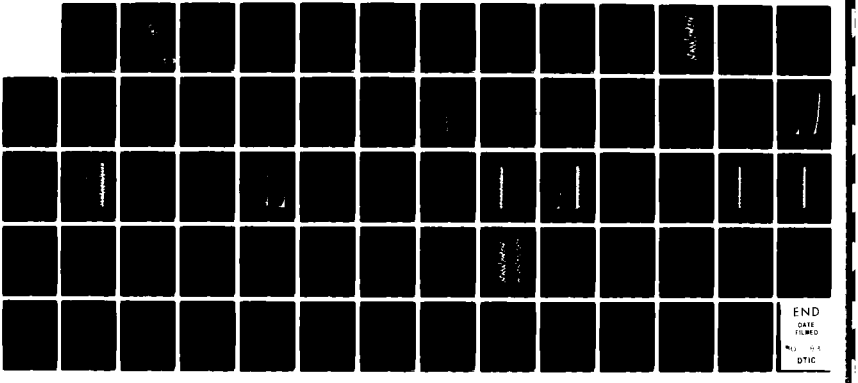
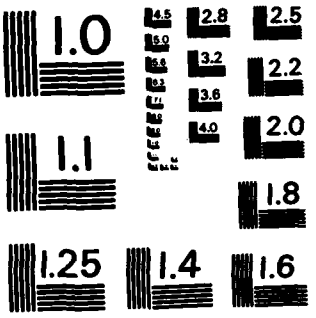
AD-A132 619 ANALYSIS OF A 12-YEAR RECORD OF SEA-SURFACE TEMPERATURES OFF PT SUR CALIFORNIA(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA L C BREAKER ET AL. 1/1

UNCLASSIFIED JUN 83 NPS55-83-018 F/G 8/10 NL

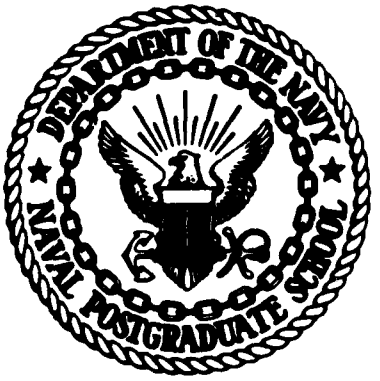




NPS55-83-018

NAVAL POSTGRADUATE SCHOOL

Monterey, California



ANALYSIS OF A 12-YEAR RECORD OF SEA-SURFACE TEMPERATURES OFF PT. SUR, CALIFORNIA

by

L. C. Breaker
P. A. W. Lewis
E. J. Orav

June 1983

Prepared for: Chief of Naval Research
Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral J. J. Ekelund
Superintendent

D. A. Schrady
Provost

This work was supported in part by the Office of Naval Research under Grants RR031-03-01; RR033-02-0K and 083-275(647-001X) and RR014-05-01.

L. C. Breaker
Oceanographer
Department of Oceanography

P. A. W. Lewis
Professor
Department of Operations Research

E. J. Orav
Adjunct Professor
Department of Operations Research

Reviewed by:

Alan R. Washburn, Chairman
Department of Operations Research

Released by:

William M. Tolles
Dean of Research

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: NPS55-83-018	2. GOVT ACCESSION NO.: AD-A132619
	3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): ANALYSIS OF A 12-YEAR RECORD OF SEA-SURFACE TEMPERATURES OFF PT. SUR, CALIFORNIA	5. TYPE OF REPORT & PERIOD COVERED: Technical
	6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): L. C. Breaker; P. A. W. Lewis; E. J. Orav	8. CONTRACT OR GRANT NUMBER(s):
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, CA 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61153N; RR014-05-01; N0001483WR30026
11. CONTROLLING OFFICE NAME AND ADDRESS: Chief of Naval Research, Arlington, VA 22217	12. REPORT DATE: June 1983
	13. NUMBER OF PAGES: 62
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sea-surface Temperatures
Coastal Upwelling
El Nino
Statistical Model
Spectral Analysis
Deterministic Components
Autoregressive Model
Simulation
Prediction

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

Daily observations of sea-surface temperature (SST) taken just north of Pt. Sur, California over the past 12 years are examined for statistical and oceanographic content. Spectral analysis of the data reveals an annual cycle and two harmonics, but no obvious influence from the local winds or tides. A statistical model has been constructed based on a decomposition of the data into its various deterministic and random components. As expected, the model provides forecasts which are significantly better than those that would be obtained from a simple climatological treatment of the data alone. The model

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

is also used to generate simulated series, which in turn are used as a basis to estimate adequate sampling rates.

In 6 years of the 12 year record, an abrupt major decrease in SST was observed between February and April and has been identified as the spring transition to coastal upwelling. One such transition was found to occur almost simultaneously at three locations, including Pt. Sur, along the California coast between Pt. Conception and San Francisco. The data also indicate that a well-defined annual temperature cycle does exist at Pt. Sur although its range is reduced and its maximum occurs later in the year than would normally be expected at this latitude. Three El Nino events are clearly evident in the 12-year record including the present El Nino; SSTs have increased by almost 3C at Pt. Sur as a result of the present El Nino warming. Over the whole 12 year period an average increase in temperature of 1.6C was found.

ANALYSIS OF A 12-YEAR RECORD OF SEA-SURFACE TEMPERATURES OFF PT. SUR, CALIFORNIA

L.C. Breaker P.A.W. Lewis E.J. Orav

Naval Postgraduate School
Monterey, California 93940

ABSTRACT

Daily observations of sea-surface temperature (SST) taken just north of Pt. Sur, California over the past 12 years are examined for statistical and oceanographic content. Spectral analysis of the data reveals an annual cycle and two harmonics, but no obvious influence from the local winds or tides. A statistical model has been constructed based on a decomposition of the data into its various deterministic and random components. As expected, the model provides forecasts which are significantly better than those that would be obtained from a simple climatological treatment of the data alone. The model is also used to generate simulated series, which in turn are used as a basis to estimate adequate sampling rates.

In 6 years of the 12 year record, an abrupt major decrease in SST was observed between February and April and has been identified as the spring transition to coastal upwelling. One such transition was found to occur almost simultaneously at three locations, including Pt. Sur, along the California coast between Pt. Conception and San Francisco. The data also indicate that a well-defined annual temperature cycle does exist at Pt. Sur although its range is reduced and its maximum occurs later in the year than would normally be expected at this latitude. Three El Nino events are clearly evident in the 12-year record including the present El Nino; SSTs have increased by almost 3C at Pt.Sur as a result of the present El Nino warming. Over the whole 12 year period an average increase in temperature of 1.6C was found.

June 20, 1983

ANALYSIS OF A 12-YEAR RECORD OF SEA-SURFACE TEMPERATURES OFF PT. SUR, CALIFORNIA

I. Introduction

Sea-surface temperatures (SSTs) are collected on a regular basis (SIO Reference 81-30) along the California coast at approximately 25 locations. In view of the increasing interest in coastal circulation with respect to the dispersal of pollutants and a continuing interest in fisheries biology, it is surprising that few systematic studies have been conducted to examine these records, particularly within the past 15 years. Perhaps the availability of high-quality temperature data along the continental shelves from new and sophisticated instruments has caused us to overlook this resource in recent years. However, the coastal observations, unlike those acquired from moored arrays, often extend over many years and thus provide a unique opportunity to examine coastal variability over relatively long periods. For locations where the measuring site has a good exposure to the adjacent continental shelf and slope, measurements of SST may be particularly revealing with respect to some of the physical processes that occur regionally as well as locally. In this regard a 12-year record of SSTs off the central California coast at Granite Canyon, 15km south of Monterey, has been chosen for the initial study. These data have been acquired daily by the Marine Culture Laboratory of the California Fish and Game Commission since March of 1971. The Granite Canyon site where the observations are made (see Fig. 1) has an excellent exposure to the deep ocean with the continental shelf extending less than 10km offshore.

Previous investigations along the U.S. West Coast indicate that the oceanographic information content of coastal SST's is usually high. SSTs have been shown (a) to be useful indicators of ocean temperature

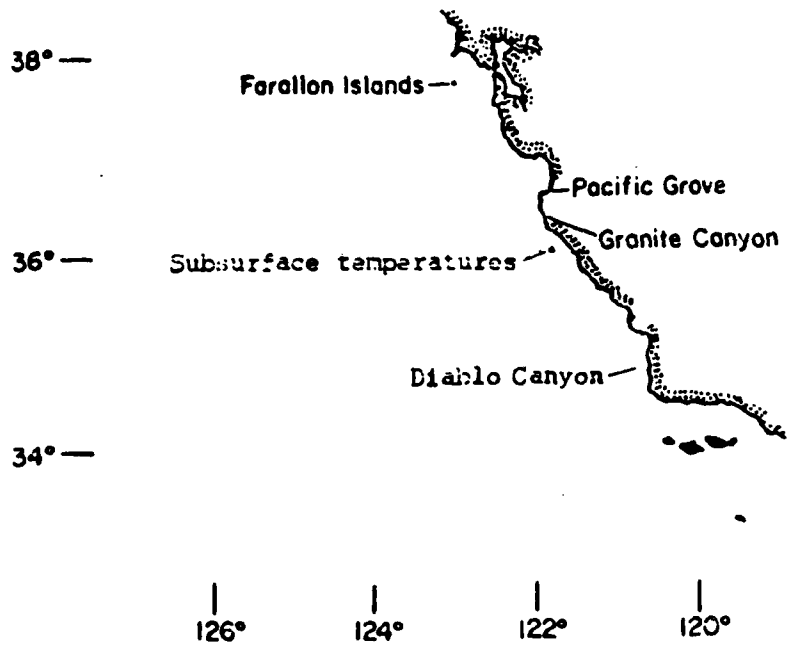


FIGURE 1

Measurement Locations. The principal time series considered here was taken at Granite Canyon. Qualitative comparisons are made to fragmentary series of subsurface temperatures close by, and to a short time series at Diablo Canyon. Comparisons of the Granite Canyon SSTs will be made to extensive time series of SSTs at the Hopkins Marine Station at Pacific Grove and to SSTs at the Farallon Islands.

variability, (b) to be often representative of phenomena occurring over wide regions, (c) to be related to other ocean and atmospheric variables, and (d) to have consistent internal structures. Based on surface and subsurface temperatures along the California coast, Reid, Roden and Wyllie (1958) indicated two different causes of seasonal variability in SSTs. The first is coastal upwelling. North of Pt. Conception the upwelling process decreases the seasonal range in temperature and increases the duration of cooler temperatures. The second cause stems from the importance of the coastal countercurrent. This flow to the northwest transports warm water from the south along the California coast, increasing surface temperatures during the fall. Also, Reid, Roden and Wyllie (1958) found that variations in temperature along the U.S. West Coast show considerable coherence over great distances and that major anomalies in temperature persist for several months.

Using coastal SSTs (and sea level) Stewart, Zetler, and Taylor (1958) verified the existence of anomalous conditions off the U.S. West Coast during 1957 and 1958. Robinson (1957) indicated the value and representativeness of coastal SSTs. She found that coastal SSTs and air temperatures were highly correlated and that SST anomalies identified at coastal locations could be correlated with SST anomalies up to 250km offshore. Roden (1963) found that non-annual extreme temperature fluctuations were coherent and in phase over distances of 200-300km. He also found that the average persistence of non-annual minimum temperatures was 3-4 months while the average persistence of non-annual maximum temperatures was 3-5 months along the coast of northern California. Spectral analysis indicated the presence of an annual cycle only, however.

More recently List and Koh (1976) analyzed SSTs along the California coast over several years using spectral analysis and found that many of the

coastal stations exhibited strong spectral peaks at 14.7 days and that these peaks were probably the result of tidal aliasing and/or the spring and neap tides. They also found that a well-defined seasonal variation in SST was essentially absent between Pt. Conception and Monterey Bay and that certain fluctuations occurred simultaneously at most coastal stations.

Surface and subsurface temperatures across Monterey Bay were collected on a regular basis by the Hopkins Marine Station from 1929 to 1974. Results of these measurements have been reported by Skogsberg (1936), Skogsberg and Phelps (1946), and Bolin and Abbott (1963). Much of the data presented in these references were time-averaged providing monthly mean thermal conditions over the Bay. Skogsberg used these data to define three separate oceanic seasons for the Monterey Bay area.

Our purpose in this paper is threefold: first, to model (i.e. describe mathematically and statistically) the behavior of the 12 year Granite Canyon data set; second, to use the model and other statistical techniques to project or predict the data and estimate time scales of variability in SSTs, and third, to provide a descriptive interpretation of the Granite Canyon data from an oceanographic viewpoint.

