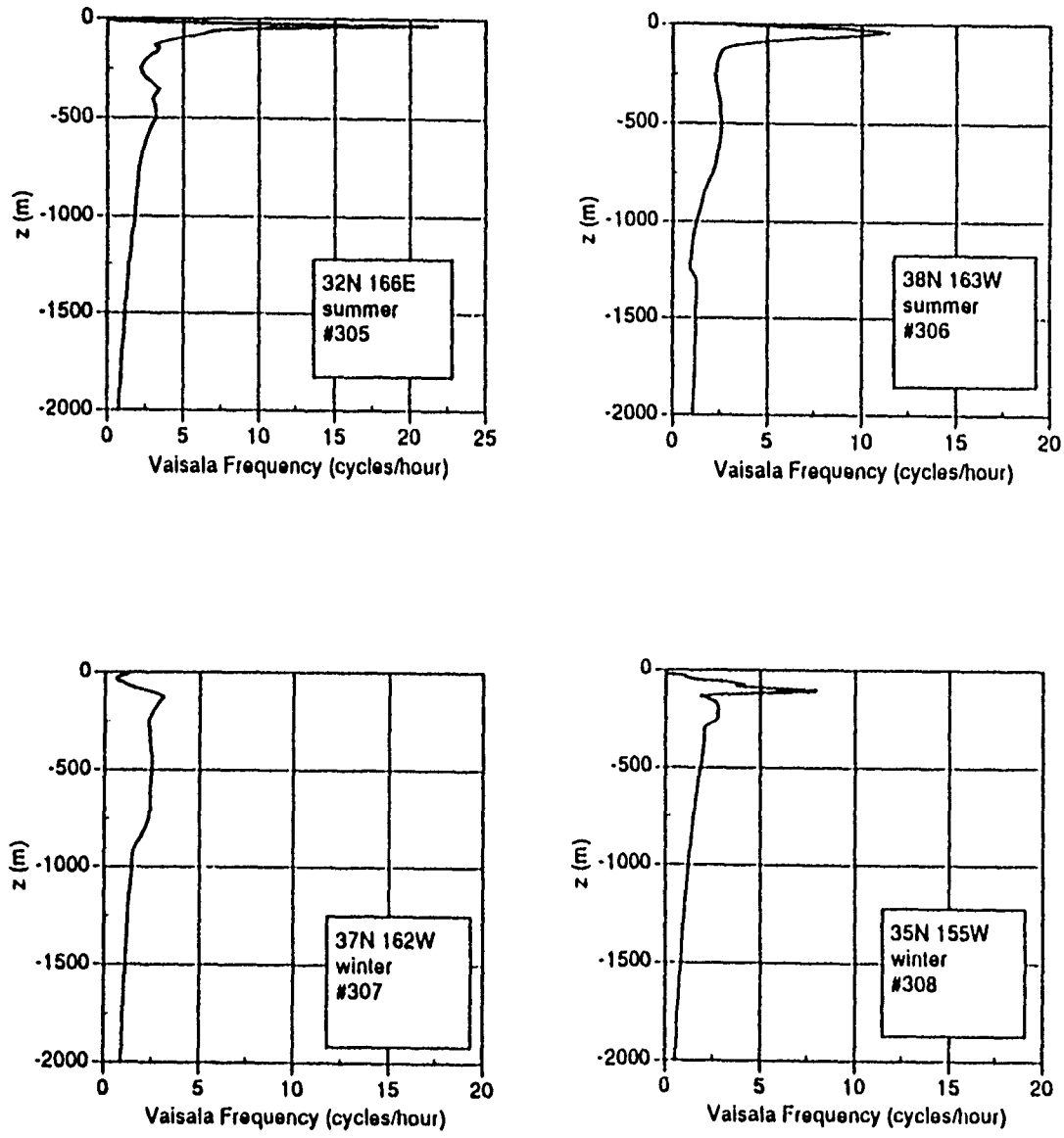


Fig. 2 Seasonally averaged Vaisala frequency profiles.

Fig. 3 The energy transfer rate to the IW field as a function of wind speed.

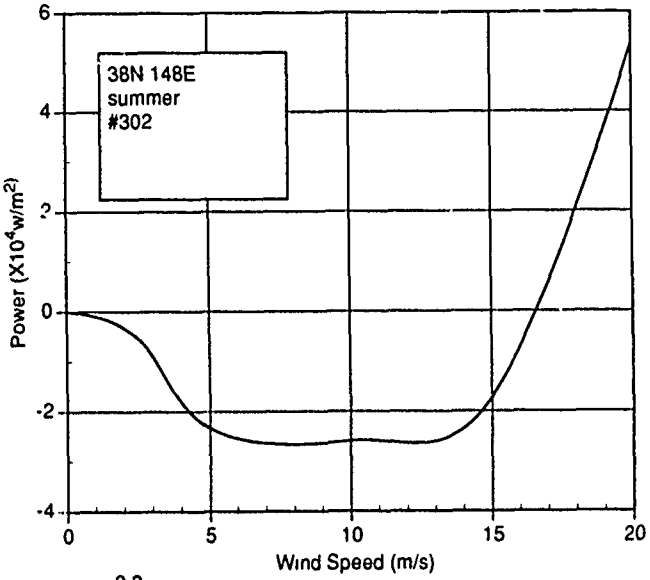
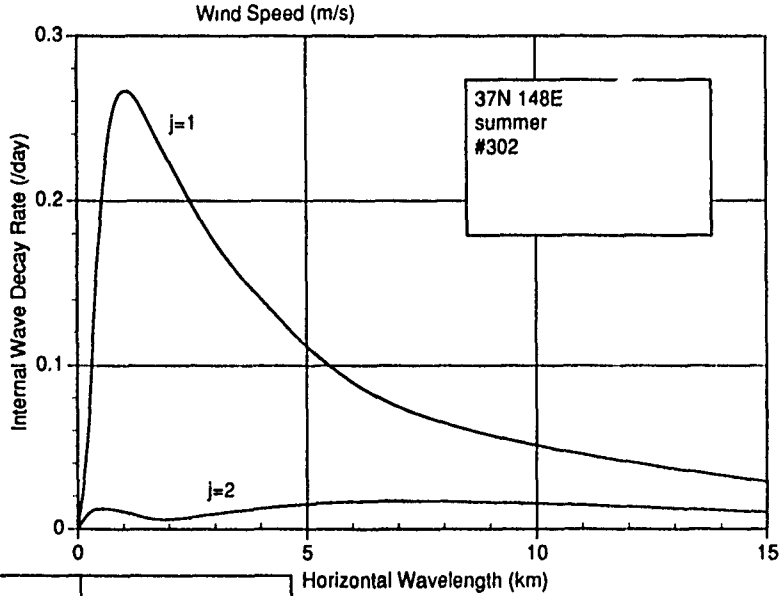


Fig. 4 The e-folding rate for energy loss by the IW field for a summer Vaisala profile.



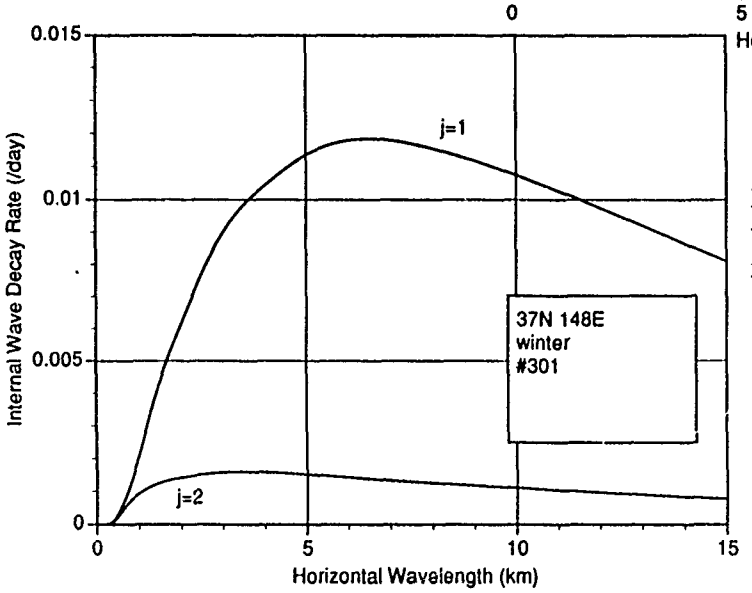


Fig. 5. The e-folding rate for the energy loss by the IW field for a winter Vaisala profile.

References

Banner, M.L., 1990: Equilibrium spectra of wind waves. J. Phys. Oceanogr. 20, 966-983.

Bell, T. H., 1975: Topographically generated internal waves in the open ocean. J. Geophys. Res. 80, 320-327.

Bell, T. H., 1978: Radiation damping of enertial oscillations in the upper ocean. J. Fluid Mech. 88, 289-308.

Donelan, J. J. Hamilton and W. Hui, 1985: Directional spectra of wind generated waves. Phil. Trans. Roy Soc. London. A315, 509-562.

Dysthe, K. B. and K. P. Das, 1981: Coupling between a surface wave spectrum and an internal wave: Modulation interaction. J. Fluid Mech. 104, 483-503.

Flatte', S. M., F. S. Henyey and J. A. Wright, 1985: Eikonal calculations of short wavelength internal wave spectra. J. Geophys. Res. 90, 7265-7272.

Gargett, A. E., P. J. Hendricks, T. B. Sanford, T. R. Osborn and A. J. Williams, 1981: A composite spectrum of vertical shear in the upper ocean. J. Phys. Oceanogr. 11, 1258-1271.

Garrett, C. and W. Munk, 1972a: Space-time scales of internal waves. Geophys. Fluid Dyn. 2, 225-264.

Garrett, C. and W. Munk, 1972b: Ocean mixing by breaking internal waves. Deep-Sea Res. 19, 823-832.

Greg M., 1989: Scaling dissipation in the thermocline. J. Geophys. Res. 19, 823-832.

Gregg, M. E., D. Asaro, T. Shay and N. Larson, 1986: Observations of persistent mixing and near-inertial internal waves. J. Phys. Oceanogr. 16, 856-885.

Kanthu, L. H., 1979: On generation of internal waves by turbulence in the mixed layer. J. Dyn. Atmos. Oceans. 3, 39-46.

Lueck, R. G., W. R. Crawford and T. R. Osborn, 1983: Turbulent dissipation over the continental slope off Vancouver Island. J. Phys. Oceanogr. 13, 809-1818.

McComas, C. H., 1978: Equilibrium mechanisms within the oceanic internal wave field. J. Phys. Oceanogr. 7, 836-845.

McComas, C. H. and P. Muller, 1981: The dynamic balance of internal waves. J. Phys. Oceanogr. 11, 970-986.

Olbers, D. J. and K. Herterich, 1979: The spectral energy transfer from surface waves to internal waves. J. Fluid Mech. 92, 349-380.

Osborn, T., 1978: Measurements of energy dissipation adjacent to an island. J. Geophys. Res. 83, 2939-2957.

Phillips, O. M., 1985: Spectral and statistical properties of the equilibrium range in wind-generated gravity waves. J. Fluid Mech. 156, 505-531.

Pomphrey, N., J. D. Meiss and K. M. Watson, 1980: Description of nonlinear wave interactions using Langevin methods. J. Geophys. Res. 85, 1085-1094.

Watson, K. M., 1985: Interaction between internal waves and mesoscale flows. J. Phys. Oceanogr. 15, 296-1311.

Watson, K. M., 1990: The coupling of surface and internal gravity waves: Revisited. J. Phys. Oceanogr. 20, 1233-1248.

Watson, K., B. West and B. Cohen, 1976: Coupling of surface and internal gravity waves: a mode coupling model. J. Fluid Mech. 77, 85-208.

THE SATURATION OF MIDDLE-ATMOSPHERE GRAVITY WAVES

Colin O. Hines

Arecibo Observatory, P.O. Box 995, Arecibo, Puerto Rico 00613

1. INTRODUCTION

A broad spectrum of gravity waves propagates through the middle atmosphere, analogous to that found in the oceans. As in the oceans, this spectrum exhibits a degree of universality over a decade or more in vertical wavenumber m: the wind power spectral density is approximately of the form $KN_o^2m^{-3}$ over the universal range, where N_o is the unperturbed buoyancy frequency and K is a constant variously taken to lie between 0.1 and 0.5. (In this paper, frequencies and wavenumbers are to be measured in radians per second and per meter, respectively, though it matters not for this particular purpose.) This "saturated" portion of the spectrum is found irrespective of meteorological conditions, time, place and height (Balsley and Carter 1982; Dewan et al. 1984; Tsuda et al. 1989; VanZandt 1982; Vincent 1984). The cause of saturation in the atmosphere is still a matter for debate, as is that in the oceans; the two may have — indeed, are likely to have — similar origins and differences of detail only.

The most frequently employed theory of saturation in the atmosphere attributes it to linear instability (Dewan and Good 1986, henceforth DG86; Smith et al. 1987, henceforth SFV87), though a recent paper by Weinstock (1990) makes an important challenge based on the diffusive dissipation imposed by nonlinear combination of the smaller-scale waves. The present paper outlines yet another mechanism: wave-wave interaction imposed by the advective nonlinearity of the Eulerian fluid-dynamic relations. It adapts and summarizes a three-part set of papers recently accepted, and a fourth part recently submitted, for publication (Hines, 1991a,b,c,d, henceforth H91a,b,c,d, respectively). The three parts correspond to Sections 2, 3 and 4 respectively, with the fourth part included in Section 3, and an over-all discussion is presented in Section 5.

To set the stage for oceanographers, I should remark here on two important differences between the atmospheric and oceanic cases: (1) The wave spectrum in the middle atmosphere is believed to be dominated by waves propagating their energy upward from sources at lower levels (such as winds blowing over mountains, moving cold fronts, shear in jet streams, and tropospheric convection that penetrates to or through the tropopause). Partial internal reflections may occur in the middle atmosphere, but there is no strongly reflecting single level analogous to the surface of the ocean. Instead, the upper levels act primarily as a dissipative region, the effective diffusion coefficients increasing with height z. (2) Wave amplitudes tend also to increase with z, in response to the decrease of gas density. (The amplitudes

tend to grow as exp z/2H while the density decreases as exp - z/H, where the scale height H is about 7 km.) In consequence, the wave-induced wind variance would be expected to grow by some five or six orders of magnitude on ascending from the tropopause (at heights of 8 - 16 km) to the turbopause (at heights of 100 - 110 km), were it not for the limiting effects of saturation and dissipation; in fact it increases by about three orders of magnitude only, but this still constitutes a wide range. The factor K in the saturated portion of the spectrum is found to remain relatively constant despite these wide ranges of anticipated and observed intensities, but the saturated range of wavenumber m is found to shift, with at least the lower bound progressing to smaller m with increase of height (e.g., SFV87). These features open to the atmospherist a degree of freedom not available to the oceanographer, one that may inspire new modes of attack on the problem of saturation and one that provides for a testing of any new theory.

2. CRITIQUE OF LINEAR-INSTABILITY THEORY

The linear-instability theory of DG86 and SFV87 attributes saturation to instability engendered by the wave spectrum in consequence of the latter's growth with height. The theory is a descendant of an analysis by Hodges (1967) in which a monochromatic, upgoing gravity wave was considered. I shall represent such a wave as having phase variations given by [kx + ly - mz - ωt], with z the upward coordinate, m and ω positive, and I shall term it a single "mode", there being no overlying reflector to produce a downgoing complement.

Hodges determined the Richardson number Ri as a function of phase for a single mode under approximations appropriate to much of the atmospheric spectrum. These approximations, which I also adopt, produce the dispersion relation

$$(\omega/h)^2 = (N_o/m)^2 \qquad (1)$$

as in an incompressible medium (neglecting Earth's rotation), h = + $(k^2 + l^2)^{1/2}$ being the horizontal wavenumber, taken to be << m. He found that the minimum Ri in a single mode could fall below 1/4 (and so produce dynamic instability) only if it in fact fell below 0 at an appropriate phase (and so produced, at that phase, static -- or convective -- instability). The required condition for instability could be stated as $\sigma_s^2 = 0.5$, where σ_s is the standard deviation of the wave-induced shear, nondimensionalized (here and henceforth) by division by N_o. At greater heights, the wave amplitude was expected to remain constant (be "saturated", in later parlance), or perhaps even be reduced, in consequence of the transfer of wave energy to turbulence energy.

In DG86, arguments previously presented by Phillips (1977) for ocean waves were adapted to the atmospheric case. The wave-induced winds of the saturated portion of the atmospheric spectrum were taken to be produced by a succession of wave groups (structured in the vertical) having a range of vertical wavenumbers Δm proportional to m, and these were taken to become unstable -- statically or dynamically -- when their σ_s reached some critical value taken

to be of order unity. With this critical value independent of m, DG86 found that the corresponding wind variance must be proportional to $N_o^2\ m^{-2}$, and so the power spectral density within the Δm group must be proportional to $N_o^2\ m^{-3}$, the factor of proportionality being assumed to be of order unity. They then attributed this spectral form to the observable saturated wind spectrum as a whole, and thereby provided an explanation for the observations. A second argument by DG86 obtained the same spectral form on dimensional grounds, with the assumption that N_o represented the only background atmospheric parameter and m (taken to be $\gg$ h) the only spectral parameter relevant to instability: the dimensions of the power spectral density of the winds then required that the density itself be proportional to $N_o^2\ m^{-3}$ as before. The same spectral form was adopted by SFV87, based only on a reference back to DG86, not on any further argument.

The first argument of DG86 (and so that of Phillips) is subject to severe criticism even if one accepts the assumptions that go into it: it in fact requires yet another assumption -- a hidden assumption to date -- that the postulated wave groups must enter the observations with equal frequency of occurrence across the m spectrum. I see no basis for such an assumption, even if the $\Delta m \propto m$ assumption might be justified on some scaling grounds (which I doubt), and so I cannot accept that the first argument is relevant. The second, dimensional argument will be valid if its assumptions are, but one of those assumptions is that wave instability is indeed the mechanism that shapes the spectrum: consistency is found, which is a necessary but not a sufficient finding to establish wave instability as the relevant mechanism. The theory of stratified turbulence, for example, as in Lumley (1964), leads to the same spectral form, as does the diffusive wave nonlinearity of Weinstock (1990). Below, I shall argue that the total wind standard deviation σ_T is a relevant parameter, thereby negating the purely dimensional argument, and indeed I shall argue that the m^{-3} form is only an approximate accident anyway, subject to some change from case to case and plausibly from atmosphere to ocean.

In DG86, the multiplier of $N_o^2\ m^{-3}$ was left unspecified other than that it be roughly of order unity (probably). In SFV87, on the other hand, a critical value of 0.5 was adopted for σ_s^2 by analogy with the case of the single, monochromatic wave. This critical value was first applied to a wave group with $\Delta m \sim m$ and shown to produce $N_o^2/2m^3$ as the requisite spectral power density. Subsequently, a model spectrum of the form $(1 + [m/m_*]^3)^{-1}$ was assumed, m_* being a characteristic vertical wavenumber determining the transition from an unsaturated (flat) portion of the wind spectrum at smaller m to the saturated, large-m, m^{-3} tail spectrum. With a further assumption as to the length of the tail — it was taken to extend to about 200 m_* — and of the form of the frequency spectrum (assumed separable), the shear power spectral density was integrated over m and set equal to the assumed critical shear variance, 0.5. This procedure determined the absolute intensity of the spectral form and led to a wind power spectrum of $N_o^2/6m^3$ in the saturated tail. This spectrum was claimed to be in good agreement with observations — better agreement, in fact, than that given by the $N_o^2/2m^3$ result of the $\Delta m = m$ "narrow-band" group. The agreement was taken as evidence not only in favor of the instability theory, but also of the requirement to integrate the shear over the full spectrum when establishing the condition for instability —

thereby denying the individual-group model on which the form of the spectrum had first been based.

Here again, the favorable conclusion is one of compatibility rather than determinacy. It could be turned quite the other way around: the wave spectrum might be established in the model form by processes other than instability, and the SFV87 calculation would then merely establish that the modeled spectrum, if extended to the assumed maximum m, would indeed be unstable.

Since the linear-instability theory maintains consistency with observation, it cannot be disproven by the criticisms I have made here; but the strength of the arguments in favor of that theory should, I think, have been undermined. In any event, given that theory, a further point must be made: the linear-instability theory has been developed to date on the basis of linear wave theory, but the model adopted to illustrate it reveals that the waves are highly nonlinear.

If, for example, one adopts as the wind power spectral density WPSD the form

$$WPSD = (N_o^2/6m_*^3) / (1 + [m/m_*]^3), \tag{2}$$

favored by SFV87, one finds upon integration over m the wind variance

$$\sigma_T^2 = (\pi/9\sqrt{3}) \ (N_o^2/m_*^2) = 0.20\,(\omega/h)_*^2, \tag{3}$$

where σ_T is the wind standard deviation as before and $(\omega/h)_*$ is the horizontal trace speed of a wave with the characteristic vertical wavenumber m_*, use having been made of (1). This reveals that, at m_*, the horizontal trace speed is little more than twice the standard deviation of the wave-induced wind field through which the wave is propagating. Waves having greater m will have proportionately smaller horizontal trace speeds, via (1). At a point in space where the wave-induced wind $\underline{V}$ has a component V_h (assumed horizontal, for the moment) in the direction of the wave's propagation, the total time derivative of the Eulerian fluid-dynamic equations is

$$D/Dt = \partial/\partial t + \underline{V}.\nabla = (1 + V_h/[\omega/h])\,\partial/\partial t, \tag{4}$$

with the $V_h/[\omega/h]$ term representing a nonlinear interaction between the chosen wave and the whole of the wave-induced wind field (plus any background wind, in general). This interaction will clearly be of import to all waves except, perhaps, to those with $\omega/h \geq 2\ V_h$, which will be roughly those with $m \leq m_*$. Integration of the (nondimensionalized) shear power spectral density corresponding to (2), namely

$$SPSD = (m^2/6m_*^3) / (1 + [m/m_*]^3), \tag{5}$$

now shows that these relatively immune waves contribute only 0.04 to σ_s^2. If, as has been supposed, the critical value of σ_s^2 is of order unity (e.g., 0.5 as assumed by SFV87), it is evident that the requisite shears must come primarily from the tail portion of the wave spectrum, the portion that suffers strong nonlinear interaction. To take an extreme: at the tip of the tail, where $m \simeq 200\ m_*$ according to the rough estimate of SFV87, the wind standard deviation would be about 72 times the horizontal trace speed. This confirms for the middle atmosphere a conclusion about the relevance of nonlinearity reached many years ago for the ocean (e.g., Holloway 1980, 1981, Munk 1981). (My own estimate of the length of the tail — to about 23 m_*, as is derived below — leads to a horizontal trace speed at the tip of the tail equal to about $\sigma_s/8$. This corresponds closely to Holloway's 1980 statement, that oceanic internal waves are too energetic by two orders of magnitude to be considered weak waves.)

These considerations reveal that, even if one wanted to pursue an instability theory of saturation, one would in principle be forced to pursue that theory nonlinearly, specifically with the effects of the advective nonlinearity $\underline{V}.\nabla$ being taken into account. As I shall argue below, the effects of this nonlinearity seem to be adequate in themselves to shape the tail into something like an m^{-3} form, at least if there is a dissipative process acting strongly at large m. This (secondary) process could be instability, if the tail extends to large enough m such that σ_s^2 attains the critical value, whatever that may be, but it could alternatively be molecular diffusion (as I shall argue it is, above the turbopause) or the nonlinear, wave-wave diffusion examined by Weinstock (1990).

The analysis of H91a includes a derivation of the probabilities of insipient instability (i.e., the probabilities of finding $Ri < 1/4$ for dynamic instability and < 0 for static instability) for an azimuthally isotropic, Gaussian distribution of wave-induced winds. It shows that each probability increases continuously from 0 as σ_s increases from 0 and reaches appreciable levels (of order 0.1) for σ_s of order unity, thereby confirming the criteria adopted by DG86 and SFV87 for their critical shears, and providing a firm base for future application.

3. DEVELOPMENT OF DOPPLER-SPREAD THEORY

The intent of this section is to make an analytic estimate, necessarily crude, of the consequences of the advective nonlinearity in a spectrum of waves such as the middle atmosphere supports. For the purpose, I assume an atmosphere that is wind-free and nearly isothermal in the absence of the waves, gravitationally stratified and nonrotating. I further assume a power spectral density of x-component (u) and y-component (v) of wave-induced wind given by

$$Q_u^2 = (k/h)^2\ Q^2\{k, l, m\} = \cos^2\alpha\ Q^2\{\alpha, h, m\} \tag{6}$$

and

$$Q_v^2 = (l/h)^2\ Q^2\{k, l, m\} = \sin^2\alpha\ \ Q^2\{\alpha, h, m\} \tag{7}$$

respectively, where $\alpha \equiv \arccos k/h = \arcsin l/h$ is the azimuth of wave propagation and Q^2 is the power spectral density of the wave-induced wind in the azimuth of the wave's own propagation. The notation within braces is intended to indicate that Q^2 may be thought of as a function of horizontal-component wavenumbers k and l, or azimuth α and horizontal wavenumber h, in addition to vertical wavenumber m.

The spectral components are taken to be randomly phased relative to one another, and the spectrum is taken to be broad, in which case the Central Limit Theorem establishes that the horizontal wind components will have a Gaussian distribution: the probability of finding u between u and u + du is given by

$$P_u\{u\}\, du = \frac{1}{\sqrt{2\pi V_u}} \exp\{-u^2/2V_u\}\, du, \tag{8}$$

where V_u is the variance of u, given by

$$\begin{aligned} V_u = \sigma_u^2 &= \iiint (k/h)^2\, Q^2\{k,l,m\}\, dk.dl.dm \\ &= \iiint \cos^2\alpha\, Q^2\{\alpha,h,m\}\, h\, d\alpha.dh.dm \end{aligned} \tag{9}$$

and likewise for the y-component. (The spectrum is taken to contain only upgoing waves — that is, waves with positive m, under present convention — and the integration over m is correspondingly restricted.) For convenience, I shall assume an azimuthally isotropic spectrum, so that Q^2 is independent of azimuth and the integration over α may be conducted trivially. Then each of the two variances is equal to half the total (horizontal) wind variance σ_T^2:

$$2V_u = 2V_v = \sigma_T^2 = 2\pi \iint Q^2\{h,m\}\, h\, dh.dm \tag{10}$$

and

$$P_u\{u\} = P\{u\} \equiv \frac{1}{\sqrt{\pi}\,\sigma_T} \exp\{-u^2/\sigma_T^2\}. \tag{11}$$

Similarly for $P_v\{v\}$.

I now assume that a spectrum of the type assumed is incident on the middle atmosphere from below, where the waves can be taken to be non-interacting — i.e., linear wave theory applies. The Q^2 of the spectrum there will be denoted by Q_i^2, the "i" indicating "initial". The task now is to determine how this spectrum will alter with height, as wave amplitude increases and the advective nonlinearity comes into play.

For the purpose, I consider a small packet of waves in the middle atmosphere having wavenumbers in the bin $\Delta k.\Delta l.\Delta m$ at {k,l,m}. This is not to be construed as a physical wave packet, engendered by some particular source, but

rather just a mathematical construct deriving from the waves that happen to be at hand. I take it to be propagating through the irregular wind system provided by the full wave spectrum, but I take the effect of that wind system to be introduced only by the horizontal component of wind and to be treatable, in its statistical consequences, as if that component were a "background" wind, unvarying in time or horizontal position. These restrictions will limit the accuracy and to some extent the validity of the form of subsequent relations, to be sure, but it is to be hoped that they do not invalidate the general (statistical) tendencies that will be revealed.

Given these assumptions, the chosen wave packet will retain its α and h unchanged; I shall take $\alpha = 0$ for convenience at the start. In the underlying region, the packet's intrinsic frequency will be ω_i and its vertical wavenumber will be

$$m_i = N_o k/\omega_i \tag{12}$$

from (1). At some height of interest, however, where the wave-induced wind field has x-component u, the intrinsic frequency will be Doppler shifted to

$$\omega = \omega_i - ku \tag{13}$$

and the dispersion relation (1) then establishes that

$$m^{-1} = m_i^{-1} - u/N_o. \tag{14}$$

That is to say, the spectral energy located in the bin $\Delta k.\Delta l.\Delta m$ at $(k,0,m)$ in wavenumber space, at a point in physical space-time where the x-component of the total perturbation wind is u, will have arrived there as a consequence of Doppler shifting from the vicinity of $(k,0,m_i)$ with m_i defined by (14). Though (13) permits ω and ω_i to differ in sign, such an occurrence would indicate Doppler shifting through the critical condition $\omega = 0$ at which extreme dissipation is anticipated; it will not be admitted in the present work. Correspondingly, with m_i restricted to positive values, m will be likewise restricted.

It is known from previous work (e.g., Hines and Reddy 1967) that $(u')^2\ m^{-1}$ must be height-invariant as a single mode (with perturbation wind u', vertical wavenumber m) propagates through a background wind system, if reflections are ignored (as in a WKBJ treatment). I take the spectral analogue of that conclusion to be that $Q^2\ m^{-1}\ dm$ must be invariant under Doppler shifting in the present case. (The standard exponential growth with height has not been taken into account here. It applies uniformly across the whole spectrum and so plays no explicit part in deforming the spectrum. It will be reintroduced later.) Thus, were the u of (14) the only u ever encountered, we would find that

$$Q^2 m^{-1}\ dm = Q_i^2\ m_i^{-1}\ dm_i \tag{15}$$

or

$$Q^2 = Q_i^2\ m\ m_i^{-1}(dm_i/dm) = Q_i^2\ m^{-1}\ m_i, \tag{16}$$

where Q_i^2 is the Q^2 of the initial spectrum and the derivative has been taken from (13) subject to u being held constant.

In fact, one finds a whole range of u's at a given point of space in the course of time; (16) is found only with a certain probability, given by $P_u(u)\,du$ for a small range du about u. This small range corresponds to a range dm_i about the initial m_i such that

$$du = |du/dm_i|\, dm_i = N_o\, m_i^{-2}\, dm_i, \tag{17}$$

in which the derivative has been taken from (14) subject to m being held constant. In combination with (11) and (16), and with integration over all contributing dm_i, this implies that the Q^2 appropriate to (k,0,m) is given by

$$Q^2(k,0,m) = \int Q_i^2(k,0,m_i)\, m^{-1}\, m_i^{-1}\, N_o\, P_u\{N_o(m_i^{-1} - m^{-1})\}\, dm_i. \tag{18}$$

In the present case of an azimuthally isotropic spectrum of waves, this may be rewritten in terms of $Q^2(h,m_i)$ and $Q_i^2(h,m_i)$, with P_u the corresponding isotropic probability, thus:

$$Q^2(h,m) = \int Q_i^2(h,m_i) \frac{N_o}{\sqrt{\pi}\,\sigma_T m m_i} \exp -\{N_o^2 (m_i^{-1}-m^{-1})^2/\sigma_T^2\}\, dm_i. \tag{19}$$

If the vertical wavenumbers are nondimensionalized via multiplication by σ_T/N_o, the transformation may be rewritten as

$$Q^2(h,M) = \int Q_i^2(h,M_i) \left[\frac{e^{-(M_i^{-1}-M^{-1})^2}}{\sqrt{\pi}\, M M_i}\right] dM_i, \tag{20}$$

where

$$M \equiv m\sigma_T/N_o \tag{21}$$

and similarly for M_i.

The factor in brackets in (20) may be thought of as a transfer function, leading from the initial to the observable spectrum in the scaled units. It is independent of h as well as azimuth, and so applies equally to the one-dimensional (in m) spectrum obtained by integration over h, the spectrum that exhibits an approximately m^{-3} portion in the middle atmosphere. Its consequences are indicated here with the aid of Figures 1 and 2.

Figure 1 exhibits the transfer function itself, as a function of M, for a number of values of M_i in a range about $M_i = 1$ (for which $\omega_i/h = \sigma_T$). This

function may be thought of as the statistically observable spectrum Q_δ^2 derived from a delta-function input spectrum at $M_i = M_\delta$, this position being marked by the abscissa of the heavy dot on each curve in turn. This observable spectrum is seen to be broadened -- or Doppler spread --, with its peak moving to an M value somewhat lower than M_δ and a tail extending to higher M. From the form of (20), it is apparent that this tail asymptotes to the form M^{-1}, and yet, for M_δ values somewhat smaller than 1, it exhibits a markedly stronger variation before reaching the asymptotic form. Indeed, for $M_\delta = 1/2$, it exhibits something close to an M^{-3} form over at least a decade in M, which would be consistent with what we know of the actually observed spectra. (The percentages associated with the individual curves will be explained shortly.)

Figure 2 exhibits the convolution Q_c^2 of the transfer function with a step-function input -- with an input spectrum that is flat (white noise) up to some cutoff value M_c (labeled and with abscissa marked by the heavy dot on each curve in turn) and is then cut off to zero. Again a tail is found, with the M^{-3} form occurring again and extending over at least a decade in M, now for $M_c = 1/2$.

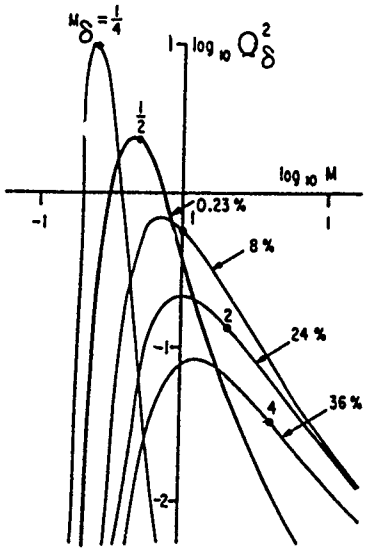


Figure 1

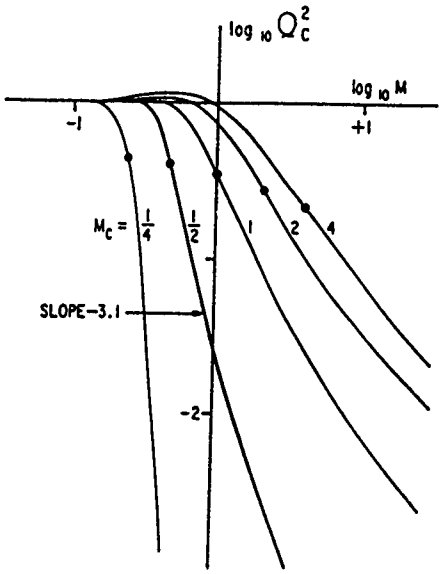


Figure 2

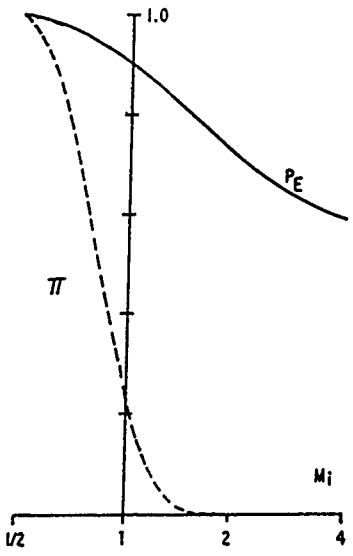


Figure 3

It would seem from these examples that the advective nonlinearity has a propensity for producing a semi-universal spectrum of the observed form, at least if some mechanism exists that would prevent incident waves having $M_i >$ 1/2 or 1 playing any substantial role in the observations. I suggest that the required mechanism is to be found in the approach of such waves to critical-layer interactions at and below the height of observation. As an approximation, I assume that waves with vertical wavenumber exceeding some maximum value m_M (scaled alternatively to M_M) are simply obliterated.

The probability of a wave with incident wavenumber M_1 having a local wavenumber less than m_M, and so escaping obliteration, is given by

$$P_E = 0.5 + 0.5\ erf\ \{N_o/m_i\sigma_T - N_o/m_m\sigma_T\}, \\ = 0.5 + 0.5\ erf\ \{M_i^{-1} - M_M^{-1}\} \qquad (22)$$

adapted from (2.22) of H91b, erf being the error function,

$$erf(x) = 2\pi^{-\frac{1}{2}} \int_o^x e^{q^2} dq .$$

This probability is depicted by the continuous curve in Figure 3 for the case $M_M = 11.5$, which will be justified below, although the critical-level case $M_M = \infty$ is not much different. (The percentages marked on Figure 1 represent the complementary probability, namely that waves of the respective M_δ values will indeed be obliterated.)

This P_E curve applies only locally, at the height currently of interest. To have survived to this height, the waves must have avoided obliteration at all lower levels. The probability of this successful escape is roughly the product of the probabilities of escape through each of the statistically independent underlying slabs of atmosphere. I take the depth of each such slab to be $\lambda_c/2$, where $\lambda_c = 2\pi/m_c$ is the cutoff vertical wavelength of the incident wave spectrum — the transition wavelength between the small-m body and the large-m tail spectra — and I number the underlying slabs with the index n, counting downward from the current height of interest. I also take σ_T to vary with height as exp (z/4H) and $m_M\sigma_T$ to be height-independent (see below). The probability of successful escape through all underlying slabs is then the product of individual-layer factors given by (22) for each in turn:

$$\Pi = \prod_{n=1}^{n_*} (0.5+0.5\ erf\ [N_o \exp\ (n\lambda_c/8H)/m_i\sigma_T - N_o/m_M\sigma_T]) , \qquad (23)$$

where n_* is the number of slabs between the height of observation and the source height. The Π curve of Figure 3 represents this probability for $M_M = 11.5$ again, with the use of $\lambda_c = 1.07$ km and H = 7 km, which are representative of the middle stratosphere (ca. 25 km height), and $n_* = 30$, corresponding to a source 16 km below. This curve clearly exhibits a sharp cutoff that might be approximated by a step function sited at some M_c in the range 0.5 – 1, as was wanted. (This calculation and these curves are newly presented here; they are not contained in H91b but are now submitted for publication as Part IV of the in-press series.)

A finite M_M, such as I have adopted, can be imposed by molecular diffusion (viscosity and heat conduction), and in the next section I take it to be so imposed above the turbopause. Below the turbopause, where (by definition) turbulence is to be found, linear instability might well provide the relevant mechanism. But it would be the mechanism, now, for limiting the length of the tail, not for determining the tail's form or intensity. Examination of this possibility in any detail analytically requires an analytic model spectrum.

In H91b, following VanZandt and Fritts (1989), I adopt the "modified Desaubies" spectral form

$$Q^2\{m\} = Km/(1+[m/m_c]^4)\,, \tag{24}$$

where m_c is a characteristic vertical wavenumber corresponding closely to m_* in Section 2, in that it marks the onset of the tail portion of the spectrum. The m_c notation is justified because this m_c is analogous also to the cutoff of incident wavenumbers in Figure 2, where too the tail may be said to begin. I take as my choice $m_c = N_o/2\sigma_T$ (or $M_c = 1/2$, in Figure 2), on the empirical grounds that it seems to represent as close to observed reality as a step-function approach to the cutoff is likely to come, but with legitimacy being provided by the Π curve of Figure 3. This choice combines with (24), from which σ_T^2 can be inferred by integration, to establish the value of K as $N_o^2/\pi m_c^4$. (The integral is only weakly dependent on the upper limit of integration m_M, provided that $m_M \gg m_c$ as I take it to be, and so integration to infinity is appropriate.) Then (24) becomes

$$Q^2\{m\} = N_o^2 m/\pi m_c^4 (1 + [m/m_c]^4)\,, \tag{25}$$

which approximates to $N_o^2/\pi m^3$ in its large-m tail. The intensity of this tail is fully consistent with observations (cf. $N_o^2/2m^3$ and $N_o^2/6m^3$ cited in Section 2), and it was derived with no assumption being needed as to the critical value of shear or even the existence of instability (as the linear-instability derivations required). Those assumptions are replaced here by the assumption of $N_o/2\sigma_T$ as an effective cutoff of incident wavenumbers -- an assumption justified by the continuous curve of Figure 3, at least for the middle stratosphere.

I would argue, further, that in the limit of small m the wave field would be unaffected by dissipation or Doppler spreading and hence the spectral density in that limit should grow with height as exp z/H. This assumption, imposed upon (25), leads directly to

$$m_c = m_o \exp -z/4H \tag{26}$$

and so also to

$$\sigma_T = (N_o/2m_o) \exp z/4H, \tag{27}$$

with m_o a constant that would be determined by the initial energy sources, specifically those producing small m_i. These height variations are in accord with observations as I know them; specifically, the wind variance now grows with height at half the rate appropriate to a nonsaturating, nondissipating spectrum, as it was said to do observationally in Section 1.

The (nondimensionalized) shear spectrum can be obtained from (25) via multiplication by m^2/N_o^2, and then integrated to obtain the shear variance σ_s^2. The integral is logarithmically divergent and must be terminated at some upper

bound m_M, beyond which the spectrum is of turbulence, not of waves. Then

$$\sigma_s^2 = (4\pi)^{-1} \ln \{1+[m_M/m_c]^4) = \pi^{-1}\ln \{m_M/m_c\} \quad (28)$$

where $m_M >> m_c$ has been assumed. Now, if instability is to determine m_M, as being m_{Minst} say, this shear variance must equal the critical value σ_{Scrit}^2 for the marginal maintenance of instability, whatever that may be. Hence,

$$m_{Minst} = m_c \exp \{\pi \sigma_{Scrit}^2\}. \quad (29)$$

If I take σ_{Scrit} to be 1, as suggested in Section 2, then (29) yields $m_{Minst}/m_c = \exp \pi = 23$: the tail extends over a 23-fold range in m. This is the value that leads to the P_E curve in Figure 3. The result $m_{Minst}/m_c = 23$ is consistent with the observations known to me, and specifically with those of Dewan et al. (1984), though perhaps for the unfortunate reason that the data do not extend reliably over a greater range. It is possible that (29) would be used best in the opposite direction, as a means of determining σ_{Scrit}^2 empirically from observed values of m_{Minst}/m_c.

In H91b, an appendix establishes that inclusion of the Coriolis force associated with Earth's rotation leads to a tail spectrum that asymptotes to m^{-2} (rather than m^{-1} as in (19)). A second appendix outlines a failed attempt to reach an m^{-3} form by the further inclusion of the wave-induced vertical wind field in the advective nonlinearity, and a third appendix deals briefly with azimuthally anisotropic conditions. I consider my failure to produce an m^{-3} asymptotic form to be serious only from one point of view, to be discussed in Section 5; specifically, I do not consider it to be serious for purposes of comparison with data. This is because I have obtained here, at least for the cases $M_\delta = 1/2$ (Figure 1) and $M_c = 1/2$ (Figure 2), the wanted m^{-3} form as a transitional form bridging the gap between low wavenumbers, which are unaffected by Doppler spreading, and high wavenumbers, which will be subject to dissipative processes and so will (probably) be altered from whatever asymptotic form they might otherwise have achieved. So far as I am aware, that is all that the observations demand.

4. FORMATION OF THE TURBOPAUSE

Molecular diffusion increases with height through the atmosphere and imposes an ultimate cutoff of the wave spectrum. I suggest that turbulence terminates — the turbopause is formed — when the cutoff imposed in this manner, which occurs at some m_M given roughly by

$$m_{Mmol} = (N_o h/2\pi\eta)^{\frac{1}{3}}, \quad (30)$$

limits the length of the tail to a value just less than that required for the occurrence of instability. (See H91c; η is the molecular kinematic

viscosity.) Equality of m_{Mmol} and m_{Minst} at the turbopause then requires there that

$$\sigma_{Scrit}^2 = \pi^{-1} \ln \{2\sigma_T N_o^{-2/3} \lambda_h^{-1/3} \eta^{-\frac{1}{3}}\}, \quad (31)$$

$\lambda_h = 2\pi h^{-1}$ being a (representative) horizontal wavelength. One would like to solve this equation for the height at which it is satisfied, thereby determining the height of the turbopause, but σ_{Scrit}^2 is as yet unknown to the accuracy required for a meaningful conclusion. Instead, I use it the other way around, as a means of estimating σ_{Scrit}^2.

The turbopause is typically taken to lie at a height of 100 - 110 km, where $\eta \simeq 100$ m^2/s and $N_o \simeq 0.02$ s^{-1}. If I adopt a representative $\lambda_h \simeq 50$ km and a representative $\sigma_T \simeq 30$ m/s, then (28) yields $\sigma_{Scrit}^2 = 0.497$, almost exactly the value chosen by SFV87. This seems to be a perfectly reasonable value of σ_s to accept as a condition for the termination of instability, particularly since the wave spectrum might well have been narrowed to a nearly monochromatic wave by the time turbopausal heights are reached. The same estimates combine with the modified Desaubies spectrum (25) to imply a spectral peak (which occurs at $m_p = 3^{-1/4} m_c = 0.76 m_c$) at a vertical wavelength of $\lambda_p = 25$ km. Such a value is frequently reported as dominating the spectrum at turbopausal heights, often with the (possibly false) identification of the observed wave as the diurnal 1,1 tidal mode. The maximum m, m_M, corresponds to a vertical wavelength of about 4 km, which is consistent with observations.

With the model spectrum taken to be applicable down to the tropopause, it can now be employed to estimate the wave spectrum in the middle stratosphere, where it can be compared with other data. There, both λ_p and σ_T will be decreased by the fourth root of the atmospheric density ratio, as given by (26) and (27), which root is approximately 18, yielding $\lambda_p \simeq 1.4$ km and $\sigma_T \simeq 1.7$ m/s. If m_{Minst}^{-1} is similarly scaled, it yields a transition from waves to turbulence at a vertical wavelength of about 220 m, whereas if σ_{Scrit} is raised to 1 before the scaling is done it yields a vertical wavelength of 46 m for the transition. The observations of Dewan et al. (1984), for example, are said to exhibit a tail with log-log slope of -3.0 extending from about 1 km down to about 200 m in vertical wavelength, with some curvature (consistent with that in Figure 1 above) at smaller wavelengths, down to about 40 m. The observations are said to be unreliable outside this range, but the theory is clearly compatible with the reliable observations. Moreover, the mean wind power spectral density at a vertical wavelength of 1 km was found to be 3.42 $(m/s)^2/(c/m)$, which converts to present units as giving a tail spectrum of $0.306 N_o^2/m^3$. This is to be compared with $N_o^2/\pi m^3 \simeq 0.318 N_o^2/m^3$ in the present theoretical model.

It should be specially noted that, above the turbopause, molecular diffusion simply replaces instability as the mechanism of dissipation, but the Doppler-spread theory continues on, otherwise unaffected in principle (until the spectrum becomes so narrow that a statistical treatment is inappropriate). There is, however, one side-effect that comes into play. At these great

heights, the statistically independent slabs of atmosphere are much thicker, so there will be fewer of them in the height range 4H over which σ_T decreases substantially (on moving downward from slab to slab). Consequently, the product probability Π in (23) is unlikely to be decreased from P_E in (22) by anything like the degree illustrated by the transition from the one curve to the other in Figure 3. The cutoff of incident waves will then not be as sharp: a step-function approximation may be inappropriate and the spectral form of the tail might be more like one of the forms depicted for values of M_c equal to or greater than 1 in Figure 2.

5. DISCUSSION

The present analysis has adopted a necessarily approximate approach to the estimation of the effects of Doppler spreading in the middle-atmosphere spectrum of gravity waves. Despite its failings, however, it seems to have shown that these effects are important -- are even necessary to include in any alternative theory of the spectral tail -- and on their own are of a nature that accords with observation, without the necessity for an alternative theory.

The importance of the advective nonlinearity that gives this Doppler spreading has been recognized in oceanographic studies for more than a decade now (e.g., reviews by Holloway 1980, 1981 and Munk 1981), and some progress has been made in numerical studies that incorporate its effects (e.g., Flatte et al. 1985). The present analysis might well be carried over to oceanographic studies, and perhaps improved, to account for the approximately m^{-3} spectral form in velocity (or m^{-1} spectral form in shear) that is found there over a middle range of vertical wavenumbers.

In making the transition, one would have to include both upgoing and downgoing waves, but that change will be of minimal operational consequence. One would also have to drop the exponential growth of wave amplitude with height, a change that would have at least two important consequences, one operational and one conceptual.

Operationally, the step that led from the continuous to the broken curve of Figure 3 will not have the exponential change with height that is allowed for in (23). Instead, the full ocean depth D will act uniformly to limit the probability of escape from obliteration, and that depth will be measured in some height-independent characteristic wavelength λ_c (~ 10 m) to give a very large number of statistically independent slabs. The relevant probability for escape from obliteration will be given by a modified form of (23) in which all factors are identical and the required probability becomes simply the probability of escape from obliteration in one slab raised to the power $4D/\lambda_c$, a very large number and one that will produce an even closer approach to a step-function form than that given by the Π curve in Figure 3. (The relevant power will be even greater than $4D/\lambda_c$, if it is held that the borderline waves can propagate up and down more than once before being obliterated. The relevant power may, moreover, change somewhat with D from case to case and so perhaps produce cutoffs and spectral slopes that similarly change somewhat, observationally. These questions open the way to further investigation.)

Conceptually, the physical growth of wave amplitude with height that occurs in the atmosphere must be replaced, in oceanographic work, by the imagined growth that would have occurred, had the spectrum of waves been "turned on" with a gradually increasing intensity, a gradually increasing nonlinearity. Or, alternatively, it might be replaced simply by a straightforward calculation of the steady-state nonlinear result, if such a calculation can be said to be straightforward.

Happily enough, such a basis for calculation has in fact been laid already by Allen and Joseph (1989), but it was developed in different terms. That work adopts (in its Case III) a canonical spectrum of waves as described in a Lagrangian formulation and evaluates the appearance of the spectrum as it would be seen in Eulerian coordinates. The waves are fully linear in the Lagrangian description -- their frequencies and wavenumbers lie on the dispersion-equation surface in 4-space -- but are allowed to appear nonlinear in the Eulerian description, with the nonlinearity arising from the advective nonlinearity of the Eulerian equations only. That is precisely the transition that I have attempted to model here, beginning with the initial spectrum from which nonlinearities were excluded and ending with an observable spectrum in which the effects of the advective nonlinearity are included (albeit for reasons of amplitude growth with height, rather than a transition from a Lagrangian to an Eulerian description).

The Allen and Joseph work reaches (amongst other things) an m^{-3} form for the ultimate tail spectrum, a thing I have been unable to do. As noted above, inclusion of the Coriolis force (which Allen and Joseph included) would have led me to an m^{-2} form, but that is not enough. Allen and Joseph automatically included also the part of the advective nonlinearity that comes from vertical advection, a thing I have been unable to achieve successfully, and I suppose my failure is a consequence of my oversimplification of the means of handling the Doppler shifts. This leaves a missing link in the chain connecting my work to theirs, but conceptually the two approaches seem identical in the intended transition they incorporate.

The Allen and Joseph work as developed to date does not give the form of transition from the small-m body to the large-m tail of the wave spectrum, which is what my own analysis does succeed in doing, however approximately. It is this transition, rather than an asymptotic tail (which would in any event be deformed by dissipative processes), that is seen in the middle atmosphere, I believe. Perhaps it is this same transition in the oceans, as well.

Acknowledgments: I wish to thank Carmen Torres for her assistance in the production of this typescript. The Arecibo Observatory is operated by Cornell University with funding provided under a cooperative agreement with the National Science Foundation.

REFERENCES

Allen, K. R., and R. I. Joseph, 1989: A canonical statistical theory of oceanic internal waves. *J. Fluid Mech.*, **204**, 185–228.

Balsley, B. B., and D. A. Carter, 1982: The spectrum of atmospheric velocity fluctuations at 8 km and 86 km. *Geophys. Res. Lett.*, **9**, 465–468.

Dewan, E. M., and R. E. Good, 1986: Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere. *J. Geophys. Res.*, **91**, 2742–2748.

Dewan, E. M., N. Grossbard, A. F. Queseda and R. E. Good, 1984: Spectral analysis of 10m resolution scalar velocity profiles in the stratosphere. *Geophys. Res. Lett.*, **11**, 80–83 and 624.

Flatte, S. M., F. S. Henyey and J. A. Wright, 1985: Eikonal calculations of short-wavelength internal-wave spectra. *J. Geophys. Res.*, **90**, 7265–7272.

Hines, C. O., 1991a: The saturation of gravity waves in the middle atmosphere. Part I: Critique of linear-instability theory. *J. Atmos. Sci.*, in press.

Hines, C. O., 1991b: The saturation of gravity waves in the middle atmosphere. Part II: Development of Doppler-spread theory. *J. Atmos. Sci.*, in press.

Hines, C. O., 1991c: The saturation of gravity waves in the middle atmosphere. Part III: Formation of the turbopause and of turbulent layers beneath it. *J. Atmos. Sci.*, in press.

Hines, C.O., 1991d: The saturation of gravity waves in the middle atmosphere. Part IV: The cutoff of the incident wave spectrum. Submitted for publication, *J. Atmos. Sci.*

Hines, C. O., and C. A. Reddy, 1967: On the propagation of atmospheric gravity waves through regions of wind shear. *J. Geophys. Res.*, **72**, 1015–1034.

Hodges, R. R. Jr., 1967: Generation of turbulence in the upper atmosphere by internal gravity waves. *J. Geophys. Res.*, **72**, 3455–3458.

Holloway, G., 1980: Oceanic internal waves are not weak waves. *J. Phys. Ocean*, **10**, 906–914.

Holloway, G., 1981: Theoretical approaches to interactions among internal waves, turbulence and fine structure. *Nonlinear Properties of Internal Waves* (Ed: B. West). American Institute of Physics.

Lumley, J. L., 1964: The spectrum of nearly inertial turbulence in a stably stratified fluid. *J. Atmos. Sci.*, **21**, 99–102.

Munk, W., 1981: Internal waves and small-scale processes. *Evolution of Physical Oceanography* (Eds: B. A. Warren and C. Wunsch). MIT Press.

Phillips, O.M., 1977: The Dynamics of the Upper Ocean, 2nd ed., Cambridge University Press, New York.

Smith, S. A., D. C. Fritts and T. E. VanZandt, 1987: Evidence for a staurated spectrum of atmospheric gravity waves. *J. Atmos. Sci.*, 44, 1404-1410.

Tsuda, T., T. Inoue, D.C. Fritts, T.E. Van Zandt, S. Kato, T. Sato and S. Fukao, 1989: MST radar observations of a saturated gravity wave spectrum. *J. Atmos. Sci.*, 46, 2440-2447.

Van Zandt, T.E., 1982: A universal spectrum of buoyancy waves in the atmosphere. *Geophys. Res. Lett.*, 9, 575-578.

Vincent, R. A., 1984: Gravity wave motions in the thermosphere. *J. Atmos. Terr. Phys.*, 46, 119-128.

Weinstock, J., 1990: Saturated and unsaturated spectra of gravity waves, and scale dependent diffusion. *J. Atmos. Sci.*, 47, 2211-2225.

DIAGNOSING DIAPYCNAL MIXING

Kraig B. Winters and Eric A. D'Asaro

Applied Physics Laboratory
Departments of Applied Mathematics and Oceanography
University of Washington, Seattle, WA 98105

ABSTRACT

The diapycnal mixing associated with a small scale mixing event is discussed. A direct numerical simulation of a wave packet propagating through a density stratified shear flow and breaking at a critical level is used to illustrate the issues involved in diagnosing mixing. An indirect approach, based on the density variance equation, is shown to be an ambiguous indicator of mixing. A direct approach, based on the simple idea that net mixing implies a change in the volume of fluid enclosed by a given pair of isopycnals is presented, along with a potential energy analysis and an estimate of the mixing efficiency of the event.

INTRODUCTION

The role of diapycnal mixing in maintaining the large scale heat balance is an important unresolved issue in small scale oceanography. Although the relative contributions to the overall mixing from the interior and at boundaries is in question, much of the diapycnal mixing in the ocean interior is thought to result from isolated, intermittent turbulent "events". Quantifying the mixing associated with a given event is not an easy task. Observational efforts have attempted to determine the diapycnal diffusivity of mass K_d. Generally, measurements of the dissipation rates of kinetic energy or temperature variance are used to infer net vertical fluxes or diffusivity, based on assumed dynamical balances. A review of these techniques can be found in Gregg (1987) or in Moum (1989). The validity of the assumed balances, however, has not yet been well established for the variety of dynamical mechanisms capable of producing turbulent mixing events. From a theoretical viewpoint, it is not clear what type of sampling and averaging is necessary to separate the reversible small-scale internal wave fluxes from the turbulent fluxes.

Large scale ocean circulation models depend critically on details of the parameterization of small scale mixing. Understanding the diapycnal mixing of individual events is an important early step toward the eventual parameterization of small-scale processes and predicting their effects on the large-scale ocean circulation. In this paper, we will focus on the issue of diagnosing the diapycnal mixing of an isolated turbulent mixing event. A high resolution, three-dimensional primitive equation model will be used to generate a mixing event, similar to what may be seen in the ocean

interior. The simulation is similar in spirit to the two-dimensional calculation of Winters and D'Asaro (1989). In the simulated flow, a spatially isolated internal wave packet propagates vertically into a horizontal shear flow. The wave packet is refracted by the background shear, ultimately reaching a critical level where the horizontal phase speed matches the ambient flow speed. Near the critical level, the wave becomes highly nonlinear, developing localized overturns and high shears. The wave breaks through a three-dimensional instability and energy is driven to small scales, where it is rapidly lost to dissipation and diffusion. The instability mechanism is discussed in Winters and Riley (1991). Our interest here is in the diffusive mixing associated with the event.

We will first attempt to diagnose the mixing indirectly, by appealing to the density (or temperature) variance equation. This approach is shown to be problematic, as the production of density variance can be accomplished both by adiabatic (nonmixing) as well as diabatic (mixing) effects. We then illustrate a direct method in which the two effects are isolated. The method is conceptually simple, based on the idea that mixing implies changing the volume of fluid between given isopycnals. This idea leads naturally to the concepts of available and background potential energy which are also useful for diagnosing diffusive mixing.

FLOW SIMULATION

A mixing event is simulated by numerically solving an initial value problem in which a wave packet propagates toward a critical level, where it breaks down into much smaller scale motions. The equations of motion for the (dimensionless) velocity vector $\vec{u}(\vec{x},t) = (u,v,w)$ and the perturbation density ρ, with respect to an ambient linear profile $\Theta(z)$, are listed below.

$$\frac{\partial \vec{u}}{\partial t} + \vec{u} \cdot \nabla \vec{u} = -\nabla p - Ri\,\rho \hat{z} + Re_*^{-1} \nabla^6 \vec{u} \tag{1a}$$

$$\frac{\partial \rho}{\partial t} + \vec{u} \cdot \nabla \rho - w = Pr_*^{-1} Re_*^{-1} \nabla^6 \rho \tag{1b}$$

$$\nabla \cdot \vec{u} = 0 \tag{1c}$$

The parameter Ri is the bulk Richardson number defined as $Ri = [NL/U]^2$ with constant $N^2 = -g/\rho_0\, d\Theta/dz$. The equations are solved in the unit cube $0 \le x,y,z < 1$ with periodic boundary conditions in all space dimensions. The grid is uniform, with 32 points in each horizontal direction and 200 points in the vertical. A pseudo-spectral numerical algorithm, with second order Adams Bashforth time stepping is used to evolve the flow field in time.

Sub-grid scale model:

A sub-grid scale model is incorporated through the inclusion of "hyper" viscous and diffusive operators. Mathematically, the physical Laplacian operators and their coeffecients have been replaced with nonphysical ∇^6 operators and new coefficients. The magnitude of the coefficients

depends on the grid resolution. For high resolution grids, these parameters can be extremely small since the sixth order derivative introduces a factor of wavenumber to the sixth power. By choosing these coefficients to be small, $O(10^{-17})$, the viscous/diffusive effects can be confined to the narrow region of wavenumber space near the resolution limit. This leaves the remainder of wavenumber space to be treated essentially inviscidly and nondiffusively. To contrast, a second approach would be to leave the physical terms in the equations and choose an unphysically small Reynolds number. The form of sub-grid scale model chosen preserves the diffusive character of the small scale processes and is convenient to implement in a spectral scheme. It does, however, have important physical implications and needs some justification for the present study. We note that *any* diffusive mixing that occurs in the simulation will be done by "hyperdiffusion" and hence must be regarded as occurring at subgrid scale. In employing this form of sub-grid model, two important assumptions are made. First, we assume that the overall rate of dissipation and diffusion is controlled fundamentally by the rate of downscale energy transfer and not by the specific form of the viscous/diffusive operators. Second, we assume that the transfer of energy back upscale from the sub-grid scales is unimportant.

Initial conditions:

An ambient horizontal shear flow $U(z)$ and a downward propagating internal wave packet are specified as initial conditions. The wave packet is specified to be two-dimensional, with variations in the x and z directions only, with nondimensional wavelengths of 1 and 1/8 respectively. The packet is localized in the vertical by a slowly varying Gaussian envelope. The form of the ambient flow U is chosen so that a critical level is present at $z_c = 0.54$. A broad band spectrum of small amplitude, three-dimensional "noise" is also initialized, allowing the flow to evolve into three dimensions.

Flow evolution:

The equations are integrated forward in time for 50 buoyancy periods. The wave packet propagates downward from a region of no shear into the ambient flow. Refraction of the wave by the mean shear impedes its progress and reduces the intrinsic scale of the wave motion. The wave steepens near the critical level and creates regions of overturned isopycnals and strong shear before "breaking" into small scales which then dissipate and diffuse.

Figure 1 shows profiles of displacement taken at a fixed horizontal location at several points in time. The initial Eulerian phase speed of the wave packet is zero. The intrinsic right going phase propagation is balanced near the top of the figure by a left flowing mean current. The current speed is zero at z_c, where the wave becomes critically refracted. The prescribed wave envelope is nearly zero in the region shown in the figure, thus the displacements are initially small. Later, the wave has propagated into the region from above, resulting in finite displacements. Figure 2 shows contours of the density field sampled at the same horizontal location. Note the overturns in the isopycnals near the critical level at about t=20. Clearly, the most interesting flow dynamics occur in this "wave breaking" region. The remainder of the paper will concentrate on the diagnosis of the diffusive mixing associated with the flow in this region.

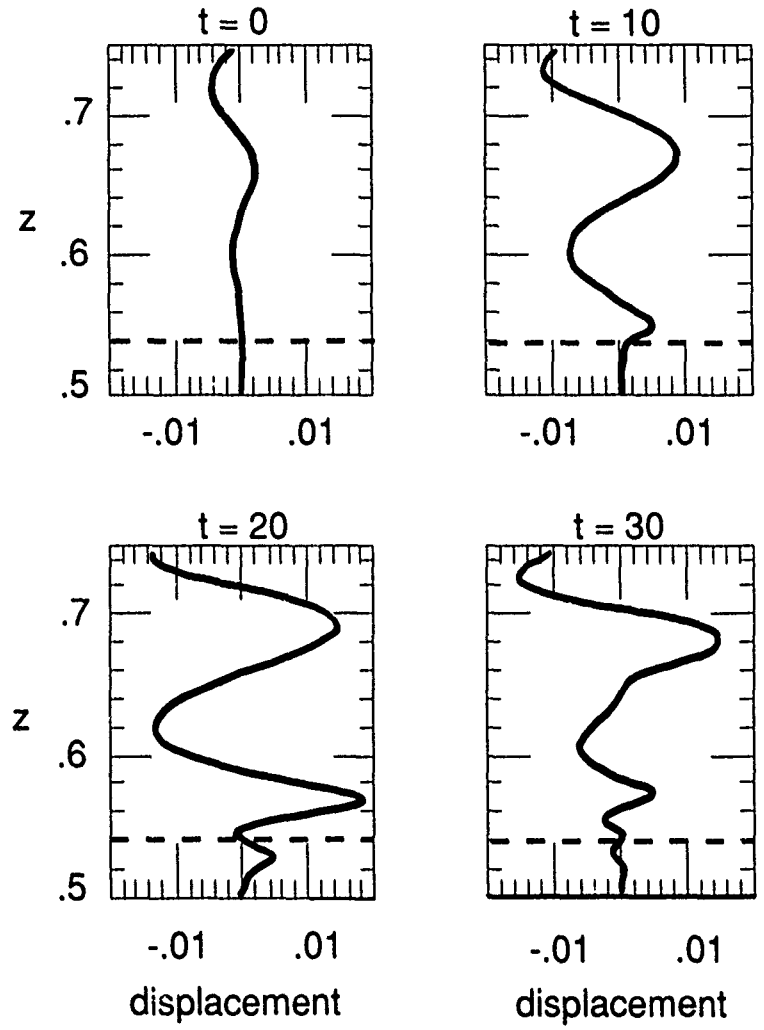


Figure 1. Vertical profiles of (dimensionless) isopycnal displacement are shown at several times t (in buoyancy periods). The wave packet propagates downward toward a critical level, depicted by the broken line.

DENSITY VARIANCE AS AN INDICATOR OF MIXING

We begin our discussion of mixing by looking at the dynamical balance of the density variance equation. The equation for the density perturbation from an ambient linear gradient is

$$\frac{\partial \rho}{\partial t} + \vec{u} \cdot \nabla \rho - w = D(\rho) \tag{2}$$

where $D(\rho)$ is simply the "hyperdiffusion" term appearing in Eq. (1b). The perturbation density ρ can be further decomposed into horizontal mean and fluctuating components; $\rho = \bar{\rho} + \rho'$. Multiplying Eq. (2) by the fluctuating component ρ' and horizontally averaging, denoted by an overbar,

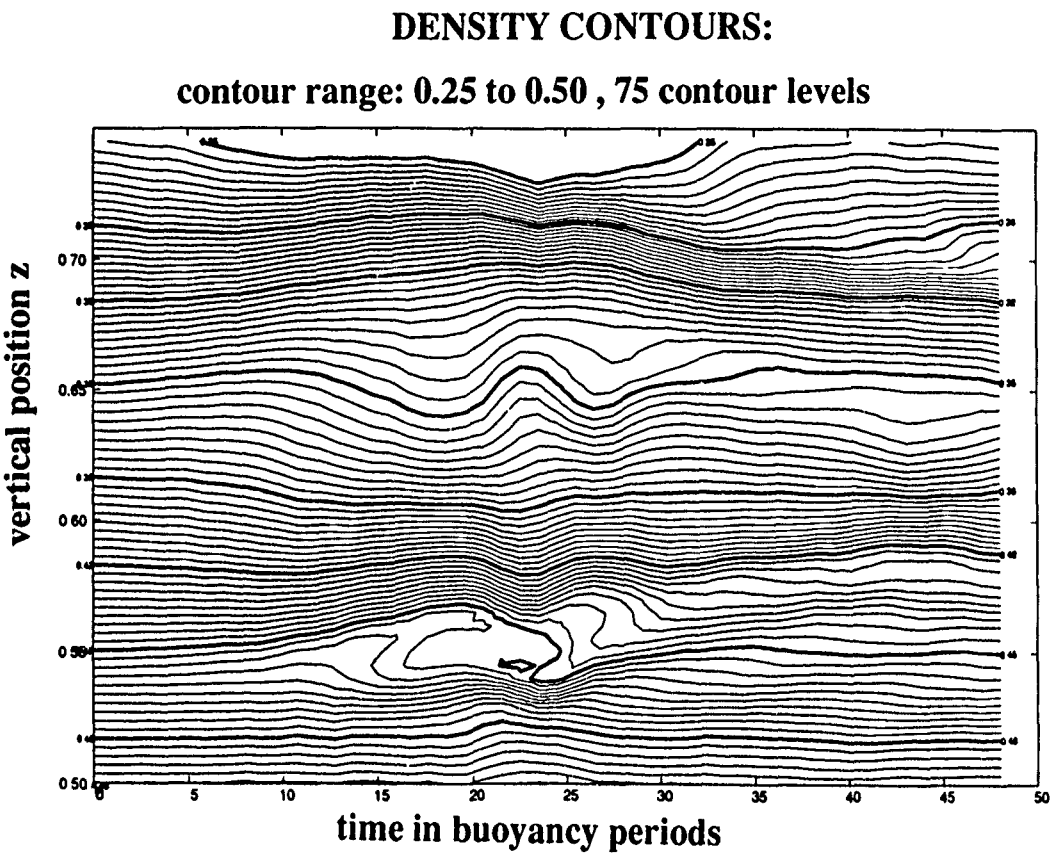


Figure 2. The density profiles obtained at a fixed horizontal location are contoured in the depth-time plane. Overturned isopycnals are apparent near $t = 20$ when the wave breaks.

gives an equation governing the production of density variance.

$$\frac{\partial}{\partial t}\overline{\rho'^2} = -\frac{\partial}{\partial z}\overline{\rho'^2 w} - 2\overline{\rho w}\frac{d\rho_{tot}}{dz} + \overline{\rho' D(\rho')} \qquad (3)$$

The horizontally periodic boundary conditions have been taken into account in the derivation of Eq. (3). The quantity ρ_{tot} is the horizontally averaged total density. The terms on the right hand side of this equation can be thought of as forcing terms, producing density variance. Eq. (3) states that density variance can be generated by advection across the vertical boundaries, buoyancy flux, or diffusion. Similar equations can be derived for temperature variance or "turbulent" kinetic energy. Indirect microstructure measurement techniques are based on approximate balances within these equations, which are assumed to be valid in the ocean interior when "ensemble averaged" over many profiles. Here we focus on the behavior of Eq. (3), integrated over the depth interval of interest, enclosing the critical level and the wave breaking region.

Figure 3 shows the volume integrated ρ'^2 as a function of time. Note that its production is strongly time dependent; there is no interval of time when Eq. (3) can be reasonably approxi-

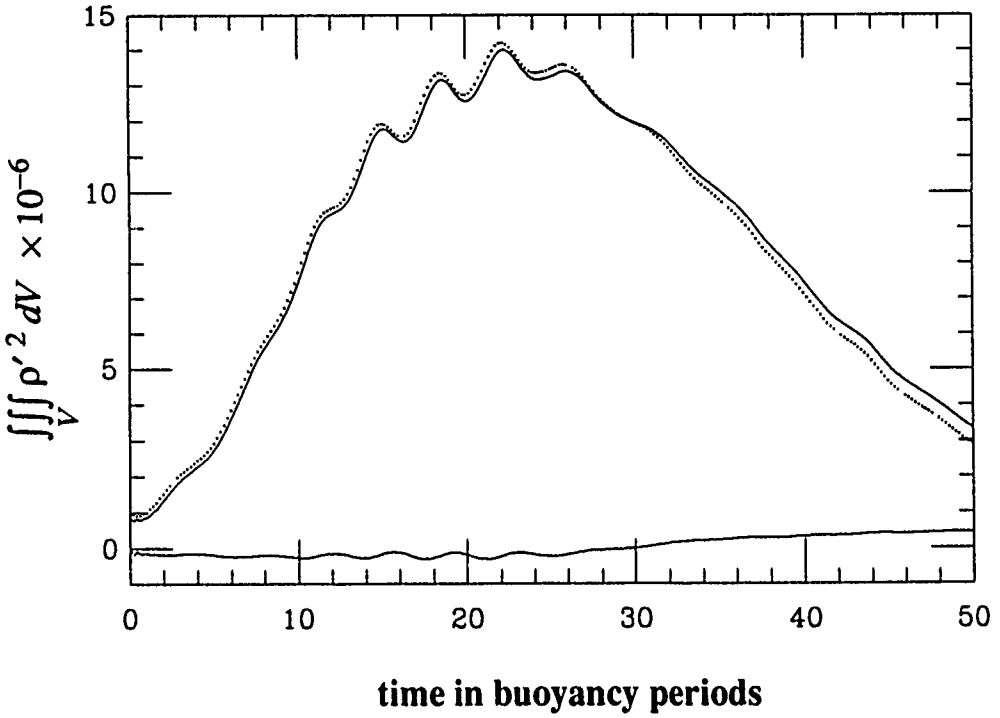


Figure 3. The upper solid curve is the volume integral of ρ'^2. The dotted curve was obtained by integrating the production terms on the right-hand side of Eq. (3) in time. The small residual between the two calculations is also shown.

mated as steady state. Note also that changes in density variance are not irreversible, density variance increases rapidly and then decreases just as efficiently, with a relatively small residual left behind when the calculation was terminated. Normally, when we think of mixing by an isolated turbulent event, we think of a nearly irreversible process, with a localized mixed layer perhaps eventually diffusing away, but only on very long time scales after the turbulence dies out. It seems reasonable to conclude that the production of density variance is influenced, or even dominated by, nonmixing processes. It is not clear how to use Eq. (3) to diagnose the mixing associated with this event.

The role of the buoyancy flux

Further insight can be obtained regarding the temporal behavior of Eq. (3) by looking at the role played by the buoyancy flux ρw. Figure 4a shows the vertically integrated buoyancy flux as a function of time. The buoyancy flux shows strongly oscillatory behavior on a short time scale approximately corresponding to the initial wave period. The magnitude of these oscillations can be quite large, even causing the signal to oscillate sign. It appears obvious that there is a strong wave component to the buoyancy flux. This is not too surprising, it would be reasonable to look at this signal temporally averaged to remove the intrinsic wave oscillations. The broken curve is a crude approximation of the time averaged signal. The averaged buoyancy flux is positive until about $t = 20$, when the wave breaks. It then changes sign and remains negative throughout the rest of the calculation.

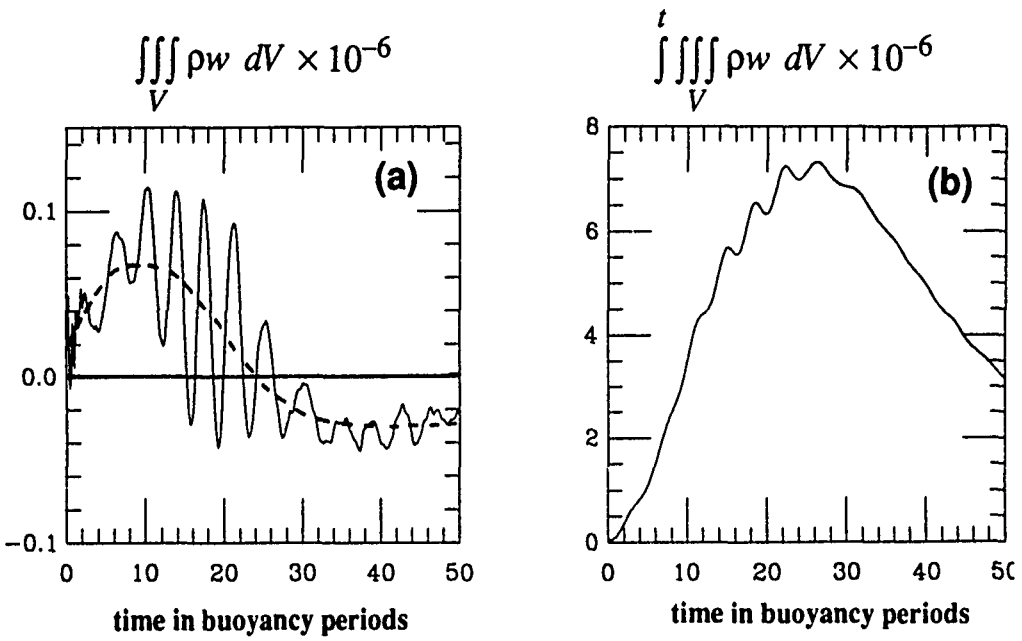


Figure 4. The volume integrated buoyancy flux is given by the solid curve in (a). The broken curve is an estimate of the time averaged signal removing the wave oscillations. The integral in time of the solid curve is given in (b).

Even the time averaged buoyancy flux is dominated by wave effects. As the wave approaches the critical level, the initially undisplaced density field is disturbed. Fluid parcels are displaced away from their equilibrium locations, resulting in a positive buoyancy flux. This process continues until about t=20, the time of wave "breaking". Two processes occur after the wave breaks. On the average, the displaced fluid parcels return towards their equilibrium positions, and some diffusion occurs at small scales as they do so. Parcels returning towards equilibrium give rise to a negative or counter-gradient buoyancy flux while diffusion results in mixing.

To illustrate the dominance of the wave effects on the density variance dynamics, we have integrated the net buoyancy flux of Figure 4a with respect to time and shown the result in Figure 4b. The similarity between Figures 3 and 4b implies that wave dynamics, through the buoyancy flux term, plays the dominant role in the density variance equation. For this event at least, using the density variance equation is not a particularly clean way to diagnose the mixing of the event. In essence, we conclude that spatial averaging does not adequately separate mixing from nonmixing processes. Loosely, we have not been able to separate "waves" from "turbulence".

DIRECT VIEW OF MIXING

We will now attempt to look at the mixing of the wave breaking event in a more direct manner. We will exploit a simple and intuitive concept, namely that diapycnal mixing results in changes in the volume of fluid found between a given pair of isopycnal surfaces. We can think of a stratified fluid at some time t_0 as a continuous distribution of isopycnal surfaces, along with an

associated velocity field. By discretizing the density range, we can think in terms of a finite number of these surfaces, each pair enclosing a finite volume of fluid. Suppose the fluid now evolves, undergoing both diabatic and adiabatic processes. At some later time t_1, the topology of these surfaces may be extremely complicated. To diagnose the amount of mixing, however, we are interested only in a single variable, the total volume of fluid between each pair. Figure 5a shows a schematic of the density field of a fluid that evolves from a state of uniform density gradient to a nonuniform, mixed state. Initially there is an equal volume in each of the density "bins". Later the volume in the middle bin has increased at the expense of the neighboring bins.

Obviously, the volume between the isopycnal surfaces has a functional dependence on the density field $\rho(\vec{x},t)$. This dependence, however, is very different than the usual horizontal or temporal averages we normally employ. Thinking of the total volume of fluid as a collection of fluid par-

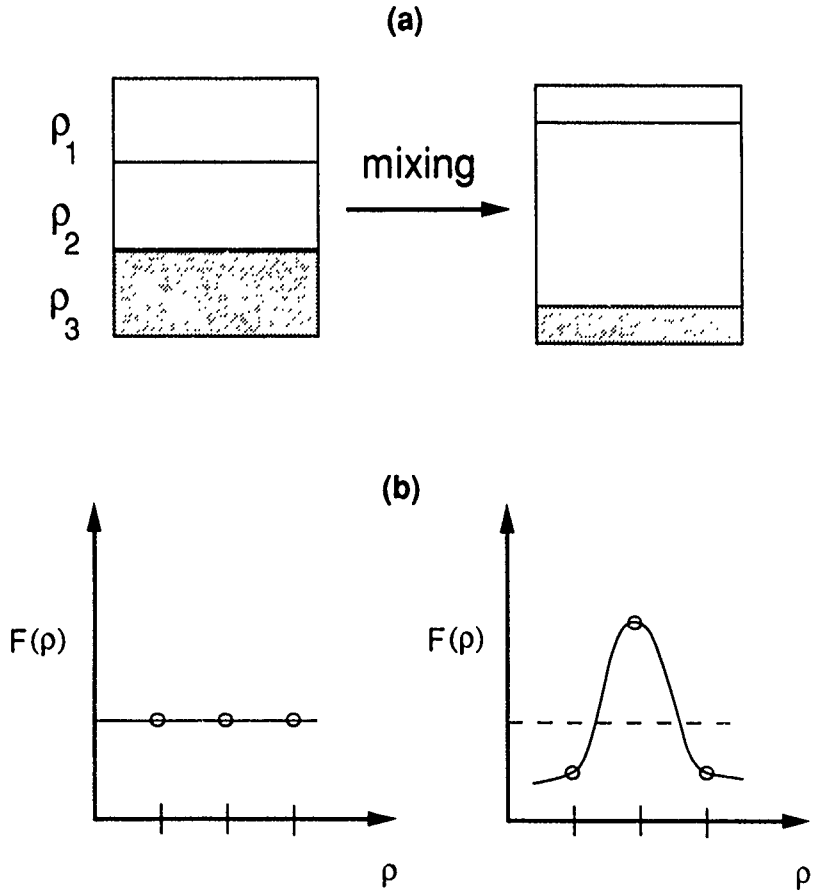


Figure 5. A schematic of a density stratified fluid is given in (a). The density range has been discretized into three "bins". Initially, the volume of fluid within each density bin is equal. After a mixing event, the volume of fluid in the central bin has increased while the volumes of the neighboring bins have decreased. Regarding the fluid as an ensemble of parcels, (b) shows the probability F of a randomly selected parcel having a density value within a given bin. Initially, the probabilities are equal. Later, the probability of finding a value within the central bin is greater. The function F depends on density but not on position.

cels, each with a density and velocity value, we can denote the probability of a randomly selected parcel having a density value of ρ as $F(\rho)$. Figure 5b shows the probability distribution F of the hypothetical density field of Figure 5a. Initially, there is an equal probability of occurrence for all values while later it is much more probable that a given parcel will have a value near ρ_2. The main difference between this type of functional dependence on density and, say, the horizontal average, is the lack of dependence on the vertical position of each parcel.

This simple idea can be incorporated into a straightforward algorithm for diagnosing mixing; at least in numerical simulations. For a given discretization of the density range, the volume of fluid can be computed directly for each pair of neighboring isopycnals. Each differential volume can be "spread out" over a constant area A, yielding a differential thickness for each "bin". The bins are then "stacked", with the densest fluid on the bottom at some fixed reference height, in order of decreasing density. By keeping track of the thickness of each bin, the locations of the edges of each bin are also determined. One can think of this algorithm as producing the "pseudo-positions" z_* of the given set of isopycnal values, *i.e.* a one-dimensional "sorted" density field which happens to be specified on an unevenly spaced grid.

Figure 6 shows contours of density in the (t, z_*) plane. In the absence of mixing, these contours would remain flat with a uniform vertical spacing. Deviations from flat isopycnals imply diapycnal mixing. The net spreading of isopycnals near $t = 20$ and $z_* = 0.54$ indicates the formation of a mixed layer. This mixed layer is bordered, both above and below, by regions of net isopycnal convergence where the background density gradient is enhanced. The lowest several isopycnals remain approximately flat, implying a lower level across which there is no mixing.

AVAILABLE AND BACKGROUND POTENTIAL ENERGY

These same ideas can be formulated in terms of energetics using the concepts of available and background potential energy. The concept of available potential energy has been discussed and applied by many authors including Lombard (1989), Dillon (1984) and Holliday and McIntyre (1981). The object is to define a one-dimensional "background" state for density that is insensitive to adiabatic dynamics and compute the potential energy associated with it. Changes in the background potential energy can then only occur through diffusive mixing. At some time t, the potential energy of a flow, in dimensionless form, is given by

$$E_p(t) = Ri \iiint_V \rho(\vec{x},t)(z - z_0)\, dV \tag{4a}$$

where z_0 is an arbitrary reference location.

Let z_* be the "pseudo-position" variable, defined by computing differential volumes, "sorting" and "stacking", in the limit as the isopycnal spacing goes to zero. The background potential energy E_b is simply the potential energy of the flow in pseudo-position space.

$$E_b(t) = Ri \iiint_V \rho(\vec{x},t)(z_* - z_0)\, dV \tag{4b}$$

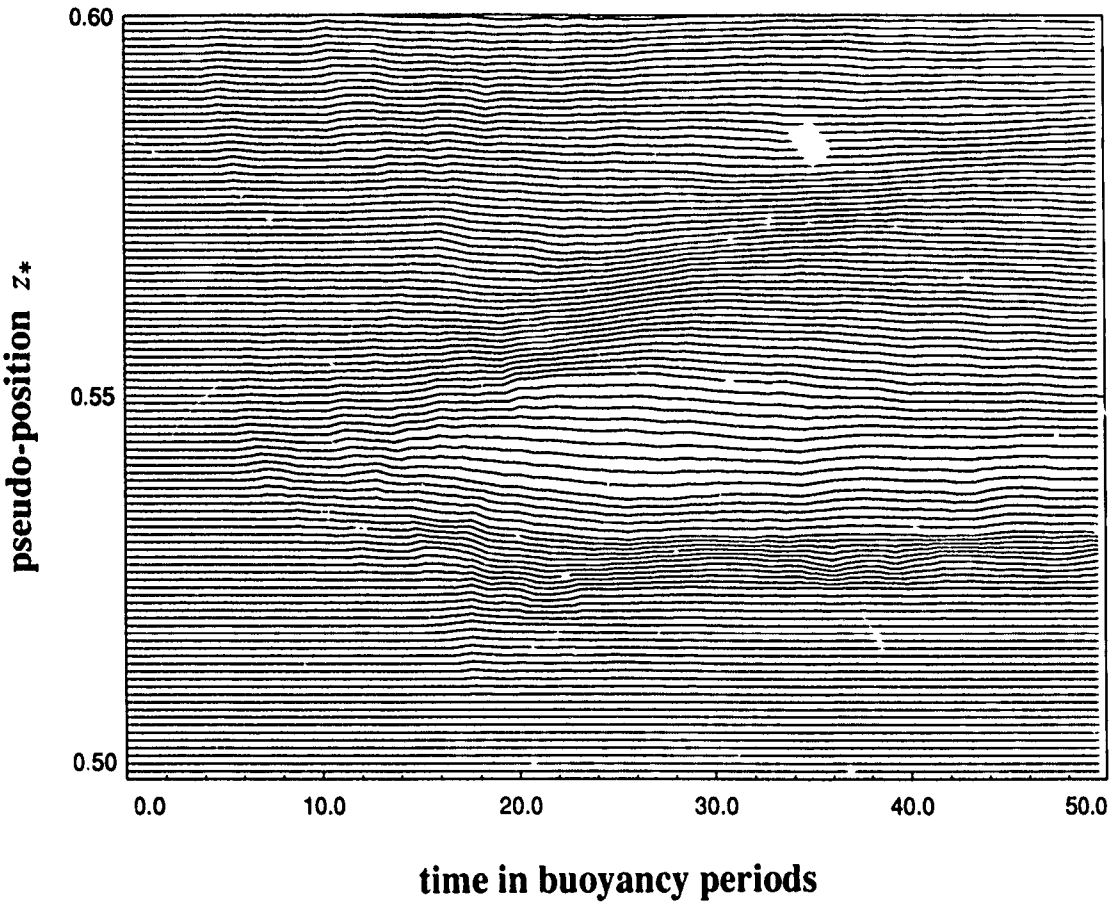


Figure 6. Contours of density in the (t, z_*) plane. Deviation from flat, regularly spaced contours indicates diapycnal mixing of the fluid. The main feature is the region of strong isopycnal spreading corresponding to the time and (true) position of the wave packet breaking near the critical level. Regions of net isopycnal convergence are also evident.

It is clear from this definition that the background potential energy is independent of true position z , depending only upon the ordered position z_* . For a given density value, z_* remains constant in time in the absence of mixing and can only change if mixing occurs.

The available potential energy E_a is defined as the total energy released by the flow in attaining the sorted state adiabatically.

$$E_a(t) = Ri \iiint_V \rho(\vec{x},t)(z - z_*)\, dV \qquad (4c)$$

Eqs. (4) define a unique decomposition of potential energy into available and background parts, *i.e.* $E_p = E_a + E_b$.

We denote the cumulative change in time of the background potential energy by $E_{mix} = E_b(t) - E_b(0)$. E_{mix} would seem to be a good indicator of diffusive mixing in a stratified fluid.

There are some subtleties, however. A volume must be specified over which the analysis will be applied. For the present discussion, we will work in an Eulerian reference frame and specify a fixed volume in space. It is instructive to take the time derivative of Eq. (4) to see how the potential energy in such a volume can change.

$$\frac{d}{dt} E_p = (z - z_0)\,\overline{\rho w}\,|_{z_2}^{z_1} + Ri \iiint_V \rho w \, dV + \iiint_V (z - z_0)\, D(\rho)\, dV \qquad (5)$$

We immediately encounter a difficulty. The first term on the right hand side of Eq. (5) indicates that the total potential energy changes as a result of net mass flux across the upper and lower boundaries. Although the volume remains constant in an Eulerian framework, the mass is not necessarily conserved within the volume. Care must be taken in interpreting a potential energy budget of an open system in which the mass is changing.

Note that this difficulty arises in this problem because we choose to apply our energetics analysis to only a small domain in which the wave breaking occurs. Had we used the entire computational domain, we would have had boundary conditions on the top and bottom to impose. For some boundary conditions, *i.e.* no slip wall conditions, the vertical velocity is zero, the mass flux terms are zero, and there is no problem. Spectral methods, however, are commonly used for simulating stratified turbulence and can create difficulties in analyzing potential energy even when the entire domain is used. Often, periodic boundary conditions are specified for velocity and *perturbation* density (from, say, a linear ambient profile) in all space directions. Such conditions are often interpreted loosely as "waves that propagate out through one boundary propagate back in through the paired boundary". Note, however, that this implies that parcels from near the bottom of the domain can be interchanged with parcels from near the top by wavelike oscillations. In other words, the total mass is not conserved and care must be used in quantifying the potential energy of the flow.

In general, the analysis could be performed in isopycnal coordinates. For a volume bounded by two isopycnals, across which no mixing occurs, both the volume and the mass remain constant. If two such bounding isopycnals can be found, the analysis can, in principle, be performed cleanly.

For our particular problem, however, we can compensate for the nonconservation of mass in one of two ways. The cumulative changes in potential energy due to net mass flux can be computed explicitly and the potential energy budget "corrected" for this effect. A second approach takes into account the fact that the mass flux occurs primarily at the top boundary. The lower boundary is located below the critical level, which prevents most of the wave energy from reaching it. The difficulty, then, is the net interchange of parcels near the top boundary but inside V, with parcels of slightly different density outside V. From Eq. (5) we see that the problem can be minimized by selecting the reference location z_0 to coincide with the top boundary z_2. Exchanging parcels with different densities at exactly this level then results in no change in potential energy

while exchanges of parcels near this value have only small effects. Either of these corrections are adequate for our purposes here. We have chosen the second technique for the energetics balance below.

Figure 7 shows the available and background potential energies as a function of time. The available potential energy behaves much like the net integrated buoyancy flux of Figure 4b. E_a increases prior to wave breaking as parcels are displaced on the average, and decreases after the wave breaks, as parcels return towards their equilibrium values. The background potential energy, however, behaves quite differently. Prior to wave breaking, it remains approximately constant as the flow contains no appreciable energy at the diffusive scales. After wave breaking, the fluid mixes as small scale features diffuse while returning towards equilibrium.

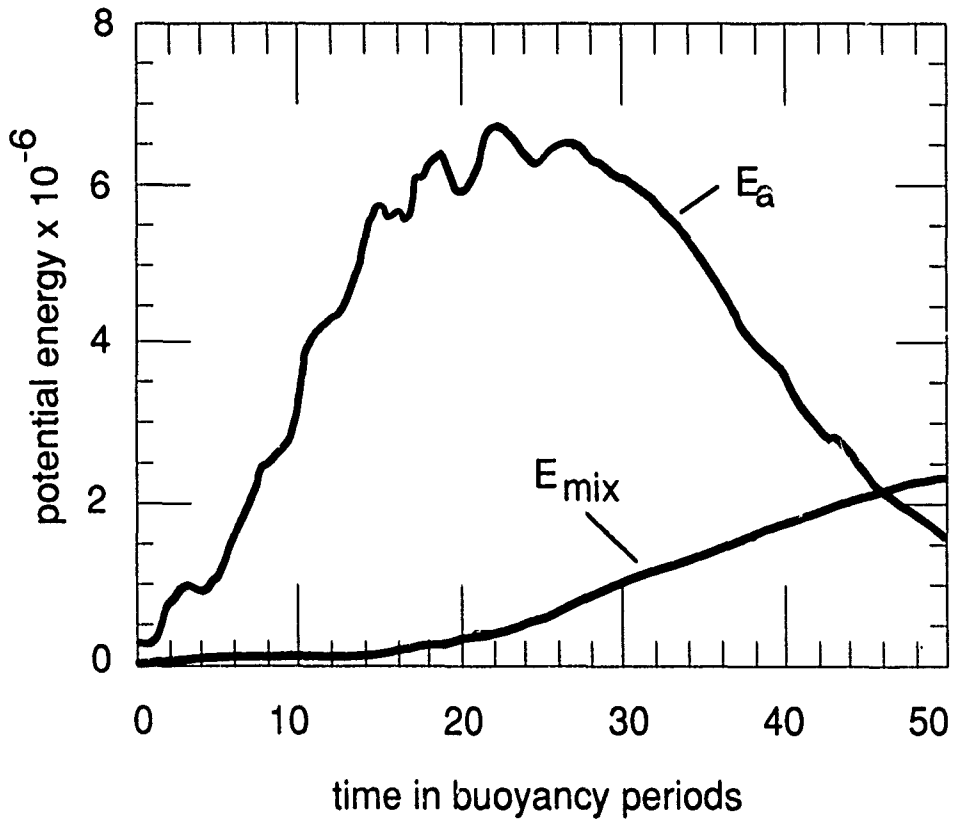


Figure 7. The available potential energy E_a and the cumulative change to the background potential energy E_{mix} are shown as functions of time. E_a increases initially as the packet propagates into the volume and displaces the stratification and decreases after wave breaking as parcels return towards their equilibrium positions. The increase in E_{mix} at later times indicates that diffusive mixing takes place as the fluid restratifies.

VERTICAL DIFFUSIVITY K_d

The function $\rho(t,z_*)$ is an interesting and useful quantity. We can think of it as "mean" density field where the averaging operator filters out small-scale details of the flow, like the displacements of fluid parcels due to internal wave motion, but retains the overall effects of small-scale mixing. It is this sort of averaging that is implicit in the formulation of meso to large-scale circulation models. In such models, internal waves and turbulence occur at scales much too small to resolve, but the overall mixing accomplished by these motions must be incorporated in some manner.

Earlier, we interpreted the spreading and convergence of the isopycnals in Figure 6 as a qualitative indicator of diffusive mixing. We can also use $\rho(t, z_*)$ in a more quantitative manner by computing the rate of spreading. Suppose we introduce a model evolution equation for the "mean" density field.

$$\frac{\partial}{\partial t}\rho(t,z_*) = \frac{\partial}{\partial z_*}Q(t,z_*) = \frac{\partial}{\partial z_*}K_d\frac{\partial \rho}{\partial z_*} \tag{6}$$

Although the mean state changes as a result of complicated three-dimensional motions at small scales, Eq. (6) collapses the net effect onto a single vertical diffusivity function K_d . K_d can be computed directly from the mean density field and the overall character of the mixing event can be examined with respect to pseudo-position z_* and time.

$$K_d(t,z_*) = \frac{\int^{z_*}\frac{\partial \rho}{\partial t}(t, z_*')\,dz_*'}{\frac{\partial \rho}{\partial z_*}} \tag{7}$$

Figure 8 shows contours of the (dimensionless) diffusivity function K_d. The main feature in the figure is located near $z_* = 0.54$ and $t = 20$, which corresponds to the time and place that the strongest wave breaking occurred. The diffusivity is positive in this region, corresponding to local spreading of isopycnals. This feature is responsible for most of the mixing of the event. It is relatively brief in duration, lasting only about 5 buoyancy periods. After the main mixing event, the diffusivity field becomes much more complicated, with both positive and negative features appearing. The negative values correspond with the regions of converging isopycnals in Figure 6. Figure 9 shows the result of time averaging the diffusivity field.

MIXING EFFICIENCY

We define the mixing efficiency of the event γ_{mix} as follows.

$$\gamma_{mix} = \frac{\text{rate of increase of background potential energy}}{\text{rate of dissipation of kinetic energy}} \tag{8}$$

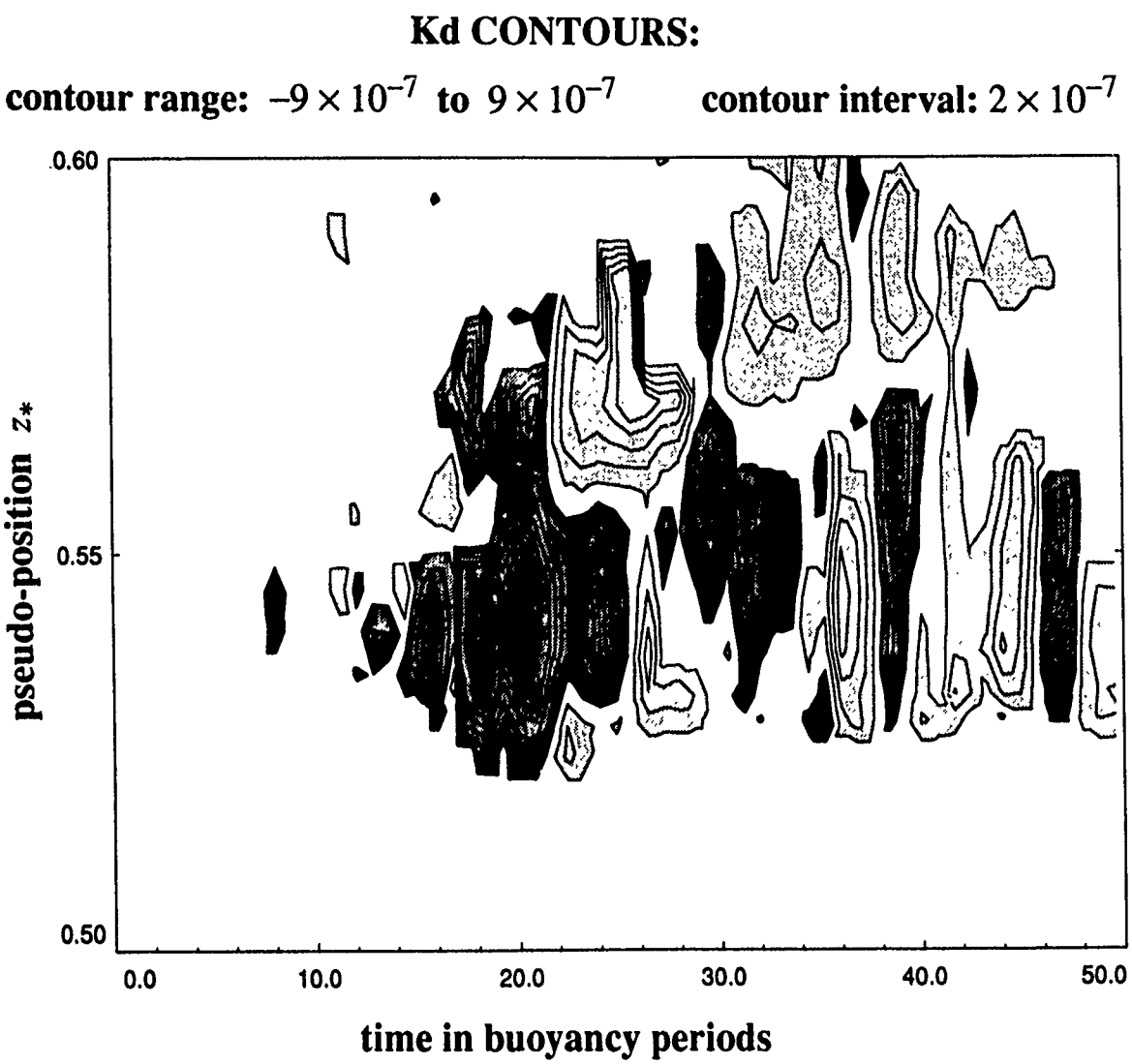


Figure 8. Contours of the vertical diffusivity of mass K_d defined in Eq. (7). A strong positive feature is seen to correspond with the initial region of wave breaking. At later times, the field becomes more complicated with both positive and negative diffusivities appearing.

The rate of increase of background potential energy can be obtained by estimating the nearly constant slope from Figure 7. The dissipation rate of kinetic energy, ε is shown in Figure 10 along with the rate of dissipation of density variance χ. Taking a representative value of ε over the time period between $t = 20$ and $t = 50$, allows the ratio in Eq. (8) to be computed, yielding

$$\gamma_{mix} \approx 0.38 . \qquad (9)$$

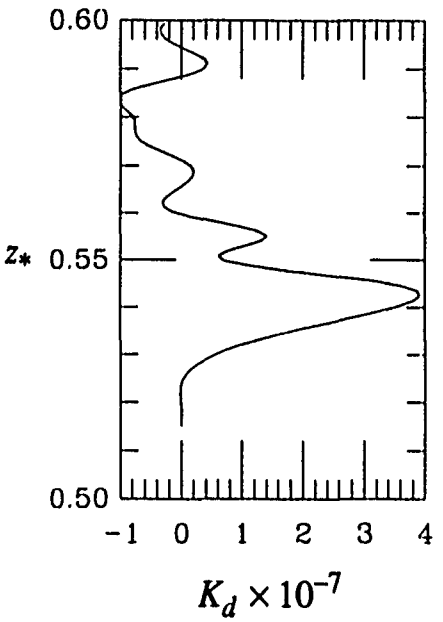


Figure 9. The time averaged profile of the diffusivity K_d.

Perhaps not surprisingly, approximately the same value is obtained if the mixing efficiency is defined as the ratio of representative values of χ to ε. We are encouraged that direct simulations can produce mixing efficiencies near the expected value, but more study is required to determine the sensitivity of this result to changes in the sub-grid scale model.