2. The Granite Canyon Data

SSTs at Granite Canyon are measured daily at approximately 0800 local time. The raw data are shown in Figure 2. Because the data are taken only once every 24 hours, the possibility of aliasing by the predominant semidiurnal tide exists. This problem will be addressed later in the analysis. As mentioned before the location has an excellent exposure to the deep ocean. Temperatures are read to the nearest 0.1C using a calibrated VWR

RAW SEA-SURFACE TEMPERATURE AT GRANITE CANYON-3/1/71 TO 3/1/83
MEASUREMENTS TAKEN AT 0800 PST

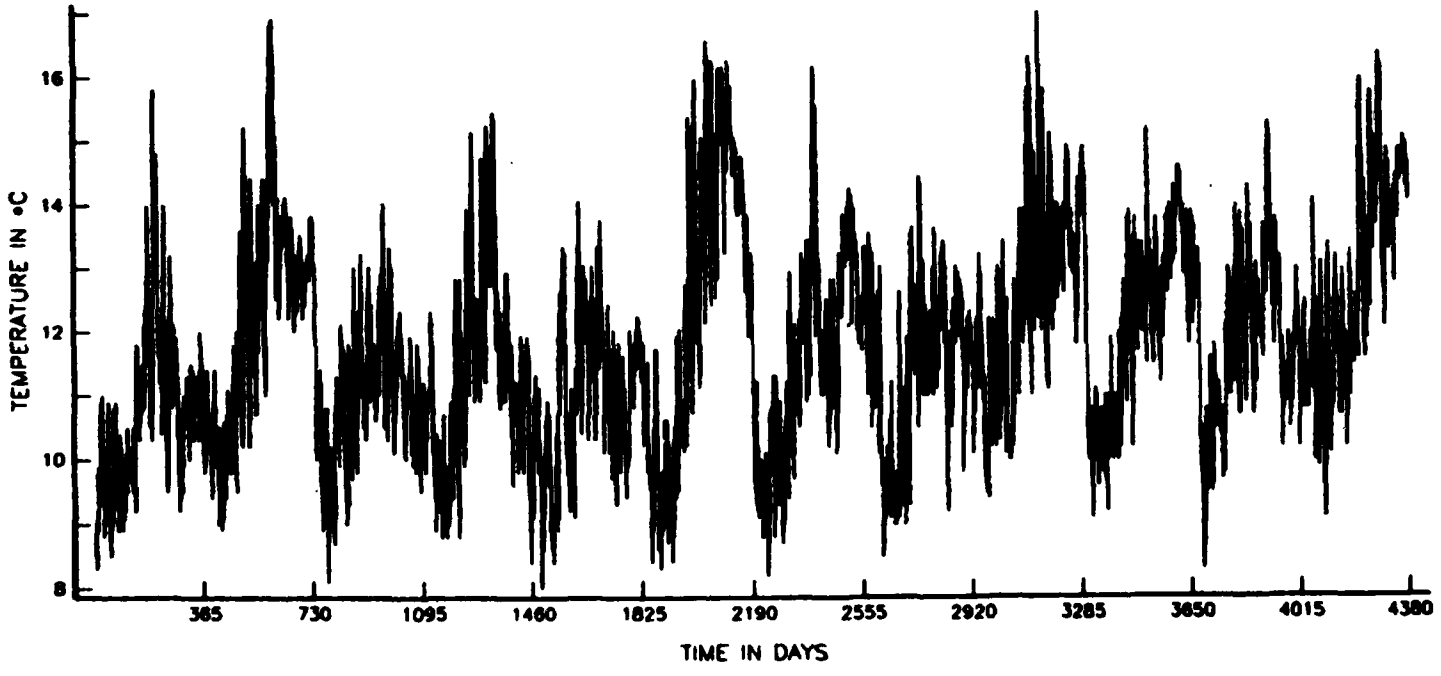


FIGURE 2

thermometer, and are considered accurate to about ± 0.2C (SIO Reference 81-30). The observations span 12 years starting March 1, 1971 and ending on February 28, 1983. SSTs over the same time period were acquired for the Hopkins Marine Station at Pacific Grove and the Farallon Islands (see Figure 1 for location). The accuracy of these data is also about ± 0.2C.

Since the data from Granite Canyon (GC) have only been available since 1971 and since they have not been compared with other coastal temperatures, it was of interest initially to consider the representativeness of these observations. Consequently we have compared the Granite Canyon data with several other data sets for periods where comparisons could be made. The annual variations in SST were generally similar at Granite Canyon, Diablo Canyon and the Farallons, but temperatures were consistently higher at Pacific Grove during the spring and summer months than they were at the other three locations.

Temperature data at four depths between 113 and 510m were acquired from a moored array approximately 50km south of Granite Canyon for the period 03/05/80 to 04/07/80 and compared with the SST data at Granite Canyon, see Figure 3. The period was chosen because it includes a major event tentatively identified as the spring transition to upwelling, a phenomenon which has been recently identified further north, along the coast of Oregon, by Huyer et al, (1979). This event can be readily identified in a number of the annual cycles in the raw data (Figure 2, and Table 1). The drop in temperature associated with this event during March of 1980 can also be seen at each depth in Figure 3, down to 510m.

Figures 4, 5, 6 show comparisons of SSTs between Granite Canyon and Diablo Canyon, between Granite Canyon and Pacific Grove, and between Granite

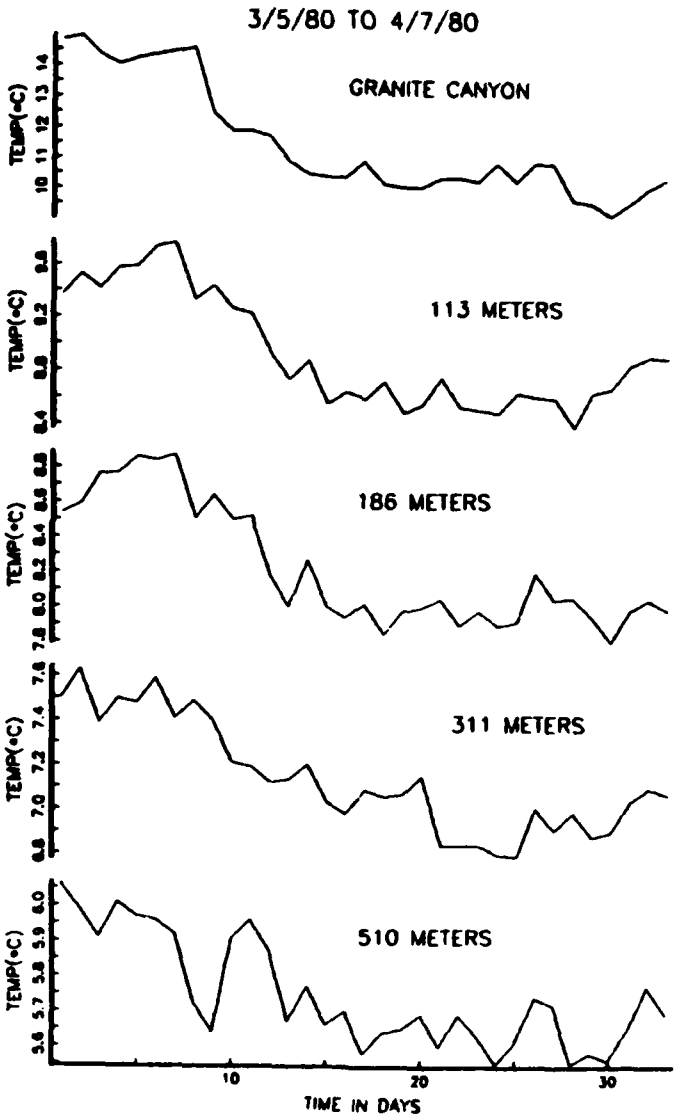


FIGURE 3

Comparison of a section of the Granite Canyon SST data with subsurface temperature data at 4 depths, from a moored array 50 Km south of Granite Canyon. (These data were provided by J.B. Wickham and S.P. Tucker, Oceanography Department, Naval Postgraduate School.)

GRANITE CANYON AND DIABLO CANYON COMPARED

1/1/80 TO 1/1/81

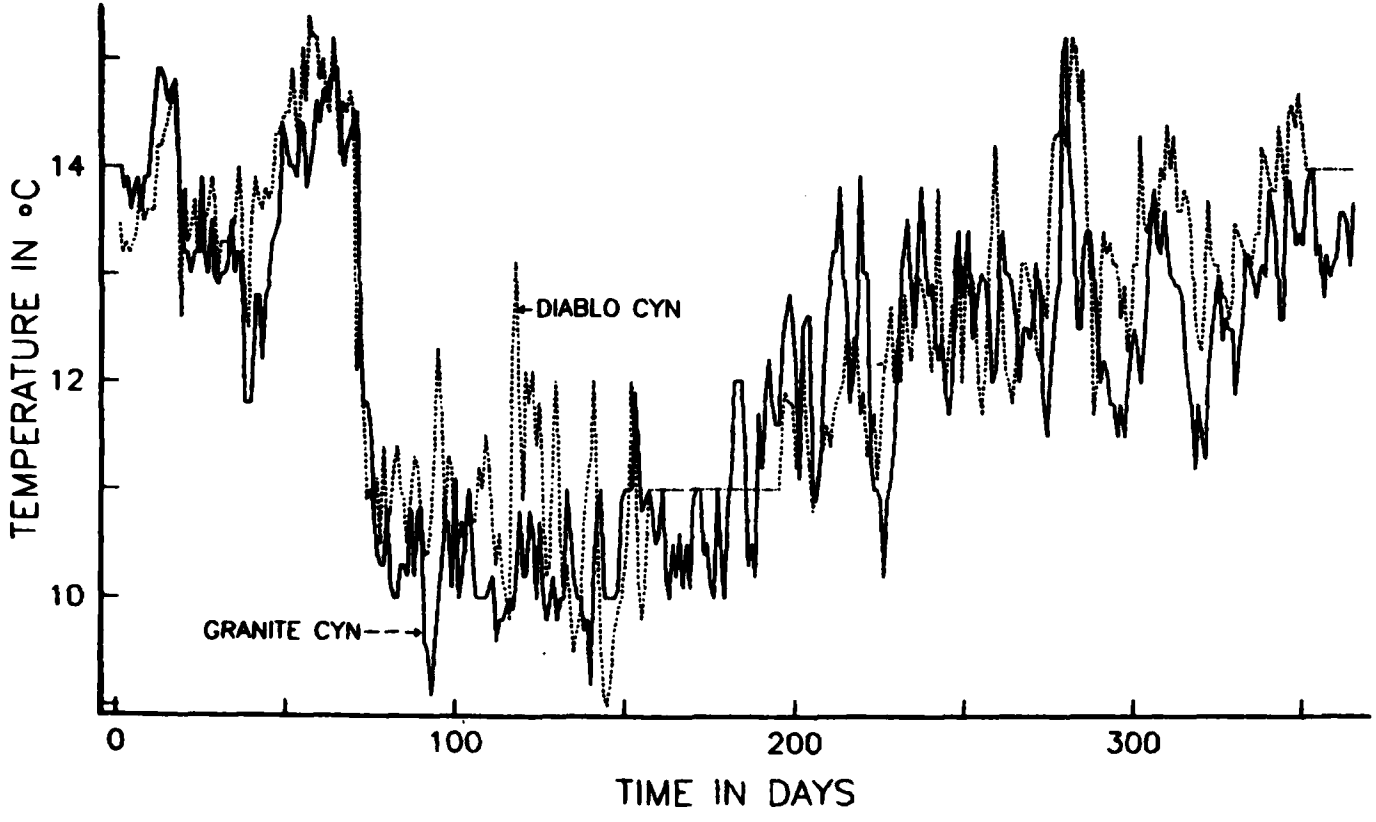


FIGURE 4

GRANITE CANYON AND PACIFIC GROVE COMPARED

1/1/80 TO 1/1/81

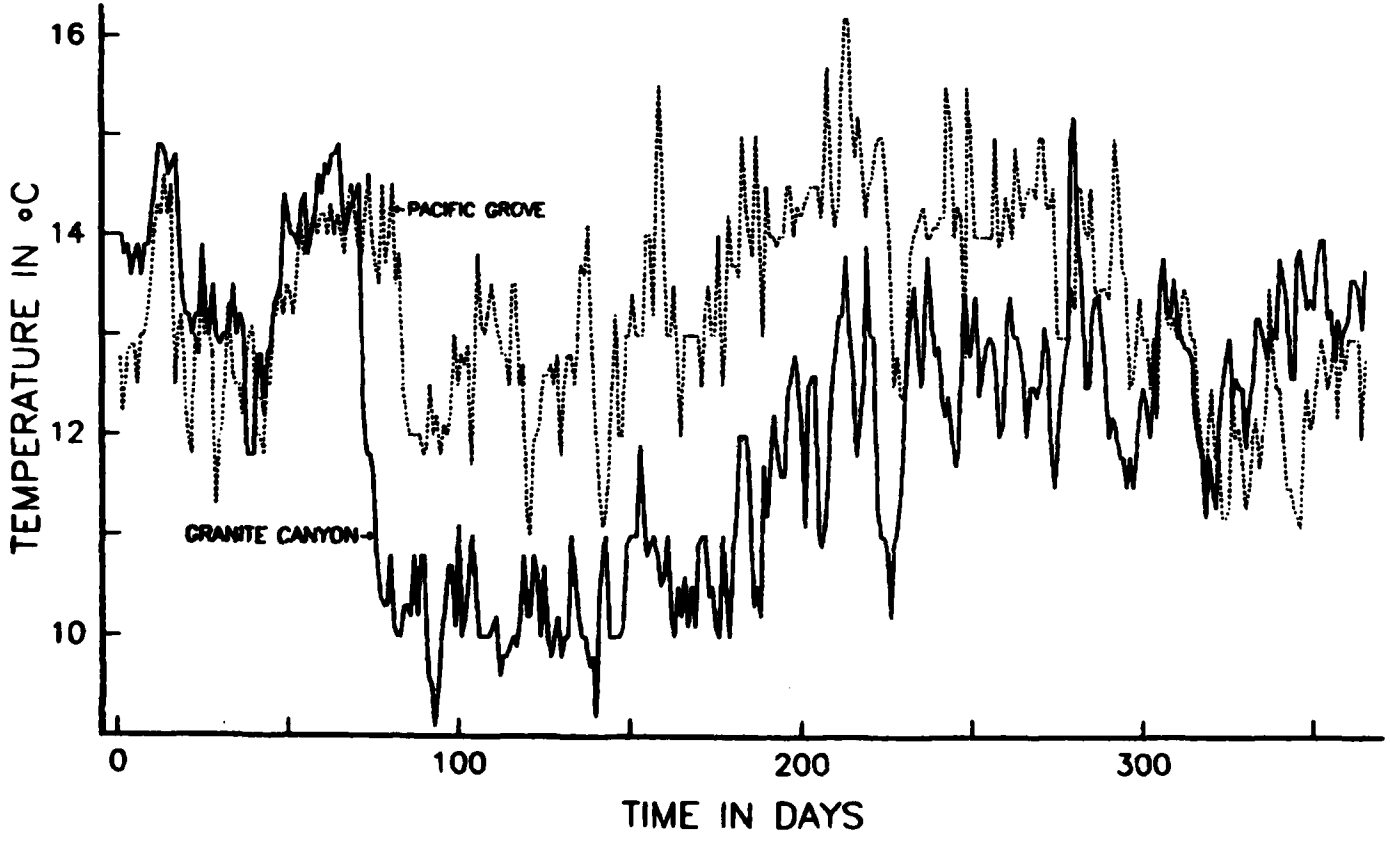


FIGURE 5

GRANITE CANYON AND THE FARALLONS COMPARED

1/1/80 TO 1/1/81

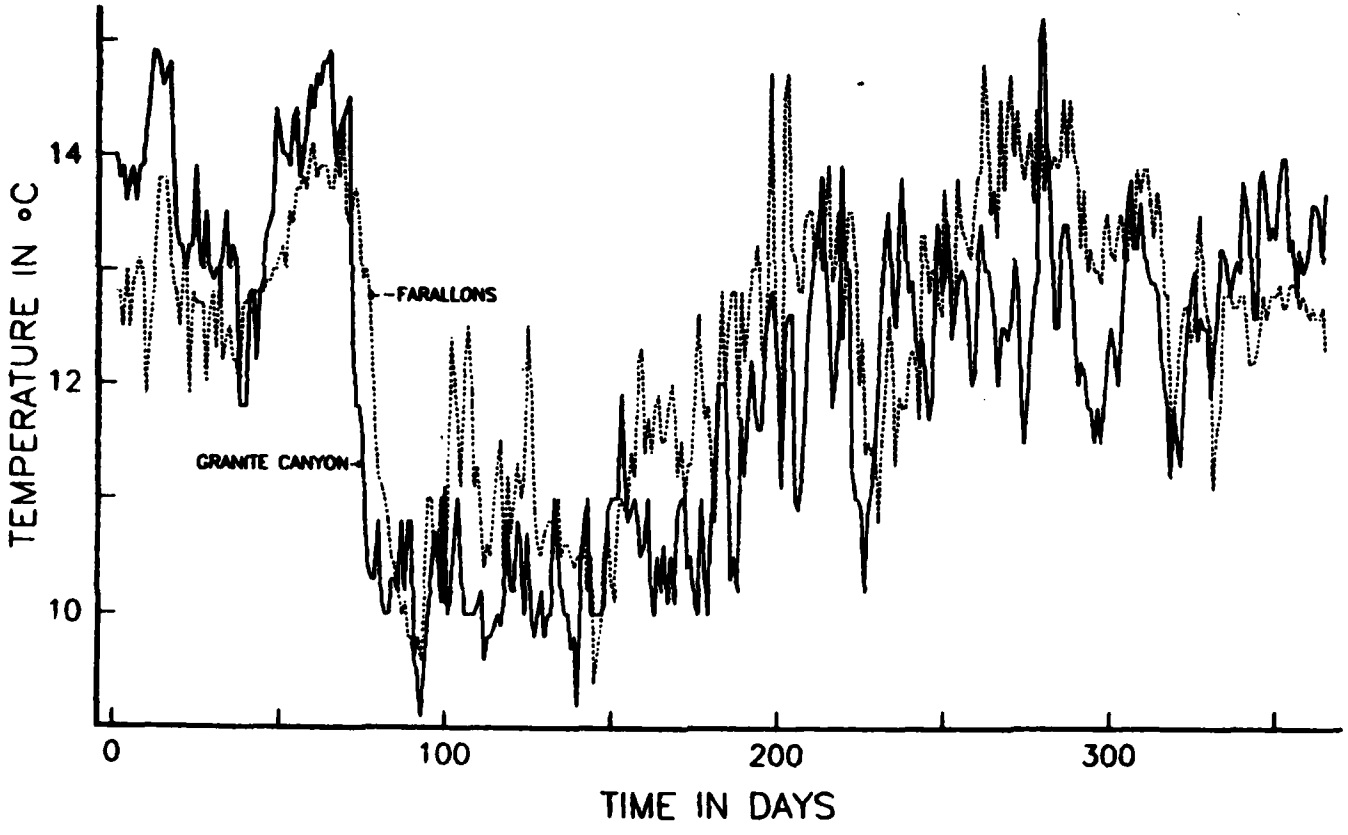


FIGURE 6

Canyon and the Farallon Islands respectively for the year 1980. Diablo Canyon is located about 200 Km south of Granite Canyon and the data from that location were recorded at 5m below the surface on the continental shelf, just off the PG&E nuclear power plant located on the adjacent coast. The temperatures at Pacific Grove are collected near the shore in a sheltered area inside Monterey Bay. The Farallon Islands data contained numerous multi-day gaps which were filled by linear interpolation. The same spring transition shown in Figure 3 is seen to be present in these figures. The overall patterns in the annual cycle at these four locations are seen to be generally similar at least during 1980.

3. Gross trends in the Granite Canyon SST's

We discuss here gross features of the Granite Canyon SST record. These features are of interest per se and are also a prelude to the more quantitative decomposition and time series analysis performed in the subsequent sections.

Figure 7 shows a scatter plot of the temperature data against day number. In most years fewer than three observations were missing from the raw data. Where missing data were found, linearly interpolated values were substituted. Since the points in Figure 7, unlike those in Figure 3, are not connected, the quantization of the data (0.1C) can be clearly seen.

The most interesting feature of Figure 7 is a definite upward average tilt to the data. A least-squares fit of the linear trend equation $y(t) = \alpha + \delta t$ to the data gives the fit $y(t)=10.9 + 0.00037t$, where t is in days, and $y(t)$ is water temperature in ^{o}C on day t. This indicates an "average" temperature rising from 10.9 to 12.52C over the twelve year measurement period ($t = 4380$). Note that this linear trend is simply a convenient way

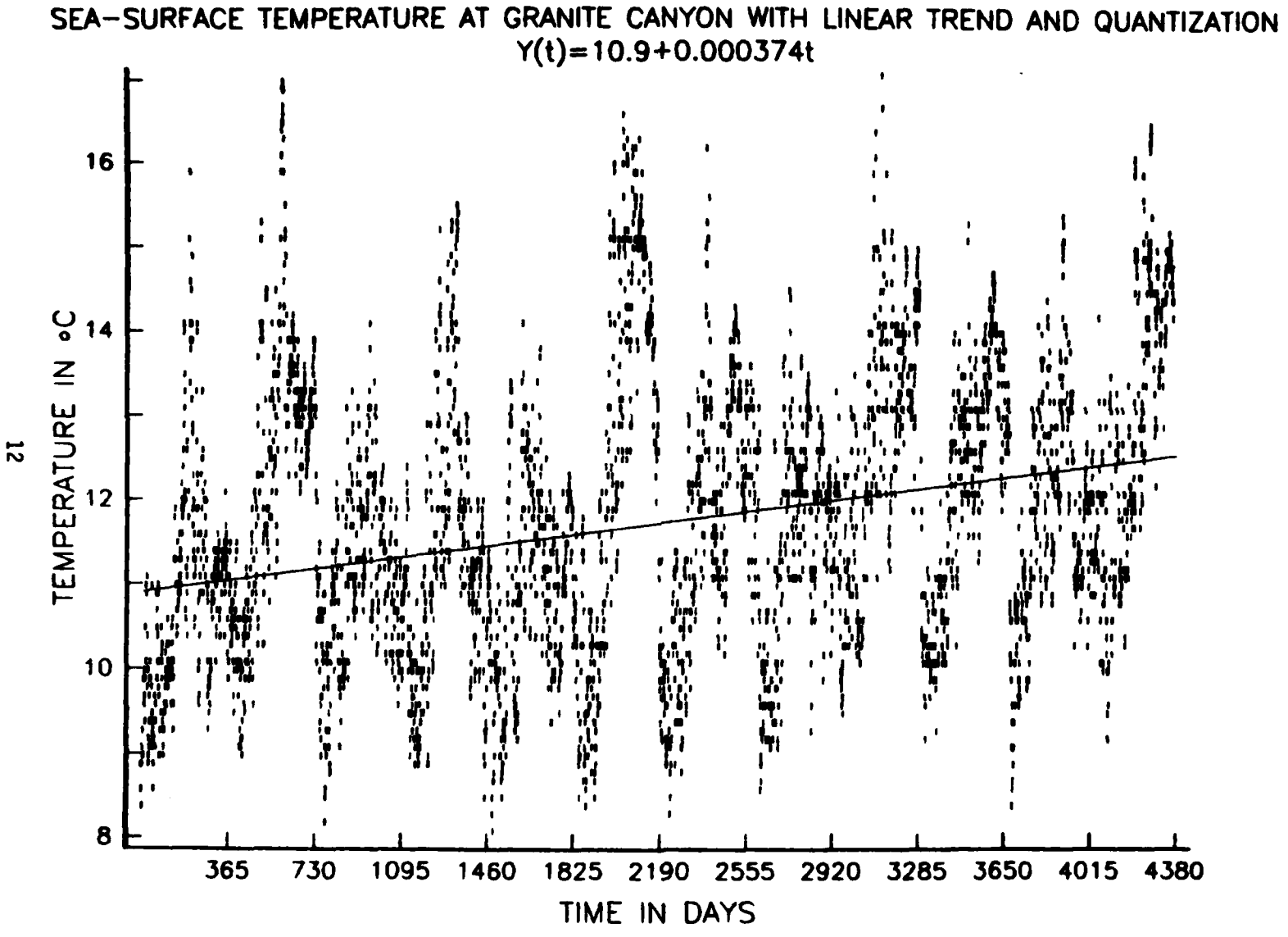

SEA-SURFACE TEMPERATURE AT GRANITE CANYON WITH LINEAR TREND AND QUANTIZATION
Y(t)=10.9+0.000374t
TEMPERATURE IN °C
16
14
12
10
8
365
730
1095
1460
1825
2190
2555
2920
3285
3650
4015
4380
TIME IN DAYS

FIGURE 7

of representing the overall increase in temperature over the 12-year period. Higher order polynomial fits provide only slight improvement. We may, however, be observing only a part of a much longer cycle of fixed or variable period which would be revealed in a much longer data series. The statistical significance of the trend parameter $\hat{\delta}$ =0.00037 (i.e. whether it is really different from zero) is complicated by the presence of cycles and possible short term correlation in the data. Thus standard tests that $\delta = 0$ based on independent data are invalid. However, using simulation techniques described later, we have found the standard deviations of δ to be 0.000011 and hence reject the hypothesis that $\delta = 0$ at the 95% level of confidence.

The gross features of the Granite Canyon SST's clearly include annual cycles and it is apparent from Figures 2 and 7 that these annual cycles contain inter-annual variability as well. In particular the oceanic El Niños of 1972, 1976, and 1983 are clearly visable. The possibility of a temperature anomaly in 1979 is also visable. Although the effects of El Niño events are often restricted to equatorial regions and to regions along the South American coast, their influence does upon occasion extend at least as far north as central California (Bretschneider and McLain, 1983).

Underlying patterns in the data can be exposed by performing local smoothing. (Since the computations in this paper were done using the APL language (see e.g. Gilman & Rose, 1976) with an experimental APL Graphics package, GRAFSTAT3, the smoothing was performed using the program FILTER from Anscombe (1981, p392).) The smoothing computes a cosine-weighted moving average of extent m to give a smoothed time series $\bar{y}(t)$:

$$\bar{y}(t) = \sum_{j=-(m-1)}^{m-1} y(t)w(t-j) \quad , \tag{3.1}$$

where the $w(k) = \frac{1}{m+1}\left(1-\cos[2\pi(k+1)/(m+1)]\right)$, for $k=0,\pm1,\pm2,..,\pm(m-1)$, (3.2)

are equal-spaced ordinates of 1-cos(2t). The gain of this filter is given by

$$G(\omega) \sim \frac{\sin[\omega(m+1)2]}{\omega(m+1)[1-\omega^2(m+1)^2(2\pi)^2]} \quad . \tag{3.3}$$

Details are given in Anscombe (1981, p.156).

Smoothings of the Granite Canyon SST's for various values of (2m+1) together with the number of filter weights used are given in Figure 8. It is clear that 61 weights (or 81 weights) resolve the gross details adequately. It is interesting to note that there is not only a sharp drop in temperature in the spring, but also a rather clearly defined and relatively abrupt rise in temperature in the fall. We note that both the spring and fall transitions occur over periods relatively short in comparison with the annual cycle.

Since well-defined spring transitions could be clearly identified in at least 6 out of the 12 years, it was decided to tabulate these events with respect to their dates of occurrence, duration, and associated changes in temperature. The results are shown in Table 1 below. Because the spring transitions add complexity to the annual temperature cycle, the harmonic content of this data is undoubtedly richer by their presence. We will say more about the annual variation in temperature at Granite Canyon in Section 10.

SMOOTHED VERSION OF ORIGINAL TIME-SERIES
COSINE FILTER

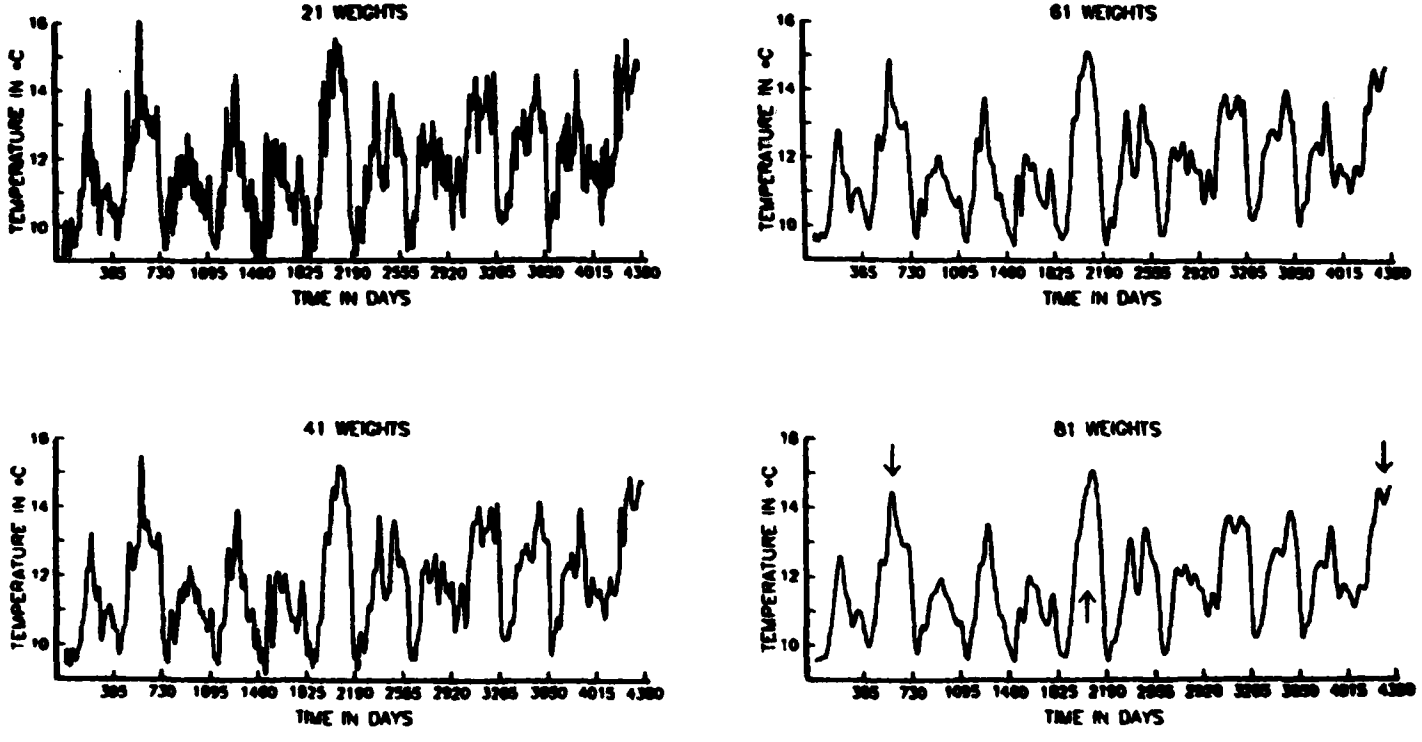


FIGURE 8

Table 1*

YEAR	STARTING DATE	DURATION	DELTA T	dT/dt
1973	March 9	7 days	13.0 - 10.0C	0.43^{o}/Day
1974	March 31	5 days	12.3 - 10.0	0.46
1977	Feb. 22	10 days	13.1 - 9.3	0.38
1978	April 25	9 days	13.0 - 8.5	0.50
1980	March 11	7 days	14.5 - 10.4	0.59
1981	March 23	8 days	13.1 - 10.0	0.39

* A lesser transition was observed on December 19, of 1981. Whether or not this event corresponds to a "spring" transition is not known.