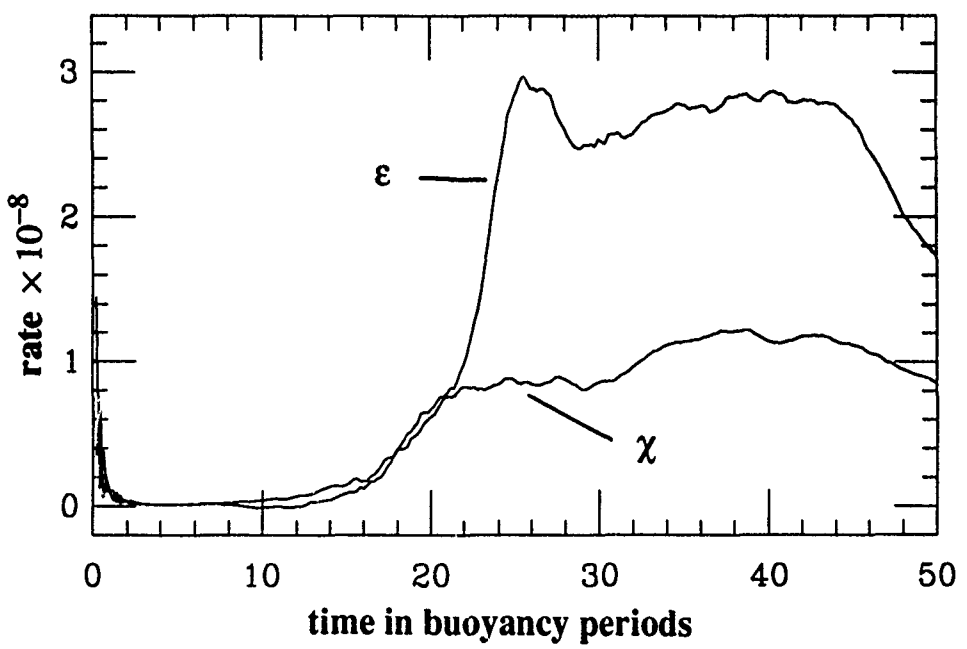


Figure 10. The dissipation rates of kinetic energy ε and density variance χ.

SUMMARY

We have examined the diapycnal mixing associated with a numerically simulated mixing event similar to what may be seen in the ocean interior. A high resolution, three-dimensional primitive equation model was used to simulate a wave packet breaking at a critical level. The temporal behavior of the density variance equation was discussed in relation to the diagnosis of the mixing of the event. The wave component of the buoyancy flux was shown to dominate the dynamics, obscuring the diagnosis. A direct approach, based on the computation of fluid volume between discrete isopycnal surfaces, was shown to be a less ambiguous indicator of the mixing dynamics. The results of this calculation were used to estimate the vertical diffusivity of mass K_d. Coupling this analysis with an energetics budget based available and background components of potential energy, a reasonably complete picture of the mixing dynamics emerges.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation, grant OCE 8700108. Funds to attend the Hawaiian Winter Workshop were provided by the Applied Physics Laboratory at the University of Washington.

REFERENCES:

Dillon, T.M., 1984: The energetics of overturning structures: implications for the theory of fossil turbulence. *J. Phys. Ocean. 14*, 541-549.

Gregg, M. C., 1987: Turbulence in stratified fluids: A review. *J. Geophys. Res.* **92,** 5249-5286.

Holliday, D. and M. E. McIntyre, 1981: On potential energy density in an incompressible stratified fluid. *J. Fluid Mech.* **107,** 221-225.

Lombard, P. N., 1989: Energetics of a stably stratified turbulent flow. *Masters Thesis, University of Washington* 157 pp.

Moum, J. N., 1989: Measuring turbulent fluxes in the ocean - The quest for K_ρ. *Proceedings of the 1989 Hawaiian Winter Workshop* 145-156.

Winters, K. B. and E. A. D'Asaro, 1989: Two-dimensional instability of finite amplitude internal gravity wave packets near a critical level, *J. of Geophys. Res.*, **94:C9,** 12,709-12,719.

Winters, K. B. and J. J. Riley, 1991: Instability of internal waves at critical levels, *Dynamics of Atmospheres and Oceans* (in press).

ENERGY TRANSFERS ACROSS AN INTERNAL WAVE /VORTICAL MODE SPECTRUM

Dave Ramsden, Daleth Research, 1255 Oscar St. Victoria BC, V8V2X6 Canada

Greg Holloway, Institute of Ocean Sciences, Sidney, BC, V8L4B2 Canada

ABSTRACT

Nonlinear interactions among internal gravity waves are investigated by direct numerical experiments. In the first part of this paper, cases which solve the full (3D) Navier-Stokes equations are compared to cases in which variability is suppressed in one horizontal direction. It is found that the 3D and 2D simulations exhibit certain similarities. We find both in 3D and in 2D that the transfer of kinetic energy (KE) from large to small scales is less efficient than the transfer of potential energy (PE). The imbalance between these transfers leads to a characteristic buoyancy flux spectrum which is negative (KE to PE) at large scales and positive (PE to KE) at small scales. Integrated over all scales, the buoyancy flux is very small for a wide range of flow regimes. However, results concerning buoyancy flux are sensitive to assumptions about the manner of energetic forcing.

These sensitivities are taken up in the second part of the paper where energy is introduced in a 'surface' layer and removed via a bottom absorption layer. It is found that it is kinetic energy which is radiated down the fluid column and the mechanism for the radiation is the divergence of the pressure-velocity correlation. For these cases, buoyancy flux in the radiated region is on average negative. The buoyancy flux spectra, however, retain the same tendency to become positive at the smallest scales.

Energy balances in the radiation region for both kinetic and potential energy are shown to be qualitatively achieved between transport, buoyancy flux and dissipation terms without the need to posit a 'mean shear' extraction term. Buoyancy flux is shown to be of the same order as dissipation rate of kinetic or potential energy. Kinetic and potential energy dissipation rates are about equal.

INTRODUCTION

The objectives of this study are to attempt to determine the role of nonlinear transfers in the evolution of internal waves near dissipation scales. The first part of this work follows a paper recently submitted; *Ramsden and Holloway* [1990] (hereafter RH) on statistically homogeneous forced cases, so we will be mainly emphasizing the highlights of those investigations. The second part of this study is on surface layer forcing. This work is new and still under active investigation. In the spirit of 'Aha, we present the preliminary results.

Part 1: STATISTICALLY HOMOGENEOUS FORCED CASES

The starting point for RH was to attempt to isolate the effects of nonlinearities in the evolution of internal waves. In addition, they also wanted to consider the role of the zero frequency, potential vorticity containing component of the motion field [*Holloway*, 1983], the so called vortical mode [*Muller et al., 1988*]. Interaction of the vortical mode with the internal wave field is not well understood. In particular, the role of nonlinear interactions in rearranging wave and vortical energies at order 1–50 m scales is unknown. RH were also interested in determining the effects of nonlinear interactions on other flow parameters such as buoyancy flux and what effect restriction of the 3D motion field to vertical planar 2D would have.

There are two reasons to compare 2D results to 3D. The first is that it may be useful to run 2D simulations because the range of accessible scales is greater for given computer capacity. If the results are qualitatively similar to 3D, a great deal of time and effort can be saved by performing 2D simulations. The second reason is to ask what part(s) of the results from 2D theoretical treatments of internal waves may be applicable to the full 3D problem.

RH summarizes and follows up on numerical experiments reported in *Shen and Holloway* [1986] (non-rotating statistically stationary 2D internal waves), *Ramsden and Holloway* [1987] (2D with rotation), and *Holloway and Ramsden* [1988] (initial presentation of 3D results).

METHOD

The governing equations are the 3D nondimensionalized Navier-Stokes equations under the Bousinesq approximation;

$$\partial_t \mathbf{u} + \mathbf{u}\cdot\nabla\mathbf{u} = -\nabla p - \rho \mathbf{e}_z + \nu\nabla^2\mathbf{u} - 2\boldsymbol{\Omega}\times\mathbf{u} + \mathbf{F}_\mathbf{u} \tag{1}$$

$$\partial_t \rho - W + \mathbf{u}\cdot\nabla\rho = \kappa\nabla^2\rho + F_\rho \tag{2}$$

$$\nabla\cdot\mathbf{u} = 0 \tag{3}$$

In (1) to (3), $\mathbf{u}$ is velocity (U,V,W) on basis vectors $(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z)$, ∇ is the gradient operator, p is pressure, ρ is the departure from a mean background gradient of density, $\bar{\rho}$, ν is kinematic viscosity, $\boldsymbol{\Omega}$ is the earth's angular velocity, set to (0,0,f/2), $\mathbf{F}_\mathbf{u}$ and F_ρ are external forcing, and κ is mass diffusivity. Time has been scaled such that the bv frequency $N = \left(\frac{g}{\rho_0}\frac{\partial\bar{\rho}}{\partial z}\right)^{1/2}$ (g gravity, ρ_0 reference density) is unity. Vorticity is defined to be $\Lambda=\nabla\times\mathbf{u}$. (1) to (3) are the equations employed by *Riley et al* [1981] in the pioneering study of 3D stratified turbulence, with the addition of rotation and forcing terms. Full details of the nondimensionalization may be found in *Riley et al.*

Spatial derivatives are calculated by the spectral transform method after *Orszag* [1971] in which each variable $\mathbf{u}, \rho$, etc. is expressed as a truncated Fourier series, a procedure which implies

periodic boundary conditions. On wavevectors $\mathbf{k} = (k_x, k_y, k_z)$ denote the Fourier transform of any field $a(\mathbf{x})$ to be $a_\mathbf{k}$:

$$\mathbf{u}(\mathbf{x}) = \sum_{\mathbf{k}} \mathbf{u}_\mathbf{k} e^{\iota \mathbf{k}\cdot\mathbf{x}} \tag{4}$$

$$\rho(\mathbf{x}) = \sum_{\mathbf{k}} \rho_\mathbf{k} e^{\iota \mathbf{k}\cdot\mathbf{x}} \tag{5}$$

A centered leapfrog method with a *Robert* [1966] filter to suppress the leapfrog mode was used to timestep the equations.

The 2D Analog

If the prognostic variables are subjected to the constraint that $\partial_x = 0$, the equations of motion can be written:

$$\partial_t \zeta + J(\psi, \zeta) - f\partial_z U + \partial_y \rho = F_\zeta + \nu\nabla^2\zeta \tag{6}$$

$$\partial_t \rho + J(\psi, \rho) - \partial_y \psi = F_\rho + \kappa\nabla^2\rho \tag{7}$$

$$\partial_t U + J(\psi, U) + f\partial_z \psi = F_U + \nu\nabla^2 U \tag{8}$$

With the planar velocities $\mathbf{u} = (V, W)$ given by streamfunction ψ, $V = -\partial_z\psi$, $W = \partial_y\psi$. Also, $\zeta = \Lambda \cdot \mathbf{e}_\mathbf{x} = \nabla^2\psi$ and J is the nonlinear operator $J(A, B) = \partial_z A \cdot \partial_y B - \partial_y A \cdot \partial_z B$. Forcing takes the same form as the 3D simulations.

Notes on the Simulations: The Problem of Forcing

The 3D runs were performed at grid resolution 32^3 and the 2D simulations were performed at 128^2. These modest resolutions were chosen in order to be able to obtain statistically stationarity states and subsequent adequate averaging periods for a number of flow regimes. In both cases, suites of experiments were performed at Kolmogorov dissipation scales chosen to be the smallest resolvable scale $1/n$ where n is the number of grid points and the Kolmogorov scale is defined as $(\nu^3/\epsilon)^{\frac{1}{4}}$, where ϵ is the kinetic energy dissipation rate per unit volume or area $\epsilon = -\nu\mathbf{u}\cdot\nabla^2\mathbf{u}$. The Prandtl number ν/κ was unity for all cases.

So far, the form of the forcing has not been specified. The major problem is that due to limited computer storage, the forced scales are not well separated from the (Kolmogorov) dissipation scale. This will necessarily mean that results will be influenced by the form of the forcing. It

would be desirable to introduce energy at GM levels, but the requisite scales are much larger than those available to the model. At scales close to dissipation, there is some evidence for velocity isotropy [*Gargett et al* 1981]. Also, *Muller et al* [1988] infer significant vortical mode energy near dissipation (order 10m-1m) scales. For the purposes of this study, it was decided to take as a constraint a maximum entropy principle where kinetic and potential energies are forced in a 2:1 ratio (inviscid equipartition ratio) with all fields forced isotropically.

The forcing functions, *F(k)* were chosen to have energy spectra of

$$E[F(k)] = \frac{k^3}{\left(k^2 + k_0^2\right)^6} \tag{9}$$

for both 3D and 2D kinetic and potential energies; k_0 was given the value 7.

$$F(k) = A\sqrt{k}/\left(k^2 + 7^2\right)^3 \ (for\ 3D) \tag{10}$$

$$F(k) = Ak/\left(k^2 + 7^2\right)^3 \ (for\ 2D) \tag{11}$$

with A's amplitude factors. At each timestep in the running of the models, (10) and (11) were multiplied by a random phase factor, $e^{i\theta}$, $0 < \theta \leq 2\pi$. Kinetic and potential energy were forced in a 2:1 energy ratio and a separate random phase was chosen at each timestep with the proviso that the forcing of the 3D velocity field be nondivergent. The forcing function above has a maximum at wavenumber 4. It provides the criteria of being large scale, isotropic and forces KE:PE at energy equipartition rates. It also excites wave and vortical modes (see the section on waves and vortices).

The basic range of the experiments covered ratios of nonlinear to linear timescales as characterized by an r.m.s. turbulence Froude number (bv frequency, N, was unity) In 3D this is defined as $\sqrt{\left(\frac{\partial U}{\partial z}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2}$. In the 2D simulations, Froude number is defined as $\sqrt{\left(\frac{\partial V}{\partial z}\right)^2}$, the achievement of a particular Froude number meant adjusting A and ν until statistically stationary states at the desired Froude numbers and Kolmogorov scales were reached whereupon time integrations of many Brunt-Vaisala periods were performed in order to calculate average statistics.

Figure 1a shows instantaneous density slices from 3D simulations at r.m.s. turbulence Froude numbers of 0.1, 0.3, 1 and 3. Figure 1b shows the same Froude numbers for 2D. These 4 cases are chosen to span the flow regimes likely to be encountered in the ocean, from weakly wavelike to strongly overturning. The cases presented in Figure 1 are without rotation. It may be seen from Figure 1 that overturning has begun by an r.m.s Froude number of about 1 for both 2D and 3D.

Table 1 gives some of the relevant parameters for the runs shown in Figure 1. The weaknesses of the direct simulations are seen in that for the 3D cases, the Reynold's numbers $|u|l/\nu$ based on the Taylor microscale

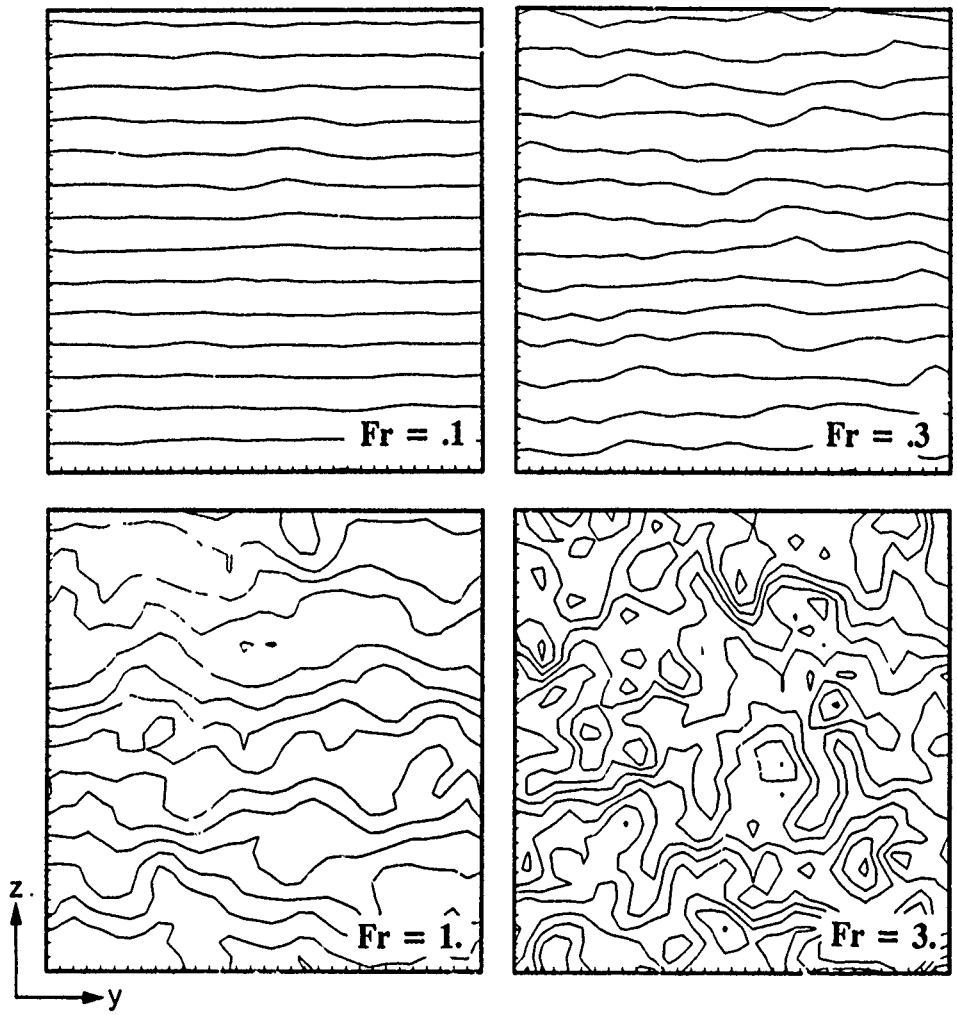


Figure 1a: Instantaneous density slices for (left to right, top to bottom), r.m.s. turbulence Froude numbers (Fr) of 0.1, 0.3, 1 and 3 for the 3D cases, all with f=0

$$l = \sqrt{\frac{15|u|^2}{|\Lambda|^2}} \tag{12}$$

are around 15–25. Any other measure of turbulent 'activity' indicates a highly viscous regime. Box sizes are small with all simulations near dissipation scales. On the other hand, the small flow domain justifies the use of a constant background gradient of density. Highly turbulent patches in a stratified environment are frequently of vertical size 10 m or less (e.g. *Moum* [1989]). Also, the measure of kinetic energy residence time KE/ϵ for the 3D Froude number 1 case is around 0.33 bv, which compares favorably with oceanic turbulence measurements [*Crawford*, 1987] .

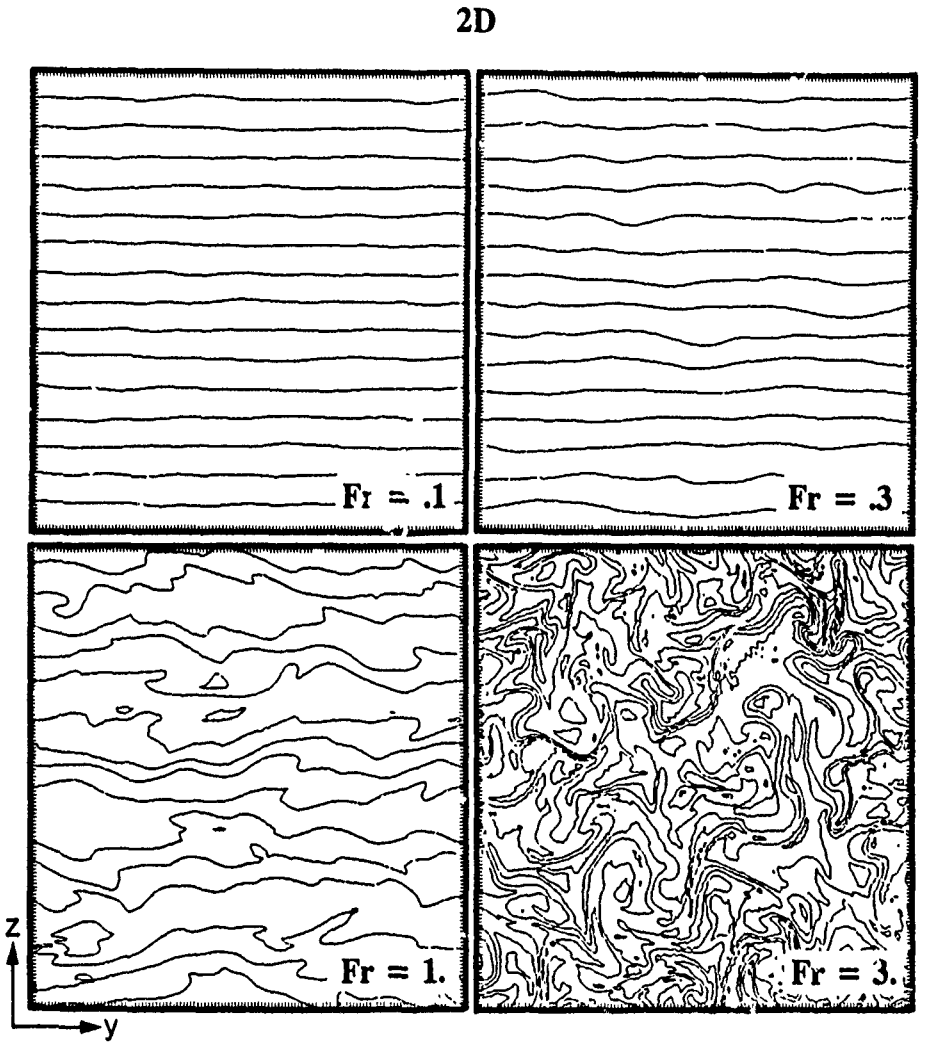


Figure 1b: Same as Figure 1a for 2D cases.

Table 1. Reynold's numbers and box sizes for the basic simulations

Froude number	3D Reynold's number	3D box size (m)	2D box size (m)
0.1	27.6	4.0	16.0
0.3	23.0	2.4	11.2
1	16.6	1.3	5.3
3	16.8	0.75	3.0

Nonlinear Energy Transfers

For the 3D dynamics, the nonlinear transfer of kinetic and potential energies can be expressed in the Fourier domain as:

$$\partial_t KE(k_h, k_z) = \mathbf{u}_\mathbf{k}^* \cdot (\Lambda \times \mathbf{u})_\mathbf{k} \tag{13}$$

and

$$\partial_t PE(k_h, k_z) = -\rho_\mathbf{k}^* (\mathbf{u} \cdot \nabla \rho)_\mathbf{k} \tag{14}$$

where * denotes complex conjugation and the transfers are collected as functions of horizontal and vertical wavenumbers k_h and k_z. This formulation is helpful in revealing details of the nonlinear processes.

In 2D, the nonlinear transfer rates can be written in the Fourier domain as:

$$\partial_t KE(k_y, k_z) = \psi_\mathbf{k}^* \{J(\psi, \zeta)\}_\mathbf{k} - U_\mathbf{k}^* \{J(\psi, U)\}_\mathbf{k} \tag{15}$$

and

$$\partial_t PE(k_y, k_z) = -\rho_\mathbf{k}^* \{J(\psi, \rho)\}_\mathbf{k} \tag{16}$$

where the horizontal wavenumber is now k_y.

By averaging over many buoyancy periods, the energy transfer rates have been calculated. Figure 2 shows contours of these energy transfer rates for the case of Froude number 1. The 3D nonlinear transfer rates are contoured as a function of horizontal and vertical wavenumber, k_h and k_z and the 2D as a function of k_y and k_z. Areas which are losing energy due to nonlinear interactions have dashed contours and areas which are receiving energy have solid contours. The contour increments are logarithmic with a minimum energy cutoff of 1.0×10^{-3} times the maximum value. In order to emphasize those wavenumbers which are contributing significantly to the overall energy transfer, areas which receive/lose more than 10% of the maximum energy transfer rate are shaded.

For the 2D case, the kinetic energy may be divided into two parts, one which will be called 'planar' ('KeP'$\equiv \frac{1}{2}\sum_k k^2|\psi|^2$) and the 'cross' energy ('KeC'$\equiv \frac{1}{2}\sum_k |U|^2$). PE is simply $\frac{1}{2}\sum_k |\rho|^2$. In Figure 2, both planar and U energy transfer rates are combined, hence it does not resemble a classical 2D energy transfer shape. For 3D, the distribution of KE and PE transfers are of similar shape and magnitude. For 2D, KE and PE distributions have somewhat dissimilar shapes, but the rates are again of comparable magnitude even though there is about twice as much kinetic energy as potential for all cases. The transfer of kinetic energy to small scales relative to potential energy appears suppressed in both 3D and 2D. A tentative exploration into the reasons for the similarity in net nonlinear transfer rates between 3D and 2D is presented in the section on Waves and Vortices.

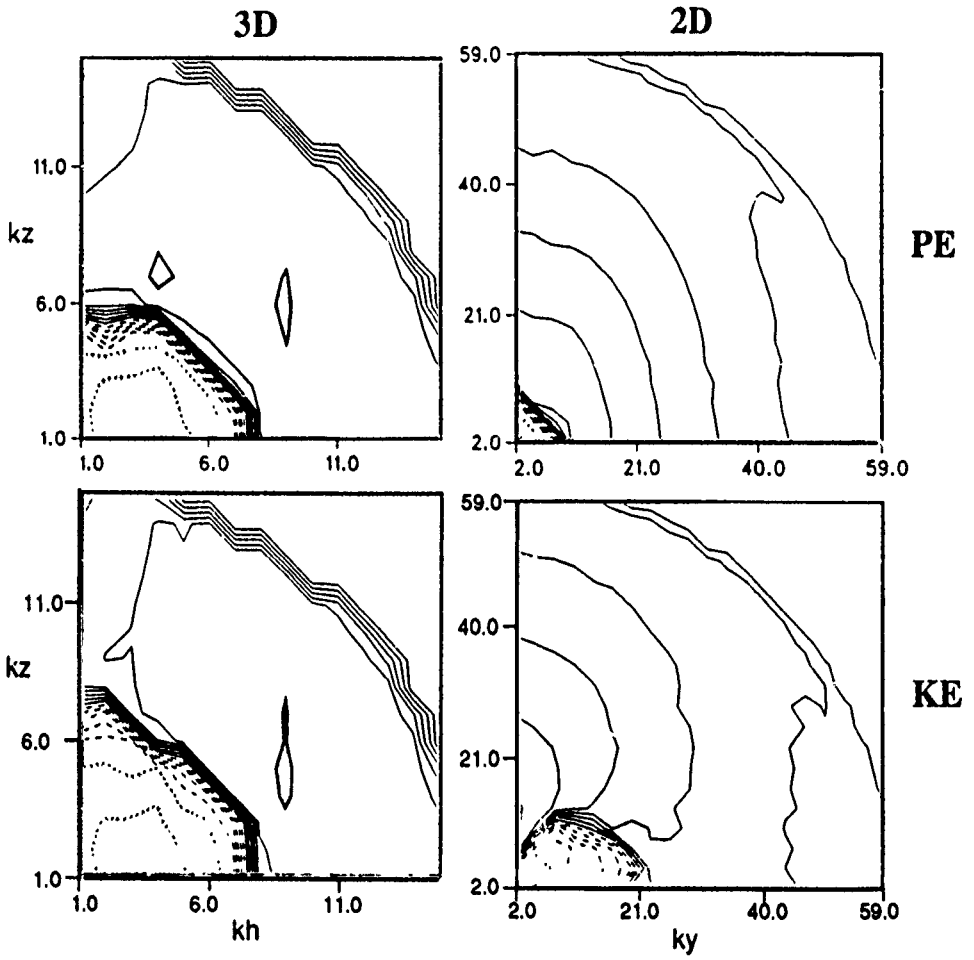


Figure 2: The net sums of potential (top) and kinetic (bottom) nonlinear transfers of energy for 3D (left) and 2D (right) both at a Froude number of 1 plotted versus horizontal and vertical wavenumbers. The logarithmic contour increments for Figures 2 and 4 are 0.71×10^{-4} for the 3D and 0.11×10^{-4} for the 2D with energy loss denoted by dashed contours and energy gain denoted by solid contours. The maximum transfer rates are 0.22×10^{-3} and 0.33×10^{-4} for 3D and 2D respectively. Areas which receive or lose more than 10% of the maximum are shaded.

Figure 3 shows the resultant energy spectra. The 3 cases presented are (Fr=0.3,f=0), (Fr=1, f=0) and (Fr=1,f=0.1). The 3D spectra are kinetic energy and twice potential energy. The 2D spectra are planar and cross kinetic energy and potential energy. For the 2D nonrotating cases, PE and planar KE would overlay exactly for a linear internal wave regime in energy equipartition. The departure from equipartition due to the nonlinear transfers is seen for the Froude number 1 case. The effect of rotation is to bring the 2D planar and U energy together with PE still greater than planar KE at the smallest scales.

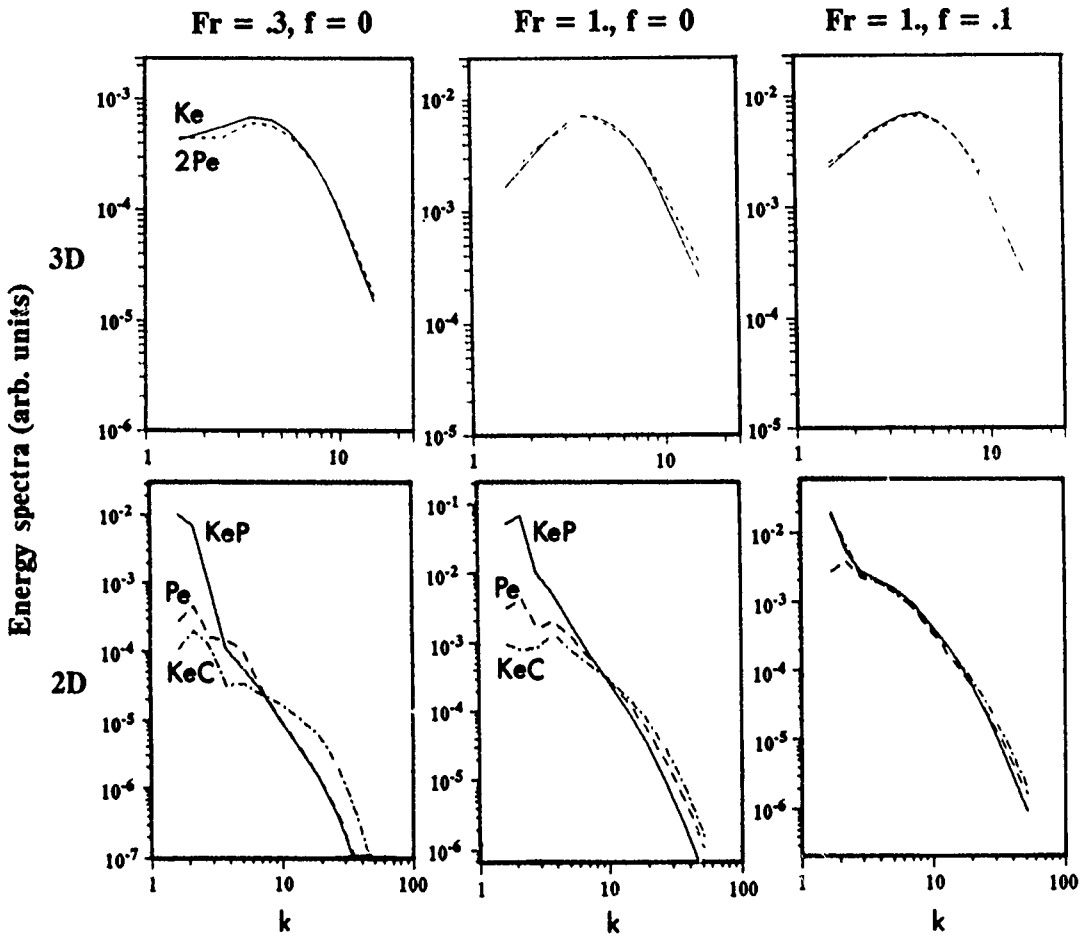


Figure 3: The 3D (top) and 2D (bottom) energy spectra for (left to right) Fr=0.3 f=0, Fr=1, f=0 and Fr=1, f=0.1 plotted versus total wavenumber. The 3D cases show kinetic (solid) and twice potential (dashed) energies. The 2D show planar (solid) and U (dashed) kinetic and potential ("Pe": dotted) energies.

For all the 3D cases, the tendency to produce an excess of potential energy relative to kinetic energy at the smallest scales (>1:2) is clearly seen with rotation having no discernible effect. For both 3D and 2D, the tendency is therefore to produce an excess of potential energy relative to kinetic energy at the smallest scales. For all cases shown in Figure 3, the low wavenumber regions have excesses of kinetic energy relative to potential energy. *Herring et al* [1982] have seen similar implications for eddy Prandtl numbers in closure models of passive scalar advection.

Figure 4 shows the buoyancy flux cross spectrum $-\langle \rho_{\mathbf{k}}^{*} W_{\mathbf{k}} \rangle$ for the same case shown in Figure 2. Time-averaged contours of this quantity are presented in the same manner as Figure 2. In both 3D and 2D, the shape of the time averaged buoyancy flux contours are negative (KE to PE conversion) at large scales and positive (PE to KE conversion) at small scales near the $k_z = 0$ axis. This shape can be interpreted as a direct result of the non-equipartition spectra seen in Figure 3.

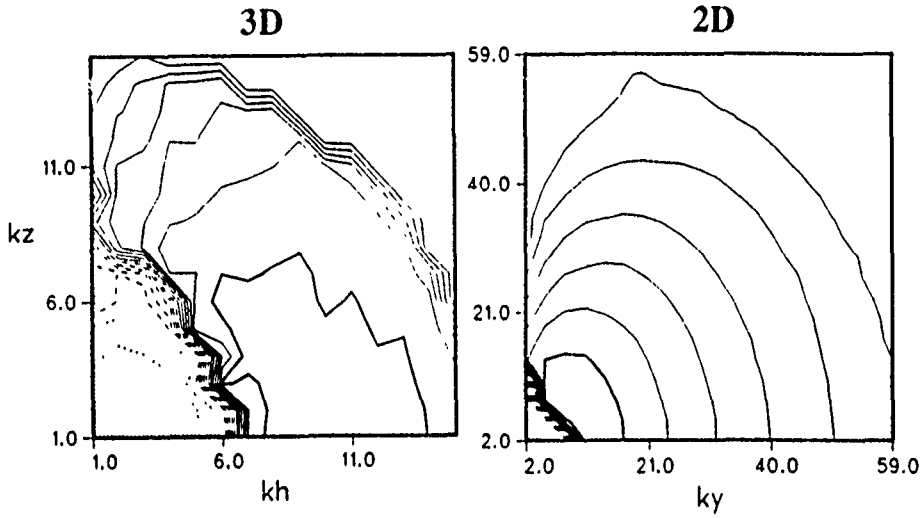


Figure 4: The buoyancy flux cross spectra from the same cases as Figure 2. The logarithmic contour increment for 3D (left) is 0.14×10^{-4} with a maximum of 0.44×10^{-4} and for 2D the increment is 0.32×10^{-5} with a maximum of 0.10×10^{-4}.

Figure 5 shows that all 3D cases have buoyancy spectra with this shape except for the weakest flow regime. *Shen and Holloway* [1986] show similar results for 2D which can also be inferred from Figure 4. *Holloway* [1988] has shown 2D spectral closure theory and simulations which take as the force term for buoyancy flux the difference between kinetic and potential energy. These yield buoyancy flux spectra not unlike those of Figure 5.

On Figure 5, the *Thorpe* [1977] overturning and buoyancy $\left(\sqrt{\epsilon/N^3}\right)$ scales have been marked where possible. Where a buoyancy scale has not been marked, it is at a scale less than the smallest available to the model. Since we have not defined 'patches' of overturning as observationalists would do, we have estimated the largest (Thorpe) scales of overturning as three times the calculated Thorpe scales. It is seen from Figure 5 that scales of mixing (negative buoyancy flux) are larger than either the Thorpe or buoyancy (where turbulence begins to dominate buoyancy) scales. The implication of this is that 'mixing' does not occur where overturning occurs, rather it occurs at the larger scales . Restratification appears to be happening at overturning scales. Kraig Winters critically addressed issues about the interpretation of mixing showing that a reordering of the ρ profile may be necessary to diagnose mixing. The reader should consider results described in the present paper from the view of *Winters* [1991].

Table 2 shows the time averaged net buoyancy fluxes, standard deviation of same and kinetic energy dissipation rate for the full range of 3D and 2D cases shown in Figure 1. As Table 2 shows, net buoyancy fluxes are on average not significantly different from zero for most 3D and 2D cases with or without rotation for these forcing ratios.

Only at a Froude number significantly greater than unity is there net negative buoyancy flux (destratification). Small amounts of overturning at a Froude number of 1 are not sufficient to

Table 2. Average net buoyancy fluxes and KE dissipation rates. All have f=0 unless noted.

Froude number	3D buoyancy flux	Standard deviation	ε	2D buoyancy flux	Standard deviation	ε
0.1	0.19 e-5	0.14 e-4	0.19 e-4	-0.25 e-6	0.14 e-4	0.13 3-5
0.3	0.15 e-4	-0.11 e-3	0.56 e-3	0.62 e-5	0.12 e-3	0.24 e-4
1	-0.47 e-4	0.11 e-2	0.25 e-1	0.15 e-3	0.16 e-2	0.16 e-2
1, f=0.1	-0.64 e-4	0.11 e-2	0.26 e-1	0.29 e-4	0.16 e-2	0.17 e-2
3	-0.20 e-1	0.13 e-1	0.73	-0.33 e-1	0.37 e-1	0.11

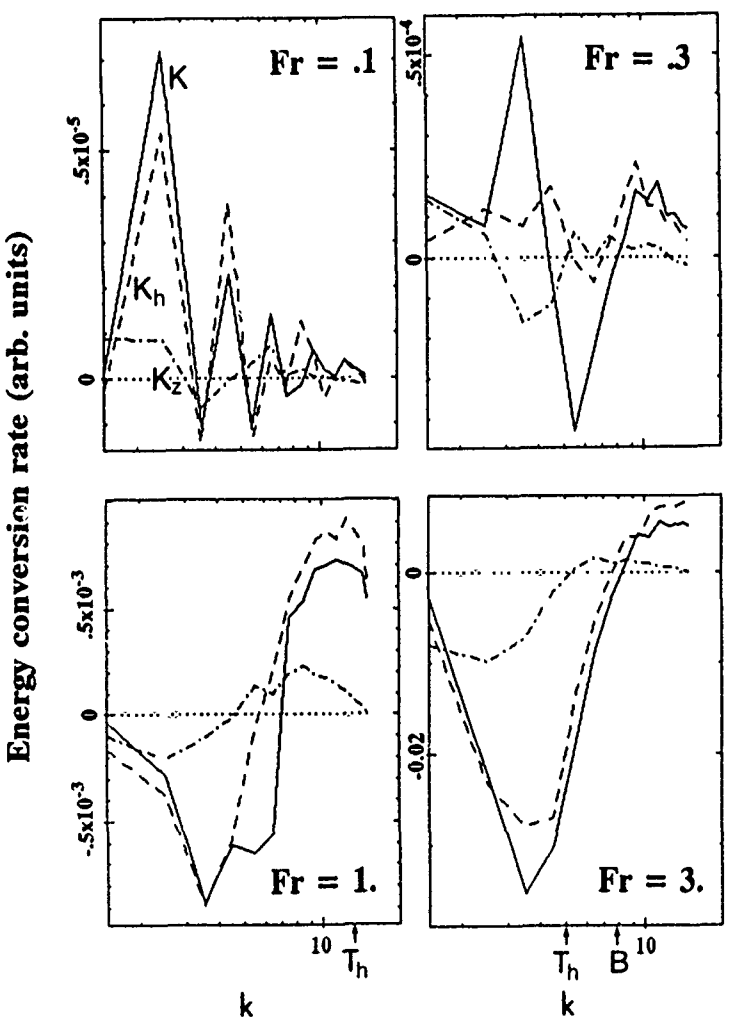


Figure 5: The 3D buoyancy flux cross spectra plotted versus total (solid), horizontal (dashed) and vertical (dot-dash) wavenumber for the 4 Froude number cases. The Thorpe and buoyancy scales are marked where possible.

cause net destratification in either 3D or 2D. A scaling of buoyancy flux from dissipation rate ($\rho' w' = 0.2\epsilon$, see *Osborne and Cox* [1972]) is only possible for one of the cases shown in Table 2 (2D, Froude number 3.). Obviously, the results are dependent on the ratio of forcing of KE:PE. *Moum* [1989] has reported directly measured values from overturning patches of $\rho'\omega' = 50\ \mathrm{W\ m^{-2}}$; $0.2\varepsilon = 420\ \mathrm{W\ m^{-2}}$. Moum cautions against taking these values too literally, but they are like the average values seen in the simulations using the energy equipartition rates of forcing.

The tendency to produce positive buoyancy fluxes at the high wavenumber end of an inertial subrange is very robust to the form of forcing. *Holloway and Ramsden* [1988] show a case without rotation which was run with forcing of kinetic energy only with correspondingly large negative net buoyancy fluxes. Even though the forced scales are very close to the Kolmogorov scale, a positive buoyancy flux 'tail' was observed. Other examples of this robustness are given in the same paper.

WAVES AND VORTICES

An attempt has been made to understand why 3D kinetic energy transfer is suppressed relative to potential energy by separating nonlinear transfers into component parts, viz: if all nonlinear, dissipative, and forcing terms are omitted from the 3D equations of motion then the ideal linearized model equations become

$$\partial_t \mathbf{u_k} = \frac{\mathbf{k}}{|k|^2}[k_z \rho_\mathbf{k} - f(k_x V_\mathbf{k} - k_y U_\mathbf{k})] - \rho_\mathbf{k}\mathbf{e}_z + f(V_\mathbf{k}\mathbf{e}_x - U_\mathbf{k}\mathbf{e}_y) \tag{17}$$

$$\partial_t \rho_k = -W_\mathbf{k} \tag{18}$$

The resultant system of equations (details may be found in RH) has eigenvalues zero and $\pm\frac{1}{|k|}\sqrt{k_h^2 + f^2 k_z^2}$ with corresponding eigenvectors

$$\begin{bmatrix} a_0 \\ a_+ \\ a_- \end{bmatrix} = \begin{bmatrix} 1/c & 0 & 1 \\ -c & d & 1 \\ -c & -d & 1 \end{bmatrix} \begin{bmatrix} v_c \\ v_a \\ \rho_\mathbf{k} \end{bmatrix} \tag{19}$$

with $c = f\frac{k_z}{k_h}$, $d = \iota\sqrt{1 + c^2}$ and v_c, v_a are velocity components orthogonal to k. The a_0 is the nonpropagating vortical mode and $a_\pm$ are wave modes propagating at frequencies $\pm\frac{1}{|k|}\sqrt{k_h^2 + f^2 k_z^2}$. The 2D system possesses a similar set of solutions.

Note that the right hand sides of equations (13) and (14) are the product of three quantities, each of which is the sum of waves and vortices. Eight contributions to the net nonlinear transfers are thus possible. These are denoted by, *e.g.*, W-W-V where the interaction of waves and vortices is projected onto the wave component of the flow field. V-W-W would be wave-wave interactions projected onto the vortical part of the flow field. The 2D has a similar decomposition.

Examination of these transfer components have shown that for the 3D case the forward transfer of potential energy to small scales is matched by a corresponding kinetic energy transfer mode with the following notable exception. Figure 6 shows the kinetic and potential energy W-W-W interactions for both 3D and 2D to the same scale as Figure 2. This transfer mode appears to be the key player in the suppression of 3D KE transfers (at least for wave energy). The rate of energy transfer is *much* greater for potential energy than kinetic energy for *both* 3D and 2D and the shapes are the same, too. The point of this section is to isolate in some degree where the suppression of KE transfer relative to PE transfer is occurring and to reemphasize the similarity of 3D and 2D results.

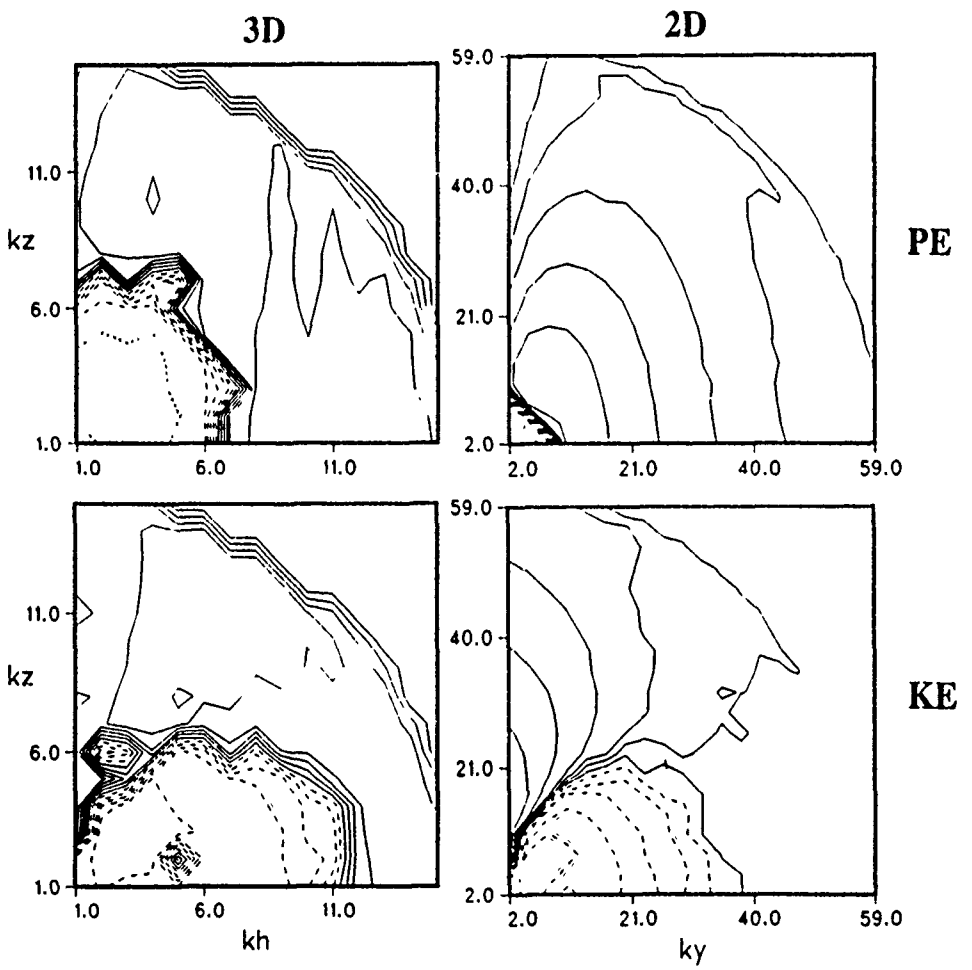


Figure 6: The W-W-W transfer component potential (top) and kinetic (bottom) energy transfers for 3D (left) and 2D (right) plotted versus vertical and horizontal wavenumbers. The scaling is as Figure 2.

CONCLUSIONS TO Part 1

The nonlinear transfer of kinetic and potential energy in an internal wave regime near dissipation scales has been investigated by direct numerical simulations. Two cases were considered, one which solves the full 3D Navier-Stokes equations and another which restricts the motion field to a vertical 2D plane. It was found that the rate of transfer of kinetic energy is less than that of potential energy for both 3D and 2D and the net transfer rates have similar spectral shapes. Although suppression of the 2D kinetic energy transfer is explained by the energy/enstrophy conservation restraint, the reason for KE suppression in 3D is not well understood.

The difference of transfer rate of KE and PE creates an imbalance in the relative amounts of kinetic and potential energies at different scales, more PE at small scales and more KE at the larger scales. The relative excesses of energy in turn drive a characteristic buoyancy flux spectrum, negative at large scales and positive at small scales, with the results from 2D resembling those of 3D. The tendency to produce a relative excess of potential energy at the high wavenumber end of an inertial range, resulting in positive (restratifying) buoyancy fluxes, is robust to the form of the forcing.

Rotation up to f=0.1 slightly alters the form of the 2D energy spectra but does not significantly affect the differential transfer rates leading to the buoyancy flux measurements. Rotation to f=0.1 has little effect on the 3D results due to the short energy residence times.

Net buoyancy fluxes are on average zero except at Froude numbers much larger than 1. There is little net mixing in a marginally overturning environment, either in 3D or 2D. An Osborne-Cox scaling to derive buoyancy flux from kinetic energy dissipation rates appears to be only possible in the most active of turbulent patches for the forcing ratios used here.

It may seem surprising that 2D dynamics resemble so nearly the full 3D regime with respect to the nonlinear behavior of internal waves near dissipation scales. It is encouraging to think that cost-effective numerical experiments in 2D may be relevant to real-world applications.

Part 2: SURFACE LAYER FORCING

One may criticize Part 1 apart from the extremely limited flow domain. Energy 'appears' by application of fictitious body forces which are intended to represent interactions with scales larger than the resolvable domain. Energy introduced at internal wave equipartition levels is moved to different scales by nonlinear processes, but no energy is created or destroyed by this process. Hence, it is perhaps not unexpected that the observed net buoyancy fluxes were small.

Holloway, [1988], [1989] has raised the issue that presumptions about gradient production and dissipation rates such as that of *Osborne and Cox* [1972] may be on shaky theoretical grounds. In order to gain insight into the relative strengths of terms in the Navier-Stokes equations, we have attempted to measure them in a simulation domain with a vertical flow of energy.

METHOD

The 2D simulation code used in Part 1 has been modified to have a free slip, rigid lid condition on ψ and ρ at grid resolution 128*256 (and preliminary 128*512). The forcing function has been masked by

$$M(z) = e^{-(\frac{3}{2}z)^2}\ ;\ -\pi < z < 0 \tag{20}$$

where the simulation domain is $-\pi < z < 0$. At each timestep, ψ and ρ have been multiplied by a bottom absorption taper function

$$T(z) = 1 - C\left\{1 - \left(10\left[1 + \frac{z}{\pi}\right]\right)^2\right\}\ ;\ -\pi < z < -0.9\pi \tag{21}$$

with C a constant.

In this way, it was attempted to define an 'energy radiation' region between the forced region and the sponge region which mimics the passage of energy down the fluid column. To simplify the energetics, the 'cross' velocity field has been eliminated (no rotation). Figures 7a and 7b show instantaneous snapshots of stream function, density and kinetic+potential energy density for one of the cases. It is seen that energy is concentrated in the 'surface' layer.

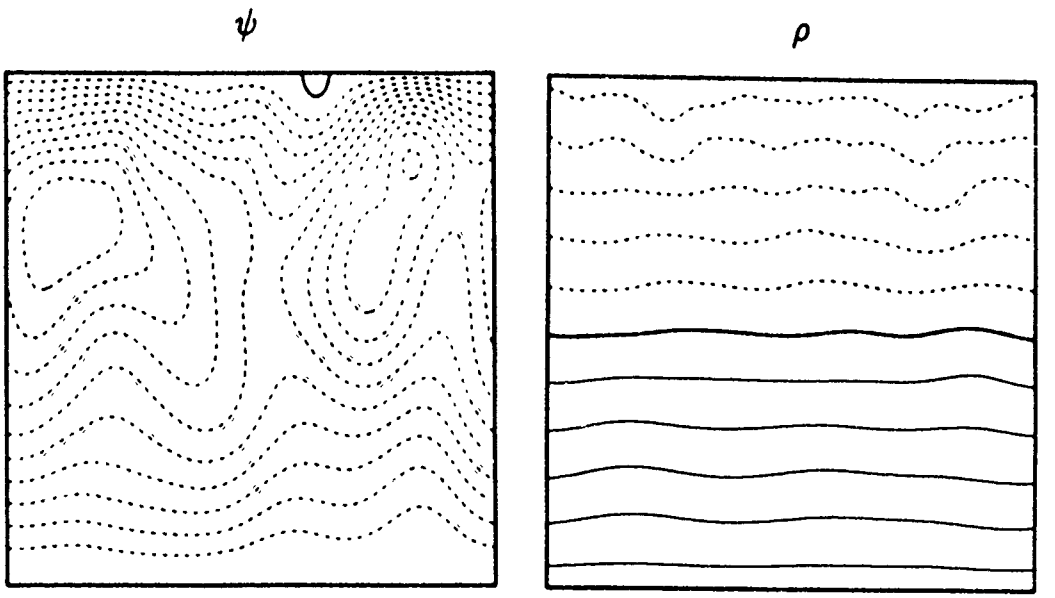


Figure 7a: Instantaneous snapshots of (left) stream function and (right) total density for a surface layer forced run.

Figure 8 shows vertical profiles of kinetic and potential energy and energy dissipation rates. It is seen that the presence of the sponge region creates gradients near the bottom which tend to enhance the dissipation rates, and there is probably some reflection of energy from the bottom.

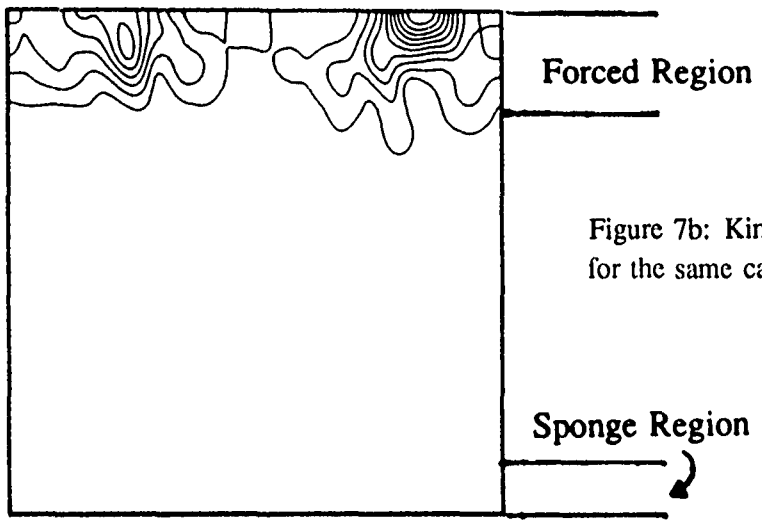


Figure 7b: Kinetic + potential energy density for the same case as Figure 7a.

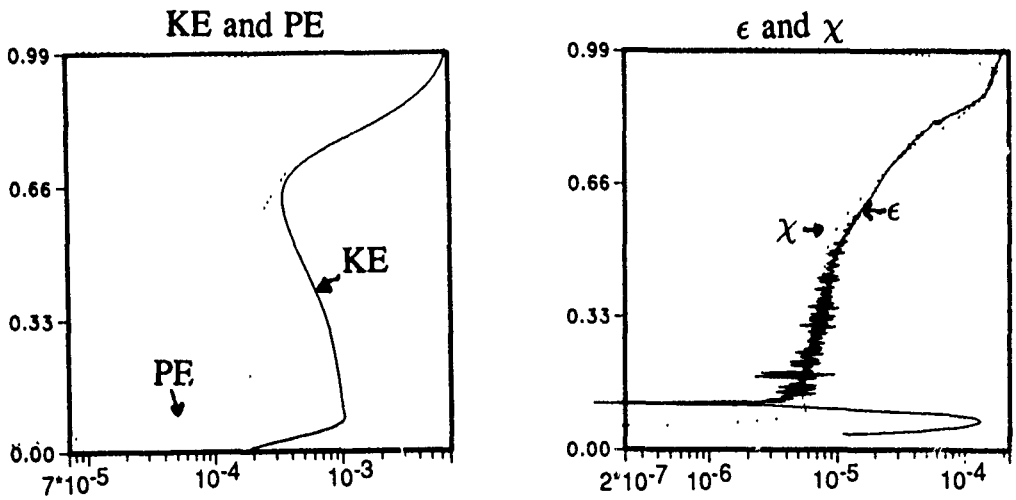


Figure 8: Time and horizontally averaged vertical profiles of (left) kinetic (solid) and potential (dotted) energy and (right) dissipation rates of kinetic (solid) and potential (dotted) energy.

Figure 9 shows energy spectra for the area away from the forced and sponged region only and a vertical profile of Froude number. The reduced area spectra indicate that there is more KE than PE away from the surface layer. This conclusion is supported by Figure 10, a vertical profile of buoyancy flux and a spectrum of buoyancy flux away from the mixed layer. For this case the net buoyancy flux is clearly negative in the radiation region but the spectrum is showing the tendency to develop a positive 'tail' as was seen in the homogeneously forced cases.

Nonlinear Energy Radiation Terms

In the radiation region, what are the sources and sinks of energy? We have attempted to answer this question by determining an energy balance in the non-forced region. From equation (1)

$$\partial_t(\mathrm{u}^2/2) = \nabla \cdot (\mathrm{up} + \mathrm{uu}^2/2) - \rho W - \epsilon \qquad (22)$$

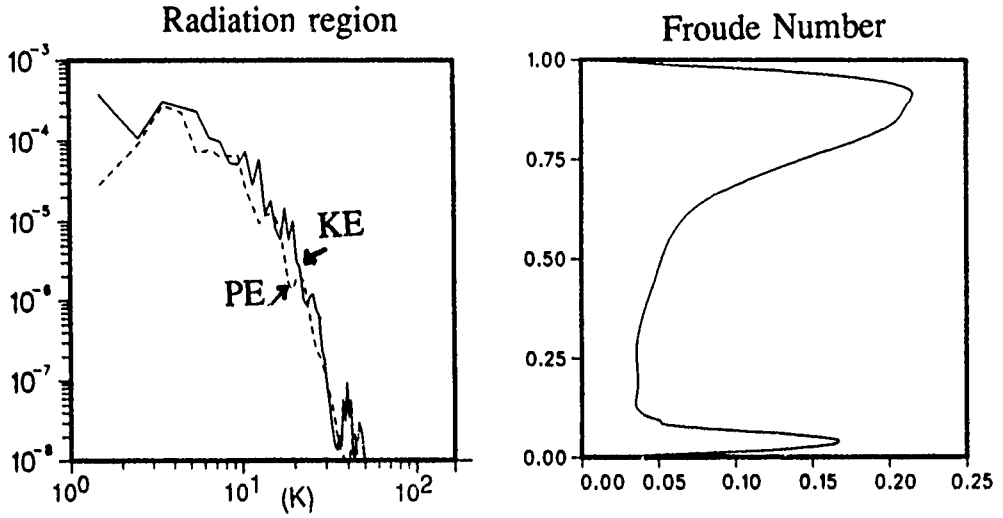


Figure 9: (left) Time averaged kinetic (solid) and potential (dashed) energy spectra plotted versus total wavenumber for radiation region only: (right) vertical profile of Froude number.

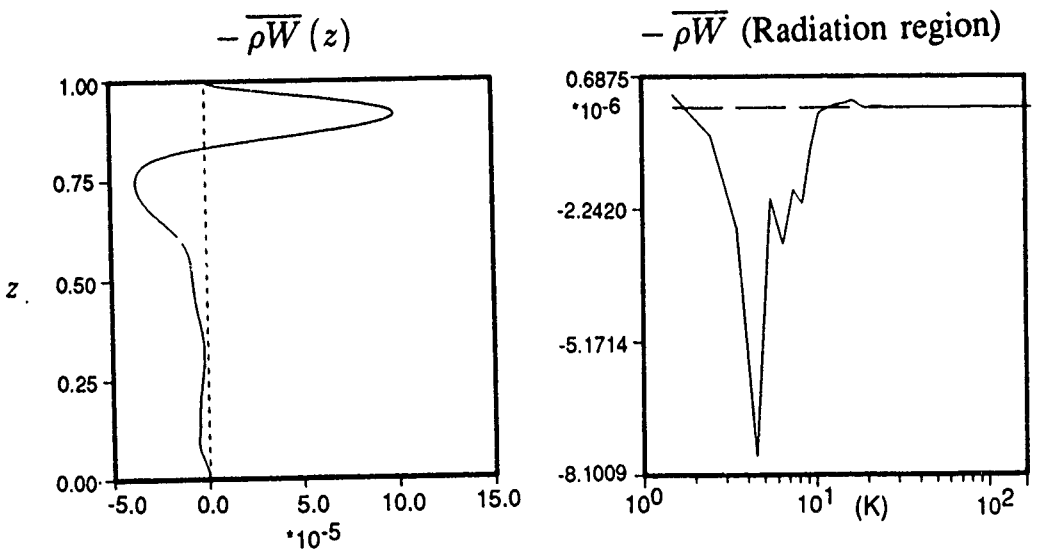


Figure 10: (left) Time and horizontally averaged profile of buoyancy flux; (right) time averaged buoyancy flux spectra from the radiation region plotted versus total wavenumber.

The nonlinear terms are composed of a divergence of a pressure-u correlation which we denote as 'up' and a turbulent flux term which we denote as 'tf'. The ρW is minus buoyancy flux and ϵ is kinetic energy dissipation rate. From equation (2)

$$\partial_t(\rho^2/2) = \nabla \cdot (\mathrm{u}\rho^2/2) + \rho W - \chi \tag{23}$$

where χ is potential energy dissipation rate and we will denote the nonlinear term as 'nl'.

Assuming that stationarity has been reached and the ∂_t terms can be dropped, which of the remaining terms in equations (22) and (23) are dominant? Figure 11 left shows vertical profiles of time and horizontally averaged 'nl', 'up' and 'tf' and Figure 11 right shows sums of terms for KE and PE. Clearly, it is the divergence of the pressure-u correlation which carries (kinetic) energy down the water column. Figure 11 right indicates a qualitative balance between nonlinear production, buoyancy flux and dissipation in the radiation region for both kinetic and potential energy (forcing and sponge region effects were not factored into the balances).

At level z=-π/2 in the radiation region, the average values of balance terms for two r.m.s Froude numbers are presented in Tables 3a and 3b. Table 3b indicates that the major source of potential energy in the radiation region is the buoyancy flux and the nonlinear term tends to remove potential energy. Table 3a and 3b indicate that buoyancy flux is of comparable magnitude to ϵ or χ (which are about equal).

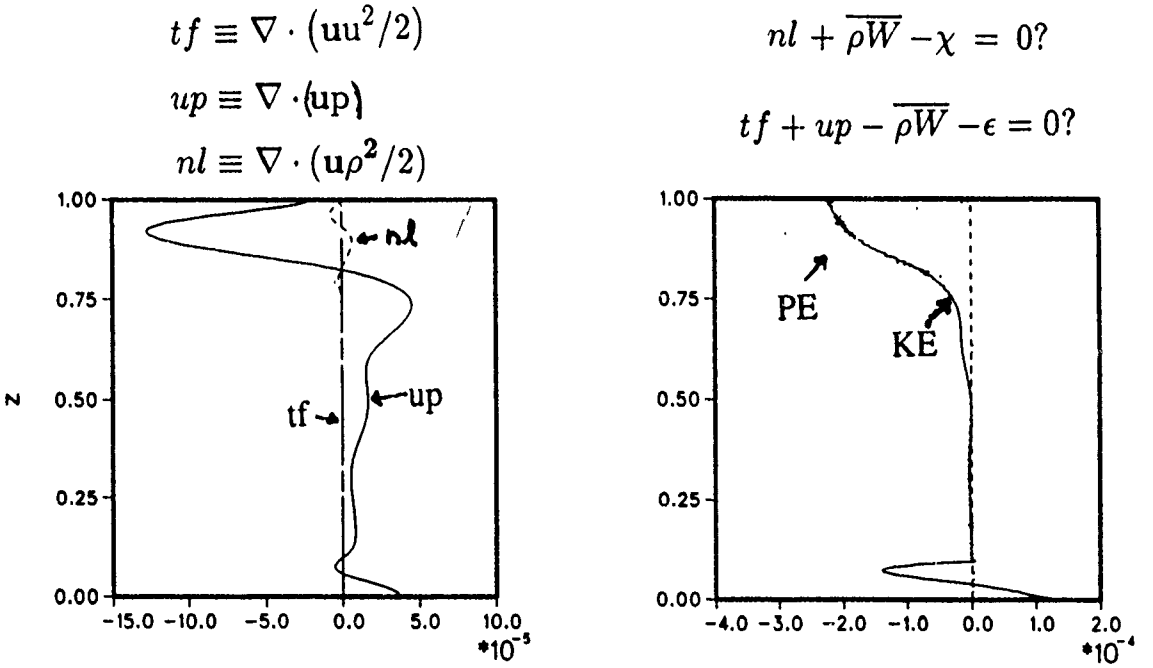


Figure 11: (left) Time averaged profiles of divergence of pressure-u correlation (solid), $\mathbf{u}u^2/2$ (dotted, not visible) and $u\rho^2/2$ (dashed). (Right) Time and horizontally averaged profiles of kinetic (solid) and potential (dotted) energy balance sums along with the terms which go into the sums.

CONCLUSIONS TO Part 2

For the cases studied here, energy from the surface forced layer moves down the fluid column by the pressure-u correlation. The turbulent flux term in the equations of motion is negligible compared to the divergence of the pressure-u correlation. The kinetic energy supplied to the fluid interior is balanced by dissipation and conversion to potential energy via buoyancy flux. The buoyancy flux is hence on average negative in the radiation region, but the spectrum

Table 3a. Kinetic energy balance terms.

Froude number	up	tf	b*w	ε	sum
0.05	0.17 e-4	-0.46 e-10	-0.74 e-5	-0.11 e-4	-0.165 e-5
0.15	0.44 e-4	-0.93 e-8	-0.22 e-4	-0.13 e-4	0.85 e-5

Table 3b. Potential energy balance terms

Froude number	nl	-b*w	χ	sum
0.05	-0.24 e-7	0.74 e-5	-0.84 e-5	-0.11 e-5
0.15	-0.18 e-5	0.22 e-4	-0.15 e-4	0.51 e-5

shows the same qualitative behavior (negative at large scales, positive at small scales) as the statistically homogeneous forced cases.

The major source of potential energy in the fluid interior is the buoyancy flux and it is roughly balanced by dissipation and (relatively small) vertical nonlinear transfer. Taken together, Parts 1 and 2 indicate the hazard of scaling buoyancy flux rates from dissipation rates. Indication is that net buoyancy flux occurs during times in which energy is 'filling' a region. When energy levels reach sufficient magnitude to cause a dissipation 'event', there may be very little net gradient flux associated with it, but this may have little to do with 'mixing' (again *Winters* [1991]). Future work will concern attempting to extend the parameter range to higher Froude number radiation regions and completing runs at the higher resolution (larger radiation region).

ACKNOWLEDGMENTS

The assistance of Patricia Kimber in drafting the figures is gratefully acknowledged. This work was supported in part by the Office of Naval Research under grant number N00014–87–J-1262.

REFERENCES

Crawford, W., A Comparison of Length Scales and Decay Times of Turbulence in Stably Stratified Fluids, *J. Phys. Ocean.*, 16, 1847–1854, 1987.

Gargett, A., P. Hendricks, T. B. Sanford, T. R. Osborn and A. J. Williams III, A Composite Spectrum of Vertical Shear in the Upper Ocean, *J. Phys. Oceanogr.*, **11**, 1258–1271, 1981.

Herring, J. R., D. Schertzer, M. Lesieur, G. R. Newman, J. P. Chollet and M. Larchevsque, A Comparative Assessment of Spectral Closures as Applied to Passive Scalar Diffusion, *J. Fluid. Mech.*, **124**, 411–437, 1982.

Holloway, G., A Conjecture Relating Oceanic Internal Waves and Small–Scale Processes. *Atmos. Ocean*, **21**, 107–122, 1983.

Holloway, G., The Buoyancy Flux from Internal Gravity Wave Breaking, *Dyn. Atmos. Oceans*, **12**, 107–125, 1988.

Holloway, G., Relating Turbulence Dissipation Measurements to Ocean Mixing, Proc. 'Aha Huliko'a Hawaiian Winter Workshop, ed. P. Muller and D. Henderson, Hawaii Inst. of Geophys., 329–339, 1989.

Holloway, G., and D. Ramsden, Theories of Internal Wave Interaction and Stably Stratified Turbulence: Testing Against Direct Numerical Experimentation, in Small-Scale Turbulence and Mixing in the Ocean, J. C. H. Nihoul and B. M. Jamart eds., Elsevier, 363–378, 1988.

Moum, J. N. , Measuring Turbulent Fluxes in the Ocean, the Quest for K_ρ, Proc. 'Aha Huliko'a Hawaiian Winter Workshop, ed. P. Muller and D. Henderson, Hawaii Inst. of Geophys., 145–156, 1989.

Muller, P., R-C. Lien and R. Williams, Estimates of Potential Vorticity at Small Scales in the Ocean, *J. Phys. Oceanogr.*, **18**, 401–416, 1988.

Orszag, S. A., Numerical Simulation of Incompressible Flows Within Simple Boundaries. I. Galerkin (spectral) Representations, *Stud. in Appl. Math.*, 50, 293–297, 1971.

Osborne, T. R. and C. S. Cox, Oceanic Fine Structure, *Geophys. Fluid Dyn.*, 3, 321–345, 1972.

Ramsden, D. and G. Holloway, Direct Simulation and Turbulence Closure Evaluation for Large Amplitude Internal Wave Interactions in the Vertical Plane, *Proc. of Third Intl. Sympos. on Stratified Flows*, Cal. Inst. of Technology, Vol. 1, 1987.

Ramsden, D. and G. Holloway, Energy Transfers across an Internal Wave/Gravity Mode Spectrum, preprint, 1990.

Riley, J., R. W. Metcalfe, and M. A. Weissman, Direct Numerical Simulations of Homogeneous Turbulence in Density stratified Flows, in Nonlinear Properties of Internal Waves, AIP Conf. Proc. No. 76, 79–112, 1981.

Robert, A., The Interpretation of a Low Order Spectral Form of the Primitive Meteorological Equations, *J. Met. Soc. of Japan*, 44, 237–244, 1966.

Shen, C. Y., and G. Holloway, A Numerical Study of the Frequency and Energetics of Nonlinear Internal Gravity Waves, *J. Geophys. Res.*, 91, C1, 953–973, 1986.

Thorpe, S. A., Turbulence and Mixing in a Scottish Loch, *Phil. Trans. Roy. Soc. London*, A286, 125–181, 1977.

Winters, K., Diagnosing Mixing in Numerical Models, Proc. 'Aha Huliko'a Hawaiian Winter Workshop, ed. P. Muller and D. Henderson, Hawaii Inst. of Geophys., in press, 1991.

LARGE-EDDY SIMULATION OF INTERNAL WAVE MOTIONS

DAVID A. SIEGEL

Computational Oceanography Group, Department of Geography
University of California at Santa Barbara, Santa Barbara, California, 93106

ABSTRACT

Large-eddy simulation (LES) techniques are used to numerically investigate the motions and mixing processes of the oceanic internal wave field. LES's solve the three-dimensional, nonlinear, time dependent Navier-Stokes equations for the resolved scales of motion while parameterizing the effects of the unresolved subgrid scales. This allows oceanographically relevant length and time scales to be numerically simulated using present-day supercomputers.

Here, the results (some preliminary) of two recent LES experiments of stably-stratified ocean turbulence are presented. The first experiment is the study of the decay of stably-stratified turbulence at oceanographically relevant length and time scales. The initial wavenumber distribution is consistent with the Garrett-Munk (GM) internal wave spectrum extrapolated to scales smaller than 10 m. This experiment clearly shows many features of the so-called "turbulent collapse" (i.e., the transition from a fully turbulent flow to an internal wave dominated field). However, the present results illustrate several important differences from previous laboratory and *direct* numerical results. In particular, highly anisotropic, "pancake" structures are not formed coincident with the onset of the turbulent collapse. We hypothesize that the lack of two-dimensionalization of the kinetic energy structures is due to the relatively large Reynolds numbers of the LES experiments and suggest that the results of laboratory and *direct* numerical experiments be cautiously applied to the internal wave field.

The second experiment is a preliminary attempt to use LES techniques to directly determine values of vertical eddy diffusivities for steady thermocline motions. The Navier-Stokes equations are forced to maintain constant energy, consistent with GM amplitudes, for the lowest resolved wave modes (50 m). The phases of the forced wave mode are determined dynamically by solving the equations of motion. The long time integration (≥30 Nt) and steady forcing enables the domain averaged buoyancy flux and hence, the vertical eddy diffusivity (K_{BF}) to be determined directly. Values of K_{BF} are consistent with oceanographic inferences (6×10^{-6} to 6×10^{-4} m^2 s^{-1}) although their dependency upon the buoyancy frequency is quite anomalous (going as N^{+2}!!). This anomalous dependency is thought to be caused by the nature of the numerical forcing. The direct determinations of K_{BF} are compared with several indirect turbulent variance dissipation rate estimates. The forced LES experimental results are preliminary as many detailed analyses still need to be performed.

MODELING THE OCEANIC INTERNAL WAVE FIELD

A variety of scientific questions may be addressed using "data" obtained from the numerical simulation of the ocean's internal wave field. For instance, important parameters that are very difficult to directly measure *in situ*, such as diapycnal and isopycnal eddy diffusivities, may be directly

determined numerically. This would enable parametric dependencies to be directly investigated. Diapycnal eddy diffusivities for active (temperature and salinity) and passive (tracer) scalars may be compared to examine the importance of buoyancy on the mixing of scalars. Similarly, the parametric dependence of turbulent variance dissipation rates can be evaluated. In addition, the consistency of the linear internal gravity wave dispersion relation with the simulated "data" may be evaluated. Knowing the spatial structure and temporal evolution of the internal wave field, the potential of many oceanographic applications may be evaluated. For example, the effects of internal wave motions upon the transmission of acoustic signals can be determined using simulated "data" (e.g., T. Ewart, this volume). The sampling characteristics of oceanographic instrumentation may also be evaluated (e.g., Ledwell and Watson, 1991; Siegel and Plueddemann, 1991).

The point is that all of the above scientific questions and technical needs require "data" from the internal wave field that are difficult to obtain from current observational techniques. Here, we illustrate the application of large-eddy simulation (LES) techniques to the oceanic internal wave field as a method for *supplementing* our present observational tool set. The idea is to use *in situ* observations of the internal wave field to drive a LES model in order to calculate those properties which are difficult to obtain observationally (such as, eddy diffusivities or dissipation rates). This approach has proven to be quite successful in studies of the atmospheric boundary layer (e.g., Deardorff, 1973; Wyngaard, 1984; Ebert *et al.*, 1989).

In order to realistically simulate the internal wave field, several rather stringent requirements must be satisfied. First, it is anticipated that the motions of the internal wave field are fully three-dimensional and must be modeled that way. Second, characteristic nondimensional numbers (cf., Reynolds, Froude and Rossby) are important and must be correctly considered in any investigation. Third, relevant spatial scales range from 10's of cm's to many 100 m's, while temporal scales range from minutes to days (e.g., Garrett and Munk, 1972; 1975; Munk, 1981; Gargett *et al.*, 1981; 1984). Fourth, the internal wave field is comprised not only of *linear* internal waves, but also many turbulent and vortical motions (e.g., Holloway, 1983; Müller *et al.*, 1986; Müller, 1988). Thus, the full suite of allowable motions must be included. The most straightforward way to insure that these requirements are met is to directly solve the Navier-Stokes equations for the proper time and space scales.

Unfortunately, computational resources are finite. Only a limited number of spatial modes can be numerically simulated using present-day supercomputers (O(100) in each direction). This puts rather severe restrictions on the spatial domain that one can investigate by *directly* solving the Navier-Stokes equations. When directly solving the Navier-Stokes equations, *all* scales relevant to the transport and dissipation of energy *must* be explicitly resolved. The bandwidth of required spatial scales is described by the Reynolds number (e.g., Tennekes and Lumley, 1972). For the oceanic internal wave field, Reynolds numbers ($u'h/\nu$; where ν is the kinematic viscosity = 10^{-6} m^2 s^{-1}) based upon a 1 m overturning length scale (h) and a velocity scale (u') of 0.1 m s^{-1} are O(10^5). This indicates that molecular processes are not directly relevant to the internal wave field's evolution. Values of the Reynolds number (Re_λ) based upon the Taylor length scale ($\lambda=u'/(\partial u/\partial x)'$ for an isotropic field) are often used for comparing dissimilar approaches of studying turbulence (e.g., Lesieur, 1987). For the oceanic internal wave field, values of Re_λ are O(10^5) (using $u'=0.1$ m s^{-1} and $\varepsilon_{KE}=10^{-8}$ m^2 s^{-3}). However for laboratory and *direct* numerical simulations of stably-stratified turbulence, reported

values of Re_λ are always less than 100 (e.g., Riley *et al.*, 1981; Stillinger *et al.*, 1983; Itsweire *et al.*, 1986; Holloway and Ramsden, 1988; Métais and Herring, 1989; Lienhard and Van Atta, 1990; Ramsden and Holloway, 1991). Thus, it is clear that both laboratory and *direct* numerical investigations can not adequately simulate the oceanic internal wave field. This is not to say that these experiments have not given valuable insights into the innerworkings of stably-stratified turbulence, just that they cannot realistically simulate the ocean's internal gravity wave field.

SIMULATING THE LARGE EDDIES

To get around the finite resolution afforded by present computational resources, we apply the large-eddy approximation (e.g., Deardorff, 1970; 1973; Schumann, 1975; Ferziger, 1983; Wyngaard, 1984; Rogallo and Moin, 1984; Lesieur, 1987). In a large-eddy simulation (LES), the large, or grid scale (GS), motions are directly simulated while the effects of the subgrid scales (SGS) are parameterized in terms of the GS. Most importantly, the modeling of the SGS processes allows oceanographically relevant Reynolds numbers to be simulated in a numerically achievable wavespace. The application of the large-eddy approximation means that the GS motions contain the energy of the flow and perform the mixing (that is, support a buoyancy flux), while the SGS eddies act to dissipate energy as part of the turbulent cascade. This separation of the dynamical roles of the GS and SGS motions provides a useful means of examining the consistency of a LES. That is, GS energy levels should be much greater than the SGS, while the GS dissipation rates must be much smaller than the SGS rates of dissipation. These simple concepts are helpful for evaluating LES model performance.

In a LES, the cutoff scale between GS and SGS motions is generally made within the inertial subrange of turbulence. This is done for two reasons. First, motions at these spatial scales should be locally isotropic. Second, the rate at which kinetic (KE) and potential (PE) energies are transferred across the SGS cutoff must be equal to the rate at which they are dissipated by molecular processes (the KE and PE dissipation rates, ε_{KE} and ε_{PE}). These facts make the parameterization of SGS processes simpler than second moment turbulence closures which must hold over all scales of motion.

It is important to recognize that the essential difference between a large-eddy simulation and a *direct* simulation of any turbulent flow is the development of a parameterization for the effects of the SGS processes in terms of the resolved scale dynamics. In some sense, the degree that the large-eddy simulation conforms to its intended reality is a function of how well the SGS parameterization performs its task. This suggests that one must carefully choose a SGS parameterization method. Unfortunately, there have been very few comparison studies of SGS model performance (e.g., Clark *et al.*, 1979; Schumann, 1991). However, these results, as well as some reported here, indicate that the choice of SGS models makes little difference to the resolved scale energetics. This result is not really unexpected as the GS cutoff is made within the inertial subrange of turbulence and the SGS motions have "forgotten" the orientation of the GS motions that are driving them. This means that a simple statistical accounting of the GS energy sink due to SGS energy cascade, such as by a uniform eddy viscosity, may be enough to successfully simulate the large scale motions (e.g., Lesieur, 1987).

LES techniques have long been applied to atmospheric science and engineering problems (e.g., Deardorff, 1970; 1973; 1980; Schumann, 1975; 1991; Ferziger, 1983; Wyngaard, 1984; Rogallo and Moin, 1984; Moeng, 1984; Eidson, 1985; Lesieur, 1987; Moeng and Wyngaard, 1988; Schmidt and

Schumann, 1989; Mason and Derbyshire, 1990). However, they have only been recently applied to oceanographic flows (e.g., Siegel, 1988; 1991; Gallacher, 1990; McWilliams *et al.*, 1990; Siegel and Domaradzki, 1991a; 1991b). One should note that the modeling of SGS processes is similar in spirit to the parameterization of SGS processes in eddy-resolving oceanic general circulation models (e.g., Holland and Lin, 1975). The major difference is that more is known about the dynamics of 3-D, locally isotropic turbulence than about submesoscale (O(1-10 km)) ocean motions.

Here, we will illustrate the development of a LES model and its application for investigating the interactions among internal gravity waves and turbulence within the thermocline. One of our goals is to use LES techniques as a predictive tool for evaluating the magnitude and parametric dependence of diapycnal and isopycnal eddy diffusivities for use in regional and global scale ocean circulation models. Here, we will describe the results of two LES experiments. The first experiment describes the decay of stably-stratified turbulence at oceanic space and time scales, while the second applies a LES model to directly determine values of the diapycnal eddy diffusivity. This work is preliminary, but clearly illustrates a "proof of concept".

DEVELOPMENT OF A LES MODEL FOR THE INTERNAL WAVE FIELD

The motions of the resolved scale eddies are determined by solving the 3-D Navier-Stokes equations for an incompressible fluid, satisfying the Boussinesq approximation, where temperature is the active scalar. Using these assumptions, the equations of motion are

$$\frac{\partial u_i}{\partial t} = -\frac{\partial u_i u_j}{\partial x_j} - \frac{\partial p}{\partial x_i} + \alpha g \delta_{i3} T + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + F_i \tag{1a}$$

$$\frac{\partial T}{\partial t} = -\frac{\partial T u_j}{\partial x_j} - u_3 \frac{dT_s}{dx_3} + \kappa \frac{\partial^2 T}{\partial x_j \partial x_j} + F_T \tag{1b}$$

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1c}$$

where t is time, x_i is the spatial vector ($\mathbf{x}$; x_1,x_2,x_3; x,y,z), u_i is the velocity vector ($\mathbf{u}$; u_1,u_2,u_3; u,v,w), p is the pressure perturbation from the hydrostatic balance (normalized by the mean density), T is the perturbation temperature from the horizontal mean temperature ($T_s(x_3)$), which describes the stable background profile), g is the magnitude of the vertical gravitational acceleration (9.8 m s^{-2}), α is the coefficient of thermal expansion (0.025 K^{-1}), ν is the kinematic viscosity (10^{-6} m^2 s^{-1}), and κ is the thermal conductivity (10^{-6} m^2 s^{-1}), δ_{i3} is the Kronecker-delta function and F_i and F_T represent external forcing functions. The degree of stable stratification is characterized by the buoyancy frequency, N, which is equal to

$$N = \left(\alpha g \, \delta_{i3} \frac{dT_s}{dx_i} \right)^{1/2} \tag{2}$$

The domain is a triply-periodic cube with sides of length, L. The boundary conditions are

$$u_i(x_j) = u_i(x_j + nL) \tag{3a}$$

$$T(x_j) = T(x_j + nL) \tag{3b}$$

where n is any integer. This allows spectral numerical techniques to be applied, enabling the spatial terms to be expanded using Fourier series (Orszag, 1971). Periodic boundary conditions are appropriate for the numerical simulation of statistically homogeneous flows (Lesieur, 1987).

The equations of motion are nondimensionalized using the box size, $L/2\pi$, as the length scale, the buoyancy period, N^{-1}, as the time scale, and α^{-1} as the temperature scale. The factor of 2π in the definition of the length scale enables spatial wavenumbers to be defined as integer values. Hereafter, the nondimensionalized forms of the equations of motion will be utilized.

In developing a LES model, the equations of motion are *volume* averaged to define the grid scale motions distinct from the SGS. This filtering operation partitions a fluid variable ($f(\mathbf{x};t)$) into GS ($\bar{f}(\mathbf{x};t)$) and SGS ($f'(\mathbf{x};t)$) components (Leonard, 1974),

$$f(\mathbf{x};t) = \bar{f}(\mathbf{x};t) + f'(\mathbf{x};t) \tag{4}$$

The GS value of $f(\mathbf{x};t)$ is evaluated as its convolution with a spatial filtering function ($G(\mathbf{x})$), or

$$\bar{f}(\mathbf{x};t) = \int_D G(\mathbf{x}-\mathbf{x}')\, f(\mathbf{x}';t)\, d\mathbf{x}' \tag{5}$$

where the integration is taken over the entire domain, D. A boxcar average in physical space is often used for $G(\mathbf{x})$, or

$$G(\mathbf{x}) = \begin{cases} 1 & |\mathbf{x}| < \Delta \\ 0 & |\mathbf{x}| > \Delta \end{cases} \tag{6}$$

where Δ is the physical space grid-scale ($\Delta=L/N_X$; where N_X is the numerical resolution in the x-direction). This boxcar averaging procedure is implicit in the pioneering large-eddy simulations of Deardorff (1970) and Schumann (1975) and is appropriate when the SGS fluxes and stresses are parameterized in physical space. There are many other filtering operators that may be applied in physical space (e.g., Leonard, 1974; Ferziger, 1983). Alternatively, the filtering may be preformed in spectral space, following

$$\bar{f}(\mathbf{k};t) = \int_D G(\mathbf{k}-\mathbf{k}')\, f(\mathbf{k}';t)\, d\mathbf{k}' \tag{7}$$

where $\mathbf{k}$ is the vector wavenumber (k_i; k_1,k_2,k_3) and $G(\mathbf{k})$ is the spectral filtering function which is often a sharp-cut filter ($G(\mathbf{k})=1$ if $|\mathbf{k}|<k_c$; $G(\mathbf{k})=0$, elsewhere; where k_c is the GS cutoff wavenumber=π/Δ). Spectral filtering is appropriate when the SGS parameterization is made in spectral space. Gaussian filtering functions (which are the same in spectral and physical space) may also be used for $G(\mathbf{x})$ (e.g., Mansour *et al.*, 1979; Ferziger, 1983; Piomelli *et al.*, 1989). For the present application, the advantages of these procedures do not appear to outweigh the cost, complexity and possible ambiguities in the interpretation of the resulting simulations (Clark *et al.*, 1979; Ferziger, 1983; Eidson, 1985). However, this may not be true for inhomogeneous flows, such as wall-bounded flows (Ferziger, 1983; Piomelli *et al.*, 1989).

Application of either of the averaging procedures (eqns. 5 or 7) to the nondimensionalized equations of motion gives

$$\frac{\partial \bar{u}_i}{\partial t} = -\frac{\partial \overline{u_i u_j}}{\partial x_j} - \frac{\partial \bar{p}}{\partial x_i} + g\,\delta_{i3}\bar{T} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \bar{F}_i \tag{8a}$$

$$\frac{\partial \bar{T}}{\partial t} = -\frac{\partial \overline{u_j T}}{\partial x_j} - \bar{u}_3 \frac{dT_s}{dx_3} + \kappa \frac{\partial^2 \bar{T}}{\partial x_j \partial x_j} + \bar{F}_T \tag{8b}$$

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \tag{8c}$$

where the summation convention is used. The nonlinear flux terms also need to be partitioned into GS and SGS components.

$$\overline{u_i u_j} = \overline{\bar{u}_i \bar{u}_j} + \overline{u'_i \bar{u}_j} + \overline{\bar{u}_i u'_j} + \overline{u'_i u'_j} = \bar{u}_i \bar{u}_j + \Lambda_{ij} - \tau_{ij} \tag{9a}$$

$$\overline{u_i T} = \overline{\bar{u}_i \bar{T}} + \overline{u'_i \bar{T}} + \overline{\bar{u}_i T'} + \overline{u'_i T'} = \bar{u}_i \bar{T} + \Lambda_{\theta i} - \tau_{\theta i} \tag{9b}$$

where Λ_{ij} is the Leonard momentum flux ($\overline{\bar{u}_i \bar{u}_j} - \bar{u}_i \bar{u}_j$), $\Lambda_{\theta i}$ is the Leonard heat flux ($\overline{\bar{u}_i \bar{T}} - \bar{u}_i \bar{T}$), τ_{ij} is the SGS momentum flux ($\overline{u'_i u'_j} + \overline{u'_i \bar{u}_j} + \overline{\bar{u}_i u'_j}$), and $\tau_{\theta i}$ is the SGS heat flux ($\overline{u'_i T'} + \overline{u'_i \bar{T}} + \overline{\bar{u}_i T'}$). The Leonard fluxes are a consequence of the nested averaging procedure (Leonard, 1974) and will be neglected presently. Leonard fluxes are identically zero when a sharp-cut spectral filter is used and are very small when boxcar averaging in physical space is applied (e.g., Leonard, 1974; Clark *et al.*, 1979; Yoshizawa, 1982; Ferziger, 1983; Eidson, 1985). However, the Leonard fluxes must be considered when a Gaussian filter is used.

Applying the above definitions of the SGS fluxes and writing the GS nonlinear terms in the momentum equation in vorticity form (to insure numerical stability), the GS equations of motion may be expressed as

$$\frac{\partial \bar{u}_i}{\partial t} = \varepsilon_{ijk} \bar{u}_j \bar{\Omega}_k - \frac{\partial \bar{p}}{\partial x_i} - \frac{\partial (\bar{u}_j \bar{u}_j)/2}{\partial x_i} + g \delta_{i3} \bar{T} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} + \bar{F}_i \tag{10a}$$

$$\frac{\partial \bar{T}}{\partial t} = -\frac{\partial \bar{u}_j \bar{T}}{\partial x_j} - \bar{u}_3 \frac{dT_s}{dx_3} + \kappa \frac{\partial^2 \bar{T}}{\partial x_j \partial x_j} - \frac{\partial \tau_{\theta j}}{\partial x_j} + \bar{F}_T \tag{10b}$$

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \tag{10c}$$

where $\bar{\Omega}_i$ is the GS vorticity vector ($\bar{\Omega}_i = \varepsilon_{ijk} \partial \bar{u}_k / \partial x_j$) and ε_{ijk} is the alternating tensor. These equations are then solved, for specific forcing conditions, to give the evolution of the GS velocity and temperature fields. First the SGS momentum and heat fluxes (τ_{ij} and $\tau_{\theta j}$) need to be parameterized in terms of the GS velocity and temperature fields.

SUBGRID SCALE PARAMETERIZATION METHODOLOGIES

Many different methodologies have been employed for parameterizing SGS fluxes in LES models. SGS parameterizations that are applied in physical space range from very simple "production equals dissipation" SGS energy budget approaches (e.g., Smagorinsky, 1963; 1990; Clark *et al.*, 1979; Eidson, 1985; Mason and Derbyshire, 1990; Siegel and Domaradzki, 1991a) to detailed evaluations of the entire SGS energy budget (e.g., Deardorff, 1980; Moeng, 1984; Schmidt and Schumann, 1989; Gallacher, 1990). On the other hand, spectral SGS parameterizations have been developed based upon turbulence closure theory calculations of the rate of energy transfer across the GS cutoff and are

usually applied in wavenumber space (e.g., Kraichnan, 1976; Chollet and Lesieur, 1981; Lesieur, 1987; Domaradzki *et al.*, 1987; Yakhot *et al.*, 1989). Unfortunately, detailed intercomparison studies of these SGS parameterization methods are lacking for stably-stratified turbulent flows (see Schumann, 1991 for a recent exception).

Here, we discuss and apply two different SGS parameterization methods, the physically-based Smagorinsky SGS eddy viscosity and the spectral Chollet and Lesieur (1981; hereafter CL81) eddy viscosity. Preliminary comparisons that we have made indicate that the evolution of the GS energetics is essentially the same for the two different SGS eddy viscosities. However, the GS energy spectra produced using the CL81 spectral eddy viscosity reflect spectral levels and slopes closer to theoretical Kolmogorov "5/3" spectra than do those produced using the Smagorinsky eddy viscosity (fig. 1). Before our LES results are discussed, it is important to address how SGS parameterizations are made.

The SGS fluxes and stresses are most often represented in flux-gradient form, or

$$\tau_{ij} = - K_{SGS} \bar{S}_{ij} + \frac{\tau_{kk}}{3} \delta_{ij} \qquad (11a)$$

$$\tau_{\theta i} = - \frac{K_{SGS}}{Pr_t} \frac{\partial \bar{T}}{\partial x_i} \qquad (11b)$$

where K_{SGS} is the SGS eddy viscosity (a function of x and t), $\bar{S}_{ij}$ is the GS rate of strain tensor $(\partial \bar{u}_i/\partial x_j + \partial \bar{u}_j/\partial x_i)$, and Pr_t is the SGS turbulent Prandtl number (the ratio of the SGS eddy viscosity to the SGS eddy diffusivity). The final equations of motion are

$$\frac{\partial \bar{u}_i}{\partial t} = \varepsilon_{ijk} \bar{u}_j \bar{\Omega}_k - \frac{\partial \bar{\pi}}{\partial x_i} + g \delta_{i3} \bar{T} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \frac{\partial}{\partial x_j}\left(K_{SGS} \bar{S}_{ij}\right) + \bar{F}_i \qquad (12a)$$

$$\frac{\partial \bar{T}}{\partial t} = - \frac{\partial \bar{u}_j \bar{T}}{\partial x_j} - \bar{u}_3 \frac{dT_s}{dx_3} + \kappa \frac{\partial^2 \bar{T}}{\partial x_j \partial x_j} + \frac{\partial}{\partial x_j}\left(\frac{K_{SGS}}{Pr_t} \frac{\partial \bar{T}}{\partial x_j}\right) + \bar{F}_T \qquad (12b)$$

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0 \qquad (12c)$$

where $\bar{\pi}$ is the pressure head which includes the GS $(\bar{u}_k \bar{u}_k/2)$ and SGS $(\tau_{kk}/3)$ normal stresses.

Once the SGS eddy viscosity and forcing functions are specified, this system of equations can be solved directly for the evolution of the GS flow. We employ pseudo-spectral numerical techniques, based upon Fourier series expansions, to solve for the spatial terms of the equation of motion and advance time using the leap-frog method (e.g., Orszag, 1971; Siegel, 1988; Canuto *et al.*, 1988). These methods are highly accurate and are standard in numerical turbulence investigations. A complete description of the numerical methods used may be found in Siegel (1988).

THE SMAGORINSKY SGS EDDY VISCOSITY

The Smagorinsky SGS eddy viscosity is derived assuming that K_{SGS} depends only upon the smallest resolvable scale (Δ) and the rate at which KE is transferred across the GS cutoff (Δ). For the inertial subrange of turbulence, the spectral transfer of KE is equal to the KE dissipation rate, ε_{KE}. Applying only dimensional arguments, K_{SGS} is given by

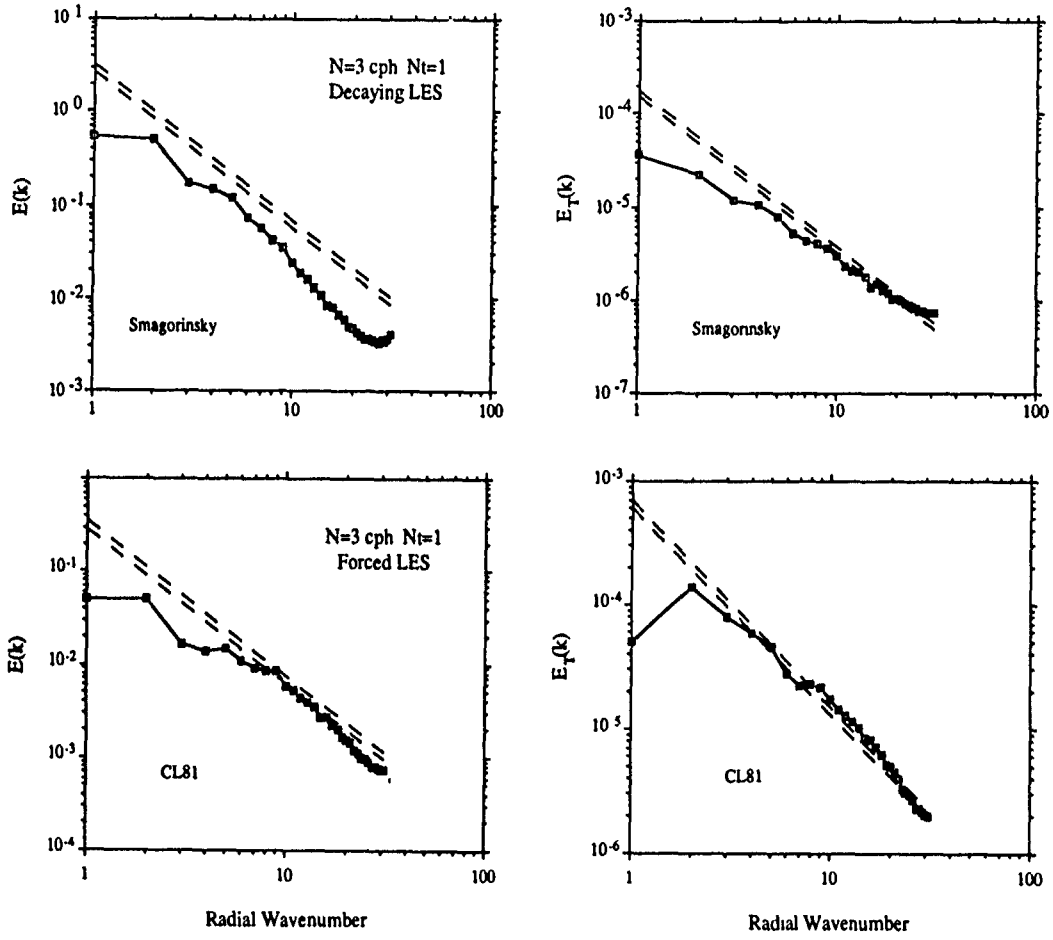


Figure 1: Comparison of radial wavenumber kinetic energy and temperature variance spectra calculated using the Smagorinsky SGS eddy viscosity (upper) and the CL81 eddy viscosity (lower). The dotted lines are theoretical inertial subrange Kolmogorov spectra for KE ($E(k) = \alpha_K \varepsilon^{2/3} k^{-5/3}$) and for temperature variance ($E_T(k) = \beta_K \chi \varepsilon^{-1/3} k^{-5/3}$, where χ is the temperature variance dissipation rate). The two lines express the magnitude of experimental uncertainty in the values of the Kolmogorov constants, α_K and β_K.

$$K_{SGS} = (c_s \Delta)^{4/3} \varepsilon_{KE}^{1/3} \tag{13}$$

where the constant of proportionality (c_s) is referred to as the Smagorinsky constant. To evaluate the value of K_{SGS} at any point in space, the energy transfer rate (ε_{KE}) must be known at that location. In general, ε_{KE} depends on all scales of motion. However, the SGS fluxes must be parameterized in terms of the GS motions alone. For a statistically homogeneous flow, such within the turbulent inertial subrange, the dissipation of KE is equal to its production, or

$$\varepsilon_{KE} = -\frac{1}{2} \bar{S}_{ij} \tau_{ij} + g \delta_{i3} \tau_{\theta i} \tag{14}$$

Here, the dissipation of KE is balanced by the turbulent production by the GS shear ($\bar{S}_{ij}$) and by the

SGS buoyancy flux ($g\,\delta_{i3}\,\tau_{\theta i}$). Applying the flux-gradient forms of the SGS fluxes (eq. 11) results in

$$\varepsilon_{KE} = \frac{1}{2} K_{SGS}\,\bar{S}^2 - \frac{g}{Pr_t} K_{SGS} \frac{\partial \bar{T}}{\partial x_3} \tag{15}$$

where $\bar{S}^2$ is the mean square rate of strain ($=\bar{S}_{ij}\bar{S}_{ij}$). Finally, K_{SGS} solved for by substituting the scaling relation for ε_{KE} (eq. 14) into the SGS kinetic energy balance (eq. 15), or

$$K_{SGS} = \frac{(c_s\,\Delta)^2}{2^{1/2}} \left(\bar{S}^2 - \frac{2g}{Pr_t} \frac{\partial \bar{T}}{\partial x_3} \right)^{1/2} \qquad \bar{S}^2 > \frac{2g}{Pr_t} \frac{\partial \bar{T}}{\partial x_3} \tag{16}$$

The SGS eddy viscosity is set to zero if $\bar{S}^2$ is less than $2g/Pr_t \partial\bar{T}/\partial x_3$. However, this SGS "critical Richardson number" condition did not occur during our simulations. Further, the inclusion of SGS buoyancy production did not significantly affect the resulting values of K_{SGS} (Siegel, 1988) consistent with the assumption that Δ lies within the inertial subrange of turbulence.

The choices of values for c_s and Pr_t were made primarily by comparing simulated GS spectral energy distributions with theoretical Kolmogorov spectra for high Reynolds number turbulence, although a thorough literature survey was also made. The best results from this comparison are shown in figure 1. Six numerical experiments were performed to examine the degree of correspondence between the LES determined and theoretical radial wavenumber spectra and its relationship with variations in the values of c_s and Pr_t (Siegel, 1988). The final values of $c_s = 0.15$ and $Pr_t = 1.0$ best represented the theoretical spectra, especially for the temperature variance spectra (fig. 1). However, the LES KE spectra underestimate the Kolmogorov spectra by a factor of roughly five and exhibit a sharp "roll-up" as the GS cutoff wavenumber is approached (k_c=31). This "roll-up" indicates that the Smagorinsky SGS eddy viscosity is not adequately transferring KE across the GS cutoff and that there KE is "piling-up".

Spectral distributions of the magnitude of K_{SGS}, as well as the energy transfers due to the parameterized SGS fluxes, were also examined (Siegel and Domaradzki, 1991b). A "cartooned" representation of the spectral distribution of K_{SGS} (normalized by $(E(k_c)/k_c)^{1/2}$; where $E(k_c)$ is the value of the KE spectrum at the cutoff wavenumber, k_c) is shown in figure 2. The normalized K_{SGS} is to first order spectrally uniform, isotropically distributed and invariant in time. Similarly, SGS spectral energy transfers are white, indicating a uniform loss of energy at all scales. These results suggest that the Smagorinsky SGS eddy viscosity may be replaced by an isotropic, spectral SGS parameterization that scales as $(E(k_c)/k_c)^{1/2}$. A detailed analysis of the Smagorinsky SGS eddy viscosity and its energy transfers is presented in Siegel and Domaradzki (1991b).

THE CHOLLET AND LESIEUR (1981) SPECTRAL EDDY VISCOSITY

The Chollet and Lesieur (1981; CL81) spectral eddy viscosity has been derived using the Eddy Damped Quasi-Normal Markovian (EDQNM) turbulence closure theory calculation of the energy transfers across the GS cutoff (Chollet and Lesieur, 1981; Chollet, 1984; Lesieur, 1987). SGS eddy viscosities derived from alternative spectral closure formulations (Kraichnan, 1976), as well as from the results of direct isotropic turbulence simulations (Domaradzki *et al.*, 1987), are similar. The CL81

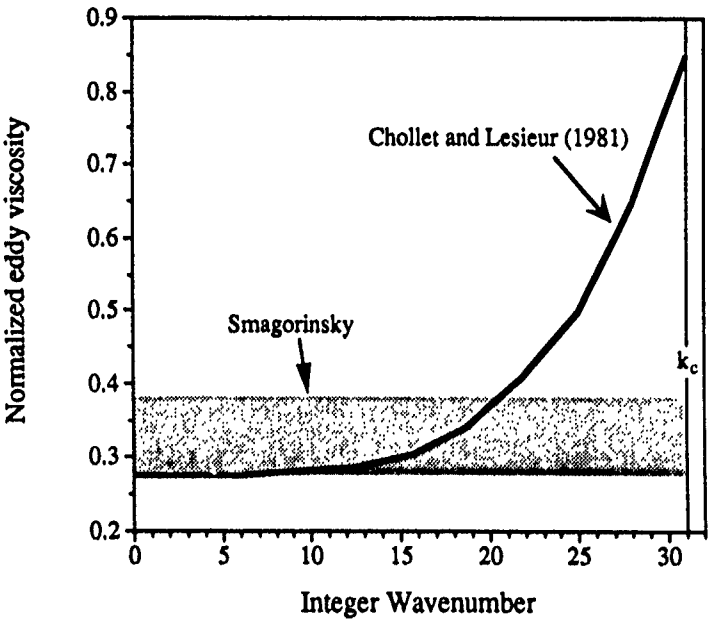


Figure 2: Spectral distribution of the CL81 eddy viscosity (Chollet and Lesieur, 1981) and the spectral distribution of the Smagorinsky SGS eddy viscosity calculated from the decaying simulations of stratified turbulence (based upon data from Siegel and Domaradzki, 1991b).

eddy viscosity has been used by Lesieur and his collaborators for examining stably-stratified turbulence (Métais, 1985; Lesieur, 1987; Lesieur *et al.*, 1988).

The CL81 eddy viscosity, $\nu_e(\mathbf{k})$, is formulated by the EDQNM evaluation of the spectral energy transfer across the GS cutoff (Δ) due to triad interactions which are both local (due to wavenumbers near the cutoff wavenumber, k_c) and nonlocal (from spectrally distant wavenumbers). The form of the CL81 spectral eddy viscosity (normalized by $(E(k_c)/k_c)^{1/2}$) is shown in figure 2. For low radial wavenumbers ($k<10$), the normalized CL81 eddy viscosity is equal to ~0.28 while it increases rapidly as the GS cutoff wavenumber is approached. This high wavenumber spectral cusp indicates the importance of local interactions to the spectral transfer of energy across the GS cutoff.

The CL81 SGS eddy viscosity is applied in spectral space where the term in (12) containing K_{SGS} is replaced with the CL81 spectral eddy viscosity ($\nu_e(\mathbf{k})$),

$$\frac{\partial}{\partial x_j}\left(K_{SGS}\,\bar{S}_{ij}\right) \quad \Rightarrow \quad -\nu_e(\mathbf{k})\,k^2\,\bar{u}_i \tag{17}$$

where $\nu_e(\mathbf{k})$ is equal to the normalized SGS eddy viscosity shown in figure 2 multiplied by the scaling factor, $((E(k_c)/k_c)^{1/2})$. A similar transformation is made for the heat equation. The CL81 spectral eddy viscosity ($\nu_e(\mathbf{k})$) and diffusivity ($\kappa_e(\mathbf{k})$) are related by a constant SGS Prandtl number (Pr_t=0.6; Chollet, 1984). Note that the normalized value of the Smagorinsky SGS eddy viscosity spectra, although uniform, is greater than the normalized CL81 viscosity at low wavenumbers and less at higher wavenumbers (fig. 2). In a sense, the Smagorinsky SGS eddy viscosity accounts only for the

nonlocal turbulent interactions represented by the low wavenumber plateau region of the CL81 SGS viscosity, but at a level higher than the CL81 plateau to account for its inefficiency of local energy transfers (Lesieur, 1987; Siegel and Domaradzki, 1991b).

The GS KE spectra calculated using the CL81 eddy viscosity are quantitatively similar to the theoretical Kolmogorov energy spectra (fig. 1) for the higher simulated wavenumbers (k>10). This subjective measure of SGS parameterization performance indicates that the CL81 eddy viscosity clearly outperforms the Smagorinsky SGS eddy viscosity. Another indication is the lack of a "roll-up" in the CL81 energy spectra near the cutoff wavenumber, indicating that the CL81 cusp is effective at removing energy near k_c. In addition to an increase in performance, the use of the CL81 eddy viscosity represents a substantial improvement in computational efficiency (about a factor of 2). This speedup comes from the fact that CL81 SGS parameterization does not require that $\partial(K_{SGS}\,\overline{S}_{ij})/\partial x_j$ be evaluated at each time step, which eliminates more than half of the 3-D fast Fourier transforms required to implement the Smagorinsky SGS parameterization.

LES EXPERIMENTS

We present the results of two LES experiments to investigate the dynamics and kinematics of oceanic internal gravity wave fields. The first experiment is the study of the decay of stably-stratified turbulence at oceanographically relevant length and time scales (Siegel, 1988; Siegel and Domaradzki, 1991a; 1991b). The initial wavenumber distribution is consistent with an extrapolation of the Garrett-Munk (GM) internal wave spectrum to scales smaller than 10 m. This experiment clearly shows many features of the so-called "turbulent collapse" (the transition from a fully turbulent flow to an internal wave dominated field). However, the present results illustrate several important differences from previous laboratory and *direct* numerical results. In particular, highly anisotropic, "pancake" structures are not formed coincident with the onset of the turbulent collapse. We speculate that these differences may be attributed to the extreme differences in Reynolds numbers for the two cases.

The second experiments are the preliminary investigations of the use of LES techniques to directly determine eddy diffusivities for a steady GM ocean (Siegel, 1991). Here, the lowest resolved wave modes are forced to maintain constant energy, consistent with GM amplitudes while the phases of the forced waves are determined dynamically. The long time integration (≥30 Nt) and steady forcing enables time/space mean values of the buoyancy flux and hence, the vertical eddy diffusivity (K_{BF}) to be directly determined. Values of vertical eddy diffusivity averaged over long time (≥20 Nt) and large space (125,000 m^3) scales are appropriate for use in oceanic circulation models. The direct K_{BF} determinations are compared with the results of several indirect turbulent energy dissipation rate methods (Osborn and Cox, 1972; Osborn, 1980). In addition, the dependency of K_{BF} upon N, the buoyancy frequency, is also investigated. These LES results are still preliminary as further experiments and detailed analyses are required. However, this successful "proof of concept" experiment suggests that LES techniques, when explicitly forced to *in situ* observations, may be a useful tool for examining mixing processes within the thermocline.

LES OF DECAYING OCEAN TURBULENCE

The first LES experiment to be discussed is the study of the decay of stratified turbulence at oceanographically relevant space and time scales. The temporal evolution and spatial structure of the decaying stratified turbulent flow are investigated where the initial conditions are represented by an extrapolation of the GM spectrum. These results will be addressed briefly here as they are discussed in detail in Siegel and Domaradzki (1991a) and (1991b).

The modeling domain (L) is chosen to be 10 m and the numerical resolution is 64^3. This results in a GS cutoff (Δ) equal to 15.6 cm. The choice of L (=10 m) was made to insure that the GS cutoff lies within the inertial subrange of turbulence. The separation between GS and SGS motions is done with an implicit boxcar filter and the Smagorinsky SGS eddy viscosity is used. Three simulations were made with buoyancy frequencies of 1, 3 and 10 cph and the equations of motion were integrated forward in time for 10 units of buoyancy time (Nt). Each experiment required 6.8 CPU hours of Cray XMP/48 time. For brevity, only the 1 cph case will be shown.

The initial conditions are based upon an extrapolation of the GM internal wave spectrum to scales less than 10 m. Of course the GM spectrum was not developed to correctly predict the variance at these scales. However, there does not exist a unified 3-D velocity and temperature finestructure spectrum that one could use to initialize the LES model, although several advances towards this goal have been made recently (e.g., Kunze *et al.*, 1990; M.C. Gregg, this volume). As it is recognized that the GM spectrum should not hold for scales less than 10 m, there is little reason to believe that the phases of the individual internal wave modes will follow linear theory either (e.g., Holloway, 1983; Müller *et al.*, 1986; Shen and Holloway, 1986). Hence, we randomized not only the internal wave mode phases, but the phase of each *component* of each wave mode. This will obviously result in an initial flow field that is highly unstable and its initial decay will be rapid and turbulent. The use of a fully random initial field is consistent with numerical turbulence procedures (e.g., Orszag and Patterson, 1972; Rogallo and Moin, 1984; Métais and Herring, 1989). Further details concerning the initialization procedure and its consequences may be found in Siegel and Domaradzki (1991a).

The temporal evolution of the domain averaged energetics and length scales is shown in figure 3. All of the component energy levels decay rapidly as expected. Rapid exchanges of vertical kinetic energy (VKE) and potential energy (PE) are observed during the decay's latter stages. These component energy exchanges are driven by internal gravity wave motions which give rise to a reversible buoyancy flux (BF). However, the internal waves cannot be linear as the domain-averaged BF is nonzero (Stewart, 1969). Evaluation of SGS energetics, fluxes and dissipation indicates that SGS processes regulated energy dissipation, but made negligibly small contributions to the total energetics and fluxes, consistent with the LES assumptions.

The evolution of the the Ozmidov (L_o) and the vertical integral (h) length scales provide additional insights into the state of decaying stratified turbulence. The Ozmidov scale is equal to

$$L_o = \left(\frac{\varepsilon}{N^3}\right)^{1/2} \tag{18a}$$

which represents the vertical scale where buoyancy and inertial forces equally influence the evolution of vertical momentum. That is, vertical scales larger than L_o will be affected by the stable

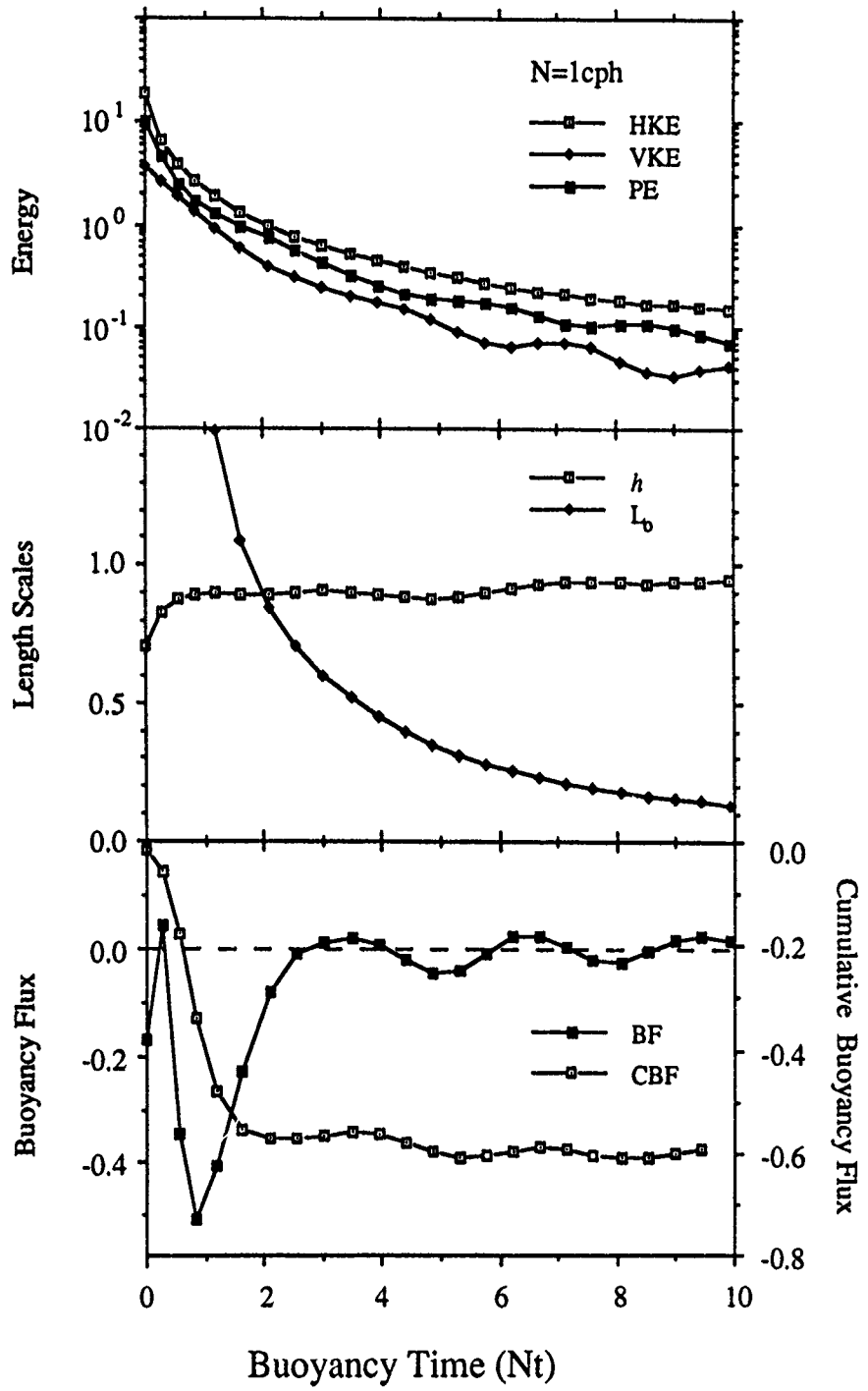


Figure 3: Temporal evolution of the domain averaged HKE, VKE and PE (upper panel), the vertical energy containing length scale, h, and the Ozmidov scale, L_o, (middle panel), and the buoyancy flux, BF, and cumulative buoyancy flux, CBF (lower panel). The mean stratification is 1 cph and the experiment is conducted for 10 buoyancy periods, Nt. ((from Siegel and Domaradzki, 1991a).

stratification. Whereas for scales smaller than L_o, these eddies will remain unaffected. Integral length scales are calculated by evaluating the most probable length scale from the kinetic energy spectra, or

$$h = \frac{\pi}{6}\left(\sum_{k_3=1}^{N/2} \hat{u}_i(k_3)\hat{u}_i(-k_3)/|k_3|\right)\left(\sum_{k_3=1}^{N/2} \hat{u}_i(k_3)\hat{u}_i(-k_3)\right)^{-1} \qquad (18b)$$

where $\hat{u}_i(k_j)$ here represents Fourier amplitudes of the i^{th} component of the velocity field in the j^{th} wavenumber direction and index summation is assumed. This definition for h is somewhat arbitrary as many other choices could have been made (Métais and Herring, 1989). It should be noted that h is related to the vertical extent of the kinetic energy structures and makes no distinction whether these structures are turbulent eddies or internal waves.

The vertical energy-containing scale (h) remains nearly constant while the Ozmidov scale (L_o) decreases rapidly with time (fig. 3). Initially, L_o is much larger than h indicating that the vertical motions of the energy-containing eddies are regulated primarily by inertial, or nonlinear, processes. However as L_o decreases, buoyancy forces become more important. This transition occurs at $Nt \cong 2$, indicating that the Froude number ($Fr \equiv (L_o/h)^{2/3}$) is equal to one. Note that this dynamical transition does not manifest itself in the evolution of h while synchronous oscillations in VKE and PE are most apparent after the transition. This behavior is also observed with different stratification intensities (Siegel and Domaradzki, 1991a).

Evidence of the turbulent collapse is most apparent in the temporal evolution of the buoyancy flux (BF; fig. 3). When BF is negative, VKE is converted to PE, effectively raising the fluid's center of mass. If this conversion happens irreversibly, the fluid's center of mass is permanently raised indicating that diapycnal mixing has occurred. A reversible BF shows up as internal wave motions. Irreversible aspects of BF are best illustrated by examining the cumulative buoyancy flux (CBF), or

$$CBF(t) = \int_0^t BF(\tau)\, d\tau \qquad (19)$$

which is shown in the bottom panel of figure 3. Early in its evolution ($Nt<2$), values of CBF decrease monotonically. After this period, CBF stops accumulating negative values and begins to oscillates in time. The transition from an irreversible accumulation of PE to reversible BF oscillations indicates that the initially turbulent flow evolves rapidly into a field of nonlinear internal waves.

The transition from a fully turbulent to a buoyancy dominated flow may be visualized by examining the temporal evolution of vertical slices of the absolute temperature ($\bar{T}(x;t) + T_s(z)$). An example of the temporal evolution in the absolute temperature is shown in figure 4 for Nt of 1, 2, 3 and 4. These spatial distributions give the appearance of a turbulent flow gradually evolving into an internal wave-dominated flow, with the general features described above. However, this transition does not appear to occur suddenly and these data cannot be used to visually determine the transition time. We have also analyzed 3-D topological structure of isotherm distributions which also provide useful visual evidence of the collapse (color plates may be found in Vasilopoulos, 1991).

The term "turbulent collapse" refers to the process of decay of stably-stratified, fully developed turbulence to a state where the vertical turbulent scales are suppressed by the background stratification (e.g., Dickey and Mellor, 1980; Stillinger *et al.*, 1982; Itsweire *et al.*, 1986; Lesieur, 1987; Hopfinger, 1987). This vertical scale suppression is thought to cause the initially 3-D eddies to

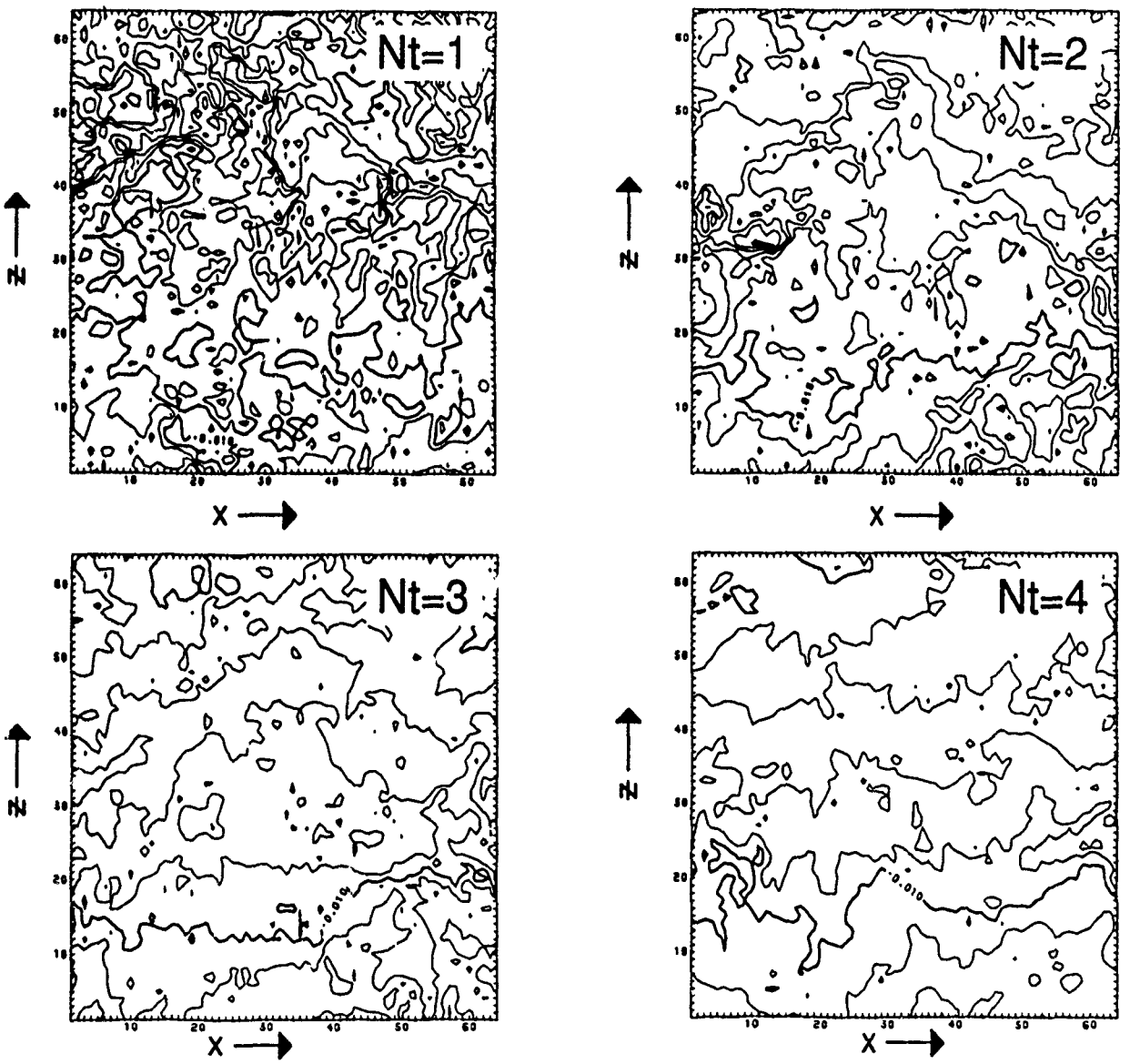


Figure 4: Temporal evolution of the vertical-horizontal sections of the absolute temperature ($T_s(z)$ + $\bar{T}(x,t)$) for buoyancy periods of 1, 2, 3, and 4. The mean stratification is 1 cph. (from Siegel and Domaradzki, 1991a).

"collapse" into nearly horizontal "pancake-like" vortices where their vertical scale is limited by L_o. Observations of these "pancake-like" eddies have been made primarily from laboratory flow visualizations (e.g., Lin and Pao, 1979; Hopfinger, 1987; Browand *et al.*, 1987), while reductions in energy decay rates and/or a vanishing buoyancy flux are often used as proxies of the collapse's occurrence (Dickey and Mellor, 1980; Stillinger *et al.*, 1982; Itsweire *et al.*, 1986; Hopfinger, 1987).

The LES results appear to give conflicting information about the occurrence of the turbulent collapse. For example, energy decay rates do not decrease significantly coincident with the onset of the turbulent collapse. Further, the size of the vertical integral scale (h) remains unaffected by the collapse, roughly

the same size as the horizontal energy-containing scale (l; not shown; Siegel and Domaradzki, 1991a). These results are unexpected as classical length scale phenomenology predicts that the vertical length scale of the *turbulent* eddies, and their KE, should decrease in a manner similar to the Ozmidov scale. Thus, we are not seeing a "pancaking" of the energy-containing eddies coincident with the collapse. However, the change in the nature of the BF indicates that the collapse has occurred (fig. 3).

The invariance of integral length scales implies that the dynamics of these energy-containing motions must have evolved from a turbulent state to one where buoyancy forces are important while retaining their size and isotropy. This lack of "pancaking" may be examined by addressing the evolution of the total vertical energy (TVE), the sum of VKE and PE. As the buoyancy flux acts only to exchange VKE and PE, BF will have absolutely no influence upon TVE. This means that vertical length scales characteristic of TVE will remain constant during the collapse. Similarly, the nonlinear energy cascades of TVE and hence, its energy decay rate will remain unaffected by the collapse. Thus, the onset of the collapse only indicates that the evolution of PE and VKE will be intertwined by a reversible BF and energy decay rates, as well as integral length scales, should remain unaffected by this dynamic transition This will remain true as long as molecular dissipation processes are *not* important for the evolution of the energy-containing motions ($Re_\lambda >> 1$).

Here, we suggest that this apparent lack of correspondence between the LES results and previous laboratory and *direct* numerical experiments is caused by extreme differences in the Reynolds numbers for the two cases. As described earlier, Reynolds numbers based upon Taylor length scales (Re_λ) corresponding to the oceanic internal wave field are $O(10^5)$ consistent with the LES results. However, laboratory or *direct* numerical experiments rarely give *initial* Re_λ values as great as 100. It should not be expected that these low Re_λ experiments can realistically simulate the interactions found within the oceanic internal wave field as molecular processes will have far too important roles in the evolution of the energy-containing eddies.

This Reynolds number dependency of the turbulent collapse is consistent with the results of a recent scaling analysis by Gargett (1988). Gargett's scaling indicates that for a high Reynolds number flow characteristic length scales will be isotropically distributed, whereas for a low Re_λ flow the vertical scales would be restricted. Her scaling holds only when the Froude number is O(1) and hence, only at the onset of the collapse. However, it is not clear whether the assumptions used in Gargett's scaling are valid (e.g., Van Atta, 1990; Yamazaki, 1990).

Unfortunately, we cannot state with certainty that this observed lack of "pancaking" is caused by the large LES-simulated Reynolds numbers. This is because there are many questions concerning the roles of the SGS method and the initial conditions upon the LES simulated flow field. One should note that the evolution of the buoyancy flux (a second order moment) is predicted consistent with the collapse onset while the spatial energy distributions (i.e., first order statistics) are clearly inconsistent with previous low-Re observations. It seems odd that the choice of the SGS parameterization, even an isotropic one like the Smagorinsky SGS eddy viscosity, would "correctly" predict the BF evolution, but not the flow energetics. Other SGS methods (CL81) which we have tried give nearly identical results. We believe that the reason for this is that the simulated high wavenumber motions are within the inertial subrange of turbulence and are locally isotropic (Siegel and Domaradzki, 1991b). This is consistent with the basic tenets of the large-eddy approximation described above. Therefore, the

application of an isotropic SGS eddy viscosity to parameterize the SGS energy fluxes in stratified turbulence may be appropriate.

Similarly, it seems unlikely that the initial conditions have much influence upon the final evolution of the flow field. We have initialized the LES experiments with a highly anisotropic energy distribution that rapidly becomes isotropic. As discussed above, this rapid reorganization of spectral energy is caused by the choice of random phases among wave components. Thus, it is unlikely that the initial anisotropic flow is controlling much of the flow's evolution. Further arguments supporting our conjecture of the importance of Reynolds numbers in the evolution of stratified turbulence is beyond the scope of this contribution. However, detailed experiments and analyses are required before the turbulent collapse can be considered to be an "understood" phenomena, particularly for high Reynolds number flows.

We suggest that the absence of "pancaking" in our LES results is due to the large simulated Reynolds numbers compared with previous laboratory and *direct* numerical experiments. If our LES results are truly representing the turbulent collapse, this suggests that the fundamental nature of decaying stably-stratified turbulence may be different at oceanographically relevant space and time scales compared with previous laboratory observations. This indicates that one should be cautious when applying low Reynolds number results to the oceanic internal wave field.

LES OF STEADY THERMOCLINE MOTIONS

Here, we illustrate preliminary results of a LES model of a steady GM ocean for directly determining values of the vertical eddy diffusivity. The quasi-steady forcing and long time integration enables values of the buoyancy flux (<BF>, where <•> represents a time-space average) and hence, values of the vertical eddy diffusivity (K_{BF}) to be calculated over long time scales ($\geq$ 20 Nt). The LES calculated K_{BF} values are determined over many hours and spatial scales of 50 m and are representative of vertical eddy diffusivities appropriate for use in oceanic general circulation models. Please note that these results are preliminary.

The GS equations of motion are numerically forced to maintain constant energy, consistent with GM spectral amplitudes, for the 8 wave modes where $|\mathbf{k}| = (1,1,1)$ (that is, for $\mathbf{k}=(1,1,1)$, $\mathbf{k}=(1,1,-1)$, $\mathbf{k}=(1,-1,1)$, etc.). The forcing amplitude is based upon the GM internal wave spectrum while the phases are determined by solving the equations of motion. This forcing method has been used in the numerical simulation of steady isotropic turbulence (Siggia and Patterson, 1978; Kerr, 1985). It has been shown that the nature of the forcing method (i.e., VKE to PE ratio, spectral weighting of the forcing function, etc.) can have significant impact on the simulated flow fields (e.g., Shen and Holloway, 1986; Holloway, 1988; Holloway and Ramsden, 1988). By selecting *a priori* the forcing wavenumber ($|\mathbf{k}| = (1,1,1)$), the ratio of VKE to PE (2/3) (and PE to HKE; 1/2) for the forced waves is fixed by the GM spectrum and is independent of the buoyancy frequency. Whether this choice of forcing is representative of the oceanic internal wave field, as well as its role in determining K_{BF}, is unknown at this time.

The domain is a triply periodic cube with dimensions of 50 m on a side. The numerical resolution is 96^3 resulting in a GS cutoff (Δ) of 52 cm. The Chollet and Lesieur (1981; CL81) spectral eddy viscosity is used to parameterize the SGS processes. Near-perfect correspondence is observed

between the theoretical Kolmogorov spectra and the LES calculated energy spectra suggesting that the CL81 SGS eddy viscosity is performing adequately (fig. 1). The flow field is initialized as in the decaying experiments and is numerically forced as described above. Three mean stratifications (1, 3 and 10 cph) are used and the GS equations of motion are integrated for 30 buoyancy periods (Nt). Statistical quantities (time/space means; $< \bullet >$) are calculated for the last 20 Nt of the simulations when the flow is nearly stationary. Statistical quantities are re-cast into dimensional units for ease in comparison with field observations. Each experiment requires about 12 CPU hours of Cray YMP 8/864 time (~20 CPU minutes per Nt).

Temporal evolution of the energetics and buoyancy flux are shown in figure 5 for the N = 3 cph experiment (the other experiments are qualitatively similar). A high degree of temporal reorganization is occurring while the energy levels are generally decreasing in time. This is due primarily to the energy variations of the zero wavenumber components, although there is some evidence of small changes in spectral energy levels after Nt = 10. As before with the decaying simulations, significant oscillations in VKE and PE are observed driven by a reversible buoyancy flux. However, the BF oscillations are not strictly periodic, indicating the existence of several internal wave modes.

Here, our goal is to use the results from the quasi-steady LES experiments to directly determine relevant mixing parameters. Space/time averages for kinetic and potential energy dissipation rates ($<\varepsilon_{KE}>$ and $<\varepsilon_{PE}>$) are given in Table 1. Values of $<\varepsilon_{KE}>$ range from 1.0×10^{-8} to 5.1×10^{-7} $m^2 s^{-3}$ consistent (although a bit high) with observed mean values (e.g., Gregg, 1987; 1989; Gregg and Sanford, 1988). The variation of the time/space mean values of $<\varepsilon_{KE}>$ and $<\varepsilon_{PE}>$ with buoyancy frequency, N, is shown in figure 6. Both $<\varepsilon_{KE}>$ and $<\varepsilon_{PE}>$ show a $N^{+7/4}$ dependency similar to the scaling results of Gargett and Holloway (1984) and the ε_{KE} parameterization of Gregg (1989). A direct comparison of the LES results with the Gregg's (1989) parameterization is presently underway. The ratio of $<\varepsilon_{PE}>$ to $<\varepsilon_{KE}>$ defines the mixing efficiency, η. Values of η are equal to 0.3 independent of N (Table 1), fairly consistent with *in situ* observations (e.g., Oakey, 1982).

Similarly, the vertical eddy diffusivity, K_{BF}, can be directly determined knowing the space/time averaged buoyancy flux, <BF>, or

$$K_{BF} \equiv - \frac{<BF>}{N^2} \qquad (20a)$$

The LES-calculated K_{BF} values all are positive indicating that on the average there is a down gradient flux of heat. Specifically, values of K_{BF} range from 6.0×10^{-6} to 6.7×10^{-4} $m^2 s^{-1}$ (Table 1) for the stratifications investigated and are generally consistent with the "abyssal recipes" value of 10^{-4} $m^2 s^{-1}$

TABLE 1: SPACE/TIME MEAN QUANTITIES FOR THE GM FORCED LES EXPERIMENTS

N (cph)	BF (m^2/s^3)	ε_{KE} (m^2/s^3)	ε_{PE} (m^2/s^3)	η	K_{BF} (m^2/s)	K_ε (m^2/s)	K_χ (m^2/s)
1	1.83×10^{-11}	1.01×10^{-8}	3.02×10^{-9}	0.30	6.01×10^{-6}	6.65×10^{-4}	9.93×10^{-4}
3	2.33×10^{-9}	5.77×10^{-8}	1.74×10^{-8}	0.30	8.48×10^{-5}	4.20×10^{-4}	6.35×10^{-4}
10	2.05×10^{-7}	5.11×10^{-7}	1.65×10^{-7}	0.32	6.73×10^{-4}	3.36×10^{-4}	5.43×10^{-4}

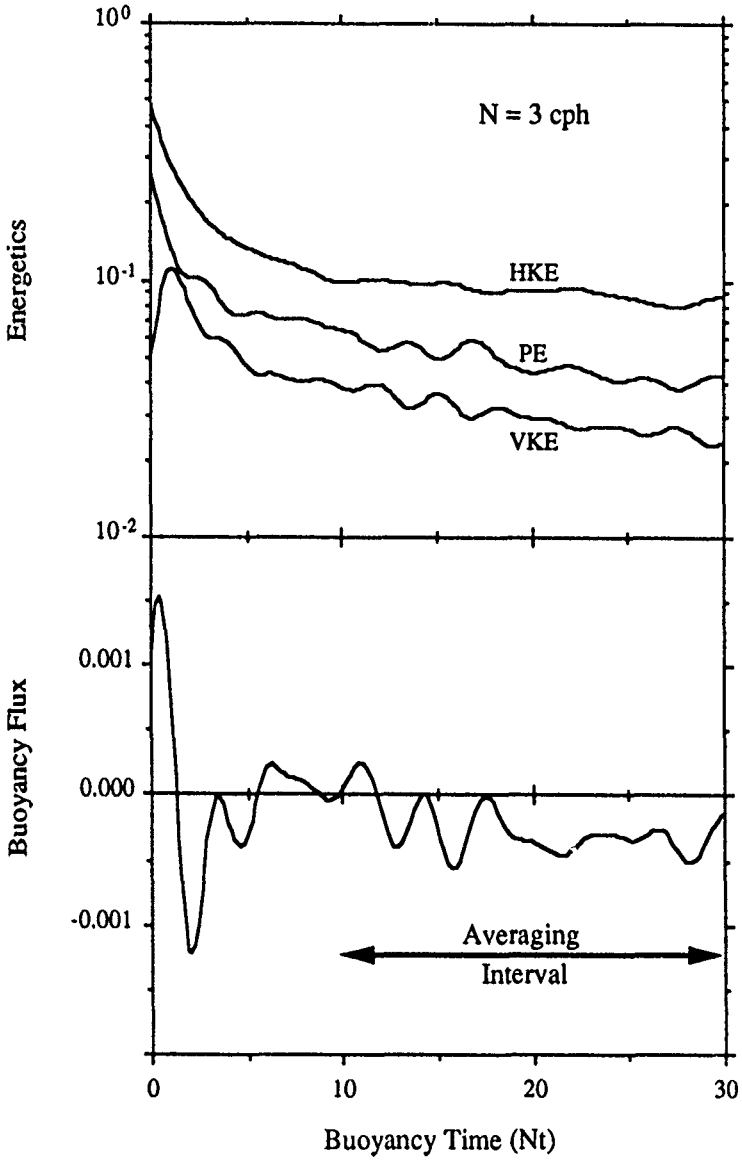


Figure 5: Temporal evolution of energetic quantities (HKE,VKE and PE), as well as the buoyancy flux (BF), for the forced GM LES experiment (N=3 cph). Averaging for the statistical calculations is made only for the last 20 Nt.

(Munk, 1966). However, values of K_{BF} increase dramatically with N (going as N^{+2}; fig. 7). At this point, we do not know exactly what is causing this anomalous dependency on N. However, it seems likely that the nature of the numerical forcing can have an undue influence upon K_{BF}. Indications of this are apparent when one compares the space/time averaged <BF> and $<\varepsilon_{PE}>$ for the N=10 cph experiment (Table 1). For this experiment, values of <BF> are greater than $<\varepsilon_{PE}>$ indicating that more energy is being transferred into PE from VKE than is being removed by diffusive processes. For a simplified "production (+ forcing) equals dissipation" PE budget to hold, there must be a mechanism to transfer this energy somewhere else. We think that the "culprit" may be the forcing mechanism

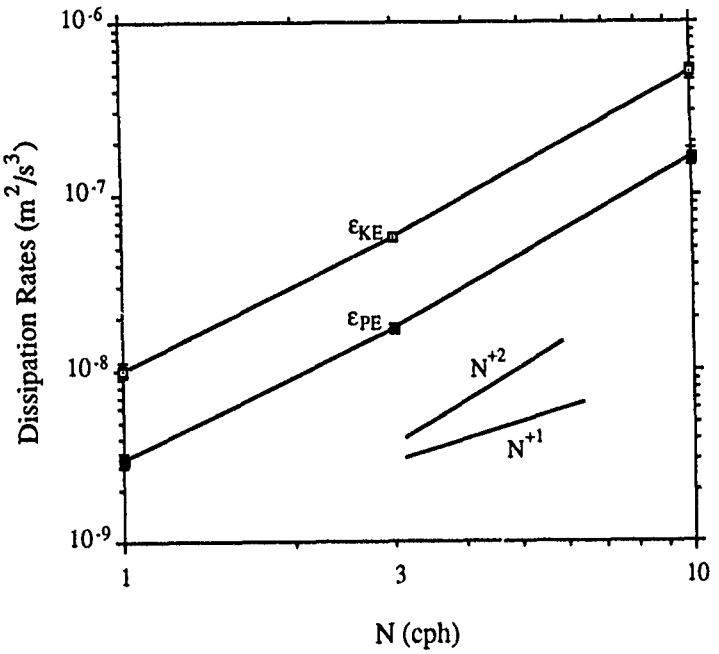


Figure 6: Dependency of space/time mean values of $\langle \varepsilon_{KE} \rangle$ and $\langle \varepsilon_{PE} \rangle$ upon buoyancy frequency, N. Standard deviations are shown by the error bars. Spectral slopes of N^{-1} and N^{-2} are shown.

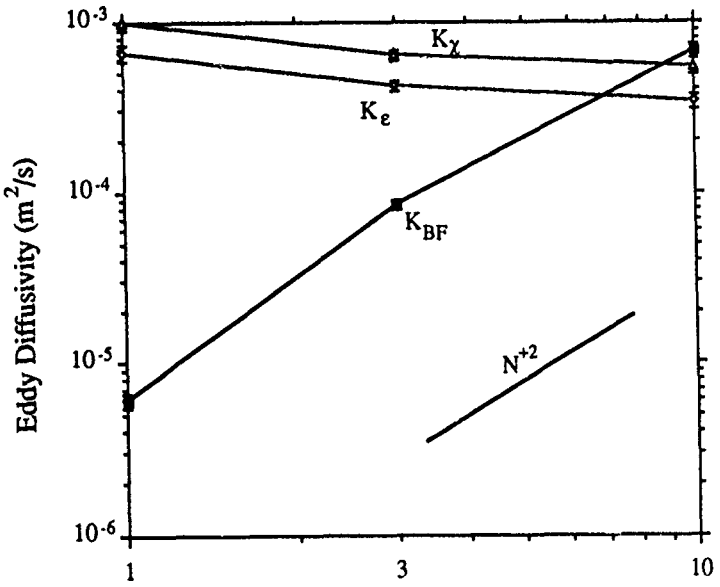


Figure 7: Dependency of space/time mean values of LES determined vertical eddy diffusivities (K_{BF}, K_ε and K_χ) upon buoyancy frequency, N. Standard deviations are shown as the error bars. A spectral slope of N^{+2} is also shown.

where PE is actually removed by the forcing. We are presently investigating this in detail. It should be mentioned that the time averaged BF will give the correct irreversible exchange of KE to PE (e.g., Lombard *et al.*, 1990) as it is not necessary to calculate the change in the background stratification in order to determine mixing coefficients (e.g., Winters and D'Asaro, 1990).

Several parameterization methods have been developed to predict values of vertical eddy diffusivity from measurements of turbulence dissipation rates. These indirect methods solve simplified turbulence variance budgets to determine values of vertical eddy diffusivity (e.g., Osborn and Cox,

1972; Osborn, 1980; Moum, 1990). For example, if the temperature variance dissipation rate, χ, is known, the temperature variance budget may be solved to give the eddy diffusion of heat, K_χ, (Osborn and Cox, 1972), or

$$K_\chi \equiv \frac{<\chi>}{(dT_s/dz)^2} \qquad (20b)$$

where dT_s/dz is the mean temperature gradient. Another method involves solving a simplified turbulent kinetic energy budget to provide an *upper* bound for values of the vertical eddy diffusivity (Osborn, 1980). After the ratio of shear to buoyancy production of turbulent kinetic energy is decided upon, values of ε_{KE} are used to provide the eddy diffusivity, K_ε, or

$$K_\varepsilon \equiv 0.2 \frac{<\varepsilon_{KE}>}{N^2} \qquad (20c)$$

In the past, eddy diffusivity parameterization assumptions have been tested by evaluating variance budgets, not by direct comparison to measurements of the buoyancy flux (Garrett, 1979; Washburn, 1987). Only recently have direct determinations of vertical velocity fluctuations enabled these comparisons to be made (Moum, 1990).

Values of K_χ and K_ε for the forced LES experiments range from 3.4×10^{-4} to 9.9×10^{-4} $m^2 s^{-1}$ and hence, vary considerably less than the two-decade range of variation observed for K_{BF} (Table 1). Generally, K_χ and K_ε are both greater than the directly determined K_{BF}. This was also observed in the *in situ* K_{BF} determinations by Moum (1990). However, the dependencies of K_χ and K_ε on N are very different than the dependency observed for K_{BF} (going like $N^{-1/4}$ instead of N^{+2}; fig. 7).

Obviously, there is a discrepancy in the vertical eddy diffusivities predicted by the direct and the variance dissipation methods. These differences are particularly apparent in the dependency of the estimated eddy diffusivities on the background stratification intensity, N. As discussed previously, the interactions of the numerical forcings with the determination of the BF are not well understood. However, the fact that the indirect methods do such a poor job in predicting K_{BF}, regardless of the what's forcing the large-scale (50 m) internal waves, indicates that these indirect methods are not adequately parameterize vertical eddy diffusivities for the conditions numerically simulated.

The existence of domain averaged BF oscillations in the present numerically forced LES experiments may have some interesting observational consequences. That is, although the domain-averaged BF is calculated over a volume equal to 125,000 m^3, its value is **not** stationary (fig. 5). Thus, instantaneous determinations of the buoyancy flux, even when averaged over a large spatial scales, may not adequately provide the necessary data to determine vertical eddy diffusivities for use in general circulation models. These large-scale ocean models require that small-scale internal wave "noise" be filtered out in terms of an effective eddy diffusivity. The present observations suggest that long time and large space integrations or *in situ* observations are necessary before relevant eddy diffusivities are to be determined.

WHAT'S NEXT?

The GM forced LES experiments have successfully replicated many of the dynamical quantities characteristic of the ocean's internal wave field. Specifically, direct calculations of vertical eddy

diffusivity, turbulent dissipation rates, and mixing efficiency have been surprisingly well in line with oceanographic thinking. Frankly, the fact that the *sign* of the eddy diffusivity was correct was a cause for celebration in Santa Barbara!! However, the anomalous dependency of K_{BF} on N raises doubts that our LES model is *realistically* simulating the internal wave field. Of course, there are many possibilities. The details of the numerical forcing, the omission of planetary vorticity (and hence, near-inertial waves), the restricted spatial scales simulated (< 50 m), the triply-periodic boundary conditions, possible lack of stationarity, as well as the subgrid scale parameterization method all could have contributed to the anomalous dependency of K_{BF} on N. Presently, we think that the "culprit" is the forcing mechanism and we are investigating, in detail, the energy transfers at the forcing wavenumber. This is not to say that each of the possibilities listed above do not need to be improved upon. In order for LES techniques to be of any use to oceanographers, the assumptions and mechanisms used to produce the simulations *must* be continually tested and evaluated.

We are presently implementing planetary vorticity (the f-plane approximation) into our LES model. This in the view of many is a serious omission in producing realistic simulations of the internal wave field. The inclusion of planetary vorticity (f-plane approximation) will enable us to force a LES with prescribed near-inertial waves. *In situ* observations have shown that much of the velocity finestructure is associated with near-inertial waves (Kunze *et al.*, 1990) and that these near-inertial waves can give rise to patches of turbulence (Gregg *et al.*, 1986). This improvement in the realism of the forcing mechanism should only help the LES model better replicate the internal wave field. An interesting hypothesis is that dependency of ε_{KE} on N, or for that matter K_{BF} on N, may depend upon the nature of how energy is fed into the internal wave field.

Eventually, it is envisioned that LES models can be used with *in situ* observations to provide "data" of the internal wave field that are presently unobtainable. Such a coupling would enable eddy diffusivities to be estimated for specific field experiments. This approach is similar to how hydrographic inverse or regional circulation models are used to give information about fluxes, vorticity and energy transfers that are, at best, difficult to determine from the field observations alone. Thus, the combination of field and numerical approaches may prove to be highly synergistic, allowing many new questions concerning the ocean's internal wave field to be addressed.

ACKNOWLEDGMENTS - The author would like to thank Peter Müller for inviting him to the 1991 'Aha Huliko'a Winter Workshop. Comments on this manuscript by Libe Washburn, Sally MacIntyre and Andrzej Domaradzki were extremely helpful. As always, the support, encouragement and guidance of Andrzej Domaradzki and Tommy Dickey is gratefully appreciated. Programming assistance from Ms. Xiaoning Duan has been instrumental in the success of the forced GM experiments which have been supported by the Office of Naval Research (N00014-90-J-1914). Supercomputer time has been provided by grants from the San Diego Supercomputer Center which is supported by the National Science Foundation.

REFERENCES

Browand, F.K., D. Guyomar, and S.-C. Yoon, 1987: The behavior of a turbulent front in a stratified fluid: experiments with an oscillating grid. *J. Geophys. Res.*, **92**, 5329-5341.

Canuto, C., M.Y. Hussaini, A. Quarteroni, and T.A. Zang, 1988: *Spectral Methods in Fluid Dynamics*. Springer-Verlag, p 567.

Chollet, J.P., and M. Lesieur, 1981: Parameterization of small scales of three-dimensional isotropic turbulence utilizing spectral closures. *J. Atmos. Sci.*, **38**, 2747-2757.

Chollet, J.P., 1984: Two point closures as a subgrid scale modeling for large eddy simulations. In: *Turbulent Shear Flows IV*, F. Durst and B. Launder (eds.), 62-72.

Clark, R.A., J.H. Ferziger and W.C. Reynolds, 1979: Evaluation of subgrid-scale turbulence models using a fully simulated turbulent flow. *J. Fluid Mech.*, **91**, 1-16.

Deardorff, J.W., 1970: A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers. *J. Fluid Mech.*, **41**, 453-480.

Deardorff, J.W., 1973: Three-dimensional numerical modeling of the planetary boundary layer. *Workshop on Micrometeorology*, D. Haugen, ed., American Meteorological Society, 271-311.

Deardorff, J.W., 1980: Stratocumulus-capped mixed layer derived from a three-dimensional model. *Bound.-Layer Meteor.*, **18**, 495-527.

Dickey, T.D., and G.L. Mellor, 1980: Decaying turbulence in neutral and stratified fluids. *J. Fluid Mech.*, **99**, 13-31.

Domaradzki, J.A., R.W. Metcalfe, R.S. Rogallo and J.J. Riley, 1987: Analysis of subgrid-scale eddy viscosity with use of results from direct numerical simulations. *Phys. Rev. Lett.*, **58**, 547-550.

Ebert, E.E., U. Schumann and R.B. Stull, 1989: Nonlocal turbulence mixing in the convective boundary layer. *J. Atmos. Sci.*, **46**, 2178-2207.

Eidson, T.M., 1985: Numerical simulation of the turbulent Rayleigh-Bénard problem using subgrid modelling. *J. Fluid Mech.*, **158**, 245-268.

Ferziger, J.H., 1983: Higher-level simulations of turbulent flows. In: *Computational Methods for Turbulent, Transonic, and Viscous Flows*. J.A. Essers (ed.), 93-182, Hemisphere Pub., Washington, D.C.

Gallacher, P.C., 1990: Large-eddy simulations of the turbulent boundary layer in the upper ocean. *Trans. Amer. Geophys. Union, EOS*, **71**, 1354 (abstract only).

Gargett, A.E., 1988: The scaling of turbulence in the presence of stable stratification. *J. Geophys. Res.*, **93**, 5021-5036.

Gargett, A.E., and G. Holloway, 1984: Dissipation and diffusion by internal wave breaking. *J. Mar. Res.*, **42**, 15-27.

Gargett, A.E., P.J. Hendricks, T.B. Sanford, T.R. Osborn, and A.J. Williams III, 1981: A composite spectrum of vertical shear in the upper ocean. *J. Phys. Oceanogr.*, **11**, 1258-1271.

Gargett, A.E., T.R. Osborn, and P.W. Nasmyth, 1984: Local isotropy and the decay of turbulence in a stratified fluid. *J. Fluid Mech.*, **144**, 231-280.

Garrett, C.J.R., 1979: Mixing in the ocean interior. *Dyn. Atmos Oceans*, **3**, 239-265.

Garrett, C., and W. Munk, 1972: Space-time scales of internal waves. *Geophys. Fluid Dyn.*, **2**, 225-264.

Garrett, C., and W. Munk, 1975: Space-time scales of internal waves: A progress report. *J. Geophys. Res.*, **80**, 291-297. (also, corrigendum *J. Geophys. Res.*, **80**, 3924)

Gregg, M.C., 1987: Diapycnal mixing in the thermocline: A review. *J. Geophys. Res.*, **92**, 5249-5286.

Gregg, M.C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.*, **94**, 9686-9698.

Gregg, M.C., and T.B. Sanford, 1988: The dependence of turbulent dissipation on stratification in a diffusively stable thermocline. *J. Geophys. Res.*, **93**, 12,381-12,392.

Gregg, M.C., E.A. D'Asaro, T.J. Shay, and N. Larson, 1986: Observations of persistent mixing and near-inertial waves internal waves. *J. Phys. Oceanogr.*, **16**, 856-885.

Holland, W.R. and L.B. Lin, 1975: On the origin of mesoscale eddies and their contribution to the general circulation of the ocean. I. A preliminary numerical experiment. *J. Phys. Oceanogr.*, **5**, 642-657.

Holloway, G., 1983: A conjecture relating oceanic internal waves and small-scale processes. *Atmos.-Ocean*, **21**, 107-122.

Holloway, G., 1988: The buoyancy flux from internal gravity wave breaking. *Dyn. Atmos. Oceans*, **12**, 107-125.

Holloway, G., and D. Ramsden, 1988, Theories of internal wave interaction and stably-stratified turbulence: Testing against direct numerical experimentation. In: *Small-Scale Turbulence and Mixing in the Ocean*, J.C.J. Nihoul and B.M. Jamart (eds.), 363-377, Elsevier, New York.

Hopfinger, E.J., 1987: Turbulence in stratified fluids: A review. *J. Geophys. Res.*, **92**, 5287-5303.

Itsweire, E.C., K.N. Helland, and C.W. Van Atta, 1986: The evolution of grid-generated turbulence in a stably stratified fluid. *J. Fluid Mech.*, **162**, 299-338.

Kerr, R.M., 1985: Higher-order derivative correlations and the alignment of small-scale structures in isotropic numerical turbulence. *J. Fluid Mech.*, **153**, 31-58.

Kraichnan, R.H., 1976: Eddy viscosity in two and three dimensions. *J. Atmos. Sci.*, **33**, 1521-1536.

Kunze, E., M.G. Briscoe and A.J. Williams III, 1990: Interpreting shear and strain fine structure from a neutrally buoyant float. *J. Geophys. Res.*, **95**, 18,111-18,125.

Ledwell, J.R., and A.J. Watson, 1991: The Santa Monica Basin tracer experiment: A study of diapycnal and isopycnal mixing. In press: *J. Geophys. Res.*

Leonard, A., 1974: Energy cascades in large eddy simulations of turbulent fluid flows. *Advances in Geophysics*, Vol. 18, Academic Press, 237-248.

Lesieur, M., 1987: *Turbulence in Fluids*. Martius Nijhoff Publ., Dordrecht, The Netherlands, p 286.

Lesieur, M., and R. Rogallo, 1989: Large-eddy simulation of passive scalar diffusion in isotropic turbulence. *Phys. Fluids A*, **1**, 718-722.

Lesieur, M., O. Métais, and H. Laroche, 1988: Numerical simulations of turbulent stably-stratified and free shear flows. (extended abstract), 57-60, In: *Eighth Symposium on Turbulence and Diffusion*, American Meteorological Society, Boston MA.

Lienhard V, J.H, and C.W. Van Atta, 1990: The decay of thermally stratified turbulence. *J. Fluid Mech.*, **210**, 57-112.

Lin, J.T., and Y.H. Pao, 1979: Wakes in stratified fluids. *Ann. Rev. Fluid Mech.*, **11**, 317-338.

Lombard, P.N., D.D. Stretch, and J.J. Riley, 1990: Energetics of a stably stratified mixing layer. *Proc. 9th Symp. Turb. Diff.*, 202-205.

Mansour, N.N., P. Moin, W.C. Reynolds, and J.H. Fertziger, 1979: Improved methods for large eddy simulations of turbulence. In: *Turbulent Shear Flows I*, F. Durst, B. Launder, F. Schmidt, and J. Whitelaw (eds.), Springer-Verlag, 386-401.

Mason, P.J., and S.H. Derbyshire, 1990: Large-eddy simulation of the stably-stratified atmospheric boundary layer. *Bound.-Layer Met.*, **53**, 117-162.

McWilliams, J.C., P.C. Gallacher, C.-H. Moeng, and J. Wyngaard, 1990: Large-eddy simulations of oceanic boundary layers. Presented at the International Workshop on Large Eddy Simulation, St. Petersburg, FL, December, 1990.

Métais, O., 1985: Evolution of three dimensional turbulence under stratification. In: Lesieur, M., 1987: *Turbulence in Fluids.* (also presented at Turbulent Shear Flows VI).

Métais, O., and J.R. Herring, 1989: Numerical simulations of freely evolving turbulence in stably stratified fluids. *J. Fluid Mech.*, **202**, 117-148.

Moeng, C.-H., 1984: A large-eddy simulation model for the study of planetary boundary-layer turbulence. *J. Atmos. Sci.*, **41**, 2052-2062.

Moeng, C.-H. and J.C. Wyngaard, 1988: Spectral analysis of large-eddy simulations of the convective boundary-layer. *J. Atmos. Sci.*, **45**, 3573-3587.

Moum, J. N., 1990: The quest for Kρ - Preliminary results from direct measurement of turbulent fluxes in the ocean. *J. Phys. Oceanogr.*, **20**, 1980-1984.

Müller, P., 1988: Vortical Motions. In: *Small-Scale Turbulence and Mixing in the Ocean,* J.C.J. Nihoul and B.M. Jamart (eds.), 285-301, Elsevier, New York.

Müller, P., G. Holloway, F. Henyey, and N. Pomphrey, 1986: Nonlinear interactions among internal gravity waves. *Rev. Geophys.*, **24**, 493-536.

Munk, W.H., 1966: Abyssal recipes. *Deep-Sea Res.*, **13**, 707-730.

Munk, W., 1981: Internal waves and small-scale processes. In: *Evolution of Physical Oceanography,* B. Warren and C. Wunsch (eds.), The MIT Press, Cambridge, MA, 264-291.

Oakey, N.S., 1982: Determination of the rate of dissipation of turbulent energy from simultaneous temperature and velocity shear microstructure measurements. *J. Phys. Oceanogr.*, **12**, 256-271.

Orszag, S., 1971: Numerical simulation of incompressible flows within simple boundaries: Accuracy. *J. Fluid Mech.*, **49**, 75-112.

Orszag, S., and G.S. Patterson, 1972: Numerical simulation of three-dimensional homogeneous isotropic turbulence. *Phys. Rev. Lett.*, **28**, 76-79.

Osborn, T.R., 1980: Estimates of the local rate of vertical diffusion from dissipation measurements. *J. Phys. Oceanogr.*, **10**, 83-89.

Osborn, T.R., and C.S. Cox, 1972: Oceanic fine structure. *Geophys. Fluid Dyn.*, **3**, 321-345.

Piomelli, U., P. Moin, and J.H. Ferziger, 1989: Model consistency in large-eddy simulation of turbulent channel flows. *Phys. Fluids*, **31**, 1884-1891.

Ramsden, D. and G. Holloway, 1991: Vortex-wave interactions in stably stratified turbulence. Submitted to *J. Geophys. Res.*

Riley, J.J., R.W. Metcalfe, and M.A. Weissman, 1981: Direct numerical simulations of homogeneous turbulence in density-stratified fluids. *AIP Conf. Nonlinear Properties of Internal Waves.* B. West (ed.), 79-112.

Rogallo, R., and P. Moin, 1984: Numerical simulations of turbulent flows. *Ann. Rev. Fluid Mech.*, **16**, 99-137.

Schmidt, H., and U. Schumann, 1989: Coherent structure of the convective boundary layer derived from large-eddy simulations. *J. Fluid Mech.*, **200**, 511-562.

Schumann, U., 1975: Subgrid-scale model for finite difference simulations of turbulent shear flows in plane channels and annuli. *J. Comp. Phys.*, **18**, 376-404.

Schumann, U., 1991: Subgrid length-scales for large-eddy simulation of stratified turbulence. In: *Theoretical and Computational Fluid Dynamics*, T. Gatski and C. Speziale (eds.), Springer-Verlag, Berlin.

Shen, C.Y., and G. Holloway, 1986: A numerical study of the frequency and the energetics of nonlinear internal gravity waves. *J. Geophys. Res.*, **91**, 953-973.

Siegel, D.A., 1988: Large-eddy simulation of the decay of a small-scale oceanic internal gravity wave field, Ph.D. Dissertation, Department of Geological Sciences, University of Southern California, p 187.

Siegel, D.A., 1991: Large-eddy simulation of steady thermocline motions. In preparation.

Siegel, D.A., and J.A. Domaradzki, 1991a: Large-eddy simulation of the decay of stably-stratified oceanic turbulence, I: Model development and observations of temporal variability. Submitted to *J. Phys. Oceanogr.*

Siegel, D.A., and J.A. Domaradzki, 1991b: Large-eddy simulation of the decay of stably-stratified oceanic turbulence, II: Characterization of the kinematic and dynamic structure. Submitted to *J. Phys. Oceanogr.*

Siegel, D.A., and A.J. Plueddemann, 1991: The motion of a solid sphere in an oscillating flow: An evaluation of remotely-sensed Doppler velocity estimates in the sea. *J. Atmos. Oceanic Tech.*, **8**, 296-304.

Siggia, E.D., and G.S. Patterson, 1978: Intermittency effects in a numerical simulation of stationary three-dimensional turbulence. *J. Fluid Mech.*, **86**, 567-592.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Mon. Wea. Rev.*, **91**, 99-164.

Smagorinsky, J., 1990: Some historical remarks on the use of non-linear eddy viscosities in geophysical models. Presented at the International Workshop on Large Eddy Simulation, St. Petersburg, FL, December, 1990.

Stewart, R.W., 1969: Turbulence and waves in a stratified atmosphere. *Radio Sci.*, **4**, 1269-1278.

Stillinger, D.C., K.N. Helland, and C.W. Van Atta, 1983: Experiments on the transition of homogeneous turbulence to internal waves in a stratified fluid. *J. Fluid Mech.*, **131**, 91-122.

Tennekes, H., and J.L. Lumley, 1972: *A First Course in Turbulence*, The MIT Press, Cambridge, MA., p 300.

Van Atta, C., 1990: Comment on "The scaling of turbulence in the presence of stable stratification" by A.E. Gargett. *J. Geophys. Res.*, **95**, 11,673-11,674.

Vasilopoulos, A., 1991: Volume imaging: A new world view. *Comp. Graphics World*, April, 63-72.

Washburn, L., 1987: Two-dimensional observations of temperature microstructure in a coastal region. *J. Geophys. Res.*, **92**, 10,787-10,798.

Winters, K.B., and E.A. D'Asaro, 1990, Diapycnal mixing associated with a wave breaking event. *Trans. Amer. Geophys. Union, EOS*, **71**, 1360 (abstract).

Wyngaard, J. (ed.), 1984: *Large-Eddy Simulation: Guidelines for its Application to Planetary Boundary Layer Research.* Report from the Working Group on Large Eddy Simulation, Boulder, CO., p 122.

Yakhot, A., S.A. Orszag, V. Yakhot, and M. Israeli, 1989: Renormalization group formulation of large-eddy simulations. *J. Sci. Comp.*, **4**, 139-158.

Yamazaki, H., 1990: Stratified turbulence near a critical dissipation rate. *J. Phys. Oceanogr.*, **20**, 1583-1598.

Yoshizawa, A., 1982: A statistically derived subgrid scale model for large eddy simulation of turbulence. *Phys. Fluids*, **25**, 1532-1538.

NUMERICAL MODELING OF THE LARGE-SCALE DYNAMICS OF INTERNAL WAVES

David Rubenstein

Science Applications International Corp.
PO Box 1303, McLean, VA 22102

ABSTRACT

Several issues related to large-scale modeling of internal waves are discussed. First, we consider the question of the linear internal wave response of the ocean to surface forcing. Past attempts at analyzing the linear response are briefly reviewed. Then we discuss a numerical experiment, in which a nonlinear model of internal waves generates a time-evolving field of motions in a vertical plane. The model flow is initially at rest, and is forced with a simple surface-layer body force with a long (40 km) wavelength. Despite the fact that the model was not initialized with a particular spectrum, it develops a frequency-wavenumber spectrum with features that are similar to that of Garrett and Munk. After a saturated spectrum develops, the dispersion of a cloud of Lagrangian tracer particles is analyzed. The linear dimensions of the cloud expand roughly as the square root of time. Based on the expansion rate we estimate values for eddy diffusivity coefficients. We find that these coefficients are independent of length scale, but vary approximately linearly with the internal wave horizontal kinetic energy. For a kinetic energy level equivalent to that of the Garrett-Munk spectrum and a stratification N~3 cph, we calculate K_x~0.26 m^2/sec, K_z~2×10^{-4} m^2/sec.

1. INTRODUCTION

Near-inertial waves are the most energetic component of the internal wave spectrum. Therefore it seems natural to focus (at least initially) on near-inertial waves, to understand their generation, dynamics, interactions, and dissipation. Near-inertial internal waves have rather long horizontal length scales (100's of meters to 10's of kilometers) in comparison with higher frequency waves. Therefore a numerical model of near-inertial waves should span this broad range of length scales.

The problem is even more difficult than this. Surface forcing of near-inertial waves due to atmospheric events covers an even broader range of length scales—up to 100's of kilometers. Certainly a wide variety of mesoscale, upper surface-layer and bottom boundary layer processes interact with near-inertial waves in this regime of spatial scales. Therefore it is important to try to sort out the relative magnitudes of all of the different interactions.

My overall objectives are to try to answer three important questions: 1) How are large-scale internal waves generated? 2) Once they are generated, how are their space and time scales maintained? 3) How do these waves interact with, and contribute to the mixing of, mesoscale flows?

Because the questions are still largely unanswered, I am following a multi-pronged approach. The first approach, discussed in Section 2, is to determine the extent to which linear dynamics are capable of generating the internal wave spectrum. A body of circumstantial evidence indicates that many of the characteristics of the internal wave field can be explained using linear mechanisms. The next approach, discussed in Section 3, is to determine whether wave-wave interactions are

sufficient to generate the internal wave spectrum, and over what time scales does such a spectrum develop and decay. Then results of some experiments related to the mixing of a passive tracer are presented in Section 4. The space-plus-time trajectories of Lagrangian tracer particles are calculated, and give us some insights into the mixing process.

2. LINEAR DYNAMICS OF GENERATION

It is known that wind-forcing is an important generator of near-inertial motions. Simple linear models of wind-induced inertial motions are sometimes able to explain a very large percentage of the inertial energy in the surface mixed layer. Sometimes these models fail in their deterministic predictions, for a variety of reasons. The reasons may include, for example, incomplete knowledge of the full wind field (its history and its spatial scales), mesoscale motions which interact with the inertial motions, inertial waves which have propagated from a distant source, and so on.

An important question which arises is, to what extent can *linear* dynamics explain the generation of near-inertial waves, and perhaps, the generation of the *entire* spectrum of internal waves. As far as near-inertial waves are concerned, there is a body of evidence that linear dynamics are sufficient. For example, Rubenstein (1983) showed that observed features related to vertical phase propagation and interchange of energy between the surface layer and the thermocline can be explained with linear dynamics. Kundu and Thomson (1985) showed that linear theory can explain the observed intermittency of near-inertial waves, and the horizontal phase structure of surface-layer oscillations. Gill (1984) presented a linear theory which explained many other features of near-inertial waves, such as the tendency for horizontal and vertical scales to decrease after a storm has passed. Eriksen (1988) examined the linear response of near-inertial waves to a wind stress which deposits a body force as a step function in the vertical; uniformly within a surface layer, and zero below. He found that the vertical wavenumber spectral response is the same as that predicted by the GM81 empirical spectrum.

In a sense, our understanding of the vertical scales of internal waves is much better than that of the horizontal scales. High-quality vertical profiles are easier to produce than horizontal profiles—especially profiles of velocity. Also, there are major gaps in our 2D+time description of the wind stress field. There have been several attempts to jump over these gaps.

Several years ago I performed a little study, in which I looked at the patterns of near-inertial motions generated in the surface layer. I digitized a couple of months of 6-hr synoptic weather maps in a 2600-km square region of the North Atlantic, roughly from 35°N to 60°N, and from 15°W to 50°W. I gridded the isobars and computed the geostrophic wind, adjusted for friction and for turning effects. Figure 1 shows an example of one such map. I used a very simple slab model for the mixed layer (uncoupled in the horizontal), and observed the patterns of near-inertial oscillations. Figure 2 shows a single snapshot of the u-component of velocity, after 40 days of evolution. There seems to be a notable difference between the northern and southern regions of the model domain. The northern half is dominated by propagating low-pressure systems. The wave pattern orientations and length scales are variable, on synoptic time scales. In contrast, the southern half of the model domain is dominated by a climatological high pressure system. As a result, the wave pattern is more static. The wave patterns tend to be elongated along latitude lines, because of the beta effect.

D'Asaro (1989) used an imaginative technique to bridge the wind-field problem. He used a Seasat scatterometer wind stress field, with 25-km resolution, and advected the field across the modeled

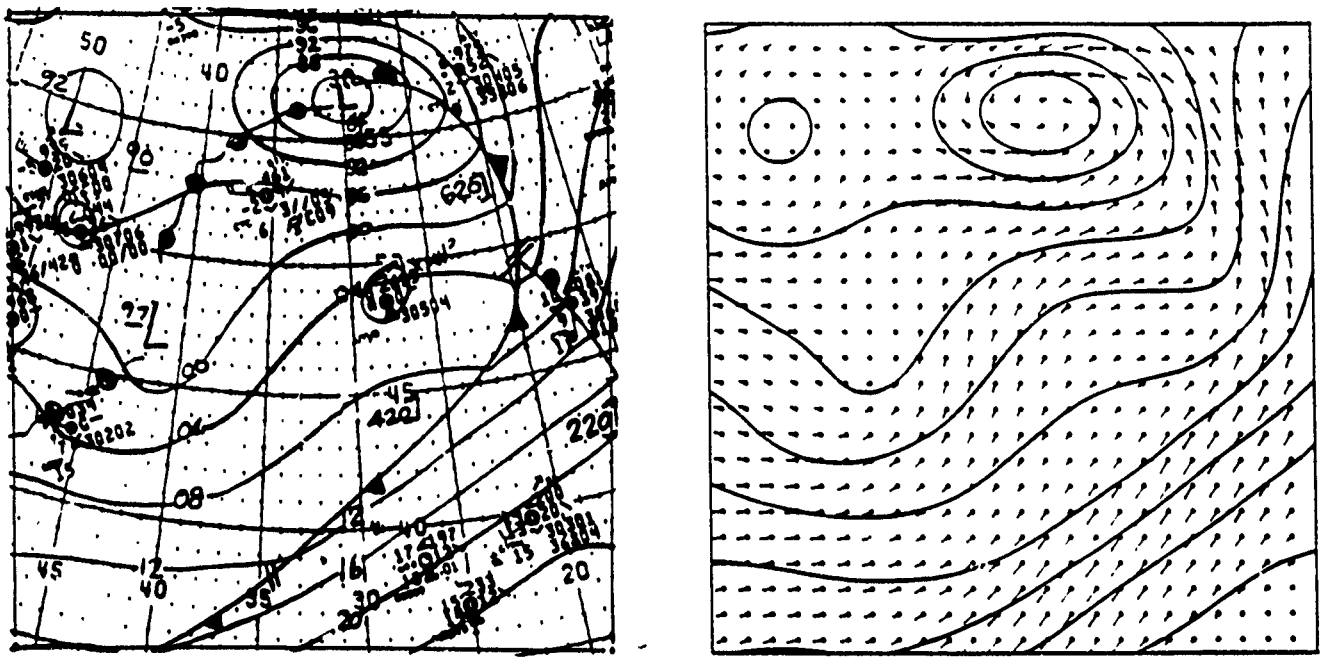


Figure 1. Left panel: Sample weather map over central North Atlantic. Right panel: Digitized, contoured isobars, and geostrophic wind adjusted for friction and turning effects.

ocean. In this way he was able to estimate the relative contributions of the wind field, the advection speed, and the beta effect, in generating smaller horizontal scales of near-inertial waves.

The problem with these approaches is that mesoscale and smaller-scale variance in the wind stress field is deficient. I've been developing a possible remedy to this problem. The approach is to start with an atmospheric mesoscale model, and to perform a stochastic interpolation in time and space, to resolve higher wavenumbers and frequencies. This is a sort of engineering approach, but if it yields a statistically realistic evolving wind field, it could give meaningful results.

3. NONLINEAR MODEL

If we apply an intermittent, large-horizontal-scale wind-induced surface forcing to a nonlinear model, what sort of internal wave spectrum develops? Over what time scale does such a spectrum develop, and after forcing is removed, over what time scale does it dissipate?

In an attempt to answer these questions, we formulate a two-dimensional model in a vertical x-z plane. Motion is allowed in three directions, but the model is invariant in the y-direction. We assume an initially motionless ocean, with an exponentially decaying buoyancy frequency, and a fluid which satisfies the Boussinesq approximation. The lateral, top and bottom boundary conditions are reflecting.

The coordinates (x,y,z) are defined with z positive upward, with the origin at the ocean bottom. A channel of depth D=1 km (corresponding to a flat ocean bottom) and width X=20 km contains the flow. The equations of motion are

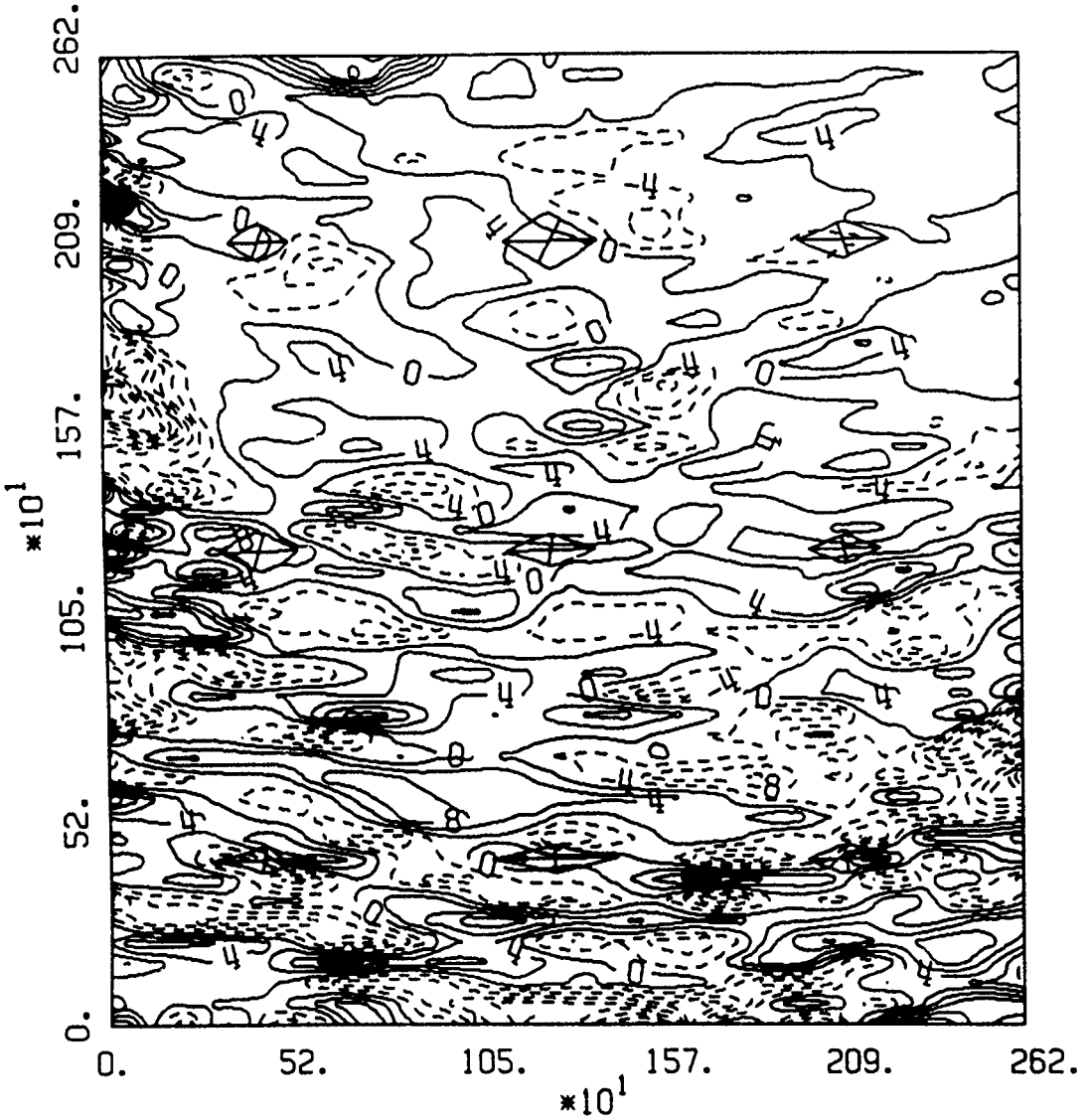


Figure 2. U-component of velocity in the surface layer, predicted by a simple slab model. The spatial domain is a 2,620 km square region, as shown in Fig. 1. The contour interval is 4 cm/sec. The elongation of structures along latitude lines is due to the beta-effect. The longer length scales in the northern half of the domain is due to the rapid, successive propagation of atmospheric fronts.

$$\partial_t u + u\partial_x u + w\partial_z u - fv = -\partial_x p + \kappa_x \partial_{xx} u + \kappa_z \partial_{zz} u + F \tag{1}$$

$$\partial_t v + u\partial_x v + w\partial_z v + fu = + \kappa_x \partial_{xx} v + \kappa_z \partial_{zz} v \tag{2}$$

$$\partial_t b + u\partial_x b + w\partial_z b + N^2 w = + \kappa_x \partial_{xx} b + \kappa_z \partial_{zz} (b + B) \tag{3}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \tag{4}$$

$$\frac{\partial p}{\partial z} = b \tag{5}$$

$$N^2 = \frac{\partial B}{\partial z}, \quad B_T = B + b \tag{6}$$

The velocity components are (u, v, w), p is a reduced perturbation pressure, $N=(3\ \text{cph})\times e^{-z/1300}$ is the buoyancy frequency profile, and f is the Coriolis parameter. The total buoyancy B_T is split up into a steady, initial profile $B(z)$ and a perturbation buoyancy b. The κ_x and κ_z terms are diffusivity coefficients, and are uniform in space. A rigid lid leads to an additional constraint, that

$$\int_0^D u\,dz = 0 \tag{7}$$

The model is driven by a body force distributed in the surface layer of depth H. The body force has a single, large-scale horizontal sinusoidal component, and is given by

$$F(x,z,t) = F(t)\sin(\frac{\pi x}{X})\frac{1}{2}[1+\sin(\pi\frac{z-(D-H)}{2H})] \ . \tag{8}$$

The behavior of $F(t)$ simulates a short-duration impulse. During an initial spinup period, these impulses were applied at regular intervals of 5.5 inertial periods. After the wavenumber spectrum is saturated, the forcing is turned off, and the flow is allowed to decay.

The diffusion coefficients κ_x and κ_z are chosen to have the minimum possible values, and still maintain numerical stability. During the spinup period, we set $\kappa_x=0.1$ m^2/sec and $\kappa_z=10^{-4}$ m^2/sec, and during the decay period we set $\kappa_x=0.05$ m^2/sec and $\kappa_z=5\times10^{-5}$ m^2/sec. We find that during the decay period, energy dissipates with an e-folding time scale of about 30 days. This time scale falls within the 15-44 days time scale estimated for the replenishment of the internal wave field (Gregg, 1987).

Spatial derivatives are approximated by centered differences, and time integration is performed using the two-step Lax Wendroff technique. The finite difference equations are solved on a 128×128 grid, using a time step of 12 seconds. The artificial viscosity associated with this numerical technique is given by (Roache, 1972)

$$\kappa_{art} = u^2\,\Delta t/2 \ . \tag{9}$$

In the model runs presented here, this artificial viscosity is less than a quarter of the explicit diffusion coefficients, in both the horizontal and the vertical.

Figure 3 shows spectra of kinetic energy in two projections; ω-k_x and ω-k_z, where ω is frequency, k_x is horizontal wavenumber, and k_z is vertical wavenumber. We first consider the ω-k_z spectrum. The strongest ridge parallel to the k_z axis corresponds to near-inertial waves. The next-strongest

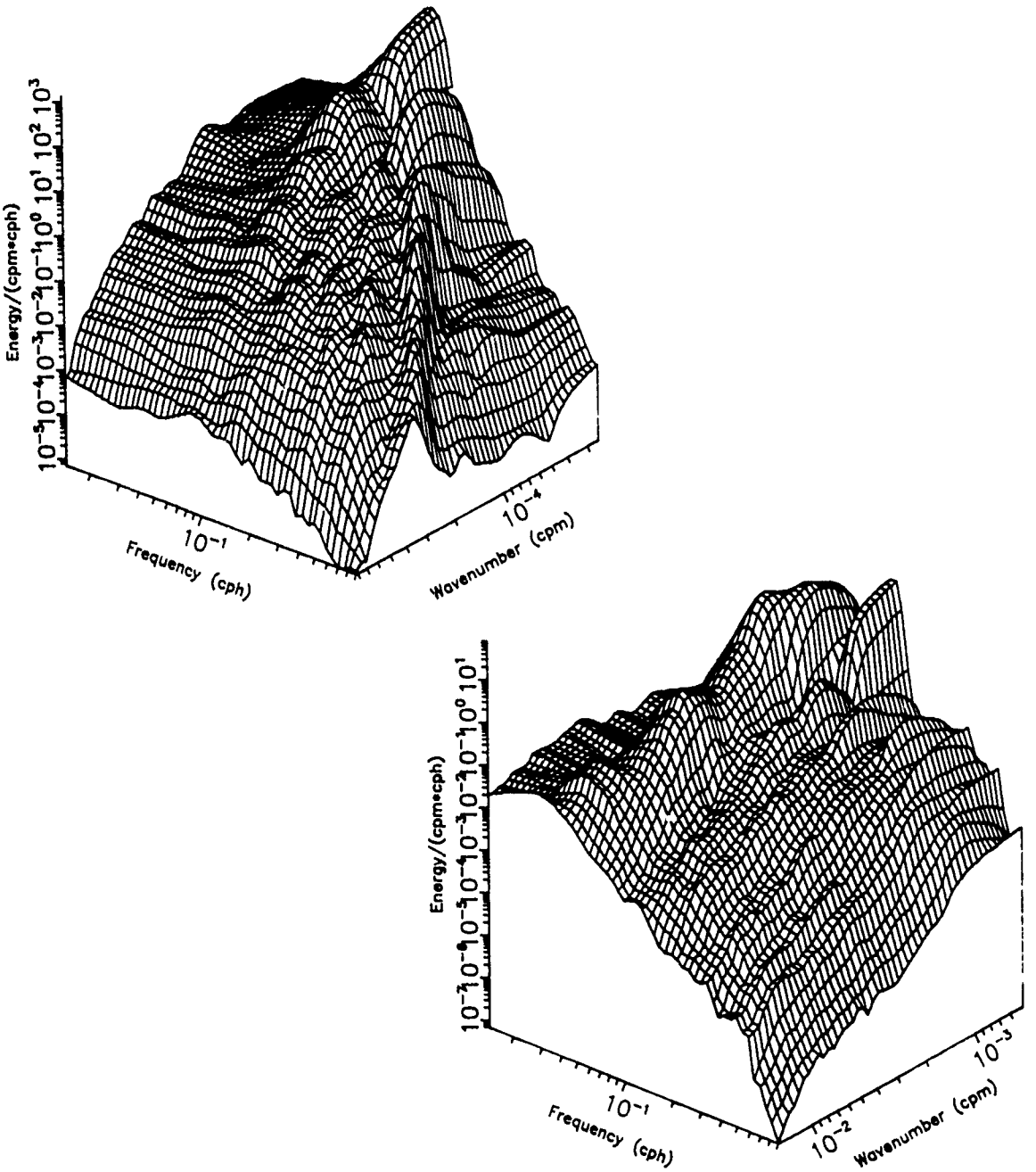


Figure 3. Kinetic energy spectrum, predicted by the nonlinear numerical model. The top panel shows a projection in ω-k_x space, and the bottom panel shows ω-k_z.

ridge, at a slightly higher frequency, corresponds to the first vertical mode. The weaker, higher-frequency ridges correspond to harmonics. In the ω-k_x spectrum, we see some of the same features, but in addition we see a set of radial ridges. The radial ridges are the projections of discrete vertical modes onto the ω-k_x spectrum.

It is also interesting to note that the ω-k_z spectrum appears to be nearly (but not exactly) separable between ω and k_z, while the ω-k_x spectrum is not remotely separable in ω and k_x. This feature of separability in ω-k_z and non-separability in ω-k_x is similar to that of the Garrett-Munk spectrum (for example, Munk, 1981). The feature is somewhat at odds with the slant-wavenumber frequency spectrum observed by Pinkel (1984). Pinkel's wavenumber spectra in the upper ocean exhibit slopes which flatten with increasing frequency.

On the other hand, both the modeled ω-k_z spectrum and Pinkel's (1984) observed spectrum show a set of ridges parallel to the k_z axis. In the case of the model, these ridges represent harmonics between the lowest vertical modes of the near-inertial frequency waves. In the case of Pinkel's observations, these ridges represent possible harmonics between tidal and inertial frequency motions. In both model and observations, these ridges stand out at low wavenumber, but not at high wavenumber.

4. DISPERSION OF LAGRANGIAN TRACER PARTICLES

Eddy viscosity and diffusivity coefficients are employed by ocean circulation models, for the purpose of parameterizing sub-grid scale mixing processes. Much of this sub-grid scale mixing is due to the internal wave field. Several techniques have been developed for the estimation of diffusivities. Direct observations of internal waves (Ruddick and Joyce, 1979; Brown and Owens, 1981; Kunze, 1986; Hebert, 1987) use the eddy correlation technique to derive eddy viscosity and diffusivity. Sometimes even the sign of the derived viscosity or diffusivity coefficients is in doubt. These direct observations are limited because in the thermocline, eddy correlation signals are weak, and they are disturbed by non-stationarity of the mean flow.

Other techniques for estimating viscosity and diffusivity parameters are extensively reviewed by Gregg (1987). These techniques include inverse methods over regional domains, flux estimations from microstructure measurements, compilations of statistics of mixing patches, and direct measurements of dye patches (Ewart and Bendiner, 1981; Ledwell, 1989).

One technique that has not yet received much attention is numerical simulations. Numerical simulations have the advantage that they are capable of separating the diffusion due to internal waves from that of mesoscale motions. They also allow sensitivity studies to systematic parameter variations, and of course, allow complete understanding of the underlying flow field. On the other hand, full nonlinear models have the significant disadvantage of covering a limited range of length scales. Another limitation is the lack of direct control over the spectral level and shape.

To help understand mixing due to internal waves, we consider the dynamics of Lagrangian particles. Figure 4 shows the trajectories of nine individual Lagrangian particles over an 8-day interval. The particles are tracked between the grid cells using bilinear interpolation. The horizontal extents Δx of the tracer trajectories are on the order of 1 to 4 km, and the vertical extents Δz range from 30 to 80 m.

We are not interested only in the absolute dispersion of a tracer particle; we are interested in the *relative* dispersion of a cluster of particles. The four panels in Figure 5 show the positions of 1000 tracer particles at four instances in time; 1.6, 3.3, 6.6, and 11.6 days. The thick circle in each panel (actually an ellipse in physical space; major axis=312 m in the horizontal, minor axis=15.6 m in the vertical) indicates the initial distribution of the particles. In each panel, the portrayed positions of the particle clouds have been translated to keep the centers of mass coincident with the centers of the panels. At the early times, the particles remain organized in stringy clumps. At the later times, the particles separate from one another, and their positions become more random.

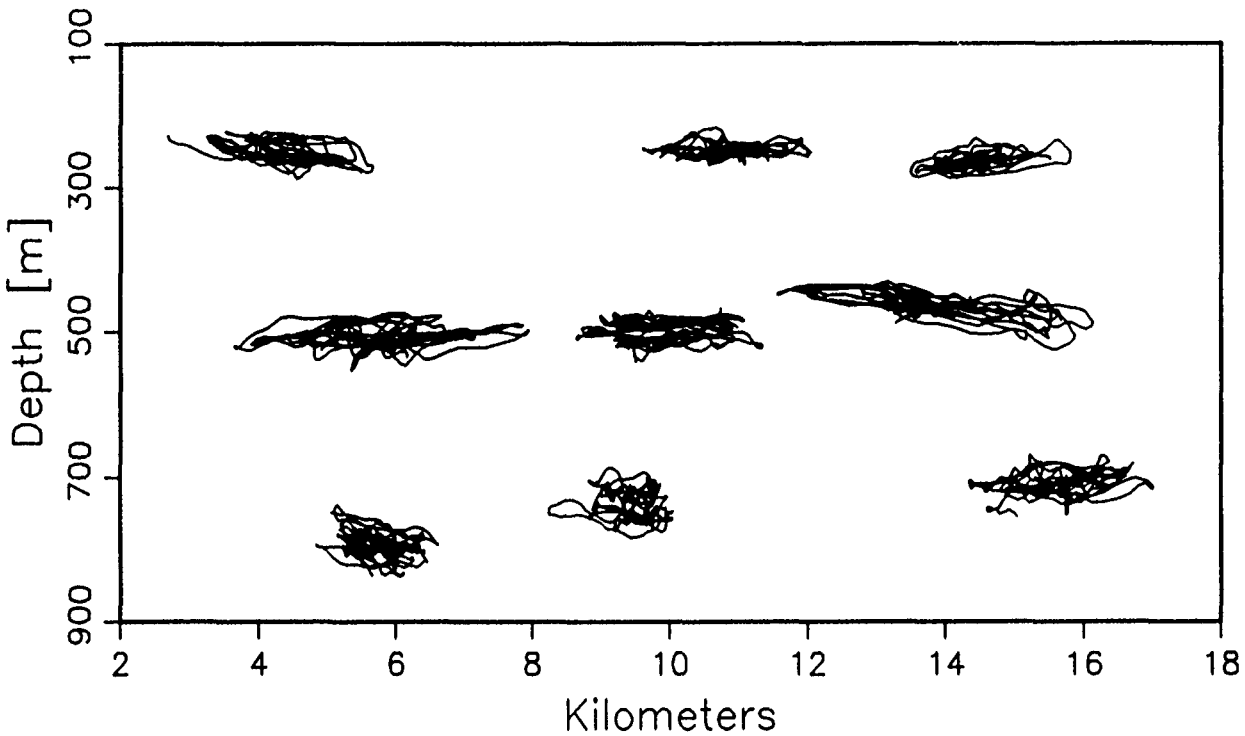


Figure 4. Trajectories of nine individual Lagrangian particles in a vertical plane, due to modeled internal waves over an 8-day interval when the model was quite energetic.

In each panel, a thin-line ellipse is drawn to indicate the standard deviation of the particle distribution in two orthogonal directions. The size of the ellipse tends to increase with time. By 11.6 days, the ellipse has increased by a factor of 5.5 relative to its initial size. The eccentricity does not have a noticeable trend with time.

To obtain a statistical average, we constructed a hierarchical grid of clouds of Lagrangian particles. A 5×5 grid of cloud centers was placed in the central area of the model domain. These clouds were initially circular in the finite grid space, and therefore elliptical in physical space. A hierarchy of 5 initial cloud sizes allowed a larger range of spatial scales to be examined. We traced the trajectory of each particle, and computed the ensemble average variance of the cloud distributions (segregated by initial cloud size) in the horizontal and vertical directions.

Figure 6 shows the average cloud distribution variances in the horizontal and vertical directions. Both panels in the figure show a set of five curves, each corresponding to a different initial cloud radius. The thin, oscillating curves represent the ensemble-averaged distribution variances, and the thick, smooth curves represent low-pass filtered time series, with a 1-day time constant. For the first 8 days, the variance increases approximately linearly with time. From this rate of increase, we can estimate the effective eddy diffusion coefficients in the horizontal and in the vertical,

$$K_x = \frac{1}{2}\frac{d\sigma_x^2}{dt} \approx 0.7\ \mathrm{m^2/sec}\ , \qquad (10)$$

$$K_z = \frac{1}{2}\frac{d\sigma_z^2}{dt} \approx 5\times10^{-4}\ \mathrm{m^2/sec}\ .$$

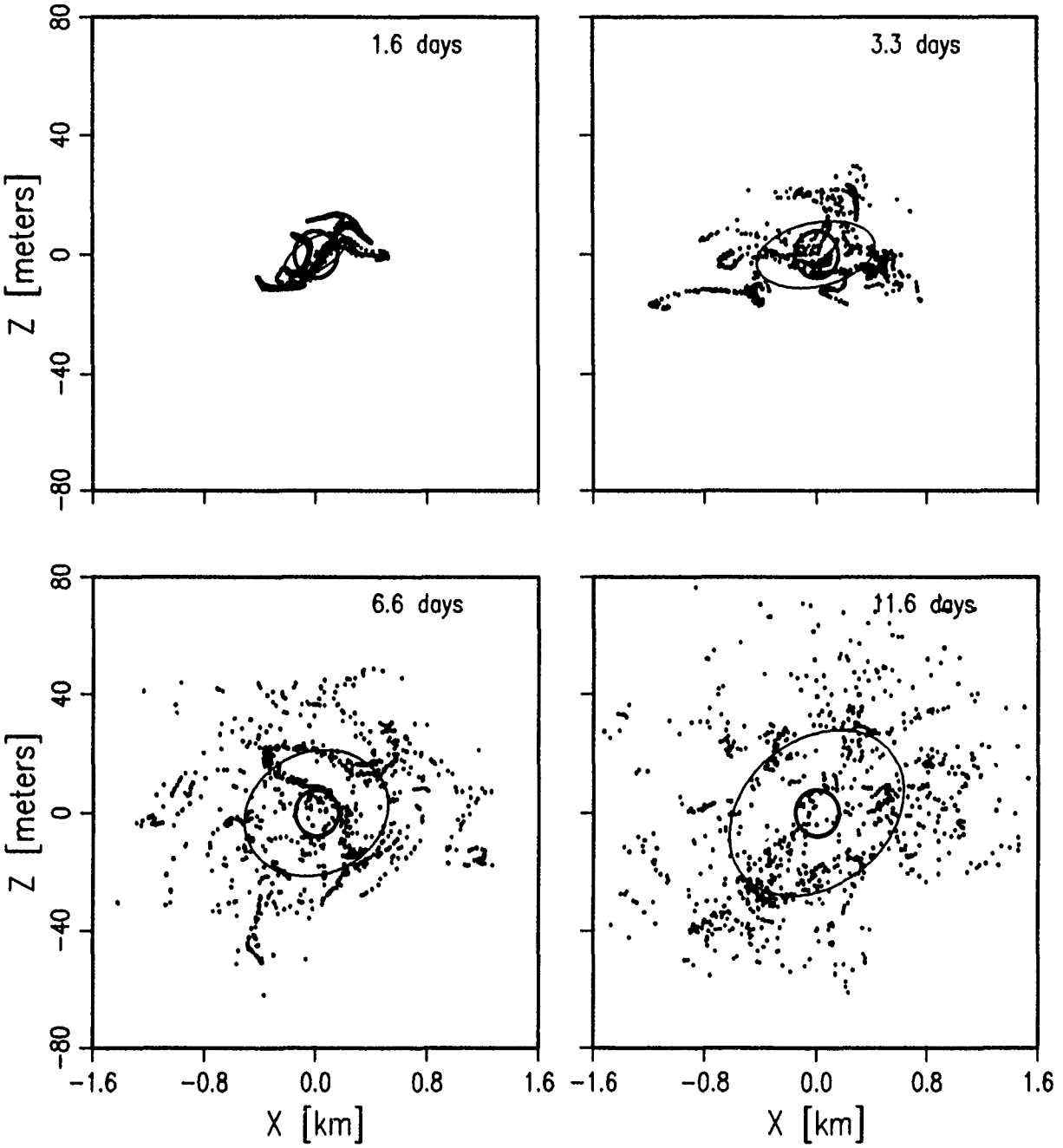


Figure 5. Relative dispersion of 1000 tracer particles in a vertical plane at four instants in time; 1.6, 3.3, 6.6, and 11.6 days. The thick curve in each panel (appears as a circle in this compressed coordinate space) is the locus of the initial distribution of particles. The thin-line ellipse in each panel indicates the standard deviation of particle distribution, in two orthogonal directions.

After 8 days, the variance increases more slowly (K_x~0.3 m^2/sec, K_z~3×10^{-4} m^2/sec). This occurs because kinetic energy is slowly dissipating (the rms u-component of velocity has decreased from 7.5 to 6 cm/sec), and the rate of expansion of the cloud size is very sensitive to the kinetic energy spectrum.

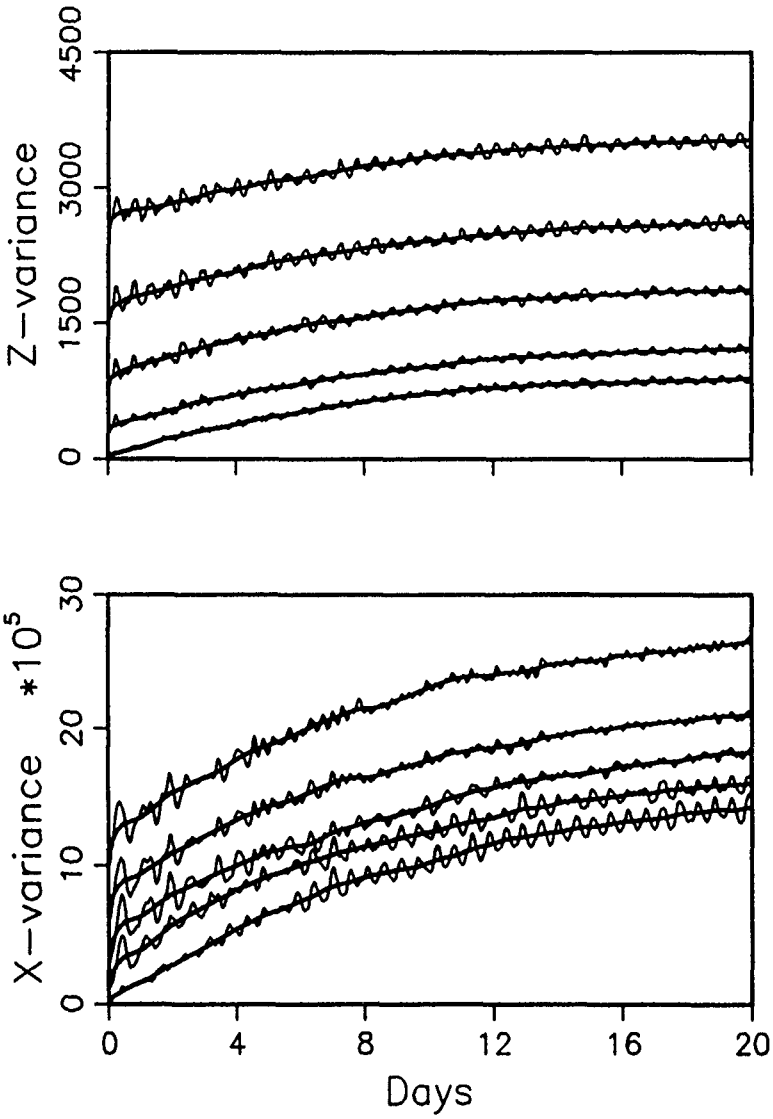


Figure 6. Thin curves: Ensemble average tracer cloud distribution variances in the horizontal and vertical directions. Thick curves: Time-average (1 day time constant) which smooths out the oscillations due to isopycnal deformations.

Despite the gradual decrease in expansion rate of the tracer clouds, the distribution variances for the various initial cloud sizes are all parallel to one another. The implication is that although the rates of increase $d\sigma_x^2/dt$ and $d\sigma_z^2/dt$ are sensitive to the kinetic energy spectrum, they are independent of σ_x^2 and σ_z^2. In other words, the effective eddy diffusivities are independent of the length scales involved. This result is in contrast with Okubo's (1971) finding that horizontal diffusivity increases with length scale to the 1.1 power. We must keep in mind that the observations compiled by Okubo include the effects of the entire, complicated spectrum of ocean velocities. The numerical prediction that eddy diffusivity is independent of length scale is largely due to the facts that only internal wave motions are included, and that tracer cloud length scales fall within the internal wave spectral continuum, which has a rather constant slope over a wide range of wavenumbers.

The short term oscillations in Fig. 6 represent deformations of the tracer clouds, associated with isopycnal straining and tilting. The ensemble average over sets of independent tracer clouds has removed a portion of the isopycnal deformations. The time-averaged, smooth curves have removed the residual deformations, and represent true diapycnal (across density contours) and isopycnal (along density contours) mixing.

Horizontal and vertical diffusion coefficients are plotted in Figure 7, as a function of rms horizontal velocity. These coefficients were estimated at different times during the model integration, as the energy level dissipated. Best fitting quadratic coefficients were estimated;

$$K_x = 54\,\sigma_u^2\,, \qquad K_z = 0.042\,\sigma_u^2\,, \tag{11}$$

where σ_u is the rms horizontal velocity. At the buoyancy frequency N=3 cph, the Garrett-Munk spectral energy level yields σ_u~7 cm/sec, and from Eq. (11) we get K_x~0.26 m^2/sec and K_z~2×10^{-4} m^2/sec. A quadratic fit of K_x and K_z to σ_u seems to be reasonable; therefore the eddy diffusion coefficients vary approximately linearly with horizontal kinetic energy.

The results of this study can be compared with recent measurements of ocean tracers. Ledwell (1989) measured the diapycnal spreading rate of a tracer, and deduced an approximate value K_z~3×10^{-5} m^2/sec, in the Santa Monica Basin. During the experiment, the buoyancy frequency was about 1 cph, and the rms internal wave velocity (excluding the low-mode, semidiurnal internal tide component) was approximately σ_u=3 cm/sec. Substituting this value into Eq. (11), we get K_z=3.8×10^{-5} m^2/sec, in good agreement with the measured value.

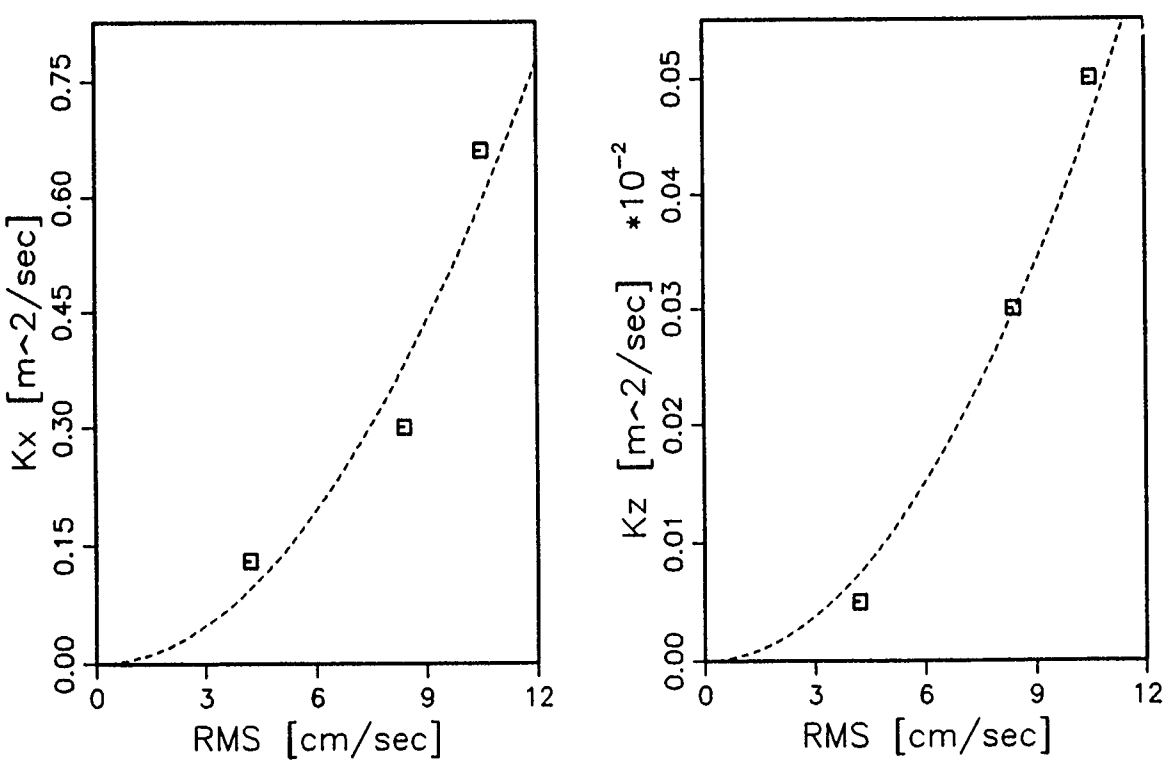


Figure 7. Horizontal and vertical diffusion coefficients, as a function of rms horizontal velocity associated with internal waves.