A detailed statistical analysis of the time series of SSTs will now be given. This will be used in Sections 8 and 9 to give a more detailed examination, statistical and oceanographic, of the SSTs.

4. Cyclic fit and decomposition of the data: General ideas

We proceed to decompose the SST time series into an additive model of the form

$$y(t) = m(t) + s(t) + \varepsilon(t), \qquad (4.1)$$

where the trend m(t) can be defined loosely as a long term change in the mean (Chatfield, 1980, p.13), s(t) consists of seasonal and cyclic changes, and $\varepsilon(t)$ is a stationary random sequence which describes the irregular fluctuations. The $\varepsilon(t)$ sequence is assumed to have mean zero, (i.e. $E(\varepsilon(t))= 0$), constant variance, and possibly some serial dependence. That m(t), s(t) and $\varepsilon(t)$ are treated as additive implies the gross assumption that this dependency structure in the $\varepsilon(t)$, as well as the variance, is not affected by m(t), or s(t).

The model (4.1) is essentially a conceptual tool for thinking with and its usefulness and approximate validity must be established from the data. Thus to define "scales" in the data these different components must be identified.

The term m(t) was discussed in Section 3 and here we let

$$y_1(t) = y(t) - m(t) = s(t) + \varepsilon(t) \tag{4.2}$$

insofar as it can be observed. With the estimate $\hat{m}(t) = 10.9 + 0.000374t$ removed from $y(t)$ we denote the resultant as the first residual, $r_1(t)$, ignoring for the present possible interaction between the estimation of m(t), s(t) and $\varepsilon(t)$.

In this section we estimate the structure of s(t) and $\varepsilon(t)$ by analyzing $r_1(t)$. Note that $s(t)$ includes definite seasonal terms, such as the obvious annual effect in $E(y(t))$ (and its harmonics), and terms with periods longer than a year which might be interpreted as global fits to the obvious long-term modulation of the annual cycle which can be seen in figures 2 and 6. Again there may be cycles at shorter periods associated with the tides and winds; these are generally difficult to distinguish from quasi-cycles "generated" by highly correlated errors $\varepsilon(t)$. Thus we assume that

$$s(t) = \sum_{j=1}^{J} \{\gamma_j \cos(\omega_j t) + \beta_j \sin(\omega_j t)\} \tag{4.3}$$

$$= \sum_{j=1}^{J} A_j \cos(\omega_j t + \theta_j) , \tag{4.4}$$

where $\omega_j = 2\pi f_j = 2\pi/T_j$ are (unknown) frequencies corresponding to periods T_j, $A_j = (\gamma_j^2 + \beta_j^2)^{1/2}$ is the amplitude, $\theta_j = \tan^{-1}(-\beta_j/\gamma_j)$ is the phase, and J is the number of cycles which comprise $s(t)$. Again we note that cycles with periods longer than 1 year may represent long-term random fluctuations. The representation (4.3) of the component $s(t)$ of $E(y(t))$ as a sum of deterministic cycles with fixed frequencies is a convenience only. In particular it might be hazardous to predict very long term behavior on the basis of this model for $s(t)$ and $m(t)$. Also, while our model may identify cycles with periods shorter than one year, such cycles may not represent real cyclic behavior, but rather they may serve only to accommodate the non-sinusoidal nature of the year cycle and the very sharp transition in that cycle that often occurs during spring (see Table 1). Because our data extends for only 12 years, it will be impossible to detect any cycles with periods longer than half that length, or 6 years.

To find "hidden periodicities" in $s(t)$ is a formidable task, even if the spectrum of the random component $\varepsilon(t)$, namely

$$f(\omega) = \frac{1}{\pi}\left\{1 + 2\sum_{k=1}^{\infty} \rho(k)\cos(k\omega)\right\}, \qquad 0 \leq \omega \leq \pi, \tag{4.5}$$

where $\rho(j) = \mathrm{corr}(\varepsilon(t+j)\varepsilon(t))$, were constant (flat) i.e. the $\varepsilon(t)$ are "white noise". (This is the classical problem of "hidden periodicities" going back to the original work of Schuster (1898) and Fisher (1929).) To find these periodicities in the presence of a non-flat spectrum for the $\varepsilon(t)$'s, we proceed as follows:

(1) Examine the periodogram of the first residuals, $r_1(t)$, to find possibly significant periodicities. Estimate $s(t)$ as $\hat{s}(t)$. Given the chosen frequencies ω_j, this involves obtaining least

squares estimates of the γ_j's and β_j's in the linear model (4.3). The least squares estimates of the amplitudes A_j are directly proportional to the periodogram values at the given frequencies.

(2) Examine the second residuals $r_2(t)=y(t)-\hat{s}(t)-\hat{m}(t)=r_1(t)-\hat{s}(t)$ to determine the properties of $\varepsilon(t)$. If $\varepsilon(t)$ can re represented, say, as a pth-order autoregressive process,

$$\varepsilon(t) = a_1\varepsilon(t-1) + \ldots + a_p\varepsilon(t-p) + n(t), \qquad (4.6)$$

where the $n(t)$ are uncorrelated. Then we estimate a_1 through a_p by least squares to be $\hat{a}_1$ through $\hat{a}_p$ and "pre-whiten" the first residuals, $r_1(t)$, as

$$r_3(t) = r_1(t) - \hat{a}_1 r_1(t-1) \ldots -\hat{a}_p r_1(t-p). \qquad (4.7)$$

We then examine the periodogram of these third residuals to ascertain whether the $\hat{s}(t)$ chosen in step (1) is still a "reasonable" representation of the seasonal and cyclic changes (global fit) for the data. The "prewhitening" is performed because highly autocorrelated data will elevate spectral values near zero frequency, making it difficult to decide whether a given spectral peak is due to the correlation or due to a cycle in the data. Once the autocorrelated component is removed, we obtain a clearer picture of what cycles should actually be included in $\hat{s}(t)$.

(3) Using $r_3(t)$ to point out a possibly refined $\hat{s}(t)$, re-examine the structure of $\varepsilon(t)$ via the (possibly refined) second residuals, $r_2(t)=y(t)-\hat{m}(t)-\hat{s}(t)$. Knowledge of the structure of $\varepsilon(t)$ is important for determing "local scales" and obtaining the "whitened" residuals $r_4(t)$. Thus, if it is determined that $\varepsilon(t)$ has pth

order autoregressive structure, possibly different from that in (2) above, compute fourth residuals

$$r_4(t) = r_2(t) - \hat{a}_1 r_2(t-1) \ldots - \hat{a}_p r_2(t-p). \qquad (4.8)$$

These residuals are used firstly to establish the overall goodness of fit of the estimated model, since the $r_4(t)$ process will have a flat spectrum and a zero level correlogram, (i.e., "white noise") if the model is reasonable. Secondly they can be used (e.g. Tsiao and Box, 1982) to establish cross-correlations of the Granite Canyon data with other SST data. Thirdly, they provide an estimated distribution of the n(t) which will be useful later for modelling and for checking assumptions during statistical testing.

5. Global fit and cyclic components

Figure 11 shows, in the top panel, a cyclic fit $\hat{s}(t)$ with J = 13 terms to the first residuals (original data minus long term linear trend $\hat{m}(t)$). The data, since leap days have been removed, consists of N = 4380 daily observations of temperature. Thus if angular frequency is re-expressed as discrete frequencies of the form $\omega_p = 2\pi p/N$, where p is the "number of cycles per period of 4380 days", then the cycles removed from the data as components of s(t), have frequencies p = 12,24,94, corresponding to periods of 1 year, 6 months and approximately 1.5 months, with the remaining p = 2,3,4,5,6,7,8,9,11,18.

The determination of these frequencies follows the steps in Section 4. First, Figure 10 shows the estimated serial correlations of the first residuals

$$\hat{\rho}(k) = \sum_{t=1}^{N-k} [y(t) - \hat{m}(t)]\,[y(t+k) - \hat{m}(t)] / \{(N-k)S^2\} \qquad (5.1)$$

$$= \sum_{t=1}^{N-k} r_1(t) r_1(t+k) \quad \{(N-k)S^2\}$$

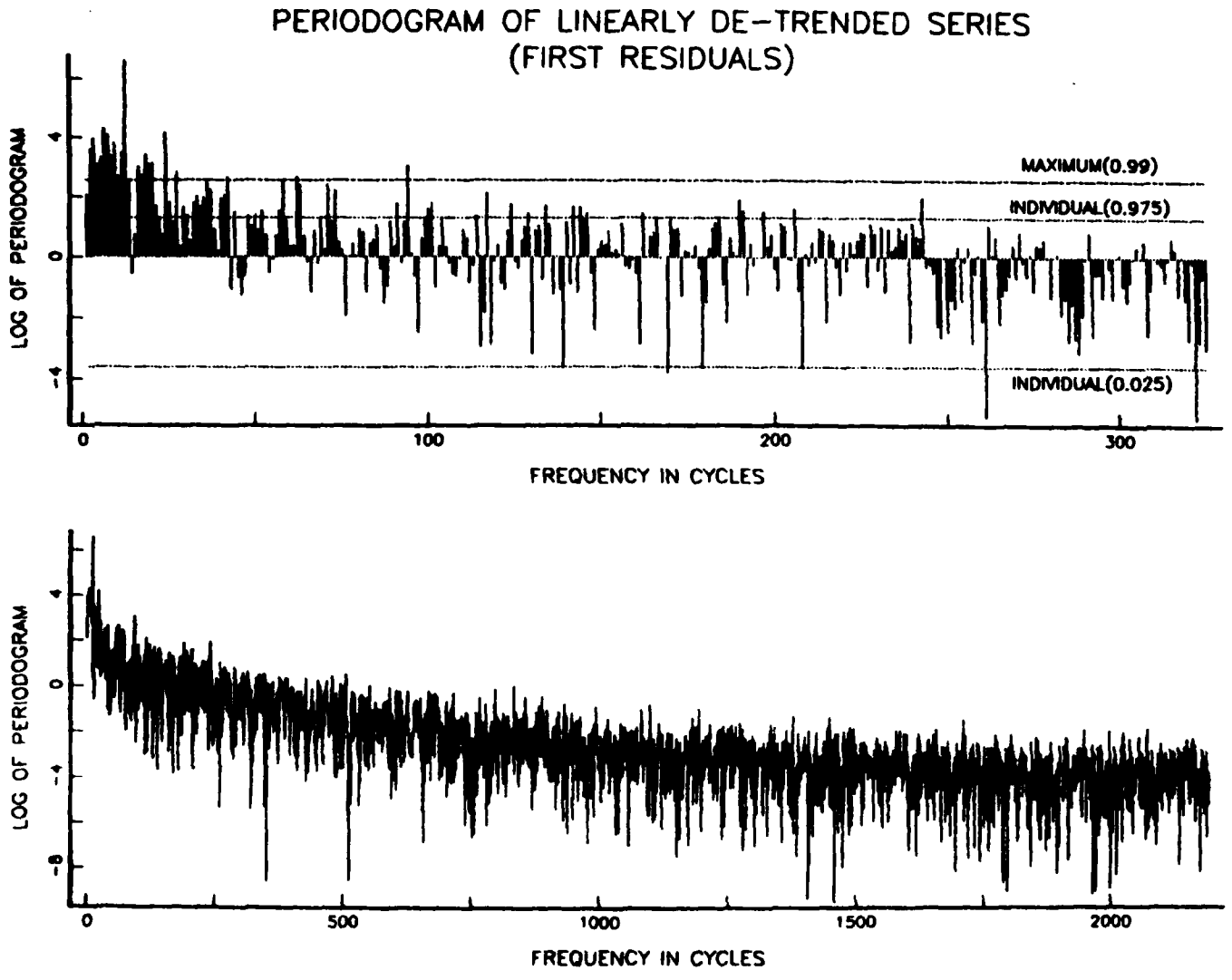

PERIODOGRAM OF LINEARLY DE-TRENDED SERIES
(FIRST RESIDUALS)
LOG OF PERIODOGRAM
4
0
-4
MAXIMUM(0.99)
INDIVIDUAL(0.975)
INDIVIDUAL(0.025)
0
100
200
300
FREQUENCY IN CYCLES
LOG OF PERIODOGRAM
4
0
-4
-8
0
500
1000
1500
2000
FREQUENCY IN CYCLES

FIGURE 9

AUTOCORRELATION OF LINEARLY DETRENDED DATA
(FIRST RESIDUALS)