In another dye spreading experiment performed by Ewart and Bendiner (1981), several types of diffusion were computed. Assuming Fickian diffusion, estimates for K_x ranged from 0.03 to 0.12 m^2/sec. A local estimate of $K_z \sim 10^{-6}$ m^2/sec was observed associated with edge gradients, and a global value of about 5×10^{-5} m^2/sec was associated with total patch thickness. Assuming a Garrett-Munk energy level at $N=1$ cph, these estimates should be compared with model predictions of $K_x=0.08$ m^2/sec and $K_z=7 \times 10^{-5}$ m^2/sec.

The randomization of particle positions shown in Fig. 5 implies that the perimeters of the particle clouds become increasingly convoluted with time. Figure 8 shows the perimeters of four clouds of particles, initially concentric ellipses. The length scales eventually become shorter than the grid cell resolution. Even though the perimeter of a tracer cloud becomes increasingly convoluted with time, the buoyancy field does not. As soon as a buoyancy field kink develops to a sufficiently short scale, the explicit diffusivity parameterization in Eqs. (1)-(3) smooths the kink out. In this way the tracer particles, which are not directly affected by the explicit diffusivity, become disassociated from any specific isopycnal surface. Strong internal wave motions generate these "kinks" more rapidly than do weak motions, and therefore diapycnal mixing is more rapid. The exact values of the explicit coefficients κ_x and κ_z are not critical to the effective eddy diffusivities. In an experiment where κ_x and κ_z were doubled in value, the energy level decayed more rapidly, so the effective eddy diffusivities K_x and K_z likewise decreased in magnitude.

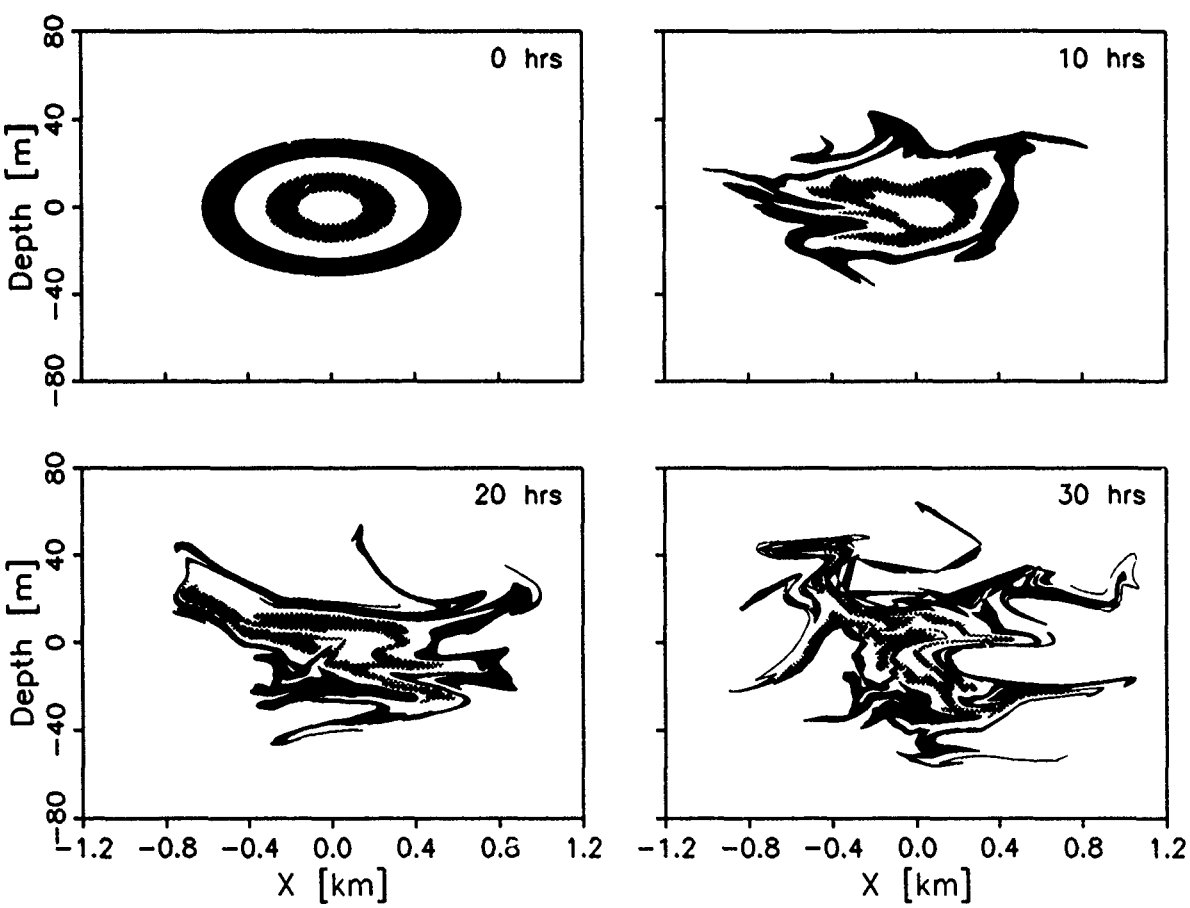


Figure 8. Four clouds of tracer particles were released into the numerical model, in initial concentric elliptical distributions. The outlines of these clouds are indicated by the inner and outer perimeters of the shaded areas.

In order to attain physical intuition into the diffusion mechanism, Young et al. (1982) developed analytic models of advection-diffusion. Young et al. modeled cases where a vertically sheared horizontal velocity field advects a tracer. Given tracer diffusivities κ_x and κ_z, an additional effective horizontal diffusivity K_x develops due to the interaction of the vertical shear and κ_z. They assumed a value of $\kappa_z \sim 10^{-5}$ m^2/sec, and for a particular model shear spectrum derived

$$K_x - \kappa_x \approx 1300\,\kappa_z \approx 0.013\ \text{m}^2/\text{sec} \tag{12}$$

The physical mechanisms incorporated into the numerical model are different from those analyzed by Young et al. (1982). The numerical model explicitly includes vertical velocities, and implicitly develops an effective vertical diffusivity K_z. If we were simply to substitute the value of the diffusion coefficient $\kappa_z = 5\times10^{-5}$ m^2/sec—which was used in Eqs. (1-3)—into Eq. (12), then we would obtain an effective diffusivity K_x much smaller than numerically computed values. On the other hand, if we combine the expressions in Eq. (11), we get the relationship

$$K_x = 54/0.042\,K_z = 1286\,K_z\ . \tag{13}$$

If we acknowledge that nonuniform, nonstationary vertical velocities in the numerical model generate an effective eddy diffusivity K_z in addition to κ_z, then we see that the results expressed in Eqs. (12) and (13) are quite comparable.

ACKNOWLEDGMENTS: This work is a result of research sponsored by the Office of Naval Research, under Contract N00014-88-C-0225. Thanks go to Michael Brill and David Bacon for stimulating discussions from which this study benefitted.

REFERENCES

Brown, E.D., and W.B. Owens, 1981: Observations of the horizontal interactions between the internal wavefield and the mesoscale flow. *J. Phys. Oceanogr.*, **11**, 1474-1480.

D'Asaro, E.A., 1989: The decay of wind-forced mixed layer inertial oscillations due to the ß effect. *J. Geophys. Res.*, **94**, 2045-2056.

Eriksen, C.C., 1988: On wind forcing and observed oceanic wave number spectra. *J. Geophys. Res.*, **93**, 4985-4992.

Ewart, T.E., and W.P. Bendiner, 1981: An observation of the horizontal and vertical diffusion of a passive tracer in the deep ocean. *J. Geophys. Res.*, **86**, 10,974-10,982.

Gill, A.E., 1984: On the behavior of internal waves in the wakes of storms. *J. Phys. Oceanogr.*, **14**, 1129-1151.

Gregg, M.C., 1987: Diapycnal mixing in the thermocline: A review. *J. Geophys. Res.*, **92**, 5249-5286.

Hebert, D., 1987: An estimate of the effective horizontal eddy viscosity in the Gulf Stream due to internal waves. *J. Phys. Oceanogr.*, **17**, 1837-1841.

Kundu, P.K. and R.E. Thomson, 1985: Inertial oscillations due to a moving front. *J. Phys. Oceanogr.*, **15**, 1076-1084.

Kunze, E., 1986: The mean and near-inertial wave velocity fields in a warm-core ring. *J. Phys. Oceanogr.*, **16**, 1444-1461.

Ledwell, J.R., 1989: A strategy for open ocean mixing experiments. *Parameterization of Small-Scale Processes,* Proceedings Hawaiian Winter Workshop, Univ. of Hawaii, P. Muller and D. Henderson, Eds., pp. 157-163.

Munk, W.H., 1981: Internal waves and small scale processes. *Evolution of Physical Oceanography,* B.A. Warren and C. Wunsch, Eds., The MIT Press, 264-290.

Okubo, A., 1971: Oceanic diffusion diagrams. *Deep-Sea Res.,* **18,** 789-802.

Pinkel, R., 1984: Doppler sonar observations of internal waves: The wavenumber frequency spectrum. *J. Phys. Oceanogr.,* **14,** 1249-1270.

Roache, P.J., 1972: On Artificial Viscosity. *J. Computational Physics,* **10,** 169-184.

Rubenstein, D., 1983: Vertical dispersion of inertial waves in the upper ocean. *J. Geophys. Res.,* **88,** 4368-4380.

Ruddick, B.R., and T.M. Joyce, 1979: Observations of interaction between the internal wavefield and low-frequency flows in the North Atlantic. *J. Phys. Oceanogr.,* **9,** 498-517.

Young, W.R., P.B. Rhines, and C.J.R. Garrett, 1982: Shear-flow dispersion, internal waves and horizontal mixing in the ocean. *J. Phys. Oceanogr.,* **12,** 515-527.

PARAMETERIZING MIXING IN INVERSE MODELS

Trevor J McDougall

CSIRO Division of Oceanography, GPO Box 1538, Hobart, TAS 7001, Australia.

ABSTRACT

The conservation equations for scalars are developed in forms that are suitable for use in inverse models and with emphasis on the ways in which mixing processes are represented. The difference between the lateral Lagrangian and Eulerian velocity vectors is discussed and a suitable parameterization for the difference velocity (the Stokes drift) is proposed. The conservation equation for potential density is developed and is shown to be much more complicated than the normal diapycnal advective-diffusive balance. The extra terms that have not previously been included are often as large as the more obvious terms that have been retained in the past. In order to be able to draw comparisons with previous inverse models, the scalar conservation equations are developed with respect to potential-density surfaces rather than neutral surfaces. The differences that arise here are due to the lateral mixing being directed along neutral surfaces rather than along potential-density surfaces, and again the extra terms can be surprisingly large. The conservation statements for spiciness and Veronicity have their own set of problems since these variables are as nonlinear as is potential density.

The special problems that can beset box inversions are then described, including the need to include vertical diffusion as well as interfacial advection. For the purpose of deducing information about vertical mixing, perhaps the most serious failing of previous box-model inversions has been the masking of the information contained in the scalar conservation equations (particularly the salinity equation) by noise in the continuity equation. This has led previous workers in this field to conclude that vertical mixing was not needed by their inversions even when the results had strong diapycnal advection. A solution is suggested to overcome this problem with the divergence form of the conservation statements.

INTRODUCTION

Due to increasing societal pressure to predict the impact of the Greenhouse Effect on future climate, the role of the ocean in climate change is becoming more widely acknowledged and the subject of greater study. On the timescales of decades to centuries, the intermediate and deep waters of the ocean play a crucial role in the complicated feedbacks between the ocean and the atmosphere, and the circulation at these depths is intimately linked to the existence of vertical mixing processes (Bryan, 1987). We do not yet know how the strength of this mixing is regulated nor whether it is widely distributed in space or localized in small regions such as near topography. For these reasons it is increasingly important to use all methods at our disposal to deduce the strength of vertical mixing processes. One such method is to apply inverse procedures to hydrographic and tracer data to infer the strength of mixing processes that must have been present to cause the observed three-dimensional tracer patterns.

Much of the effort on oceanographic inverse models has been directed to determining the time-averaged circulation of the ocean, and so the models have concentrated on using the thermal wind relation in

conjunction with the continuity equation or the conservation of potential vorticity to solve for the lateral velocity vector for every station-pair at some (deep) reference level. Much less effort has been spent on using these techniques to deduce the magnitude of mixing processes in general, and the vertical diffusivity in particular. Three recent papers that have made substantial progress in this direction are Olbers, Wenzel and Willebrand (1985) and Hogg (1987) that have both been based on the β-spiral technique, and Schlitzer (1988) that has used the box-inverse technique. Another paper that was not only successful in detecting oceanic mixing, but was also able to distinguish between two different types of vertical mixing processes (salt-fingering and vertical eddy diffusion), is that by Bauer and Siedler (1988). This paper used the form of the conservation equations that did not explicitly include dianeutral advection, and so ensured that there could be no confusion in their model between dianeutral advection and dianeutral diffusion. While their work is not a full inversion in the sense that it did not solve for the velocity components simultaneously with the mixing parameters, I hope that it is typical of the results we can expect when inverse models are set up specifically with mixing processes in mind.

There is a distinct advantage in determining vertical diffusivities from hydrographic data as compared with determining the strength of lateral mixing processes. This advantage is due to the large separation of vertical scales between the small vertical scale at which the mixing processes act (of order 1 m), and the large vertical scale on which either an inverse model or a forward model is constructed (of order 100 m). Because of this large separation of scales we can expect that any answer that we believe for the vertical diffusivity from an inverse model will also apply to a forward model since the vertical scales in both models are so much larger than the scales of the mixing processes. A similar favorable ratio of relevant scales does not apply to lateral mixing processes so that one cannot be sure that a lateral eddy diffusivity that one obtains from an inversion using data at a certain horizontal spacing will be applicable to a forward model that has its grid points at a different horizontal resolution.

The paper begins with a review of the conservation equations for scalars in the neutral surface framework and goes on to derive a parameterization for the lateral bolus transport or Stokes velocity. Then, in order to compare what has been done by previous inverse modellers, the conservation equations for scalars are derived when mixing is performed with respect to potential-density surfaces. Many extra terms arise due to the different slopes and curvatures (in x,y,z space) of potential-density surfaces and neutral surfaces. These terms are often not negligible. An additional set of problems can arise when using the divergence form of the conservation equations, as is necessary in box-model inversions. The straightforward use of a salinity conservation equation in these box models seldom finds mixing processes to be important, even though the β-spiral techniques, using data from the same region, are able to detect the influence of ocean mixing. We show that this is most likely because the salinity conservation equation is dominated by uncertainty in the net volume flux into or out of the box. A suggestion is offered to overcome this problem, thereby increasing the rank of the solution and allowing mixing information to be recovered from the inversion.

A REVIEW OF SUITABLE CONSERVATION EQUATIONS FOR INVERSE MODELS

Since the lateral mixing by mesoscale eddies is believed to occur along neutral surfaces, it is convenient to develop the conservation equations in this reference frame. The vertical turbulent property fluxes are then parameterized by a vertical diffusivity, D, acting on the vertical property gradients, but the lateral turbulent fluxes are parameterized separately on the basis of a further Reynolds decomposition performed in this section. The continuity equation, $\nabla \cdot \mathbf{u} = 0$, and the divergence form of the conservation statement for a tracer, C, are integrated over the thickness, h, between two neutral surfaces (see Figure 1), to obtain

$$\frac{\partial h}{\partial t} + \bar{\nabla}_n \cdot (h\mathbf{V}) + e^u - e^\ell = 0 \tag{1}$$

and

$$\frac{\partial (hC)\big|_n}{\partial t} + \nabla_n \cdot (h\mathbf{V}C) + [eC]_\ell^u = [DC_z]_\ell^u . \tag{2}$$

Here C is any conservative scalar variable (including S and θ), e is the vertical velocity of fluid through the neutral surface (henceforth called the dianeutral velocity), $\mathbf{V}$ is the instantaneous two-dimensional horizontal velocity vector, and ∇_n is the two-dimensional lateral gradient operator for properties measured in the neutral surface. The exact nature of the geometrical transformation between the Cartesian and neutral surface reference frames is explained in McDougall (1987a). The superscripts u and ℓ refer to the upper and lower neutral surfaces, as in Figure 1.

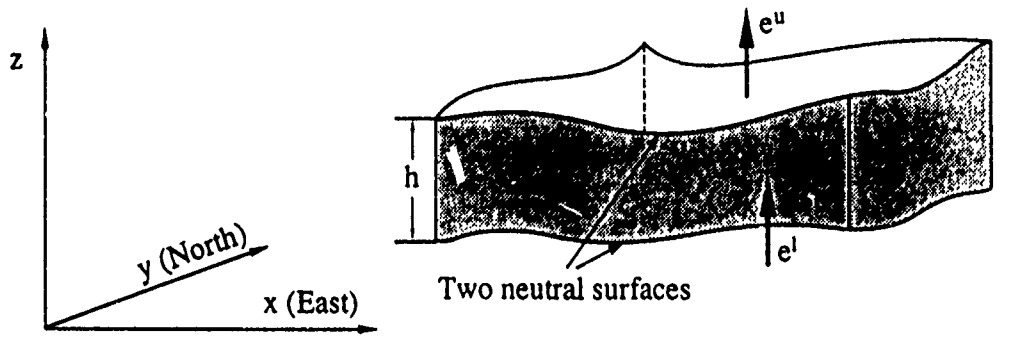


Fig. 1. Sketch showing two neutral surfaces separated by a height $h(x,y)$. Mixing processes can cause fluid to flow vertically through these neutral surfaces with the velocities e^u and e^ℓ.

Now we specifically allow for the lateral eddy motions by performing a Reynolds decomposition on the lateral velocity vector and on the scalar fields. The quasi-Lagrangian velocity vector is defined as

$$\bar{\mathbf{V}}^L \equiv \bar{\mathbf{V}}^E + \frac{\overline{h'\mathbf{V}'}}{\bar{h}}, \tag{3}$$

(called "quasi-Lagrangian" because, while $\bar{\mathbf{V}}^L$ follows the average position of marked fluid parcels along a neutral surface, it doesn't follow the dianeutral motion). $\bar{\mathbf{V}}^E$ is the usual Eulerian average velocity vector as would be measured by a current meter at a fixed position. Reynolds decomposing all the variables in (1) and (2) and averaging yields

$$\frac{\partial \bar{h}}{\partial t} + \nabla_n \cdot \left(\bar{h}\bar{\mathbf{V}}^L\right) + e^u - e^\ell = 0 \tag{4}$$

and

$$\frac{\partial \left(\bar{h}\bar{C}\right)\big|_n}{\partial t} + \frac{\partial \left(\overline{h'C'}\right)\big|_n}{\partial t} + \nabla_n \cdot \left(\bar{h}\bar{\mathbf{V}}^L\bar{C}\right) + \left[e\bar{C}\right]_\ell^u = -\nabla_n \cdot \left\{\bar{h}\,\overline{C'\mathbf{V}'} + \overline{h'C'}\,\bar{\mathbf{V}}^E + \overline{h'C'\mathbf{V}'}\right\} + \left[D\bar{C}_z\right]_\ell^u . \tag{5}$$

Note that the mean value of C that appears in Eq. (5) is the Eulerian average value, not the quasi-Lagrangian value that would be defined in a similar way to Eq. (3). We are very likely to be justified in

assuming that the triple correlation term on the right-hand side of Eq. (5) is quite small, and probably also $\overline{h'C'}\,\overline{\mathbf{V}}^E << \overline{C'\mathbf{V}'h}$ (Rhines, 1982) but in any case, we will assume that the lateral diffusivity, K, of a scalar is defined by

$$-\bar{h}K\nabla_n\bar{C} \equiv \left\{\bar{h}\,\overline{C'\mathbf{V}'} + \overline{h'C'}\,\overline{\mathbf{V}}^E + \overline{h'C'\mathbf{V}'}\right\}. \tag{6}$$

If the term $\partial\left[\overline{h'C'}|_n\right]/\partial t$ is taken to be zero, the divergence form, Eq. (5) simplifies to

$$\boxed{\frac{\partial(hC)|_n}{\partial t} + \nabla_n\cdot\left(h\mathbf{V}^L C\right) + [eC]^u_\ell = \nabla_n\cdot\left(hK\nabla_n C\right) + \left[DC_z\right]^u_\ell,} \tag{7}$$

where the overbars have been dropped from all the variables, while the continuity equation is

$$\boxed{\frac{\partial h}{\partial t} + \nabla_n\cdot\left(h\mathbf{V}^L\right) + [e]^u_\ell = 0.} \tag{8}$$

The advective form of the conservation equation for C is found by letting h tend to zero so that $[eC]^u_\ell \to h[eC]_z$, $[e]^u_\ell \to he_z$, and $\left[DC_z\right]^u_\ell \to h\left[DC_z\right]_z$, and combining Eq. (7) and Eq. (8), so that

$$\boxed{C_t|_n + \mathbf{V}^L\cdot\nabla_n C + eC_z = h^{-1}\nabla_n\cdot\left(hK\nabla_n C\right) + \left[DC_z\right]_z.} \tag{9}$$

This conservation statement (and also the divergence form, Eq. (7)) holds for any conserved scalar quantity including salinity and potential temperature. The conservation statements for potential vorticity take different forms since potential vorticity is not mixed vertically in the same manner as a conservative tracer. Potential vorticity conservation statements are not discussed in this paper.

The spatial gradients of θ and S along a neutral surface and also their temporal derivatives on a neutral surface are related through the thermal expansion coefficient, α, and the saline contraction coefficient, β, so that

$$\alpha\theta_t|_n = \beta S_t|_n, \quad \text{and} \quad \alpha\nabla_n\theta = \beta\nabla_n S. \tag{10}$$

Using these relations and the advective conservation statements for S and θ, Eq. (9), one finds

$$\boxed{\begin{aligned}\theta_t|_n + \left[\mathbf{V}^L - h^{-1}\nabla_n(hK)\right]\cdot\nabla_n\theta &= K\nabla_n^2\theta + DgN^{-2}\theta_z^3\beta\frac{d^2S}{d\theta^2} \\ &\quad + KgN^{-2}\theta_z\left\{C_b\nabla_n\theta\cdot\nabla_n\theta + T_b\,\nabla_n\theta\cdot\nabla_n p\right\},\end{aligned}} \tag{11}$$

and

$$\boxed{\left[e - D_z\right]\left[\alpha\theta_z - \beta S_z\right] = D\left[\alpha\theta_{zz} - \beta S_{zz}\right] - K\left\{C_b\nabla_n\theta\cdot\nabla_n\theta + T_b\nabla_n\theta\cdot\nabla_n p\right\},} \tag{12}$$

where C_b and T_b are the cabbeling and thermobaric parameters respectively, defined by

$$C_b = \left[\frac{\partial\alpha}{\partial\theta} + 2\frac{\alpha}{\beta}\frac{\partial\alpha}{\partial S} - \frac{\alpha^2}{\beta^2}\frac{\partial\beta}{\partial S}\right] \quad \text{and} \quad T_b = \left[\frac{\partial\alpha}{\partial p} - \frac{\alpha}{\beta}\frac{\partial\beta}{\partial p}\right].$$

Typical values are $10^{-5}\ \mathrm{K}^{-2}$ for the cabbeling parameter and $2.6\times10^{-8}\ \mathrm{K}^{-1}\ (\mathrm{db})^{-1}$ for the thermobaric parameter.

If one were able to measure the Lagrangian rate of change of a fluid property, say potential temperature, following a fluid parcel, vertical mixing with diffusivity D would cause the parcel's potential temperature to change at the rate $\left[D\theta_z\right]_z$. However, since we are only able to observe the ocean after a series of mixing processes has occurred, this Lagrangian approach cannot be pursued. By using the neutral surface reference frame, the vertical advection caused by vertical mixing can be accurately identified (Eq. 12) and then eliminated from the conservation equations in the form of Eq. (11).

Vertical mixing processes affect the conservation statement both through the non-advective diffusion of tracer and also through dianeutral advection. The ratio of dianeutral diffusion to dianeutral advection is, in general, different for different tracers.

It is important to realize that the dianeutral advection, e, is not a separate mixing process in its own right. Rather, e occurs simply because of the vertical diffusivity, D, and also due to cabbeling and thermobaricity. Once these physical mixing processes are specified and the hydrographic fields are known, e follows directly from Eq. (12). Eq. (11) is the appropriate form of the θ conservation equation for water-mass analysis since such studies are performed in property-property space and are inherently not fully Lagrangian. Vertical mixing by small-scale (isotropic) turbulence causes observable changes of fluid properties only if the relevant property-property plots are curved. Eq. (11) is sometimes called a water-mass transformation equation, since the temporal change of potential temperature on a neutral surface is equivalent to a change of the S-θ curve of a water-mass.

The reason for concentrating on the advective form of the conservation equations in neutral surfaces is because scalar properties are usually measured with high precision and so one can detect quite small changes in property-property diagrams (like the S-θ curves), and a change in such a property-property diagram is equivalent to a change on some kind of density horizon such as a neutral surface. In their divergence forms, the conservation equations contain not only the information of how mixing processes change properties along neutral surfaces, but they are also affected by the temporal change of the volume of the control volume, and the uncertainty in the volume fluxes into and out of a control volume bounded by two neutral surfaces. This volume is never well known in oceanographic observations because internal waves and mesoscale eddies cause large vertical excursions of neutral surfaces and so cause a large temporal variability of the height between neutral surfaces. However, these same wave-like processes do not change the property-property diagrams at all, and so these diagrams and the advective conservation equations in the neutral surface framework contain information that is sensitive to mixing processes and insensitive to the vertical heaving motions caused by internal waves and the like.

A simple scale analysis of the potential temperature conservation equation, Eq. (11), can reveal the magnitude of cabbeling and thermobaricity in relation to the epineutral mixing of θ by the lateral Laplacian term. At a thermoclinic front the magnitude of the epineutral Laplacian scales as the lateral gradient divided by the half-width of the front, L, that is, $|\nabla_n^2\theta| \approx |\nabla_n\theta| / L$. Setting half the epineutral θ and p contrasts across the front equal to $\Delta\theta$ and Δp respectively ($\Delta\theta = L|\nabla_n\theta|$ and $\Delta S = L|\nabla_n S|$), we find from Eq. (11) that

$$KgN^{-2}\theta_z\left\{C_b\nabla_n\theta\cdot\nabla_n\theta + T_b\,\nabla_n\theta\cdot\nabla_n p\right\}\Big/\left(K\nabla_n^2\theta\right) = \frac{R_\rho}{\left[R_\rho - 1\right]}\left\{\frac{C_b}{\alpha}\Delta\theta \pm \frac{T_b}{\alpha}\Delta p\right\}$$

$$\approx \frac{R_\rho}{\left[R_\rho - 1\right]}\left\{0.1\Delta\theta \pm 3x10^{-4}\Delta p\right\}. \qquad (13)$$

In the Antarctic Circumpolar Current where $R\rho$ is about 2, $\Delta\theta$ is about 1 K, and Δp is about 500 db (and in the same sense as $\Delta\theta$), Eq. (13) is about 0.5, implying that the peak contribution of thermobaricity and cabbeling to water-mass conversion is 50% of the peak contribution of the epineutral Laplacian term. However thermobaricity and cabbeling are much more important than this comparison suggests because the epineutral Laplacian term changes sign across the front and so averages to zero in the frontal region, while cabbeling and thermobaricity contribute a term of the same sign across the whole front.

SUGGESTED PARAMETERIZATION FOR THE BOLUS TRANSPORT

The conservation equations that have been developed above have all included the bolus transport, $\overline{h'\mathbf{V}'}$, of volume along neutral surfaces. This is equivalent to recognizing that in both the advective and divergence forms of tracer conservation statements, the relevant velocity vector is the Lagrangian-mean velocity, not the Eulerian-mean velocity (Rhines, 1986, p 121). The difference between Lagrangian and Eulerian velocities in the dianeutral direction does not arise since the coordinate frame is already quasi-Lagrangian in this direction. The prefix quasi- is used to indicate that the coordinate frame does not quite move with the fluid: the difference being the dianeutral velocity, e. This means that the dianeutral velocity, e, is inherently a Lagrangian velocity component with respect to neutral surface coordinates. The lateral velocity vector that a current meter would measure at a fixed point is the Eulerian lateral velocity, $\mathbf{V}^E$, whereas the mean lateral velocity of a patch of dye or a cluster of floats is the Lagrangian velocity, $\mathbf{V}^L$, which is the sum of the Eulerian velocity and the Stokes drift, $h^{-1}\overline{h'\mathbf{V}'}$. In Cartesian coordinates, this bolus transport arises as a skew diffusion tensor when the diffusion tensor contains antisymmetric terms (Haidvogel and Rhines (1983), Rhines (1986), and Middleton and Loder (1989)).

The geostrophic and thermal wind equations relate the Eulerian velocity to the pressure and in-situ density fields, while the above equations show that it is the Lagrangian velocity vector that appears in the continuity and tracer equations (in both the advective and the divergence forms). To date inverse models have not distinguished between the Eulerian and Lagrangian lateral velocity vectors, and we do not yet know whether the distinction between these velocity vectors is significant in the ocean. In this section a parameterization is suggested for the difference velocity, the Stokes drift, $\mathbf{V}^S = \mathbf{V}^L - \mathbf{V}^E$. By trying this parameterization in inverse models we may be able to deduce the importance of the Stokes drift in the conservation of tracers.

In a quasi-geostrophic layered eddy-resolving model, Holland and Rhines (1980) have found that the eddy flux of potential vorticity is directed down the lateral gradient of mean potential vorticity, and that this eddy flux is dominated by the thickness flux. The only exception to this was in the western boundary current where the eddy relative vorticity flux was larger than that due to the bolus transport. The eddy flux of relative vorticity was much smaller, had smaller spatial scales and was both up and down-gradient at different locations. Brown, Owens and Bryden (1986) have used the LDE (Local Dynamics Experiment) current meter data of POLYMODE to determine both the eddy flux of relative vorticity and the thickness flux of potential vorticity at depth of about 600 m.. They found that the thickness flux of potential vorticity was an order of magnitude larger than the eddy flux of relative vorticity, and that the thickness flux of potential vorticity was directed almost exactly down the large-scale lateral gradient of potential vorticity (i.e. in the sense of a positive diffusivity, that is, in the direction of $-\nabla_\sigma\left(f/h^\sigma\right)$). The lateral diffusivity of potential vorticity that is implied by their measurements is about 2600 $m^2\ s^{-1}$. Since $\overline{q'\mathbf{V}'} \approx -qh^{-1}\overline{h'\mathbf{V}'}$, the Stokes velocity, $h^{-1}\overline{h'\mathbf{V}'}$, was directed up the epineutral gradient of potential vorticity. Their observed thickness flux of potential vorticity is equivalent to a Stokes velocity, $h^{-1}\overline{h'\mathbf{V}'}$, of about 1.5 mm s^{-1}.

The observation from both quasi-geostrophic ocean models and from the LDE measurements of Brown et al (1986) that the thickness flux of potential vorticity is directed down the lateral q gradient suggests a parameterization for the eddy flux of potential vorticity as being minus a diffusivity times the epineutral gradient of potential vorticity. Because of the difficulty in Reynolds decomposing $q = f / h$, (since h is in the denominator), and because $\bar{h}$ and h' do arise unavoidably in the continuity equation, I denote the reciprocal of q as a new variable, r, and assert that the thickness flux of r is directed down its lateral gradient, so that

$$\overline{r'\mathbf{V}'} = -B\nabla_n \bar{r} \;=\; -B\nabla_n\left[\bar{h}/f\right], \text{ where } r \equiv h/f,\;\; \bar{r} = \bar{h}/f,\;\text{ and } r' = h'/f. \tag{14}$$

For small amplitude perturbations, this is equivalent to the more obvious $\overline{q'\mathbf{V}'} = -B\nabla_n \bar{q}$ parameterization, but at large amplitude when $h'/\bar{h}$ is not small, it is not clear how to link q' and $\bar{q}$ to h' and $\bar{h}$. With this parameterization of the epineutral flux of r due to eddies, the bolus flux is given by

$$\boxed{\mathbf{V}^S = \mathbf{V}^L - \mathbf{V}^E = \bar{h}^{-1}\overline{h'\mathbf{V}'} = B\nabla_n\left[\ell n\, \breve{q}\right] \;= -B\nabla_n\left[\ell n\, \bar{h}\right] + \left[\beta/f\right]B\mathbf{j},} \tag{15}$$

where $\breve{q} \equiv \bar{r}^{-1} = f/\bar{h}$. The overbars will henceforth be omitted from averaged variables. Since the magnitude of the neutral-surface potential vorticity generally increases towards the poles, the Stokes velocity will often be directed poleward.

The lateral advection term in the scalar conservation equations contains the combination of terms, $\mathbf{V}^L - h^{-1}\nabla_n(hK)$, and this can be most readily expressed as (using the geometrical identity $\nabla_n[\ell n h] = \partial[\nabla_n \mathcal{N}]/\partial z$ where $\mathcal{N}[x,y]$ is the height of a neutral surface)

$$\mathbf{V}^L - h^{-1}\nabla_n(hK) = \mathbf{V}^E - [B+K]\partial[\nabla_n\mathcal{N}]/\partial z - \nabla_n K + [\beta/f]B\mathbf{j}. \tag{16}$$

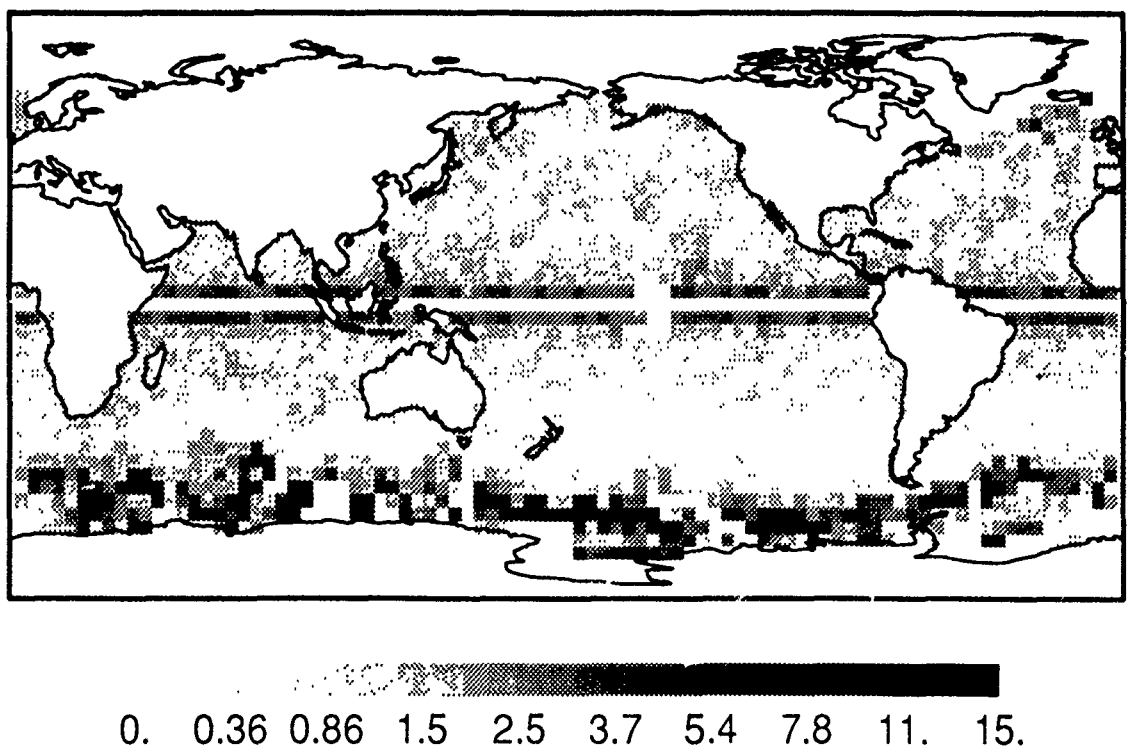


Fig. 2. Map of the magnitude of the lateral Stokes velocity, $\mathbf{V}^S$, as given by Eq. (15), on a neutral surface that lies at about 1200 m in the equatorial Pacific (where σ_θ is about 27.5), and with the lateral diffusivity, B, equal to 1000 m^2 s^{-1}. The grey-scale covers the range from 0-15 mm s^{-1} and is evenly spaced in $\ell n(1 + |\mathbf{V}^S|)$.

Following Rhines and Holland (1979) we take the lateral diffusivity of potential vorticity to be equal to that of passive scalars, that is, $B = K$, so that

$$\boxed{\left[\mathbf{V}^L - h^{-1}\nabla_n(hK)\right] = \mathbf{V}^E - 2K\partial[\nabla_n\mathcal{N}]/\partial z - \nabla_n K + [\beta/f]K\mathbf{j}.} \tag{17}$$

This is the effective lateral velocity that advects tracers, whereas $\mathbf{V}^E$ is the lateral velocity that is evaluated from the thermal wind equation plus the reference level velocity.

The difference between Lagrangian and Eulerian lateral velocity vectors has not yet been incorporated into any inverse model of the ocean circulation. Rather, inverse models to date have assumed that the same lateral velocity vector appears in the tracer conservation equations as in the thermal wind equation. In fact, we have very little insight from either oceanographic theory or observations about the magnitude of the bolus transport. The deduction of a Stokes drift of 1.5 mm s^{-1} from the work of Brown et al (1986) is a rare insight into the importance of this process.

Since the contours of potential vorticity and of other tracers are often nearly parallel (although with the gradients often pointing in opposite directions), the Stokes drift or bolus transport can be expected to have a substantial component *across* the mean epineutral gradients of other tracers (including S and θ), and its direction will be up the epineutral tracer gradients as often as down them. Because the lateral velocity vector generally points almost along epineutral isolines of tracers rather than down the epineutral tracer gradients, the Stokes drift will be much more important for the conservation of tracers than a simple velocity estimate of say 1.5 mm s^{-1} would suggest. For example, a lateral velocity of 1 mm s^{-1} down an epineutral potential temperature gradient of 10^{-6} K m^{-1} (one degree of potential temperature change in 10° of latitude) creates the same amount of water-mass conversion as a dianeutral velocity of 10^{-7} m s^{-1} acting on a vertical potential temperature gradient of 10^{-2} K m^{-1}. Since this dianeutral velocity is as large as that expected from vertical mixing processes, it is concluded that the Stokes drift is likely to be important in tracer balances in the ocean.

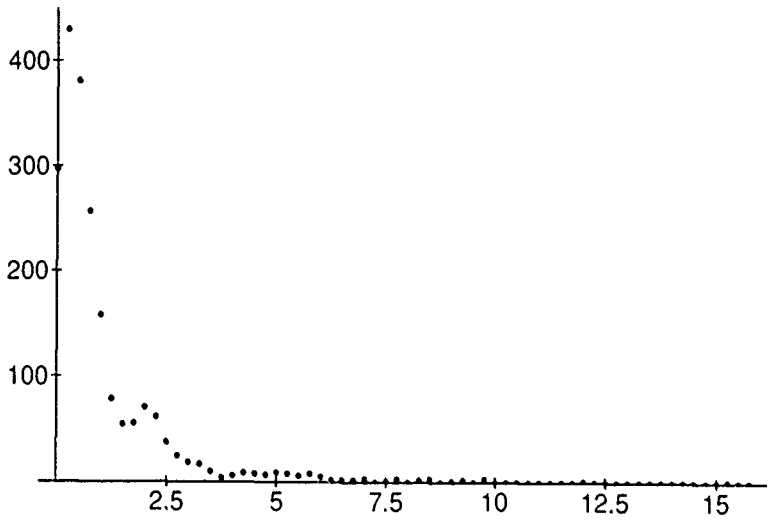


Fig. 3. Histogram of the number of occurrences of various magnitudes of the Stokes velocity (from 0 to 15 mm s^{-1}) in the map of Figure 2.

The Levitus (1982) atlas has been used to generate two neutral surfaces throughout the world ocean that were at depths of 1100 m and 1200 m respectively in the equatorial Pacific. The height, h, between them was used in Eq. (15) to estimate the Stokes velocity using 1000 $m^2\ s^{-1}$ for the lateral diffusivity. The magnitude of this estimate of the Stokes drift is shown in Figure 2 and the distribution of these estimates is shown in the histogram of Figure 3 where it is seen that a value of order 1 mm s^{-1} is typical. It remains to be seen whether the incorporation into inverse models of the difference between the lateral velocity vectors in the tracer equations and in the thermal wind equation, using the above parameterization Eq. (15) for the Stokes drift, makes any significant differences.

MIXING PRESCRIPTIONS IN β-SPIRAL INVERSIONS: The Advective Forms of Conservation Statements in Potential-Density Surfaces

Inversions of the β-spiral type are local in nature, in that a solution is found for the reference-level lateral velocity components and the mixing coefficients from the information at just one location. The inversion procedure can be cast in terms of the slopes of the surfaces of the conserved quantity (Schott and Stommel, 1978) or in terms of the conservation equation of this scalar quantity (e.g. Olbers et al., 1985). In either case a conservation statement is written in the advective form as opposed to the flux, or divergence, form. Here the equation for the conservation of potential density is derived and compared with that used by previous researchers. Important differences are found. Then the conservation equations of potential temperature and spiciness will be discussed.

The Potential Density "Conservation" Equation

In a potential-density surface the temporal and lateral gradients of salinity and potential temperature are related by (McDougall, 1987b)

$$\tilde{\alpha}\theta_t\big|_\sigma = \tilde{\beta} S_t\big|_\sigma, \text{ and } \tilde{\alpha}\nabla_\sigma\theta = \tilde{\beta}\nabla_\sigma S, \tag{18}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are the values of the thermal expansion and saline contraction coefficients evaluated at the reference pressure of the potential density variable, that is, $\tilde{\alpha} = \alpha(S,\theta,p_r)$ and $\tilde{\beta} = \beta(S,\theta,p_r)$. The material derivative in the advective form, Eq. (9), of the tracer conservation equation can be written with respect to either neutral surfaces or potential-density surfaces, so that

$$C_t\big|_n + \mathbf{V}^L\cdot\nabla_n C + eC_z = C_t\big|_\sigma + \mathbf{V}^L\cdot\nabla_\sigma C + w^d C_z = h^{-1}\nabla_n\cdot(hK\nabla_n C) + \left[DC_z\right]_z, \tag{19}$$

where w^d is the vertical velocity of fluid through the potential-density surface (henceforth called the diapycnal velocity). This equation applies for both potential temperature and salinity, and these equations can be multiplied by $\tilde{\alpha}$ and $\tilde{\beta}$ respectively and then subtracted so as to eliminate the temporal and isopycnal advective terms, obtaining (after rearrangement) the following equation for the diapycnal velocity , (since $\tilde{\beta}S_t\big|_\sigma = \tilde{\alpha}\theta_t\big|_\sigma$, $\tilde{\beta}\nabla_\sigma S = \tilde{\alpha}\nabla_\sigma\theta$, and $(\rho_\theta)^{-1}(\rho_\theta)_z = \tilde{\beta}S_z - \tilde{\alpha}\theta_z$, see McDougall, 1987b),

$$\begin{aligned} w^d\left[\tfrac{1}{\rho_0}\tfrac{\partial\rho_0}{\partial z}\right] = {} & \left[D\tfrac{1}{\rho_0}\tfrac{\partial\rho_0}{\partial z}\right]_z \\ & + D\tfrac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2 \\ & + \tilde{\alpha}(c-1)\tfrac{1}{h}\nabla_n\cdot(hK\nabla_n\theta) + \tfrac{\tilde{\beta}}{\beta}K\{C_b\nabla_n\theta\cdot\nabla_n\theta + T_b\nabla_n\theta\cdot\nabla_n p\}. \end{aligned} \tag{20}$$

This is the full conservation equation for potential density. The first term on the right-hand side is the only one that would occur if potential density were a conservative variable. If so desired, the left-hand side can be written with respect to Cartesian coordinates (as in Olbers et al (1985)) as

$$\left.\frac{\partial \rho_\theta}{\partial t}\right|_z + \mathbf{V}^L \cdot \nabla_2 \rho_\theta + w \left.\frac{\partial \rho_\theta}{\partial z}\right|_{x,y} = w^d \left.\frac{\partial \rho_\theta}{\partial z}\right|_{x,y} . \qquad (21)$$

The usual approach with the potential-density equation is to equate the diapycnal advection of potential density, $w^d(\partial\rho_\theta/\partial z)$, to (minus) the diapycnal flux divergence, $\partial[D\,\partial\rho_\theta/\partial z]/\partial z$. That is, normally one sees the left-hand side of Eq. (20) being balanced by only the first term on the right. Here we investigate the magnitudes of the other terms that are usually neglected. The second term on the right of Eq. (20) arises because while the vertical gradient of potential density is given by $-(\rho_\theta)^{-1}(\rho_\theta)_z = \tilde{\alpha}\theta_z - \tilde{\beta}S_z$, when casting the right-hand side of Eq. (20) in terms of $(\rho_\theta)_{zz}$, the extra term, $D(\partial\tilde{\alpha}/\partial\theta)\theta_z^2$ appears. (The relatively small variation of $\tilde{\alpha}$ with S and the variation of $\tilde{\beta}$ with both θ and S have been ignored.) McDougall and You (1990) have shown that this extra term is larger than the correct term, $-D\left[\tilde{\alpha}\theta_{zz} - \tilde{\beta}S_{zz}\right]$, in much of the subtropical thermoclines of each of the world's ocean basins (see their Figure 9). This is illustrated in Figures 4-6 which show three different combinations of terms mapped on a neutral surface in each of the three major oceans. On much of the neutral surface of Figure 4 in the North Atlantic, the correction term in Eq. (20), $D\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2$, is larger than the sum of the first two terms in Eq. (20), $-D\left[\tilde{\alpha}\theta_{zz} - \tilde{\beta}S_{zz}\right]$, (ignoring D_z), implying that the straightforward term in Eq. (20), $D\left[(\rho_\theta)^{-1}\partial\rho_\theta/\partial z\right]$, has the opposite sign to the correct right-hand side of Eq. (20) in these places. In the Pacific in Figure 5 the nonlinear term is up to 40% of the correct term, while in the Indian Ocean in Figure 6 it is as large as 30% of the correct term. It is concluded that serious errors have occurred in the conservation equation of potential density in present inverse models. From these figures it can be seen that in Hogg's (1987) study of the circulation and mixing in the central North Atlantic, the second term in Eq. (20), $D(\partial\tilde{\alpha}/\partial\theta)\theta_z^2$, which Hogg omitted, was 50% of the correct term, $D\left[\tilde{\alpha}\theta_{zz} - \tilde{\beta}S_{zz}\right]$, on the upper of the two σ_1 surfaces he considered. In parts of the North Atlantic, the relative error is more than 100%, and this type of error was built into the Olbers et al (1985) model.

In the third term in Eq. (20), the parameter c is defined by $c = [\alpha/\beta]\left[\tilde{\alpha}/\tilde{\beta}\right]^{-1}$ and is unity at the reference pressure, p_r, of the potential density. This third term in the equation scales as $K\tilde{\alpha}[c-1]\nabla_n^2\theta$ and the magnitude of this term can be estimated as follows. Consider the use of potential density (referenced to the sea surface) down to a depth of 2000 m (as in Olbers et al., 1985). At that depth, c is about 1.5 and $K\nabla_n^2\theta$ can be taken to be the same order as $e\theta_z$ where the dianeutral velocity is 1×10^{-7} m s^{-1} and the vertical gradient of potential temperature is 3×10^{-3} K m^{-1}. When divided by the vertical gradient of potential density at 2000 m in the North Atlantic, this third term in Eq. (20) amounts to a contribution of 0.3×10^{-7} m s^{-1} to the diapycnal velocity. Again, this is not an insignificant diapycnal velocity, although it is rather unusual these days to find researchers using potential-density surfaces as far as 2000 m away from their reference pressure.

The last term in Eq. (20) is due to thermobaricity and cabbeling, (see the corresponding terms in the dianeutral velocity equation, Eq. (12)). Because the buoyancy frequency is related to the vertical gradient of potential density by $N^2 = \mu\left(\beta/\tilde{\beta}\right)\left[-(g/\rho_\theta)(\partial\rho_\theta/\partial z)\right] \approx \mu\left[-(g/\rho_o)(\partial\rho_\theta/\partial z)\right]$, (see McDougall, 1987b and Eq. (23) below for a definition of μ) the *diapycnal* velocity induced by cabbeling and thermobaricity will in fact be larger by the factor μ than the *dianeutral* velocity, e, caused by these processes. McDougall and You (1990) assumed a lateral diffusivity, K, of 1000 m^2 s^{-1} and have shown that in parts of the North Atlantic and in the Southern Ocean, cabbeling causes a downward dianeutral velocity as large as -1×10^{-7} m s^{-1}, so that the contribution of the last term in

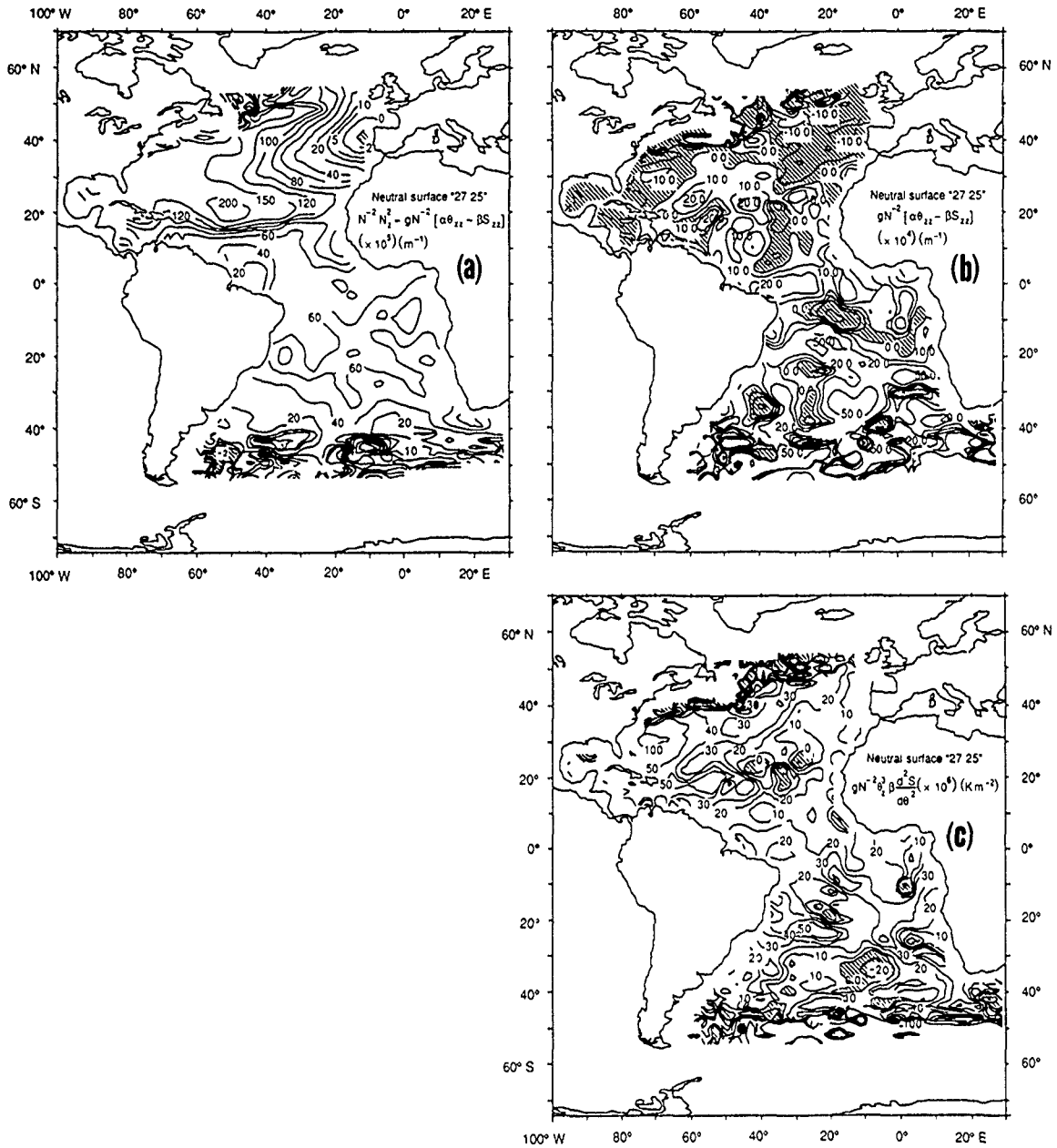


Fig. 4. Graphs of various terms on a neutral surface in the Atlantic at a depth of approximately 800 m in mid-latitudes. (a) The term plotted here, $N^{-2}N_z^2 - gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, is approximately equal to $gN^{-2}\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^3$ in the upper 1000 m of the ocean. (b) When multiplied by the vertical diffusivity, D, this term, $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, is the rate at which vertical mixing contributes to the dianeutral velocity. (c) When multiplied by the vertical diffusivity, D, this term, $gN^{-2}\theta_z^3\beta\frac{d^2S}{d\theta^2}$, is the rate at which vertical turbulent mixing causes potential temperature to change on a neutral surface. (From McDougall and You (1990) Figure 1).

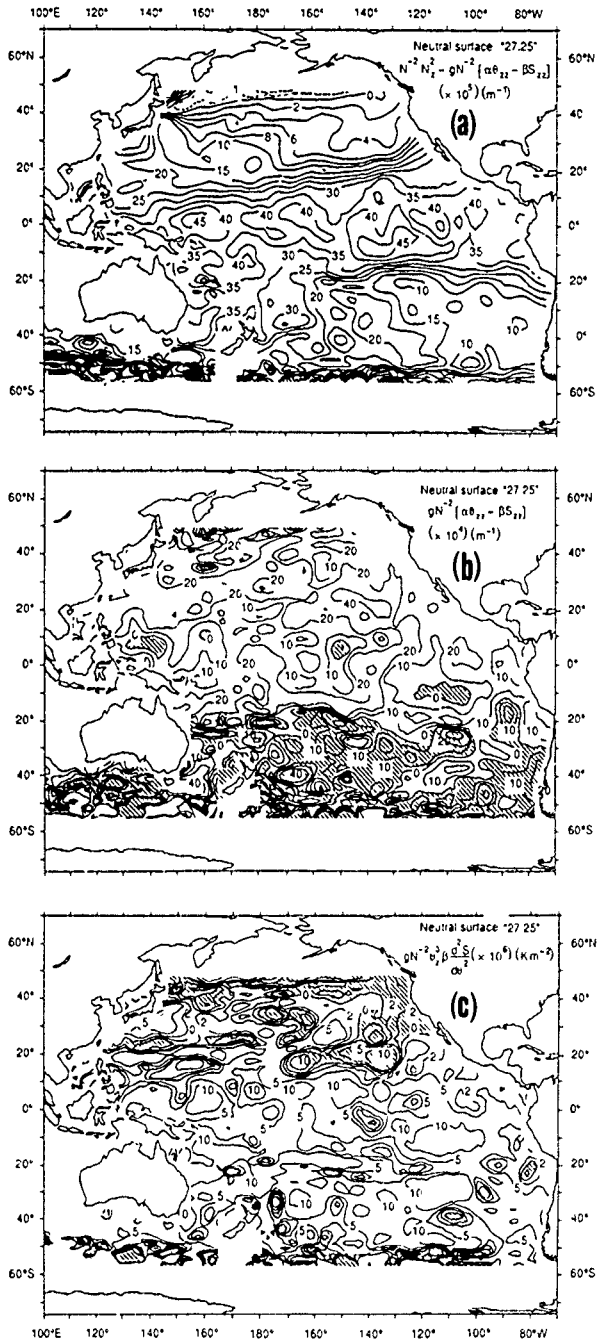


Fig. 5. Graphs of various terms on a neutral surface in the Pacific at a depth of approximately 900 m in mid-latitudes. (a) The term plotted here, $N^{-2}N_z^2 - gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, is approximately equal to $gN^{-2}\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^3$ in the upper 1000 m of the ocean. (b) When multiplied by the vertical diffusivity, D, this term, $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, is the rate at which vertical mixing contributes to the dianeutral velocity. (c) When multiplied by the vertical diffusivity, D, this term, $gN^{-2}\theta_z^3\beta\frac{d^2S}{d\theta^2}$, is the rate at which vertical turbulent mixing causes potential temperature to change on a neutral surface. (From McDougall and You (1990) Figure 4).

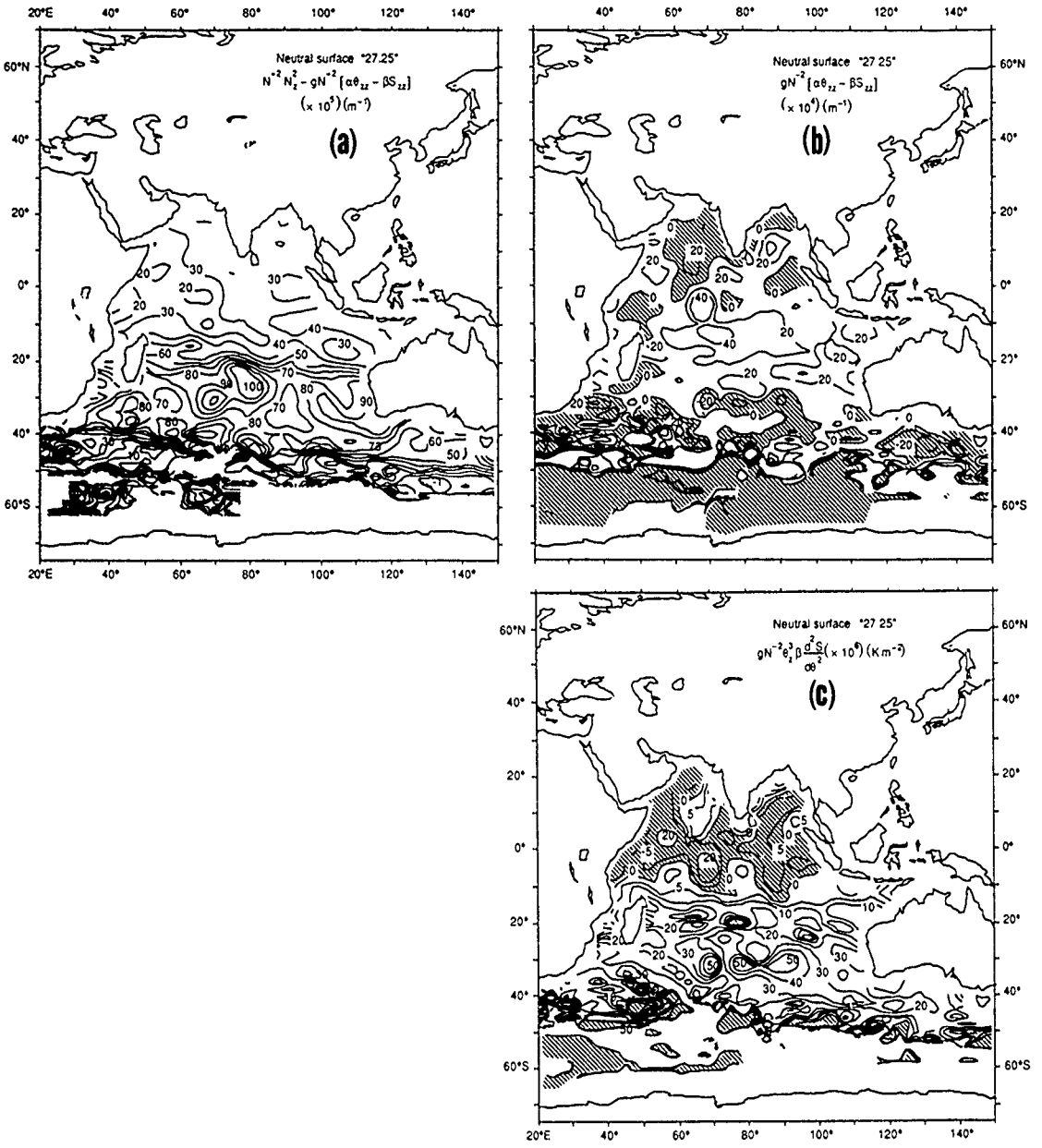


Fig. 6. Graphs of various terms on a neutral surface in the Indian at a depth of approximately 950 m in mid-latitudes. (a) The term plotted here, $N^{-2}N_z^2 - gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, is approximately equal to $gN^{-2}\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^3$ in the upper 1000 m of the ocean. (b) When multiplied by the vertical diffusivity, D, this term, $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, is the rate at which vertical mixing contributes to the dianeutral velocity. (c) When multiplied by the vertical diffusivity, D, this term, $gN^{-2}\theta_z^3\beta\frac{d^2S}{d\theta^2}$, is the rate at which vertical turbulent mixing causes potential temperature to change on a neutral surface. (From McDougall and You (1990) Figure 6).

Eq. (20) to the diapycnal velocity will be larger than this in these locations. The neglect of this cabbeling term in their potential density conservation equation would have caused significant errors (of order -1×10^{-7} m s^{-1} in the diapycnal velocity) in Hogg (1987) and in Olbers et al. (1985). Following the results of McDougall and You (1990), thermobaricity is as large as (and in the same sense as) cabbeling in the Southern Ocean and together they cause large negative (downward) diancutral and diapycnal velocities. The proper inclusion of these processes in the potential density conservation equation of Olbers and Wenzel (1989) must be expected to lead to large changes to the implications of mixing processes in this model.

Notice that only one of the additional terms in the diapycnal velocity equation, Eq. (20), relies on the difference between the in-situ pressure and the reference pressure; this being the term proportional to the epineutral flux divergence of potential temperature, $h^{-1}\nabla_n\cdot(hK\nabla_n\theta)$. The other three terms (cabbeling, thermobaricity and $D\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2$) contribute to the nonconservation of potential density even if $p = p_r$ and $c = 1$. The fact that the "e" equation, Eq. (12), does not contain the epineutral Laplacian of potential temperature is an attractive feature in comparison with the w^d equation, Eq. (20), since the epineutral Laplacian is relatively noisy to evaluate. This is an added argument in favour of using neutral surfaces rather than potential-density surfaces.

Where thermobaricity is significant in the ocean, it makes a positive contribution to Eq. (20) (McDougall and You, 1990), so that three of the four additional terms in Eq. (20) are positive. Assuming that the term due to the epineutral flux divergence of potential temperature is either positive, or if it is negative, it is small, then the sum of the four extra terms in Eq. (20) will be positive. The simplistic conservation equation of potential density which regards potential density as a conservative variable then results in a diapycnal velocity that is too large (for a given vertical diffusivity, D), by as much as 100%, or by more than 1×10^{-7} m s^{-1}. This error in the diapycnal velocity then feeds into the conservation equations of other tracers like potential temperature and salinity so that the interior water-mass conversion in these models does not occur at a rate proportional to the curvature of the vertical S-θ curve, but rather is dominated by the diapycnal advection term (if it is positive). If one assumes that the inverse model retains the correct diapycnal velocity for dynamical reasons, the use of the simplistic equation for the diapycnal velocity proves to be equivalent to a corresponding underestimation of the diapycnal diffusivity, D (for a positive w^d). Hence we expect that the omission of the conservative terms in the conservation equation of potential density has resulted in either an overestimate of the diapycnal velocity or an underestimate of the vertical diffusivity (for $w^d > 0$), or, equivalently, a combination of both.

The Potential Temperature Conservation Equation

The conservation equation of potential temperature with the total material derivative written with respect to potential-density surfaces is (from Eq. (19))

$$\begin{aligned}\theta_t\big|_\sigma + \mathbf{V}^L\cdot\nabla_\sigma\theta + w^d\theta_z &= \tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n\theta\right) + \left[D\theta_z\right]_z \\ &\neq \tfrac{1}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma\theta\right) + \left[D\theta_z\right]_z .\end{aligned} \tag{22}$$

The lateral flux divergence of potential temperature in a neutral surface framework is not the same as that in a potential density framework, and here we derive the differences.

The lateral gradient of potential temperature in a neutral surface, $\nabla_n\theta$, is related to the lateral gradient in the intersecting potential-density surface, $\nabla_\sigma\theta$, by the simple expression (McDougall, 1987b)

$$\boxed{\nabla_\sigma\theta = \mu\nabla_n\theta,} \qquad \text{where} \qquad \boxed{\mu = c\left[R_\rho - 1\right]/\left[R_\rho - c\right],} \tag{23}$$

c is defined by $c = \left[\alpha/\beta\right]\left[\tilde{\alpha}/\tilde{\beta}\right]^{-1}$ and R_ρ is the density ratio $\alpha\theta_z / \beta S_z$. The same multiplying factor, μ, applies to the temporal derivatives of potential temperature in the two reference frames. In order to find the relationship between the lateral mixing terms in the two different reference frames in Eq. (22), the following relationship between the lateral differential operators is used (McDougall and Jackett, 1988)

$$\nabla_\sigma = \nabla_n + \left[\nabla_\sigma\mathcal{R} - \nabla_n\mathcal{N}\right]\frac{\partial}{\partial z}\Big|_{x,y}, \tag{24}$$

where $\mathcal{R}[x,y]$ and $\mathcal{N}[x,y]$ are the heights of a particular potential-density surface and the intersecting neutral surface respectively, so that $\nabla_\sigma\mathcal{R} = \left[\mathcal{R}_x, \mathcal{R}_y\right]$ and $\nabla_n\mathcal{N} = \left[\mathcal{N}_x, \mathcal{N}_y\right]$. Applying the differential operator, Eq. (24), to $\nabla_\sigma\theta$ $(= \mu\nabla_n\theta)$ we find that

$$\nabla_\sigma^2\theta = \mu\nabla_n^2\theta + \nabla_n\mu\cdot\nabla_n\theta + \left[\nabla_\sigma\mathcal{R} - \nabla_n\mathcal{N}\right]\cdot\left(\nabla_\sigma\theta\right)_z, \tag{25}$$

and the potential temperature conservation equation can be rewritten in the form

$$\begin{aligned}\boxed{\theta_t\big|_\sigma + \mathbf{v}^L\cdot\nabla_\sigma\theta + w^d\theta_z = {} & \tfrac{1}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma\theta\right) + \left[D\theta_z\right]_z \\ & - \left[\mu - 1\right]\tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n\theta\right) - K\nabla_n\mu\cdot\nabla_n\theta \\ & - K\left(\left[\mu^2 - \mu\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta\right)_z + \mu\nabla_n\theta\cdot\left(\nabla_n K - \nabla_\sigma K\right),}\end{aligned} \tag{26}$$

where the following geometrical relationships between the slopes of potential-density surfaces and neutral surfaces have been used (McDougall, 1988),

$$\left[\nabla_\sigma\mathcal{R} - \nabla_n\mathcal{N}\right] = \left[\mu - 1\right]\nabla_n\theta/\theta_z, \tag{27}$$

$$\left[\nabla_\sigma\mathcal{R} - \nabla_n\mathcal{N}\right]_z = \tfrac{1}{h^\sigma}\nabla_\sigma h^\sigma - \tfrac{1}{h}\nabla_n h. \tag{28}$$

The usual conservation equation of potential temperature in a potential density framework involves just the terms on the first line of Eq. (26). The additional terms on the second and third lines arise because of the different lateral gradients of potential temperature in the two coordinate frames and because of the different slopes of potential-density surfaces and neutral surfaces. As an example of the importance of the first term on the second line, consider the use of potential-density surfaces down to a depth of 2000 m in the North Atlantic, as in the inversion of Olbers et al (1985). At 1500 m in the North Atlantic, μ is generally greater than 1.5 so this term, $-\left[\mu - 1\right]\tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n\theta\right)$, represents a correction of 50% of the correct lateral mixing term that appears on the right of this equation (see Eq. (22)). In addition, the next term, $-K\nabla_n\mu\cdot\nabla_n\theta$, is of similar magnitude at this depth in frontal regions of the North Atlantic.

The third line of Eq. (26) contains the term $-K\mu_z\left[2\mu - 1\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta$, which does not go to zero as μ tends to 1; that is, as the in-situ pressure tends to the reference pressure. Neither does the $-K\nabla_n\mu\cdot\nabla_n\theta$ term go to zero as p tends to p_r. These terms arise because while the two types of surface are tangent where μ is 1, they do not have the same curvature in space. Even though Hogg's (1987) study used σ_1 surfaces rather than σ_θ surfaces so that μ was close to unity, the $-K\mu_z\left[2\mu - 1\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta$ term is estimated to contribute 10^{-9} K s^{-1} to the rate of change of potential temperature in the southern part of his western box; a magnitude that is as large as that of any of his mixing or advection terms. In the Antarctic Circumpolar Current $-K\nabla_n\mu\ \nabla_n\theta$ is as large as any other term in (26) even if $p_r = p$, due to the pressure dependence of c through the term $-K\frac{\partial\mu}{\partial c}\frac{1}{\tilde{\alpha}}T_b\nabla_n\theta\cdot\nabla_n p$.

When the diapycnal velocity is estimated using just the first line of Eq. (20), the constraint that the potential temperature conservation equation places on the temporal and epipycnal changes of θ, $\theta_t|_\sigma + \mathbf{V}^L \cdot \nabla_\sigma \theta$, is not due to the (correct) mixing terms that would result from inserting the full expression Eq. (20) for w^d into the correct conservation equation for θ, (the top line of Eq. (22) or all of Eq. (26)), but has additional terms due to the errors made in estimating the diapycnal velocity. Using Eq. (20) and Eq. (26), it can be shown that

$$\begin{aligned}\theta_t|_\sigma + \mathbf{V}^L \cdot \nabla_\sigma \theta \;=\; & \tfrac{1}{h^\sigma} \nabla_\sigma \cdot \left(h^\sigma K \nabla_\sigma \theta\right) + \left[D\theta_z\right]_z - \theta_z \left[\tilde{\alpha}\theta_z - \tilde{\beta}S_z\right]^{-1} \left[D\left(\tilde{\alpha}\theta_z - \tilde{\beta}S_z\right)\right]_z \\ & + \tfrac{\beta}{\tilde{\beta}} \mu g N^{-2} D \tfrac{\partial \tilde{\alpha}}{\partial \theta} \theta_z^3 \\ & - K \nabla_n \mu \cdot \nabla_n \theta - K\left(\left[\mu^2 - \mu\right] \theta_z^{-1} \nabla_n \theta \cdot \nabla_n \theta\right)_z + \mu \nabla_n \theta \cdot \left(\nabla_n K - \nabla_\sigma K\right) \\ & + \mu K g N^{-2} \theta_z \left\{C_b \nabla_n \theta \cdot \nabla_n \theta \;+\; T_b \, \nabla_n \theta \cdot \nabla_n p\right\}.\end{aligned} \tag{29}$$

The top line of the right-hand side of this equation represents the way in which the mixing processes parameterized by D and K affect the temporal and epipycnal changes of potential temperature in such a truncated model, while the additional terms are corrections that result partially from the fact that θ has been mixed along potential-density surfaces rather than along neutral surfaces (the many error terms in Eq. (26)) and partially due to the inaccurate estimation of the diapycnal velocity from just the first term in Eq. (20), that is, as $\left[\tilde{\alpha}\theta_z - \tilde{\beta}S_z\right]^{-1}\left[D(\tilde{\alpha}\theta_z - \tilde{\beta}S_z)\right]_z$ rather than using the full Eq. (20). Note that the divergence of the epineutral flux of potential temperature terms that appear in both Eq. (20) and Eq. (26) have cancelled and do not appear in Eq. (29). The $\beta\tilde{\beta}^{-1}\mu g N^{-2} D \frac{\partial \tilde{\alpha}}{\partial \theta}\theta_z^3$ term in Eq. (29) is important in the upper ocean, while thermobaricity and cabbeling tend to be larger at thermoclinic fronts deeper in the water column, and the other terms involving the lateral diffusivity, K, have been illustrated following Eq. (26). The discussion following Eq. (13) above indicates how serious the omission of thermobaricity and cabbeling is in the water-mass conversion equation, Eq. (11) or Eq. (29), at a thermoclinic front such as the Antarctic Circumpolar Current (cf Olbers and Wenzel, 1989).

This section has shown that even though potential temperature, θ, is a conservative variable, the use of a potential density rather than a neutral surface reference frame for the lateral mixing process results in many additional terms in the conservation equation (see Eq. (26)). None of these terms have previously been included in inverse models. Two of these terms, $-K\nabla_n\mu \cdot \nabla_n \vartheta$, and $-K\mu_z\left[2\mu - 1\right]\theta_z^{-1}\nabla_n\theta \cdot \nabla_n\theta$ remain even when the reference pressure of the potential density is equal to the in-situ pressure. We have also seen how a mistaken estimate of the diapycnal velocity feeds through into the tracer conservation statements so that these equations represent an incorrect balance between epipycnal advection (LHS) and mixing (RHS). This leads to an error in the estimated lateral flow field. A similar misestimate of the lateral velocity field occurs in the linear vorticity equation through the vortex stretching term, part of which is due to the difference of the diapycnal velocity through neighboring potential-density surfaces.

The ρ_θ "Conservation" Equation as a Linear Combination of the θ and S Equations

Rather than carrying a conservation equation for potential density in an inverse model, one can instead carry conservation equations for both potential temperature, θ, and salinity , S. The linear combination of these equations using the multiplying coefficients, $\tilde{\alpha}$ and $\tilde{\beta}$ is the potential density equation. It would actually be desirable to replace either the θ or the S equation with this linear combination of the S and θ equations because it is independent of any errors in the lateral velocity vector and so yields a more direct link (better signal to noise) between the mixing processes and the consequent diapycnal

advection (see the discussion around Eq. (41) below). Another real advantage of approaching the potential density equation by using the θ and S equations is that one does not leave out the nonlinear term due to vertical mixing, $D\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2$ that is commonly omitted from Eq. (20). Also, the nonlinear terms due to lateral mixing are smaller (at least when at the reference pressure). To see this we need to first derive the conservation equation for salinity that is analogous to Eq. (26) for potential temperature. This is done by applying the differential operator, Eq. (24), to $\nabla_\sigma S = \frac{\mu}{c}\nabla_n S$, obtaining

$$\begin{aligned} S_t\big|_\sigma + \mathbf{V}^L\cdot\nabla_\sigma S + w^d S_z \;=\; & \tfrac{1}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma S\right) \;+\; \left[DS_z\right]_z \\ & -\left[\tfrac{\mu}{c}-1\right]\tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n S\right) - K\nabla_n\left(\tfrac{\mu}{c}\right)\cdot\nabla_n S \\ & -K\left(\tfrac{\tilde{\alpha}}{\tilde{\beta}}\left[\mu^2-\mu\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta\right)_z \;+\; \tfrac{\mu}{c}\nabla_n S\cdot\left(\nabla_n K - \nabla_\sigma K\right). \end{aligned} \tag{30}$$

Multiplying this equation by $\tilde{\beta}$ and subtracting $\tilde{\alpha}$ times Eq (26) we find

$$\begin{aligned} w^d\left[\tfrac{1}{\rho_0}\tfrac{\partial\rho_0}{\partial z}\right] \;=\; & \tfrac{\tilde{\beta}}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma S\right) - \tfrac{\tilde{\alpha}}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma\theta\right) + \tilde{\beta}\left[DS_z\right]_z - \tilde{\alpha}\left[D\theta_z\right]_z \\ & +\tilde{\alpha}(c-1)\tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n\theta\right) \;+\; \tilde{\alpha}\tfrac{\mu}{c}K\nabla_n c\cdot\nabla_n\theta \\ & -K\tfrac{\partial\tilde{\alpha}}{\partial\theta}\left[\mu^2-\mu\right]\nabla_n\theta\cdot\nabla_n\theta \;-\; \tfrac{\tilde{\beta}}{\beta}\left[\tfrac{\mu}{c}-1\right]K\left\{C_b\nabla_n\theta\cdot\nabla_n\theta + T_b\nabla_n\theta\cdot\nabla_n p\right\}. \end{aligned} \tag{31}$$

The terms in the second and third lines of this equation are due to having assumed that the lateral diffusion occurs along potential-density surfaces rather than along neutral surfaces. In contrast to the σ_θ equation, Eq. (20), there are no correction terms in Eq. (31) due to vertical mixing. If the reference pressure is equal to the in-situ pressure, the only error term in Eq. (31) comes from the $\tilde{\alpha}\frac{\mu}{c}K\nabla_n c\cdot\nabla_n\theta$ term and is $KT_b\nabla_n\theta\cdot\nabla_n p$, that is, thermobaricity, whereas in Eq. (20) there is also the cabbeling term, $KC_b\nabla_n\theta\cdot\nabla_n\theta$, in this situation. Of course, if the lateral mixing had been parameterized as being epineutral there would be no correction terms in Eq. (31). It is concluded that carrying both the θ and S conservation statements, even in their truncated forms (i.e. the top lines of Eq. (24) and Eq. (30)), is more desirable than using the top line of the conservation equation for potential density, Eq. (20). As noted above, signal-to-noise considerations mean that it is preferable to carry either the S or the θ equations, and the linear combination $\tilde{\beta}$ times Eq. (30) minus $\tilde{\alpha}$ times Eq. (26).

The Spiciness "Conservation" Equation

Olbers et al., (1985) and Olbers and Wenzel (1989) chose to use Veronicity, τ^V, instead of salinity or potential temperature in their beta-spiral inversions of the North Atlantic and Southern Ocean. Here it is shown that just as potential density is not a conservative variable, neither is Veronicity, and the non-conservative nature of Veronicity has introduced significant errors in these inversions. The non-conservative nature of Veronicity can be noted from the fact that contours of constant τ^V are not straight lines on the S-θ diagram (see Figure 7) so that mixing between two water parcels that have the same Veronicity will produce a parcel with a different value of Veronicity. The curvature of the τ^V isolines on this diagram is similar to that of the σ_θ isolines, so nonlinear effects of a similar magnitude may be expected. For a given (and arbitrary) value of the relative scaling of the two axes, Veronicity has the property that its isolines are orthogonal to lines of constant potential density on the S-θ diagram. Jackett and McDougall (1985) have proposed a different definition of orthogonality that does not depend on the scaling of the axes of the S-θ diagram, and here a conservation equation is developed for their variable, spiciness, τ.

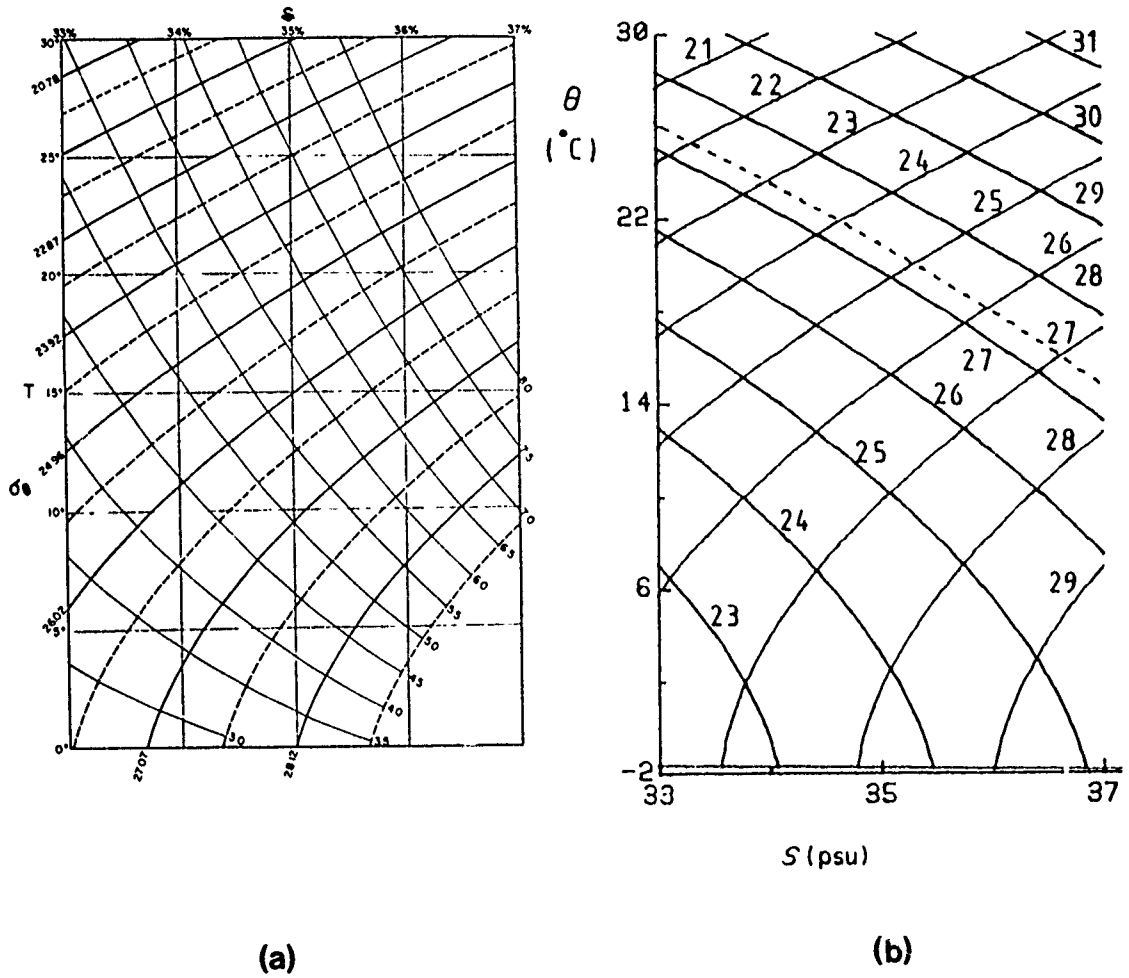


Fig. 7. Contour plots of, (a) Veronicity, τ^V and (b), spiciness, τ, on the S-θ diagram. The other lines shown are of potential density referenced to the sea-surface.

To a good approximation, variations of spiciness are related to those of potential temperature and salinity by

$$d\tau = \tilde{\beta} dS + \tilde{\alpha} d\theta. \tag{32}$$

For this total differential to define a path-independent function, τ, one needs $\partial\tilde{\beta}/\partial\theta = \partial\tilde{\alpha}/\partial S$ which is not true of the real equation of state of seawater. One can imagine an equation of state that did satisfy this constraint (by for example, having $\partial^2\rho/\partial\theta\partial S = 0$), and then a linear combination of the conservation statements of potential temperature and salinity can be taken to arrive at the following conservation equation for τ

$$\tau_t\big|_z + \mathbf{V}^L \cdot \nabla_2 \tau + w\tau_z = \tau_t\big|_\sigma + \mathbf{V}^L \cdot \nabla_\sigma \tau + w^d \tau_z = h^{-1}\nabla_n \cdot (hK\nabla_n \tau) + \left[D\tau_z\right]_z - K\frac{\partial\tilde{\alpha}}{\partial\theta}\nabla_n\theta \;\; \nabla_n\theta - D\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2 . \tag{33}$$

Just as in Eq. (20) above where the nonlinear mixing terms in the potential density conservation equation were found to be large in the thermoclines of many ocean basins, so too in these regions will the above terms, $-D\frac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2$ and $-K\frac{\partial\tilde{\alpha}}{\partial\theta}\nabla_n\theta\cdot\nabla_n\theta$ be important in Eq. (33).

In practice, the lateral mixing of τ in previous inverse models has been along potential-density surfaces, rather than along neutral surfaces, and the differences can be found by applying the differential operator, Eq. (24) to $\nabla_\sigma\tau\,(=2\tilde{\alpha}\nabla_\sigma\theta)$ so that Eq. (33) can be written as

$$\begin{aligned}\tau_t\big|_\sigma+\mathbf{V}^L\cdot\nabla_\sigma\tau+w^d\tau_z=\;&\tfrac{1}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma\tau\right)+\left[D\tau_z\right]_z\\&-D\tfrac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2\\&-\tilde{\alpha}\left[2(\mu-1)-(c-1)\right]\tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n\theta\right)\;-\;2\tilde{\alpha}K\nabla_n\mu\cdot\nabla_n\theta\\&-2\tilde{\alpha}K\left(\left[\mu^2-\mu\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta\right)_z\;+\;2\tilde{\alpha}\mu\nabla_n\theta\cdot\left(\nabla_nK-\nabla_\sigma K\right)\\&-2\mu^2\tfrac{\partial\tilde{\alpha}}{\partial\theta}K\nabla_n\theta\cdot\nabla_n\theta\;+\;\tfrac{\tilde{\beta}}{\beta}K\left\{C_b\nabla_n\theta\cdot\nabla_n\theta+T_b\nabla_n\theta\cdot\nabla_np\right\}.\end{aligned}\tag{34}$$

The mixing terms that were included by Olbers et al (1985) and Olbers and Wenzel (1989) were the first two, on the top line of this equation; all the other terms are due to the nonconservative nature of both potential density and spiciness. There are now five terms that do not disappear when p_r is equal to p (i.e. when $c=\mu=1$). These are $-2\tilde{\alpha}K\nabla_n\mu\cdot\nabla_n\theta$, $-2\tilde{\alpha}K\mu_z\left[2\mu-1\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta$, and the three terms on the last line of the equation. Since $C_b\approx\frac{\partial\tilde{\alpha}}{\partial\theta}$, cabbeling will effectively appear as $-\left[2\mu^2-1\right]KC_b\nabla_n\theta\cdot\nabla_n\theta$. All of these terms have been discussed above and they were found to each be large in different parts of the ocean, notably at thermoclinic fronts where either $\nabla_n\theta\cdot\nabla_n\theta$ or $\nabla_n\theta\cdot\nabla_np$ is large.

Since the diapycnal velocity is commonly estimated with a truncated conservation equation, the effective constraint that the spiciness conservation equation would maintain between the epipycnal advection of τ and the mixing processes is illustrated better by the following τ conservation equation,

$$\begin{aligned}\tau_t\big|_\sigma+\mathbf{V}^L\cdot\nabla_\sigma\tau=\;&\tfrac{1}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma\tau\right)+\left[D\tau_z\right]_z-\tau_z\left[\tilde{\alpha}\theta_z-\tilde{\beta}S_z\right]^{-1}\left[D\left(\tilde{\alpha}\theta_z-\tilde{\beta}S_z\right)\right]_z\\&+\tfrac{2\mu}{\left(R_\rho-1\right)}D\tfrac{\partial\tilde{\alpha}}{\partial\theta}\theta_z^2\;-\;2\tilde{\alpha}K\nabla_n\mu\cdot\nabla_n\theta\\&-2\tilde{\alpha}K\left(\left[\mu^2-\mu\right]\theta_z^{-1}\nabla_n\theta\cdot\nabla_n\theta\right)_z\;+\;2\tilde{\alpha}\mu\nabla_n\theta\cdot\left(\nabla_nK-\nabla_\sigma K\right)\\&-2\mu^2\tfrac{\partial\tilde{\alpha}}{\partial\theta}K\nabla_n\theta\cdot\nabla_n\theta\;+\;\tfrac{2R_\rho}{\left(R_\rho-c\right)}\tfrac{\tilde{\beta}}{\beta}K\left\{C_b\nabla_n\theta\cdot\nabla_n\theta+T_b\nabla_n\theta\cdot\nabla_np\right\}\end{aligned}\tag{35}$$

Again, only the top line was included in the previous inversions. Notice that the $h^{-1}\nabla_n\cdot\left(hK\nabla_n\theta\right)$ term has again cancelled from this equation. The nonlinear mixing term due to vertical diffusion has changed sign between Eq. (34) and Eq. (35) and has increased in magnitude by the factor $2\mu\left(R_\rho-1\right)^{-1}=2c\left(R_\rho-c\right)^{-1}$. At 1500 m in the North Atlantic this factor is about 3. Similarly, cabbeling and thermobaricity has a different multiplying factor, $2R_\rho\left(R_\rho-c\right)^{-1}$, and this is about 5 in the North Atlantic at 1500 m. These multiplying factors increase the magnitude of these nonlinear terms which were already as large as the physically correct terms in the conservation statements that have been considered above. It is the compounding of the errors from the potential density conservation statement, and the assumption that spiciness is a conservative variable that has led to these large multiplying factors.

The simple lesson for us to learn here is that just as potential density is not a conservative variable, neither is Veronicity, τ^V, or spiciness, τ, and so conservation statements for these variables must take this into account. What many will find surprising is the magnitude of the many nonconservative terms in relation to the traditional terms in the conservation statements, and also the fact that the difference between epipycnal and epineutral mixing remains even when the potential density's reference pressure is equal to the in-situ pressure.

MIXING PRESCRIPTIONS IN BOX MODEL INVERSIONS: The Divergence Forms of Conservation Statements in Potential-Density Surfaces

The Nonconservative Production Terms in the Divergence Forms of the Conservation Equations

As far as the conservation equations are concerned, the distinguishing feature between β-spiral inversion methods and box-inverse methods is that the β-spiral methods use advective conservation statements while box-model methods use divergence (or flux) conservation statements. In addition, the β-spiral method is often used with data that have been laterally smoothed in some way, whereas the box models are generally used with unsmoothed data. The β-spiral method examines the balance of terms at a point while the box-inverse method is concerned with the balance of fluxes through the six faces of a box. The streamfunction method is a sub-set of the box-model approach that expresses the Eulerian lateral velocity vector in terms of a geostrophic streamfunction. Note that the Lagrangian lateral velocity vector does not satisfy the required relationship to be expressible as the derivative of a streamfunction. However, since inverse methods have not yet addressed this question of the difference between the Lagrangian and Eulerian mean flows, this aspect has not yet been an issue. Some models can be a little hard to classify as a box inversion or as a β-spiral inversion. For example, Hogg (1987) used a divergence form for the lateral gradient operator (using a streamfunction) but evaluated the vertical derivatives on a potential-density surface rather than between a pair of surfaces. In this way his model is like a β-spiral method in the vertical while being like a box model in the lateral directions.

In neutral surface coordinates, the divergence form of the general conservation statement for a conservative tracer, C, is given by Eq. (7), and when integrated over the sides of a box of volume V and lateral area A, is

$$V\frac{\partial(hC)|_n}{\partial t}+\left[L_{NS}hu^LC\right]_W^E+\left[L_{EW}hv^LC\right]_S^N+\left[AeC\right]_\ell^u=\left[L_{NS}hKC_x\right]_W^E+\left[L_{EW}hKC_y\right]_S^N+\left[ADC_z\right]_\ell^u, \qquad (36)$$

where L_{EW} and L_{NS} are the horizontal length scales of a face of the box in the east-west and north-south directions respectively, and the lateral derivatives of C are evaluated in neutral surfaces. Here u^L and v^L are the x and y components of the Lagrangian-mean velocity vector

To date, box models have used potential-density surfaces as their upper and lower boundaries rather than neutral surfaces, and here we examine the errors so introduced in the lateral transport of properties. To do so, divergence forms of conservation statements are needed in the potential-density surface coordinate frame. These are obtained by noting that the continuity equation can be written with respect to potential-density surfaces as (cf Eq. (8)),

$$\frac{\partial h^\sigma}{\partial t}+\nabla_\sigma\cdot\left(h^\sigma\mathbf{V}^L\right)+\left[w^d\right]_{\sigma_\ell}^{\sigma_u}=0, \qquad (37)$$

and taking C times this equation plus h^σ times the advective form of the C conservation equation, (Eq. (9) with the advection terms on the left-hand side written with respect to potential-density surfaces), one finds

$$\boxed{\frac{\partial\left(h^\sigma C\right)\big|_\sigma}{\partial t} + \nabla_\sigma\cdot\left(h^\sigma \mathbf{V}^L C\right) + \left[w^d C\right]_{\sigma_\ell}^{\sigma_u} = \nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma C\right) + \left[DC_z\right]_{\sigma_\ell}^{\sigma_u} + h^\sigma\left\{\tfrac{1}{h}\nabla_n\cdot\left(hK\nabla_n C\right) - \tfrac{1}{h^\sigma}\nabla_\sigma\cdot\left(h^\sigma K\nabla_\sigma C\right)\right\}.} \tag{38}$$

The extra terms due to the different coordinate frames appear in the second line of this equation and exactly correspond to the extra terms in the advective forms of the equation, for example, in Eq. (22) and Eq. (26) in the case of potential temperature, and in Eq. (30) for salinity. The extra terms in the conservation equations for potential density (Eqs. 20 or 31) and spiciness (Eq. 34) due to the non-conservative nature of these variables also carry directly across to the divergence forms. We conclude that all the extra terms due to the non conservative nature of potential density and spiciness that have been found in the previous section of this paper for the advective forms apply equally to the divergence forms of these conservation equations. For this reason we do not need to repeat the error analysis for these extra terms. Suffice it to say that these terms that have been missing from previous inverse studies are often as large as the straightforward mixing terms in various parts of the model domains.

The Impact of Uncertainty in the Continuity Equation on the Tracer Equations

The single most important aspect of many past box-inverse models that has precluded them from being influenced by and detecting oceanic mixing processes is the overwhelming effect of errors in the continuity equation (the incompressibility equation) on the tracer conservation equations. In order to demonstrate this, the left-hand side of Eq. (38) is rewritten as the linear combination of Eq. (9) and Eq. (37) as

$$C\left\{\frac{\partial h^\sigma}{\partial t} + \nabla_\sigma\cdot\left(h^\sigma\mathbf{V}^L\right) + \left[w^d\right]_{\sigma_\ell}^{\sigma_u}\right\} + h^\sigma\left\{C_t\big|_\sigma + \mathbf{V}^L\cdot\nabla_\sigma C + w^d C_z\right\} = \tfrac{h^\sigma}{h}\nabla_n\cdot\left(hK\nabla_n C\right) + \left[DC_z\right]_{\sigma_\ell}^{\sigma_u}. \tag{39}$$

The physical constraint that we wish to glean from this equation is the connection between the mixing of tracer C (the right-hand side) and the Lagrangian advection of C (the second curly bracket on the left). The first curly bracket is included so as to be able to write the left-hand side of Eq. (38) in a divergence form. When integrated over the sides of a box of a typical oceanic box model, the uncertainty in the continuity equation (first bracket in Eq. (39)) is typically 1 Sv (10^6 m^3 s^{-1}) and this is so large as to swamp any signal that represents the physically interesting "advective-diffusive" balance of C-stuff. As an example, consider the conservation of salinity in box 11 of Wunsch and Minster (1982). From their Figures 1c and 3a the epipycnal and vertical advection of salt amount to only 1% of the mean salinity (35 psu) times the volume flux imbalance into the box. That is, the solution has the left-hand sides of our Eq. (38) and Eq. (39) being dominated, by a factor of a hundred, by the error in the continuity equation. The consequence of this is that the signature of mixing processes and the balancing advection of C are simply small terms in the equation so that even when the mixing and advection are modelled correctly, the variance of the salt equation can only decrease by two percent. Effectively, the salt conservation equation becomes simply another continuity equation and consequently (i) the rank of the system of equations suffers because of the nearly collinear nature of the equations, (ii) the information that is contained in the salinity contours in three-dimensional space is not

imposed as a constraint on the solution, and (iii) no information on the mixing processes can be obtained from the inversion.

The same comments apply to many other box-inverse papers including Wunsch (1984), Joyce Wunsch and Pierce (1986), and Wunsch, Hu and Grant (1983). In all of these papers, the balancing advective-diffusive terms in the salinity conservation equations were typically 1% of the residual error left in the salt conservation equation by the fact that the continuity equation was not satisfied identically. Similar comments apply to other conservation statements (e.g. potential temperature, dissolved oxygen or silicate) but the situation is worst for salinity simply because its mean value is quite large in relation to its variations in the ocean.

The Wunsch and Minster (1982) paper is especially interesting in this regard because it carried both epipycnal and diapycnal diffusivities, and surprisingly, the lateral diffusivities of the model were of order 1-10 $m^2\ s^{-1}$, or a factor of between a hundred and a thousand less than what we believe for the ocean. One would think that any inconsistencies in a model's data or its equations would cause the magnitude of its parameters to increase so as to soak up some of the noise. Why then were the lateral diffusivities close to zero? I believe the answer is again related to the volume flux imbalances in the continuity equation. When using all the eigenvectors, in addition to satisfying the equations identically, the SVD solution procedure also minimizes the norm of the solution vector. When some of the eigenvectors are discarded (the ones with the smallest eigenvalues) the solution norm becomes even smaller and the equations are no longer satisfied exactly. The method then represents a tension between satisfying the equations while also having a small solution vector. This competition between minimizing the equation errors and the solution norm is more mathematically obvious in the ridge regression or tapered least squares procedure. Because the error in the salinity conservation statement is dominated by the error in the continuity equation for each box, a respectable value of the lateral diffusivity (of order 1000 $m^2\ s^{-1}$) would only reduce the error variance in the salinity equation by 2% so the solution procedure chooses instead to have a small solution vector, and in particular, to have small lateral diffusivities. In this situation, the solution will be sensitive to the column weighting, and presumably a different choice of column weights could also have yielded very small vertical diffusivities as well. This explanation of the unrealistically small values of the lateral diffusivity found by Wunsch and Minster (1982) is due to the combination of (i) the signature of mixing and advection of salinity being well below the allowable error in the salinity conservation equation, combined with (ii), a solution procedure that prefers small values of the diffusivities.

The inverse model of Wunsch, Hu and Grant (1983) is, I believe, another example of a study that has been unwittingly plagued by the salt equation being effectively another continuity equation. The paper concludes that there is no need for vertical mixing in their model, even though the downward diapycnal velocities were of order 10^{-7} m s^{-1} (see their Figure 11a). The necessity for diapycnal advection but not for vertical diffusion was argued because the imbalances in the tracer conservation equations (notably salinity) were almost completely explained (a posteriori) by interfacial advection with little apparent need for vertical diffusion. However 99% of the interfacial advection that was added into the salt conservation equation went into correcting the first curly bracket in Eq. (39) by adding the vertical part of the volume flux divergence, leaving a comparatively small salt residual (called ε_2 by the authors) that resembles the physically interesting Lagrangian advection of salinity. From Figure 5b, one finds a value of $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$ of about -1×10^{-3} m^{-1} in the region of the South Pacific between the two Scorpio sections at a depth of about 1000 m. At this depth Wunsch et al (1983) find a (downward) diapycnal velocity of about -0.7×10^{-7} $m^2\ s^{-1}$, which is consistent with a vertical diffusivity of 0.7×10^{-4} $m^2\ s^{-1}$ acting on the above value of $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$. However, deeper in the water column at 1400 m in this region, McDougall and You's (1990) figure 5 shows that $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$ is positive so that a positive vertical diffusivity would not be consistent with the downward diapycnal

velocity found by Wunsch et al (1983) at this depth. Since the salinity equations in Wunsch et al (1983) were functionally linearly dependent with the continuity equations, the reference level velocity vectors that came out of the inversion were effectively not influenced by the advective-diffusive salt balances. Only by allowing this advection of salinity to emerge as signal in a conservation equation, rather than being hidden in the last few percent of an equation's noise, and at the same time balancing this advection of salt with a parameterized mixing term, will we be able to say whether mixing processes are required by the data.

Is there a procedure by which mixing processes can be made to stand out above the noise in tracer equations in box inverse models? The answer is yes, and the key is to minimize the influence of the first curly bracket in Eq. (39) by writing a conservation equation, not for the tracer itself, but for the tracer anomaly, C', from some fixed value, $\overline{C}$. The reason why this procedure works is that the advective form of the conservation statement applies equally well to C' as to C, since it involves only various derivatives of the tracer. Hence in forming Eq. (38) and Eq. (39) we can multiply Eq. (37) by the primed tracer variable instead of by C, and Eq. (38) and Eq. (39) become exactly the same equations with C replaced by C'. This procedure has in fact been used to good effect by several researchers, including Hogg (1987), and Lee and Veronis (1991). This procedure is easiest to justify if the same mean tracer value is subtracted from every box, but in the last subsection of this paper it is argued that it may be beneficial to subtract a different mean tracer value for each box.

Memery and Wunsch (1990) were able to balance the volume fluxes into and out of their boxes to within about 0.02 Sv without having an excessively large solution norm. This was most likely due to the fact that they used the Levitus (1982) data set that is that is temporally and horizontally averaged and so does not contain the complicating features such as internal waves and mesoscale eddies that are found in real cruise data. Presumably the salinity conservation equations were not unduly correlated with the continuity equations in the Memery and Wunsch (1990) study, and that if vertical diffusivities were added to the model, the inversion would have found them to be necessary. Conservation equations for potential temperature should also be added to the system in order to provide added information on the mixing processes (see the section below). Schlitzer (1987, 1988) also used the Levitus data set and was able to specify the volume flux imbalances to as iittle as 0.001 Sv, and he resolved both diapycnal advection and diapycnal diffusivities. This lends support to the present claim that the errors in the continuity equation are what have been precluding most box-inverse models from resolving mixing processes.

Many box-inverse models of the Atlantic and Indian Oceans (e.g. Wunsch and Grant (1982), Rintoul and Wunsch (1991) and Fu (1986)), have found that salinity conservation did not add any information to their inversions. Since these oceans contain substantial variations of salinity along neutral surfaces, some of which information has previously been used to determine flow directions of water types by the Wüstian tongue method, it would be quite incredible if the salinity field contained no information on mixing and advection: the challenge is to extract this information.

The Need for Vertical Diffusion as well as Interfacial Advection

It has been quite common in box inversions to include a diapycnal flux of volume but not to have any diffusive flux of tracers such as potential temperature or salinity. Since diapycnal advection occurs only in response to mixing processes (see Eq. (20)), it is clearly dangerous to include one part of the effect of mixing processes (the advection of tracer) without at the same time carrying the other part (the diffusion of tracer). For example, in the solutions for the North Atlantic circulation presented by Wunsch and Grant (1982) and by Wunsch (1984), many of the isopycnal interfaces had diapycnal

velocities of more than 10×10^{-7} m s^{-1} passing through them. That is, the diapycnal velocity was more than ten times the canonical value of the upwelling velocity in the deep ocean. On the face of it, this implies that the vertical diffusivity must be about ten times its canonical value, or about 10×10^{-4} m^2 s^{-1} in order to fuel this seemingly large (but probably not impossibly large) diapycnal flow. Wunsch (1984) has obviously wrestled with this aspect of his model, as the discussion around his equation (18) shows (and the corresponding discussion in Wunsch, Hu and Grant (1983)). There it was proposed to regard the model's interfacial velocity as a combination of a true interfacial velocity and a diffusive flux of density. But a scale analysis shows that the eddy diffusion of density is a very small fraction (less than 0.3%) of the advection of density (i.e. $D \partial \rho_\theta / \partial z << \rho_\theta w^d$). We are obviously never going to be able to account for terms in conservation equations to this accuracy.

Many of the box-inversion papers write conservation equations for "density" or "mass", however, in the divergence form, a conservation equation for "density" is almost the same as a conservation statement for volume transport. This is apparent from a scale analysis of the two curly brackets on the left of Eq. (39), using the in-situ density (or indeed any other kind of density) for C. Since the imbalance in the volume transport in and out of a box is allowed to be say 0.3 Sv, the second curly bracket in Eq. (39) (the advection of density) amounts to only 0.3% of the first bracket, hence the "density" conservation equation is equivalent to the incompressibility condition. There is an extra pedantic twist to this issue in that McDougall and Garrett (1991) have shown that while the divergence of the instantaneous velocity vector, $\nabla \cdot \mathbf{u}$, is directly related to the instantaneous Lagrangian change of density, the divergence of the mean velocity vector, $\nabla \cdot \overline{\mathbf{u}}$, is unaffected by the divergence of the *turbulent* fluxes of heat and salt (even though the *molecular* flux divergences of heat and salt do contribute to $\nabla \cdot \overline{\mathbf{u}}$!). In this way, $\nabla \cdot \overline{\mathbf{u}} = 0$ is actually a better embodiment of the continuity equation in a turbulent ocean than is $\nabla \cdot (\overline{\rho}\ \overline{\mathbf{u}}) = 0$, although, as we have just shown, the differences are of order 0.3% and so are quite negligible.

The separate roles of dianeutral advection and dianeutral mixing in causing water-mass conversion have been illustrated by maps of the relevant vertical derivatives of hydrographic variables on some neutral surfaces from each of the world's oceans in McDougall and You (1990). Maps of the dianeutral advection caused by small-scale mixing, $gN^{-2}\{\alpha\theta_{zz} - \beta S_{zz}\}$, and the rate at which vertical mixing changes the potential temperature on a neutral surface, $gN^{-2}\theta_z^3 \beta \frac{d^2S}{d\theta^2}$, (both terms are normalized by the vertical diffusivity, D), are plotted in Figures 4-6 . These maps show well-defined patterns in the magnitude and sign of these terms and there are many regions where the terms have the same sign. In these locations the rate of water-mass conversion achieved by vertical mixing is of opposite sign to that caused by the dianeutral advection alone. An inverse model that carries only vertical advection in the salinity or potential temperature conservation equation will tend to force the dianeutral advection to be the opposite sign to reality. This is explained in more detail by McDougall and You (1990) and will not be repeated here.

I wish to emphasize that there is no problem at all with the neglect of eddy diffusion terms from the continuity equation (called by Wunsch the "mass" or the "density" conservation equation), rather it is the omission of vertical diffusive terms from the tracer conservation equations that causes concern. In particular, if an inverse model carries just thermal wind equations and continuity equations, without considering any tracer conservation equations, then there can be no argument with the interpretation of the diapycnal velocities that one obtains from the model. One realizes that there must be some vertical mixing going on to cause this diapycnal flow, but since the diapycnal diffusivity does not appear in the two types of equations in one's model (thermal wind and incompressibility) the model is quite consistent without having to include the diapycnal diffusion. The problems referred to in the previous paragraph and described more fully in McDougall and You (1990) arise when a tracer or several tracers

are carried by the model. Then one has no choice but to include both interfacial diffusion and advection in order to construct a physically consistent model.

When is it Redundant to Carry both S and θ Conservation Equations?

One has a tendency to think that since the epineutral (and epipycnal) gradients of salinity and potential temperature are related, the conservation equations of S and θ must be linearly dependent. This is not the case, as can be readily appreciated by inspecting the form of Eq. (9). While the temporal and epineutral derivatives of S and θ are proportional (through the ratio α/β) the dianeutral advection terms are in the different ratio S_z/θ_z and the dianeutral diffusion terms are related by a third different ratio. It is only when the dianeutral advection is eliminated from the S and θ conservation statements to arrive at the form Eq. (11) that the S and θ equations are redundant. Another way in which a real redundancy can arise is if some kind of density conservation equation is used together with both the θ and S equations in the form of Eq. (9). Since density conservation equations (such as the "e" equation, Eq. (12)), are simply a linear combination of the S and θ equations, it is obvious how the linear dependance arises when all three equations are carried.

Wunsch and Minster (1982) carried a continuity equation and both S and θ equations and found that all three types of equations were effectively linearly dependent. As explained above, such linear dependence should not have been expected since their model carried both diapycnal advection and diapycnal diffusion. The reason for the near collinearly in their case would have been the dominance of both the S and θ equations by the same imbalances in the continuity equation:- both the S and θ equations were essentially repeated continuity equations with the mixing information buried in the noise. I believe that the simple procedure of subtracting a suitable mean salinity and a suitable mean potential temperature of each box before writing down the divergence forms of the conservation statements would have yielded very different results in their study: the rank of the model would have risen substantially, the lateral diffusivities would have been much larger, and quite possibly, the vertical diffusivities and diapycnal advection would have been better-determined.

Why use a neutral surface coordinate scheme rather than Cartesian coordinates?

The issue of the surfaces in which one assumes the lateral mixing to occur has been addressed earlier in this paper, and will not be repeated here. Rather, here some observations are offered on the advantages of casting the left-hand sides of the conservation statements in the neutral surface framework. These remarks will apply equally well to the advective and the divergence forms of the conservation statements and so apply to both the β-spiral and box inversions. Consider the steady-state conservation statement for say potential temperature, θ, in a region where the epineutral gradient of potential temperature is small in relation to the horizontal gradient, $\nabla_2\theta = \theta_x\mathbf{i} + \theta_y\mathbf{j}$, and where the neutral surface slopes significantly. The three-dimensional advection of θ is the same in both coordinate frames so that (from Eq. (9))

$$\left\{\mathbf{v}^L\cdot\left[\nabla_2 C - \nabla_n C\right] + \left[w-e\right]C_z\right\} + \mathbf{v}^L\cdot\nabla_n C + eC_z \;=\; h^{-1}\nabla_n\cdot\left(hK\nabla_n C\right) \;+\; \left[DC_z\right]_z. \qquad (40)$$

The terms in the curly brackets here sum to zero exactly, but if the neutral surface is significantly sloped, this cancellation can, and often does, represent the difference between two large numbers. In this situation, any uncertainty in the lateral velocity components will cause an unnecessarily large uncertainty in the left-hand side of this equation and so potentially upset the desired balance between advection (LHS) and diffusion (RHS) in Eq. (40). When the equations are weighted by their row-

norms, the equation error that is allowed in the inversion procedure will be unnecessarily large in the Cartesian formulation.

This is especially obvious in the linear combination of the S and θ conservation equations that is the 'e' equation, Eq. (12) (in loose terminology, this is the advective form of the 'density' conservation equation). In neutral surface coordinates, this represents a very direct relationship between the dianeutral advection, e, and mixing processes. Since the lateral velocity vector does not appear in this equation in the neutral surface framework, the inherent uncertainty in the inversion's lateral velocity vector can neither upset this balance nor can it contribute to the expected error of the equation. In this way the 'density' or 'e' equation is almost guaranteed to contain separate information to the other conservation statements that contain the lateral velocity vector. However, in the Cartesian framework this same equation does contain large balancing terms that do involve the lateral velocity vector, as in the curly bracket below,

$$\left\{ \mathbf{V}^L \cdot \left[\alpha \nabla_2 \theta - \beta \nabla_2 S \right] + \left[w - e \right] \left(\alpha \theta_z - \beta S_z \right) \right\} + e \left(\alpha \theta_z - \beta S_z \right) = \left[D_z \left(\alpha \theta_z - \beta S_z \right) \right]_z - K \left\{ C_b \nabla_n \theta \cdot \nabla_n \theta + T_b \nabla_n \theta \cdot \nabla_n p \right\}. \tag{41}$$

The terms in the curly brackets on the left of this equation sum to exactly zero. The individual terms in this bracket are frequently much larger than the other term on the left, namely, $e\left(\alpha\theta_z - \beta S_z\right)$, as the magnitude of $[w - e]$ is frequently much larger than that of e. The same point can be made regarding the equation for the diapycnal velocity, w^d, in the potential density framework compared with the Cartesian framework in that Eq. (20) does not involve the lateral velocity components whereas in the Cartesian form, Eq. (21), it does contain the do-nothing combination of terms, $\left\{ \mathbf{V}^L \cdot \nabla_2 \rho_\theta + \left[w - w^d \right] \frac{\partial \rho_\theta}{\partial z} \right\}$. Another appealing feature of the e equation in the neutral surface framework is that it does not contain the lateral Laplacian of any property. Since the lateral Laplacian is a relatively noisy quantity to estimate from data, this feature of the e equation, Eq. (12), augments the absence of the lateral velocity vector to suggest that the equation is a relatively noise-free connection between dianeutral advection and (mainly vertical) diffusion.

The above discussion of the merits of performing inversions in neutral surface coordinates has been focused on the implications for determining the strength of mixing processes. But even in models without mixing the different slopes of various surfaces affects the reference level velocities and other outputs. For example, Schott and Zantopp (1979) showed that the β-spiral technique gave reference level velocities at 1000 m that differed by 5 mm s^{-1} depending on whether potential density or steric anomaly (specific volume anomaly) was conserved by the inversion. In a box model inversion of sections in the North Atlantic, Rintoul and Wunsch (1991) have compared their model which used interfaces that were a close approximation to neutral surfaces with a previous model that used surfaces that were a coarser approximation to neutral surfaces. They found that the differences between the surfaces caused the poleward flow of intermediate water to decrease by 2.4 Sv and the equatorward flow of deep water to be reduced by the same amount. The poleward heat flux changed by 0.1×10^{15} W.

Some recommendations for box inversions

The above issues for box inversions are all intimately linked because (i) the tracer (especially salinity) conservation statements have been dominated by noise from the "mass" conservation equation (really the incompressibility equation), (ii) diapycnal diffusion coefficients have not been resolvable from the models and hence (iii) it has been pointless to add a conservation equation for potential temperature

since this would also have been linearly dependent with the existing two sets of equations. It is to be hoped that the above recommendation of forming divergence conservation statements for (a) volume and (b) for the deviations of tracer concentrations from suitable average values, will cure all three problems simultaneously. This simple procedure should make the tracer conservation statements become linearly independent of the continuity equations, giving the advective-diffusive balance of say salinity a chance of constraining both diffusivity coefficients and the mean velocity field. This will also result in an increase in the rank of the system of equations. Conservation equations of potential temperature and of other tracers should also be included in the inversions so as to extract the further independent information on mixing processes that these equations contain.

The interfaces that separate the boxes should be neutral surfaces in order to avoid the many error terms in Eq. (26) and Eq. (30) that arise due to mixing laterally along potential-density surfaces rather than along neutral surfaces. The previous subsection also shows that the neutral surface framework should be superior from the signal-to-noise viewpoint. In particular, the direct relationship between dianeutral advection and mixing processes can be obtained by taking the linear combination of α times Eq. (36) with the anomaly of potential temperature as the tracer, minus β times Eq. (36) with the salinity anomaly as the tracer.

Previous box inverse models have carried continuity equations for each box (despite the different labels of these equations as "density" or "mass" conservation statements), whereas β-spiral methods do not enforce the continuity equation. Perhaps this is the key difference between the two methods. Box models should continue to carry the volume conservation equations for each box (in the form of Eq. (36) with $C = 1$), and it may be necessary to recognize that the Lagrangian velocity components that appear in this equation contain the Stokes drift (Eq. (15)) in addition to the Eulerian-mean velocity that is obtained from thermal wind.

<u>On the choice of mean tracer value for each box</u>

Consider forming the salinity anomaly for each box by subtracting a mean salinity (say 35 psu) from all the salinities in all the boxes of a box model. The maximum value of S' for any particular box may be say 1 psu so that the uncertainty in the salinity conservation equation due to the continuity imbalance is improved by a factor of 35. However, this may not be enough of an improvement to guarantee that the salinity conservation statement enforces an advective-diffusive salinity balance. In the examples quoted above, the advective-diffusive salt balance would improve from being 1% of the residual equation error to being 35%. If this proved to be not enough of an improvement, then one would need to use tracer anomaly values that were referenced to a mean tracer value that is closer to the average tracer value of each box. In order to extract the advective-diffusive tracer balance from the divergence form of the conservation equations, the most appropriate mean value of the tracer, $\overline{C}$, to subtract from C to form the new variable, C', is the mean value along all six faces of the box. The diffusive terms on the right-hand sides of Eqs. (36), (38) or (39) would still be evaluated with the original values of the tracer variable since there is no gain in accuracy to be had by changing variables here, but the left-hand sides of these equations would be evaluated using the refined tracer anomaly variable. In the case of salinity, the variation of the salinity over the six sides of a box may be 0.1 psu so that the influence of the continuity equation in the salinity balance will be reduced by a further factor of ten or more, so that instead of the continuity uncertainty accounting for 99% of the error in this equation, it could account for no more than 20% of the error. In this way the salinity conservation statement will tend to represent the desired advective-diffusive balance.

By choosing the box-average tracer value as the mean that is subtracted from each box, one has essentially forced the divergence form of the conservation statements to have the same balance that is present in β-spiral methods, namely the advective-diffusive balance. Unlike the β-spiral methods however, this proposed box-inverse method also carries the continuity equation (Eq. (36) with $C = 1$). In this way, the method recommended above can be interpreted as a β-spiral method, but with the inclusion of the continuity constraint. One could of course add the normal β-spiral equations to a box inverse method, but the grids on which the velocity components are evaluated are different for an advective and a divergence grid. The above procedure achieves the same physical balances in the equations as would a combined β-spiral and box inverse model, but without any complications due to the different grids, since one uses the box-model grid throughout.

When using anomalies from box-averaged data rather than from a single constant value, the sum of the tracer conservation equations over more than one box does not have a physical interpretation (Rintoul and Bindoff, personal communications, 1991). Using salinity as an example, if a single mean salinity of say 35 psu is used to form the salinity anomaly variable for the left-hand side of the conservation equations, the sum of these equations over many boxes does not amount to the conservation of salt over the boxes, but it does represent the conservation of a different variable, namely (S - 35 psu). But with a mean salinity that varies from box to box, there is no such interpretation. Should one worry about this? I think that this is not a concern for the following two reasons. First, by regarding the above procedure as a β-spiral method with the added continuity constraint, the issue of what the sum of the salinity conservation statements represents does not seem so pressing since this has not been part of past β-spiral methods. Second, if after performing an inversion one calculates the total salt imbalance (not salt anomaly imbalance) summed over all the boxes, the answer will be dominated by the errors in the continuity equation for each box multiplied by the mean salinity of that box. This will be very close to the global-averaged salinity times the global-averaged volume flux imbalance, no matter what mean salinity is chosen for the mean value of each box in the inversion. For these reasons I believe that using a salinity anomaly that is defined differently for each box is a viable procedure. Of course, proof of the pudding will be in the eating.

CONCLUSIONS

This paper has derived the conservation equations for scalars (including potential density and spiciness) with respect to both neutral surfaces and potential-density surfaces, and in both the advective form which is applicable to the β-spiral method, and in the divergence form that is used in box-model inversions. In the limited space available, conservation statements for potential vorticity were not addressed. The salient findings of the paper are listed below.

• Scalars are advected by the Lagrangian-mean velocity vector rather than the Eulerian-mean velocity that appears in the geostrophic and thermal wind relations, and it may well be important to recognize the distinction between these velocities. A simple parameterization for the difference velocity (the Stokes drift) is proposed and this should be easy to implement in inverse models. Even when this Stokes drift is relatively small, it may well be significant for the lateral advection of tracers because it will be much more closely aligned with the epineutral tracer gradient than will be the Eulerian-mean lateral velocity vector. The magnitude of the estimated Stokes drift is displayed in Figures 2 and 3 for a single neutral surface in the world ocean, and values in excess of 1 mm s^{-1} are common. A lateral velocity of this magnitude down the epineutral tracer gradient causes as much water-mass conversion as vertical mixing

processes. This highlights the importance of the distinction between Lagrangian and Eulerian velocities.

• The nonlinear nature of the equation of state has been shown to cause significant errors in the diapycnal velocity when it is deduced from the commonly used advection-diffusion balance for potential density. The potential density variable is significantly nonconservative and more care must be taken when writing a conservation equation for potential density. The terms that have been omitted from the potential density conservation equation in the past are (see Eq. (20)), (i) a term proportional to the vertical diffusivity that is largest in the upper 1000 m of the water column, (ii) a term that depends on the epineutral flux divergence of potential temperature and arises because potential density varies along neutral surfaces, (iii) cabbeling and (iv), thermobaricity. Each of these terms are too large to be ignored in various regions of the ocean.

• The omission of the nonconservative terms in the conservation equation of potential density results in either an overestimate of the diapycnal velocity or an underestimate of the vertical diffusivity (for $w^d > 0$), or, quite probably, a combination of both. The use of a simplistic potential density equation also affects the way mixing processes are extracted from the potential temperature or salinity conservation equations, as is illustrated in Eq. (29).

• Thermobaricity and cabbeling were found to be quite strong in the North Atlantic and Southern Oceans, causing contributions to the dianeutral downwelling velocity of order -1×10^{-7} m s^{-1}. These in turn make a large impact on the conservation equations of scalars in these regions and probably also cause significant vortex stretching in the conservation equation for potential vorticity. Of the two processes, thermobaricity is rather smaller than cabbeling except in the Antarctic Circumpolar Current where it is at least as large as cabbeling and also of the same sign.

• The contributions of both thermobaricity and cabbeling to water-mass conversion at a thermoclinic front have been compared with the epineutral mixing of potential temperature. The ratios of these processes were found to be significant and to be proportional to the changes in potential temperature and pressure across the front (see Eq. 13). In the Antarctic Circumpolar Current each of these three processes were found to be equally important, especially because the net water-mass conversion achieved by the lateral diffusion term averages to zero whereas cabbeling and thermobaricity have a consistent sign across the front.

• Many of the pitfalls with forming a potential density conservation equation are avoided if one uses the S and θ equations separately rather than attempting to form a "density" conservation equation. The remaining differences are due to the difference between the epineutral flux divergences and the epipycnal flux divergences of S and θ, as in Eq. (31).

• Conservation equations for conservative variables (like potential temperature and salinity) are affected by the nonlinear nature of the equation of state when a model's lateral mixing is directed along potential-density surfaces rather than along neutral surfaces. The differences between these two lateral mixing parameterizations are explored and are documented in Eqs. (26) and (30). Several terms arise that are significant when the reference pressure of the potential density is significantly different to the in-situ pressure, and two terms remain even when this pressure difference is zero. The difference between epineutral and epipycnal mixing of tracers is important at regions of large epineutral gradient of potential temperature (i.e. at thermoclinic fronts).

• The 'orthogonal' variables on the S-θ diagram, spiciness and Veronicity, are significantly nonconservative variables and the nonlinear terms that appear in their conservation equations are of similar magnitude to those that appear in the potential density equation. It has been shown that one cannot afford to write conservation equations for either potential density or spiciness as though they were conservative variables.

• Box inverse models need their conservation equations cast in the divergence (or flux) form, and it is shown that in this form, the conservation equation for 'density' or 'mass' is really simply a volume integral of the incompressibility equation.

• One often reads that a box inverse model has not required vertical mixing in order to explain the data. Here it has been argued that these models have been set up in such a way that the signatures of all types of mixing processes are well hidden behind the noise due to errors in the continuity equation. The imbalance in the continuity equation is directly reflected in the tracer conservation equations, and is a consequence of the need to write the conservation equations in the divergence form. This causes the tracer equations, and particularly the salinity equation, to be linearly dependent with the continuity equation, so causing a reduced rank in the system of equations. This is similar to the well known problem that, if the mass flux across an ocean section is not zero, the heat flux across the section is different if the temperatures are measured in Kelvins rather than in degrees Celsius.

• A very simple solution to this deficiency of present box-inverse models is proposed: it is to subtract a suitable mean value from the values of a tracer before the conservation equations are evaluated. This should have the effect of (i), increasing the rank of the solution, (ii), extracting information on the advection and diffusion of salinity, thereby constraining both the interfacial advection and diffusion, and (iii), making the conservation equation of potential temperature independent of that of salinity.

• It is argued that the neutral surface framework provides the best link between advection and diffusion of tracers in that the individual terms that comprise the advection of tracer are not large in comparison with the total advection of tracer. In this way, uncertainties in the lateral velocity vector do not introduce as much uncertainty into the left-hand side of these conservation equations as they do in Cartesian coordinates. This is especially true of the e equation, Eq. (12) that is totally independent of the lateral velocity vector. This e equation has the added advantage that it does not have a lateral Laplacian term (which is relatively noisy) on the right-hand side.

Acknowledgments

It is a pleasure to thank Drs Peter McIntosh, Nathan Bindoff and Stephen Rintoul for helpful conversations and Dr David Jackett for kindly providing Figures 2 and 3.

REFERENCES

Bauer, E., and G. Siedler, 1988: The relative contributions of advection and isopycnal and diapycnal mixing below the subtropical salinity maximum. *Deep-Sea Res.*, **35**, 811-837.

Brown, E., W. B. Owens and H. L. Bryden, 1986: Eddy-potential vorticity fluxes in the Gulf Stream recirculation. *J. Phys. Oceanogr.*, **16,** 523-531.

Bryan, F., 1987: Parameter sensitivity of primitive equation ocean circulation General Circulation Models. *J. Phys. Oceanogr.*, **17,** 970-985.

Fu, L-L., 1986: Mass, heat and freshwater fluxes in the South Indian Ocean. *J. Phys. Oceanogr.*, **16,** 1683-1693.

Haidvogel, D. B., and P. B. Rhines, 1973: Waves and circulation driven by oscillatory winds in an idealized ocean basin. *Geophys. and Astrophys. Fluid Dyn.*, **25**, 1-63.

Hogg, N. G., 1987: A least-squares fit of the advective-diffusive equations to Levitus atlas data, *J. Mar. Res.*, **45,** 347-375.

Holland, W. R., and P. B. Rhines, 1980: An example of eddy-induced ocean circulation. *J. Phys. Oceanogr.*, **10,** 1010-1031.

Jackett, D. R., and T. J. McDougall, 1985: An oceanographic variable for the characterization of intrusions and water masses. *Deep-Sea Res.*, **32**, 1195-1207.

Joyce, T. M., C. Wunsch and S. D. Pierce, 1986: Synoptic Gulf Stream velocity profiles through simultaneous inversion of hydrographic and acoustic Doppler data. *J. Geophys. Res.*, **91**, 7573-7585.

Lee, J. H., and G. Veronis, 1991: On the difference between tracer and geostrophic velocities obtained from C–SALT data, *Deep-Sea Res.*, in press.

Levitus, S., 1982: Climatological atlas of the world ocean. *NOAA Prof. Pap.* **13**, Rockville, Md., 173pp.

McDougall, T.J., 1987a: Thermobaricity, cabbeling, and water-mass conversion. *J. Geophys. Res.*, **92**, 5448-5464.

McDougall, T.J., 1987b: Neutral surfaces. *J. Phys. Oceanogr.*, **17,** 1950-1964.

McDougall, T. J., 1988: Neutral-surface potential vorticity. *Prog. Oceanogr.*, **20**, 185-221.

McDougall, T. J., and C. J. R. Garrett, 1991: Conservation equations in a turbulent ocean. submitted to *Deep-Sea Res.*

McDougall, T. J., and D. R. Jackett, 1988: On the helical nature of neutral trajectories in the ocean. *Prog. Oceanogr.*, **20**, 153-183.

McDougall, T. J., and Y. You, 1990: Implications of the nonlinear equation of state for upwelling in the ocean interior. *J. Geophys. Res.*, **95**, 13263-13276.

Memery, L., and C. Wunsch, 1990: Constraining the North Atlantic circulation with tritium data. *J. Geophys. Res.*, **95**, 5239-5256.

Middleton, J. F., and J. W. Loder, 1989: Skew fluxes in polarized wave fields. *J. Phys. Oceanogr.*, **19,** 68-76.

Olbers, D. J., M. Wenzel and J. Willebrand, 1985: The inference of North Atlantic circulation patterns from climatological hydrographic data. *Revs. Geophys.*, **23**, 313-356.

Olbers, D. J., and M. Wenzel, 1989: Determining diffusivities from hydrographic data by inverse methods with applications to the circumpolar current. in *Oceanic Circulation Models: Combining Data and Dynamics*, D. L. T. Anderson and J. Willebrand (eds), pp95-139.

Rhines, P. B., 1982: Basic dynamics of the large-scale geostrophic circulation, 1982 Summer Study Program in Geophysical Fluid Dynamics, Woods Hole Oceanographic Institution, WHOI-82-45, pp.1-47.

Rhines, P. B., 1986: Lectures on ocean circulation physics. in *Large-Scale Transport Processes in Oceans and Atmospheres*, J. Willebrand and D. L. T. Anderson (eds), pp 105-161.

Rhines, P. B., and W. R. Holland, 1979: A theoretical discussion of eddy-driven mean flows, *Dyn. Atmos. Oceans*, **3**, 289-325.

Rintoul, S. R., and C. Wunsch, 1991: Mass, heat, oxygen and nutrient fluxes and budgets in the North Atlantic Ocean. *Deep-Sea Res.*, in press.

Schlitzer, R., 1987: Renewal rates of East Atlantic Water estimated by in ersion of ^{14}C data. *J. Geophys. Res.*, **92,** 2953-2969.

Schlitzer, R., 1988: Modeling the nutrient and carbon cycles of the North Atlantic. 1. Circulation, mixing coefficients, and heat fluxes. *J. Geophys. Res.*, **93,** 10,699-10,723.

Schott, F., and H. Stommel, 1978: Beta spirals and absolute velocities in different oceans. *Deep-Sea Res.*, **25**, 961-1010.

Schott, F., and R. Zantopp, 1979: Calculation of absolute velocities from different parameters in the Western North Atlantic. *J. Geophys. Res.*, **84**, 6990-6994.

Veronis, G. 1972: On properties of seawater defined by temperature, salinity and pressure. *J. Mar. Res.*, **30**, 227-255.

Wunsch, C., 1984: An eclectic Atlantic Ocean circulation model. I. The meridional flux of heat. *J. Phys. Oceanogr.*, **14,** 1712-1732.

Wunsch, C. and B. Grant, 1982: Towards the general circulation of the North Atlantic Ocean. *Prog. Oceanogr.*, **11**, 1-59.

Wunsch, C, D. Hu and B. Grant, 1983: Mass, heat, salt and nutrient fluxes in the South Pacific Ocean. *J. Phys. Oceanogr.*, **13,** 725-753.

Wunsch, C., and J-F Minster, 1982: Methods for box models and ocean circulation tracers: mathematical programming and nonlinear inverse theory. *J. Geophys. Res.*, **87**, 5647-5662.

OCEANIC MODEL TESTING

Claude Frankignoul
Laboratoire d' Océanographie Dynamique et de Climatologie,
Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France

1. Introduction

In general, it is difficult to objectively compare oceanic model results to observations, as observations are usually noisy and inaccurate, and the sample too limited, resulting in a small signal-to-noise ratio. The noise of the observed signal results from both measurement uncertainties and real disturbances, like mesoscale eddies when studying the general circulation, vortical modes when studying internal waves, or interannual fluctuations when considering the seasonal cycle; this noise mostly has large correlation scales, so that simple point-by-point comparisons with model predictions are difficult to interpret, and a multivariate viewpoint is needed. Moreover, and this is specific to the oceanic case, most motions are *externally forced* but precise information is lacking on the forcing and on the initial conditions. Thus, even a perfect oceanic model will not provide predictions that are fully consistent with the observations, i.e., within their uncertainties, and it is necessary to distinguish between model inadequacies and the model response uncertainties caused by poor knowledge of the input data.

This applies in particular to tropical motions, which are primarily wind forced and could be simulated deterministically for the most part if the wind stress were accurately known. However, information on surface wind is sparse and noisy, and the bulk formulae used to estimate the wind stress are rather inadequate. As equatorial model simulations have become increasingly realistic, visual comparisons, which reveal obvious differences, are unable to unambiguously identify model inadequacies, and more refined validation procedures are needed. Frankignoul et al. (1989) have thus developed a model testing method based on multivariate statistical analysis that could take into account explicitly all the observational uncertainties. So far, the method has been used to test and intercompare simulations of the tropical ocean. However, the approach is general and could be used to test dynamical models of the internal wave field, as suggested below.

In the next section, the model testing method is described in a general manner. A possible application to internal wave studies is then briefly discussed. In section 4, the method is illustrated by summarizing the testing and intercomparison of model simulations of the tropical Atlantic.

2. The multivariate model testing method

The simulation of oceanic motions can often be represented as an input-output problem. A model is driven by a prescribed forcing field $\mathbf{f}(\mathbf{x},t)$, which could include the initial conditions, and it predicts in particular the space/time behavior of a variable that is also observed. We denote the model prediction by $\mathbf{m}$, with $\mathbf{m} = \mathbf{L}(\mathbf{f})$. To test the model, $\mathbf{m}$ is considered in the whole $\mathbf{x}$-t domain where the model is believed to be realistic and observations, denoted by $\mathbf{d}$, are reliable, which is normally a highly dimensional space. The observations are usually inaccurate, so a probability region more properly describes the true oceanic state (Fig.1). As the forcing field is generally not

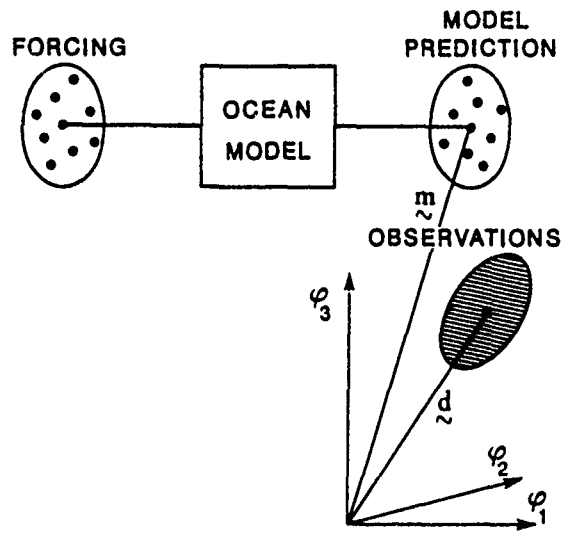


Fig. 1 Sketch of the effects of data uncertainties on the comparison between model predictions and observed data, from a multivariate statistical viewpoint.

known accurately, it should be considered as a (noisy) realisation of the true forcing field, and other equally plausible input data often are available (e.g., different wind stress products) or could be generated by Monte Carlo technics. This is represented in Fig.1 by a probability region for the true forcing <**f**>. Thus, several equally plausible model responses can be generated, and the model prediction **m** really should be considered as the sum of the (unknown) model response to the true forcing field, say <**m**>, and that resulting from the input errors, which is again represented by a probability region in Fig. 2. If the multinormal assumption holds, these uncertainties can be described by their error covariance matrix, denoted by **D** for the observations and **M** for the model predictions. The model testing problem can now be viewed as that of comparing two noisy vectors with unequal error covariance matrices. It should be stressed that a crucial, and often cumbersome, step in the procedure is the evaluation of the two error covariance matrices; this may require much data analysis and many simulations.

The agreement between the observations and the simulations is characterized by the misfit

$$T^2 = (\mathbf{m} - \mathbf{d})' (\mathbf{M} + \mathbf{D})^{-1} (\mathbf{m} - \mathbf{d}) \qquad (1)$$

which provides a measure of the differences between the two fields, weighted by the data uncertainties. If the null hypothesis that <**m**> = <**d**> holds, (1) is the appropriate test statistic which is distributed as a χ^2 variable (or a Hotelling T^2 variable if sample estimates of the error covariance matrices are used), and the usual acceptance and rejection rules can be applied. If the null hypothesis does not hold, T^2 is approximately distributed as a non-central χ^2 variable, and confidence intervals can be constructed to compare different models or model versions. Details are given in Frankignoul et al. (1989).

In practice, however, this cannot be done unless the dimensionality of the system is first strongly reduced. Indeed, the error models normally are very approximate (they are either estimated

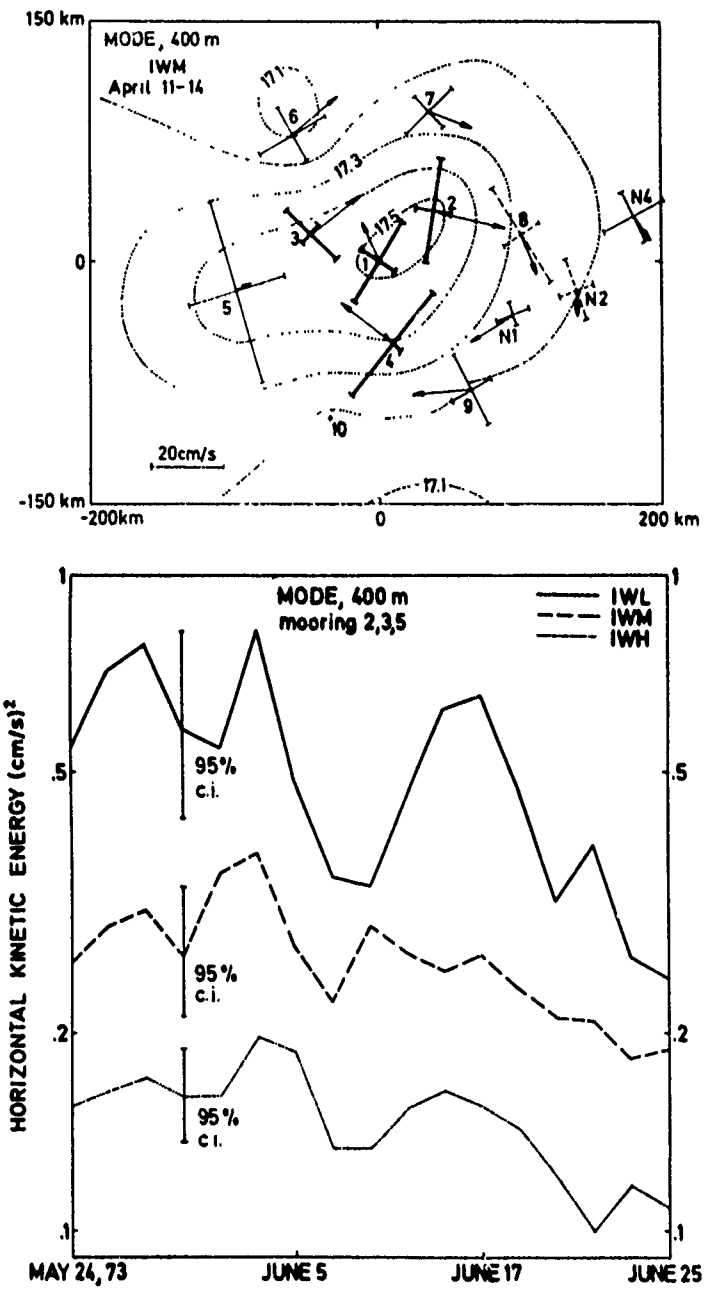


Fig. 2. *Top*: Ellipse principal axes of the internal wave horizontal kinetic energy (HKE) at 400 m depth for the internal wave frequency band (0.2-0.4 cph) in a 4-day period during MODE (the central mooring is at 28°N, 69°40'W). Arrows show 4-day mean current, and dotted lines corresponding isotherms. *Bottom*: Time series of HKE averaged over the line of moorings 2, 3 and 5 for three frequency bands in the internal wave continuum. From Frankignoul (1976).

from a limited sample or from a necessarily oversimplified error analysis), so that the inverse in (1), even if it exists, is dominated by unreliable information. A strong data compression is thus needed for the problem to be well-posed. This can be done efficiently using principal component analysis, as illustrated below. The main point is to perform the model-reality comparison in a sufficiently small orthonormal space where the main "trustable" features of both model simulations and observations are well represented, and the error models sufficiently reliable. Noisy details are filtered out, but of course so is a small part of the signal. Note that the test is very stringent as no model is expected to be perfect. However, much progress may be expected from an understanding of the discrepancies between models and observations, and the method is very efficient at distinguishing the true differences from those due to data uncertainties.

3. Application to internal wave studies

These concepts could be applied to the internal wave case. Suppose for example that we want to test a model of the dynamics of the internal wave field, which requires understanding their sources, sinks, and main interactions. Theoretical models of internal waves do not yet predict how the various energy sources control their spectral distribution, but plausible generation mechanisms have been suggested (wind forcing, topographic scattering, interaction with the mesoscale shear flow, forcing by the barotropic tides, etc.), and predictions of the spatial and temporal evolution of the averaged internal wave properties could be obtained. In particular, the variations of the total internal wave energy E(x,t), should obey the radiation balance equation (Muller and Olbers, 1975) which can be written in the approximate form

$$(d_t + \mathbf{U}\cdot\nabla)\, E = S^i(\mathbf{U}) + S^i(\text{atmosphere}) + S^i(\text{bottom}) + \ldots + S^d \qquad (2)$$

where $\mathbf{U}(x,t)$ is the mean current, S^i denotes the source terms, and S^d the dissipation. If models are available for the source terms, they can be tested by comparing the predictions from (2) with corresponding observations. Internal wave spectra have been shown to be modulated both on the seasonal scale (Fig. 2), and on short time and space scales (Fig. 3); also, evidence of energy propagation and/or transfer has been found (Fig. 4). These observations contain critical clues on internal wave energy sources and sinks, but their interpretation has been disappointing, as the dynamics is complex and the signal-to-noise ratio very low.

We believe that the key to a successful interpretation will be a multivariate model testing approach where predicted changes are compared to the observed ones, as this strongly enhances statistical significance in model-observation comparisons (see, e.g., Hasselmann, 1979). However, the input in (2) will be poorly known: the wind stress is difficult to observe at the internal wave scales, the mean shear can only be coarsely estimated from current-meter data, the bottom topography and bottom currents are inaccurately known, and little information will be available on the boundary conditions for the region of interest. Thus, there will be a large uncertainty in the forcing data that will have to be considered in addition to that of the internal wave spectral estimations, when testing the theoretical predictions. The method of section 2 should then be applicable, even if some adaptation to the problem will be needed. On the other hand, neglecting the effects of data uncertainties or the multidimensionality of the fields could result in erroneous conclusions.

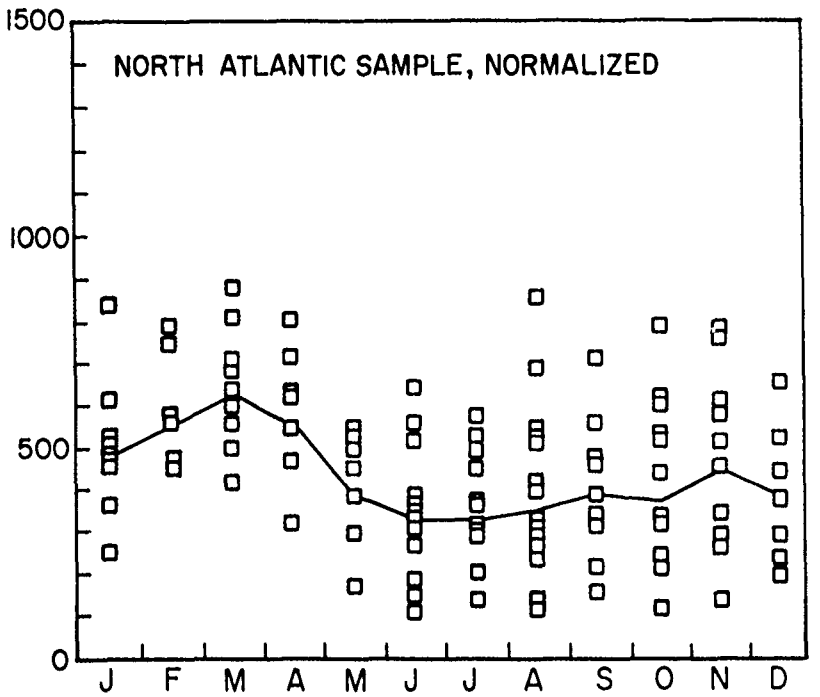


Fig. 3. Monthly high-frequency (0.1-2 cph) HKE from 17 data sets in the North Atlantic, normalized by the Brunt-Väisälä frequency, versus time of year. The solid line connects the monthly averages. From Briscoe (1984).

One remark is in order. Dynamical models of the internal wave field will unavoidably contain parameterizations and adjustable parameters (e.g., relaxation or dissipation time). A poor choice in the values of the arbitrary parameters may also be responsible for model-observation discrepancies, without impairing the model validity. Thus, it is of interest to include model tuning in the statistical method. If there is only one or two parameters, their influence on the misfit (1) can be easily found, with the optimal choice corresponding somehow to its minimum (see an example in Frankignoul et al., 1989). However, if there are many parameters, the problem becomes tedious. Recent efforts at combining model testing and parameter optimization by inverse methods are underway, and they may also prove useful in the internal wave case.

4. Application to the seasonal variability of the tropical Atlantic

To illustrate the model testing method, let us briefly consider its application to the numerical modeling of the tropical circulation. Specifically, we want to verify whether the OPA general circulation model of LODYC is able to simulate the mean seasonal changes in the surface dynamic topography of the tropical Atlantic, approximated by the 0-400 db dynamic height. The observed changes have been recently estimated from historical data by Duchêne and Frankignoul (1991), who performed an extensive error analysis to also estimate their error covariance matrix. The observations are compared in Fig. 5 (left) to the prediction of the OPA model (right). The simulation is the mean of a run forced by 1982-1988 monthly observed winds (Morlière and

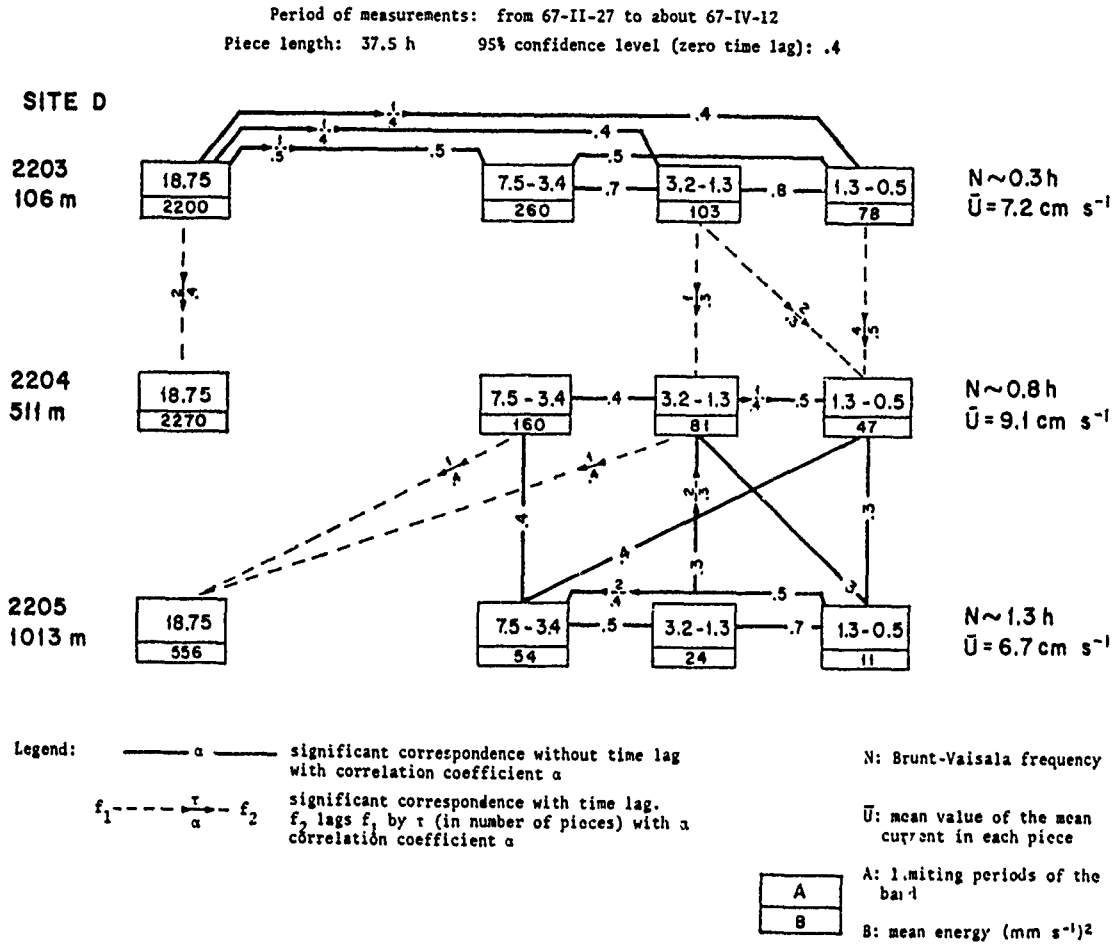


Fig. 4. Schematic representation of energy correspondences in frequency-depth space observed at Site D (39°20'N, 70°W) during a two-month period. From Frankignoul (1974).

Duchêne, 1991), so that the effects of the random wind stress errors and the interannual variability are automatically represented in the error covariance matrix that is calculated from the 7-year model response. The error associated with our lack of knowledge of the drag coefficient for the wind stress is also taken into account in a simplified manner by assuming a 20% random uncertainty in its value, as in Frankignoul et al. (1989).

The problem is to establish whether the two series of 12 monthly maps that describe the mean seasonal changes in space and time are consistent with each other, i.e., are within error bars. The overall dimension is high (number of grid points times number of months), and the noise correlated (due to data interpolation, forcing errors, etc.). Also, the sample used to estimate **M** is small, and **D** only is an idealized error model. Thus, the details of **D** and **M** are unreliable and a strong data compression is required. It is efficiently done by using common principal component analysis: the first four common empirical orthogonal functions (EOFs) (Fig. 6) are sufficient to represent about 90% of the mean simulated changes around the annual mean and 80% of the (more noisy) observed ones. Only two additional orthonormal vectors are needed to also represent the annual averages,

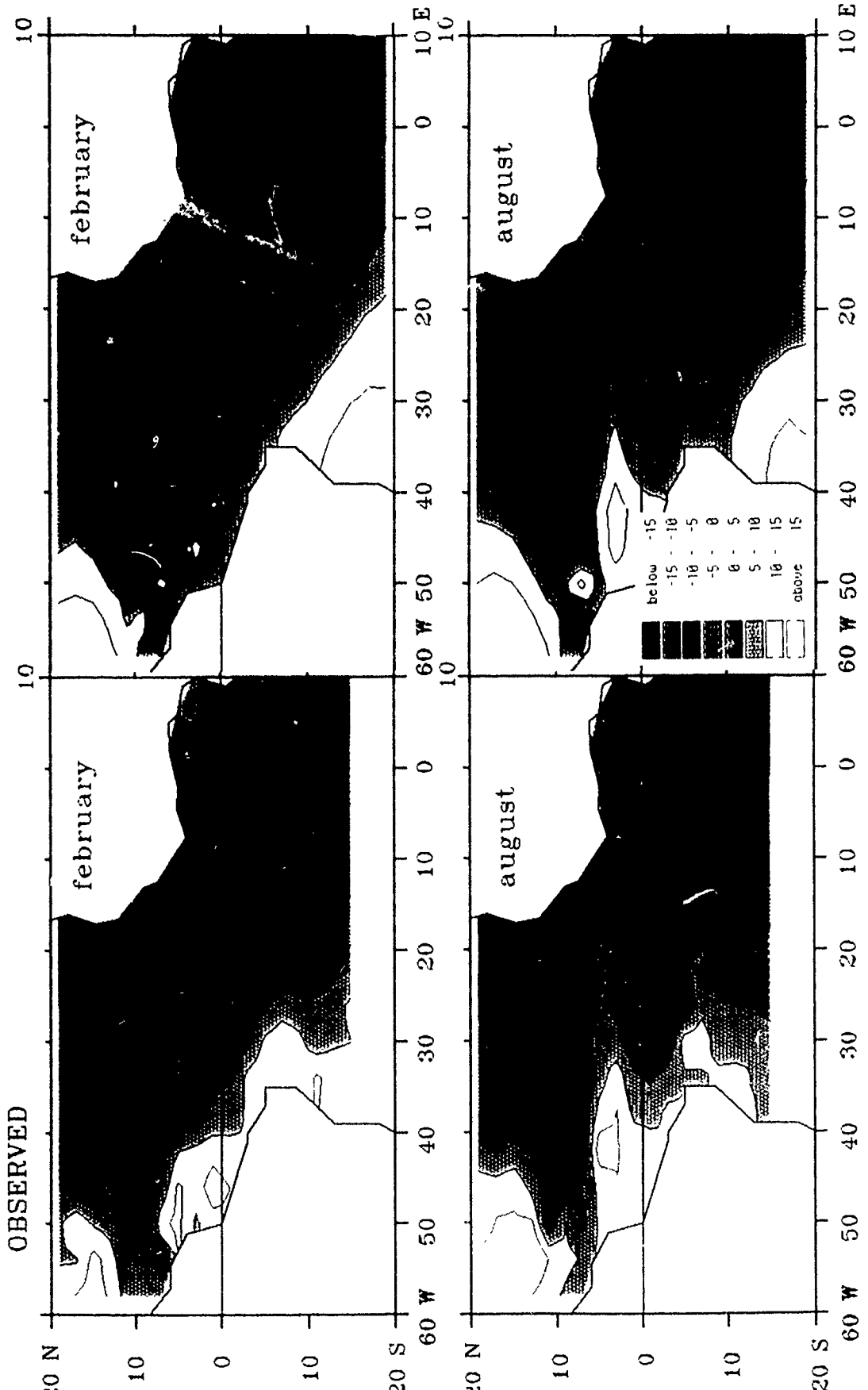


Fig. 5. *Left*: mean observed 0-400 db dynamic topography in dyn. cm. for February (top) and August (bottom). *Right*: Corresponding prediction using the OPA general circulation model. In both cases, the space-time mean has been subtracted.

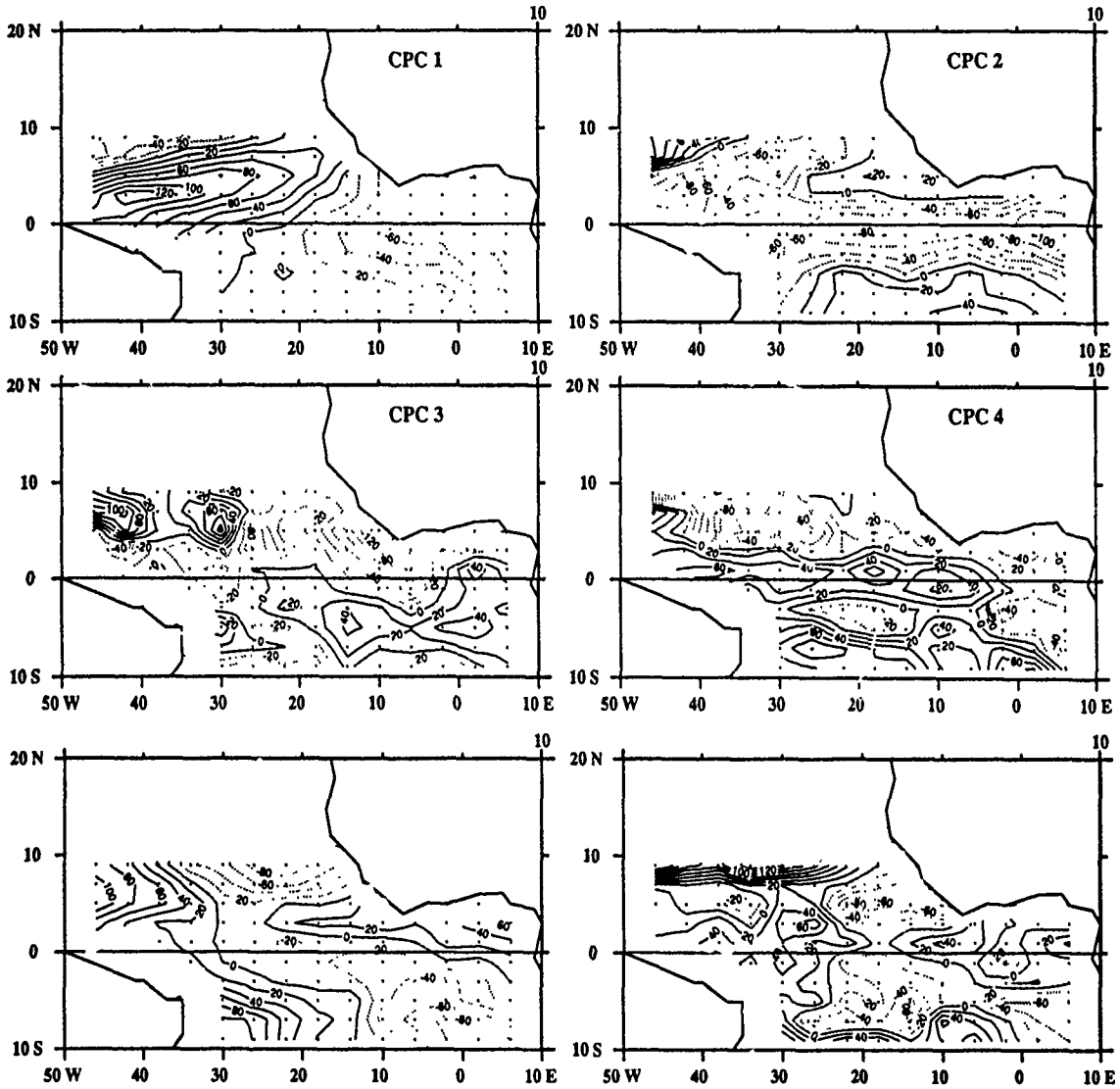


Fig. 6. The six orthonormal basis vector for the model-observation comparison. The top four vectors are the common EOFs of the seasonal variations.

which project poorly onto the common EOFs. In Fig. 7, the observed (dashed line) and simulated (continuous line) seasonal variations are represented in the reduced space, together with 95% confidence intervals estimated from the diagonal terms of the error covariance matrices, assuming normality. Note that the latter are based on univariate statistics and do not represent error correlation. Nonetheless, Figure 7 suggests that the differences between the observations and the simulations in Figure 5 cannot be entirely explained by the data uncertainties. This is confirmed by the results of the statistical test in Figure 8 with (left) and without (right) the annual mean. Note that a further data compression was done in the time domain by considering only four seasons (the dimension is thus 4×4 and 6×4, respectively).

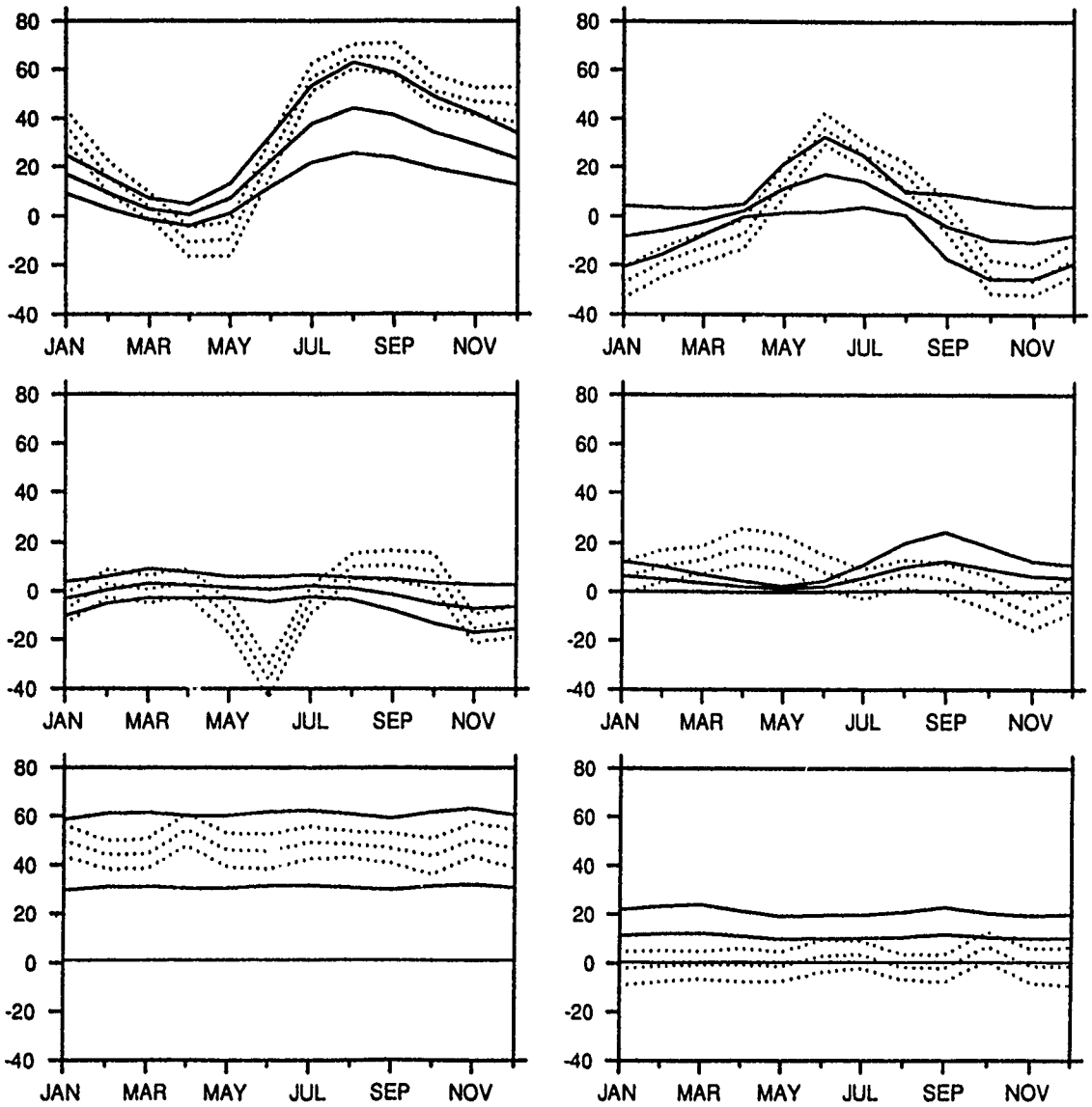


Fig. 7. Observed (dotted line) and modeled (continuous line) seasonal variations in the reduced space. The error bars are univariate estimates of the 95% confidence intervals.

The results show that the model is not consistent with the observations, as the misfit is much larger than the critical value T^2_{crit} for perfect consistency at the 5% level. Recall that (1) is a measure of the (square) difference between simulations and observations, normalized by the data uncertainties, and not an absolute measure of the fit; if the data uncertainties had been larger, T^2 would be smaller but the model would not perform better. The test only shows that the model-observation differences are about three (no mean) or four (with mean) times (i.e. $(T^2/T^2_{crit})^{1/2}$) larger than expected from the data uncertainties, at the 5% level. This indicates that there remains substantial room for model improvements. The results also show that the GCM performs better for

the mean seasonal changes around the annual mean than for the whole signal. This has also been found for all other models and variables that have been considered, and presumably reflects the fact that the long-term mean depends more on the representation of the dissipation processes, which is generally very approximate.

Figure 8 also represents for comparison the test results for two simpler models, the linear, multi-mode model of Cane (1984) and the 2-layer nonlinear LODYC model (Février, in preparation). The test values show that the tropical Atlantic data are accurate enough to distinguish between the performances of different oceanic models, and the OPA model performs *significantly* better then the two other models, consistently with its much higher sophistication. This illustrate the usefulness of the model testing approach, and the interest of quantifying as realistically as possible all the data uncertainties.

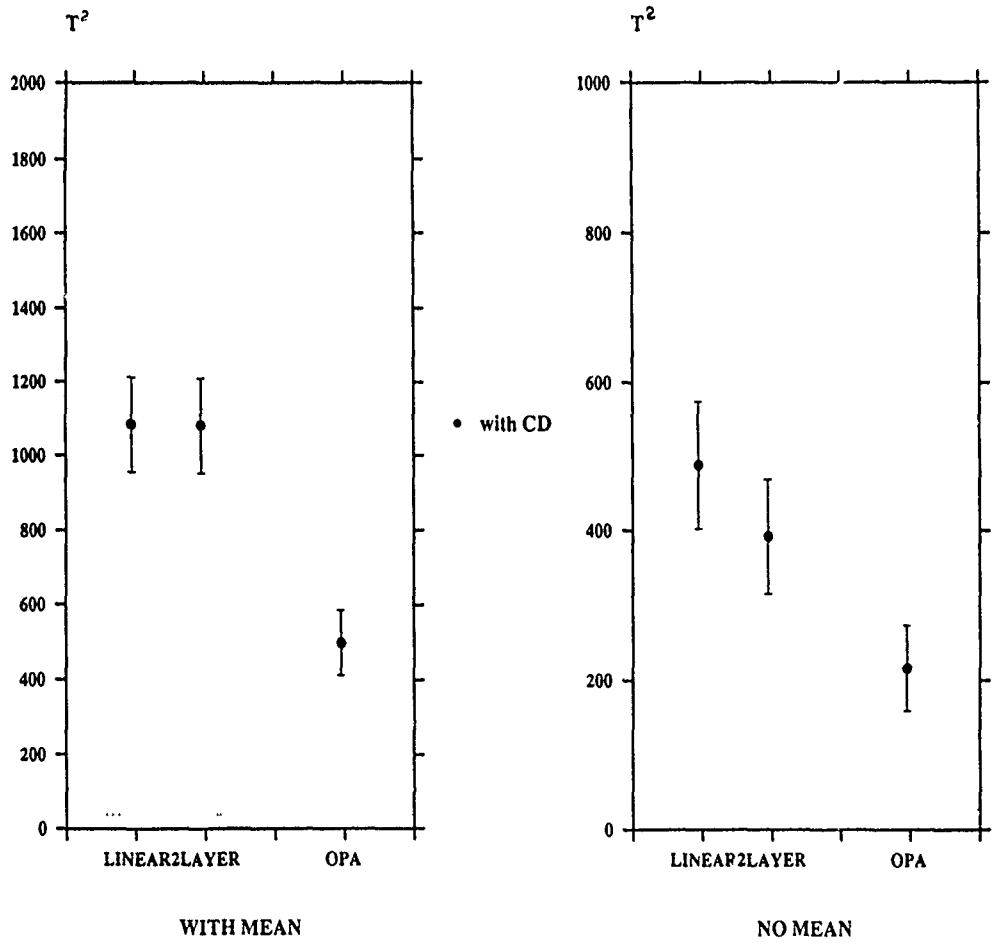


Fig. 8. Value of T^2 for three models (see text) used to simulate the seasonal cycle of surface dynamic topography in the tropical Atlantic with (left) and without (right) annual mean. The error bars represent approximate 95% confidence intervals, the dashed line the critical value for T^2 at the 5% level.

Acknowledgments

Some of the results reported here come from unpublished work performed at LODYC by Christine Duchene, Alain Morlière and Sabine Février who are gratefully acknowledged. I would also like to thank Peter Müller for stimulating discussions.

References

Briscoe, M.G., 1984: The monthly variability of the upper ocean internal wave energy: a progress report on the correspondence with wind stress. *Internal Gravity Waves and Small-scale Turbulence*, Proceedings of the 'Aha Hulikoa'a Winter Workshop, P. Muller and R. Pujalet, eds. Hawaii Institute of Geophysics, Honolulu, po. 129-150.

Cane, M. A., 1984: Modeling sea level during El Niño. *J. Phys. Oceanogr.*, **14**, 1864-1874.

Duchêne, C., and C. Frankignoul, 1991: Seasonal variations of surface dynamic topography in the tropical Atlantic: Observational uncertainties and model testing. *J. Marine Res.*, to appear.

Frankignoul, C, 1974: Preliminary observations of internal wave energy flux in frequency, depth-space. *Deep-Sea Res.*, **21**, 895-909.

Frankignoul, C., 1976: Observed interaction between oceanic internal waves and mesoscale eddies. *Deep-Sea Res.*, **23**, 805-820.

Frankignoul, C., C. Duchêne and M. Cane, 1989: A statistical approach to testing equatorial ocean models with observed data. *J. Phys. Oceanogr.*, **19**, 1191-1208.

Hasselmann K., 1979: On the signal-to-noise problem in atmospheric response studies *Meteorology of the Tropical Ocean*, D.B. Shaw, ed., Roy. Meteor. Soc., 251-259.

Morlière, A. and C. Duchêne, 1991: Progress in simulating and objectively evaluating the surface currents in the tropical Atlantic. *J. Geophys. Res.*, submitted.

Muller, P. and D.J. Olbers, 1975: On the dynamics of internal waves. *J. Geophys. Res.*, 80, 3848-3860.

ACOUSTIC IMPLICATIONS OF A NEW MODEL

Terry Ewart and Stephen Reynolds
Applied Physics Laboratory, University of Washington, MS HN10
Seattle, WA 98115

ABSTRACT

We attempt to show that the use of stochastic inverse methods allows us to precisely test proposed models of oceanic dynamical structure. It is the integrative property of the propagating acoustic wave that enables us to "view" ocean dynamics on scales that would be impossible with traditional ocean instrumentation. Unfortunately, because there has been a lack of propagation experiments conducted where the ocean dynamics are well understood, we are unable to fully demonstrate the use of the concepts we present. Additional investigations are required. Theoretical and numerical studies of acoustics in "numerical oceans" can provide new information on the scales of ocean dynamics that are important for sound propagation. In turn, this information will tell us how thoroughly we need to model ocean variability in order to predict propagation characteristics. We show that the important scales of ocean variance are much larger than the acoustic wavelength, when the ranges correspond to standard range/frequency combinations. Finally, we present some ideas for future work using acoustics to verify a "new model," and discuss the temporal and spatial scales for a possible experiment.

INTKODUCTION

Our comprehension of wave propagation in random media has progressed to the point where theoretical predictions of the fluctuations in sound waves that have passed through a medium with a known autocorrelation function of the acoustic index of refraction are quite accurate. Thus, the focus of the 'Aha Huliko'a meeting on "new" ocean dynamical and internal wave models is not only an important step in improving our understanding of ocean processes, but could herald a significant advance in our ability to test acoustic scattering predictions. In this paper, we have attempted to provide an overview of the elements of ocean dynamics required by the acoustician to make the ocean/acoustics link.

First, we present a brief review of the parameterization of ocean internal waves and finestructure used in acoustic propagation theory, and illustrate some of the progress being made in predicting acoustic scattering. Second, using a specific example, we demonstrate the power of ocean acoustic stochastic inverse methods. Note that, by stochastic inverse, we mean "imaging" the correlation or spectral properties of the index of refraction field rather than the index of refraction itself. We then discuss how one sets limits to the wavenumber/frequency bandwidth requirements of ocean models in the context of acoustic scattering. Finally, we present some ideas for a future coordinated ocean/acoustics research effort.

Ocean internal waves have a dramatic effect on sound propagation. Consider the numerical examples of acoustic propagation shown in Figure 1. The intensity of a sound wave propagating

in the depth/range plane is shown for depths to 4000 m and for the range interval from 35 to 65 km. Figure 1(a) has a vertical sound speed profile that is range-independent. Figure 1(b), in addition to the range-independent profile of 1(a), includes random internal wave induced fluctuations. Sharp foci of sound channel convergences are seen in 1(a). In 1(b), the internal

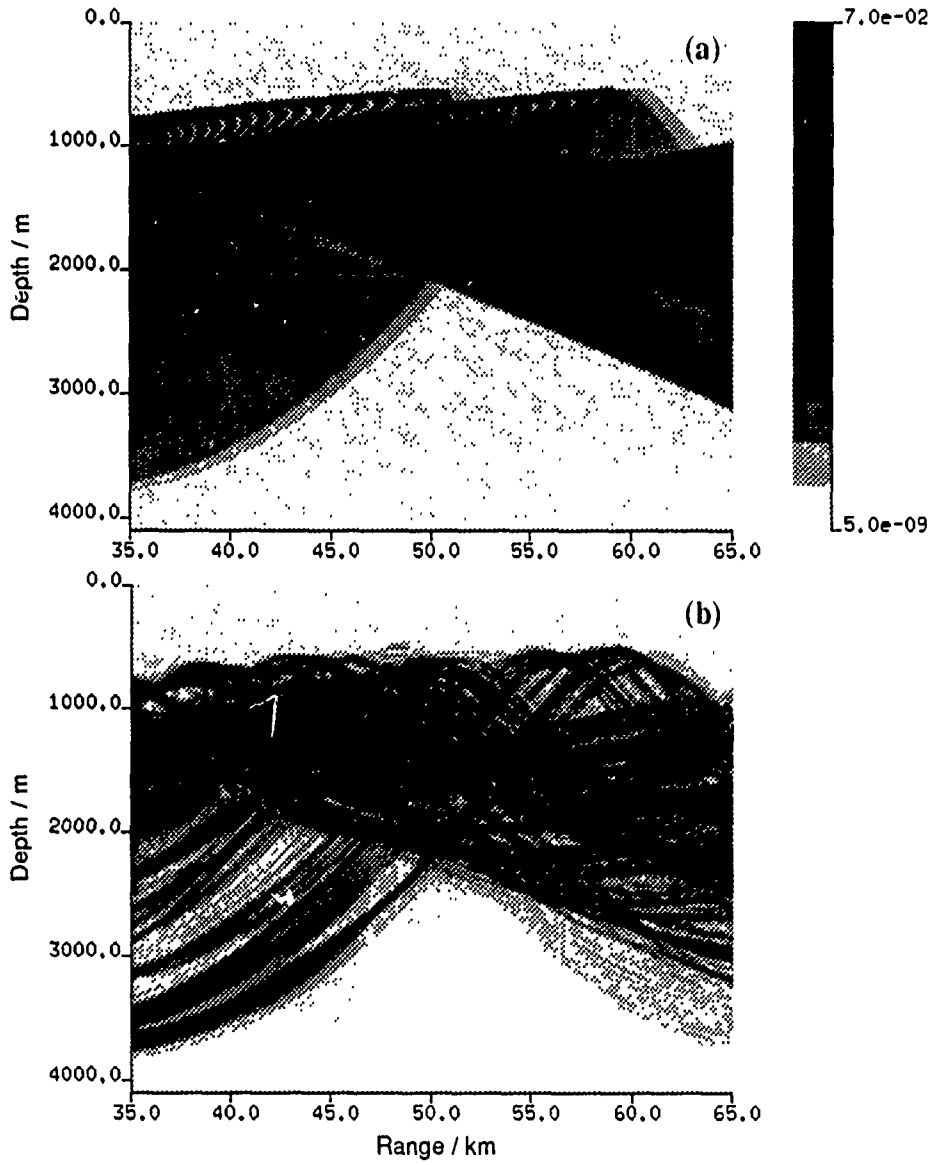


Figure 1. The intensity of a sound wave propagating in the depth/range plane is shown for depths to 4000 m and for the range interval from 35 to 65 km. The vertical sound speed is derived from the Munk canonical profile, and propagation is calculated using a wide angle parabolic equation (PE) code. The narrow beam source is above and near the axis of the sound channel. (a) Range-independent case, without internal waves. One sees sharp foci due to sound channel convergences. (b) As in (a) with internal wave variability added into the environment. The rms internal wave displacement is 7.3 m. The foci are smeared and diffuse and ribbons of intensity have formed.

wave perturbations have destroyed the foci and the sound field has broken up into ribbons of intensity. These ribbons of sound have been directly observed and are discussed in a paper by Uscinski and Potter (1988). We wish to predict the scattering statistics of these wave fields, and exploit them using stochastic inverse methods. In a discussion of internal waves, two points regarding acoustic propagation must be emphasized: (1) Because sound is the only form of energy (other than neutrinos) that can propagate long distances in the sea, it may be possible to monitor the ocean using inversion of information on acoustic travel time. (2) In order to study sound propagation in the stochastic ocean, the dynamical space/time statistics must be known.

SOUND VELOCITY FLUCTUATIONS IN A STOCHASTIC OCEAN

The coordinate system we use to discuss stochastic ocean behavior is shown in Figure 2. The two-point separation coordinates are $\xi = x_1 - x_2$, $\eta = y_1 - y_2$, $\zeta = z_1 - z_2$, $\tau = t_1 - t_2$, and the Fourier conjugate variables in the wavenumber/frequency domain are α_1, α_2, β, and ω. Propagation takes place in the x-direction.

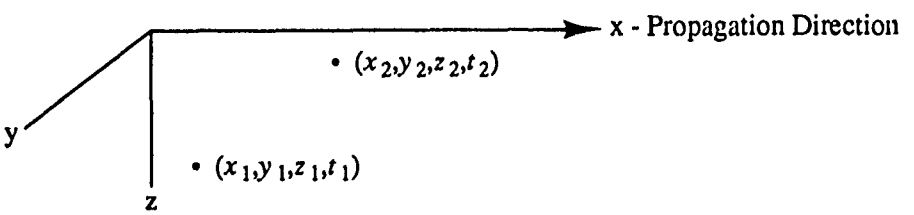


Figure 2. Coordinate system.

Following Uscinski (1986), the index of refraction, n, is written as the sum of a depth-dependent deterministic component, n_d, and a stochastic component, n_1.

$$n(x, y, z, t) = 1 + n_d(z) + \langle\mu^2\rangle^{1/2} n_1(x, y, z, t). \tag{1}$$

The rms index of refraction fluctuation, $\langle\mu^2\rangle^{1/2}$ is related to sound speed fluctuations, δC (arising from vertical displacements or velocities in the propagation direction),

$$\langle\mu^2\rangle^{1/2} = \frac{\langle\delta C^2\rangle^{1/2}}{C_o}, \tag{2}$$

where C_0 is the reference sound speed. We represent the two-point statistics of $n_1(x,y,z,t)$ by the power spectrum,

$$S(\alpha_1,\alpha_2,\beta,\omega). \tag{3}$$

In general, $\langle\mu^2\rangle$ is a function of depth. However, if a ray traverses a narrow range in depth, $\langle\mu^2\rangle$ can be considered constant. Subsequently, we will discuss experiments carried out at Cobb Seamount in the N.E. Pacific at a depth of 1000 m, and this is such a region. For the general case, the depth dependence provides little or no difficulty for simulations but adds complexity to

theoretical predictions of the moments of the acoustic field. We impose the traditional internal wave assumptions of horizontal isotropy and vertical homogeneity on the medium statistics. For simplicity, we will assume that locally the buoyancy frequency, and hence $<\mu^2>$ is depth independent. Under these conditions, the space/time spectrum of $n_1(x,z,t)$ may be written,

$$S(\alpha,\beta,\omega), \quad \text{where} \quad \alpha^2 = \alpha_1^2 + \alpha_2^2. \tag{4}$$

(We define the spectrum such that the integral over positive frequencies and wavenumbers is $\equiv 1$.) The horizontally isotropic medium correlation function is,

$$R(\xi,\eta,\zeta,\tau) = <\mu^2> \boldsymbol{F}\left\{S(\alpha,\beta,\omega)\right\}, \tag{5}$$

where $\boldsymbol{F}\{..\}$ indicates a Fourier transform. Note that $R(0,0,0,0) = <\mu^2>$. In the theoretical treatment of scattering, the medium is represented by a projection of the medium correlation function in the direction of wave propagation called the transverse correlation function (TCF). The TCF is obtained from the correlation function by

$$R_\perp(\zeta,\tau) = \int_{-\infty}^{\infty} \frac{R(\xi,\zeta,\tau)\,d\xi}{R(0,0,0)} = \int_{-\infty}^{\infty} \frac{R(\xi,\zeta,\tau)\,d\xi}{<\mu^2>}. \tag{6}$$

$R_\perp$ is the function we will discuss in the stochastic ocean context. The Fourier transform of $R_\perp$ is written $S_\perp(\beta,\omega)$, and is called the transverse spectrum. Note that η, the transverse horizontal coordinate, has been suppressed (i.e. $\eta = 0$). The η separations will not be treated here.

The ocean processes we consider are tides, internal waves and finestructure. Finestructure is the name given to the poorly understood portion of oceanic fluctuations in space and time that do not possess a wave-like dispersion relationship, but give appreciable variance in δC. As a specific example, we turn to consider the TCF used to study the acoustic propagation regime that existed during the Mid-Ocean Acoustic Transmission Experiment (MATE).

THE MATE TCF

At the 'Aha Huliko'a Meeting, our goal is to develop a "new" stochastic ocean model. The model we seek to replace is based on considerations of a spectrum of linear internal waves, and the parameters and spectral dependencies of the model were obtained by fitting the model to data sets, i.e. "data fits." Many authors have pointed out that linear internal waves alone cannot represent observations, and attempts to model the additional variance have been stymied by insufficient data in most experiments.

The Mid Ocean Acoustic Transmission Experiment (MATE) was designed to provide a detailed set of both temporal and spatial oceanographic and acoustic measurements. The MATE oceanographic setting (1000 m depth, Lat. 46°46'N., Long. 130°47'W) typifies open ocean conditions. The exception is a strong baroclinic tide caused by the presence of several seamounts. The oceanographic measurements during MATE were sufficient to overdetermine $S(\alpha,\beta,\omega)$. Details of the data analysis for finestructure and internal waves are found in Levine and Irish

(1981) and Levine et al. (1986). It was possible to obtain a model of $S(\alpha,\beta,\omega)$ for MATE based upon a fit to the various projections of S. We will now show that insight into the models can also be obtained from the MATE acoustical measurements.

The general form of the TCF is written for the case of separable vertical and time correlations as

$$R_{\perp,IW,FS}(\zeta,\tau) = \int_{-\infty}^{\infty} \frac{R_{IW,FS}(\xi,\zeta,\tau)\,d\xi}{R_{IW,FS}(0,0,0)} = L_{p;IW,FS}\,\frac{\sigma_{IW,FS}(\zeta)}{\sigma_{IW,FS}(0)}\,\frac{\psi_{IW,FS}(\tau)}{\psi_{IW,FS}(0)}, \text{ where} \qquad (14)$$

$$L_{p;IW,FS} = R_{\perp,IW,FS}(0,0).$$

For internal waves, following Uscinski (1980) we write

$$R_{\perp,IW}(\zeta,\tau) = 2\,G_o^{-1}H_o^{-1}\int_{\omega_i}^{\omega_n}\frac{G(\omega)}{r_\omega}\left[\int_{\beta_c}^{\infty}\frac{H(\beta)}{\beta}\cos(\beta\zeta)d\beta\right]\cos(\omega,\tau)\,d\omega, \qquad (15)$$

where $G(\omega)$ and $H(\beta)$ are obtained from the model presented in Levine et al. (1986); r_ω arises from the internal wave dispersion relationship,

$$r_\omega^2 = \frac{\alpha^2}{\beta^2} = \frac{\omega^2-\omega_i^2}{\omega_n^2-\omega^2}, \text{ and}$$

$$G_o = \int_{\omega_i}^{\omega_n} G(\omega)d\omega, \quad H_o = \int_{\beta_c}^{\infty} H(\beta)\,d\beta. \qquad (16)$$

ω_i and ω_n are the inertial and buoyancy frequencies, respectively. β_c is the vertical wave number corresponding to the lowest internal wave mode. The model in this form is written as

$$G(\omega) = \frac{(\omega^2-\omega_i^2)^{1/2}}{\omega^p}, \quad H(\beta) = \frac{\beta_*}{\beta_*^2+\beta^2}, \quad \beta_* = t\,(\omega_n^2-\omega^2)^{1/2}. \qquad (17)$$

With the exception of the variable spectral slope, p, this model is the same as that of Desaubies (1976). The Desaubies and GM formulations (for GM see Munk (1981)) specify $p = 3$. Also, this model uses a continuous vertical wavenumber representation rather than the modal decomposition used by GM. The parameter, t, is the bandwidth parameter of Desaubies. Note that the ζ-τ or β-ω functions are separable. This fact simplifies both theory and numerical simulation.

For finestructure, we use a form based upon that of Levine and Irish (1981). They postulated two processes: one at low wavenumbers and low frequencies that is characterized by a slow decay compared to the inertial period, and another process that is characterized by high wavenumbers and frequencies, that is modulated by the internal wave field. The high wavenumber process has a very low variance, and thus little effect on acoustic propagation. It has been discussed by Ewart et al. (1983), and will not be mentioned here except in relation to the acoustic phase correlations. The low frequency finestructure model has an asymptotic dependence in the spectral domain of α^{-2}, β^{-2}, and ω^{-2}. The forms of R_{FS} and $R_{\perp,FS}$ are

$$R_{FS}(\xi,\zeta,\tau) = \exp\left\{-\left[\left[\frac{\xi}{L_H}\right]^2 + \left[\frac{\zeta}{L_v}\right]^2\right]^{1/2}\right\} e^{-|\tau/\tau_0|}, \text{ and}$$

$$R_{\perp,FS}(\zeta,\tau) = L_{p;FS}\,\frac{\sigma_{FS}(\zeta)}{\sigma_{FS}(0)}\,\frac{\psi_{FS}(\tau)}{\psi_{FS}(0)} = L_{p,IW}\,\frac{\zeta}{L_v} K_1(\zeta/L_v)\, e^{-|\tau/\tau_0|}. \qquad (18)$$

We have taken $L_{p;FS} = 2{\cdot}L_{H,IW} = L_{p,IW}$; K_1 is the K_1 Bessel function.

Armed with the specific form of the TCF, we now proceed to discuss the observed acoustic fluctuations and stochastic inverse predictions. For a more general discussion of the forward problem, see Ewart (1986).

A PHASE STOCHASTIC INVERSE

For fifteen days during the time that MATE oceanographic measurements were made, acoustic pulses having center frequencies near 2, 4, 8, and 13 kHz were transmitted along an 18.1 km path near 1000 m depth. These transmissions were made between a fixed set of co-located transmitters and fixed, spatially separated receivers (four receivers were located at the corners of a rectangle 3 m high by 235 m in a plane transverse to the propagation path; see Ewart and Reynolds (1984)). With a maximum angle of just over 3°, the path was nearly horizontal. The transmission path, and a single realization of the density field from source to receiver is shown in Figure 3. We wish to obtain the time-varying spectrum of this density field by stochastic inverse methods.

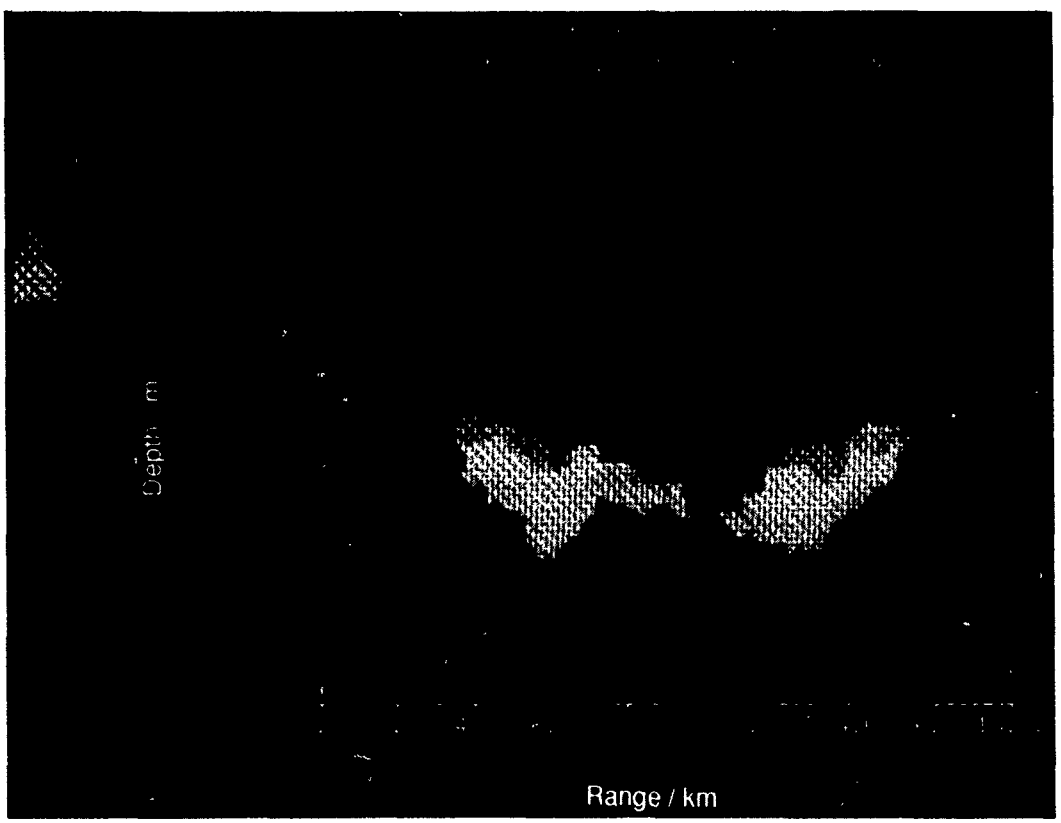


Figure 3. A single realization of the density field along the MATE transmission path from source to receiver. The measurements were obtained from a depth cycling run of SPURV. The potential density shades range from 27.3 (top) to 27.5.

We can learn a great deal from the temporal correlations of the acoustic phase. (Aside: in a geometric scattering environment, pulse travel time and phase are interchangeable.) The tidal, internal wave and finestructure processes are clearly seen in the phase spectrum from the 15 day 2 kHz data set (Figure 4). The phase spectra for the other frequencies of MATE are virtually identical out to temporal frequencies well above the buoyancy frequency, indicating the geometric nature of the phase fluctuations. The spectrum has been expressed in $<\mu^2>$ units by multiplying the measured travel times by $C_o / (Lp\ R)^{1/2}$, with $C_o = 1480$ m/s, $L_p = 4600$ m, and $R = 18.1$ km. The validity of this conversion relies upon the travel time being geometric. The integral of the spectrum is $<\mu^2>$. The diurnal, semidiurnal, and quarter-diurnal (overtone of the semidiurnal) tidal lines are evident, as is the sharp drop-off at ω_i and ω_n. The dashed lines indicate a fit to the deterministic tides using a simultaneous deterministic/stochastic inverse to a model that includes a trend function, tides, finestructure and internal waves. The finestructure and internal waves were modeled as the Fourier transforms of Equations 15 and 18, respectively. (If the phase is geometric, the phase spectrum is related to the Fourier transform of the TCF; see

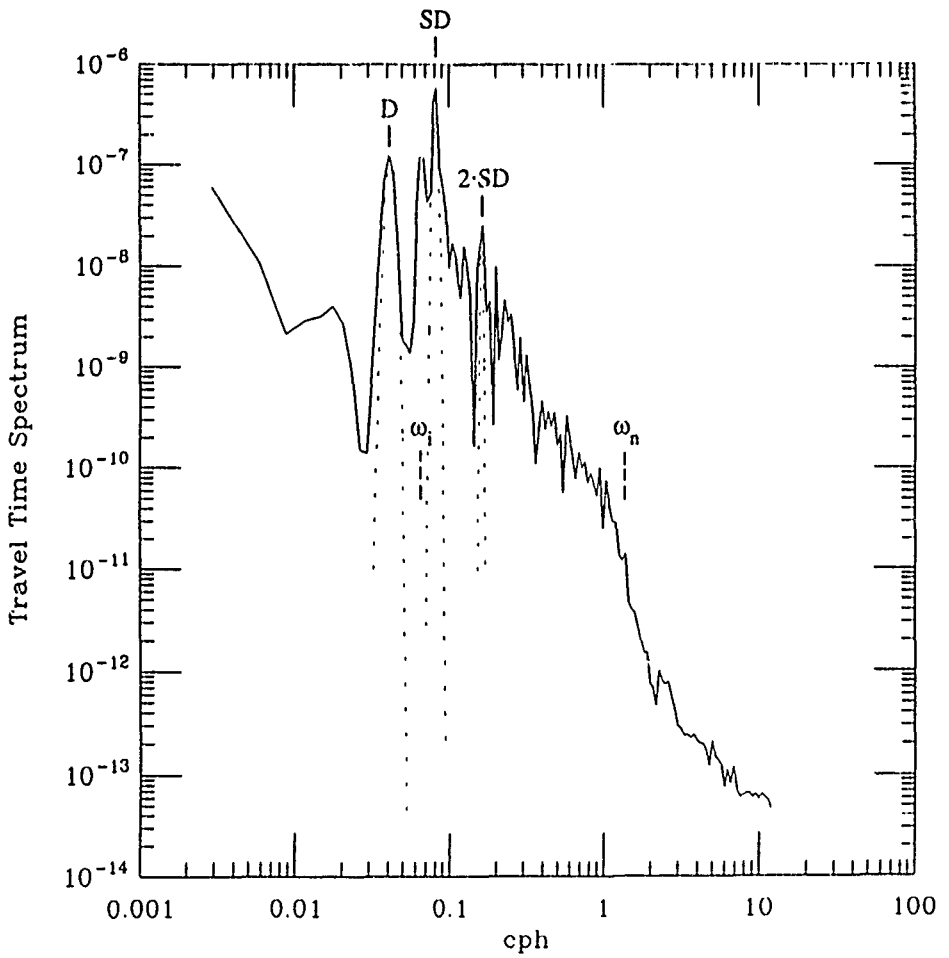


Figure 4. MATE 2 kHz travel time (phase) spectrum in $<\mu^2>$ units per cycle per hour. The spectral estimate was computed using the DPSS method described by Slepian (1978). Four discrete, prolate spheroidal windows were used. The dashed lines are the tidal components obtained by the fit described in the text; the spectral windows are evident in the tide lines.

Uscinski (1986).) The tides were modeled as a sum of sine and cosine terms with independent coefficients. This stochastic inverse method is the frequency domain equivalent to the inverse published by Ewart (1986) where the correlation function was used. The results of the inverse can be used to remove the deterministic functions from the time series. The 2 kHz acoustic phase record in the time domain is shown in Figure 5, before and after removal of the trend and tide model. Although the determinism of the tides is assumed, the large baroclinic tide is almost certainly time dependent over the 15 days.

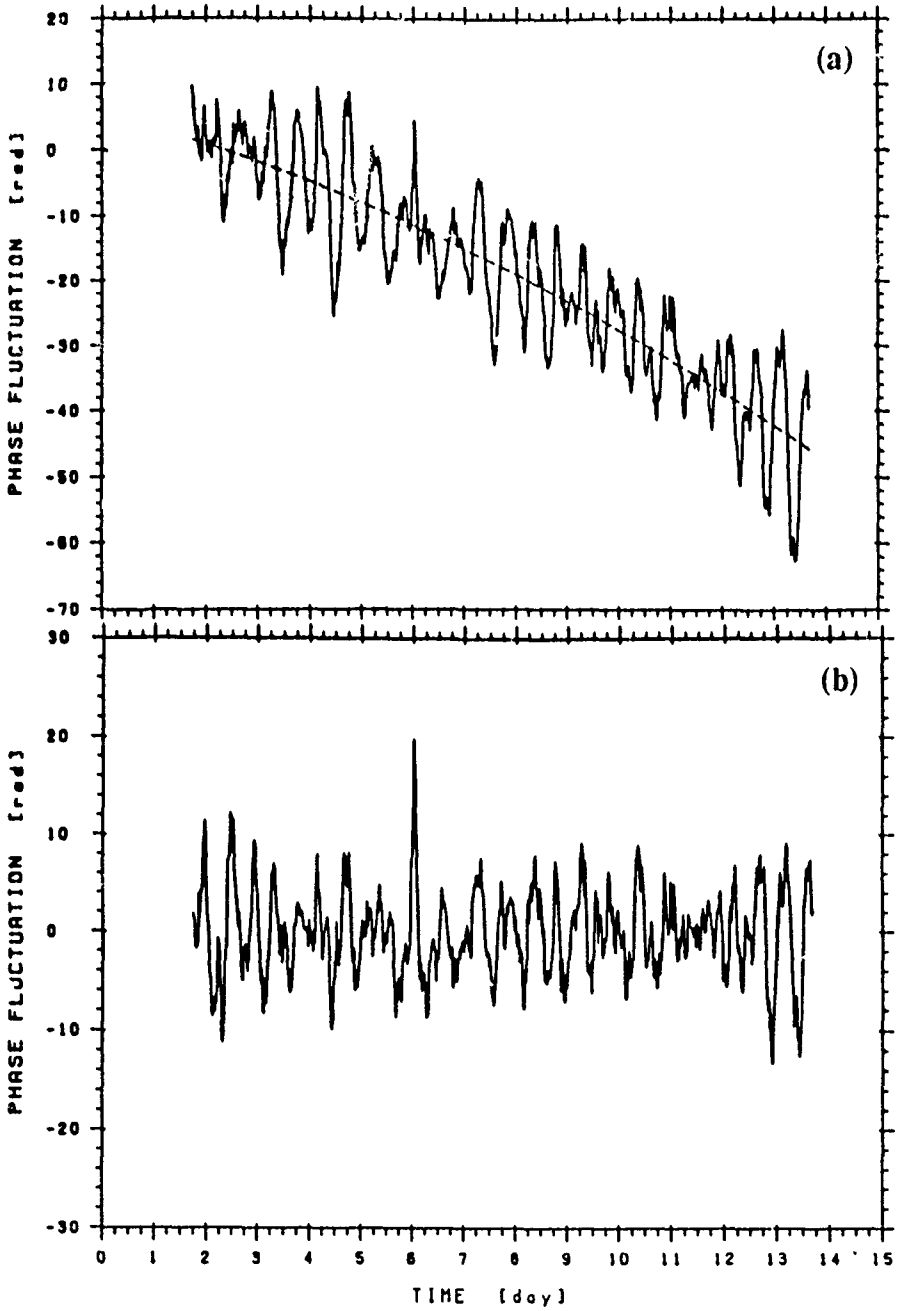


Figure 5. 2 kHz phase data in the time domain: (a) as recorded, and (b) with the deterministic trend and tidal components removed. The trend function is shown by the dashed line in (a). Note that the vertical scales are different.

To compare the travel time spectrum to the more familiar moored temperature spectrum, a similar inversion is applied to the temperature time series. The 30 day temperature record is windowed to the same 15 day period as the acoustic data and resampled to the same time grid. The modeled form of the temperature spectrum is a different integral of $S(\alpha,\beta,\omega)$, i.e., the moored spectrum:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\alpha,\beta,\omega)\, d\alpha\, d\beta\,. \tag{19}$$

To get the moored spectrum at the ray depth in $<\mu^2>$ units, we convert the temperature fluctuations to sound speed changes and use Eq. (2). Plots of both the travel time and temperature spectra in these units are illustrated in Figure 6. The plots include only the stochastic components, with the tidal components removed as described above.

Three distinct regions, separated by ω_i and ω_n are indicated. The modeled ω-dependence for each region is shown. Region I is represented by the low wavenumber/low frequency finestructure given by Eq. (18). We see that the moored and the travel time spectra are identical within statistical limits. These limits are large due to the short (15 day) time record. This supports

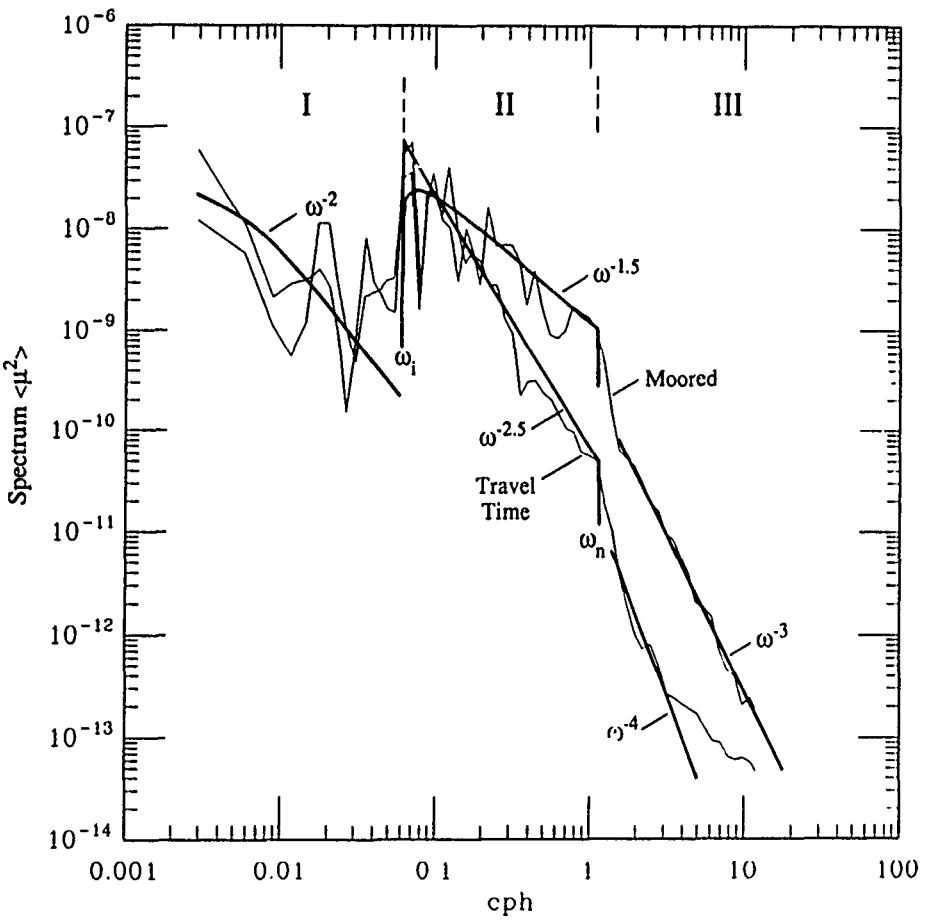


Figure 6. Plots of the travel time and moored temperature spectra in $<\mu^2>$ units. The tidal components have been removed; only the stochastic components remain in the data.

the model having no distinct dispersion relation (i.e. the moored temperature spectrum and the travel time spectra have the same spectral slope). The 500 hour value of τ_0 used in the model arises from the constraint of equal variance of finestructure and internal waves found in the β-domain (Levine and Irish, 1981). In region II, we see strong evidence that the internal wave model is correct. The differing spectral forms of $\omega^{-1.5}$ for the moored spectrum and $\omega^{-2.5}$ for the travel time spectrum support the effect of the internal wave dispersion relation (i.e. the travel time spectrum differs in spectral slope by -1 from the temperature spectrum). Also, the spectral cutoffs, the value of p, and the normalization of the model are supported. In region III, the ω^{-3} spectral slope of the moored spectrum and the ω^{-4} spectral slope of the travel time spectrum provides strong evidence that the high wavenumber finestructure is advected by internal waves (hence the effect of the internal wave dispersion relation).

We have attempted to show both the complexity of the ocean TCF as well as the large diversity of oceanographic data needed to confirm ocean spectral models. The ability of the acoustic field to give us an integral constraint on the model through phase correlations must be emphasized. It should also be emphasized that the inversion was done individually on each data set. A combined inversion is also possible (and is under proposed study). The combined inversion would impose, in a consistent manner, the relationships between the oceanographic and acoustic data sets exhibited in Figure 6. A future combined oceanographic/acoustic experiment would exploit these same relationships in a test of the new dynamical internal wave model.

AATE ACOUSTIC FIELD MEASUREMENTS

MATE demonstrated that the acoustic scattering conditions can be determined when extensive environmental measurements are made simultaneously with acoustic field measurements. Although MATE provided a long series of temporal acoustic measurements, only a few spatially separated receivers were used. To augment our understanding of the spatial characteristics of the scattered field, AATE (the AIWEX Acoustic Transmission Experiment) was designed to make both vertical and temporal measurements of the acoustic field. The transmission experiment (conducted under multi-year ice in the central Beaufort Sea) consisted of four co-located transmitters (2,4,8,16 kHz) suspended beneath the ice at 153 m depth. These were positioned 6.43 km from a depth cycling array of 3 receivers, separated by 51 m (the depth cycle was 51 m providing a 153 m vertical aperture). Simultaneous environmental measurements were made by several investigators from several institutions. A study of AIWEX moored data by Levine (1990) has resulted in an internal wave model.