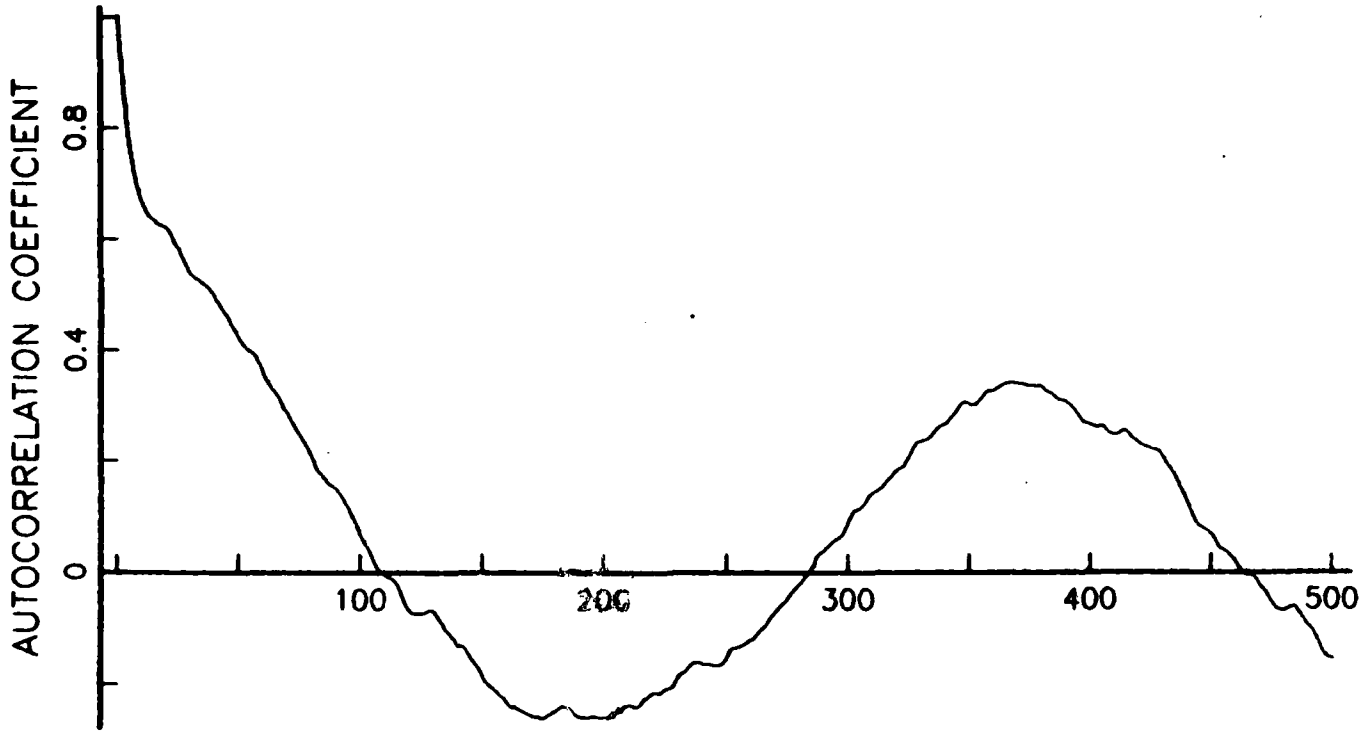


FIGURE 10

13-TERM CYCLIC FIT OF LINEARLY DETRENDED DATA (FIRST RESIDUALS)

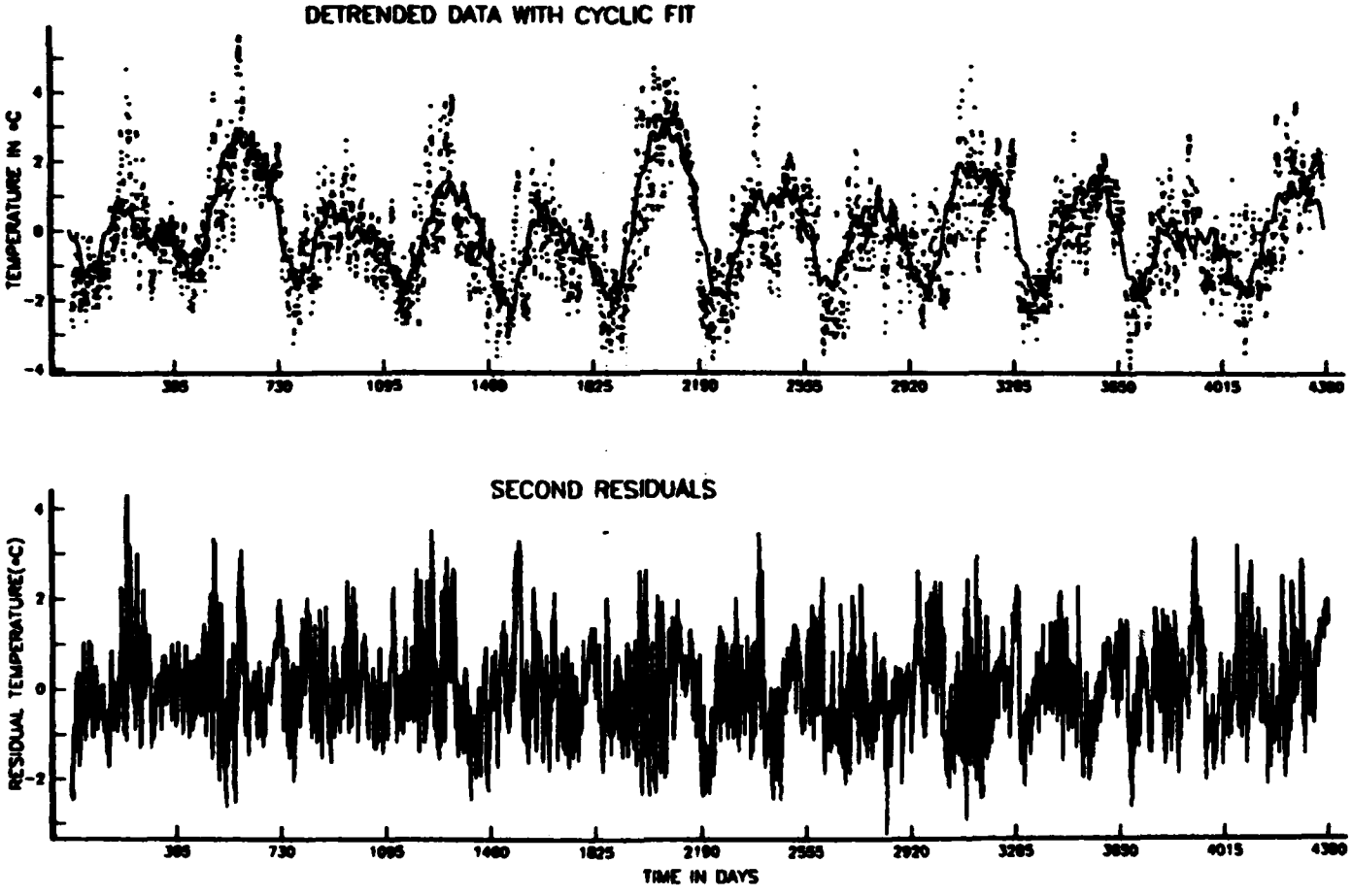


FIGURE 11

where S^2 is the sample variance of the $r_1(t)$'s. This correlogram is dominated by the yearly cycle and therefore can not be used to postulate a model for the $\varepsilon(t)$, or to identify more than the yearly cycle itself.

A harmonic spectral analysis of the $\hat{\rho}(k)$'s is given by the periodogram of the first residuals, defined as

$$I_n(\omega_p) = \{A_n^2(\omega_p) + B_n^2(\omega_p)\}/(2\pi N), \quad p = 1,..,N/2, \qquad (5.2)$$

where

$$A_N(\omega_p) = \sum_{t=1}^{N} r_1(t)\cos(t\omega_p), \text{ and } B_N(\omega_p) = \sum_{t=1}^{N} r_1(t)\sin(t\omega_p),$$

for $p = 1,...,N/2$, are the real and complex parts of the discrete Fourier transform of $r_1(t)$, $t=1, .., N$. Here $\omega_p = \frac{2\pi p}{N}$. Since the expected value of the periodogram is (asymptotically), for uncorrelated data, $\sigma^2/2\pi$, we use a normalized periodogram. Thus the normalized periodogram is defined to be the result (5.2) multiplied by $2\pi/S^2$, where S^2 is the sample variance of the data being considered, namely $r_1(t)$.

The logarithm of the normalized periodogram of the first residuals from the Granite Canyon data is shown in Figure 9 with confidence bands for individual components and for the maximum value of the periodogram components. These are based on the approximation that each value in the normalized periodogram is independently distributed as a mean one exponential random variable (chi-square with 2 degrees of freedom). The dominant frequency is $p = 12$, i.e., a period of one year. There is, in the expanded scale periodogram in the top panel of Figure 9, a significant harmonic at $p = 24$ and at $p = 94$. This latter frequency is surprising, since it is not exactly a harmonic of $p = 12$. The full periodogram in the lower panel suggests, even though it is unsmoothed, an autoregressive type spectrum for the $\varepsilon(t)$. Thus, determination of cycles at frequencies below 12 is difficult. Proceeding to the prewhitening, frequencies $p = 2,3,4,5,6,7,8,9,11$ and 18 are selected for removal, in addition to $p = 12$, 24 and 94.

Removal of these frequencies gives the second residuals whose periodogram is shown in Figure 12. On the expanded scale, note the zeros where cycles have been removed; this is because the least squares estimates of the amplitudes in (4.4) are exactly proportional to the corresponding periodogram values. Again note in the lower panel of Figure 12 the autoregressive-like spectrum. (A spectrum for a theoretical AR(2) process using parameters computed below is shown for comparison.)

Figure 13 shows the estimated autocorrelations of the second residuals as well as the first 50 correlations for four theoretical AR(P) models computed from the first four estimated correlations, via the Yule-Walker equations (Chatfield, 1980, p.48). Note in Figure 13 that the AR(P)models of order greater than 2 do not give a much better fit than the AR(2) model to the estimated autocorrelations.

Notice also that the persistent hump at about 21 days. As yet we have no explanation for this effect. The autocorrelations in Figure 12 are extended to show lags beyond 50 in Figure 13, but there is no indication of a cycle of period 21. Such a cycle would show up as a wave in the correlogram with period 21, not simply as a peak at lag 21 days.

Table 3

Estimated correlation coefficients for the residuals $r_2(t)$ and estimated coefficients for autoregressive models of orders $p = 1,2,3$ and 4. Note that the underlined values are the first four partial correlations for the data.

Correlations	Estimated coefficients	p
$\hat{\rho}(1) = 0.8511$	$\hat{a}_1 = \underline{0.8511}$	1
$\hat{\rho}(2) = 0.6973$	$\hat{a}_1 = 0.9343$; $\hat{a}_2 = \underline{-0.0975}$	2
$\hat{\rho}(3) = 0.5718$	$\hat{a}_1 = 0.9355$; $\hat{a}_2 = 0.1098$; $\hat{a}_3 = \underline{0.0129}$	3
$\hat{\rho}(4) = 0.4710$	$\hat{a}_1 = 0.955$; $\hat{a}_2 = 0.109$; $\hat{a}_3 = 0.007$; $\hat{a}_4 = \underline{0.0006}$	4