Figure 7 displays the travel time and log-intensity spectra measured over two time periods during AATE. (A strong wind event occurred between the two time periods, making interpretation of the travel time measurements across the event difficult, if not impossible.) The spectra in Figure 7 may be compared to those taken during MATE. Predictions from the weak-scattering theory of Desaubies (1978) are shown with the observations from before the wind event. The acoustic fluctuations are significantly less energetic ($\approx$ 1/50th) than those expected under canonical GM, open ocean conditions (the GM-parameters have been adjusted for the AIWEX buoyancy frequency profile). In addition, the spectral slope of the travel time spectra is a power less than the prediction. These results are similar to those obtained from measurements made at the AIWEX environmental moorings (Levine, 1990). Note that the travel time spectra observed after the wind event display a peak at the local inertial frequency. This feature is also seen in the two

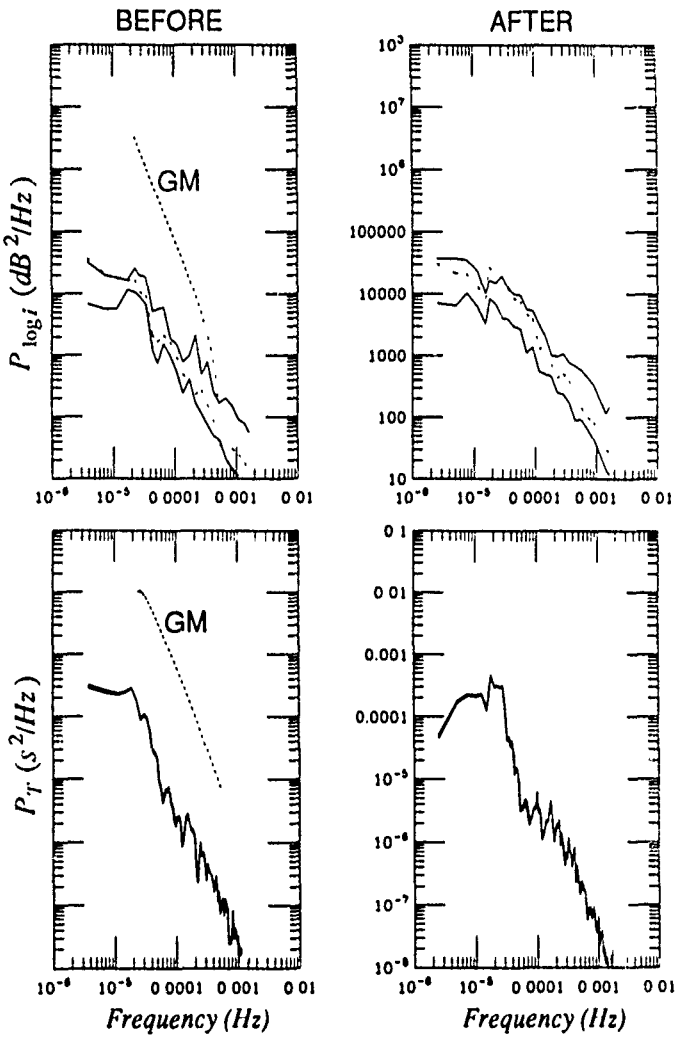


Figure 7. AATE observed log-intensity and travel time spectra for a 3 day time period before a wind event and a 4.3 day time period after the same event. To show the temporal behavior, a fixed depth data series was obtained. Spectra obtained from the 2, 4 (dashed) and 8 kHz data sets are shown. Predictions obtained using the canonical GM-model are overplotted on the left-hand graphs.

dimensional 2 kHz travel time spectra shown in Figure 8. The presence of the inertial peak after the wind event shows a serious lack of stationarity. This complicates modeling, but also indicates that additional interesting oceanographic processes were present after the wind event. Sorting out the mechanisms requires close examination of the environmental data taken during these two time periods.

A goal was to invert the AATE acoustic phase measurements and obtain a prediction for the internal wave/finestructure spectrum, analogous to what was done with the MATE results. Though the AATE travel time measurements were made to accuracies of a few microseconds out of a total travel time of 4 seconds, the travel times are contaminated by mooring motions of less than 1 cm. This contamination makes their use for an inversion problematic. However, when the

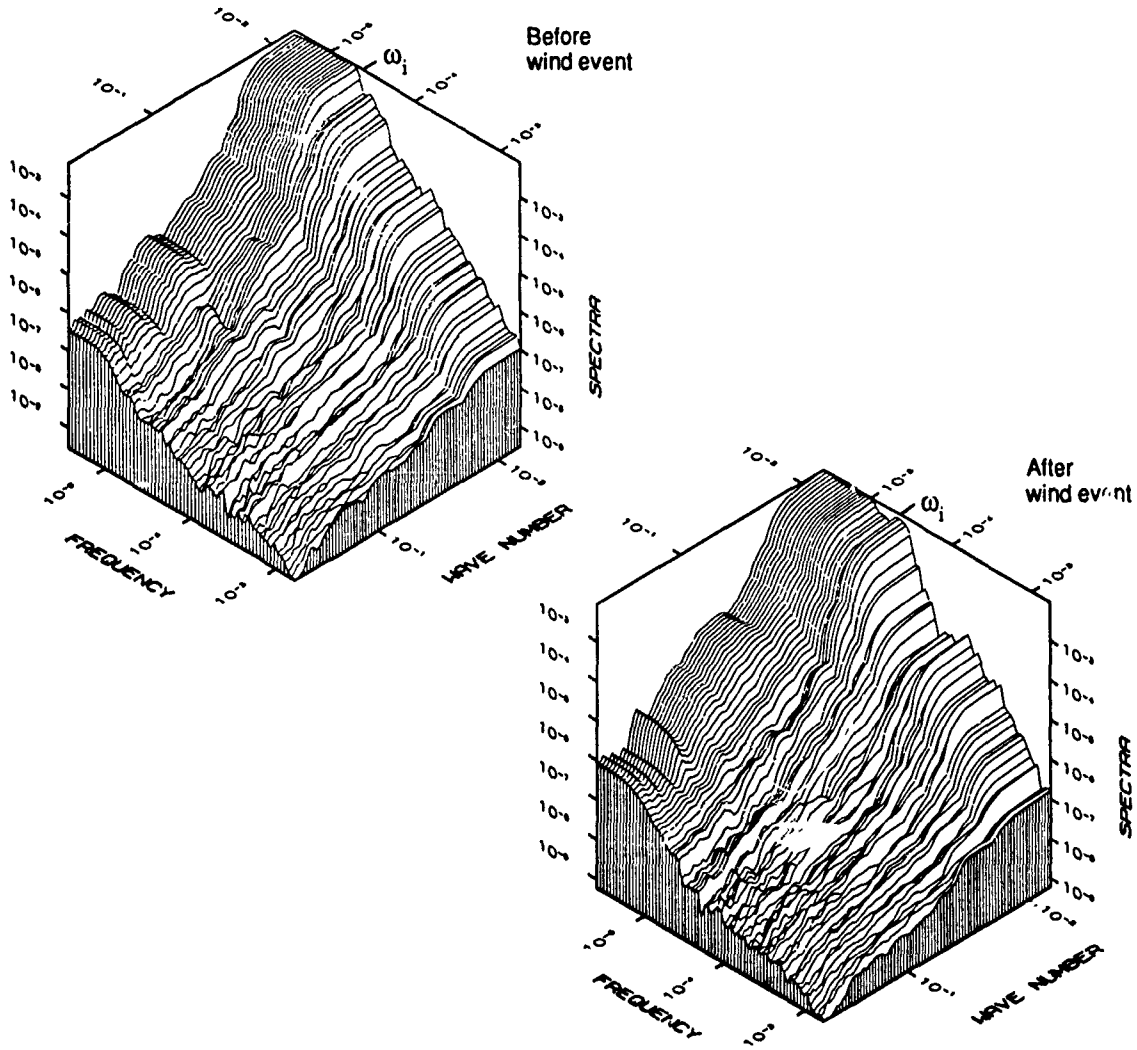


Figure 8. AATE travel time spectrum as a function of vertical wavenumber and frequency estimated from the 2 kHz measurements. A peak is observed near the inertial frequency in the spectrum from after the wind event.

scattering is sufficiently weak, stochastic inverse predictions are possible from the amplitude measurements. Under the conditions of AATE, we should be able to obtain the environmental field from inversion of the log-amplitude statistics. This analysis remains to be completed. The value of the spatial inversion is implied by the predictions shown in Figure 9. These are the wavenumber spectral filters of acoustic phase and log-amplitude obtained from the Rytov predictions of Desaubies (1978). The predicted phase and log-amplitude spectra are obtained by multiplying the medium spectrum, $S(\beta,\omega)$, by the appropriate filter function.

Note that for more traditional open ocean scattering conditions, the acoustic intensity probably cannot be used as an inversion tool. It is beyond the scope of this paper to discuss comparisons between theory and measured intensity statistics. But for propagation ranges of more than a few kilometers in most of the ocean, a multiple-scattering theory is required. When the field has been

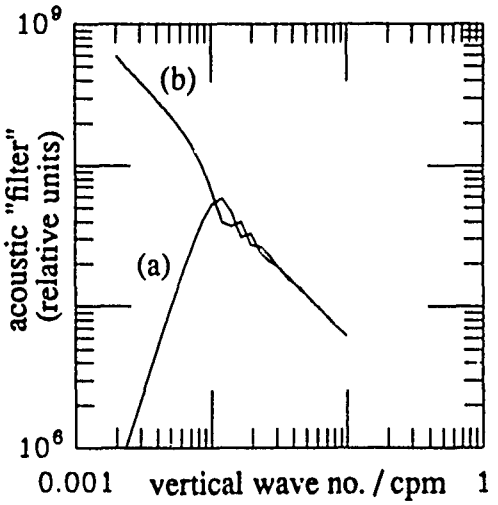


Figure 9. The acoustic filters as functions of vertical wavenumber for $\omega = .3$ cph that act on the medium spectrum for the log-amplitude (a) and phase (b). These functions were obtained using the weak-scattering theory of Desaubies (1978).

multiply scattered, the intensity statistics can no longer be understood as the action of a linear filter, like that shown in Figure 9. Fortunately though, the multiple scattering effect on the acoustic phase is small enough that it can be used as an inversion tool apparently even for long range cases. Although the AATE measurements provided measurements both in the vertical and in time, the weak acoustic phase fluctuations were masked by small mooring motions. Because virtually no other space/time measurements of the acoustic field have been made under conditions where the scattering field is known, further studies of the stochastic inverse methodology requires, at least for the near future, taking a numerical approach. Our current work presents some examples relevant to the oceanographic community and demonstrates the validity of this approach.

NUMERICAL SIMULATIONS AND THEORY

Much has been learned about acoustic fluctuations from numerical simulations. An example is its use in testing acoustic scattering theory. The technique of using parabolic equation propagation codes to test moment theoretical predictions was initiated by Macaskill and Ewart (1984). Since then, their technique has been modified to include a point source initial condition and other important physics. The moment theories are full range theories, and not asymptotic at short or long range; thus, they are important to ocean acoustics.

Figure 10 shows a contour plot of the normalized intensity variance (traditionally called the scintillation index) as a function of acoustic frequency and range (from Ewart, 1989). The input ocean model used for the TCF is $R_{\perp}(\zeta) = (1 + |\zeta|/L_v)\exp(-|\zeta|/L_v)$ which is asymptotically β^{-4} in the vertical wavenumber domain. The normalization is typical of mid-ocean internal waves with $<\mu^2> = 3.0 \cdot 10^{-9}$, $L_p = 4600$ m, and $L_v = 150$ m. For frequencies above 1 kHz, the scintillation index rises to a maximum at a location called the focus of the medium, and then decays to a value of one. The noise cutoff, plotted as a heavy solid line, demonstrates that for frequencies above 1 kHz, reaching long ranges will be difficult in the ocean.

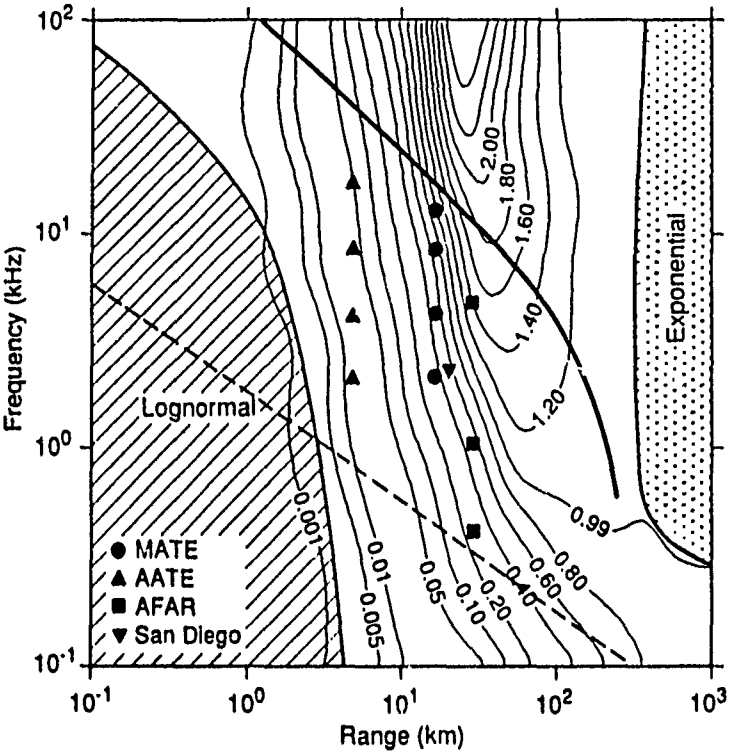


Figure 10. Contours of the normalized intensity variance vs frequency and range. Crosshatched areas are the asymptotic regions for the log-normal and exponential distributions. The solid line shows the range limit for a 200 dB source (for Sea State IV noise levels), and the dashed line is the boundary for multiple scattering.

For our purposes, it is important to note that Uscinski's full range theory for plane wave propagation is in excellent agreement with the plotted scintillation indices, and with the vertical wavenumber decomposition of the intensity variance. Similar studies of the point source initial condition display a scintillation peak that is higher. These features are also obtained with theory. Uscinski (1989) has demonstrated that when using parabolic equation propagation in polar coordinates (the natural coordinates for a point source) the predictions from moment theory and results of the simulations agree to within statistics. For that work a Gaussian TCF was used, and more work remains to be done for power law media such as the case with internal waves. But the overall agreement demonstrates that the moment theory solutions are robust over diverse scattering conditions and ranges. The underlying point is that, for a given medium transverse correlation function, the numerical experiments have shown that available theories of stochastic wave propagation can predict the second and fourth moments of the acoustic field. Our ability to predict acoustic fluctuation statistics under true ocean conditions then depends upon the validity of the TCF, and hence the importance to acoustics of the new internal wave modeling effort.

WHAT SCALES ARE IMPORTANT FOR ACOUSTICS?

Our last example is from numerical evaluation of theory. The example comes from asking the question, "If an oceanographer made vertical measurements of sound speed fluctuations, what

scales in those measurements are important in predicting acoustic volume scattering?". This question has been addressed in recent work by Ewart and Ballard (1990). They attempted to establish the small scale limit for oceanographic measurements required to predict the scintillation index, given a known TCF. Because ocean scattering is characterized by weak but multiple scattering, all scales of the medium can theoretically contribute to all scales in the prediction of the acoustic intensity. Using Uscinski's theory to predict the intensity fluctuations with depth, the following computation was carried out. Using a TCF derived from the internal wave model presented in Levine and Irish (1981), and a specific range and scattering strength, a prediction for the scintillation index was obtained for the condition of an inner scale wavenumber cut-off in the transverse spectrum of 0.1 cpm (10m). Denote this value of the scintillation index as SI_t. The cut-off wavenumber was then decreased until a scintillation index equal to .9 or 1.1 of SI_t was reached. The two regions correspond to before and after the medium focus. This process was then repeated for a wide range of scattering strengths and ranges. A contour plot of these 90% and 110% cut-off wavenumbers as functions of acoustic frequency and range is shown in Figure 11. (This parameterization requires selecting specific internal wave model parameters,

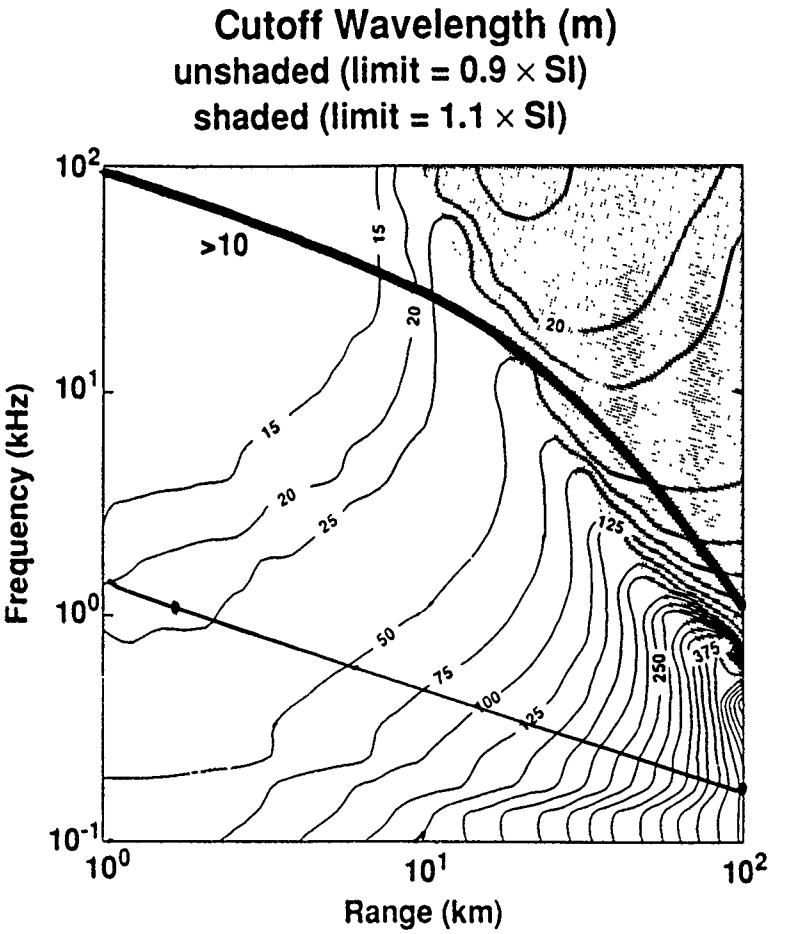


Figure 11. The vertical scale to reach 90 or 110 % of the scintillation index calculated to a limit of 10 m (see text) as a function of range and acoustic frequency. The contours are derived by selecting the MATE values of $\langle\mu^2\rangle^{1/2}$, L_p and L_v. As in Figure 10, the bold line corresponds to the noise limit for a 200 dBμPa/m source, the thin line denotes the multiple scattering boundary. In the region denoted ">10", the cutoff is at the Levine/Irish inner scale, and hence unchanging.

the ones listed above were used.) The bold line shown corresponds to the noise limit for a 200 dBμPa/m source; the light solid line denotes the multiple scattering boundary. The result is that for ranges greater than a few kilometers and frequencies less than 10 kHz, small vertical scales contribute little to the scintillation index. Note also that the smaller scales become less important as the range (and hence cumulative scattering) increases.

In the numerical simulations, the medium is treated as consisting of δ-correlated phase changing screens that are statistically described by the TCF. The question, "Is it theoretically permissible to represent the ocean statistics in this manner?" has been studied in additional research conducted by Ewart and Kaczkowski (1990). They have studied medium models that are correlated in range, versus those described by δ-correlated screens (i.e. a Markov process). Their research shows that little difference exists in the acoustic field moments for the two descriptions. This result, when considered with the observations presented in the previous paragraph on the effects of changing the inner scale, allows us to simplify oceanographic modeling for simulating acoustics. This simplification is incorporated into our discussion of dynamical modeling in the Summary section.

SUMMARY AND WHAT'S NEXT?

We have provided a brief tutorial on the connectivity between ocean stochastic modeling and predictions of the moments of an acoustic wave propagating through such an ocean. In this section we include a brief summary of our main points and mention desirable future directions for research.

Summary

- Internal wave models (new and used), are vital for predicting the statistics of acoustic propagation.
- Stochastic inverse methodology provides a tight check on proposed new models, when sufficiently careful space/time acoustic observations are made simultaneously with space/time oceanographic observations. The use of both oceanographic and acoustic measurements in combination must be emphasized.
- The missing element in research is the availability of acoustic observations in depth/time, where the medium correlation function is known.
- Numerical experiments, where dynamical models are used to define the index of refraction field, may help us to understand stochastic inverse methods, **BUT,**
- Field experiments with sufficient space/time measurement bandwidth must be coupled with acoustic measurements ----
    - We must measure the complex acoustic field $E(z,t)$ for several frequencies and ranges.
    - When a "new model" is available, a detailed experiment can be designed with numerical modeling.

3-D Dynamical Modeling

We have stated emphatically that the lack of high quality data prevents us from testing the stochastic inverse methodology on the range of space/time scales that are relevant to the

modeling goals of the 'Aha Huliko'a meeting. Using the insight we have gained from studying modeling limits, we are collaborating with Kraig Winters and Eric D'Asaro (See their contributions in this proceedings) to test the capabilities we have developed in theoretical predictions and stochastic inverse methodology with their dynamic simulation modeling. The proposed model, to be run on our Stardent Mini-Supercomputer, will have horizontal scales of 50 by 20 km in x and y respectively (x is the propagation direction) and 2000 m in the vertical. Its corresponding resolution will be 20 m in the vertical and 333 m in the horizontal with a grid of $128^3/4$ cells. The model will be initialized in several ways and run with time steps spanning many inertial periods. The resulting density fields will be the input to our PE propagation codes and will provide simulated realizations of the acoustic field. This type of modeling is very important, because it allows us to demonstrate the necessity (or lack thereof) of including dynamics in acoustic modeling.

Realizations of 3-D plus time density fields can also be produced from any proposed model, and used via Monte Carlo methods to test the stochastic inverse concepts discussed here. In verifying acoustic scattering predictions, one can test the robustness of the predictions to relaxed assumptions in the model, e.g., fully random phase versus dynamically consistent, and correlated versus Markov medium representations. All of these issues are relevant to our computer-limited ability to model acoustic propagation.

A Proposed Experiment

Eric D'Asaro has proposed an internal wave experiment designed to test many of the existing ideas on the cascade of energy from low internal wave modes to higher modes. The experiment would be conducted far from boundaries and sources of low mode internal wave energy. For example an area south of a storm region could be used to study how a changing flux of low mode internal wave energy "pumps" a local internal wave space/time spectrum. An extensive suite of dynamic and scalar oceanographic measurements would be made over a long period of time in order to develop an understanding of the linear and nonlinear processes involved. We would propose that an acoustics experiment capable of measuring the complex field at many spatially separated points be made an integral part of the overall measurement program. Many of the ideas we have presented here could be implemented. By sensing such a large volume of the experimental region, the acoustic measurements we envision would provide a severe constraint on possible space/time models.

ACKNOWLEDGMENTS

The authors would like to thank Dr. John Ballard for carrying out the evaluations of theory. Nina Triffleman edited our efforts and produced it in camera ready finished form. This research was sponsored by the Department of the Navy, Office of the Chief of Naval Research, under Grant N00014-90-J-1260. This article does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

Desaubies, Y.J.F., 1976, Analytical representation of internal wave spectra, *J. Phys. Ocean.*, **6**(6), 976-981.

Desaubies, Y.J.F., 1978, On the scattering of sound by internal waves in the ocean, *J. Acoust. Soc. Am.*, **64**(5), 1460-1469.

Ewart, T.E., 1986, Acoustic propagation, internal waves, and finestructure, *Proc. Instit. Acoust.*, **8** (Part 5), 106-122.

Ewart, T.E., 1989, A model of the intensity probability distribution for wave propagation in random media, *J. Acoust. Soc. Am.*, **86**(4), 1490-1498.

Ewart, T.E. and J. Ballard, 1990, Coupled volume and surface scattering, *J. Acoust. Soc. Am.*, Suppl. 1, **88**, S44.

Ewart, T.E. and P. Kaczkowski, 1990, Simulating stochastic scattering in ocean acoustics: How complex a model do we need ?, 1990 (unpublished manuscript).

Ewart, T.E., C. Macaskill and B.J. Uscinski, 1983, Intensity fluctuations. Part II: Comparison with the Cobb experiment, *J. Acoust. Soc. Am.*, **74**(5), 1484-1499.

Ewart, T.E. and S.A. Reynolds, 1984, The mid-ocean acoustic transmission experiment, MATE, *J. Acoust. Soc. Am.*, **75**(3), 785-802.

Levine, M.D., 1990, Internal waves under the Arctic ice pack during AIWEX: The coherence structure, *J. Geophys. Res.*, **95**(C5), 7347-7357.

Levine, M.D. and J.D. Irish, 1981, A statistical description of temperature finestructure in the presence of internal waves, *J. Phys. Ocean.*, **11**(5), 676-691.

Levine, M.D., J.D. Irish, T.E. Ewart and S.A. Reynolds, 1986, Simultaneous spatial and temporal measurements of the internal wave field during MATE, *J. Geophys. Res.*, **91**(C8),9709-9719.

Macaskill, C. and T.E. Ewart, 1984, Computer simulation of two-dimensional random wave propagation, *I.M.A. J. Appl. Math.*, **33**, 1-15.

Munk, W., 1981, Internal waves and small-scale processes, in *Evolution of Physical Oceanography*, edited by B. A. Warren and C. Wunsch, MIT Press, 264-291.

Slepian, D., 1978, Prolate-spheroidal wave functions, Fourier analysis and uncertainty - V: the discrete case, *Bell Syst. Tech. J.*, **75**(5), 1371-1430.

Uscinski, B.J., 1980, Parabolic moment equations and acoustic propagation through internal waves, *Proc. R. Soc. Lond.*, A(372), 117-148.

Uscinski, B.J., 1986, Acoustic scattering by ocean irregularities: Aspects of the inverse problem, *J. Acoust. Soc. Am.*, **79**(2), 347-356.

Uscinski, B.J., 1989, Numerical simulations and moments of the field from a point source in a random medium, *J. Mod. Optics*, **36**(12), 1631-1643.

Uscinski, B.J. and J.R. Potter, 1988, Sound ribbons in the Sea, *Acoustics Bulletin*, Oct.

INFERENCES OF GRAVITY WAVE PROCESSES FROM ATMOSPHERIC SPECTRA

David C. Fritts

Geophysical Institute, University of Alaska, Fairbanks, AK 99775-0800

ABSTRACT

Considerable information on the characteristics and variability of the atmospheric motion spectrum has emerged in the last few years. As a result, we are developing a much better understanding of the processes that maintain the spectral shape and account for spectral variability and its effects. This paper will summarize briefly some of the recent atmospheric spectral observations and their implications for gravity wave saturation processes, spectral shape, Doppler shifting, momentum fluxes, filtering and anisotropy. Also discussed are the apparent similarities and differences between the motion spectra in the atmosphere and the oceans.

INTRODUCTION

Gravity waves are now recognized by most researchers to account for the majority of the mesoscale motion field in the lower and middle atmosphere. Because of their ubiquity and their many effects, they have been the subject of considerable research efforts over the last few decades. Unlike the oceans, where the role of gravity wave motions in momentum and energy transports is at present uncertain, atmospheric gravity waves are now believed to be responsible for much of the vertical coupling of momentum and energy in the atmosphere.

Much of our knowledge of gravity wave motions in the atmosphere has come from remote sensing using ground-based radar or optical systems. Also available, however, has been a range of in situ data obtained with balloon, aircraft, space shuttle, rocket, satellite and other instrumentation. These data have yielded many insights into the structure and variability of the atmospheric motion field. Radar and lidar systems have provided data on the temporal nature and vertical structure of atmospheric motions at one location. High-resolution vertical profiles have also been obtained with balloon- and rocket-borne systems, while data on the horizontal variability and/or vertical structure of the motion, thermal, and constituent fields have been provided by balloon, aircraft, space shuttle, and satellite instrument systems. Together, these data provide an increasingly comprehensive picture of the processes controlling the atmospheric gravity wave spectrum and its various effects.

One intent of this paper is to summarize some of the more general implications of the diverse spectral observations of atmospheric motions that are now available. A second objective is to note some of the areas in which atmospheric and oceanic processes are likely to be similar and those where we should expect to see disparities. Characteristics of and inferences drawn from vertical wavenumber spectra at various levels in the atmosphere are presented in the following section. The principal result here is that several processes are acting to control the amplitude of the spectrum to a remarkable degree. We then examine the implications of frequency spectra of horizontal and vertical motions for wave propagation, Doppler shifting, and wave field anisotropy. Additional evidence of anisotropy, and of the extent to which gravity waves force the larger-scale atmospheric circulation, is provided by momentum flux and divergence measurements at a range of sites. Recent studies have also emphasized the scales at which these fluxes are preferentially contained. We conclude with a discussion of the similarities and differences of atmospheric and oceanic spectra and processes.

VERTICAL WAVENUMBER SPECTRA

Vertical wavenumber spectra of horizontal motions and density or temperature fluctuations obtained during the last few years increasingly have been interpreted as evidence of gravity wave saturation. Initially, it was argued by Dewan and Good (1986) and Smith et al. (1987) that linear instabilities should dictate a saturated spectrum at large vertical wavenumbers with horizontal velocity and fractional temperature spectra varying as ~ $N^2/6$ m^3 and $N^4/10$ g^2m^3. More recently, other saturation mechanisms have been investigated and explanations of the saturated spectrum offered based on nonlinear interactions and/or Doppler shifting (Dong and Yeh, 1988; Hines, 1988; Holloway, 1988; Dunkerton, 1989; Fritts, 1989; Fritts and Yuan, 1989a; Hines, 1991). In all cases, however, these theories have predicted amplitudes and slopes largely consistent with linear theory. This suggests that the implications of gravity wave saturation will be more dependent on spectral character than on specific saturation processes.

At lower wavenumbers, vertical wavenumber spectra are expected to have a positive slope to insure a finite vertical flux of wave action (VanZandt and Fritts, 1989). Consistent with the increase in energy density (per unit mass) with height and the amplitude limits at large wavenumbers, horizontal velocity spectra exhibit a dominant vertical wavenumber $m_* = 2\pi/\lambda_{z*}$ that decreases with increasing height from ~ 3 rad/km in the troposphere and lower stratosphere (Fritts and Chou, 1987; Fritts et al., 1988; Tsuda et al., 1989) to ~ 0.3 rad/km near the mesopause (Smith et al., 1987; Wu and Widdel, 1989). The increase in energy density with height is generally attributed to a preferential upward propagation of gravity wave energy and the associated decrease in atmospheric density. Similar slopes and amplitudes were noted in stratospheric temperature spectra (Fritts et al., 1988) obtained with high-resolution balloon soundings. Lidar data, in contrast, yield spectra with similar slopes, but with amplitudes that depart from those predicted by saturation theory in the middle stratosphere (Chanin and Hauchecorne, 1987; Kwon et al., 1990), due perhaps to the longer integration times inherent in

the lidar profiles and a lack of sensitivity to motions near inertial frequencies which contain the majority of the velocity and temperature variances.

A spectral form that appears to fit well the various observations is

$$E(m) = \frac{4E_0}{\pi m_*} \frac{m/m_*}{1+(m/m_*)^4}, \qquad (1)$$

where E_0 is the total spectral variance and m_* is the characteristic vertical wavenumber corresponding to the wavenumber of maximum spectral density in a variance content form. To be consistent with the saturated spectral amplitudes estimated by Smith et al. (1987), this implies $E_0 = \pi N^2/24\ m_*^2$ and $E_0 = \pi N^4/40\ g^2 m_*^2$ for velocity and fractional temperature fluctuation spectra, respectively.

A vertical wavenumber spectrum with the above characteristics has a number of implications for atmospheric circulation and structure. First, a form given by Eq. (1) with a decrease of m_* consistent with observations implies a growth of wave energy at small wavenumbers that is less than expected for conservative motions (Smith et al., 1987) and thus a continuous removal of energy from motions at all vertical scales. This implies, in turn, smooth variations of wave energy and momentum flux, and of their vertical divergences, with height and corresponding smooth variations in wave drag and induced diffusion. Because the saturated spectral amplitude depends on the buoyancy frequency, N, we also anticipate additional, or enhanced, saturation and effects near regions where N^2 increases with height. This implies increased wave dissipation and drag above the tropopause and the summer mesopause, in particular, that can be estimated based on the observed spectral character (VanZandt and Fritts, 1989). These effects are expected to be especially important at greater heights where the wave energy and momentum flux divergences have greater influences.

Temporal variations of the vertical wavenumber spectrum provide evidence of isolated components of the motion field and thus may suggest the processes primarily responsible for wave saturation in the atmosphere. Shown in Figure 1 are series of vertical wavenumber spectra in variance content form in the troposphere and lower stratosphere at 1-h intervals obtained with the MU radar in Japan (Fritts et al., 1988). The spectra reveal good consistency with height and significant local departures from the canonical form of the saturated vertical wavenumber spectrum. Of particular interest, however, are the persistent features both near the maxima and at higher vertical wavenumbers that suggest long-lived components of the motion spectrum propagating and dissipating in superposition, but without interaction, with other spectral components. This seems to imply a dissipation process that relies on local wave field instabilities due to superposed wave amplitudes rather than strong wave-wave interactions and spectral energy transfers.

Additional evidence of wave instability and turbulence generation is provided by measurements of spectral characteristics at smaller scales. One example of a vertical wavenumber spectrum of

neutral density fluctuations inferred with a rocket-borne positive ion probe (PIP) during the MAC/EPSILON experiment is shown in Figure 2 and exhibits three distinct spectral ranges (Blix et al., 1990). At scales larger than ~ 300 m the spectrum has a slope near -3, consistent with the large wavenumber range expected for saturated gravity wave motions. From scales of ~ 5 - 300 m, however, the spectrum has a slope near -5/3, suggesting inertial range turbulence and wave energy dissipation. Finally, at scales less than ~ 5 m, the spectrum appears to be consistent with a

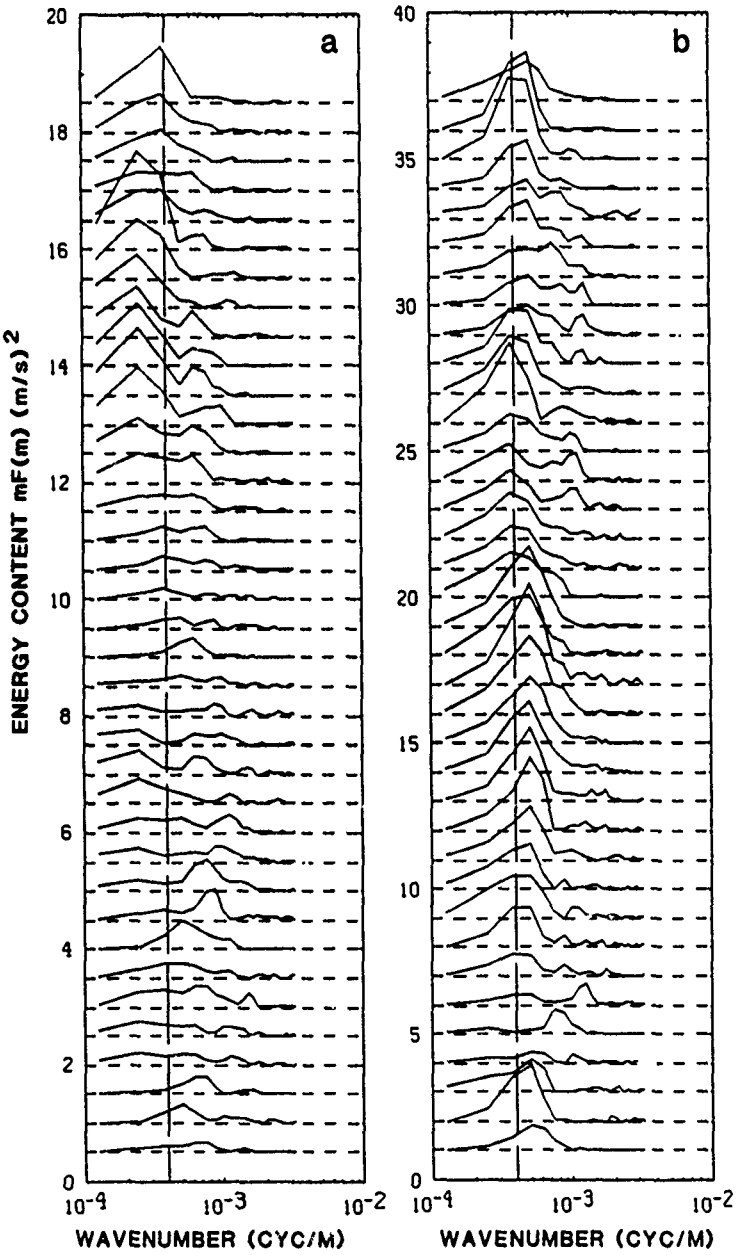


Figure 1. Hourly area-preserving spectra of northward radial velocity for (a) 5 - 13 and (b) 13 - 20.5 km heights with time increasing upward (Fritts et al., 1988). Note the consistency and persistence of spectral features in time and the departures from canonical form.

viscous subrange. Spectra obtained on other rocket flights provide similar views of small-scale dynamics and imply that these spectral features are persistent components of the motion field. These spectra imply a nearly continuous transfer of energy from gravity wave motions through inertial range turbulence to viscous scales. Estimates of the energy dissipation rates may be obtained from the transition scales and from the spectral amplitude within the inertial range, leading to estimates of ~ 0.01 W/m^2 in the middle and upper mesosphere. Estimates using radar techniques suggest comparable values, with enhancements near the summer mesopause that are consistent with predictions of saturation theory.

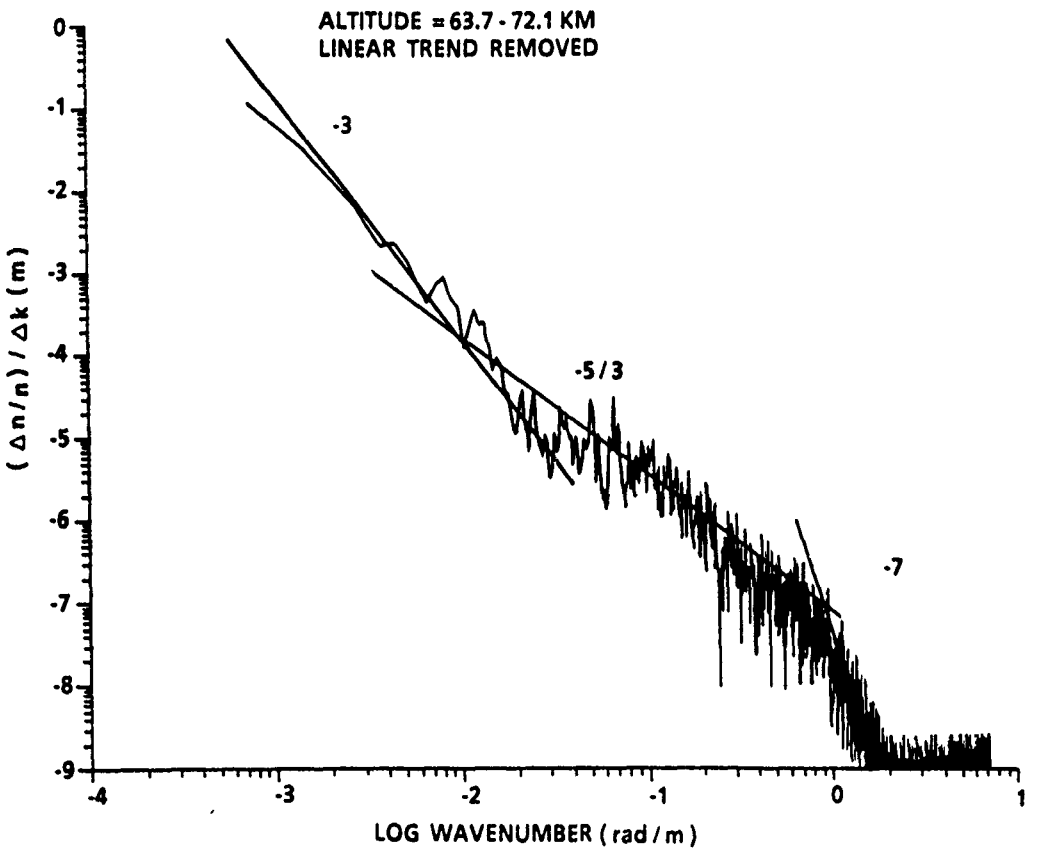


Figure 2. Spectrum of density fluctuations obtained with a positive ion probe showing three spectra ranges between scales of ~ 3 m and 8 km (Blix et al., 1990). These suggest ranges consistent with gravity wave saturation, inertial range turbulence, and viscous dissipation.

FREQUENCY SPECTRA OF HORIZONTAL AND VERTICAL MOTIONS

Measurements of frequency spectra using various techniques have revealed the general characteristics and variability of the motion field under a variety of conditions. Frequency spectra obtained using atmospheric radars have provided the greatest diversity of observations and have shown the mean frequency spectrum of horizontal motions to exhibit an amplitude growth with

increasing height and clear tidal peaks at upper levels. At frequencies greater than tidal frequencies, most observations suggest a slope near -5/3 (Balsley and Carter, 1982; Vincent, 1984) that is nearly invariant with height and mean atmospheric motions. Observations of frequency spectra of vertical motions are more varied, with slopes that are slightly positive under weak wind conditions and which become increasingly negative as mean winds increase. Examples of vertical velocity spectra under weak and strong wind conditions are shown for reference in Figure 3.

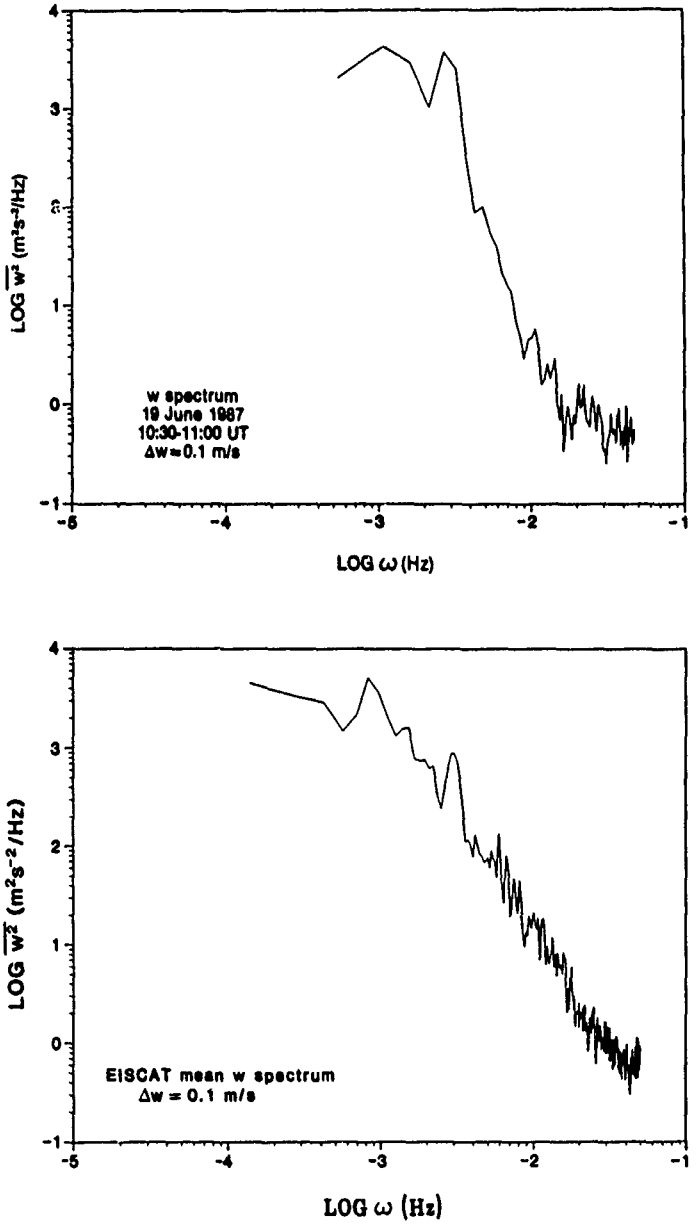


Figure 3. Vertical velocity frequency spectra during the MAC/SINE campaign (Fritts et al., 1990a). Upper plot is for light winds and exhibits a peak near N. Lower plot is a mean spectrum.

Assuming that the frequency and wavenumber spectra of gravity wave motions are separable, for which there is some evidence, and an intrinsic frequency spectrum (relative to the medium) of horizontal motions of the form ω^{-p}, with $p \sim 5/3$, we expect the intrinsic frequency spectrum of vertical velocities to vary as ω^{2-p}. These expectations are in agreement with observations under light wind conditions, but depart significantly when mean horizontal motions are large.

Departures of the observed spectra from the anticipated forms of intrinsic frequency spectra may arise for several reasons. As noted in the previous section, vertical wavenumber spectra exhibit substantial departures from the mean spectral form and suggest that similar variability might be expected to occur among frequency spectra as well. But this explanation is unable to account for departures of mean spectra from a relationship consistent with the gravity wave dispersion relation. Instead, these departures appear to be due to the differing effects of Doppler shifting on the frequency spectra of horizontal and vertical motions (Fritts and VanZandt, 1987).

The importance of Doppler shifting depends on 1) the relative motion of the intrinsic and observed frames, 2) the horizontal wavenumbers and phase speeds of the motions containing the majority of the spectral variance, and 3) the direction of wave propagation relative to the mean motion. If the mean motion is small compared to characteristic phase speeds or intrinsic frequencies are small, as is generally the case for those gravity waves accounting for most of the horizontal velocity variance, then Doppler shifting effects are relatively minor and wave propagation directions are unimportant. If, however, the mean motion is large compared to phase speeds containing the velocity variance or intrinsic frequencies are high, then there is a large potential for Doppler shifting of velocity variance throughout the spectrum and wave propagation directions are very important.

More simply stated, it is easier to Doppler shift velocity variance from high to low frequencies than from low to high frequencies (a mean flow comparable to the intrinsic phase speeds can Doppler shift high-frequency variance to $\omega \sim 0$, but not vice versa). This implies much greater differences in the intrinsic and observed frequency spectra for vertical motions (primarily at high frequencies) than for horizontal motions (primarily at low frequencies) and suggests that vertical velocity spectra will be more sensitive to such effects as a result. The effects of Doppler shifting on two-dimensional, symmetric, idealized, intrinsic horizontal and vertical velocity spectra (with $p = 2$) are illustrated in Figure 4 for various values of $\beta = u/c_*$, where u is the relative motion, $c_* = N/m_*$, and m_* is as defined with Eq. (1). These results demonstrate that the slopes of observed symmetric horizontal velocity spectra remain nearly constant at high frequencies, even for large β, while observed vertical velocity spectra may exhibit significant changes in slope and amplitude at small β.

The significant changes anticipated theoretically in observed vertical velocity spectra under different Doppler-shifting conditions provide a convenient test of these simple spectra effects. Such tests have been performed by Vincent and Eckermann (1990), VanZandt et al. (1990), and Fritts and Wang (1991) and reveal a remarkable consistency with the theory, suggesting that

Doppler-shifting effects do indeed account for much of the observed spectral variability. Vincent and Eckermann (1990) noted that Doppler-shifting appeared to explain the increase in wave variance parallel to the local flow associated with frontal circulations and VanZandt et al. (1990) found good agreement between model predictions and observed, symmetric spectra in the lower atmosphere. At greater heights, Fritts and Wang (1991) found observed vertical velocity spectra to provide evidence of Doppler shifting, wave field anisotropy, and the form of the vertical wavenumber spectrum. These results suggest that observed motions are largely consistent with gravity wave theory and differ from the predictions of geostrophic or stratified turbulence, which has been offered as an alternative explanation of the motion field (Gage, 1979; Lilly, 1983).

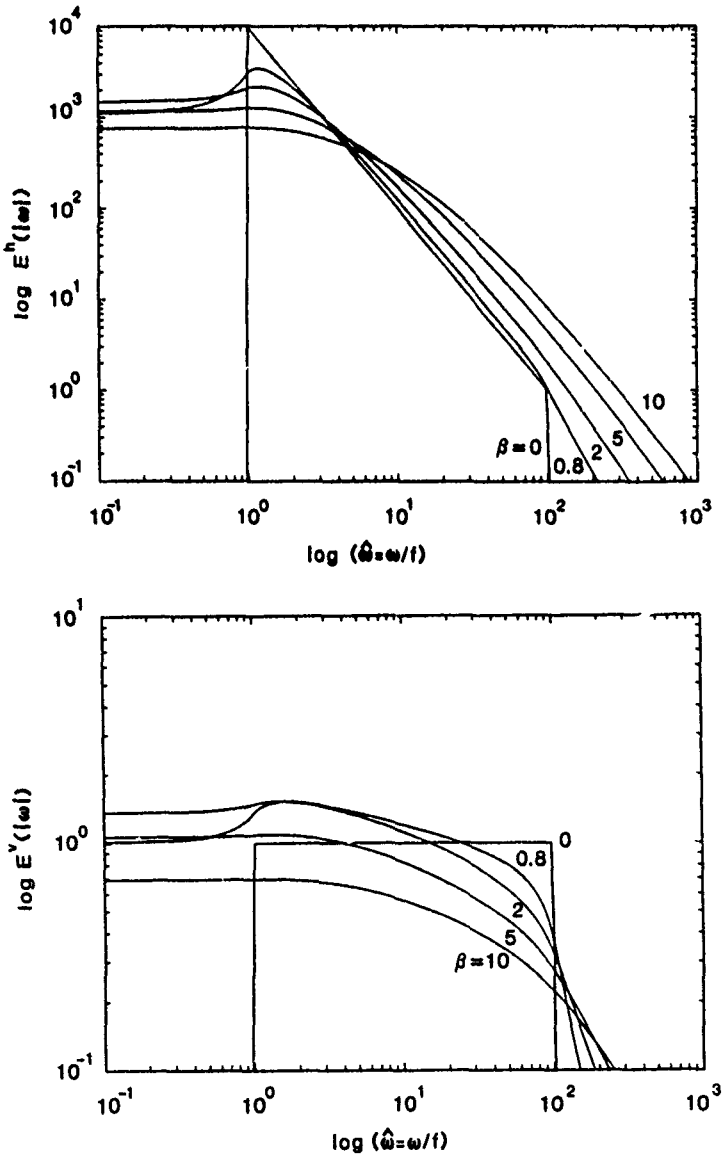


Figure 4. Effects of Doppler shifting on two-dimensional isotropic horizontal and vertical velocity spectra (Fritts and VanZandt, 1987). See text for parameters.

The observed, and inferred intrinsic, frequency spectra also have implications for the dominant transport processes within the gravity wave field. As noted by Fritts (1984), the forms of the observed frequency spectra of horizontal and vertical velocities imply that wave energy and momentum fluxes are accomplished primarily by motions with high intrinsic frequencies. While horizontal velocity variance (and thus total wave energy) resides mainly at low intrinsic frequencies, $\omega \sim f$, momentum and energy fluxes depend on the speed of vertical propagation, causing the major fluxes to be associated with motions with intrinsic periods < 1 h. These motions were estimated to provide ~ 70 % of gravity wave transports by Fritts (1984). Subsequent observations at a number of sites have served to confirm these estimates at lower and upper levels (Fritts and Vincent, 1987; Reid and Vincent, 1987; Reid et al., 1988; Fritts and Yuan, 1989b; Fritts et al., 1990b; Wang and Fritts, 1990). More will be said about the spectral character of gravity wave momentum fluxes in the following section.

MOMENTUM FLUX SPECTRA

Relative to velocity spectra, there are very few observational data at present on the spectral character of momentum fluxes. Those that are available, nevertheless, serve to confirm our expectations based on velocity spectra, integrated momentum flux measurements, and theory.

One example of the observed frequency spectrum of vertically averaged westward momentum flux computed from the difference of averaged radial velocity spectra is shown in Figure 5 in standard and flux content form (Fritts et al., 1990b). This spectrum suggests a dominance of the

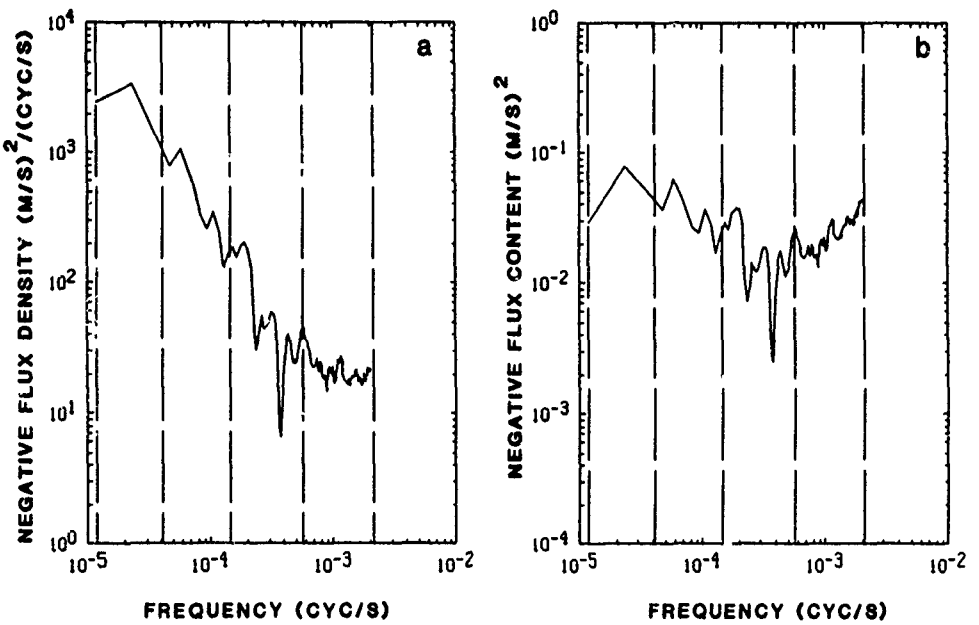


Figure 5. Frequency spectrum of vertically averaged westward momentum flux obtained with the MU radar (Fritts et al., 1990b) in (a) standard and (b) flux content form.

total momentum flux by low-frequency motions. As noted by Fritts et al. (1990b), however, the disparate radial velocity variances imply high intrinsic frequencies despite the small observed frequencies and thus wave motions and momentum fluxes largely consistent with orographic excitation. This view is supported as well by the temporal variability of the vertically averaged westward momentum flux observed in that data set.

It is not possible to obtain an accurate estimate of the intrinsic frequencies accounting for a majority of momentum fluxes at one site without knowledge of the horizontal scales and/or the relative anisotropy of the motion field. Such an estimate is possible, however, with horizontal measurements of wave activity and covariances in circumstances in which wave phase speeds may be reasonably approximated. This occurs when we expect the predominant wave motions to be due to orography, as in the observations of variance enhancements associated with orography by Nastrom and Fritts (1991). The momentum flux spectra (or cospectra of u and w) obtained from four aircraft flights during the the Global Atmospheric Sampling Program (GASP) are shown in standard and variance content form in Figure 6. These reveal, consistent with the inferences drawn from the MU radar data, that the majority of the momentum flux is due to motions with small horizontal scales and relatively high intrinsic frequencies. With the major contributions at horizontal scales of ~ 20 - 60 km and mean flows of ~ 20 m/s, we infer characteristic intrinsic periods of < 1 h. These results are also in agreement with momentum flux observations at greater heights for which orographic influences are likely to be less pronounced (Fritts and Vincent, 1987; Reid and Vincent, 1987; Fritts and Yuan, 1989b; Reid et al., 1988; Wang and Fritts, 1990).

The occurrence of the primary gravity wave momentum fluxes at small horizontal scales and high intrinsic frequencies has some important implications for the forcing of the large-scale

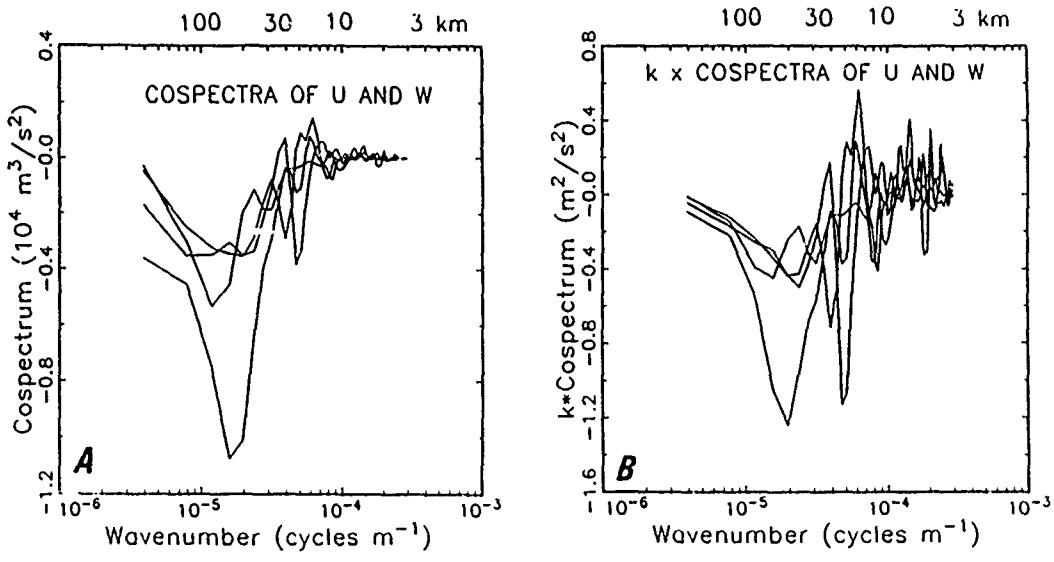


Figure 6. Momentum flux versus wavenumber in (a) standard and (b) flux content form (Nastrom and Fritts, 1991). Integrated values are negative with the major contributions at < 100 km scales.

atmospheric circulation and structure. Because such gravity waves have relatively steep propagation paths and the sources of such motions are often episodic and highly localized (Fritts and Nastrom, 1991), we might expect that the forcing of the mean state accompanying wave dissipation will also have similar attributes. This suggests that the response of the middle atmosphere to gravity wave excitation at and propagation up from lower levels will reflect the source and filtering conditions in place at that site, with horizontal coupling over large distances playing a much smaller role. These effects are particularly important in assessing the role of gravity wave forcing of the middle atmosphere because of the large mean fluxes that have been observed and the significant influences on the large-scale circulation and structure that are implied.

COMPARISON OF ATMOSPHERIC AND OCEANIC SPECTRAL IMPLICATIONS

Gravity wave spectra in the atmosphere and oceans share many features due to common source, propagation, and dissipation processes. There are also, however, some important differences between atmospheric and oceanic spectra that arise due to quantitatively different source distributions, wave-wave and boundary interactions, and characteristics of the mean wave environment.

Tidal and inertial-scale forcing of the oceanic internal wave spectrum by surface winds and currents over orographic features acts to energize near-inertial frequencies and results in a pronounced peak in wave energy near $\omega \sim f$. These motions persist for long times, may propagate large distances, and interact with other components of the spectrum on (arguably) slow time scales, resulting in an oceanic spectrum in the frequency and wavenumbers domains that is nearly universal in shape and amplitude. Because dissipation processes are slow away from boundaries and the wave spectrum is confined between two largely reflecting surfaces, there is also a tendency for the spectrum to be nearly vertically symmetrized and for net vertical fluxes of energy and momentum to be small. Under these conditions, the motion field can be conveniently expressed using a modal description.

Atmospheric gravity waves, in contrast, are excited by a wide range of source processes that appear to contribute wave energy predominantly at intermediate and high intrinsic frequencies. These include convective and frontal processes, orography, and wind shear and lead to atmospheric frequency spectra that do not include an inertial spike characteristic of the oceanic spectra. This greater fraction of total wave energy at higher frequencies in the atmosphere implies several important differences in atmospheric and oceanic spectra.

Wave motions at higher intrinsic frequencies propagate vertically more rapidly than low-frequency motions at comparable vertical scales. This results in larger vertical fluxes of energy and momentum and greater flux divergences (and associated energy dissipation and body forces) in regions of wave dissipation than would accompany lower-frequency motions. The effects of

These transports are enhanced in the atmosphere at heights above the primary wave sources due to the exponential decay of density with height, which causes wave amplitudes to increase approximately exponentially with height and energy and momentum fluxes and their divergences to increase relative to background levels. Energy dissipation rates (measured relative to local energy densities) in the stratosphere and mesosphere are consequently much larger than in the ocean and imply dissipation time scales for gravity waves of ~ 10 - 100 times less than in the oceans (~ hours - 10's of hours rather than ~ 10's of days). This likewise suggests differences in the processes acting to remove energy from the wave spectrum, requiring fast, energetic dissipation at increasing heights in the atmosphere (primarily local convective instability) and permitting slower and/or more systematic energy transfers (primarily dynamical instability or wave-wave interactions) at lower levels of the atmosphere and in the oceans. Despite the different physical processes suggested to be responsible for wave dissipation in the oceans and at various heights in the atmosphere, all those processes considered viable lead to comparable limits on wave amplitudes and corresponding saturated wavenumber spectra. Similar arguments imply that gravity wave momentum fluxes play a more significant role in forcing the large-scale circulation and structure in the middle and upper atmosphere than in the oceans.

Another difference between atmospheric and oceanic spectra that is due in part to the distribution of wave energy with frequency and in part to differences in the wave environment is the relatively greater influence of Doppler shifting in observed spectra of atmospheric motions. As noted previously, motions with intrinsic frequencies that are high are easily Doppler shifted to observed frequencies that are low, while low-frequency motions cannot be shifted to observed frequencies that are high. Additionally, typical mean motions of the atmosphere are larger than characteristic wave phase speeds, particularly at lower levels, whereas oceanic wave phase speeds are generally greater than mean currents. This results in a much greater range of observed frequency spectra, especially of vertical velocities because of their greater sensitivity (their greater energy concentration at $\omega \sim N$), in the atmosphere than in the oceans. The larger characteristic anisotropy of the atmospheric gravity wave field (due to discrete, high-frequency sources and greater departures from vertical symmetry) also contributes to variations of frequency spectra of horizontal and vertical velocities and permits inferences of wave sources, filtering, and effects that appear to be more challenging in the oceans.

In summary, the physical processes acting to excite, mould, and dissipate the gravity wave spectra in the atmosphere and oceans have some similarities and some important differences. The similarities account, broadly, for the qualitative agreement in mean observed frequency and wavenumber spectra in the two systems (spectral slopes and apparent saturation limits). But there are large differences in the potential for energy and momentum transports, dissipation processes, anisotropy, and influences on the mean state due to the unbounded and exponentially decreasing mean density in the atmosphere, the vertically contained and more vertically uniform mean structure of the oceans, and the very different intrinsic frequencies at which primary wave excitation occurs in the atmosphere and oceans.

Acknowledgments This research was supported by the SDIO/IST and managed by the Office of Naval Research under contract N00014-90-J-1271 and by the National Aeronautics and Space Administration under contract NAS8-37591.

REFERENCES

Balsley, B.B., and D.A. Carter, 1982: The spectrum of atmospheric velocity fluctuations at 8 and 86 km, Geophys. Res. Lett., 9, 465-468.

Blix, T.A., E.V. Thrane, D.C. Fritts, U. von Zahn, F.-J. Luebken, W. Hillert, S.P. Blood, J.D. Mitchell, G.A. Kokin, and S.V. Pakhomov, 1990: Small scale structure observed in situ during MAC/EPSILON, J. Atmos. Terres. Phys., 52, 835-854.

Chanin, M.-L., and A. Hauchecorne, 1987: Lidar sounding of the structure and dynamics of the middle atmosphere. A review of recent results relevant to transport processes, Transport Processes in the Middle Atmsophere, NATO Advanced Studies Workshop, G. Visconti and R. Garcia, eds., 459-477.

Dewan, E.M., and R.E. Good, 1986: Saturation and the "universal" spectrum for vertical profiles of horizontal scalar winds in the atmosphere, J. Geophys. Res., 91, 2742-2748.

Dong, B, and K.C. Yeh, 1988: Resonant and nonresonant wave-wave interactions in an isothermal atmosphere, J. Geophys. Res., 93, 3729-3744.

Dunkerton, T.J., 1989: Theory of internal gravity wave saturation, Pure Apl. Geophys., in press.

Fritts, D.C., 1984: Gravity wave saturation in the middle atmosphere: A review of theory and observations, Rev. Geophys. Space Phys., 22, 275-308.

Fritts, D.C., 1989: A review of gravity wave saturation processes, effects, and variability in the middle atmosphere, Pure Appl. Geophys., 130, 343-371.

Fritts, D.C., and H.-G. Chou, 1987: An investigation of the vertical wavenumber and frequency spectra of gravity wave motions in the lower stratosphere, J. Atmos. Sci., 44, 3610-3624.

Fritts, D.C., and G.D. Nastrom, 1991: Sources of mesoscale variability of gravity waves, II: Frontal, convective, and jet stream excitation, J. Atmos. Sci., in press.

Fritts, D.C., U.-P. Hoppe, and B. Inhester, 1990a: A study of the vertical motion field near the high-latitude summer mesopause during MAC/SINE, J. Atmos. Terres. Phys., 52, 927-938.

Fritts, D.C., T. Tsuda, T. Sato, S. Fukao, and S. Kato, 1988: Observational evidence of a saturated gravity wave spectrum in the troposphere and lower stratosphere, J. Atmos. Sci., 45, 1741-1759.

Fritts, D.C., T. Tsuda, T.E. VanZandt, S.A. Smith, T. Sato, S. Fukao, and S. Kato, 1990b: Studies of velocity fluctuations in the lower atmosphere using the MU radar: II. Momentum fluxes and energy densities, J. Atmos. Sci., 47, 51-66.

Fritts, D.C., and T.E. VanZandt, 1987: The effects of Doppler shifting on the frequency spectra of atmospheric gravity waves, J. Geophys. Res., 92, 9723-9732.

Fritts, D.C., and R.A. Vincent, 1987: Mesospheric momentum flux studies at Adelaide, Australia: Observations and a gravity wave/tidal interaction model, J. Atmos. Sci., 44, 605-619.

Fritts, D.C., and D.-Y. Wang, 1991: Doppler-shifting effects on frequency spectra of gravity waves observed near the summer mesopause at high latitude, submitted to J. Atmos. Sci.

Fritts, D.C., and L. Yuan, 1989a: A stability analysis of inertio-gravity wave structure in the middle atmosphere, J. Atmos. Sci., 46, 1738-1745.

Fritts, D.C., and L. Yuan, 1989b: Measurement of momentum fluxes near the summer mesopause at Poker Flat, Alaska, J. Atmos. Sci., 46, 2569-2579.

Gage, K.S., 1979: Evidence for $k^{-5/3}$ power law inertial range in mesoscale two-dimensional turbulence, J. Atmos. Sci., 36, 1950-1954.

Hines, C.O., 1988: The generation of turbulence by atmospheric gravity waves, J. Atmos. Sci., 45, 1269-1278.

Hines, C.O., 1991: The saturation of gravity waves in the middle atmosphere. Part I: Critique of linear-instability theory, J. Atmos. Sci., in press.

Holloway, G., 1988: The buoyancy flux from internal gravity wave breaking, Dyn. Atmos. Oceans, 12, 107-125.

Kwon, K.H., D.C. Senft, and C.S. Gardner, 1990: Airborne sodium lidar observations of horizontal and vertical wavenumber spectra of mesopause density and wind perturbations, J. Geophys. Res., 95, 13723-13736.

Lilly, D.K., 1983: Stratified turbulence and the mesoscale variability of the atmosphere, J. Atmos. Sci., 40, 749-761.

Nastrom, G.D., and D.C. Fritts, 1991: Sources of mesoscale variability of gravity waves, I: Topographic excitation, J. Atmos. Sci., in press.

Reid, I.M., R. Ruster, P. Czechowsky, and G. Schmidt, 1988: VHF radar measurements of momentum flux in the summer polar mesosphere over Andenes (69°N, 16°E), Norway, Geophys. Res. Lett., 15, 1263-1266.

Reid, I.M., and R.A. Vincent, 1987: Measurements of mesospheric gravity wave momentum fluxes and mean flow accelerations at Adelaide, Australia, J. Atmos. Terres. Phys., 49, 443-460.

Smith, S.A., D.C. Fritts, and T.E. VanZandt, 1987: Evidence for a saturated spectrum of atmospheric gravity waves, J. Atmos. Sci., 44, 1404-1410.

Tsuda, T., T. Inoue, D.C. Fritts, T.E. VanZandt, S. Kato, T. Sato, and S. Fukao, 1989: MST radar observations of a saturated gravity wave spectrum, J. Atmos. Sci., 46, 2440-2447.

VanZandt, T.E., and D.C. Fritts, 1989: A theory of enhanced saturation of the gravity wave spectrum due to increases in atmospheric stability, Pure Appl. Geophys., 130, 399-420.

VanZandt, T.E., S.A. Smith, T. Tsuda, D.C. Fritts, T. Sato, S. Fukao, and S. Kato, 1990: Studies of velocity fluctuations in the lower atmosphere using the MU radar: I. Azimuthal anisotropy, J. Atmos. Sci., 47, 39-50.

Vincent, R.A., 1984: Gravity wave motions in the mesosphere, J. Atmos. Terres. Phys., 46, 119-128.

Vincent, R.A., and S.D. Eckermann, 1990: VHF radar observations of mesoscale motions in the troposphere: Evidence of gravity wave Doppler shifting, Radio Sci., in press.

Wang, D.-Y., and D.C. Fritts, 1990: Mesospheric momentum fluxes observed by the MST radar at Poker Flat, Alaska, J. Atmos. Sci., 47, 1512-1521.

Wu, Y.-F., and H.-U. Widdel, 1989: Observational evidence of a satuated gravity wave spectrum in the mesosphere, J. Atmos. Terres. Phys., 51, 991-996.

PARADIGM LOST?

Chris Garrett

Centre for Earth and Ocean Research, University of Victoria,
Victoria, B.C. V8W 3P6, Canada

ABSTRACT

It appeared some years ago that a number of theoretical and observed features of the internal wave field can be reconciled with a diapycnal diffusivity of no more than about 10^{-5} m^2s^{-1}. This "paradigm" is re-examined here, with particular attention to the arguments concerning the alleged universality of internal waves in space and time. A key question concerns the influence of nonlinear interactions on horizontal spreading, but the simple picture still seems reasonable for the main thermocline. However, there are clearly significant departures from it in various regions, such as the upper ocean or close to the sea floor. The influence of a sloping bottom on both diapycnal fluxes and the velocity boundary condition for low frequency flows is emphasized; some general theoretical results are presente ut there is a need for further observations. Finally, a simple model for the mixing by, and decay of, an interfacial solitary wave is summarised.

1. INTRODUCTION

In January 1984 I had the pleasure of attending the Third 'Aha Huliko'a Hawaiian Winter Workshop on "Internal Gravity Waves and Small-Scale Turbulence". Figure 1 shows my attempt then to present a "zeroth order view" of the effect of internal waves on the ocean interior, summarising and perhaps oversimplifying the theoretical and observational progress made by many people. Seven years later it seems worthwhile to ask whether this paradigm should be abandoned or merely qualified.

Section 2 thus briefly reviews and extends the arguments leading to Figure 1, pointing out their weaknesses, but suggesting that the synthesis may still be appropriate for the main thermocline. It probably is not relevant, however, near the sea surface or sea floor. The role of the bottom boundary in determining the spectral shape of the internal wave spectrum is discussed briefly in Section 3. Diapycnal mixing by shear instability of bottom reflected internal waves, and the way in which this may control the velocity boundary condition for low frequency flows, is also reviewed; the need for more observational work near sloping boundaries is emphasized. Returning to the upper ocean, Section 4 briefly presents a simple model for the shear-induced decay of an internal solitary wave. Section 5

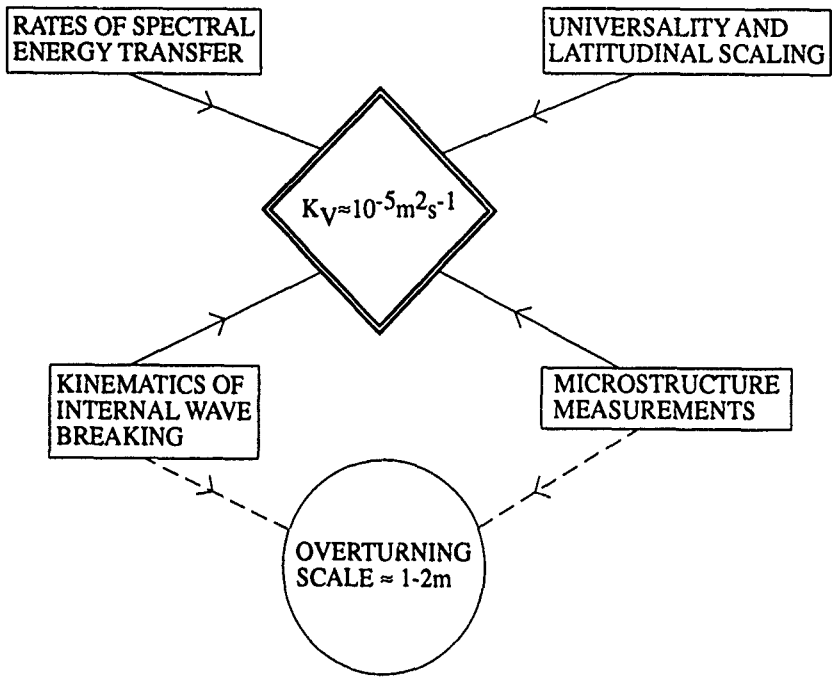


Figure 1: Schematic of a zeroth order view of various features of internal waves and turbulence that seem to be consistent with a diapycnal diffusivity of about 10^{-5} $m^2 s^{-1}$. (From Garrett, 1984).

is a personal impression of topics that need to be emphasized in future research.

2. THE PARADIGM

The purpose of Figure 1 was to demonstrate that a number of theoretical and observational properties of internal gravity waves in the main thermocline are compatible with a vertical eddy diffusivity there of no more than about 10^{-5} $m^2 s^{-1}$. Specifically:

(*i*) The rate of energy transfer through a typical internal wave spectrum has been estimated by McComas and Müller (1981b) using weak interaction theory and by Henyey *et al.* (1986) with a theory based on ray tracing. As pointed out by Gregg (1989), the two theories give essentially the same form E^2N^2 for the functional dependence of the energy flux on buoyancy frequency N and total internal wave E, but differ by a factor of nearly 7. Both theories depend on assumptions of limited validity; it remains to be established whether they are complementary, by applying in different regions of wavenumber-frequency space, and whether the factor of 7 discrepancy is an artifact of different, but hidden, assumptions.

The calculations of McComas and Müller (1981b) have been criticised for violating the

assumption of weak nonlinearity (Holloway, 1980), but the assumption does seem to be valid for the regions of the spectrum at low frequency and low vertical wavenumber which, after all, contain most of the internal wave energy. Thus, provided that these regions are not affected by back fluxes from the much less energetic regions where the assumption breaks down, it does seem reasonable to accept the McComas and Müller (1981b) estimates of the energy transfer to high wavenumber and ultimately to dissipation ϵ. For typical spectral levels, and with the increase in potential energy of the basic state being 20 to 25% of ϵ (Osborn, 1980; Oakey, 1982), the McComas and Müller (1981b) results then lead to a depth-independent K_V of less than 2×10^{-5} m^2s^{-1} (following Gregg (1989) in allowing for the energy transfer from low modes by both induced diffusion and parametric subharmonic instability).

(*ii*) Gregg's (1989) comparison of direct measurements of ϵ at various locations and depths led him to adopt a parameterisation for ϵ similar in form to those of McComas and Müller (1981b) and Henyey *et al.* (1986), but a factor of about 3 smaller than obtained by the former and a factor of 2 more than obtained by the latter. This in turn led to K_V less than about 5×10^{-6} m^2s^{-1} for typical internal wave energy levels.

(*iii*) For $K_V \simeq 10^{-5}$ m^2s^{-1}, with $\epsilon = 4$ to $5K_VN^2$ per unit mass and with an e-folding depth of 1,300 m for N, the typical vertically integrated internal wave energy level of about 4,000 Jm^{-2} is dissipated or converted to mean potential energy in about 40 to 50 days. A decay time of this order or even longer is compatible with the rather small variability in internal wave energy levels in space or time. If, for example, we think of the internal wave energy level E in an area as being governed by the equation

$$dE/dt + E/\tau = F \tag{1}$$

where τ is the decay time and F some exterior forcing, then the response to $F = F_0 + F_1 \cos\omega t$ is

$$E = \tau[F_0 + F_1(1+\omega^2\tau^2)^{-1/2}\cos\omega(t - t_0)] \tag{2}$$

where the time lag $t_0 = \omega^{-1}\tan^{-1}\omega\tau$. This may have some relevance to the finding by Briscoe and Weller (1984) and Briscoe (1984) of a lag of 2 to 3 months between internal wave energy levels and seasonal variations in wind stress. Taking $t_0 = 75$ days and $\omega = 2\pi/365$ days implies $\tau = 200$ days, somewhat longer than the 50 days or so (requiring $t_0 = 40$ days) corresponding to $K_V = 10^{-5}$ m^2s^{-1} but not much more than the 100 days

(requiring $t_0 = 60$ days) or more associated with Gregg's (1989) value of less than 5×10^{-6} m^{2}s^{-1} for K_V.

Equation (2) also implies that the variation of E about its mean should be less than the variation of the forcing by a factor $(1+\omega^2\tau^2)^{-1/2}$. For $\tau = 50, 100, 200$ days this factor is 0.76, 0.52, 0.28. Briscoe's (1984) North Atlantic data (his Figures 6b and 8) do suggest that the internal wave energy has less variation than the monthly average wind stress which is assumed to be the driving mechanism; a factor of about 0.4 seems appropriate, corresponding to a decay time of a bit more than 100 days.

Any further discussion of this should be regarded as over-interpretation, but it does seem possible that the time lag and variability of the internal wave energy level is associated with a decay time of about 100 days. Briscoe (1984) suggested that the lag represents the time for wind-forced mesoscale eddies to pass on their energy to the internal wave field, but has later also recognised the connection with the decay time of the internal wave spectrum (Mel Briscoe, personal communication) .

It has also been suggested in the past that the tendency for internal wave energy levels to be similar from place to place, in spite of spatially variable forcing, is a consequence of the rather long decay time which permits significant horizontal spreading of a patch of high energy. An important unpublished paper by Cox and Johnson (1978) discussed the way in which the spread of a patch of high internal wave energy is a diffusive process, with a horizontal diffusivity K_H proportional to the time for horizontal anisotropy to be removed by wave-wave interactions. An appropriate model in this situation, for forcing that varies in strength with wavenumber l in the north-south direction y, is

$$K_H d^2E/dy^2 - E/\tau + F_0 + F_1 \cos ly = 0, \tag{3}$$

with τ the internal wave decay time as before. The response is

$$E = \tau[F_0 + (1 + K_H\tau l^2)^{-1} F_1 \cos ly] \tag{4}$$

so that variations in forcing are smeared out if $K_H\tau l^2 \gg 1$.

For Cox and Johnson's (1978) estimate of 1.4×10^4 m^{2}s^{-1} for K_H, taking $\tau = 100$ days and with $l = 1.9 \times 10^{-6}$ m^{-1}, as for the 30° north-south wavelength of wind stress in the Atlantic (Briscoe, 1984), the factor $(1 + K_H\tau l^2)^{-1} = 0.70$. Thus significant variations in E, as reported by Briscoe (1984), should be observed, but this conclusion is highly

sensitive to K_H and hence to the time for horizontal anisotropy to be removed. Cox and Johnson (1978) took this to be 72 hours for the frequency of $1.3f$ contributing the most to their integral expression for K_H, but further analysis is long overdue.

It is also quite likely that it is dangerous to lump together all the frequencies in the inertio-internal wave spectrum. The seasonal variations reported by Briscoe and Weller (1984) and Briscoe (1984) were for high frequency internal waves (either $\omega > 1.15f$ or $\omega > 0.1$ cph in different analyses); the inertial band energy at the LOTUS site discussed by Briscoe and Weller (1984) showed a more complicated time dependence, with occasional bursts of high energy.

Spatially, Fu (1981) found that the inertial peak in the abyssal North Atlantic over smooth topography scaled as if it had originated as super-inertial energy at a lower latitude (though elsewhere it had a component due to local generation). This seems to imply that the wave does not lose much of its energy as it travels to its turning latitude, is internally reflected and travels back equatorward again. Taking y northward, with $y = 0$ at the latitude where $f = \omega$, then $f = \omega + \beta y$ with $\beta = \omega R^{-1} \cot\phi$, R being the earth's radius and ϕ the latitude. A wave of frequency ω and horizontal wave number then proceeds towards its turning latitude at a speed $\partial\omega/\partial l$, where the dispersion relation near the inertial frequency f is approximately

$$(\omega^2 - f^2)^{1/2} = (k^2 + l^2)^{1/2} c_j, \quad c_j = N_0 b (j\pi)^{-1} \tag{5}$$

for modenumber j in an ocean with a buoyancy frequency that is N_0 near the surface and decreases exponentially downwards with a scale depth b (e.g. Garrett and Munk, 1972). Hence, within the limits of ray tracing,

$$dy/dt = \partial\omega/\partial l \approx (2\beta/\omega)^{1/2} c_j (-y + y_0)^{1/2} \tag{6}$$

where $y_0 = -c_j^2 k^2 (2\omega\beta)^{-1}$ defines the turning latitude for east-west wavenumber k. Assuming that this WKB approximation is adequate to estimate the travel time T from some position $-Y$ to $-y_0$, then

$$T = [2R(Y - y_0)\tan\phi]^{1/2} c_j^{-1}. \tag{7}$$

If we take $Y - y_0 = 1{,}000$ km, $\phi = 45°$, $N_0 = 5.2 \times 10^{-3}$ s^{-1} and $b = 1{,}300$ m, then $T = 20j$ days. This suggests that any processes that would tend to remove the inertial

peak in the spectrum cannot do so in less than several tens of days.

On the other hand, as also recognised by Eric Kunze (personal communication), the latitudinal scaling of spectra is a two-way process; perhaps some of the energy at frequencies above f originates as inertial energy at a more poleward location. In fact, the higher-than-WKB inertial peak found by Fu (1981)in some locations suggests that the extra energy must be lost rather quickly.

To summarise this discussion:

(a) The temporal variability of the typical high frequency internal wave field suggests a decay time of order 100 days, with K_V no more than 10^{-5} m^2s^{-1}.

(b) Spatial variability of annual average high frequency energy levels can occur even if the decay time of the energy is 100 days or so, though this conclusion is strongly dependent on the poorly-known rate at which horizontal anisotropy is removed from a spectrum.

(c) To the extent to which the inertial peak is a latitudinal turning point effect, it also seems to imply a lifetime of several tens of days. However, regions with more elevated peaks must lose their energy in a time less than this.

(*iv*) The last main feature of Figure 1 concerns the kinematics of breaking waves. Garrett (1984) argued that typically observed shear spectra in the ocean would lead to shear instability and overturning on a vertical scale of order 1 m (as typically observed; e.g. Gregg, 1989) and that, if mixing events occur no more frequently than shear maxima, then again K_V is no more than about 10^{-5} m^2s^{-1}. These arguments seem as valid now as they did in 1984, and it does seem appropriate to point out that the estimate of the vertical scale was based on integral properties of the whole shear spectrum rather than on the vertical wavenumber of some particular kink in the spectrum!

In conclusion, to the extent to which Figure 1 represents a paradigm, it still seems reasonably valid for the ocean interior away from the surface or bottom topography where bursts of inertial energy may lead to a local increase in K_V. As before, though, it largely remains as a reminder of some of the different conceptual elements which need to be considered in developing any unified picture of the mixing produced by internal waves. It must also be remembered that processes other than internal wave breaking can lead to diapycnal mixing.

3. THE ROLE OF THE BOTTOM BOUNDARY

3.1 Effect on the Spectral Shape.

Most discussions of the internal wave wavenumber-frequency spectrum assume that it is determined by some combination of generation (most likely at the sea surface by wind) and nonlinear interaction in the ocean interior. Calculations by Eriksen (1982), Rubenstein (1988), Garrett and Gilbert (1988) and Xu (1990), however, show that reflection off bottom topography can scatter energy to higher wavenumber (though the frequency of each wave is unchanged). Thus waves which can reach the sea floor in a time short compared with their interaction time may have their energy scattered to higher wavenumbers, at the same frequency, sooner than they lose energy to other frequencies and wavenumbers by nonlinear interactions.

The vertical component of the group velocity of a wave of frequency ω and vertical wavenumber m is

$$c_{gz} = \left(\frac{N^2-\omega^2}{N^2-f^2}\right)\left(\frac{\omega^2-f^2}{\omega m}\right) \tag{8}$$

$$\simeq \left(\frac{\omega^2-f^2}{\omega m}\right) \quad \text{for } \omega \ll N \tag{9}$$

with $m \simeq j\pi N(bN_0)^{-1}$ in an ocean with an exponential N profile of scale height b (Munk, 1981), where j is the mode number. If b is significantly less than the ocean depth, the vertical travel time of waves with frequency less than N at the bottom is then $\pi\omega j(\omega^2-f^2)^{-1}$, or $\pi j\omega^{-1}$ away from f.

Now McComas and Müller (1981a, b) showed that, for a typical internal wave spectrum, the time scale for parametric subharmonic instability to drain energy from low modes with $\omega > 2f$ is given by $(0.27Em_*mN^{-2})^{-1}\omega^{-1}$ where m_* is the vertical wavenumber bandwidth. Taking $m = j\pi N(bN_0)^{-1}$, $m_* = j_*\pi N(bN_0)^{-1}$, $j_* = 3$, $N_0 = 5\times10^{-3}$ s^{-1}, $b = 1300$ m and $E = 2\pi\times10^{-5}b^2N_0N$, this time scale becomes $2\times10^3(N_0/N)j^{-1}\omega^{-1}$. Ignoring the depth dependence here by taking $N = N_0$, we see that this interaction time is greater than the vertical travel time $\pi j\omega^{-1}$ if $j < 25$.

Thus for the energetic low modes at frequencies greater than $2f$, interactions with bottom topography might be as significant as interactions with other waves in the water column. Conceivably the tendency for non-inertial waves to have a less red vertical wavenumber spectrum than inertial waves (Tom Sanford, this volume) may reflect this, although such a

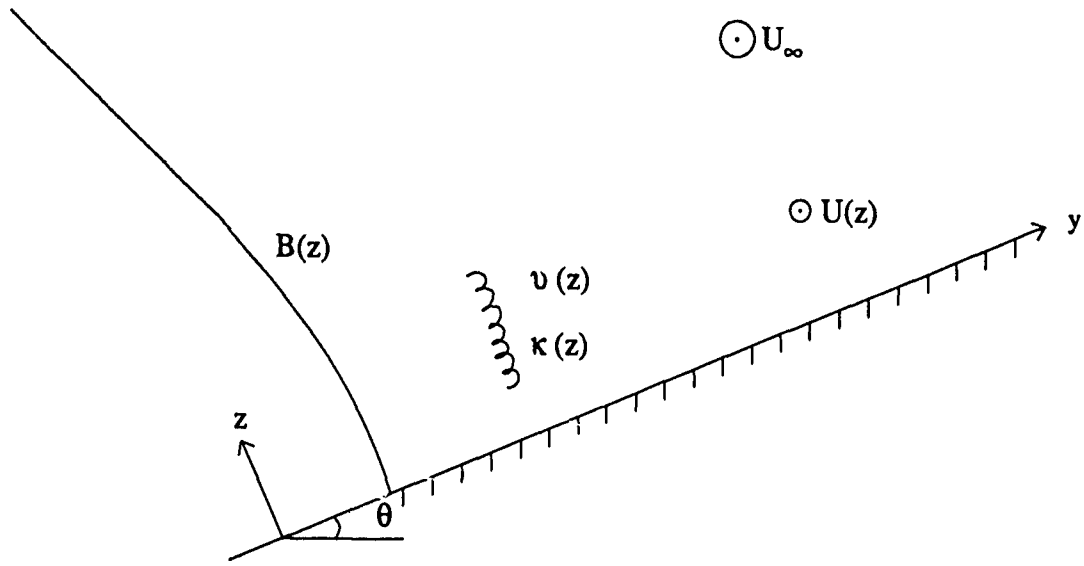


Figure 2: Definition sketch for mixing near a sloping boundary.

tendency may also be a consequence of nonlinear wave interactions (McComas and Müller, 1981a).

3.2 Mixing by Internal Waves Breaking Near a Slope

Figure 2 presents a schematic of a plane bottom slope at which internal waves are reflected, mainly into shorter vertical wavelengths with higher energy (e.g. Eriksen, 1982). This may well lead to shear instability and mixing close to the boundary (Eriksen, 1985; Garrett and Gilbert, 1988), though less so in concave regions (Gilbert and Garrett, 1989). It is assumed that the subsequent eddy fluxes of momentum and buoyancy can be represented by eddy viscosity and diffusivity profiles $\nu(z)$, $\kappa(z)$ as functions of the bottom-normal coordinate z. (This does not assume that a mixing length theory is valid; ν and κ are the eddy fluxes divided by the mean gradients). The mean buoyancy $B(z)$ is reduced near the boundary by the mixing, but $\partial B/\partial z \to N^2 \cos\theta$ as $z \to \infty$ with N^2 the constant interior stratification. In the upslope direction $\partial B/\partial y = N^2 \sin\theta$ for all z as the problem is assumed independent of y to lowest order.

The physics and mathematics of this boundary mixing problem are discussed by Phillips *et al.* (1986), Thorpe (1987) and Garrett (1990) amongst others. Here I wish merely to draw attention to some key factors that suggest the need and opportunities for observational work in this region.

Garrett (1990) showed that the vertical buoyancy flux is the value $\int_0^\infty \kappa(z) N^2 dz / \sin\theta$ that it would take if the stratification was unchanged by the mixing, multiplied by an

"effectiveness"

$$I = \int_0^\infty \kappa(z) \left[\sin^2\theta + \left(\frac{\partial B/\partial z}{N^2 \cos\theta} \right)^2 \cos^2\theta \right] dz \Big/ \int_0^\infty \kappa dz \tag{10}$$

This is 1 if $\partial B/\partial z = N^2 \cos\theta$ throughout the boundary layer but is reduced if $\partial B/\partial z < N^2 \cos\theta$. Due to secondary circulation the reduction factor in the second term in (10) is the square of what it would be due to reduced stratification alone. Now if ν_0, κ_0 are the eddy coefficients close to the boundary, then $\partial B/\partial z$ is very small over a thickness of order h_{mixed} where

$$h_{\text{mixed}} = \left(\frac{f^2}{\nu_0^2} + \frac{N^2 \sin^2\theta}{\nu_0 \kappa_0} \right)^{-1/4} \tag{11}$$

$$= h_{\text{Ekman}} (1 + S\ Pr)^{-1/4}. \tag{12}$$

Here $h_{\text{Ekman}} = (\nu_0/f)^{1/2}$ is an Ekman depth, $S = N^2 \sin^2\theta / f^2$ is the Burger number based on the ocean interior stratification and the bottom slope and $Pr = \nu_0/\kappa_0$ is the eddy Prandtl number close to the boundary.

If significant mixing dies away outside a boundary layer of thickness approximately h_{mixed}, then the boundary mixing effectiveness from (10) is small since $\partial B/\partial z$ is small where κ is large and vice versa. If, however, vigorous mixing extends well into the region above h_{mixed}, where buoyancy driven flows can restore the stratification, then I can be significant. This may be the case for internal wave breaking. The stratification is, of course, tied to the values of $\nu(z)$, $\kappa(z)$. Garrett (199.) shows that if $\nu(z)$, $\kappa(z)$ fall off to very small values over a distance much greater than h_{mixed}, but assuming $\nu/\kappa = Pr$ =constant for simplicity, then, for $z \gtrsim h_{\text{mixed}}$,

$$\partial B/\partial z \simeq N^2 \cos\theta \left(\frac{S\ Pr}{1 + S\ Pr} \right). \tag{13}$$

The nature of this region is that both the thermal wind equation for the along-slope flow and an advective (upslope)/diffusive (bottom-normal) balance for buoyancy are satisfied. For this solution, (10) gives

$$I \simeq \left(\frac{S\ Pr}{1 + S\ Pr} \right)^2 \tag{14}$$

so that the mixing is particularly effective if $S\ Pr$ is not small.

3.3 The Alongslope Flow

Steady state boundary mixing theories also lead to formulae for the alongslope mean flow $U(z)$. For arbitrary profiles of $\nu(z)$, $\kappa(z)$ it can be shown (Garrett, 1990) that

$$U(z) = \int_0^z f \cot\theta [\kappa_\infty - \kappa(\partial B/\partial z)/N^2 \cos\theta]\nu^{-1} dz \tag{15}$$

where κ_∞ is the (small) value of κ as $z \to \infty$. The bottom-normal buoyancy gradient $\partial B/\partial z$, which tends to $N^2 \cos\theta$ as $z \to \infty$, is, of course, part of the solution. For $Pr = \nu(z)/\kappa(z) =$ constant and with mixing, as before, extending over a distance H from the boundary much greater than h_{mixed}, (15) shows that as $z \to \infty$, $U(z) \to U_\infty$ given by

$$U_\infty \simeq -fH \cot\theta \left(\frac{S}{1 + S\ Pr}\right). \tag{16}$$

As argued by MacCready and Rhines (1991) and Garrett (1991), it is this downwelling-favourable alongslope flow that becomes the boundary condition for the ocean interior velocity and very slowly diffuses into the interior.

These steady state solutions for the density stratification, mixing effectiveness and alongslope flow will undoubtedly turn out to be a gross oversimplification of events occurring in the real ocean. They do, however, draw attention to the need for measurements of eddy momentum fluxes as well as eddy buoyancy fluxes near the sloping sea floor. An observational program on a sloping bottom, using acoustic doppler current profilers to measure Reynold's stresses with the techniques described by Plueddemann (1987) and Lohrmann *et al.* (1990), could be very rewarding.

4. A SIMPLE MODEL FOR THE DECAY OF AN INTERFACIAL SOLITARY WAVE

Much of the study of internal waves has been concerned with their generation, interaction and dissipation in the open ocean. They also exist, however, in the shallower water of continental shelves, where they may have very different wavenumber/frequency spectra and have significantly different behaviour and effects.

In particular, near-surface solitary waves, usually originating from nonlinear internal tides at the shelf break, are a rather common feature of stratified continental shelf waters and also occur in the deep ocean (e.g. Ostrovsky and Stepanyants, 1989). In any location they provide a possible mechanism for vertical mixing.

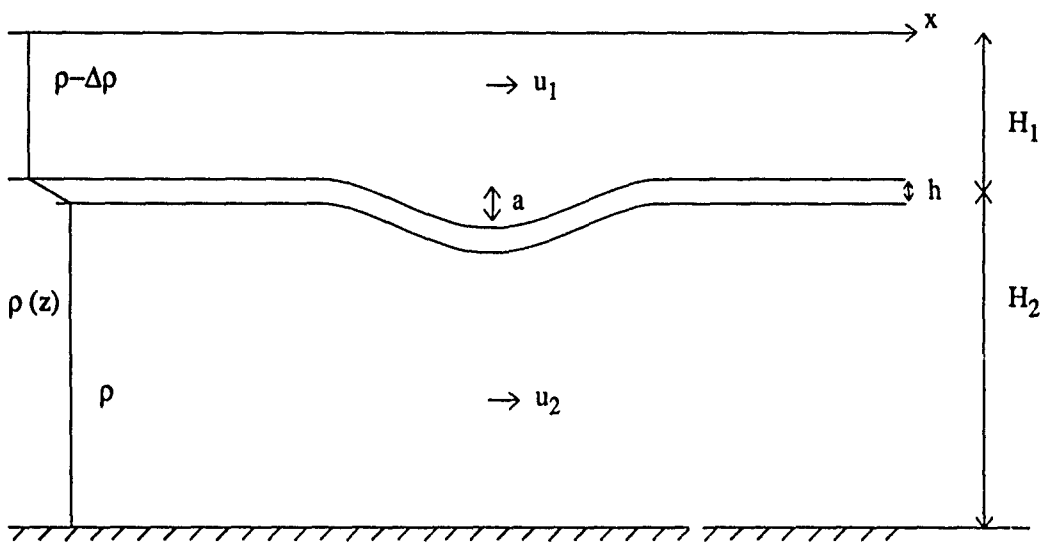


Figure 3: Schematic of an internal solitary wave in a two layer ocean with a thin interface.

In this section I summarise the simple criteria for mixing, and associated solitary wave decay, derived by Bogucki and Garrett (1991) for a fluid with two homogeneous layers separated by a thin interface. The situation envisaged is sketched in Figure 3; layers with thicknesses H_1, H_2 and densities $\rho - \Delta\rho$ and ρ are separated by a thin interface of thickness h ($\ll H_1$, H_2). A solitary wave of amplitude a generates currents u_1, u_2 in the two layers.

In a frame of reference moving with the wave speed c, continuity in the two layers requires

$$(c-u_1)(H_1+a) = cH_1, \quad (c-u_2)(H_2-a) = cH_2 \tag{17}$$

$$u_1 = ca(H_1+a)^{-1}, \qquad u_2 = -ca(H_2-a)^{-1} \tag{18}$$

if the horizontal length scale L of the solitary wave is much greater than H_1 and H_2. The Richardson number in the interface is then

$$Ri = g'h(u_1-u_2)^{-2}, \quad g' = g\Delta\rho/\rho \tag{19}$$

$$= g'h(H_1+a)^2(H_2-a)^2c^{-2}(H_1+H_2)^{-2}a^{-2}. \tag{20}$$

In the simplest situation, where $H_2 \gg H_1$ and $a \ll H_1$, we have $c^2 \simeq g'H_1$ and

$$Ri = hH_1/a^2. \tag{21}$$

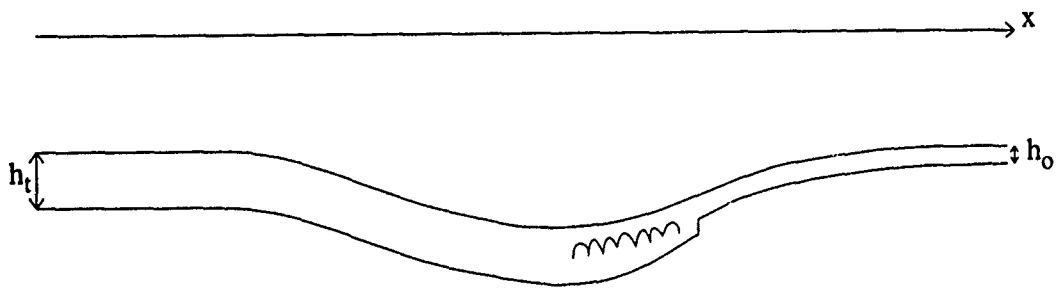


Figure 4: The internal solitary wave is progressing in the x direction and triggers shear instability at some point on the forward face. The interface quickly thickens and keeps thickening until the wave's maximum displacement is reached.

Hence $Ri < \frac{1}{4}$ for instability requires

$$a > 2(hH_1)^{1/2}. \tag{22}$$

This simple criterion for instability, that the solitary wave amplitude should be greater than twice the geometric mean of upper layer and interface thicknesses, also applies if we relax the assumption that $L \gg H_2$. In that case, as for "Benjamin-Ono" solitons (e.g. Ono, 1975) which require $L \ll H_2$, we still have $u_2 \ll u_1$ and (22) still applies subject to $a \ll H_1$ as before.

For solitary waves of amplitude greater than $2(hH_1)^{1/2}$ it seems plausible that shear instability would occur on the forward face as soon as the interfacial Richardson number drops below some critical value Ri_c which is presumably $\frac{1}{4}$. The growth and collapse of Kelvin-Helmholtz billows, and restratification of the fluid into a thicker interface layer, may all occur in a time scale given by some multiple of N^{-1}, where N is the buoyancy frequency in the interface and is given by $N^2 = g'/h$. Hence $N^{-1} = (hH_1)^{1/2}(g'H_1)^{-1/2}$. This is the time taken by the wave to travel a distance $(hH_1)^{1/2}$ which is very much less than L, indicating that the mixing proceeds rapidly compared with the passage time of the wave. Hence one might expect that, after the initial instability, the interface thickens rapidly to achieve some interfacial Richardson number Ri_t ($\simeq 0.4$ according to Thorpe (1972)). As sketched in Figure 4, the interface might then continue to grow until the wave crest passes, after which it remains at the same thickness given by

$$h_t = Ri_t\, a^2/H_1. \tag{23}$$

This thickening of the interface, from an initial thickness of h_0, has an associated increase

in potential energy of $g'\rho(h_t^2 - h_0^2)/24$. If we assume that this energy is a fraction α of the total energy lost by the propagating soliton, which is assumed to hold together in spite of the dissipation, we have

$$\alpha\, dE/dx = -g'\rho(h_t^2 - h_0^2)/24 \quad (24)$$

$$= -g'\rho Ri_t^2 H_1^{-2}(a^4 - a_0^4)/24 \quad (25)$$

where $a_0 = (H_1 h_0/Ri_t)^{1/2}$ is less than the critical amplitude $(H_1 h_0/Ri_c)^{1/2}$ at which decay of the soliton would cease.