PERIODOGRAM FOR SECOND RESIDUALS

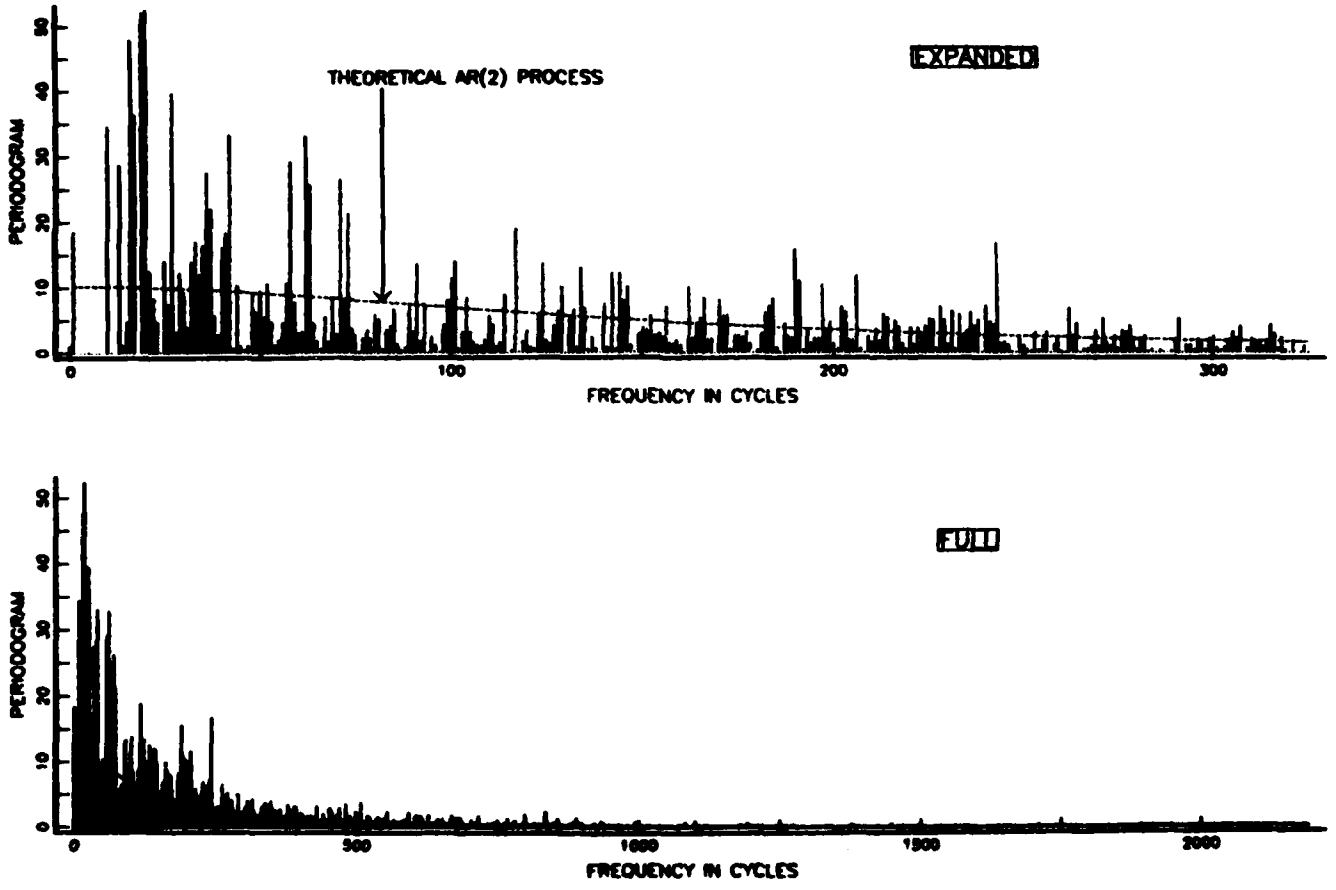


FIGURE 12

ESTIMATED AND THEORETICAL AUTOCORRELATIONS FOR SECOND RESIDUALS

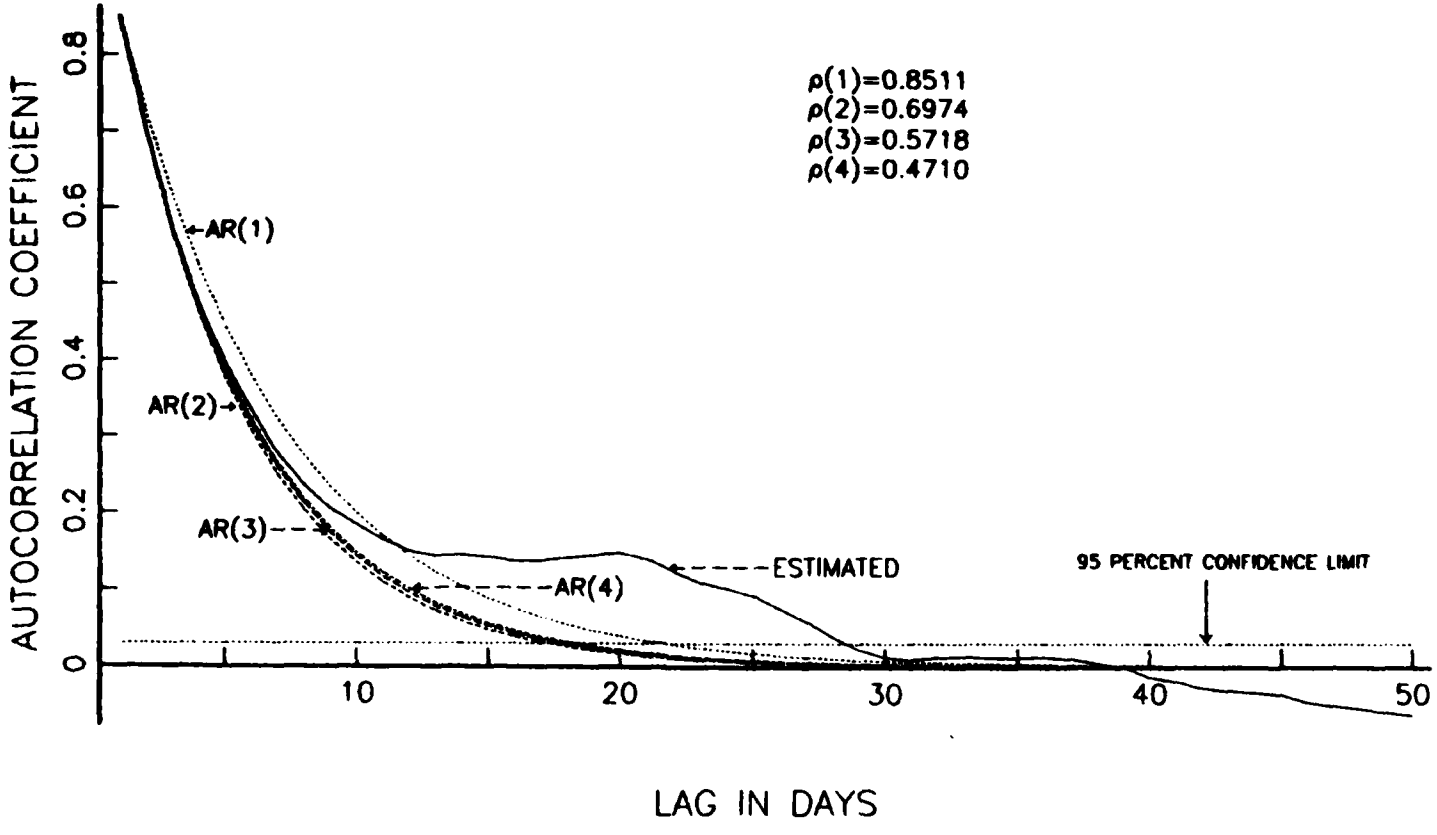


FIGURE 13

Consequently the first residuals are pre-whitened as in (4.7) above, using $p = 2$ and $\hat{a}_1 = 0.9343$ and $\hat{a}_2 = -0.0975$ derived from the values $\hat{\rho}(1) = 0.8511$ and $\hat{\rho}(2) = 0.6973$. The resulting third residuals, $r_3(t)$, and their autocorrelations are shown in Figure 15 with the logarithm of the normalized periodogram in Figure 16. The choice of frequencies removed on the basis of the first residuals is confirmed. To explain the hump in the autocorrelogram (Figure 14), several other cyclic components were removed, e.g. those at $p = 19$, 20 and 243. These changed neither the autocorrelations nor the residual variance of the data.

We discuss now details of the detrending of the data. Table 4 shows the sample variance of the original data (raw data minus grand mean), the first residuals (removal of linear trend), the second stage (removal of linear trend and $\hat{s}(t)$ to give $r_2(t)$), and the fourth residuals which will be discussed below.

Table 4

Stage	Sample Variance	Sample s.d.	% variance removed
Raw data	2.624	1.62	
			8.4%
$r_1(t)$	2.403	1.55	
			51.9%
$r_2(t)$	1.040	1.02	
			28.7%
$r_4(t)$	0.288	0.537	
			10.9% Remainder

Note that about half of the variance is accounted for by the cyclic component $s(t)$. A break-down for the $J = 13$ cyclic components is given below in Table 5. Components are listed in descending order of magnitude of their amplitudes, A_j. The variance after removal of each of these (orthogonal) components is given in the last column. The annual cycle accounts for a large part of the variation, namely $(.81)/1.36 \times 100\% = 59.6\%$ of this variance.

These estimated coefficients will be used in later sections in a simulation study of the model.

AUTOCORRELATIONS FOR SECOND RESIDUALS

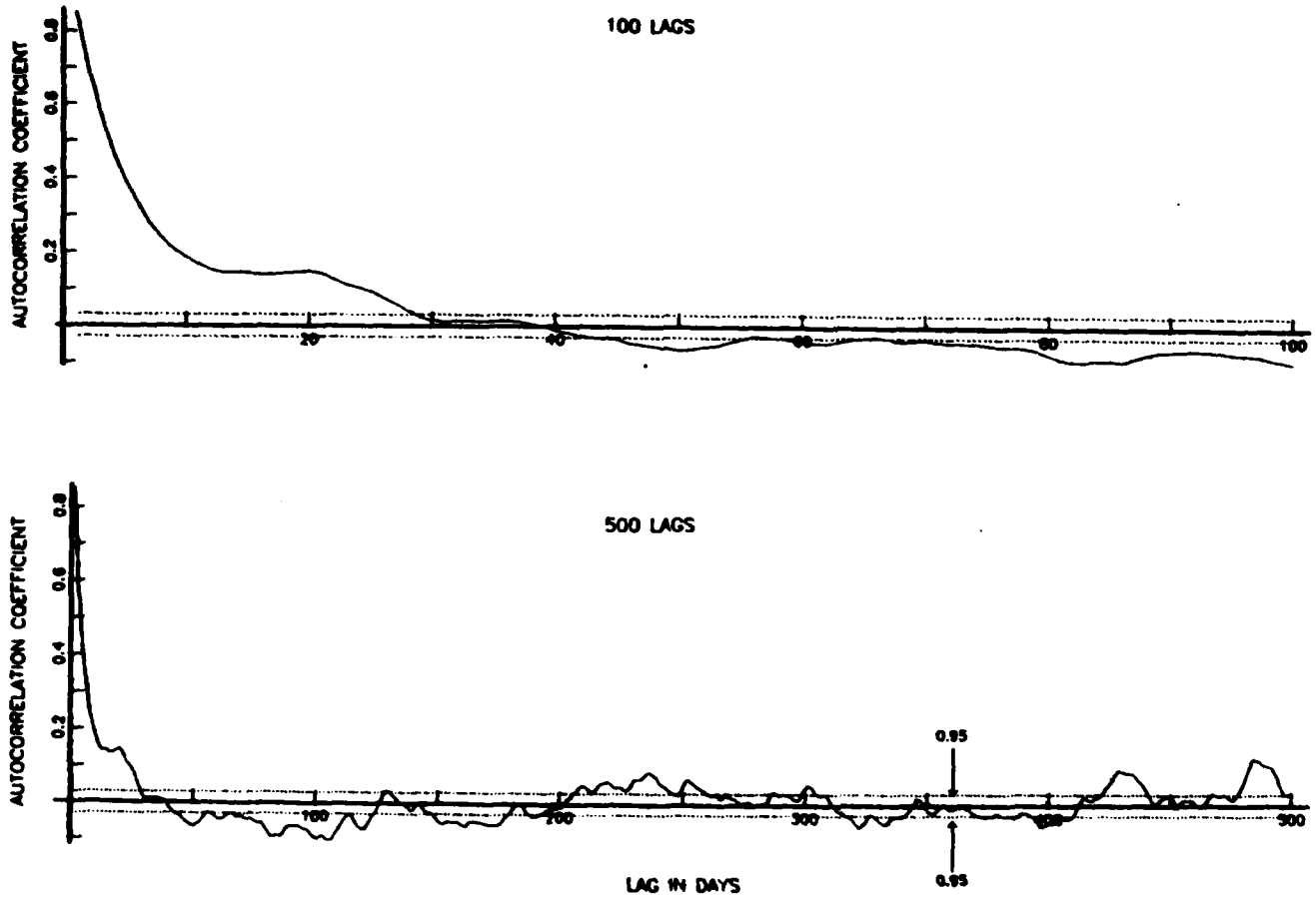


FIGURE 14

THIRD RESIDUALS FROM PRE-WHITENING[AR(2)]

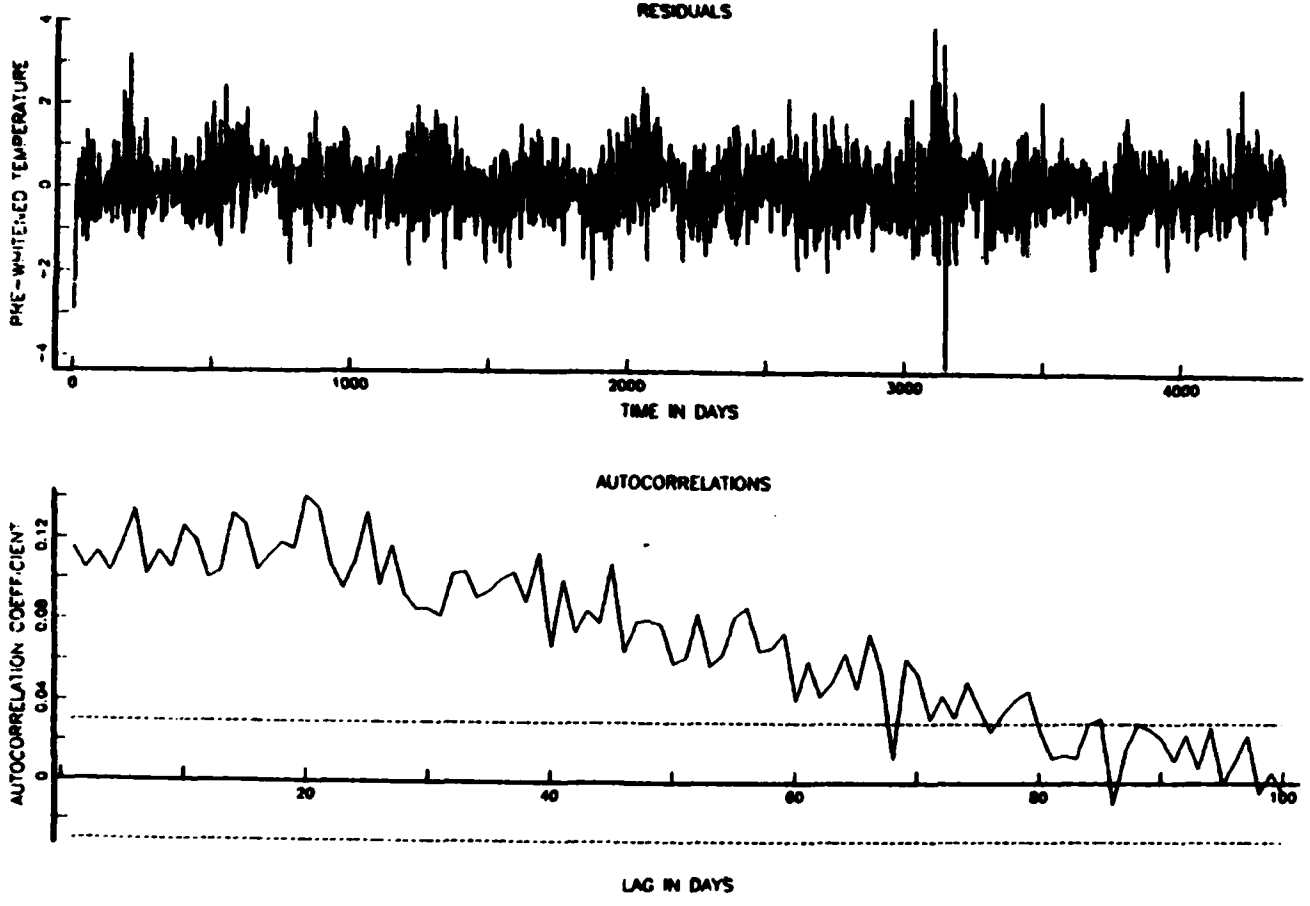


FIGURE 15

PERIODOGRAM OF THIRD RESIDUALS
(FIRST RESIDUALS–AR(2))

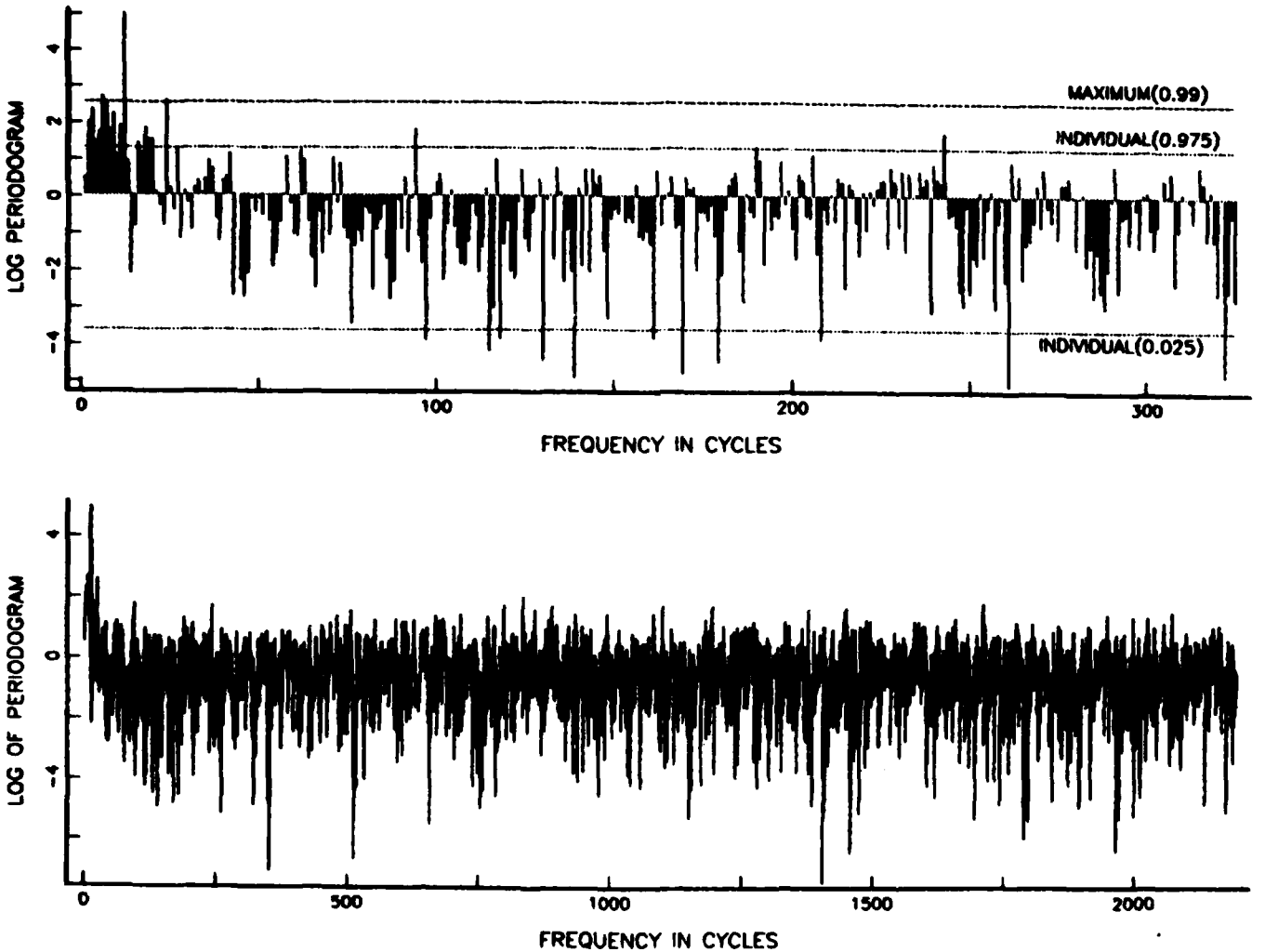


FIGURE 16

Table 5

j	p_j	T_j(years)	γ_j	β_j	A_j	θ_j	Variance
							2.40
1	12	1.00	-0.465	-1.187	1.275	111.4^0	1.59
2	6	2.00	0.273	-0.295	0.402	47.2^0	1.51
3	24	0.50	0.342	-0.162	0.378	25.4^0	1.44
4	7	1.71	0.345	0.112	0.363	342.0^0	1.37
5	3	4.00	-0.266	0.211	0.340	218.4^0	1.32
6	9	1.33	-0.101	0.302	0.318	251.6^0	1.28
7	2	6.00	0.074	0.275	0.285	285.1^0	1.23
8	11	1.09	0.097	0.250	0.268	291.2^0	1.19
9	18	0.67	-0.187	-0.180	0.260	136.1^0	1.15
10	8	1.50	0.079	-0.243	0.256	72.0^0	1.12
11	5	2.40	-0.176	-0.174	0.247	135.3^0	1.08
12	94	0.13	-0.155	-0.152	0.218	135.6^0	1.06
13	4	3.00	0.194	-0.097	0.217	26.6^0	1.04

6. Examination of the random component, $\varepsilon(t)$

The second residuals, the original data minus the linear trend and the seasonal trend $\hat{s}(t)$ containing 13 components, are whitened (4.8) under the assumption of the adequacy of an AR(2) model to give the fourth residuals

$$r_4(t) = r_1(t) - \hat{a}_1 r_2(t-1) - \hat{a}_2 r_2(t-2).$$

The first use of these residuals is as a test of the adequacy of the modelling which has been done up to this point. There are two issues which arise; (1) the physical adequacy of the model, and (2) the statistical adequacy. Any given model may satisfy either, neither or both of these requirements. For physical adequacy we require that the model describe a substantial amount of the variation in the data with as few terms as possible (i.e., the principle

of parsimony), and hope that at least the dominant terms have oceanographic interpretations. For statistical adequacy we have made the assumption that the fourth residuals are estimates of the n(t)'s and, as such, should, if the model is reasonable, be approximately uncorrelated.

The residuals themselves are plotted in the top panel of Figure 17. There is an anomaly at about 3100 days associated with isolated extremes in SSTs encountered during the fall of 1979. It is in fact a reminder that the model we have derived is a global model which will not account for transients or, say, subtle seasonal effects. Except for that anomaly, the fourth residuals generally fall in the range of ± 2.0 which is considerably smaller than the range of the original data (8.0 to 17.0). Moreover, in Table 4 we have shown that the fourth residuals, which represent the unpredictable part of the data, account for only 10.1% of the total variation. Oceanographic interpretations for some of the parameters used in constructing the model are presented later, as well as simulation results which show how well the model mimics in behavior the data.

The autocorrelation function for the fourth residuals is given in the lower panel of Figure 17. The bands are approximate 95% confidence bands for individual correlations. Again there may be an anomaly at lag 20, but since this is a maximum of 100 estimated correlations it probably is not significantly large. A more precise test that the $r_4(t)$s are uncorrelated is given in Figure 18.

The top panel is the unsmoothed, normalized periodogram and the bottom panel is the cumulative periodogram. Barely distinguishable upper and lower 95% and 99% confidence intervals for a flat spectrum based on Kolmogorov-Smirnov (KS) statistics are shown. Since the cumulative periodogram lies within these bands, the Durbin test for a flat spectrum (see Jenkins and Watts, 1969, p.234-7, or Cox and Lewis, 1966, p.169) accepts the hypothesis of

FOURTH RESIDUALS

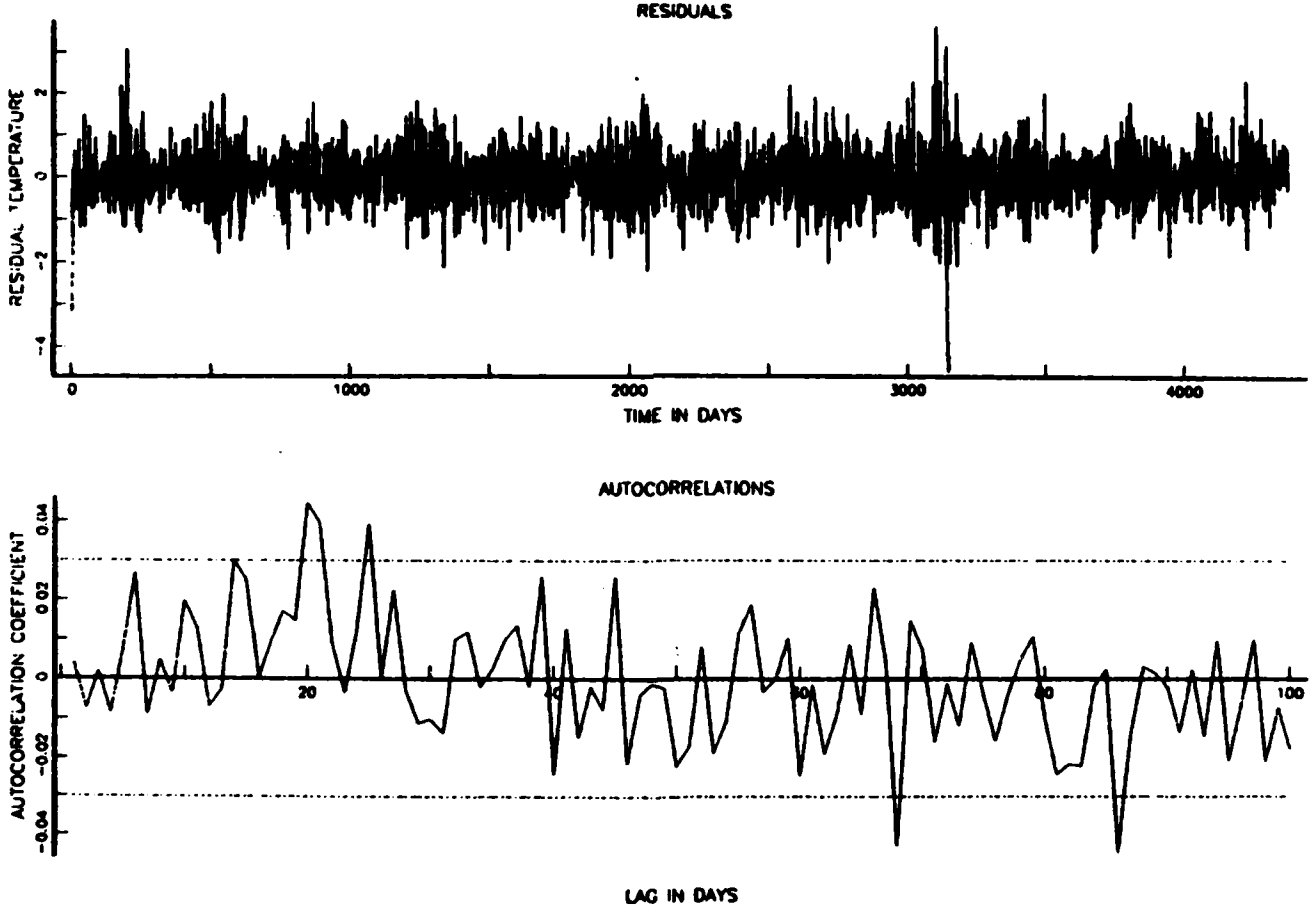


FIGURE 17

FOURTH RESIDUALS

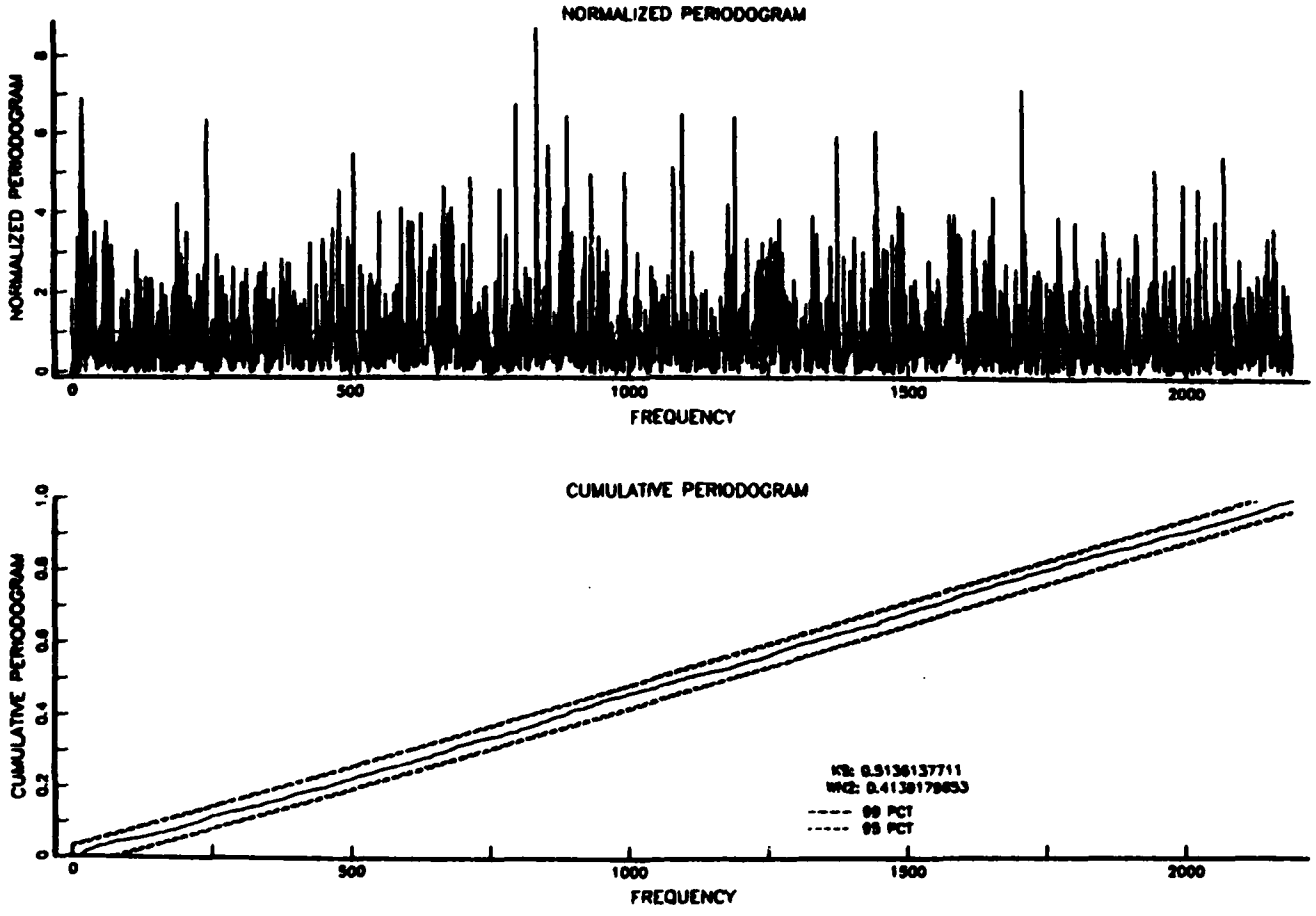


FIGURE 18

a flat spectrum. The value of the KS statistic is given as 0.5136 and compares to the upper 5% point of the statistic which has value 1.355. In addition the Anderson-Darling test statistic (value 0.9139) is given; this test statistic is more sensitive to departures at low or high frequencies (Cox and Lewis, 1966, p.169) and in this case accepts the hypothesis of no correlation.

Thus, we have established the adequacy of the model. Note that if the $\varepsilon(t)$ process is a Normal process, then a flat spectrum implies independence, but not otherwise. For this reason, and in order to complete the model, we examine the distribution of the fourth residuals.

Figure 19 shows, in the top panel, a crude histogram of the fourth residuals, together with sample statistics. The estimated kurtosis value of 3.61, which for the sample size of $N = 4380$ is significantly different from zero, indicates non-normality in the residuals. This is confirmed in the probability plot in the lower panel. This is a plot of the ordered residuals, say $X_{(i)}$, against their approximate expected value $\Phi^{-1}(i/N+1)$ where $\Phi^{-1}(.)$ is the inverse normal probability function. The curvature in this plot and the outliers indicate that the distribution of the $\varepsilon(t)$'s is somewhat non-normal, in fact pinched in the middle and long in the tails. The distribution appears symmetric; this is reflected in the estimated coefficient of skewness of -0.00147.