Further development of (25) requires a formula for the total soliton energy E. For $L \gg H_2$ the soliton interface displacement ζ as a function of x is given by $\zeta = a\mathrm{sech}^2(x/L)$ where, for $a \ll H_1$, $aL^2 = \frac{4}{3}H_1^2 H_2$ (e.g. Benney, 1966). Hence

$$E = \frac{4}{3}g'\rho a^2 L = \left(\frac{4}{3}\right)^{3/2} g'\rho H_1 H_2^{1/2} a^{3/2} \quad (26)$$

and, from (25) with $Ri_t = 0.4$ and $\alpha = 0.2$ (Oakey, 1982),

$$da/dx = -0.014 H_1^{-3} H_2^{-1/2} a^{-1/2}(a^4 - a_0^4). \quad (27)$$

If, on the other hand, $L \ll H_2$, the "Benjamin-Ono" solitons (Ono, 1975) have $\zeta = a[(x/L)^2 + 1]^{-1}$ with $aL = \frac{4}{3}H_1^2$ so that

$$E = (\pi/2)g'\rho a^2 L = (2\pi/3)g'\rho H_1^2 a. \quad (28)$$

Hence, with $\alpha = 0.2$ and $Ri_t = 0.4$ as before,

$$da/dx = -0.016 H_1^{-4}(a^4 - a_0^4). \quad (29)$$

In both situations, therefore, waves which meet the instability criterion (22) decay rather rapidly at first before eventually stabilising at $a_c = (H_1 h_0/Ri_c)^{1/2}$. As $a_c^4 = (Ri_t/Ri_c)^2 a_0^4 = 2.56 a_0^4$ for $Ri_t = 0.4$ and $Ri_c = 0.25$, the decay rate is rather insensitive to h_0; typical decay distances in either situation are several tens of kilometres (Bogucki and Garrett, 1991).

Application of these simple ideas to the real ocean may require allowance for other effects such as geometrical spreading and, more importantly, extension of the model to allow for a continuously stratified upper layer.

5. DISCUSSION

A picture of a rather persistent internal wave field, slowly losing energy to dissipation and vertical mixing with $K_V \simeq 10^{-5}$ m^{2}s^{-1}, still seems appropriate for the main thermocline. As remarked by Gregg (1987), diapycnal mixing may then be a rather unimportant process compared with other processes such as ventilation from the surface mixed layer. One part of this picture is the prediction of ϵ from nonlinear internal wave interaction theories. The validity of the model of Henyey *et al.* (1986) is still a matter of debate (e.g. Gargett, 1990), but it is, in a sense, a local model, giving the local flux of energy to small scales and hence mixing for a typical spectrum. The calculations of McComas and Müller (1981a) do seem to be valid for the energetic low-frequency, low-mode, parts of the spectrum, but it is just these spectral regions which are sufficiently long-lived to propagate into a region with a different inertial frequency or to interact with bottom topography. Thus the energy losses from the energetic part of the spectrum in the McComas and Müller (1981b) theory do not immediately lead to turbulence. They need to be integrated with consideration of other processes affecting the internal wave spectrum before we can be confident of the implications for local mixing.

In particular, the pioneering study by Cox and Johnson (1978) of the consequences of horizontal anisotropy in the internal wave spectrum need to be re-examined; the problem is central to any consideration of the evolution of a spectrum that is distorted by bottom reflection, surface generation or just the lateral spreading of a patch of high energy.

The simple scenario shown in Fig. 1 is probably not valid near the sea surface. In particular, if an elevated inertial peak is generated there which is not connected in a WKB fashion to spectra at lower latitudes, then some extra local dissipation must be occurring.

The region near the seafloor is also likely to be anomalous. It has been emphasized here that the mixing produced by internal waves near a sloping bottom may be significant not only for basin-average diapycnal fluxes, but also for the velocity boundary condition for low frequency flows. There is a need to mount more observational programs, capable of measuring eddy momentum and buoyancy fluxes, near sloping bottoms, particularly those that are convex (to avoid destructive interference of reflected internal waves) and have a Burger number greater than 1 (to emphasize slope effects).

Our developing understanding of the deep-sea internal wave field will probably be of

limited applicability on continental shelves, though there are the same concerns with surface generation and the role of a sloping bottom. A reasonably common phenomenon of shelf seas is the internal soliton; this paper has presented a simple preliminary model of the mixing produced at a thin interface by an internal soliton which is itself damped in the process.

In summary, therefore, the paradigm of Fig. 1 may be appropriate for some parts of the ocean, though some of its foundations are still rather shaky. Many interesting and important questions remain, however, about parts of the ocean where this picture is not appropriate but where internal waves have significant effects. In other words, our paradigm may hold in regions where internal waves do not matter and be inadequate in places where they do!

ACKNOWLEDGMENTS

I thank the organisers of this stimulating meeting for the invitation to participate (as "Parasite Regained"?). I also thank Mel Briscoe, Eric Kunze and Darek Bogucki for discussion and Barry Ruddick and Dan Kelley for comments. I appreciate the financial support of the Office of Naval Research and Canada's Natural Sciences and Engineering Research Council.

REFERENCES

Benney, J.D., 1966: Long non-linear waves in fluid flows. *J. Math. and Physics*, **45**, 52–63.

Bogucki, D. and C. Garrett, 1991: A simple model for the shear-induced decay of an internal solitary wave. (in preparation).

Briscoe, M.G., 1984: The monthly variability of upper-ocean internal wave energy: a progress report on the variation with wind stress. In *Internal Gravity Waves and Small-scale Turbulence, Proceedings, Hawaiian Winter Workshop*, P. Müller and R. Pujalet, Eds., 129–150. Hawaii Institute of Geophysics, Honolulu.

Briscoe, M.G. and R.A. Weller, 1984: Preliminary results from the Long-Term Upper-Ocean Study (LOTUS). *Dyn. Atmos. and Oceans*, **8**, 243–265.

Cox, C.S. and C.L. Johnson, 1978: Inter-relations of micro processes, internal waves, and large-scale ocean features. Unpublished manuscript, Scripps Institution of Oceanography.

Eriksen, C.C., 1982: Observations of internal wave reflection off sloping bottoms. *J. Geophys. Res.*, **87**, 525–538.

Eriksen, C.C., 1985: Implications of ocean bottom reflection for internal wave spectra and mixing. *J. Phys. Oceanogr.*, **15**, 1145–1156.

Fu, L.-L., 1981: Observations and models of inertial waves in the deep ocean. *Rev. Geophys. and Space Phys.*, **19**, 141–170.

Gargett, A.E., 1990: Do we really know how to scale the turbulent kinetic energy dissipation rate ϵ due to breaking of oceanic internal waves? *J. Geophys. Res.*, **95**, 15971–15974.

Garrett, C., 1984: Parameterizing the effects of internal waves: simple ideas and things we need to know. In *Internal Gravity Waves and Small-scale Turbulence, Proceedings, Hawaiian Winter Workshop*, P. Müller and R. Pujalet, Eds., 171–181. Hawaii Institute of Geophysics, Honolulu.

Garrett, C., 1990: The role of secondary circulation in boundary mixing. *J. Geophys. Res.*, **95**, 3181–3188.

Garrett, C., 1991: Marginal mixing theories. *Atmosphere-Ocean*, **29**, 313–339.

Garrett, C. and D. Gilbert, 1988: Estimates of vertical mixing by internal waves reflected off a sloping bottom.In *Small-Scale Turbulence and Mixing in the Ocean*, J.C.J. Nihoul and B.M. Jamart, Eds., Elsevier, pp. 405–423.

Garrett, C. and W. Munk, 1972: Space–time scales of internal waves. *Geophys. Fluid Dyn.*, **2**, 225–264.

Gilbert, D. and C. Garrett, 1989: Implications for oceanic mixing of internal wave scattering off irregular topography. *J. Phys. Oceanogr.*, **19**,1716–1729.

Gregg, M.C., 1987: Diapycnal mixing in the thermocline: a review. *J.Geophys. Res.*, **92**,5249–5286.

Gregg, M.C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.*, **94**, 9686–9698.

Henyey, F.S., J. Wright and S.M. Flatté, 1986: Energy and action flow through the internal wave field: An eikonal approach. *J. Geophys. Res.*, **91**, 8487–8495.

Holloway, G., 1980: Oceanic internal waves are not weak waves. *J. Phys. Oceanogr.*, **10**, 906–914.

Lohrmann, A., B. Hackett and L.P. Røed, 1990: High resolution measurements of turbulence, velocity and stress using a pulse-to-pulse coherent sonar. *J. Atmos. and Oceanic Technology*, **7**, 19–37.

MacCready, P. and P.B. Rhines, 1991: Buoyant inhibition of Ekman transport and its effect on stratified spin–up. *J. Fluid Mech.*, **223**, 631–661.

McComas, C.H. and P. Müller, 1981a: Time scales of resonant interactions among oceanic internal waves. *J. Phys. Oceanogr.*, **11**, 139–147.

McComas, C.H. and P. Müller, 1981b: The dynamic balance of internal waves. *J. Phys. Oceanogr.*, **11**, 970–986.

Munk, W.H., 1981: Internal waves and small–scale processes. In *Evolution of Physical Oceanography, Scientific Surveys in Honor of Henry Stommel*, B.A. Warren and C.

Wunsch, Eds., M.I.T. Press, pp. 264–291.

Oakey, N.S., 1982: Determination of the rate of dissipation of turbulent energy from simultaneous temperature and velocity shear microstructure measurements. *J. Phys. Oceanogr.*, **12**, 256–271.

Ono, H., 1975: Algebraic solitary waves in stratified fluids. *J. Phys. Soc. Japan*, **39**, 1082–1091.

Osborn, T.R., 1980: Estimates of the local rate of vertical diffusion from dissipation measurements. *J. Phys. Oceanogr.*, **10**, 83-89.

Ostrovsky, L.A. and Y.A. Stepanyants, 1989: Do internal solitons exist in the ocean? *Rev. Geophys.*, **27**, 293–310.

Phillips, O.M., J.–H. Shyu and H. Salmun, 1986: An experiment on boundary mixing: mean circulation and transport rates. *J. Fluid Mech.*, **173**, 473–499.

Plueddemann, A.J., 1987: Observations of the Upper Ocean Using a Multi– Beam Doppler Sonar. Ph.D.Thesis, Scripps Institution of Oceanography, 183 pp.

Rubenstein, D., 1988: Scattering of inertial waves by rough bathymetry. *J. Phys. Oceanogr.*, **18**, 5–18.

Thorpe, S.A., 1972: Turbulence in stably stratified fluids: a review of laboratory experiments. *Boundary Layer Meteorology*, **5**, 95–119.

Thorpe, S.A., 1987: Current and temperature variability on the continental slope. *Phil. Trans. R. Soc. London, Ser. A*, **323**, 471–517.

Xu, N., 1990: Interactions of internal waves with random bottom topography and a straight slope. M.Sc. Thesis, University of Hawaii, 78 pp.

A STRATEGY FOR INVESTIGATING AND MODELING INTERNAL WAVE SOURCES AND SINKS

Eric A. D'Asaro

Applied Physics Laboratory, College of Ocean and Fishery Sciences, University of Washington, Seattle, WA 98105-6698

ABSTRACT

We do not know what forces the internal wave field, except in the broadest terms. The spectral levels are remarkably uniform, but the likely sources have obvious spatial and temporal variability. This argues that long-range horizontal propagation of energy must be an important factor in "smoothing out" the irregularity of the sources. Research suggests that only low-mode, low-frequency internal waves are capable of propagating significant distances. In the short term, measurements of the directional spectra of these waves should therefore reveal the dominant energy sources for the internal wave field. In the longer term, a global model of the internal wave field and its associated diapycnal mixing could be constructed by simulating the generation, propagation, and dissipation of these low-mode, low-frequency waves. This is a worthy long-term goal for internal wave research.

WHAT DO WE WANT TO KNOW?

In the last 20 years we have learned a great deal about the spectral and spatial distribution of internal wave energy. Measurements show the spectrum is remarkably stable compared, for example, with that of surface waves. It is therefore useful to talk about a "universal" spectral shape for internal waves (Garrett and Munk, 1979). A nearly universal spectrum is consistent with theory (Müller et al., 1986), which predicts that wave–wave interactions produce an internal wave spectrum that is not very sensitive to the details of the wave sources. More recently, direct comparisons between measurements of the kinetic energy dissipation rate, ε, and values predicted by wave–wave interaction theories (see Gregg and Padman discussions, this volume) show some agreement, although the story is far from complete at present. Notably, although only small variations are seen in the *spectral level*, they correspond to much larger variations in ε, a *spectral transfer rate*. The internal wave field thus looks much more dynamic and much less universal when viewed in terms of rates rather than levels.

In contrast, only modest progress has been made in determining the sources of the internal wave field. There are variations in the energy and spectral level (Wunsch and Webb, 1979; Levine et al., 1985). Some investigators[1] have found correlations between sources and internal wave characteristics. Levine et al. (1985) find a wind forced signal in the Arctic Ocean; Padman and Dillon (1991) and Levine et al. (1983) find tidally forced signals; Kunze and Sanford (1984) find strong signals associated with mesoscale rings and fronts; and D'Asaro (1985a) finds a signal from strong storms. In general, however, it has been difficult to find any clear association between the wave field's sources and sinks and its behavior. This is not due to any lack of

[1]This is clearly an incomplete list I apologize to investigators who have not been mentioned.

potential sources and sinks. Idealized calculations indicate that wind stress (D'Asaro, 1985b), surface waves (Watson, this volume), bottom topography (Bell, 1975), and geostrophic currents (Watson, 1985), among others, could all potentially be important. The problem is distinguishing among these sources.

I propose that the next goal of internal wave research should be to find and quantify the internal wave sources. The key to the strategy is to realize that only low-mode, low-frequency internal waves can propagate substantial distances (>1000 km) in the ocean. These waves must therefore supply energy to the wave field in regions where the sources are weak. Measurements of the directional spectrum of these low-mode waves should reveal their sources, since different sources will undoubtedly be distributed differently in space and time. In the long term, we might consider working toward global models of the internal wave field which track energy from sources into propagating low modes, back into the local wave field, and finally to local dissipation and diapycnal mixing. The remainder of this paper will elaborate on these ideas.

THEORETICAL RESULTS

Scales

Internal waves span a wide range of scales and thus a wide range of dynamical regimes. Figure 1 shows the spectral domain of internal waves as a function of WKB normalized, hydrostatic, vertical wavenumber (horizontal axis), frequency (vertical axis), and horizontal

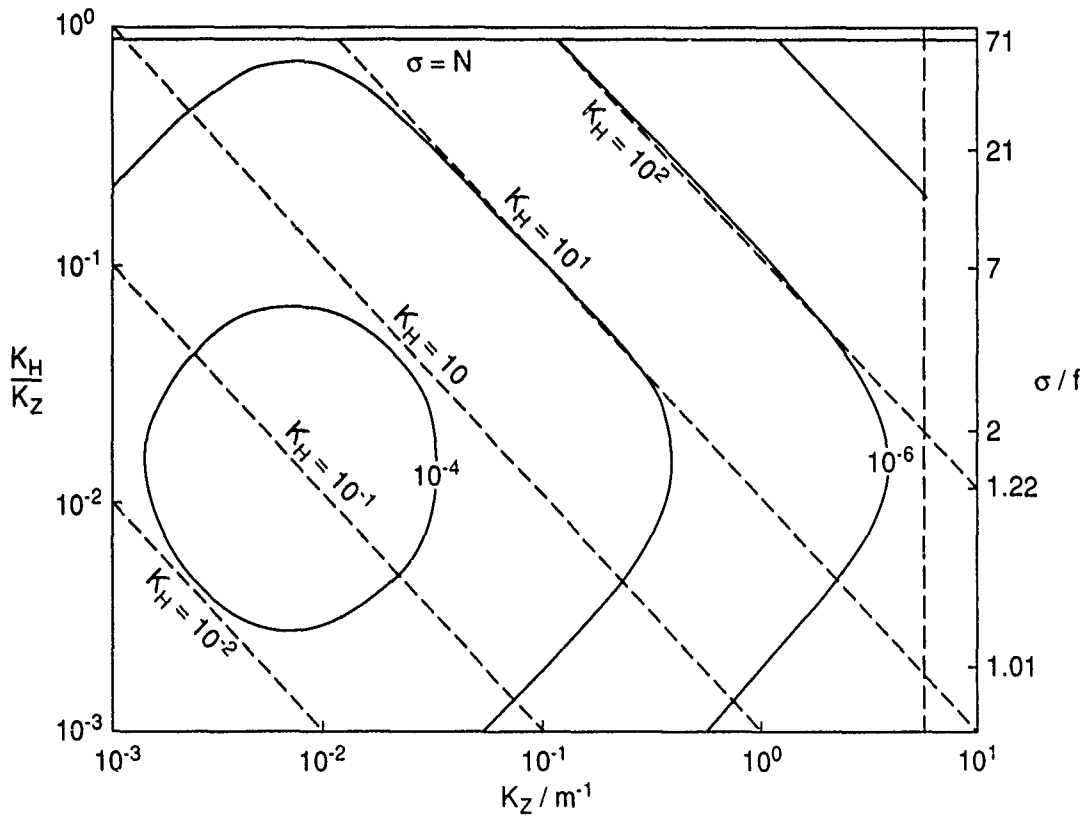


Fig. 1. Spectral density of internal wave energy according to GM76 [after McComas, 1977] as a function of vertical wavenumber K_Z (horizontal axis), frequency σ (vertical axis), and horizontal wavenumber K_H (dashed diagonal lines) using the hydrostatic approximation.

wavenumber (diagonal lines). The distribution of energy is contoured (Garrett and Munk, 1975). Most of the energy is at low wavenumbers and low frequencies.

Propagation

Internal waves can propagate horizontally and thus spread energy away from source regions. The horizontal group speed is contoured in Fig. 2 (dashed curves). Only the low modes have significant speeds.

Very nearly inertial, low-mode waves cannot exist, since the β effect changes f by more than $\sigma - f$ over their inverse horizontal wavenumber (Fu, 1981). This region is indicated by the shaded region in the lower left-hand corner of Fig. 2. Vorticity gradients in geostrophic flows are usually comparable to β and sometimes far exceed it (Kunze, 1985). The propagation of internal waves will be strongly modified by interactions with these flows. The region of limited propagation will thus be considerably larger than the "no waves" region in Fig. 2.

The low phase speed of many internal waves limits their propagation. The frequency of waves propagating in a background flow will be Doppler shifted by an amount $\Delta\sigma = \vec{k}\cdot\vec{U}$, where $\vec{k}$ is their horizontal wavenumber and $\vec{U}$ is the change in the background velocity through which they have propagated. If $\Delta\sigma$ exceeds $\sigma - f$, the waves may be Doppler shifted into a critical

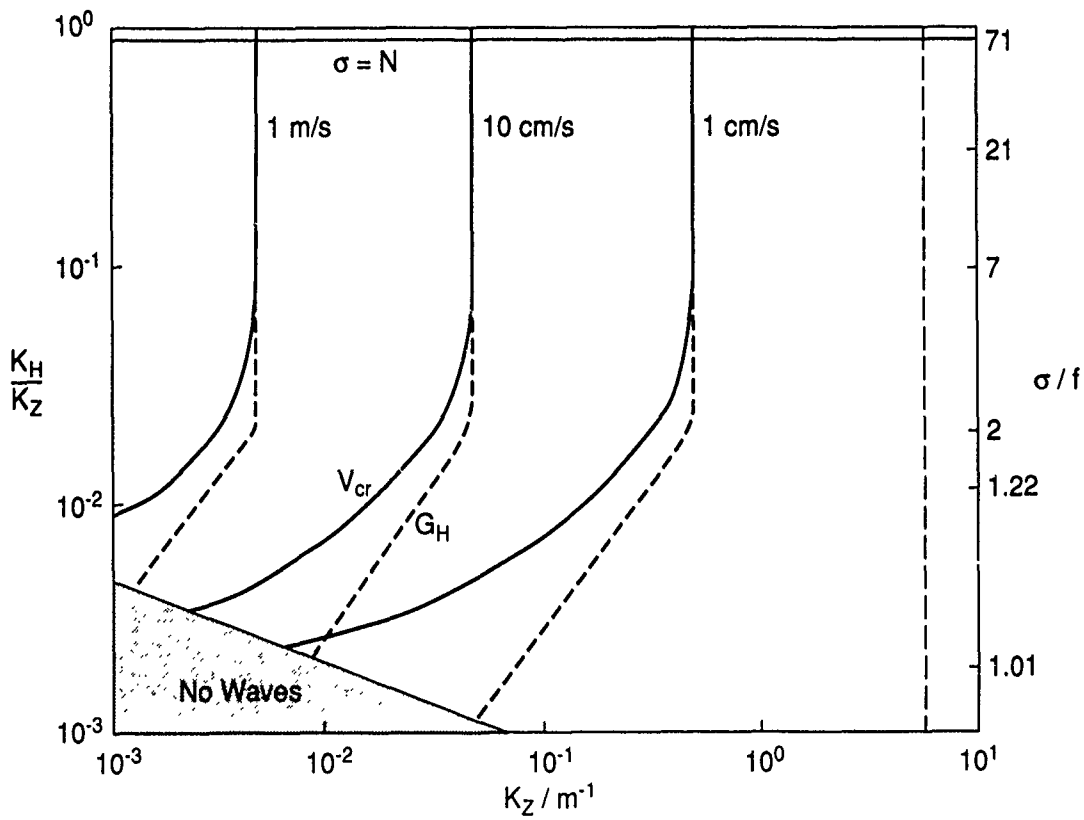


Fig. 2. Contours of hydrostatic, WKB horizontal group speed G_H (dashed curves) and U_{cr}, the background velocity necessary for critical layers (solid curves). The region of waves excluded by the β effect is shaded.

layer. Even if they are not, they will be strongly distorted. Contours of the background velocity necessary for the formation of critical layers are plotted in Fig. 2; $U_{cr} = (\sigma - f)/|\vec{k}|$. Because the internal wave field itself commonly has velocities approaching 10 cm/s, waves with $U_{cr} <$ 10 cm/s are unlikely to propagate far. This defines a region of strong wave–wave interaction at mid and high wavenumbers in which wave propagation is limited. Mesoscale geostrophic motions commonly have velocities of 10–100 cm/s. Waves with $U_{cr} < 100$ cm/s will not propagate far in regions with such a mesoscale eddy field. Oceanic velocities in excess of 100 cm/s are uncommon except in the strongest currents. The propagation of waves with $U_{cr} > 100$ cm/s will usually not be inhibited by interaction with ocean currents.

Wave–Wave Interaction

The sources of internal waves are believed to be concentrated at low wavenumbers. Wave–wave interaction theories predict a cascade of energy from small to large wavenumbers, as indicated by the arrows in Fig. 3. Although there is some disagreement (in details) between the results of resonant and eikonal approximations (Müller et al., 1986), there is consensus that these interactions feed energy to small-scale dissipation and mixing. The transfer of energy

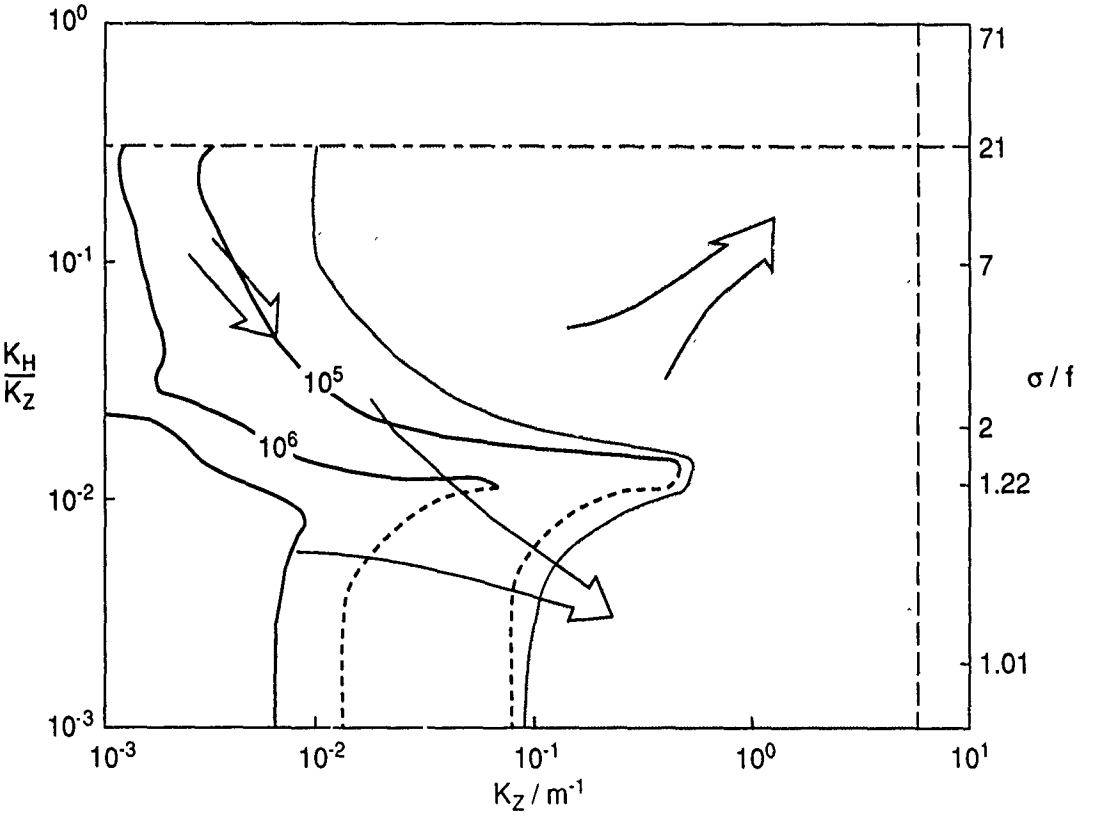


Fig. 3. Decay time for a 10% spike in the spectrum shown in Fig. 1 [from McComas, 1977]. The region where this time is shorter than a wave period is shaded. Dashed contours indicate that spike grows. Arrows indicate direction of energy flux due to nonlinear interactions: resonant theory [McComas, 1977] outside of shaded region, and eikonal [Henyey et al., 1986] inside.

between waves of different scales is due to the nonlinearity of the equations of motion, and thus becomes stronger at higher wavenumbers. If the rates are high enough, the waves interact so fast that they can no longer be described as freely propagating. Estimates of the energy transfer rates (Fig. 3) indicate that this limit is reached for waves with wavenumbers greater than 10–100 m^{-1}, as indicated by the shading. At wavenumbers smaller than this, waves can propagate for some time before they lose their energy to the rest of the wave field. Contours of this interaction time (Fig. 3) show a strong dependence on vertical wavenumber. Only the small wavenumbers (i.e., the low modes) have interaction times longer than 10 days (10^6 s).

Which Waves Can Propagate?

A "mean free path" L_{free} for horizontally propagating internal waves is estimated as the product of the interaction time τ_{iw} multiplied by the horizontal group velocity G_H. Approximate contours of L_{free} = 1000 km are shown in Fig. 4. Only a relatively small range of waves can propagate 1000 km. Another, perhaps more accurate estimate of L_{free} is given in E. Hirst's paper in this volume.

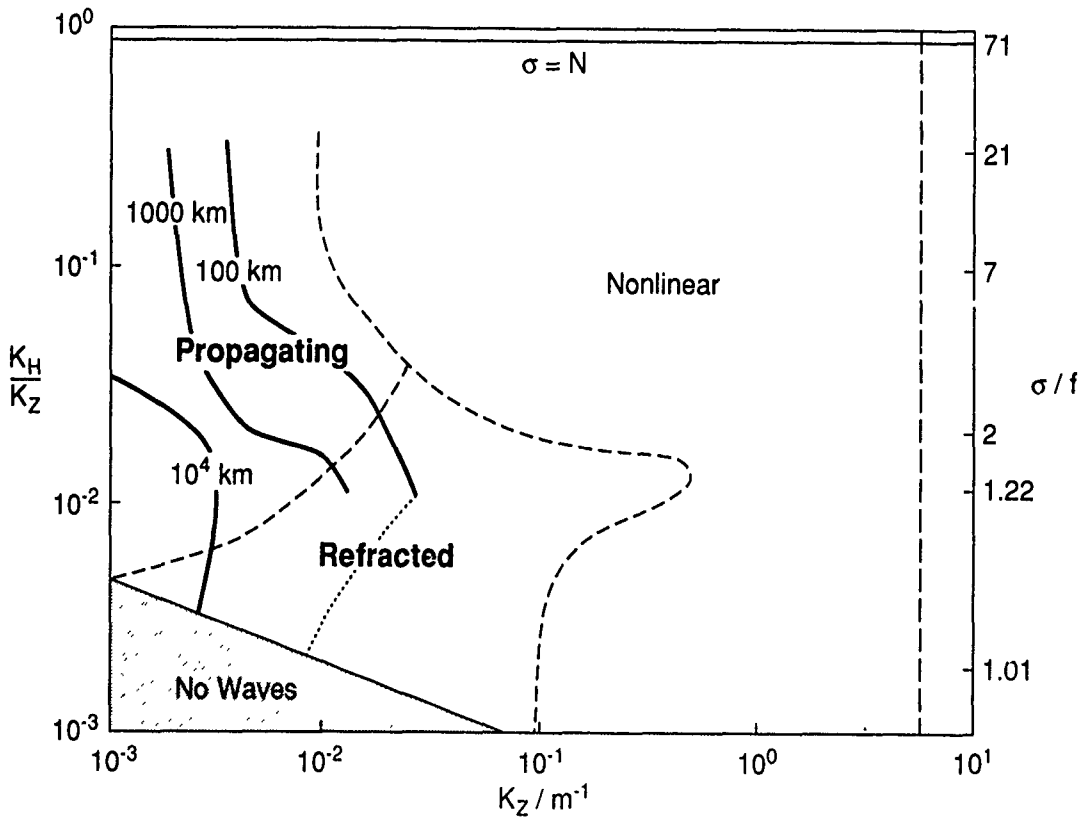


Fig. 4. Spectral location of Propagating Waves. High-wavenumber and high-frequency waves (Nonlinear) interact strongly and cannot propagate far. Many near-inertial waves (Refracted) are strongly refracted by internal waves and geostrophic currents and also cannot propagate far. The edge of this domain (dashed curve) is defined by U_{cr} = 0.2 m/s. Inertial, low-mode waves are prohibited by the β effect (No Waves). Propagation distance of low-mode, somewhat superinertial waves (Propagating), computed by hand from Figs. 2 and 3, is contoured. The end of the contours indicates inability to interpolate the figures. Dotted contour indicates that perturbations grow.

Figure 4 partitions the internal wave domain into various dynamical regions based on the above considerations. At high wavenumbers, τ_{iw} is short, and the waves are strongly nonlinear. They cannot propagate far. At lower wavenumbers, the waves become increasingly linear but still have small values of L_{free}. At low wavenumbers, but very close to the inertial frequency, wave propagation is limited by the β effect and interaction with the mesoscale flow. Only low modes with frequencies somewhat above f can propagate basinwide. I will call these "Propagating Waves."

THE ROLE OF PROPAGATING WAVES

Now consider a region of the ocean far from strong sources of internal waves such as wind, bottom topography, and internal tides generated on the shelves. Pick a region in the middle of a large ocean, over flat topography. The IWEX site might not be a bad choice! The chances are that the local sources are well below average. Nevertheless, the internal wave field is comparable to that found at more energetic sites. Wave interaction calculations and typical dissipation rates suggest that this wave field should lose most of its energy in 50–100 days, yet there is little evidence that this happens. Where does the energy come from to renew the wave field in this region? I suggest that it is supplied by the low-mode, large L_{free}, Propagating Waves discussed above. [2]

If, as indicated in Fig. 3, the major sources of internal waves are at low modes and low frequencies, much of their energy will be put into the Propagating Waves. These waves will disperse from the region where they were generated, releasing energy only slowly and over a wide region. Consequently, the Propagating Waves will flux energy away from the source regions. In a sink region, therefore, the directional spectrum of Propagating Waves will point toward the sources, much as surface-wave swell spectra point toward distant storms. Measurement of this spectrum should teach us about the distribution of sources.

Another consequence is that away from strong local sources the properties of the local internal wave field are determined by the rate at which the Propagating Waves give up energy. This rate is slow because these waves are nearly linear; thus resonant wave–wave interaction theory should be accurate. Other interactions such as scattering from topographic features may also be important. The local rate of energy loss from the Propagating Waves should equal, on average, the local rate of energy dissipation and thus govern the local rate of internal wave mixing. We should therefore observe a relationship between the field of Propagating Waves, the local internal wave spectrum, and the dissipation rate.

EVIDENCE FOR PROPAGATING WAVES

Existing evidence, although limited, supports these ideas.[3] During IWEX (Müller et al., 1978), detailed measurements were made of the internal wave field in the Sargasso Sea (27°N, 70°W), a region probably weak in local sources. During the limited time of the measurements (42 days), waves with frequencies between f and about $4f$ were significantly anisotropic, whereas those at higher frequencies were nearly isotropic. The anisotropic waves were all propagating toward the southeast, i.e., away from the North American continental shelf. Semidiurnal tidal energy, both anisotropic and dominated by low modes, looks like a Propagating Wave generated on the shelf. Interestingly, Müller et al. (1978) were so unconcerned with sources that the season of the measurements is not even mentioned!

Evidence for propagation of internal waves over long distances is provided by Fu's (1981) analysis of long-term abyssal moorings in the North Atlantic. The large-scale internal wave field was modeled by linearly propagating a Garrett-Munk-like wave field with no inertial peak northward from the equator. The model (Fig. 5) produces both near-inertial peaks and superinertial spectra which agree remarkably well with the abyssal data (given the simplicity of the model). This supports the idea that internal waves can propagate significant distances. The data are insufficient to determine the wavenumbers of the waves that propagate. The model worked poorly in the upper ocean, in the thermocline, and over rough topography. Fu (1981) models these features, with some success, as due to local sources of near-inertial waves.

Lai and Sanford (1986) observed the local generation of low modes after the passage of a hurricane near the New England continental slope. First the lowest mode was observed and then higher modes; the timing was consistent with generation at a site 80 km away.

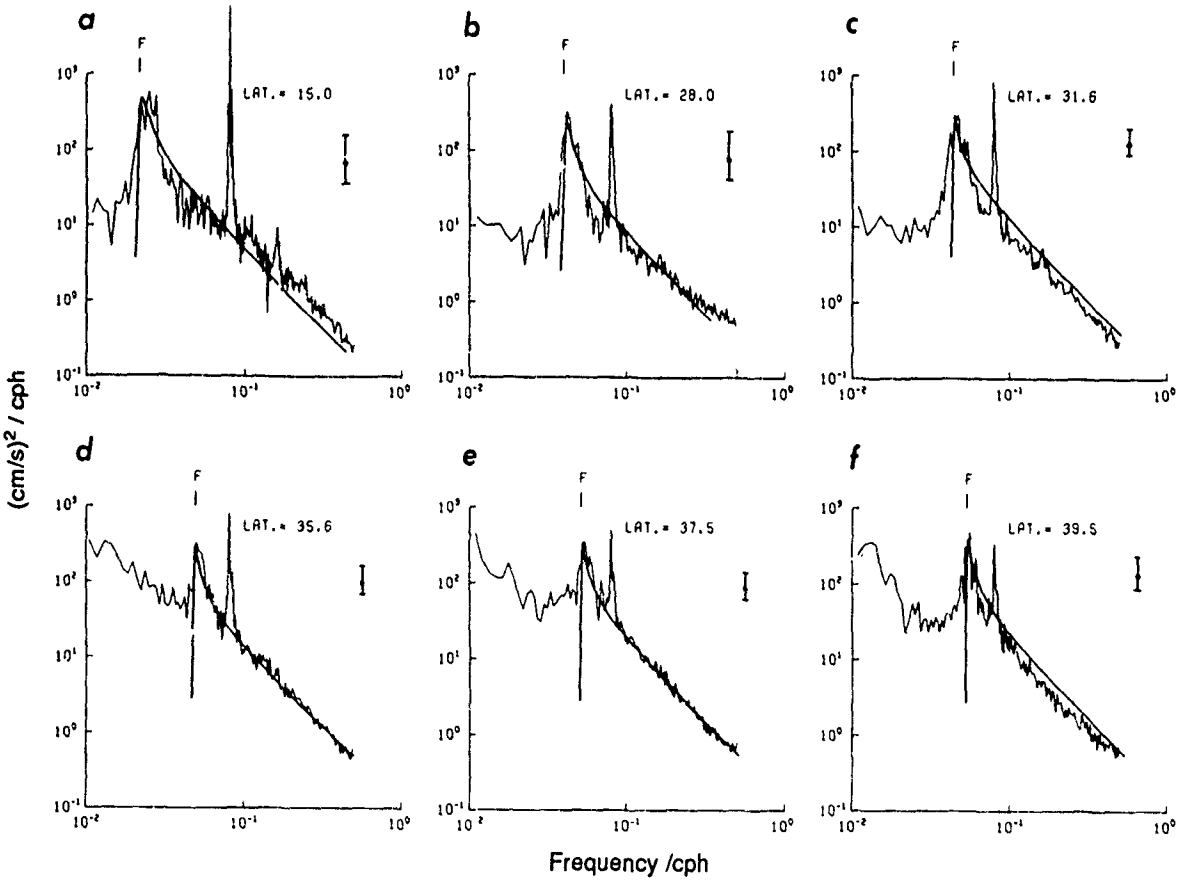


Fig. 5. Model (smooth curves) and observed spectra of horizontal kinetic energy at 4000 m measured during long term moorings in the North Atlantic. Inertial frequency and latitude are marked. [from Fu, 1981]

[2] I believe this idea originates with the unpublished Johnson and Cox manuscript.

[3] Perhaps my knowledge, rather than the evidence, is limited.

WIND FORCING

We do not know which of the possible sources of internal wave energy is most important. Rather than attempting to understand them all, we might look at just one and ask whether the data are consistent with its being the primary source. I propose that wind forcing is a primary candidate, although tidal forcing is another possibility. Internal waves in the thermocline usually show a dominance of downward propagating, near-inertial energy. Examples include data taken by Leaman and Sanford (1975), D'Asaro (1984), and Müller et al. (1978). Atmospheric forcing is the obvious source of this asymmetry.

Mixed Layer Inertial Currents

Perhaps the best understood source of internal waves is excitation of mixed-layer inertial motions by the wind. Because atmospheric wind fluctuations are typically both much larger than a Rossby radius in the ocean and move faster than the lowest baroclinic phase speed, the wind stress tends to excite mostly near-inertial waves (Gill, 1984). It has been apparent since the first near-surface current meter records that strong mixed-layer inertial currents are generated by fluctuating wind stresses (Webster, 1968). A simple slab mixed-layer model does a decent job of simulating this process under most conditions, as was first demonstrated by Pollard and Millard (1970). Mixed-layer inertial currents and the energy transferred from the wind to these currents can be estimated from wind data using this model.

How well can we predict the flux of energy from the wind to mixed-layer inertial currents? An example of such a calculation using real wind measurements is shown in Fig. 6. The required

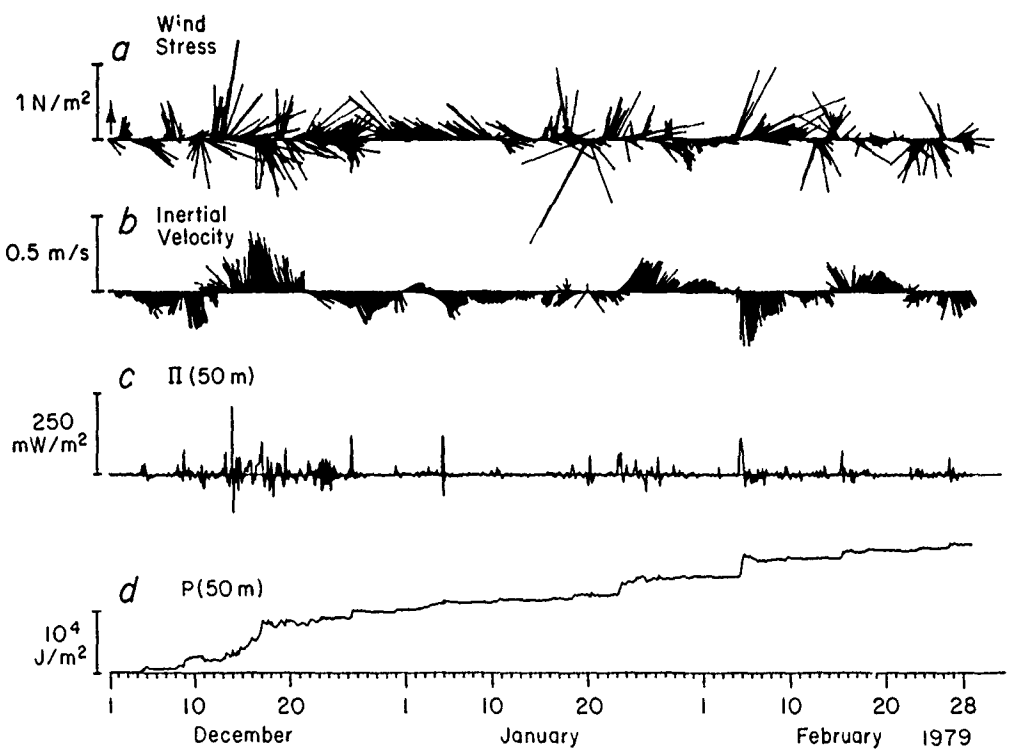


Fig. 6. Inertial-energy flux computed for NOAA Data Buoy Center buoy 46006 near 41°N, 140°W. (a) Wind stress computed with Garratt (1977) drag law. (b) Mixed-layer inertial velocity computed using 50 m mixed-layer depth. (c) Energy flux to mixed-layer inertial motions. (d) Integrated energy flux.

data are the wind stress (top panel), a drag law, and the mixed layer depth, here assumed to be 50 m. The resulting energy flux (bottom two panels) is highly intermittent and is dominated by a few storms. For 10 years of data from Ocean Weather Station P, for example, the kurtosis of the energy flux was estimated at 360 and the average November energy flux was increased 60% by a single storm (D'Asaro, 1985b). Storms are clearly important, and wind stress data are required at least every few hours. Operational weather products probably do not give these winds with sufficient accuracy over the ocean, (Thomson, 1983). Furthermore, the drag coefficient can vary by a factor of 2 during storms (Geernaert, 1988), resulting in significant changes in the computed inertial currents (D'Asaro, 1985a). Because of these factors, accurate predictions of mixed-layer inertial currents can probably not be made using operational weather products until predictions of winds over the ocean are significantly improved.

Figure 7 shows the 10 year average annual cycle of energy flux to mixed-layer inertial currents computed using a simple slab model and climatological mixed-layer depth (D'Asaro, 1985b). The flux is large in the fall and winter and small in the summer because of the stronger wind in the winter. It is largest in the early fall because of the thinner mixed layers. Although the proper calculations have not been made, the large-scale spatial pattern is probably quite similar to that of the wind stress (Fig. 8). The important point is that clear seasonal and large-scale geographical signals exist. We do not know how well operational products can predict these signals, although prediction of seasonal and regional averages is likely to be more accurate than that of local statistics.

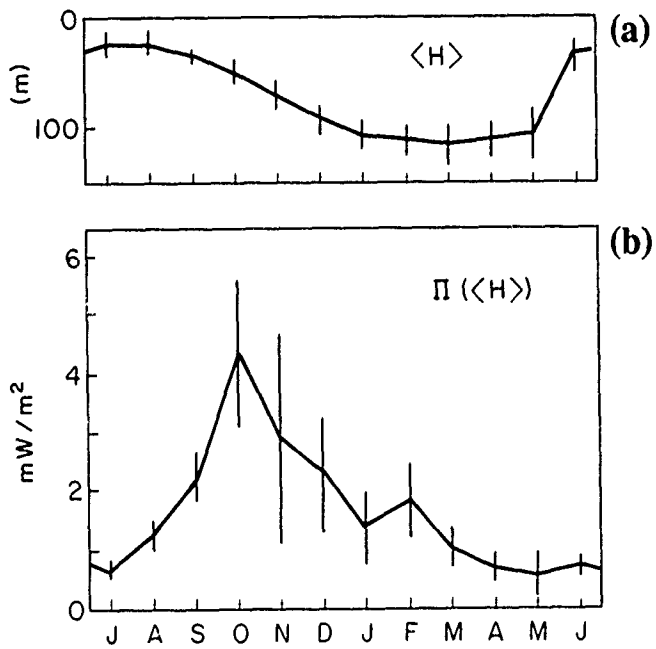


Fig. 7. Annual cycle of inertial energy flux at OWS-P (50°N, 145°W). (a) Annual cycle of mixed-layer depth. (b) Annual cycle of energy flux to mixed-layer inertial currents computed as in Fig. 6 but using mixed-layer depth from (a).

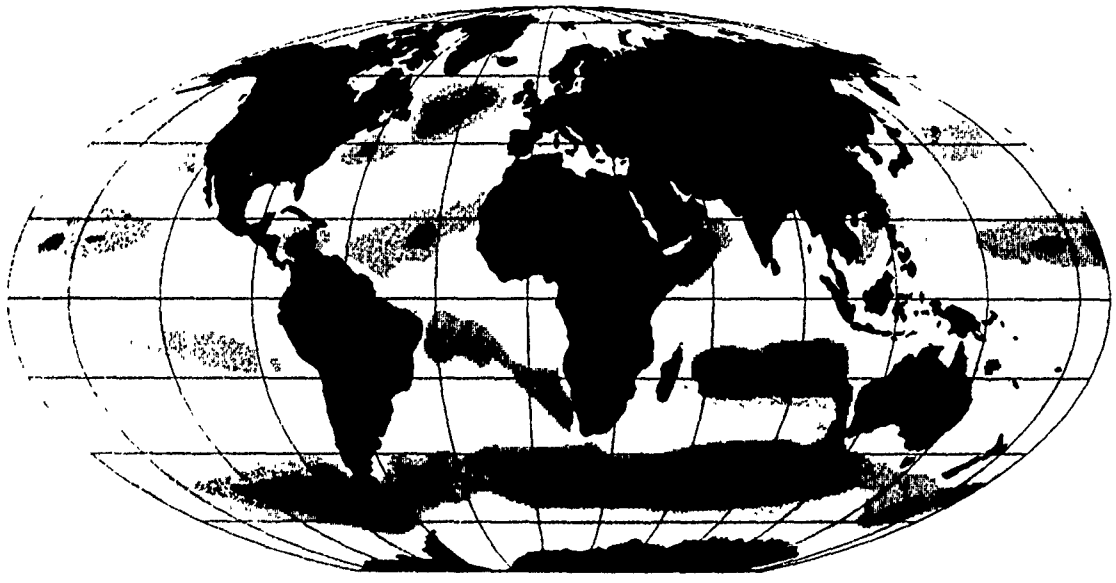


Fig. 8. Annual wind stress magnitude (Hellerman and Rosenstein, 1983). Contour level is 0.1 Pa. [from Briscoe, 1984]

Low-Mode, Near-Inertial Waves

Most of the energy in mixed-layer inertial currents probably propagates into the stratified ocean in the form of local and propagating near-inertial internal waves. Figures 9–12 show a simulation of the ocean's response to a single storm. The forcing, latitude, and stratification are tuned to match the conditions observed during early October 1987 as part of the OCEAN STORMS experiment. The model is a two-dimensional variant of Price's (1983) hurricane model, with 21 layers, full nonlinearity, hydrostatic dynamics, a β plane, and a vertical viscosity of 2×10^{-4} m^2/s, which improves its fit with the data. It is forced in the mixed layer by a Gaussian storm centered at $y = 0$ (47.5°N).

The response of the model ocean to this storm is almost entirely near-inertial. Figure 9 shows depth-time contours of the inertial amplitude as a function of depth and time in the upper kilometer. The model storm forces mixed-layer inertial currents of about 40 cm/s. For the first 10 days after the storm, the inertial energy stays in the mixed layer while the β effect increases the north-south wavenumber of the currents as described by D'Asaro (1989). There is some slight penetration of energy into the thermocline owing to the imposed vertical viscosity. Eventually, the horizontal wavenumber becomes large enough for linear wave propagation to occur. Between days 300 and 320, almost all the mixed layer energy is transferred to the upper thermocline and to low-mode, near-inertial waves. Contours of inertial speed in the latitude-time plane (Figs. 10–12) show the distinct separation and propagation of modes 1 (starting on day 306) and 2 (starting on day 320). Both modes are seen at the surface (Fig. 10) and in the upper thermocline (Fig. 11), but only mode 2 is seen at 1000 m (Fig. 12) because this is near the zero crossing of mode 1. Similar results were seen in simulations by Gill (1984).

The low modes rapidly propagate southward away from the source region. In the simulation, waves in the thermocline 2000 km south of the storm have speeds of about 1 cm/s and verti-

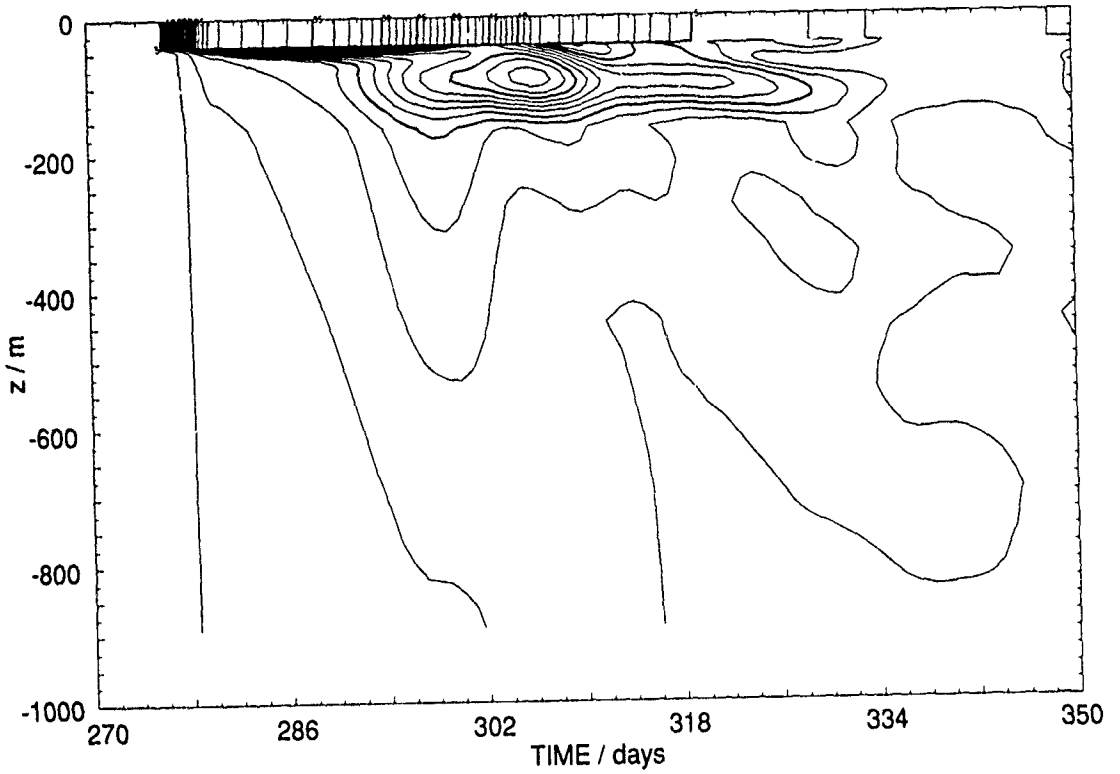


Fig. 9. Inertial speed as a function of depth and time directly beneath the storm ($y = 0$) for simulation of ocean response to a strong storm. Contour interval is 1 cm/s.

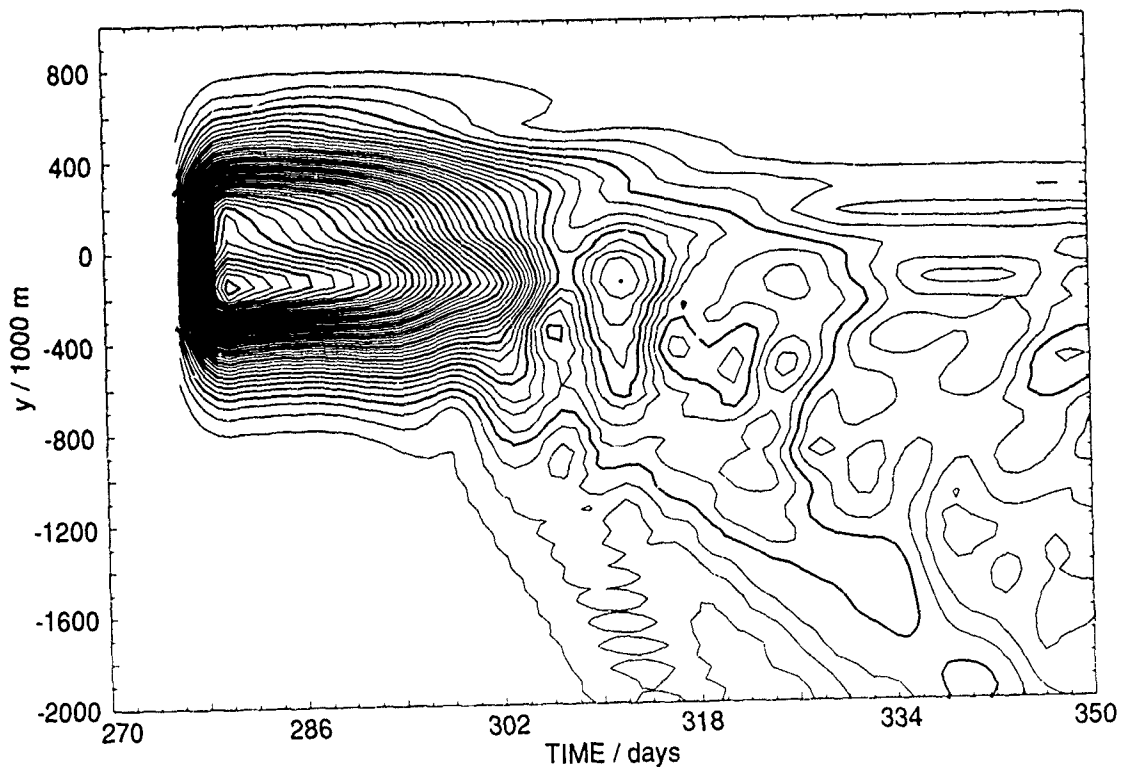


Fig. 10. Inertial speed at the surface as a function of meridional distance and time. Location and size of storm are clearly seen as is propagation of energy to the south. Contour interval is 1 cm/s.

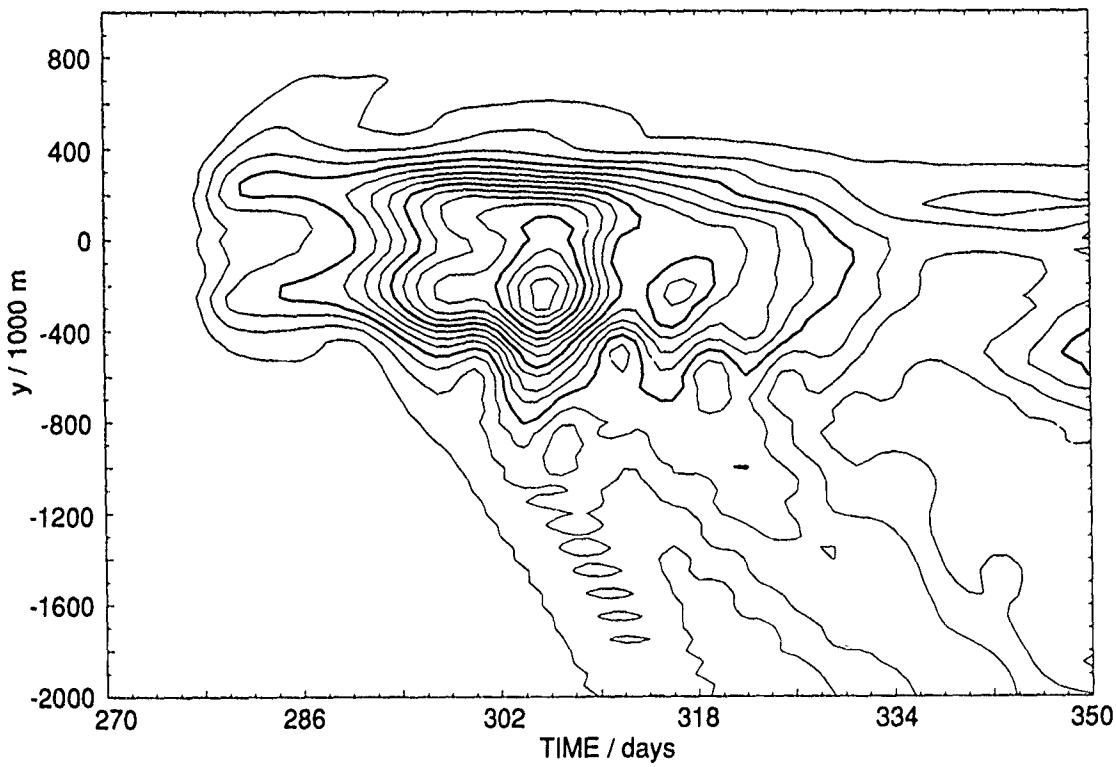


Fig. 11. Inertial speed at 100 m as a function of meridional distance and time. Comparison with Fig. 10 shows propagation of energy from the mixed layer starting near day 286 and southward propagating modes. Contour interval is 1 cm/s.

cal displacements of about 1 m for at least 20 days. These waves have distinct properties. They are low mode and have a meridional wavelength near 100 km. Their frequency is about $\beta\Delta y$, where Δy is the meridional distance from the storm, which equals $1.3f$ near 30°N. A fully three-dimensional calculation would show a zonal wavelength near U/f, where U is the advection speed of the storm. Although wave–wave interactions are not included in the model, these scales are clearly in the range of the "Propagating Waves" shown in Fig. 4. These waves should therefore not be strongly attenuated by interaction with the rest of the wave field. Their signatures are distinct and should be easily measured, given the proper instrumentation

In this simulation, the energy put into the inertial currents, averaged over the 30 day length of the ocean response, is about 10^{-3} W/m^2. The average wintertime fluxes estimated in Fig. 7 are several times this. Furthermore, because the propagating waves always go equatorward, the total southward flux is the result of all forcing north of a given location. It would not be unreasonable, therefore, to expect the peak wintertime amplitude of Propagating Waves to be several times that shown in Figs. 9–12. A far weaker signal should be expected in the summer.

SOME SUGGESTIONS FOR RESEARCH

How can we detect Propagating Waves generated by storms? The signal is so weak that the waves from a single storm will probably not be obvious unless the storm is very strong (e.g., Lai and Sanford, 1986). Measurements will therefore have to be designed to detect the signal. The key appears to be averaging either in the vertical, to isolate the low-mode structure, or in the

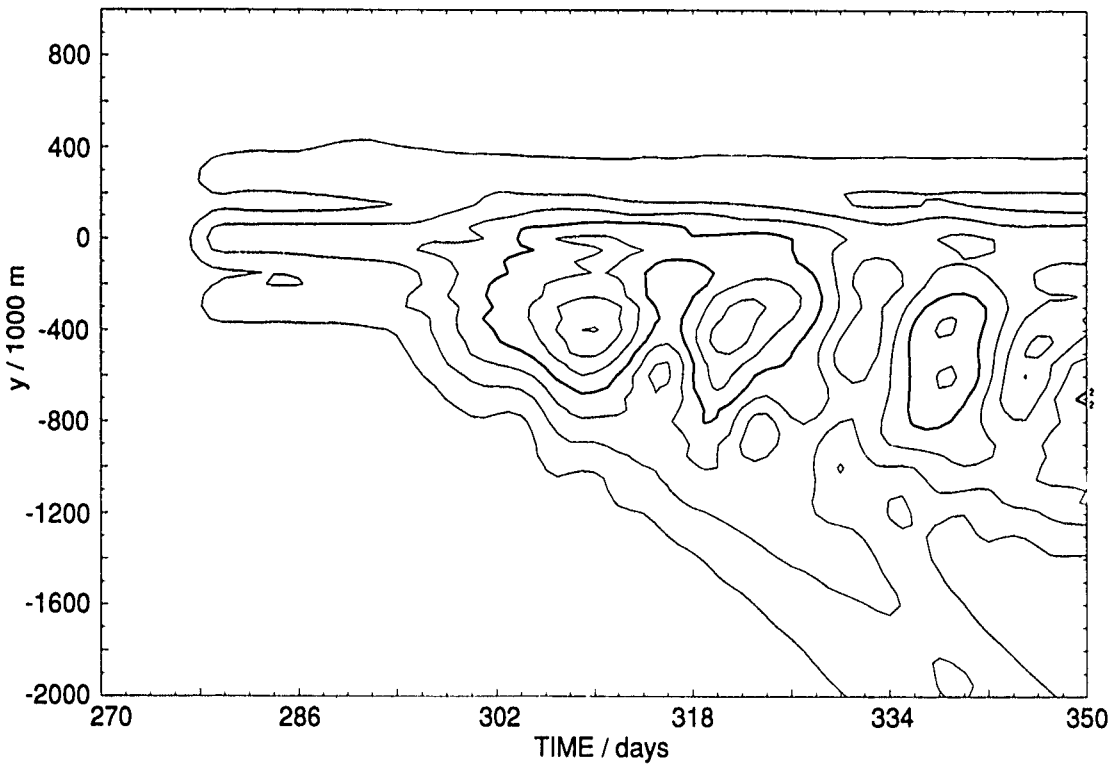


Fig. 12. Inertial speed at 1000 m as a function of meridional distance and time. Contour interval is 0.5 cm/s.

horizontal, to isolate the large horizontal scales, as well as filtering in time to isolate the distinct frequency. One approach would be to use integrating temperature sensors to filter out the high modes. Vertical or horizontal acoustic propagation is another promising technique. Given the strong seasonal cycle expected of wind forcing, a long time series of directional measurements of the low-mode waves seems like the most promising sampling scheme. A subtropical location south of the wintertime generation region that is subject to Propagating Waves generated by late summer hurricanes and a large internal tide source would seem ideal.

We also need better theoretical descriptions of the processes controlling propagation of low modes. Can we really expect to see these low mode signals? This depends on how energy gets out of them. Calculations of resonant wave–wave interaction, which should be accurate for these weakly nonlinear waves, should be reexamined to include the effects of directionality, depth, and path-varying stratification. Topographic features and mesoscale eddies may scatter energy from these propagating waves into the local internal wave field. We don't know how to estimate these effects at present.

TOWARD A GLOBAL MODEL OF THE INTERNAL WAVE FIELD

Some oceanographers consider small-scale processes a source of noise and a sink for money. But they still look to the small-scale research community for estimates of the magnitude and variability of small-scale heat, salt, and momentum transport. Although we have little to offer for momentum transport, I believe we can envision how a predictive model of diapycnal mixing due to internal waves could be constructed.

First, we need to quantify at least some of the sources of internal waves. The time and space variability of the sources is a key factor because it will allow us to distinguish various sources in data. As mentioned previously, I think we have a good start in quantifying the generation by wind. Tidal generation should also be investigated. Initially, it is probably not necessary to quantify all the possible sources since it will eventually become apparent if we are missing a major one. It certainly would be enlightening, for example, to see what an internal wave field generated only by the wind would look like!

Second, we need to understand how the energy from these sources is distributed between the local wave field and the Propagating Waves. For wind-generated energy in the absence of mesoscale eddies, we can probably estimate this. The effect of mesoscale motions is not well studied.

Third, we need to know how far the Propagating Waves travel and how they give up their energy. It seems likely, as argued here, that only a small range of low-mode waves can propagate significant distances. If true, this greatly simplifies the problem. There is ample room for both observation and theory here. Our ignorance appears to be due to inattention, not hard technical problems, so advances should be rapid.

Fourth, we need to know how the energy supplied to the local wave field by local sources and by the Propagating Waves cascades through the local wave field to mixing. Great progress is being made here, but we are not yet finished.

Acknowledgments

Thanks to Peter Müller, Mel Briscoe, Frank Henyey, and Terry Ewart for stimulating these ideas. The Winter Workshop provided a great excuse to work them out and write them down. This work was supported by ONR contract N00014-89-C-0111 and ONR grant N00014-90-J-1104.

REFERENCES

Bell, T. H., Jr., 1975: Topographically generated internal waves in the open ocean. *J. Geophys. Res.*, **80**, 320–327.

Briscoe, M. G., 1984: The monthly variability of upper-ocean internal wave energy: A progress report on the correspondence with wind stress. In *Internal Gravity Waves and Small-Scale Turbulence, Proceedings, Hawaiian Winter Workshop*, P. Müller and R. Pujalet, Eds., 129–150, Hawaii Institute of Geophysics, Honolulu.

D'Asaro, E. A., 1984: Wind forced internal waves in the North Pacific and Sargasso Sea. *J. Phys. Oceanogr.*, **14**, 781–794.

D'Asaro, E. A., 1985a: Upper ocean temperature structure, inertial currents, and Richardson numbers observed during strong meteorological forcing. *J. Phys. Oceanogr.*, **15**, 943–962.

D'Asaro, E. A., 1985b: The energy flux from the wind to near-inertial motions in the surface mixed layer. *J. Phys. Oceanogr.*, **15**, 1043–1059.

D'Asaro, E. A., 1989: The decay of wind-forced mixed layer inertial oscillations due to the β effect. *J. Geophys. Res.*, **94**, 2045–2056.

Fu, L.-L., 1981: Observations and models of inertial waves in the deep ocean. *Rev Geophys*, **19**, 141–170.

Garratt, J.R., 1977: Review of drag coefficients over oceans and continents. *Mon. Wea. Rev.*, **105**, 915–929.

Garrett, C.J.R., and W.H. Munk, 1975: Space-time scales of internal waves: A progress report. *J. Geophys Res.*, **80**, 291–297.

Garrett, C.J.R., and W.H. Munk, 1979: Internal waves in the ocean. *Ann. Rev. Fluid Mech.*, **11**, 339–369.

Geernaert, G.L., 1988: Measurements of the angle between the wind vector and wind stress vector in the surface layer over the North Sea. *J. Geophys. Res.*, **93**, 8215–8220.

Gill, A.E., 1984: On the behavior of internal waves in the wakes of storms. *J. Phys. Oceanogr.*, **14**, 1129–1151.

Hellerman, S., and M. Rosenstein, 1983: Normal monthly wind stress over the world ocean with error estimates. *J. Phys. Oceanogr.*, **13**, 1093–1104.

Henyey, F.S., J. Wright, and S.M. Flatté, 1986: Energy and action flow through the internal wave field: An eikonal approach. *J. Geophys. Res.*, **91**, 8487–8495.

Kunze, E., 1985: Near-inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.*, **15**, 544–565.

Kunze, E., and T.B. Sanford, 1984. Observations of near-inertial waves in a front. *J. Phys. Oceanogr.*, **14**, 566–581.

Lai, D.Y., and T.B. Sanford, 1986: Observations of hurricane-generated, near-inertial slope modes. *J. Phys. Oceanogr.*, **16**, 657–666.

Leaman, K.D., and T.B. Sanford, 1975: Vertical energy propagation of inertial waves: A vector spectral analysis of velocity profiles. *J. Geophys. Res.*, **80**, 1975–1978.

Levine, M.D., C.A. Paulson, M.G. Briscoe, R.A. Weller, and H. Peters, 1983: Internal waves in JASIN. *Phil. Trans. Roy. Soc. London, Ser. A*, **308**, 389–405.

Levine, M.D., C.A. Paulson, and J.H. Morison, 1985: Internal waves in the Arctic Ocean: Comparison with lower-latitude observations. *J. Phys. Oceanogr.*, **15**, 800–809.

McComas, C.H., 1977: Equilibrium mechanisms within the oceanic internal wave field. *J. Phys Oceanogr.*, **7**, 836–845.

Müller, P., D.J. Olbers, and J. Willebrand, 1978: The IWEX spectrum. *J. Geophys. Res.*, **83**, 479–500.

Müller, P., G. Holloway, F. Henyey, and N. Pomphrey, 1986: Nonlinear interactions among internal gravity waves. *Rev. Geophys.*, **24**, 493–536.

Padman, L., and T.M. Dillon, 1991: Turbulent mixing near the Yermak plateau during the Coordinated Eastern Arctic Experiment. *Geophys. Res.*, **96**, 4769–4782.

Pollard, R.T., and R.C. Millard, 1970: Comparison between observed and simulated wind-generated inertial oscillations. *Deep-Sea Res.*, **17**, 813–821.

Price, J.F., 1983: Internal wave wake of a moving storm. Part I: Scales, energy budget and observations. *J. Phys. Oceanogr.*, **13**, 949–965.

Thomson, R.E., 1983: A comparison between computed and measured oceanic winds near the British Columbia coast. *J. Geophys. Res.*, **88**, 2675–2683.

Watson, K.M., 1985: Interaction between internal waves and mesoscale flow. *J. Phys Oceanogr.*, **15**, 1296–1311.

Webster, F., 1968: Observations of inertial-period motions in the deep sea. *Rev. Geophys.*, **6**, 473–490.

Wunsch, C., and S. Webb, 1979: The climatology of deep ocean internal waves. *J. Phys. Oceanogr.*, **9**, 235–243.

OBSERVING OCEANIC INTERNAL WAVES: What have we learned? What can we learn?

Murray D. Levine

College of Oceanography
Oregon State University, Corvallis, OR 97331

ABSTRACT

An incomplete and biased assessment is made of the current state of our understanding of oceanic internal waves. The purpose of this presentation is to remind researchers of some inconsistencies that exist between the explanation and description of certain features of the internal wave field. The goal is to renew interest and guide discussion toward certain observations that may ultimately lead to an improved understanding of the dynamics controlling internal waves.

INTRODUCTION

The oceanic internal wave field has been measured by many researchers for decades. What have we learned? If condensed into a single short answer, it is that the internal wave field is remarkably constant in time and space. If one puts instruments anywhere into the ocean, then it is highly probable to expect that the observed fluctuations will follow certain characteristic spectral shapes and coherence structures. This realization led to the development of an empirical description of the wave field known as the Garrett-Munk model (Garrett and Munk, 1972, 1975; hereafter referred to as GM). This model provides a kinematically consistent framework for comparing different types of observations that may be made at different times and locations. It is actually because this first-order description works so well that it has been difficult to answer basic questions about the dynamics of the wave field, such as what are the important sources and sinks of internal waves. It is difficult to tell from where a particular wave came, and to where it is going, when its identity is obscured by a surrounding random sea of oscillations.

Despite the success of the Garrett-Munk formulation, there are notable omissions from this steady-state, climatological description. For example, waves in the near-inertial frequency band vary dramatically in space and time, and therefore are not easily characterized by a steady-state description. Much effort has gone into studying the transfer of energy from the wind into the near-inertial waves of the upper ocean. The internal tide is also not included in the GM model. The internal tide is omnipresent throughout the ocean and often represents a significant fraction of the total internal wave energy. Deviations from GM are also found near topography; either the topography is acting as a source of waves or the interactions are causing a

perturbation in the universal wave field. In unique environments, such as beneath the Arctic ice pack, dramatic differences in the overall wave energy and spectral shape have been found. Study of atypical systems may provide clues toward understanding the more typical ocean.

The second question in the title: "What can we learn?" really depends upon "What do we want to learn?" The answer will vary depending upon who asks the question; investigators in different specialties will be interested in different aspects of oceanic internal waves. For example, some acousticians may be interested in the internal wave correlation functions that describe the fluctuations of sound speed. Other acousticians may want to know higher order statistics, such as a statistical description of the occurrence of sharp gradients in velocity or sound speed. The specific internal wave information that is useful will depend on the particular acoustic study. High-frequency transmission experiments would be affected by a different scale internal wave than an acoustic tomography experiment.

Researchers studying turbulence and mixing are primarily interested in internal waves for their role in providing small-scale vertical shear. The ability to predict how the waves affect the time/space distribution of the Richardson number would be valuable. Can estimates of turbulent dissipation be made from observations of the internal wave field alone, as has been suggested (e.g., Gregg, 1989)? To determine if this is feasible, an improved description of the space/time variation of the internal wave field is needed.

Some investigators are interested in the remote sensing of internal waves. The initial emphasis in this field has been on detecting solitons and large-amplitude waves that produce a significant surface expression. Attention is also focused on processes that generate these signals, usually interaction of the tide with topography. Clearly a GM description of the wave field is less important to this group.

And there are those investigators that follow the approach of the mainstream physicist and try to understand the details of the nonlinear energy transfer among the waves. Theoreticians need, at the very least, to have an accurate description of the wave field in time/space or frequency/wavenumber in order to insure relevance to the oceanic environment. Estimates of higher-order statistical quantities would undoubtedly also be of interest.

In each of these examples a different aspect of the internal wave field is emphasized depending upon the objectives of the researcher. Certainly there are overlaps, but the labeling of certain advances in internal waves as *important* may depend on who is doing the labeling.

So, where do we go from here? Two possible paths to follow are new measurements and new data, or old observations and new analysis. There are a variety of new

instrumental developments that will be able to measure new quantities of the wave field. Better Doppler sonars, improved neutrally buoyant drifters, acoustic travel-time experiments, and longer time-series observations may provide views of the wave field from different perspectives and lead to an improved description and understanding. However, the remarks here will be confined to the second track--are there different ways to look at the old data? Given the large investment that has been made in obtaining these data it makes sense to spend some effort in exploiting the data to the fullest. Armed with new ideas, a reexamination of our data banks may reveal some interesting results, or possibly, suggest new observations that would be most fruitful. As a step toward this goal, a few topics are reexamined below to provide motivation for continuing the search for new types of analysis and new ways of looking at the data.

INTERNAL WAVE CONTINUUM

One approach to describe the internal wave field is to assume it is composed of a sum of waves from a smooth continuum of frequencies and wavenumbers. This is the basis of the GM model and provides a first-order statistical description of the wave field. This framework allows us to view the wave field in a variety of different spaces. The energy spectrum can be displayed as functions of vertical wavenumber, horizontal wavenumber or frequency. Different vantage points provide different revelations. For example, the spectrum of vertical shear ($\partial u/\partial z$) is plotted as a function of frequency ω and vertical wavenumber β (Figure 1a). The axes are logarithmically scaled, as are the contour levels. The GM shear spectrum is separable in ω-β space and increases with β before leveling off at $\beta > \beta_*$. An *ad hoc* constant upper wavenumber cutoff β_c has been included. One gets an entirely different impression of the spectrum looking at the so-called volume preserving version (Figure 1b). Here equal volumes contribute equally to the total variance. It is clear that more of the shear variance comes from frequencies near f and wavenumbers near β_c. Hence, if one is using the GM model to estimate shear for some mixing parameterization (e.g., Gregg, 1989), then the *ad hoc* choice of β_c is important. Maybe β_c is a function of frequency? Maybe β_c varies in a more complicated way (Duda and Cox, 1989)? How do uncertainties in β_c translate into uncertainties in predicting mixing from models?

Critical layer absorption is a process that is thought by some to play an important role in internal wave dynamics. A critical layer occurs at the depth where the horizontal phase speed of the wave equals the horizontal velocity of the mean flow. To get an idea of the likelihood of this occurring in the internal wave field, the magnitude of the phase speed is plotted in the same region of ω-β space as the shear spectra (Figure 1c). The phase speed is simply ω/α where α is the horizontal wavenumber and is determined from the linear dispersion relationship. Waves with lower phase speeds are more likely to encounter mean flows of comparable speed and be absorbed in critical layers. Considering Figures 1b and 1c together, it seems that the region of ω-

instrumental developments that will be able to measure new quantities of the wave field. Better Doppler sonars, improved neutrally buoyant drifters, acoustic travel-time experiments, and longer time-series observations may provide views of the wave field from different perspectives and lead to an improved description and understanding. However, the remarks here will be confined to the second track--are there different ways to look at the old data? Given the large investment that has been made in obtaining these data it makes sense to spend some effort in exploiting the data to the fullest. Armed with new ideas, a reexamination of our data banks may reveal some interesting results, or possibly, suggest new observations that would be most fruitful. As a step toward this goal, a few topics are reexamined below to provide motivation for continuing the search for new types of analysis and new ways of looking at the data.

INTERNAL WAVE CONTINUUM

One approach to describe the internal wave field is to assume it is composed of a sum of waves from a smooth continuum of frequencies and wavenumbers. This is the basis of the GM model and provides a first-order statistical description of the wave field. This framework allows us to view the wave field in a variety of different spaces. The energy spectrum can be displayed as functions of vertical wavenumber, horizontal wavenumber or frequency. Different vantage points provide different revelations. For example, the spectrum of vertical shear ($\partial u/\partial z$) is plotted as a function of frequency ω and vertical wavenumber β (Figure 1a). The axes are logarithmically scaled, as are the contour levels. The GM shear spectrum is separable in ω-β space and increases with β before leveling off at $\beta > \beta_*$. An *ad hoc* constant upper wavenumber cutoff β_c has been included. One gets an entirely different impression of the spectrum looking at the so-called volume preserving version (Figure 1b). Here equal volumes contribute equally to the total variance. It is clear that more of the shear variance comes from frequencies near f and wavenumbers near β_c. Hence, if one is using the GM model to estimate shear for some mixing parameterization (e.g., Gregg, 1989), then the *ad hoc* choice of β_c is important. Maybe β_c is a function of frequency? Maybe β_c varies in a more complicated way (Duda and Cox, 1989)? How do uncertainties in β_c translate into uncertainties in predicting mixing from models?

Critical layer absorption is a process that is thought by some to play an important role in internal wave dynamics. A critical layer occurs at the depth where the horizontal phase speed of the wave equals the horizontal velocity of the mean flow. To get an idea of the likelihood of this occurring in the internal wave field, the magnitude of the phase speed is plotted in the same region of ω-β space as the shear spectra (Figure 1c). The phase speed is simply ω/α where α is the horizontal wavenumber and is determined from the linear dispersion relationship. Waves with lower phase speeds are more likely to encounter mean flows of comparable speed and be absorbed in critical layers. Considering Figures 1b and 1c together, it seems that the region of ω-

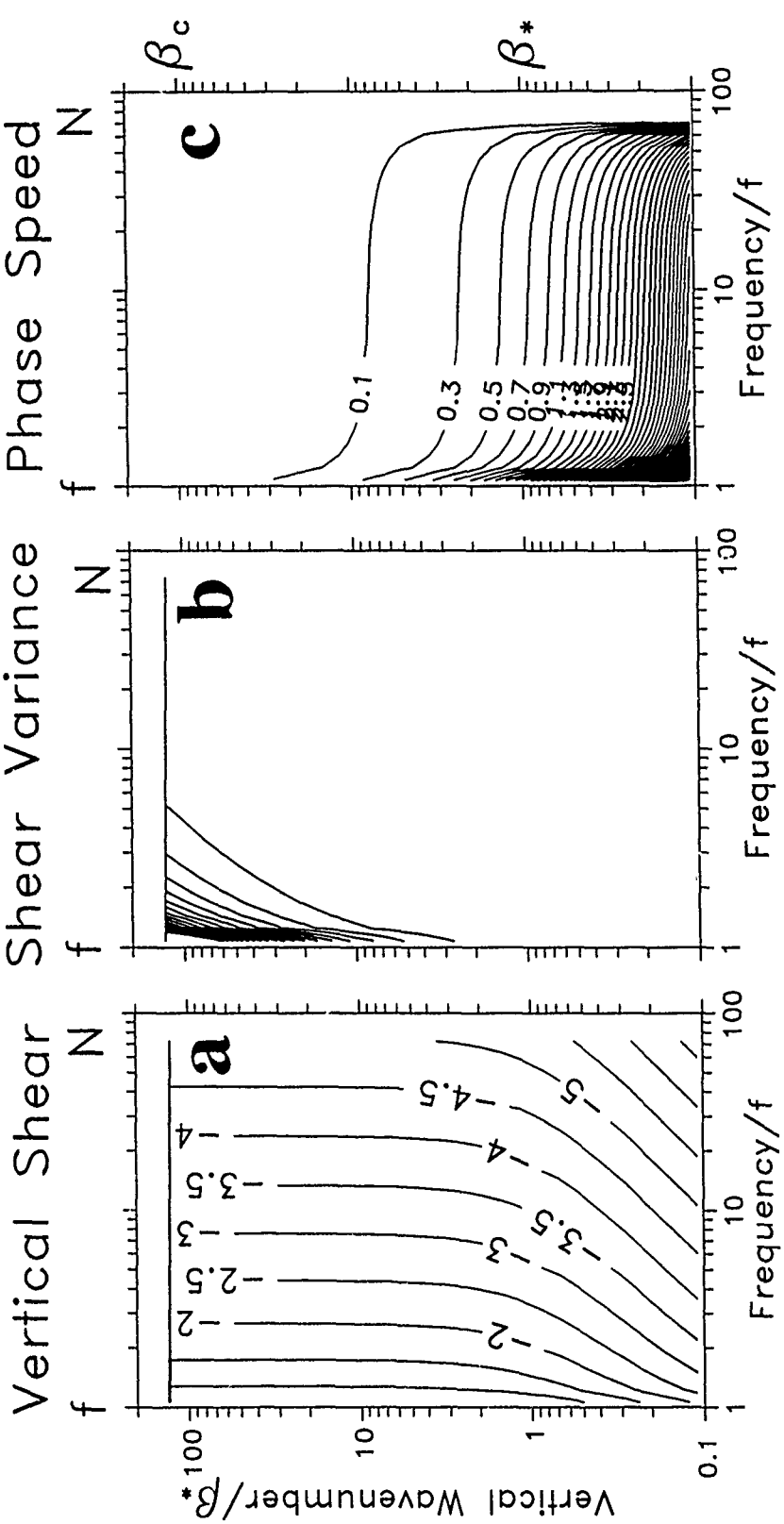


Figure 1. Vertical shear spectrum ($\partial u/\partial z$) given by the Garrett-Munk model is plotted as a function of frequency and vertical wavenumber contoured in units of spectral density (a) and in variance preserving form (b). The horizontal phase speed (ω/α) is contoured in units of meters per second in (c).

β space with the highest shear variance is the most susceptible to critical layer absorption. This realization casts some doubt on the reliability of parameterizing the vertical shear from the GM model--variations of the background shear may lead to variations in wave induced vertical shear. Even if critical layer absorption may not be important in the total energy budget of internal waves, it may be significant in the waves that contain most of the shear.

INTERNAL TIDE

The internal tide is a ubiquitous feature in the world ocean. Its existence has been documented from early in this century. While not included in the GM formulation, the energetic internal tide is important to some investigators, e.g. acousticians, because of its significant contribution to the vertical displacement and straining of the stratified ocean. However, does the internal tide play a role in the dynamics of the wave field?

The internal tide is generated by the interaction of the barotropic tide with topography. The continental shelf break is believed to be a significant source. However, the sea floor as a whole, although much less steep, may actually provide significantly more flux of internal tidal energy into the ocean than the entire continental shelf region (Baines, 1982). There are varying opinions as to whether the internal tide is a significant source of energy to the internal wave continuum. The tide is an attractive candidate for an internal wave source because it is widespread in space and steady in time. This fact is consistent with the observed universality of the internal wave continuum. Also, estimates of the energy flux **into** the internal tide from the sea floor of order 1 mW m^{-2} are comparable to that needed to maintain the entire internal wave field (Bell, 1975).

But can nonlinear interactions transfer energy out of tidal frequencies at a significant rate? Or does the internal tide propagate passively through the internal wave continuum? One theoretical estimate from weak-interaction theory suggests that the tide could not be a significant source (Olbers and Pomphrey, 1981), however this study was not comprehensive. Perhaps including non-resonant interactions will yield a different result.

It has also been argued that if the tide were a significant source, then there should be a different internal wave spectrum in non-tidal seas. The Mediterranean Sea would provide an interesting testing ground since there are regions where the tide is weak. The author has not yet located convincing evidence (although some may exist) that compares observations there to the GM spectrum. The Black Sea is another nontidal sea; some evidence (albeit limited) indicates that the spectral levels are highly variable and well below GM (Ivanov and Sereonyanyy, 1982).

There is some suggestion from observations of a link between internal tide and the continuum (Figure 2). Time series of velocity from the JASIN experiment near

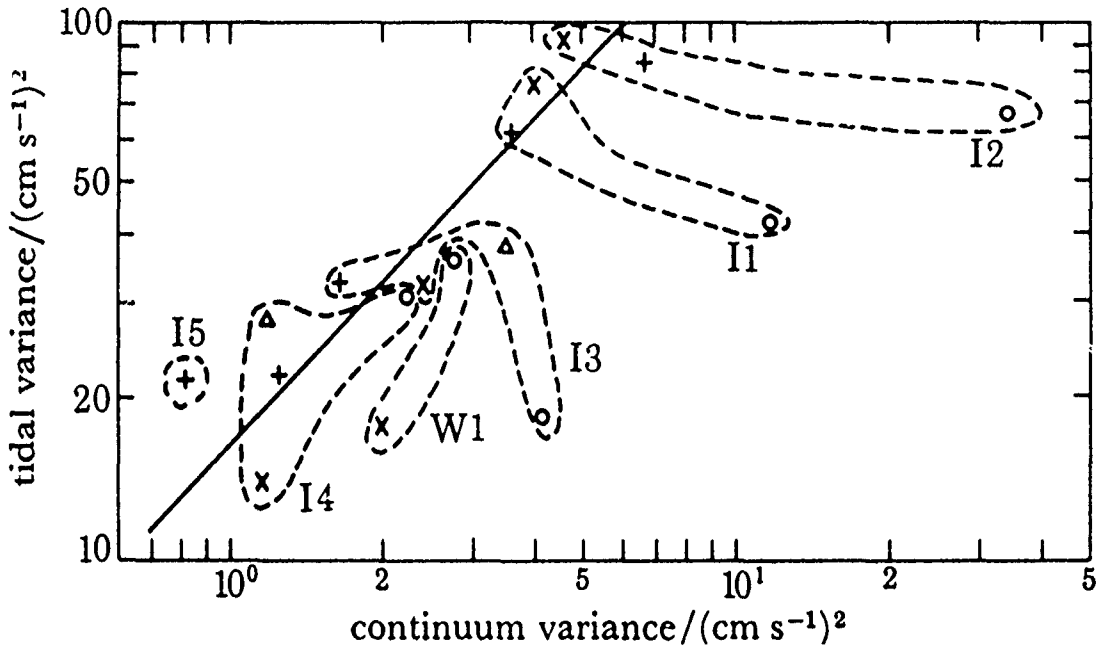


Figure 2. Variance of horizontal kinetic energy in the continuum (.1 to .3 cph) is plotted against variance in the tidal band (.076 to .090 cph) from nominal depths of 200 (×), 600 (+), 1000 (O) and 1500 m (△). The straight line indicates a constant ratio of tidal to continuum variance fitted to data at 200 and 600 m. (Adapted from Levine *et al.*, 1983.)

Scotland indicate a correlation between energy in the tidal band and in the continuum. A higher energy internal wave continuum was found with larger internal tides. This correlation, of course, does not prove cause and effect as the higher energy was found near steeper topography. Perhaps the entire internal wave spectrum was enhanced by interaction with topography--the tidal band and continuum could have no dynamic link.

Another interesting set of moored observations from the equator (Eriksen, 1985) suggests that a change in the spectral slope occurs at the semidiurnal frequency (Figure 3). If the tide did not interact with the background wave field, then there is no obvious reason for different spectral slopes above and below the semidiurnal frequency. The persuasiveness of this argument may depend somewhat on one's artistry with pen and ruler.

If energy at the semidiurnal frequency were being nonlinearly transferred, one might expect spectral peaks at harmonics of the tidal frequency. Indeed, harmonics are often observed in the open ocean in deep water. High resolution spectra suggest that the smooth continuum spectrum may actually consist primarily of harmonics of the tide and near-inertial waves (Pinkel 1981 and 1983).

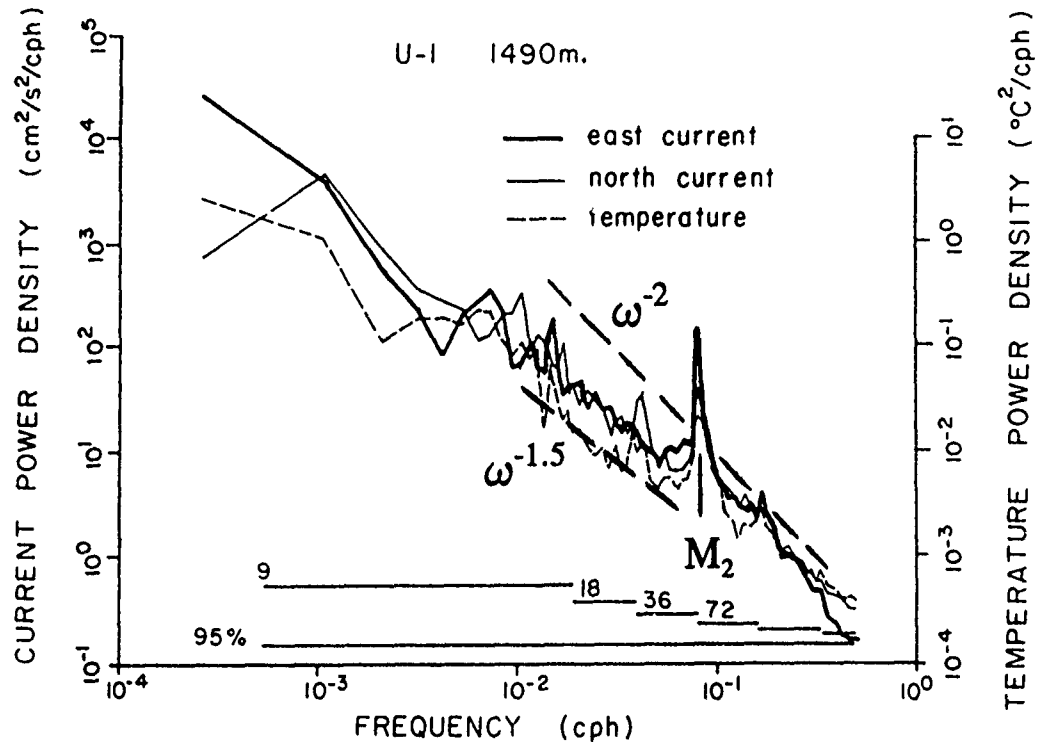


Figure 3. Autospectra of current and temperature from mooring near the equator (0°20.7'N,144°32.6'W). There is an apparent change in the spectral slope at the semidiurnal frequency. (Adapted from Eriksen, 1985.)

Even if the internal tide is not a significant energy source to the continuum, it may be worthy of study as a *test wave*; consider the ocean as a wave tank with the tide acting as a steady wave maker. It might be possible to track the propagation of the internal tide because of its high amplitude and thereby estimate its interaction with other waves. Perhaps theoretical ideas of nonlinear interaction could be tested in this way.

NEAR-INERTIAL WAVES

Waves in the near-inertial frequency band have been studied extensively for decades. Although these waves are included in the universal GM formulation, there is a large amount of space/time variability in the signal. A substantial fraction of the near-inertial energy comes from local generation (Fu, 1981) and has been modeled deterministically by many (e.g., Pollard 1970). Since most of the shear variance comes from the near-inertial band (Figure 1b), it may be risky to use the GM model to estimate vertical shear for purposes of predicting turbulent mixing. Perhaps a better way is needed to express the intermittency of these waves so that a more accurate parameterization of the wave field could be made. Some ideas of intermittency will be explored in the next section.

Can the wind-generated, near-inertial waves be the source of energy for the entire internal wave continuum? Based on a linear model, the overall near-inertial energy flux input estimated from real wind fields in the N.E. Pacific gives ~ 1 mW m^{-2} (D'Asaro, 1985), enough to make a significant contribution to the internal wave continuum. Note that most of the flux occurs in relatively few storm events--demonstrating the danger of using average values. To check for correlation, the temporal variability of the near-inertial band is compared with the higher frequency continuum in Figure 4. The mean spectrum has been divided out; hence, a contour value of 2 indicates that the spectral level is twice the average value. These observations are from 140 m on a mooring deployed in the N.E. Pacific during the Ocean Storms experiment. It appears that the fluctuations in the continuum are correlated with the near-inertial band, most of the time. Fluctuations in the high-frequency band, from 0.1 to 2 cph, seem to occur at all frequencies by about the same factor. However, the variation of the near-inertial band is usually greater than the high frequencies. It remains to determine if these observations are consistent with the notions of nonlinear interaction. Perhaps it is possible that the increased high-frequency waves are directly forced by the wind by a linear process? Or maybe the increased high-frequency signal is merely the result of the enhanced Doppler shifting of existing waves?

THE RANDOM WAVE FIELD HYPOTHESIS

The description of the internal wave field as a superposition of waves with random phase is a basic tenet of the GM model. The notion of a random wave field is also used extensively by theoreticians when modeling nonlinear interactions. How accurate is this assumption? What if there is significant correlation between some of the waves? What are the ramifications to the description and modeling of the wave field if the assumption of random phase is violated, even slightly?

Rather than tracking each wave component in α-β-ω space perhaps it is more efficient to follow the modulation or envelope of the wave field. Of course, if the wave field is modeled as a sum of random waves, then the statistical properties of the envelope can be inferred. But it still may be more straightforward to concentrate on modeling the modulation directly as it is a quantity of real interest--the wave energy. The relationship between individual waves and their envelope has been studied extensively for surface waves by Longuet-Higgins (1984) and has recently been applied to internal waves by Moustafa and Rubenstein (1990). For a random wave field with a given spectral bandwidth, say in frequency, the statistics of the temporal modulation can be predicted. Does the observed modulation of the internal wave field follow the prediction? The analysis by Moustafa and Rubenstein (1990) indicated that the temporal modulation of the internal wave field seemed consistent with the hypothesis of random phase between components. Further analysis of this type with more data and extended to include the spatial dimensions is needed to determine if this behavior is universal.

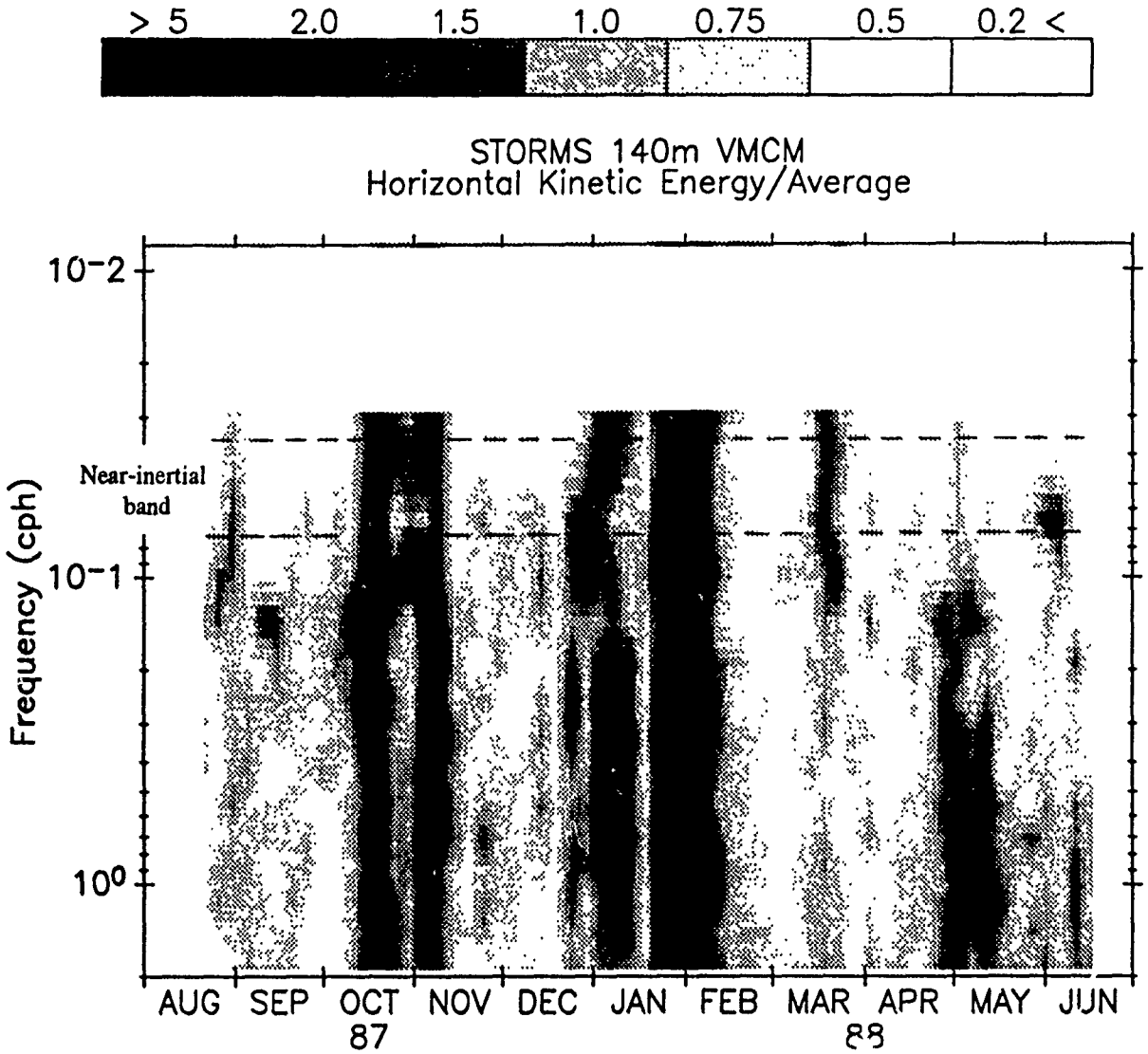


Figure 4. Frequency spectrum of horizontal kinetic energy as a function of time as observed at 140 m during OCEAN STORMS. Plotted values are the ratio of the spectral density divided by the average value. Hence, a value of 2 implies 2 times the average spectral density. The modulation of the high-frequency internal wave field is often correlated with the near-inertial variations. However, there are times, such as in early May when high-frequency waves increase without an associated increase in near-inertial energy.

An interesting extension of this approach might be to consider the consequences when all the phases between components are **not** random. Or turning the question around, what does failure of the random phase test imply about the correlation of the phases. This is undoubtedly a complicated problem since there are many ways that waves can be correlated, but only one way they can be uncorrelated.

The tracking of the modulation of the wave field rather than each wave component has been applied to observations by many investigators, e.g. Frankignoul and Joyce (1979), Briscoe (1983), and Brown and Owens (1981). The approach is usually to estimate the energy or momentum in a frequency band as a slowly varying function of time. The temporal variability can then perhaps be related to a source or sink of energy or momentum in a deterministic manner. For example, an observed temporal correlation between the internal wave energy and the large horizontal shear may indicate a cause and effect relation and lead to identification of an internal wave source.

Another approach would be to consider the modulation of energy itself as a random process and try to describe its variability. The results might not identify specific sources of wave energy, but might lead to an improved statistical description. For example, consider a formulation that tracks wave packets or groups. Suppose a wave packet consists of waves within a small frequency/wavenumber bandwidth (δk_n, $\delta\omega_n$) centered around (k_n, ω_n) (Figure 5a). The total energy of the wave group will vary in space and time because of the finite bandwidth. The statistical properties of the modulation will depend on the relative magnitudes of the waves included in the band as well as the degree to which the phases of these waves are correlated. The conjecture is that it is simpler to ignore the details of the infinite number of waves in the packet and just describe the time space variation of the packet itself. Let the total modulation $M(x,t)$ of the wave field be written as

$$M(x,t) = \sum_n \sum_m B(K_n, \Omega_m) \exp[i(K_n \cdot x - \Omega_m t)]$$ where (K_n, Ω_m) is the wavenumber and frequency of the **modulation** of the wave packet composed of individual waves with frequency/wavenumber centered at (k_n, ω_m). The quantity $B(K,\Omega)$ is a spectral distribution function, not of the waves themselves, but of the **modulation** of the waves. Note that the wave speed $c = \Omega_m / K_n$ in this case is the **group** velocity, not the phase speed, since it is the speed of the wave packet or energy. Perhaps a universal, statistical model of the wave energy can be formulated with this approach in a spirit similar to the GM framework. That is, observed spectra and coherences of the **modulation** would be used to define $< B(K,\Omega)\, B'(K,\Omega) >$, just as observed spectra and coherences were used by GM to establish $E(k,\omega)$. As an example, an estimate of the spectra of the modulation of vertical displacement in the 1 to 4 cph frequency band is given in Figure 5b. This spectrum can be characterized by an $\Omega^{-1.5}$ dependence with a hint of a few peaks. These observations are from a month-long time series recorded during AIWEX (Arctic Internal Wave Experiment). Is this universal, and what does it tell us? The feasibility and usefulness of employing this approach remains an open question.

Instead of resorting to a Fourier representation for $M(x,t)$, it may be useful to try an entirely different representation. For example, a wavelet transform of the modulation

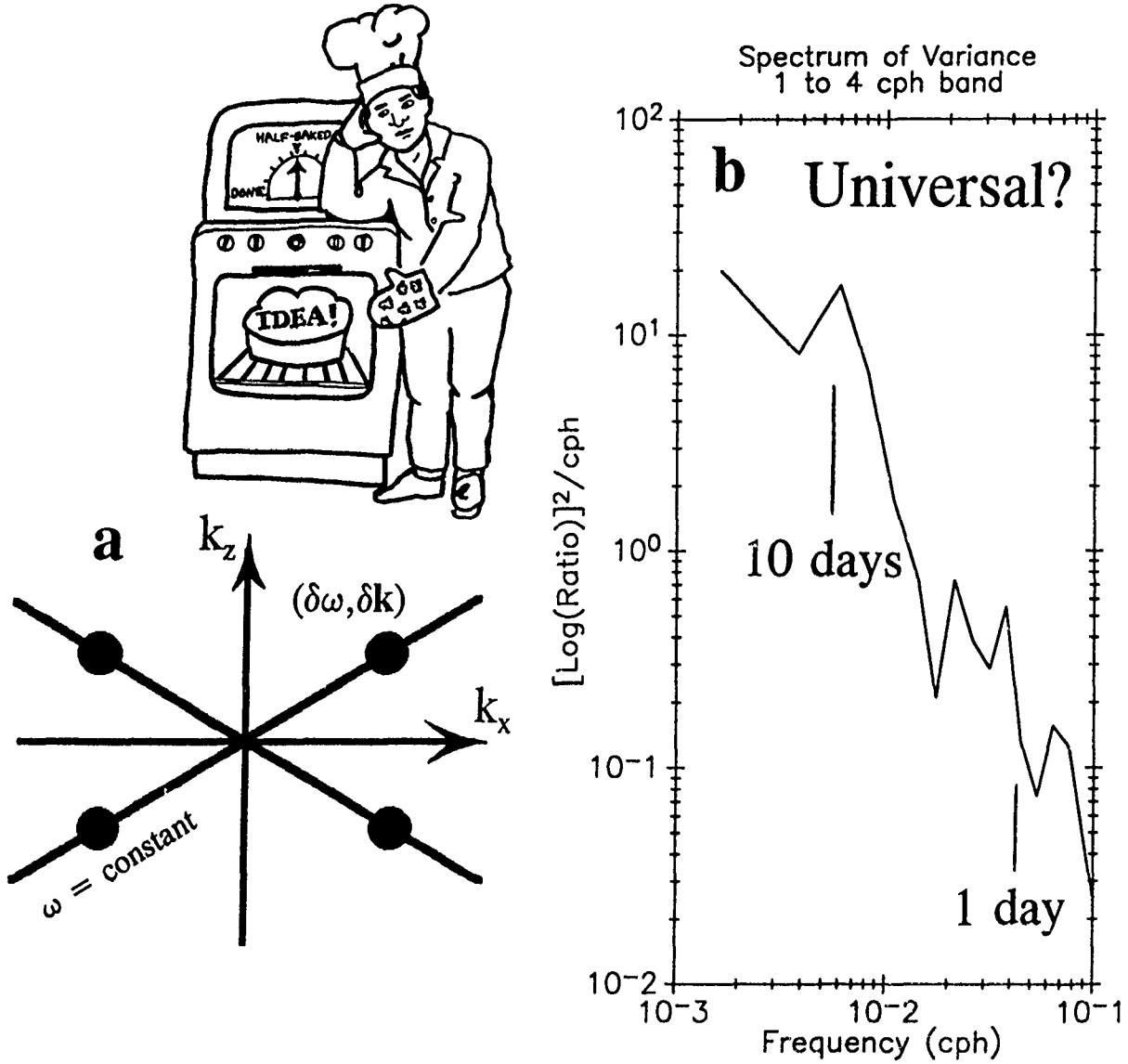


Figure 5. Schematic diagram in wavenumber space (a) of a group of waves with bandwidth $(\delta\omega,\delta\mathbf{k})$. Spectrum of the temporal modulation of the vertical displacement variance of waves is in the 1 to 4 cph band. The spectral density is expressed as the log of the ratio relative to the average variance. These data were recorded for 1 month during AIWEX (Arctic Internal Wave Experiment) in the Beaufort Sea at 257 m depth.

field may provide insight. This transform would express the modulation time series as a sum of localized "packets" of varying length. Of course, any representation of the wave field must also behave somewhat like the GM model, at least to the extent that the GM statistics mimic observations. However, a different representation may lead to significant differences in determining higher-order statistics, estimating the consistency relationships, and modeling nonlinear interactions.

CONCLUSIONS

Theoretical studies of the dynamics of oceanic internal waves rely on a description of the oceanic internal wave field. The more accurate the description, the better the chance for a realistic theory. The GM model has provided the first order statistical formulation of the wave field. However, effort is needed to improve the description of aspects of the wave field that do not follow GM. Features of the non-GM wave field that need further study include internal tides, near-inertial motion and the random wave assumption. While these areas of research are not new, they need to be updated and refined to incorporate new data and theoretical ideas.

Acknowledgments

Many thanks to Peter Müller, Crystal Miles and the other organizers of the 'Aha Huliko'a Hawaiian Winter Workshop. The meeting was physically relaxing and scientifically stimulating. The support of the Office of Naval Research (N00014-90-J-1036 & N00014-90-J-1048) is gratefully appreciated.

REFERENCES

Baines, P.G., 1982: On internal tide generation models, *Deep-Sea Res., 29,* 307-338.

Bell, T.H. Jr., 1975: Topographically generated internal waves in the open ocean, *J. Geophys. Res., 80,* 320-327.

Briscoe, M.G., 1983: Observations on the energy balance of internal waves during JASIN, *Phil. Trans. R. Soc. Lond., A 308,* 427-444.

Brown, E.D., and W.B. Owens, 1981: Observations of the horizontal interactions between the internal wave field and the mesoscale flow, *J. Phys. Oceanogr., 11,* 1474-1480.

D'Asaro, E.A., 1985: The energy flux from the wind to near-inertial motions in the surface mixed layer, *J. Phys. Oceanogr., 15,* 1043-1059.

Duda, T.F., and C.S. Cox, 1989: Vertical wave number spectra of velocity and shear at small internal wave scales, *J. Geophys. Res., 94,* 939-950.

Eriksen, C.C., 1985: Some characteristics of internal gravity waves in the equatorial Pacific, *J. Geophys. Res., 90,* 7243-7255.

Frankignoul, C., and T.M. Joyce, 1979: On the internal wave variability during the Internal Wave Experiment (IWEX), *J. Geophys. Res., 84,* 769-776.

Fu, L.-L., 1981: Observations and models of inertial waves in the deep ocean, *Revs. Geophys. Space Phys., 19,* 141-170.

Garrett, C.J.R, and W.H. Munk, 1972: Space-time scales of internal waves, *Geophys. Fluid Dyn., 2,* 225-264.

Garrett, C.J.R, and W.H. Munk, 1975: Space-time scales of internal waves: a progress report, *J. Geophys. Res., 80,* 291-297.

Gregg, M.C., 1989: Scaling turbulent dissipation in the thermocline, *J. Geophys. Res., 94,* 9686-9698.

Ivanov, V.A., and A.N. Serebryanyy, 1982: Frequency spectra of short-period internal waves in a nontidal sea. *Izv., Atmos. and Ocean. Phys., 18,* 527-529.

Levine, M.D., C.A. Paulson, M.G. Briscoe, R.A. Weller and H. Peters, 1983: Internal waves in JASIN, *Phil. Trans. R. Soc. Lond., A 308,* 389-405.

Longuet-Higgins, M.S., 1984: Statistical properties of wave groups in a random sea state, *Phil. Trans. R. Soc. Lond., A 312,* 219-250.

Moustafa, M.S., and D. Rubenstein, 1990: Intermittency of internal waves, *EOS, 71,* Abs. only, 101.

Olbers, D.J., and N. Pomphrey, 1981: Disqualifying two candidates for the energy balance of oceanic internal waves, *J. Phys. Oceanogr., 11,* 1423-1425.

Pinkel, R., 1981: Observations of the near-surface internal wavefield, J. Phys. Oceanogr., 11, 1248-1257.

Pinkel, R., 1983: Doppler sonar observations of internal waves: wave-field structure, *J. Phys. Oceanogr., 13,* 804-815.

Pollard, R.T., 1970: On the generation by winds of inertial waves in the ocean, *Deep-Sea Res., 17,* 795-812.

OVERVIEW OF OCEAN BACKGROUND CONDITIONS AND MODELING OF INTERNAL WAVES

Alex Warn-Varnas
SACLANT Research Center, Via S. Bartolomeo 400
19026 La Spezia, Italy

Steve Piacsek
Code 322, Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, MS 39529

INTRODUCTION

In this paper we would like to consider some aspects of the oceanic background and the results of some modeling studies that pertain to the generation and propagation of internal waves in the ocean. We will look at this background state as determined from the various field surveys and will consider the feasibility of modeling these background conditions and phenomena. One of the motivations for the current studies is that the propagation of internal waves is strongly affected by velocity and density gradients and that the ocean exhibits a high spatial and temporal variability of these gradients. This background description encompasses various elements, of which we consider the mixed layer region, the mesoscale variability and associated eddies, fronts, fine structure, and their interaction with internal waves.

EXAMPLES OF INTERNAL WAVE PROPAGATION AND SHEAR

Figure 1 shows the results of the simulation of a collapsing turbulent patch in a large scale oceanic shear field (Piacsek and Roberts, 1975). Note the generation of internal waves indicated by the streamfunction patterns in the form of rays. There is a prescribed mean vertically sheared flow present that moves from left to right. As the collapse progresses, more rays come out and are bent to the right. In the upper right hand side the presence of a critical layer sets up a region into which the waves cannot propagate. On the left side the rays are refracted toward the vertical and undergo reflection at the surface, setting up standing wave-like patterns.

Propagating internal waves have often been observed at the bottom of the mixed layer and in the seasonal thermocline. The results depicted in Figure 2 were obtained with a 200-meter thermistor chain during a 1987 Planet cruise in the Norwegian Sea (Sellschopp. 1987). The top portion of the figure shows the time series of the temperature profile versus depth. The corresponding isotherms are shown in the bottom portion. Note the presence of internal waves at the base of the mixed layer and in the thermocline.

It has been established that shear and internal wave activity depend on certain environmental parameters. These are the Brunt-Vaisala frequency, atmospheric forcing, geostrophic currents, and topography. The conclusions were reached on the basis of field surveys and theoretical considerations (Garrett and Munk, 1979; Ruddick and Joyce, 1979; Brown and Owens, 1981; Briscoe and Weller, 1984).

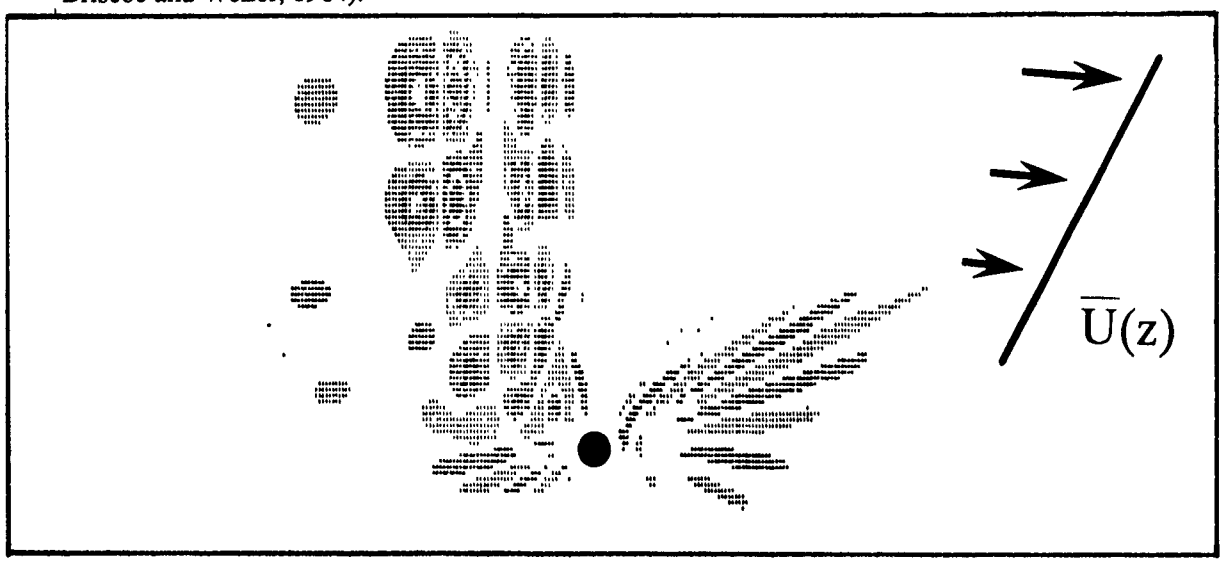


Figure 1. Internal wave generation by a collapsing turbulent patch in a stratified fluid, in the presence of a mean shear (flow from left to right) (Piacsek and Roberts, 1975).

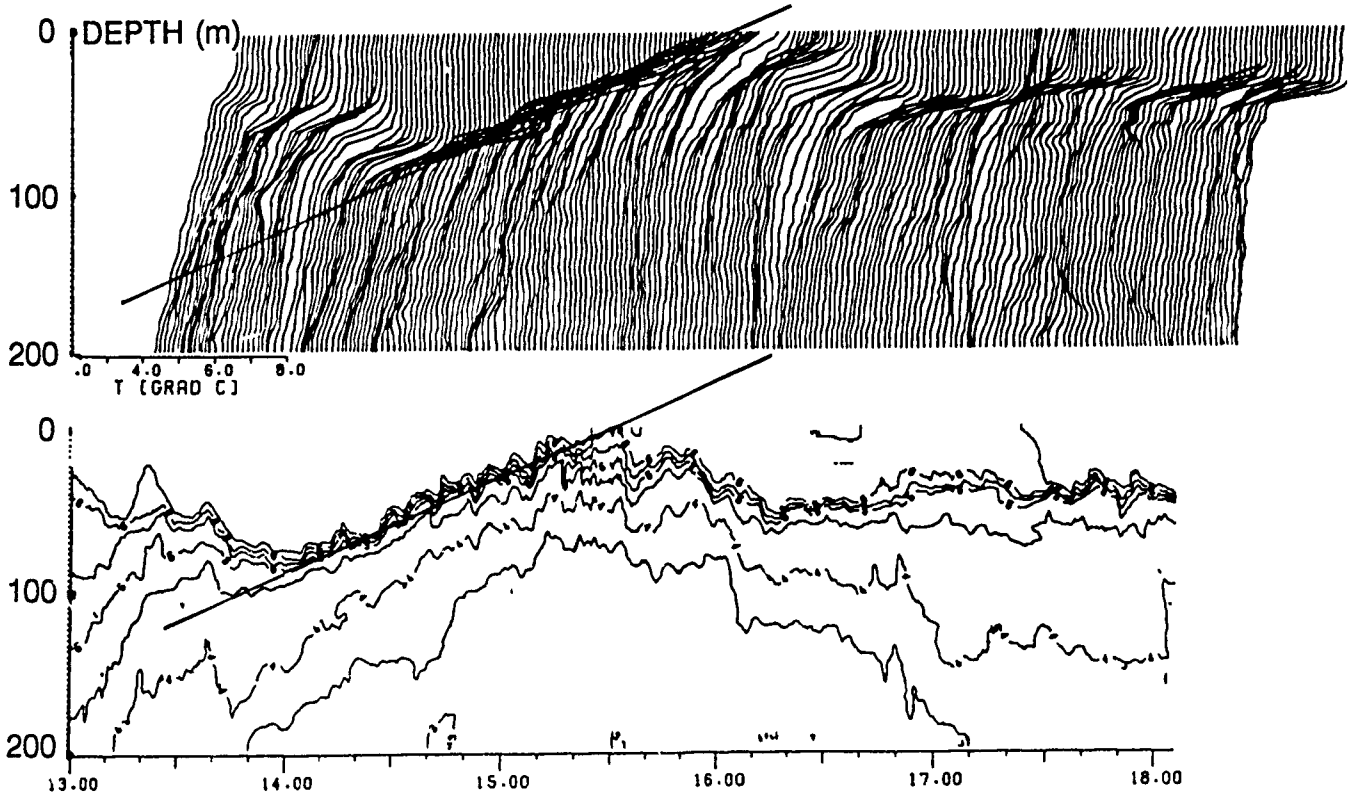


Figure 2. Internal wave propagation at the bottom of the mixed layer and in the seasonal thermocline (Selischopp, 1987).

Strong vertical geostrophic shears have been measured near fronts using an ADP (acoustic doppler profiler), during the previously mentioned cruise in the Norwegian Sea along a different track. Figure 3 shows the measured shear at the edges and middle of a warm feature. Strong positive shear occurs when passing from cold to warm water (cut a). There is a lack of shear in the middle of the eddy (cut b) and the presence of a negative shear when passing from warm to cold water (cut c). The average vertical shears are about 10^{-3} sec^{-1} in the frontal region, which is about five times smaller than the magnitude found near the base of mixed layers.

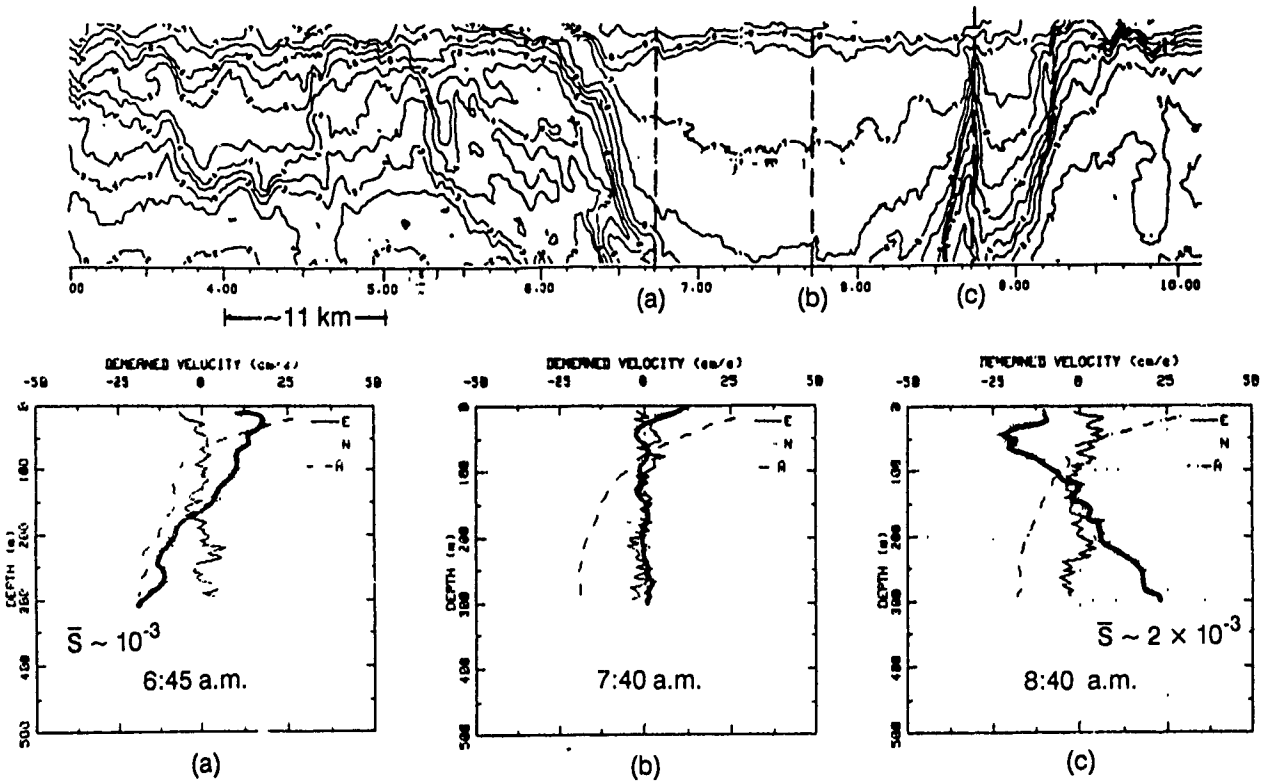


Figure 3. Examples of geostrophic shear, as observed by an acoustic Doppler profiler, in passing into and out of a warm eddy in the Faroe area (Sellschopp, 1987).

An example of strong horizontal geostrophic shear in the confluence region of the Icelandic current and the North Atlantic inflow is shown in Figure 4. Horizontal shears are about 10 cm/sec/10 km, and in some sections can be larger. The results are based on the analysis of the Icelandic Current and the Atlantic Inflow Experiment surveys, carried out by the SACLANT Research Center in 1987 (Hopkins et al. ,1989).

An analysis of the internal wave energy budget has been performed for the JASIN survey data by Briscoe (1983). The results show that sometimes the rate of change of internal wave energy correlates with the mesoscale horizontal shear, the wind, and inertial wave activity (Figure 5). The correlation is not present all the time, however. Part (b) of the figure shows the dependence of

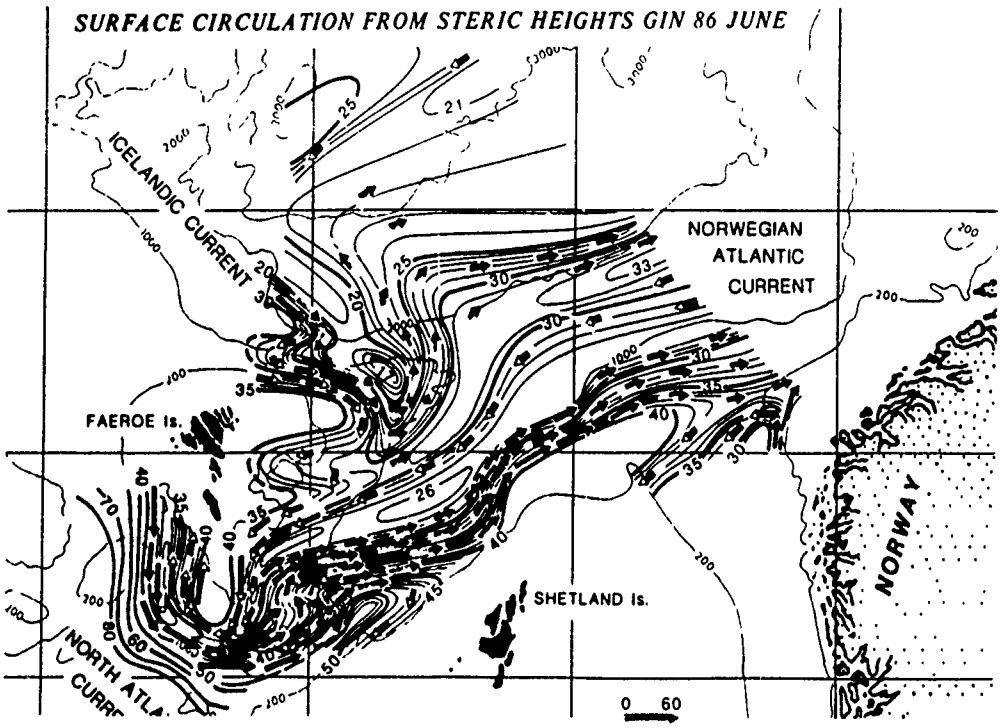


Figure 4. Example of strong horizontal geostrophic surface shears in the North Atlantic Inflow between Faroes and Shetland Islands (Hopkins et al., 1989).

internal wave energy on surface wave amplitude delayed by 11 days. The correlation is very high throughout the measurement period. Internal wave activity exhibits a seasonal variation: the analysis and interpretation of the LOTUS data demonstrate this (Figure 6) (Briscoe and Weller, 1984). Note the marked decrease in internal wave activity during the summer at the 100-500 meter depth levels. Near the surface the decrease occurs in the fall.

Figure 7 shows an example of stepped T,S structures in the Eastern Caribbean region southeast of Barbados observed during the C-SALT experiment (Boyd and Perkins, 1987). The steps gradually increase with depth from about 5 to 30 meters at the 200 to 350-meter depth levels, respectively. The temperature and salinity jumps at the profile 'steps' are thought likely to be set up by strong salt fingering processes; such fingers have actually been observed in the ocean by optical means.

FOCUS OF CURRENT STUDIES

Our current objectives are concerned with the provision of a deterministic forecast (or calculation from forecasted variables) for environmental parameters of interest at resolvable computational scales. For the unresolvable scales statistical predictions of various kinds would be used, adapted, and developed. In such fashion a combined deterministic and statistical information on environmental parameters of interest would be provided.

In the upper ocean successful prediction of the mixed layer behavior has been achieved in situations where the atmospheric forcing activity is dominant vis a vis mesoscale advection. Figure 8 displays

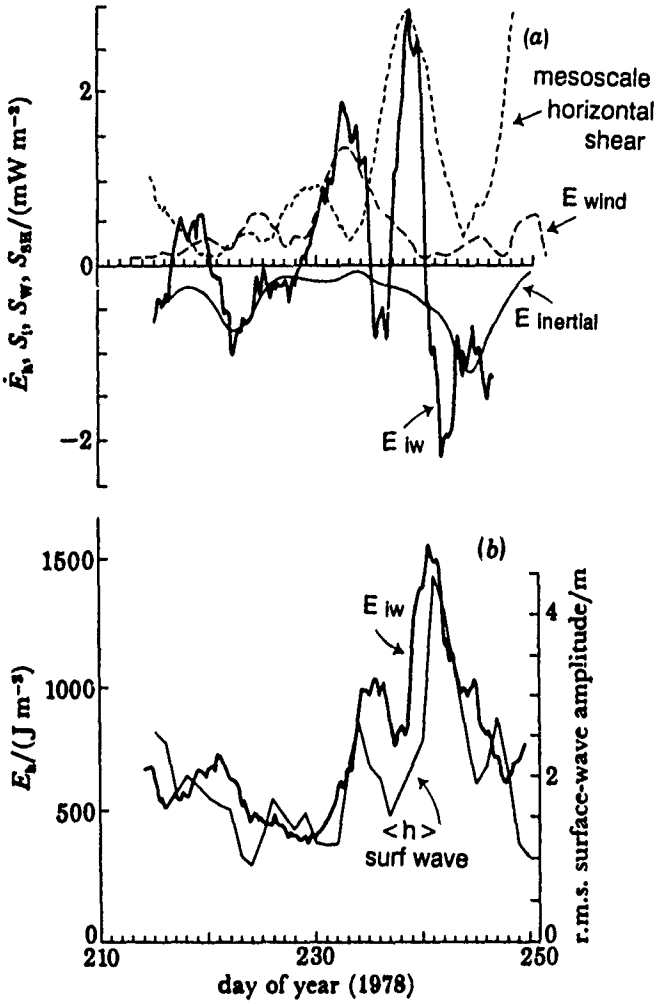


Figure 5. Dependence of internal wave energy on surface forcing during the JASIN experiment in the Northeast Atlantic (Briscoe, 1983).

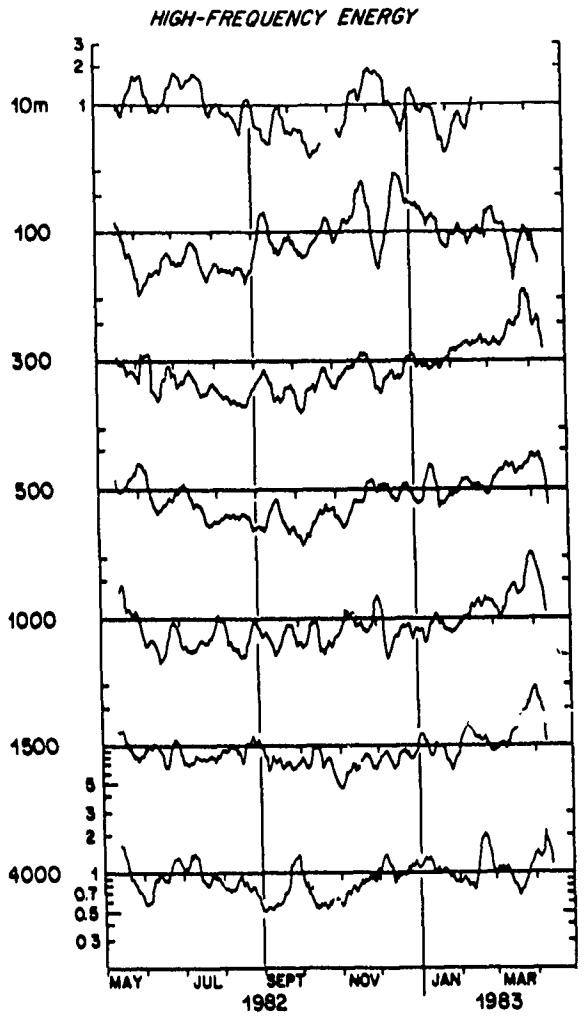


Figure 6. The depth dependence of the seasonal variation of internal wave energy during the LOTUS experiment in the Sargasso Sea (Briscoe and Weller, 1984).

mixed layer simulations with three different models at a former ocean station November (30N, 140W in the Pacific) (Martin, 1985). The Niiler (1975) and Garwood (1977) models are of the 'bulk' type (they assur ıe the existence of a mixed layer and work with the vertically integrated or 'bulk' equations), whereas the Mellor-Yamada model is of a differential type (i.e., set up by differencing on a vertical grid). Initial conditions were set to the temperature profile obtained with an XBT and the atmospheric forcing was derived from meteorological measurements. All the simulations show the seasonal trends in sea surface temperature and mixed layer depth that are observed in the data. However, detailed differences do exist in relation to the data. The models show a quicker spring transition. The ML2 model mixes down less deep in the winter than the other two and is more at variance with the data.

Simulations of mixed layer behavior that resolved the diurnal (Price et al., 1986) and synoptic time scale (Warn-Varnas et al., 1981)]were also performed. Some investigators studied the mixed layer

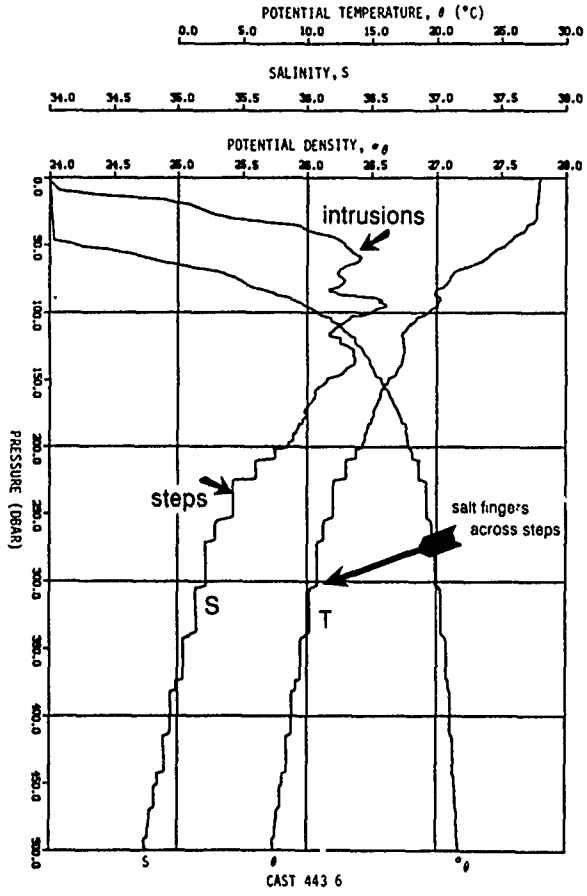


Figure 7. Step structures in temperature and salinity profiles as observed in the C-SALT experiment off Barbados Boyd and Perkins, 1987).

response to hurricanes (Martin, 1982; Price, 1981). Figure 9 shows one such study for hurricane Eloise that occurred in the Gulf of Mexico. A comparison is shown between the predicted and measured u and v velocity components. The measurements were made by the NDBO buoy EB-10. Both phase and amplitude show agreement with data with some slight differences.

A system to predict regional mesoscale behavior has been developed by the Harvard University group (Robinson et al., 1986). The dynamical model is initialized from field survey data and is forced by observed boundary conditions. Data can be assimilated into the model. The boundary conditions play a crucial role in the forecast. Figure 10 displays dynamical simulations of the California Current with such a system. Note that on day 14, the simulation with the boundary condition obtained by linear interpolation of the measurements gives agreement with the data. The persistence boundary conditions leads to substantial deviations from the actual observed state.

At present a capability has been developed to predict various environmental parameters in conjunction with a real-time upper ocean nowcasting system that is operational at FNOC (Fleet Numerical Oceanographic Center) (Clancy and Pollak, 1983). Other prediction systems and approaches for describing environmental parameters of interest exist in research environments.

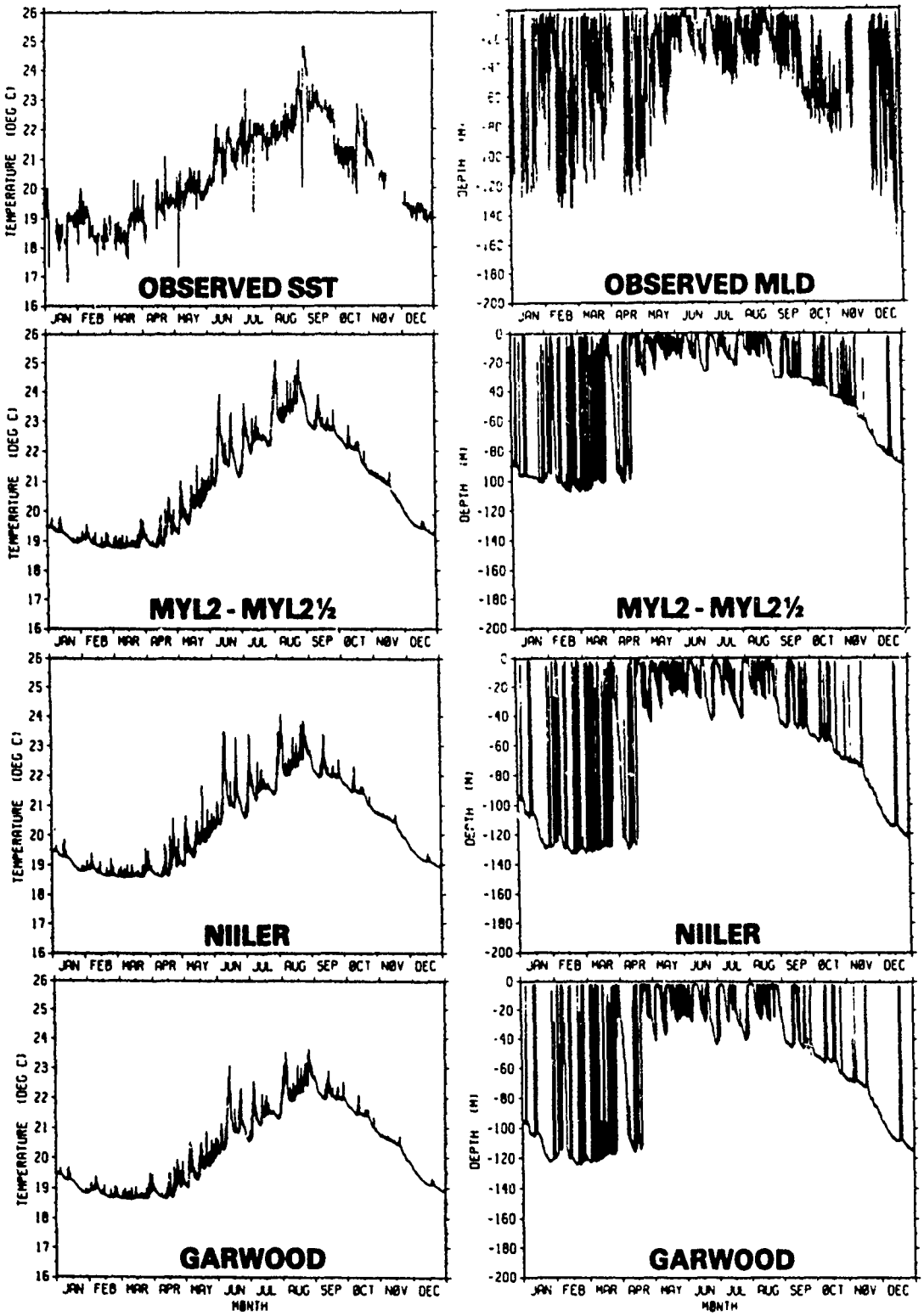


Figure 8. Mixed layer simulation with different models at weather ship November in the Pacific (Martin, 1985).

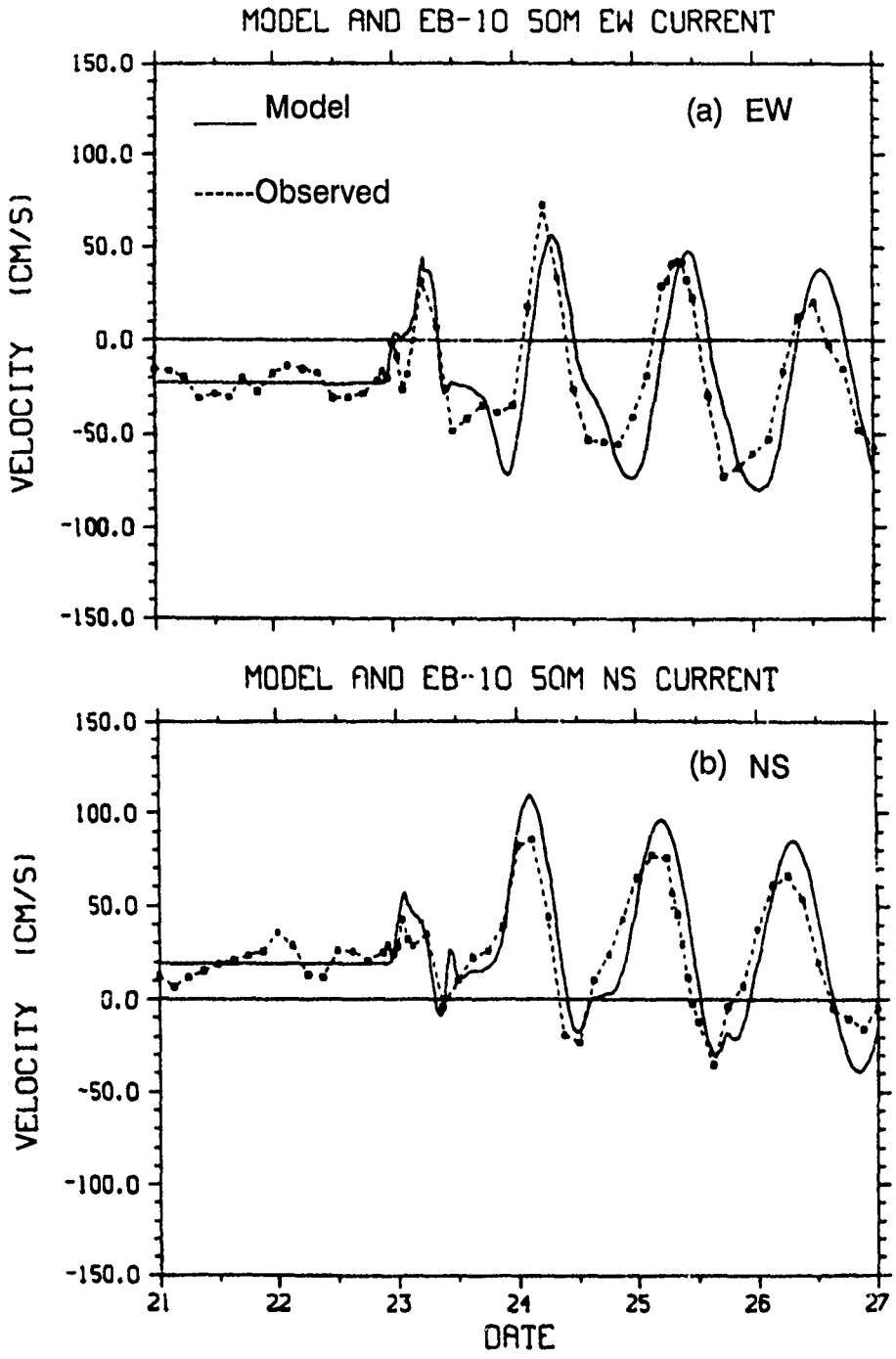


Figure 9. The mixed layer development under Hurricane Eloise (Martin, 1982).

The one that we are focusing on has the advantageous feature of being run operationally every day in conjunction with an objective analysis of all available data for the world's oceans. This system, called TOPS (thermodynamic ocean prediction system) (Clancy and Martin, 1981), encompasses a mixed layer model (formulated as a three-dimensional boundary layer model for the upper ocean), prognostic Ekman and inertial velocities (predicted by the mixed layer model), and a geostrophic velocity (provided by a diagnostic calculation or a prognostic forecast). The equations of this model (Warn-Varnas et al., 1984) take the form

$$\frac{\partial \overline{T}}{\partial t} = \frac{\partial}{\partial z}\left(\overline{-wT'} + K\frac{\partial \overline{T}}{\partial z}\right) + \frac{1}{\rho w^c}\frac{\partial \overline{F}}{\partial z}$$
$$-\frac{\partial}{\partial x}\left(u_a\overline{T}\right) - \frac{\partial}{\partial y}\left(v_a\overline{T}\right) - \frac{\partial}{\partial z}\left(w_a\overline{T}\right) + A\left(\frac{\partial^2\overline{T}}{\partial x^2} + \frac{\partial^2\overline{T}}{\partial y^2}\right) \quad (1)$$

$$\frac{\partial \overline{S}}{\partial t} = \frac{\partial}{\partial z} m\left(\overline{-w'S'} + K\frac{\partial \overline{S}}{\partial z}\right)$$
$$-\frac{\partial}{\partial x}\left(u_a\overline{S}\right) - \frac{\partial}{\partial y}\left(v_a\overline{S}\right) - \frac{\partial}{\partial z}\left(w_a\overline{S}\right) + A\left(\frac{\partial^2\overline{S}}{\partial x^2} + \frac{\partial^2\overline{S}}{\partial y^2}\right) \quad (2)$$

$$\frac{\partial \overline{u}}{\partial t} = f\overline{v} + \frac{\partial}{\partial z}\left(\overline{-w'u'} + \nu\frac{\partial \overline{u}}{\partial z}\right) - D\overline{u} \quad (3)$$

$$\frac{\partial \overline{v}}{\partial t} = -f\overline{u} + \frac{\partial}{\partial z}\left(\overline{-w'v'} + \nu\frac{\partial \overline{v}}{\partial z}\right) - D\overline{v} \quad (4)$$

where T is the temperature, S the salinity, u and v the x- and y-components of the current velocity (the x and y horizontal coordinates are defined relative to the grid), w the z-component of the current velocity, F the downward flux of solar radiation, r_w a reference density, c the specific heat of seawater, D a damping coefficient, ν a diffusion coefficient, f the Coriolis parameter, A the horizontal eddy diffusion coefficient, t the time, and z the vertical coordinate (positive upward from the level sea surface). Ensemble means are denoted by an overbar and primes indicate departure from these means. The quantities u_a, v_a are the x-, y-, and z- components of an advection current, which will be defined subsequently.

The advective terms are retained in the temperature and salinity equations and neglected in the momentum equations on the basis of scale analysis (Haney, 1974). Such an analysis shows that the advective terms in the thermal energy equations are of order unity, while the advective terms in the momentum equations are of the order of the Rossby number. Since the Rossby number is very small in most regions of the open ocean, the advective terms are dropped in the momentum equations.

Because there are no horizontal pressure gradient terms in Eqs. (3) and (4), u and v represent the wind-drift components of the current. Neglect of horizontal pressure gradients here is motivated by the fact that geostrophic currents generally do not play an important role inside the mixed-layer region which is the issue of most concern in this study.

The terms involving the damping coefficient D in Eqs. (3) and (4) represent the drag force caused by the radiation stress at the base of the mixed layer associated with the propagation of internal wave energy downward and away from the wind-forced region (Pollard and Millard, 1970; Niiler and Kraus, 1977)] The terms involving ν in Eqs. (1)-(4) account for very weak "background" eddy diffusion (due to intermittent breaking of internal waves, for example) that exists below the mixed layer. We take $D = 0.1$ day^{-1} and $\nu = 0.1$ cm^2 s^{-1} and note that these values are within the range of estimates for these quantities.

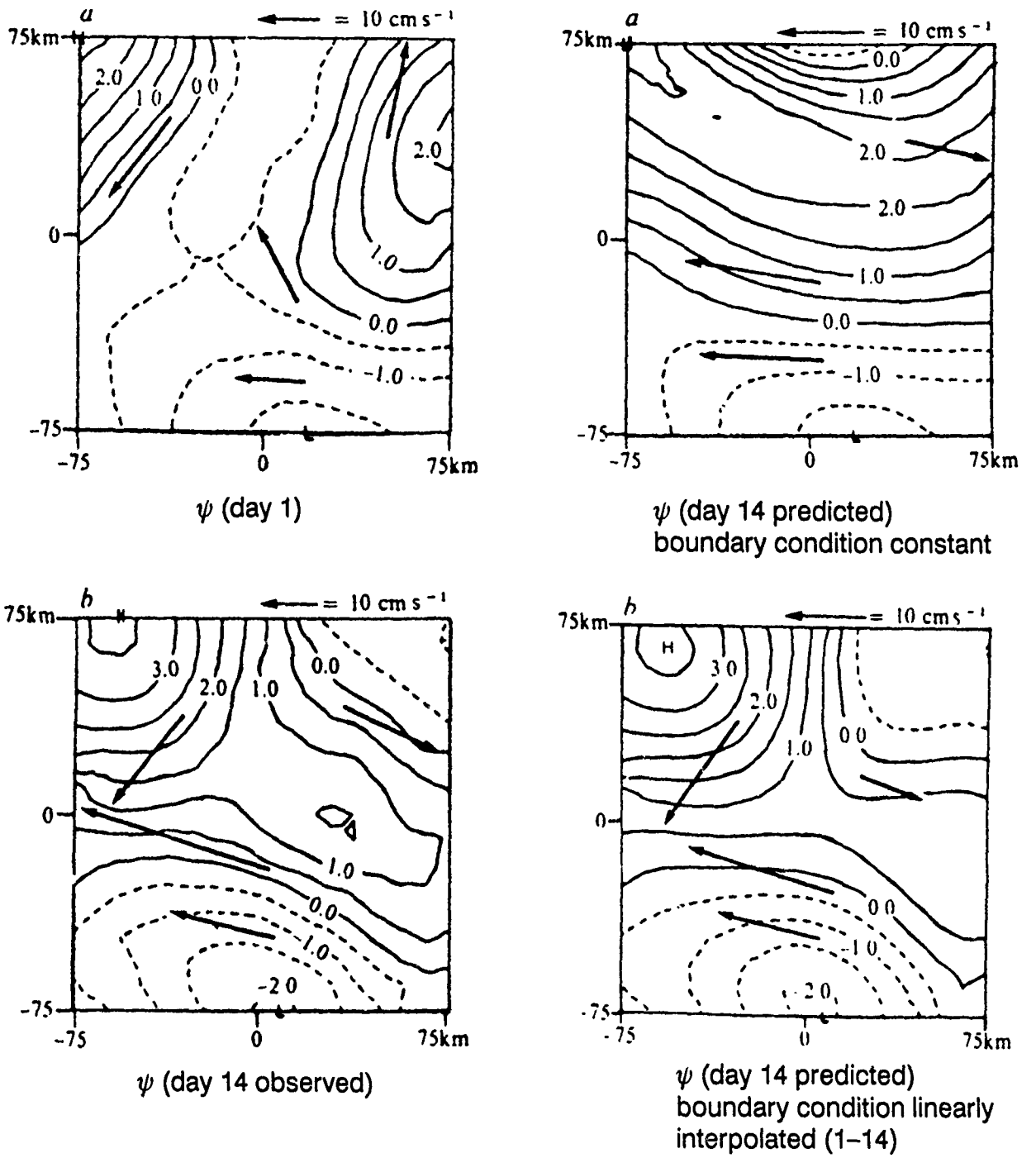


Figure 10. Example of a mesoscale ocean prediction experiment in the California Current (Robinson et al., 1986).

The Level-2 turbulence closure theory of Mellor and Yamada (1974) is used to parameterize the vertical eddy fluxes of temperature, salinity, and momentum. This turbulence model has been described in a number of papers (Mellor and Durbin, 1975; Clancy and Martin, 1981) and will not be presented here. Its energetics are essentially the same as those of Pollard et al. (1973) and Thompson (1976), with the increase in potential energy during mixed-layer deepening due to the buoyancy

flux at the layer base balanced locally by mean flow shear generation minus viscous dissipation of turbulent kinetic energy.

The horizontal eddy diffusion coefficient A is simply taken to be equal to a constant value of 10^7 cm^2 s^{-1}, and the divergence of the solar radiation flux is based on the data of Jerlov (1968) for seawater optical type 1A.

A vertically stretched grid of 17 points, extending from the level sea surface to 500 m depth, is used in the model. The horizontal grid, on which $\overline{T}$, $\overline{S}$, $\overline{u}$, $\overline{v}$, and w_a are defined, is a rectangular subset of the standard FNOC 63 x 63 Northern Hemisphere Polar Stereographic Grid.

The current used to advect the temperature and salinity is given by

$$u_a = u_i + u_g^*, \; v_a = v_i + v_g^*, \; w_a = w_i \qquad (5)$$

where u_i and v_i are the x- and y-components of the instantaneous wind-drift current, w_i is the vertical component of the current resulting from the divergence of u_i and v_i, and u_g^* are v_g^* are the components of a divergence-free geostrophic current.

The system TOPS provides information on the following environmental parameters: (a) Ekman and Inertial shears, (b) temperature and salinity profiles, (c) Brunt-Vaisala frequency, (d) Richardson number, (e) mixed layer depth, and (f) eddy coefficients

We have analyzed the archived forecasts of the TOPS prediction system for the year 1987. The analysis for environmental parameters was performed at eight former ocean station locations, Figure 11, with four locations each in the Pacific and the Atlantic, respectively. Figure 12 shows the monthly averaged mixed layer depth at the eight stations. Observe the almost simultaneous spring shallowing around the May-June period, and the more diverse fall deepening. Another striking feature of these results is the much deeper winter values of the MLD (mixed layer depth) in the Atlantic than in the Pacific, due to the presence of a halocline in the latter than tends to inhibit deep mixing in the winter. Figure 13 exhibits the simulated seasonal shear at ocean station Papa in the Atlantic. Note that the maximum shear occurs in the summer. This happens because during the summer the mixed layer is shallow and contains the momentum generated by the atmospheric forcing at the surface (in the winter the momentum is contained in a deeper mixed layer).

CONCLUSION

We advocate the use of upper ocean and mesoscale models to provide the deterministic part of the forecast of the environmental parameters. There are aspects of internal wave statistics and their non-stationarity which can be predicted by thermodynamic and hydrodynamic models of ocean circulation. Figure 14 elucidates the proposed approach. Internal wave activity depends on Brunt-

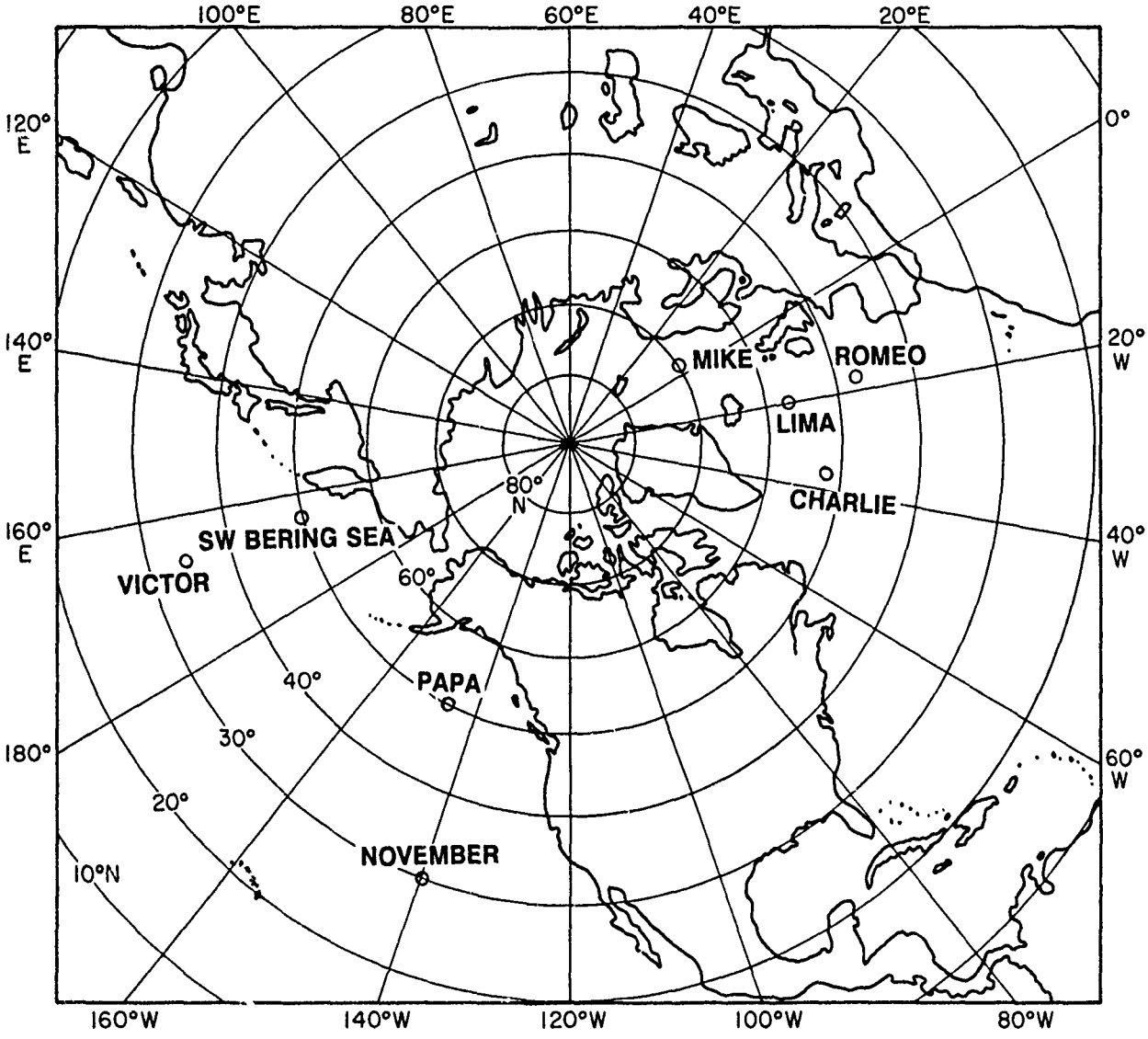


Figure 11. Locations of the three weather ships and one TOPS model grid in the North Pacific, and of the 4 weather ships in the North Atlantic, where XBT observations and routine model predictions have been carried out with the TOPS operational 3-D mixed layer model.

Vaisala frequency times a non-dimensional parameter (Briscoe and Weller, 1984; Rubenstein, 1984). Thus the model -forecasted Brunt-Vaisala frequency will enable us to calculate the internal wave activity.

Furthermore, we propose to: (a) collect statistics on internal wave and fine structure, (b) determine the critical environmental parameters on which the statistics depend, and (c) use the statistical empirical relations together with the model forecast to determine the environmental parameters of interest in terms of the mean and the deviations.

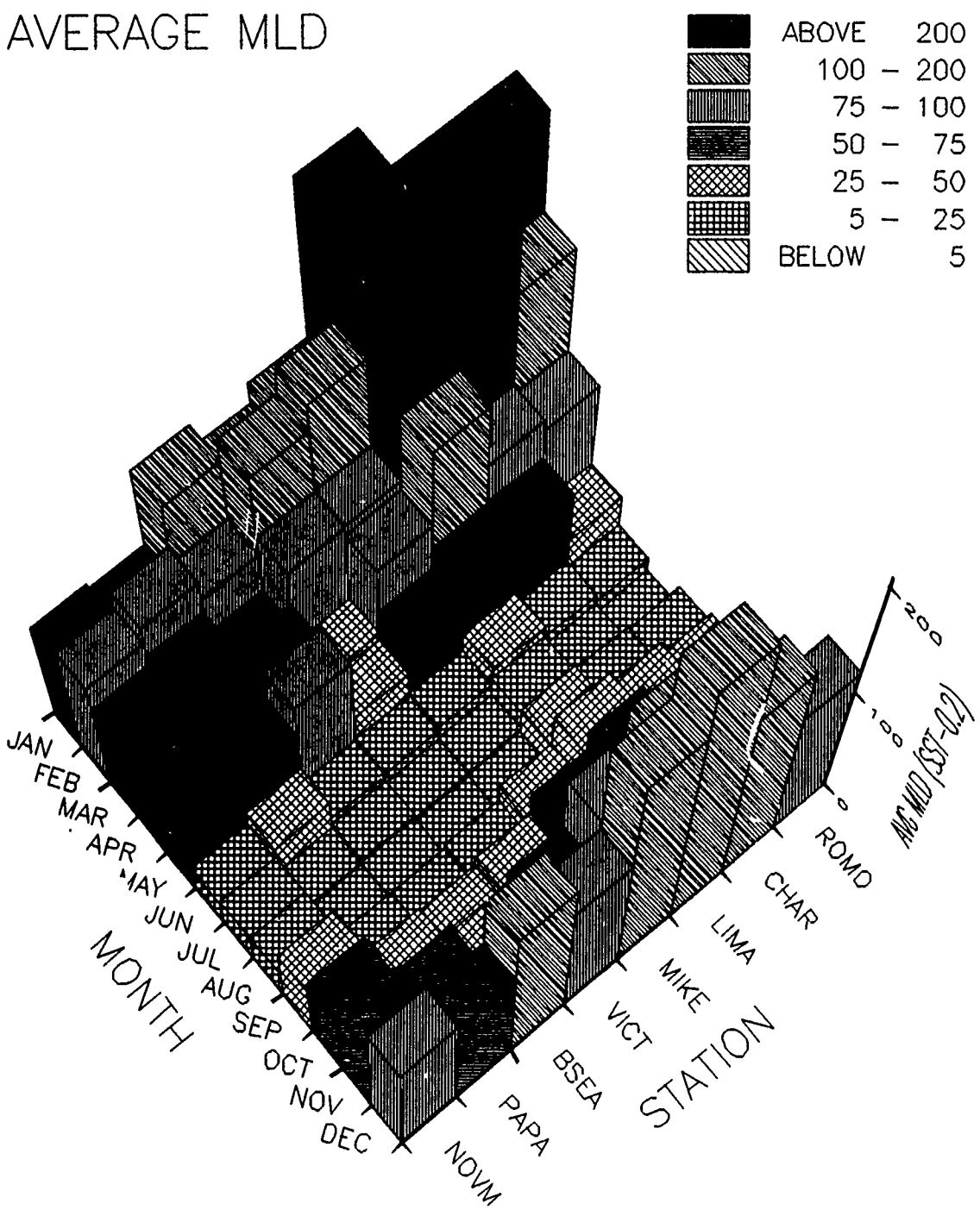


Figure 12. The average mixed layer depth (MLD) at the 8 locations illustrated in Figure 11 for the year 1985, as obtained by the operational model TOPS forced by GCM fluxes and updated by analyzed temperature fields using all available XBT's (Piacsek et al., 1988).

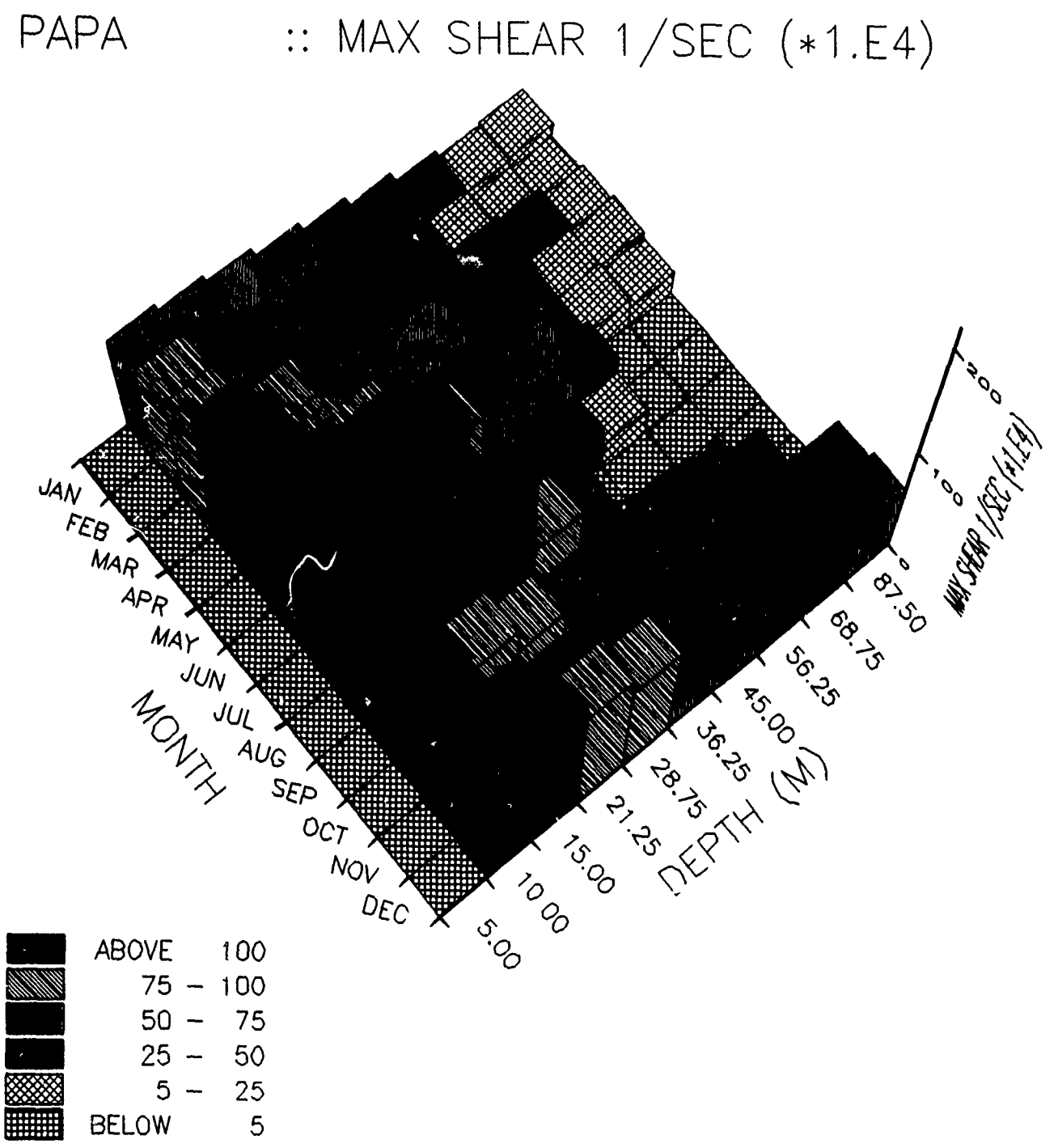


Figure 13. The maximum vertical shear at the eight locations illustrated in Figure 11 for the year 1985, as obtained by the operational model TOPS forced by GCM fluxes and updated by analyzed temperature fields using all available XBT's (Piacsek, 1988).

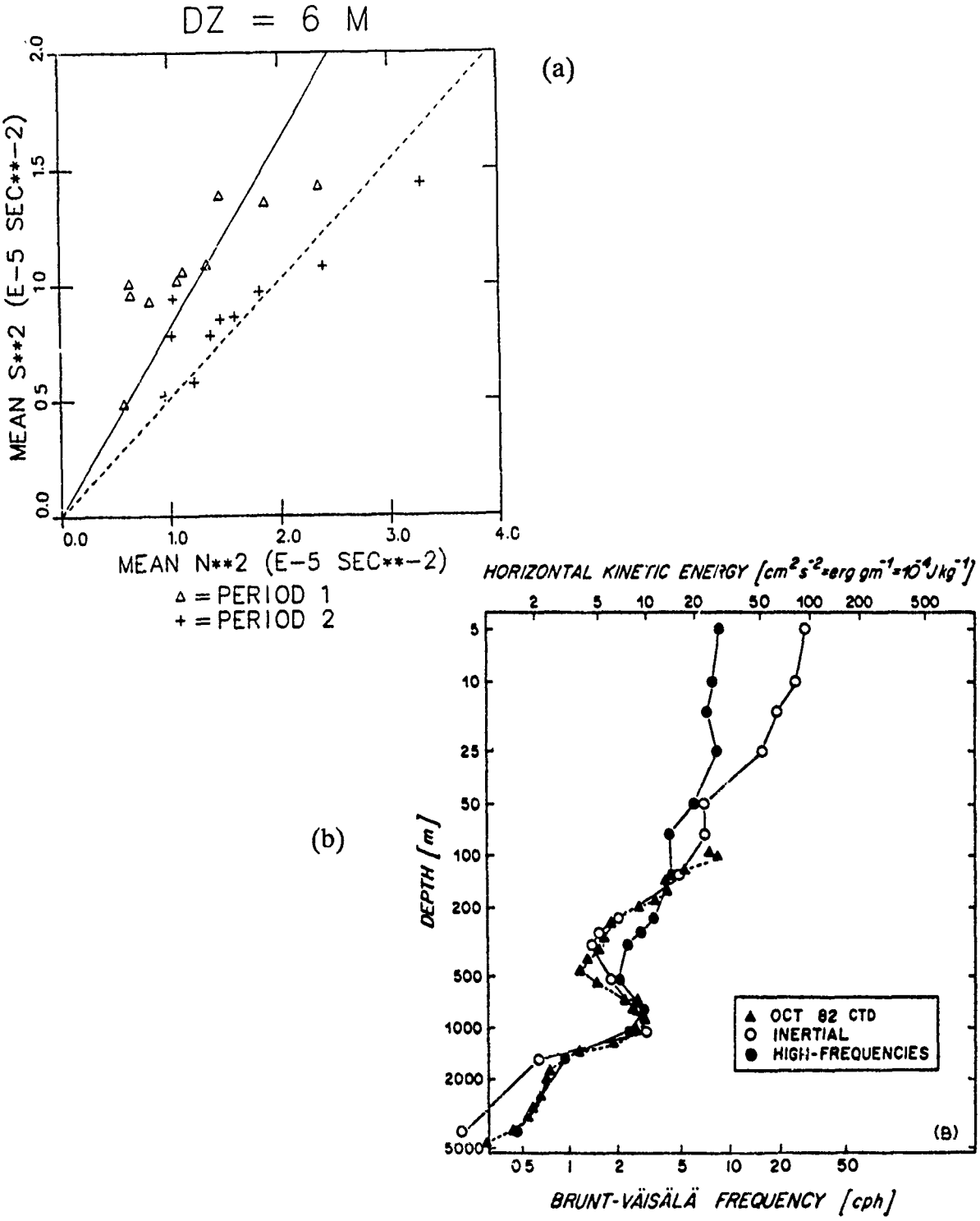


Figure 14. (a) The mean-square shear across 6m intervals vs. BV frequency during period 1 (31 July to 9 August 1978) and period 2 (23 Aug to 6 Sept 1978) (Rubinstein, 1984).
(b) Profiles of inertial and high-frequency internal wave energy together with the BV profile for October 1982 in the Hatteras abyssal plain, showing the WKB scaling of energy (Briscoe and Weller, 1984).

REFERENCES

Boyd, J.D. and H. Perkins, 1987: Characteristics of thermohaline steps off the northwest coast of South America. *Deep-Sea Res.* 337-364.

Brown, E.D. and W.B. Owens, 1981: Observations of the Horizontal Interactions between the Internal Wave Field and the Mesoscale Flow, *J. Phys. Oceanogr.* 11:1474-1480.

Briscoe, M., 1983: Observations on the energy balance of internal waves during JASIN, *Phil. Trans. Royal Soc. London*, Ser A: 308, p. 427-444.

Briscoe, M.G. and R.A. Weller, 1984: Preliminary Results from the long-term upper ocean study (LOTUS), *Dynamics of Atmospheres and Oceans* 8:243-265.

Clancy, R.M. and K.D. Pollak, 1983: A Real-Time Synoptic Ocean Thermal Analysis/Forecast System, *Prog. Oceanogr.* 12:383.

Clancy, R.M., 1987: Real-Time Applied Oceanography at the Navy's Global Center, *Mar. Tech. Soc. Journal* 21:33.

Clancy, R.M. and P.J. Martin, 1981: Synoptic Forecasting of the Oceanic Mixed Layer Using the Navy's Operational Environmental Data Base: Present Capabilities and Future Applications, Bull. AMS 62, p. 770.

Garrett, C. and W. Munk, 1975: Space-Time Scales of Internal Waves: a Progress Report, *J. Phys. Oceanogr.* 80:291-297.

Garwood, R.W., 1977: An oceanic mixed-layer model capable of simulating cyclic states, *J. Phys. Oceanogr.* 7:455-468.

Haney, R.L., 1974: A numerical study of the response of an idealized ocean to large-scale fluxes of heat and momentum flux, *J. Phys. Oceanogr.* 4:145-167.

Hopkins, T.S., P. Povero and S. Piacsek, 1989: Upper layer environmental parameters from CTD data -GIN '86 Cruise, SACLANT Research Center memorandum SM-225, December 1989.

Jerlov, N.G., 1968: *Optical Oceanography*, Elsevier Press, New York, 352 pp.

Martin, P.J., 1982: Mixed-Layer Simulation of Buoy Observations Taken During Hurricane Eloise, *J. Geophys. Res.* 87:409-427.

Martin, P.J., 1985: Simulation of the Mixed Layer at OWS November and Papa with Several Models, *J. Geophys. Res.* 90:903-916.

Mellor, G.L. and T. Yamada,, 1974: A hierarchy of turbulence closure models for planetary boundary layers, *J. Atmos. Sci.* 31:1791-1806.

Mellor, G.L. and Durbin, P.A., 1975: The structure and dynamics of the ocean mixed layer, *J. Phys. Oceanogr.* 5:718-728.

Niiler, P.P., 1975: Deepening of the wind-mixed layer, *J. Mar. Res.* 33:405-422

Niiler, P.P. and E.B. Kraus, 1977: One-dimensional models of the upper ocean. In: *Modeling and prediction of the upper layers of the ocean.* E.B. Kraus, ed., Pergamon Press, Oxford, p. 143.

Piacsek. S.A. and G.O. Roberts, 1975: Numerical Experiments on Collapsing Wakes in a Stratified Fluid, NRL Tech. Mem. #3178, 67 pp.

Piacsek, S.A., L. Henderson, P. VanMeurs, A. Warn-Varnas, and J. French, 1988: *Operational Prediction of Upper Ocean Environmental Parameters*, NORDA Tech. Note 379, April 1988, 102 pp.

Pollard, R.T. and R.C. Millard, 1970: Comparison between observed and simulated wind-generated inertial oscillations. *Deep-Sea Res.* 17:813-821.

Pollard, T.R., P.B. Rhines, and R.O.R.Y. Thompson, 1973: The deepening of the wind driven mixed layer, *Geophys. Fluid Dyn.* 3:381-404.

Price, J.F., 1981: Upper Ocean Response to a Hurricane, *J. Phys. Oceanogr.* 11:153-175

Price, J.F., R.A. Weller and R. Pinkel, 1986: Diurnal Cycling: Observations and Models of the Upper Ocean Response to Diurnal Heating, Cooling and Wind Mixing. *J. Geophys. Res.* 91:8411-8425.

Robinson, A.R., J.A. Carton, N. Pinardi, and C.N.K. Mooers, 1986: Dynamical Forecasting and Dynamical Interpolation: An Experiment in the California Current, *J. Phys. Oceanogr.* 16:1561-1579.

Ruddick, B.R. and T.M. Joyce, 1979: Observations of Interaction between the Internal Wavefield and Low-Frequency flows in the North Atlantic, *J. Phys. Oceanogr.* 9:498-517.

Rubenstein, D.M., 1984: Upper Ocean Shear During JSAIN: A Deterministic and Statistical Analysis, Science Applications, Inc. Tech. Rep. SAI-84/1027, Feb. 1984.

Sellschopp, J., 1987: Towed Thermistor Chain Data collected during the cruise "NORDMEER 87", FWG Report 1987-4, Forschungsanstalt der Bundeswehr fur Wasserschall-und Geophysik, Kiel, Germany.

Thompson, R.O.R.Y., 1976: Climatological numerical models of the surface layers of the ocean. *J. Phys. Oceanogr.* 6:496-503.

Warn-Varnas, A., G.M. Dawson, and P.J. Martin, 1981: Forecast and Studies of the Oceanic Mixed Layer During the MILE Experiment, *Geophys. Astrophys. Fluid Dyn.* 17:63-85.

Warn-Varnas, A., R.M. Clancy, M.L. Morris, P.J. Martin, and S. Horton, 1984: Studies of Large-scale Thermal Variability with Synoptic Mixed-Layer Model, in *Predictability of Fluid Motions*, G. Holloway, ed., American Institute of Physics, Conf. Proc. #106, p. 515-535.

INTERNAL GRAVITY WAVES AND MIXING

Peter Müller
Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, HI 96822

Eric D'Asaro
Applied Physics Laboratory, University of Washington, Seattle, WA 98105-6698

Greg Holloway
Institute of Ocean Sciences, Sidney BC V8L 4B2

Oceanic internal gravity waves span a range between mesoscale eddies and small-scale turbulence, providing an important link in the overall energy cascade from the large scales of generation to the small scales of dissipation. To discuss progress at understanding internal-wave dynamics, the sixth 'Aha Huliko'a Hawaiian Winter Workshop brought together oceanographers and meteorologists. Conclusions of the workshop are

- Considerable progress is being made in predicting the diapycnal mixing rates associated with the shear instability of small-scale internal waves (Figure 1).

- Large-scale internal waves are forced by the atmosphere and (possibly) other oceanic flows. These waves are the most energetic and can propagate over distances of O(1000 km).

- Intermediate-scale waves couple the large-scale waves to the scale waves by a spectral cascade. There exist a good kinematic description of these waves (Garrett-Munk model) and a reasonable account of their dynamics (nonlinear interactions).

- It is timely to proceed toward a global model to predict the internal wave field and diapycnal mixing.

This report discusses these conclusions in more detail.

Parameterization of Dissipation

Gregg (1989) suggested a parameterization of the rate of kinetic energy dissipation in terms of the internal-wave 10-m vertical shear and the buoyancy frequency, which is consistent with McComas and Müller's (1981) and Henyey et al.'s (1986) theoretical predictions of the energy cascade rate to high wavenumbers. This parameterization has considerable skill even in complex environments, as seen in Figure 2, which compares predicted with observed dissipation rates in the Florida Straits (M. Gregg). Also, the large

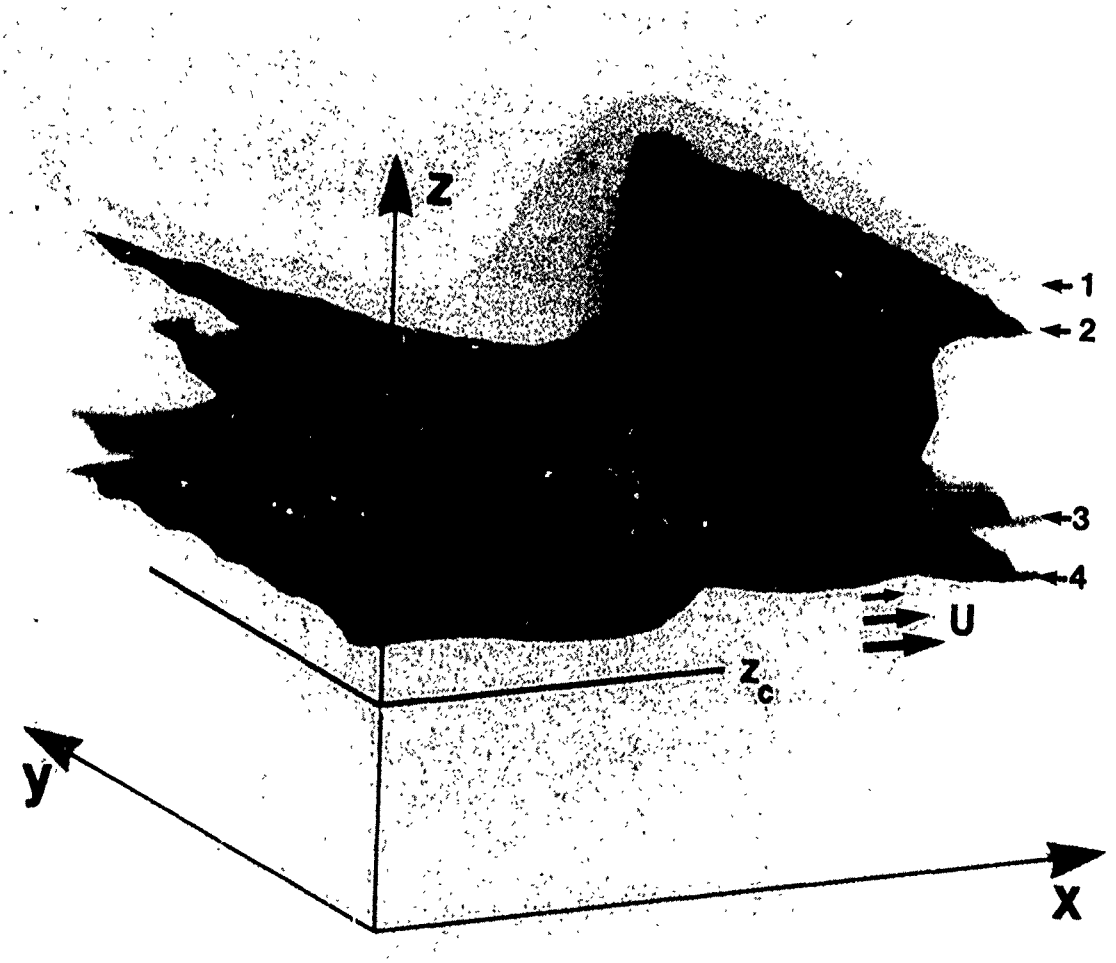


Figure 1. Visualization of an internal wave packet breaking at a critical level. Shown are four isopycnal surfaces from a numerical simulation in which a wave packet propagates downward through a horizontal shear flow $U(z)$ and encounters a critical level z_c. Near z_c, regions of high shear and overturned isopycnals develop. The wave eventually "breaks", owing to a three-dimensional instability. The results of this and other simulations are being used to investigate diapycnal mixing in the deep ocean. (Figure courtesy of Winters and D'Asaro, Applied Physics Laboratory, University of Washington.)

diapycnal diffusivity inferred from a tracer experiment in the Santa Monica Basin implies an elevated level of internal-wave shear that was indeed found (M. Gregg, E. Kunze). On the other hand, Gregg's scaling law has been questioned by Gargett (1990) and regions exist where the scaling is not observed, such as the Yermak Plateau in the Arctic Ocean where dissipation rates much higher than predicted have been found (L. Padman).

The observed scaling of the dissipation rate in terms of internal-wave parameters provides a stringent test of nonlinear wave-wave interaction theories.

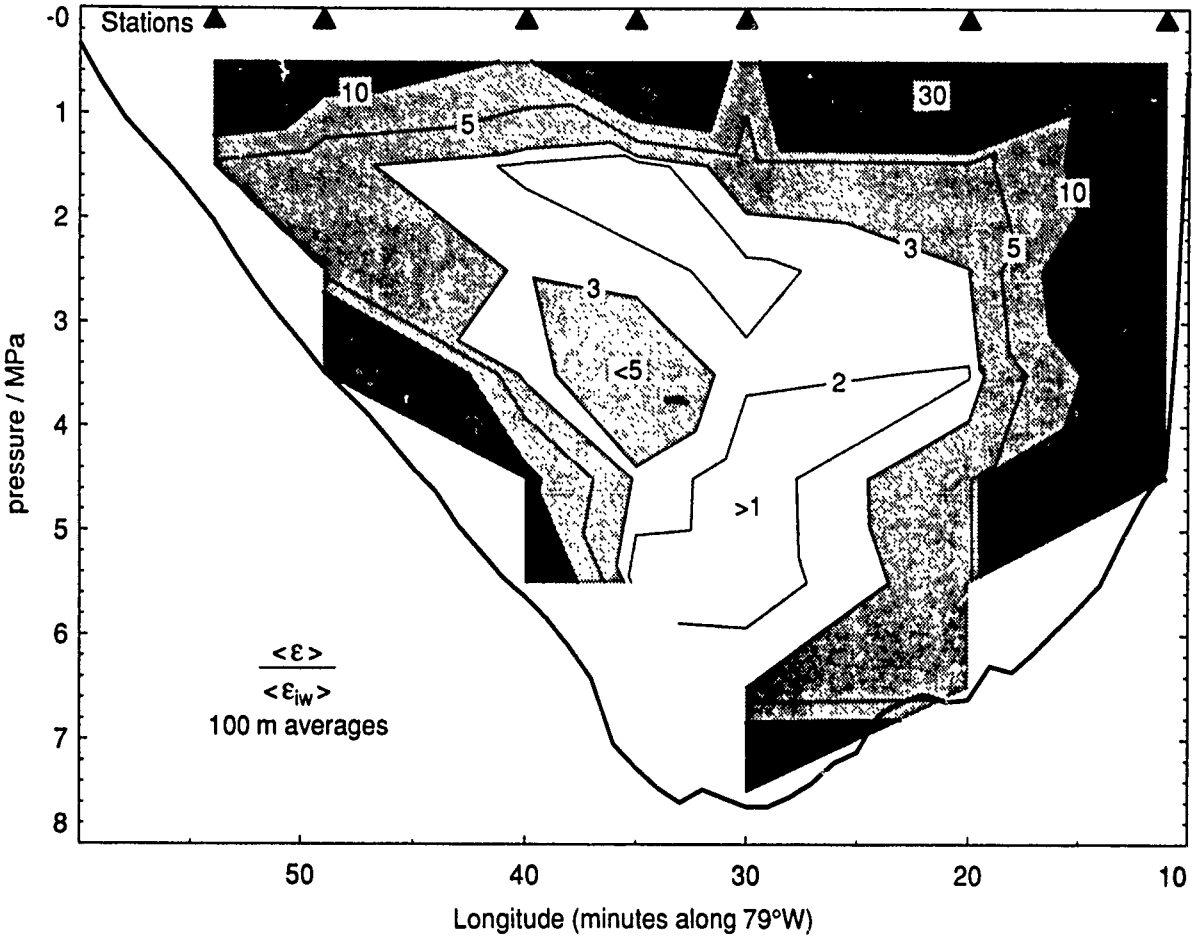


Figure 2. Contours of normalized kinetic energy dissipation rate along a section through the Florida straits (along 28° N from West Palm Beach to Little Bahama Bank). In the interior, the observed rates <ε> are comparable to the rates $<\varepsilon_{iw}>$, which are inferred from internal wave parameters using Gregg's (1989) scaling law. The higher values around the periphery are likely due to interactions with topography. (Courtesy of M. Gregg, University of Washington.)

The Saturation Range

Existing parameterizations, such as that of Gregg (1989), bridge the gap from small-scale internal waves of O(10 m) wavelengths to the dissipation scales in the centimeter range (Figure 3). Considerable progress has been made in describing these small-scale internal waves, which are not represented by the Garrett and Munk spectral model. First, a universal "saturation" range with a -1 spectral slope in the shear or strain spectra is observed between the roll-off wavenumber m_u and the Ozmidov or buoyancy wavenumber m_B. The spectral level is proportional to the buoyancy frequency N squared. Above and below this saturation range, the spectrum is much more variable. Second, the low

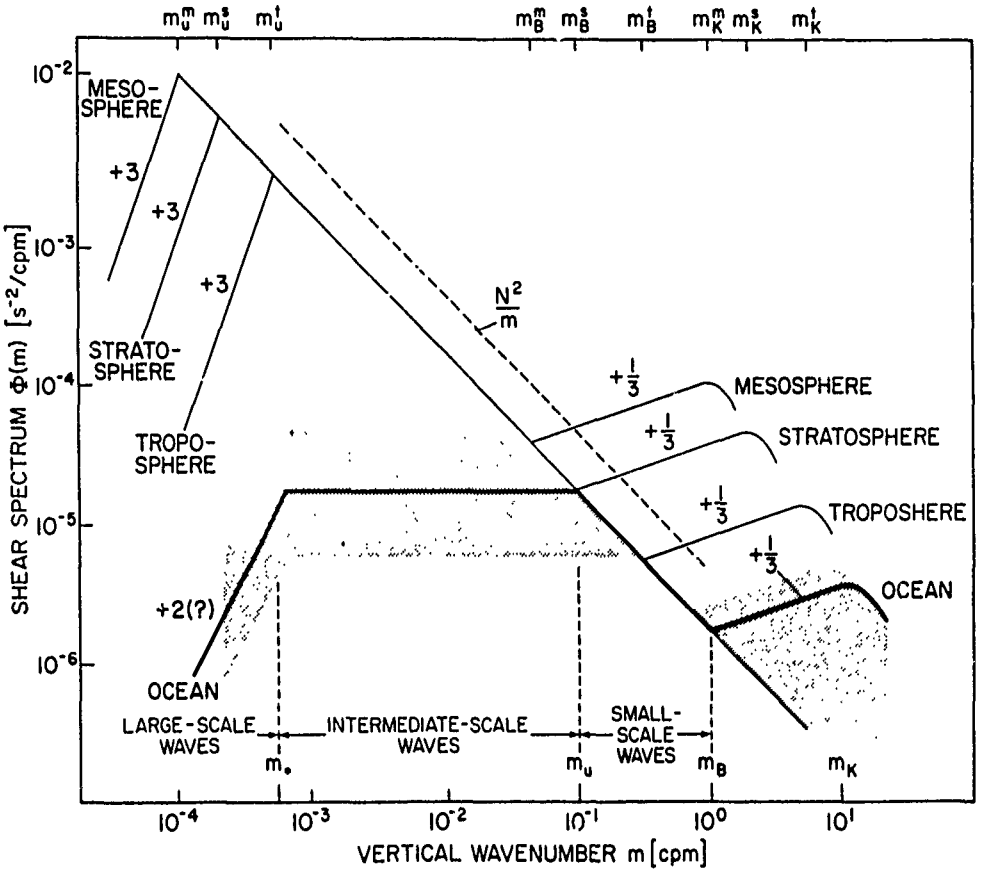


Figure 3. Schematized vertical wavenumber spectra of the vertical shear of internal gravity waves in the ocean, troposphere, stratosphere, and mesosphere, scaled to a common value of the buoyancy frequency N. The wavenumber power law ranges are indicated. The transitional wavenumbers for the ocean are the bandwidth m*, the cut-off or roll-off wavenumbers m_u, the buoyancy or Ozmidov wavenumber $m_B = (N^3/\varepsilon)^{1/2}$, and the Kolmogorov dissipation wavenumber $m_K = (\varepsilon/\upsilon^3)^{1/4}$ with ε being the kinetic energy dissipation rate and the molecular viscosity υ. The analogous wavenumbers for the tropo-, strato-, and mesosphere are indicated by superscripts t, s and m, respectively. The oceanic spectrum consists of a large-scale part ($m < m^*$) which is not well established, an intermediate- scale part ($m^* < m < m_u$) which is well described by the Garrett and Munk spectral model, a small-scale or "saturation" range between m_u and m_B with a -1 slope, and an inertial and viscous dissipation range for $m > m_B$. The variability of the spectrum is indicated by the shading. The "saturation" range is much less variable than the other ranges. The atmospheric spectra also show a saturation range with the same spectral slope and level as the oceanic spectra. The line N^2/m represents theoretical predictions by Lumley (1964) and Holloway (1983) for buoyant turbulence or nonlinear wave interactions.

wavenumber ($m < m_u$), i.e., Garrett and Munk, spectral level and the roll-off wavenumber m_u vary in such a way as to maintain the universal saturation range (M. Gregg). Third, the probability distribution function of the observed strain field is well represented by a gamma distribution. Moments at different scales are thus related to a single (dimensional) parameter (R. Pinkel).

Atmospheric spectra have also been included in Figure 3 because they share the "saturation" slope and spectral level. Atmospheric gravity waves are generated at low altitudes by topography and/or convection. When these waves propagate upward their amplitudes increase due to the decrease in density, and the wave spectrum slides along the -1 slope and "saturates" at the roll-off wavenumber (D. Fritts). Figure 1 shows that atmospheric gravity waves are a narrow-band process with a peak at m_u, whereas oceanic internal waves represent a broad-band process with a roll-off at m_u. Also, atmospheric gravity waves have shorter lifetimes than oceanic waves, and dissipation rates are higher in the atmosphere than in the ocean.

Resonant or eikonal wave-wave interaction calculations predict only the energy transfer from low wavenumbers to the roll-off wavenumber m_u, not the energy transfer across the "saturation" range to the dissipation wavenumbers. The classical "buoyancy subrange" theory of Lumley (1964) predicts the observed N^2/m spectral form on the premise that the buoyancy flux term is dominant in the turbulence kinetic energy equation. The same spectrum is also predicted by Holloway (1983) who assumes that strong wave-wave nonlinearities transfer kinetic and potential energy nearly independently to high wavenumbers. For the atmosphere a wave theory has been advanced that ascribes the "saturation" range to random refractive broadening of an upward propagating band-limited wave field (C. Hines). This wave theory also reproduces the observed spectral slope and level. On the other hand, rapid changes of the atmospheric spectrum during vertical propagation suggest that wave breaking may occur at "saturation" scales (D. Fritts). In the ocean, structures at "saturation" scales often persist for many buoyancy periods; this is indicative of wave rather than turbulent dynamics (R. Pinkel). The distinction between wave and turbulent processes may be one of degree. Numerical simulations show a smooth transition from wave refraction to vigorous overturning with increasing Froude number (D. Ramsden).

Diapycnal Mixing

Internal wave theory predicts the energy flux to mixing scales; observations provide the kinetic energy dissipation rate. Additional assumptions are needed to infer the diapycnal mixing rate. Traditionally, diapycnal mixing is inferred from the buoyancy flux, which is assumed to be equal to the dissipation rate multiplied by a mixing efficiency factor. Therefore, to infer mixing rates from internal wave parameters one needs to quantify the mixing efficiency. An increasingly useful tool to address this problem is direct numerical simulation. Both visualization of the flow field (D. Siegel, K. Winters) and diagnosis of the dynamics are now possible. It appears that spatial separation of sources and sinks, which clearly occurs in the ocean, affects the dynamical balances and must be taken into account (D. Ramsden). Multiple scale interactions and separated source and sink regions require extreme numerical resolution. To infer mixing rates one must also carefully separate changes in available and base-state potential energy (K. Winters).

Overall, present observations and theory indicate a typical diapycnal diffusivity near 10^{-5} m^2s^{-1} in the upper ocean and thermocline, with much larger values in some areas due to local wave generation (L. Padman), storm-induced mixing (C. Eriksen), or double diffusive processes (M. Gregg). This value is generally consistent with estimates from inverse models, although in some places these estimates are subject to great uncertainties due to incorrect formulation and neglect of the nonlinearities in the equation of state (T. McDougall).

Whereas the internal wave-induced diapycnal diffusivity that describes the fluxes of buoyancy and tracers across isopycnals is being quantified, there is little information about the internal wave-induced diapycnal viscosity that describes the momentum flux across isopycnals. Indeed, even the sign of the internal wave or "eddy" Prandtl number is subject to dispute. There is also little information about isopycnal transports induced by internal waves.

Boundary and Abyssal Mixing

Several distinct regions of mixing can be identified in the ocean (C. Garrett). In the ventilated part of the thermocline, vertical transport is primarily caused by advection along isopycnals, and diapycnal mixing of $O(10^{-5}$ $m^2s^{-1})$ might perhaps be secondary. Most of our measurements are from this region or from the even more variable upper ocean. In contrast, in the unventilated abyssal ocean mixing processes are likely to be more important. There are few internal-wave measurements in the deep ocean. The few that have been made indicate that the spectral distribution of internal-wave energy might be different there (T. Sanford). Characteristics of the internal-wave field in shallow seas have also not been established yet.

A special mixing regime may occur in the vicinity of benthic boundaries. Understanding such boundary mixing is important for the parameterization of mixing and the formulation of boundary conditions in large-scale numerical models. Internal waves might play a very important role in boundary mixing because both the reflection at critical angle slopes (Eriksen, 1985) and the scattering at rough bottom topography (N. Xu) result in a transfer of energy to high wavenumber waves that are more likely to break and mix. A careful search for signatures of critical angle reflection at two locations has yielded frustratingly ambiguous results (D. Gilbert).

The stratification of water mixed at the boundary is a key factor in determining how efficiently the surrounding ocean is mixed. Simple models have been developed for the restoration of stratification by buoyancy-driven secondary flows and for the net buoyancy flux (C. Garrett).

Potential Vorticity

Potential vorticity is a conserved quantity that is not carried by internal gravity waves. Theoretically we expect non-wave motions, often called "vortical" motions, to carry potential vorticity at internal-wave scales. Vortical motions can interact with internal waves (J. Riley) and may thus affect the internal-wave field. Potential vorticity

conservation expresses the invariance of fluid motion under fluid particle relabeling. A formulation of the equations of motion in terms of truncated Hamiltonian equations that exhibit this invariance has been achieved (H. Abarbanel).

Attempts to observe small-scale potential vorticity carrying motions have yielded ambiguous results. Analysis of IWEX data shows an excess of vertical relative vorticity at 25 m horizontal separation above that expected for internal waves but no such excess at 250 m (R. C. Lien). Analysis of data taken downstream of Ampere Seamount shows no contribution to potential vorticity fluctuations from relative vorticity but only from vortex stretching (E. Kunze).

Much of the velocity and density finestructure often attributed to vortical motion can be explained equally well as distortions of an internal-wave field that is vertically advected and strained. The use of vertically Lagrangian (isopycnal following) coordinates that eliminate the effects of vertical advection might simplify the kinematical description and dynamical evolution of internal waves (F. Henyey).

Forcing

Various sources have been suggested for the oceanic internal-wave field. Fluctuations in the windstress excite mixed-layer inertial oscillations and a large fraction of energy from these oscillations penetrates the ocean as near inertial, low-modenumber internal waves. Although the details of this process can be complicated by mesoscale eddies and storm-induced mixing (C. Eriksen) there is little reason to doubt that windstress fluctuations are a major source of large-scale internal waves (E. D'Asaro).

Internal waves of tidal frequency are commonly observed in the ocean and may also be a major energy source. Several clear examples of generation by the barotropic tide have been found. However, the amplitude and propagation direction of the internal tide are highly variable and this variability is not well understood (M. Levine).

Surface waves have also been suspected as being a source for internal waves. Recent calculations (K. Watson) show, however, that surface waves draw energy from the high-frequency, low-modenumber internal-wave field, except during periods of very strong surface waves. These energy transfers indicate that imbalances between the internal and surface wave fields may relax rapidly.

Despite the large energy contained in mesoscale low frequency motions there is little evidence that these motions are a major source of energy for the internal-wave field. Refraction and the resulting trapping of waves in frontal or eddy features can lead to strong localized wave dissipation and mixing (Kunze, 1985). Similar interactions in a weaker eddy field have been predicted to result in a net energy flux from the eddy field to the internal-wave field (Watson, 1985). Recently, radiation of internal waves from newly formed small coherent eddies has been observed (T. Sanford, R. Pinkel).

A Global Internal Wave Model

Sufficient progress in theory and observations has been made to outline the structure of a global model to predict the internal-wave field and diapycnal mixing. Such a model would follow the propagation of large-scale internal waves from their sources, with transfer of energy across the time-variable Garrett and Munk spectrum to its local dissipation. To achieve the understanding of the physics necessary for the construction of such a model the following issues need to be addressed:

- *Generation*: Wind and tides may be sufficient to drive the global internal-wave field. The wind source function is limited by the resolution of wind data and the accuracy of the drag coefficient. Nevertheless, the general spatial distribution and seasonal cycle can perhaps be estimated from current operational weather products. Presently, we have less confidence in the construction of the source function for internal tides.

- *Propagation*: Estimates of the mean free path from group velocity and nonlinear interaction rates indicate that wind-generated, large-scale waves propagate significant distances, O(1000 km), from their sources (E. D'Asaro, E. Hirst). The generation and propagation of large-scale waves might thus be a basically linear problem (D. Rubenstein) and can perhaps be modeled by ray tracing. The interaction of these large-scale waves with topography and the mesoscale eddy field remain challenging research areas.

- *Spectral Transfer*: We anticipate that resonant interaction theory will predict the transfer of energy from the propagating large-scale waves to the Garrett and Munk spectrum with sufficient accuracy (E. Hirst). The necessary calculations have not yet been done for a highly directional wave field. Energy transfer by scattering at topography is probably important as well, but has also not been investigated yet. The subsequent energy transfer across the Garrett and Munk spectrum, its spectral shape, and the final dissipation and diapycnal mixing rate will be controlled by more highly nonlinear interactions. To test the accuracy of nonlinear intereaction theory, comparisons with direct numerical simulations need to be carried out.

- *Dissipation*: We also anticipate that present research linking dissipation and its associated mixing to internal-wave parameters at the roll-off wavenumber m_u will converge on a parameterization with significant skill. This area remains a high priority for research since it will likely result in the ability to estimate diapycnal mixing from routine internal wave measurements. However, such measurements alone will be insufficient to prescribe mixing rates in large-scale models. A global internal wave model is needed.

To make progress on these issues requires coordinated theoretical and observational studies of specific processes. We suggest as priority issues the quantification of internal-wave sources, of the propagation of large-scale waves from their sources, and of the transfer of energy from these waves to the rest of the spectrum.

Direct numerical simulation will be a powerful tool in this effort. Although the dynamics are clearly three-dimensional (3D), some processes might be modeled in two dimensions (2D, vertical plane) affording higher resolution. By careful intercomparison of 2D and 3D models, the range of applicability of 2D can be addressed (D. Ramsden).

Observationally, the measurement of the directional spectrum of large-scale waves requires techniques that integrate vertically (or horizontally). Acoustics is a natural candidate for this (T. Ewart). Measuring the directionality of large-scale waves away from obvious sources may yield clues about the location of sources.

Finally, comparison of data and model prediction might require multivariate analysis to arrive at definite conclusions (C. Frankignoul). Actual predictive models will undoubtedly also need to be integrated with models of other oceanic processes that affect the internal-wave field (A. Warn-Varnas).

Conclusions

Internal waves and their effect on larger scales remain a basic and important issue. Recent progress suggests that we have the tools and conceptual framework to predict diapycnal diffusivities globally in the not too distant future. Such optimism is not warranted for the prediction of internal-wave induced momentum fluxes and isopycnal dispersion. In these problems the vortical mode may turn out to be important. Present understanding is focused on the upper ocean and thermocline. We are less clear about internal-wave behavior in shallow seas, the abyss and near boundaries.

Acknowledgments

The names in parentheses refer to authors of lectures given at the meeting and chapters published in the proceedings. We thank the participants of the workshop for their input into this report and for their permission to quote unpublished material. Crystal Miles is thanked for expert editorial advice. Copies of the proceedings are available from Peter Müller, Department of Oceanography, University of Hawaii, 1000 Pope Road, MSB 307, Honolulu, HI 96822. The sixth 'Aha Huliko'a Hawaiian Winter Workshop was held January 15-18, 1991, in Honolulu and is supported by Department of Navy Grant N00014-90-J-4010 issued by the Office of Naval Research. The U.S. government has a royalty-free license throughout the world in all copyrightable material contained herein. This report is School of Ocean and Earth Science and Technology contribution 2622.

References

Eriksen, C. C., 1985: Implications of ocean bottom reflection for internal wave spectra and mixing. *J. Phys. Oceanogr.*, **15**, 1145–1156.

Gargett, A. E., 1990: Do we really know how to scale the turbulent kinetic energy dissipation rate due to breaking of oceanic internal waves? *J. Geophys. Res.*, **95**, 15,971–15,974.

Gregg, M. C., 1989: Scaling turbulent dissipation in the thermocline. *J. Geophys. Res.*, **94**, 9686–9698.

Henyey, F. S., J. Wright and S. M. Flatté, 1986: Energy and action flow through the internal wave field—An eikonal approach. *J. Geophys. Res.*, **91**, 8587–8495.

Holloway, G., 1983: A conjecture relating oceanic internal waves and small-scale processes. *Atmosphere; Atmos. Ocean*, **21**, 107–122.

Kunze, E., 1985: Near inertial wave propagation in geostrophic shear. *J. Phys. Oceanogr.*, **15**, 544–565.

Lumley, J. L., 1964: The spectrum of nearly inertial turbulence in a stably stratified fluid. *J. Atmos. Sci.*, **21**, 99–102.

McComas, C. H. and P. Müller, 1981: The dynamic balance of internal waves. *J. Phys. Oceanogr.*, **11**, 970–986.

Watson, K. M., 1985: Interaction between internal waves and mesoscale flow. *J. Phys. Oceanogr.*, **15**, 1296–1311.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION School of Ocean and Earth Science and Technology 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Office of Naval Research
6c. ADDRESS (City, State, and ZIP Code) Department of Oceanography, Univ. of Hawaii 1000 Pope Road Honolulu, HI 96822	7b ADDRESS (City, State, and ZIP Code) Department of the Navy 800 N. Quincy Street Arlington, Virginia 22217
8a NAME OF FUNDING / SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-90-J-4010
8c. ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO 4222922-02 / TASK NO / WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)

Dynamics of Oceanic Internal Gravity Waves

12 PERSONAL AUTHOR(S)

Müller, Peter and Henderson, Diane (eds.)

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Workshop Proceedings	FROM 7/1/90 TO 12/31/9	November 1991	520

16 SUPPLEMENTARY NOTATION

Proceedings, 'Aha Huliko'a, Hawaiian Winter Workshop, January 1991, Honolulu, Hawaii

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	internal gravity waves, inertial oscillations, internal tides, finestructure, vortical motion, nonlinear interaction, mixing

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

These proceedings contain the lectures given at the sixth 'Aha Huliko'a Hawaiian Winter Workshop on "Dynamics of oceanic internal gravity waves" and a summary of the workshop, which appeared in Eos, Transactions, American Geophysical Union. The lectures and the summary cover the major aspects of internal wave dynamics and internal wave induced mixing.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL	22b TELEPHONE (Include Area Code) / 22c OFFICE SYMBOL