No attempt has been made to model these $\varepsilon(t)$'s in more detail. Their standard deviation is, from the last line of Table 4, estimated to be 0.537. The autoregressive parameters, a_1 and a_2, for the AR(2) process are estimated in Table 3 to be 0.9343 and -0.0975 respectively and the coefficients of $s(t)$ are given in Table 5. The long-term trend is $\hat{m}(t) = 10.9 + 0.000374t$.

MARGINAL DISTRIBUTION OF FOURTH RESIDUALS

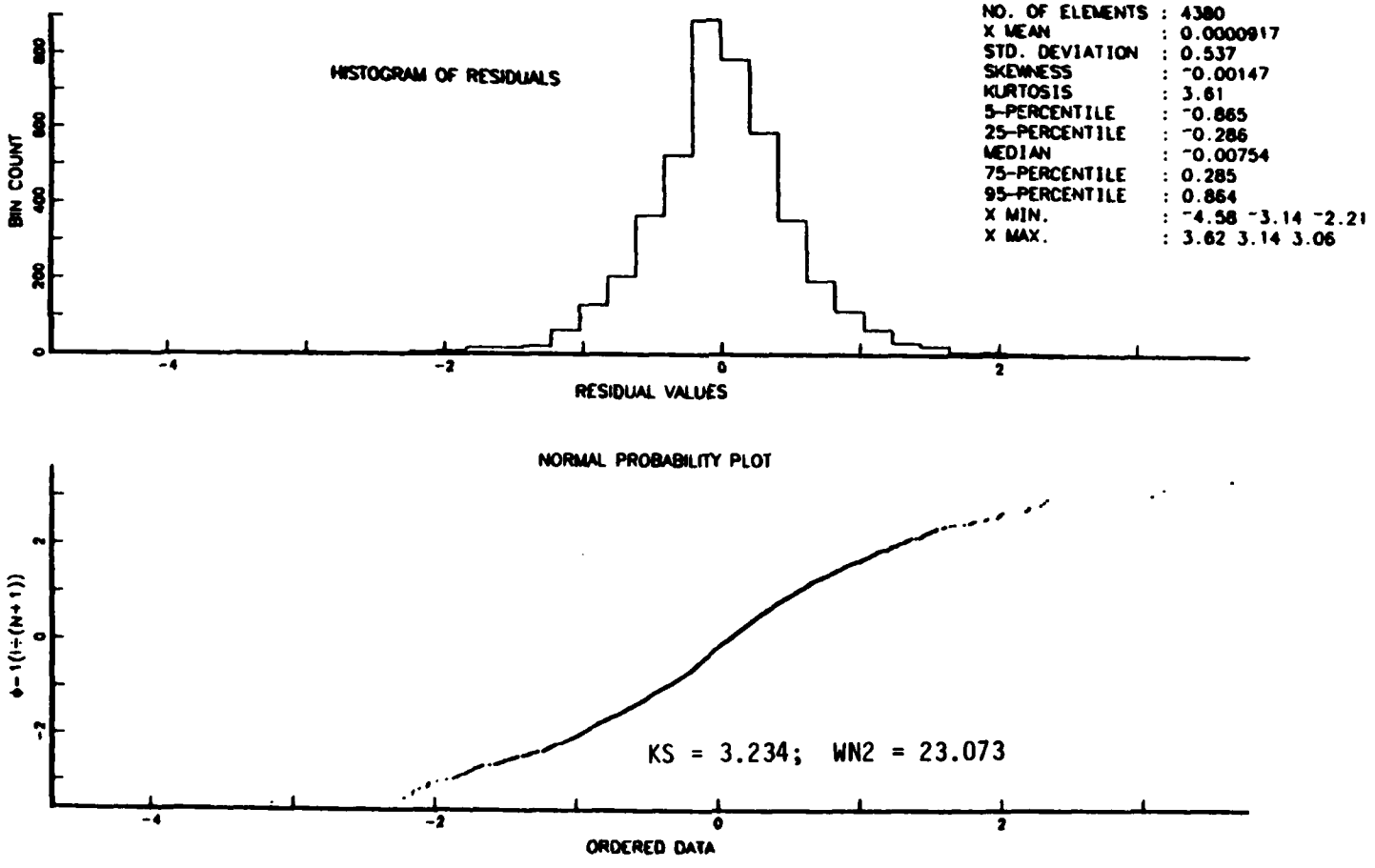


FIGURE 19

7. Comments on the model and the modelling

In Figure 8 we have given a local smoothing of the data which shows the main, gross features of the SSTs. Although there are many ways in which the data can be smoothed and there is the choice of an appropriate bandwidth, the procedure is free of the need to postulate a global model. In fact the smoothing (Figure 8) and the global fit of $\hat{s}(t)$ in Figure 11 give very similar information. Why then consider global models of $m(t)$ and $s(t)$ and the subsequent modelling of $\varepsilon(t)$?

The answer lies in our objectives.

(1) We want to use the model to identify relevant "scales" in the data? The question of scales is rather nebulous, with different answers coming out of different parts of the model. With regard to $s(t)$, there is clearly a "scale" of one year and longer scales that apply, for example, to the occurrences of El Niños. Locally, the high correlation in $\varepsilon(t)$ gives another scale, measured perhaps by the autocorrelation function of the $\varepsilon(t)$'s given in Figure 14. And, of course, if diurnal effects are of interest, then there is a scale of interest shorter than the one day sampling interval used to collect the data.

(2) With respect to scales, we want to know how often should the data be sampled? This is a question which cannot be answered without a specific objective in mind. For instance, if we want information about daily variation, we need to sample more often than once a day. If we are concerned with longer-term variation, we can relate the sampling rate to the accuracy of predictions. One component of the prediction is the cyclic and long-term trend; how accurately are they

estimated and what would be lost by sampling once every two or three days? In Section 9 we examine sampling variation in the parameters in m(t) and s(t) by generating (simulating) independent realization of SST based on the model.

(3) The model can be used to interpolate missing data. This is not a major problem in the Granite Canyon data set and thus it will not be discussed. However, gaps are a frequent problem with many oceanographic data sets (e.g. in the Farallon Island data mentioned previously).

(4) The model can be used to predict or forecast future temperatures, i.e. given the data up to today, what will they be for several subsequent days. We show in the next section how the model is used to give predictions for up to two days in advance of current data. These predictions will be shown to be better than the usual climatological prediction based on averages.

If we attempt to use a climatological model (i.e. an average across the 12 years or an average across the smoothed values for 12 years) to address the questions above, we sacrifice:

(1) The variation from year-to-year, so that events like El Niños and the long term trend represented by m(t) are masked by the averaging.

(2) The local information represented by $\varepsilon(t)$. It seems intuitively reasonable to make use of the water temperature today if we need to make a prediction for tomorrow. Quantitatively, Table 4 has shown that this local information accounts for 28.7% of the variation in the data.

(3) Decomposition of the different factors affecting sea-surface temperature, retaining only a picture of the gross behavior. Also we can no longer consider separately the yearly, 6 monthly, or tidal cycles.

Hence, we find the model to be more useful than climatology for summarizing the data. We use the next two sections to demonstrate its value.

8. Forecast of SSTs from March 1, 1983 to March 30, 1983

We now forecast SSTs, one and two days ahead, using the model developed in the previous sections. Since data are now available for the month of March, 1983, a period beyond the 12 year period used in constructing the model, predictions are made for comparison with the observations. The previous twelve years of data are used to estimate the parameters in the model. These forecast values are compared to the actual observations and to a climatological forecast using the average temperature for each given day across the twelve year ensemble.

Given an autoregressive process, i.e, the second residuals $r_2(t)$ as in equation (4.1) of Section 4, with known coefficients a_1 and a_2,

$$r_2(t) = a_1 r_2(t-1) + a_2 r_2(t-2) + \varepsilon(t),$$

and that the i.i.d residuals are normally distributed with mean zero, then the best predictors one and two steps ahead are obtained by setting $\varepsilon(t+2)$ and $\varepsilon(t+1)$ equal to their expected values of zero. We obtain

$$\text{Pred}\,(r_2(t+1)) = \hat{a}_1 r_2(t) + \hat{a}_2 r_2(t-1)\ ,$$

and

$$\text{Pred}\,(r_2(t+2)) = (\hat{a}_1^{\,2}+\hat{a}_2) r_2(t) + \hat{a}_1 \hat{a}_2 r_2(t-1)\ ,$$

where $\hat{a}_1$ and $\hat{a}_2$ are least squares estimates of a_1 and a_2 based on the observed 12 years of data. For periods far enough ahead, these predicted values would eventually converge to zero. This is simpler to see if $a_2 = 0$ so that we have an AR(1) process. Then $\hat{P}red(r_2(t+k)) = \hat{a}_1^k \, r_2(t)$.

We also have estimates $\hat{m}(t)$ and $\hat{s}(t)$ for the linear and seasonal trends for t=1 to 4380, and these can be computed beyond that range to give projections $\hat{m}(t+1)$, $\hat{m}(t+2)$, and $\hat{s}(t+1)$, $\hat{s}(t+2)$. The forecasts are then

$$Pred(y(t+1)) = \hat{m}(t+1) + \hat{s}(t+1) + Pred(r_2(t+1))$$

$$Pred(y(t+2)) = \hat{m}(t+2) + \hat{s}(t+2) + Pred(r_2(t+2))$$

These forecasts have been computed for the moving twelve year period out to March 30, 1983, together with the climatological forecast, and all are given in Table 6.

Table 6

Forecasts of one month's SSTs at Granite Canyon

Date	t	Actual y(t)	One-step Forecast	Error	Two-step Forecast	Error	One-step Climatol. Forecast	Error
March 1, 1983	4381	14.3	14.0	-0.3			11.5	-2.8
	4382	14.4	14.0	-0.4	13.7	-0.7	11.4	-3.0
	4383	14.2	14.1	-0.1	13.7	-0.5	11.5	-2.7
	4384	14.5	13.9	+0.6	13.8	-0.7	11.4	-3.1
	4385	14.3	14.2	-0.1	13.6	-0.7	11.2	-3.1
	4386	14.4	14.0	-0.4	13.8	-0.6	11.3	-3.1
	4387	14.2	14.1	-0.1	13.7	-0.5	11.4	-2.8
March 8, 1983	4388	14.8	13.9	-0.9	13.8	-1.0	11.6	-3.2
	4389	14.9	14.5	-0.4	13.6	-1.3	11.6	-3.3
	4390	14.7	14.5	-0.2	14.1	-0.8	11.5	-3.2
	4391	14.8	14.3	-0.5	14.1	-0.7	11.4	-3.4
	4392	14.0	14.4	+0.4	14.0	0.0	11.6	-2.4
	4393	14.2	13.7	-0.5	14.1	-0.1	11.1	-3.1
March 14, 1983	4394	14.5	13.9	-0.6	13.5	-1.0	11.0	-3.5
	4395	13.6	14.2	+0.6	13.7	+0.1	10.7	-2.9
	4396	13.5	13.3	-0.2	13.9	-0.4	11.0	-2.5
	4397	13.8	13.3	-0.5	13.2	-0.6	10.9	-1.9
	4398	13.7	13.6	-0.1	13.2	-0.5	10.9	-1.8
	4399	13.7	13.5	-0.2	13.4	-0.3	11.0	-2.7
	4400	13.7	13.5	-0.2	13.3	-0.4	10.9	-2.8
March 21, 1983	4401	13.5	13.5	0.0	13.4	-0.1	11.1	-2.4
	4402	13.8	13.3	-0.5	13.4	-0.4	11.0	-2.8
	4403	13.2	13.6	+0.4	13.2	0.0	11.0	-2.2
	4404	13.5	13.0	+0.5	13.4	-0.1	11.0	-2.5
	4405	13.3	13.4	+0.1	12.9	+0.4	10.8	-2.5
	4406	12.7	13.2	+0.5	13.2	+0.5	10.6	-2.1
	4407	12.6	12.6	0.0	13.0	+0.4	10.2	-2.4
March 28, 1983	4408	12.5	12.6	+0.1	12.6	+0.1	10.2	-2.3
	4409	12.4	12.5	+0.1	12.5	+0.1	10.2	-2.2
	4410	12.8	12.4	-0.4	12.5	-0.3	10.5	-2.3
					12.4			
			Av	-0.14	Av	-.38	Av	-2.7
			s.d	0.37	s.d	0.42	s.d.	0.45
			m.s.e	0.40	m.s.e.	0.57	m.s.e.	2.74

First one can see that the one-day-ahead forecast Pred(y(t+1)) and the two-day-ahead forecast, Pred(y(t+2)), are usually too low (i.e. negatively biased), the average differences being -0.14 and -0.38 respectively. The climatological forecast is even worse. Of course, this is probably due to the effect of the current El Niño. The global prediction of m(t)+s(t) cannot quite track the anomalous warming, although the autoregressive residuals

compensate to some extent. The climatological forecast is quite far off and the comparison because of the El Niño is probably not fair, but, nevertheless, the model given in this paper does give a better forecast. Note that the s.d. of the of the global forecast increases as one goes from one step ahead to two steps ahead, as does the bias. This is to be expected since less local information is available for the two-day-ahead prediction. A better over-all measure of the goodness of the forecast is the mean square error. This is given in all three cases in the last row of Table 6.

9. Simulation results

Here we use the model for the Granite Canyon SSTs to generate independent replications of the time series. These can be considered as members of an ensemble from such a process, with the proviso that the model and the model parameters used to generate the time series are those estimated from the original Granite Canyon series.

One such realization is shown in Figure 20. The top panel is the original series and the bottom panel is the simulated series. Note that the simulated series is generated out one year beyond March 1, 1983 and this projection is shown in Figure 21. This is not a true forecast, except insofar as the trend m(t) and the cycles s(t) account for a major part of the variation in the Granite Canyon SSTs. We note however, that the one year simulation (1983-84) does contain an apparent spring transition and thus introduces a note of reality into the results.

Ten independent realizations of the Granite Canyon SSTs were generated and used to compute estimated parameter values. In these simulations, the "white noise" sequence n(t) was taken to be i.i.d Normal random variables, despite the departures from Normality shown in Figure 19. An alternative would have been to "bootstrap" the process by using a random permutation of the $r_4(t)$'s.

ONE-YEAR MODEL EXTRAPOLATION AT GRANITE CANYON

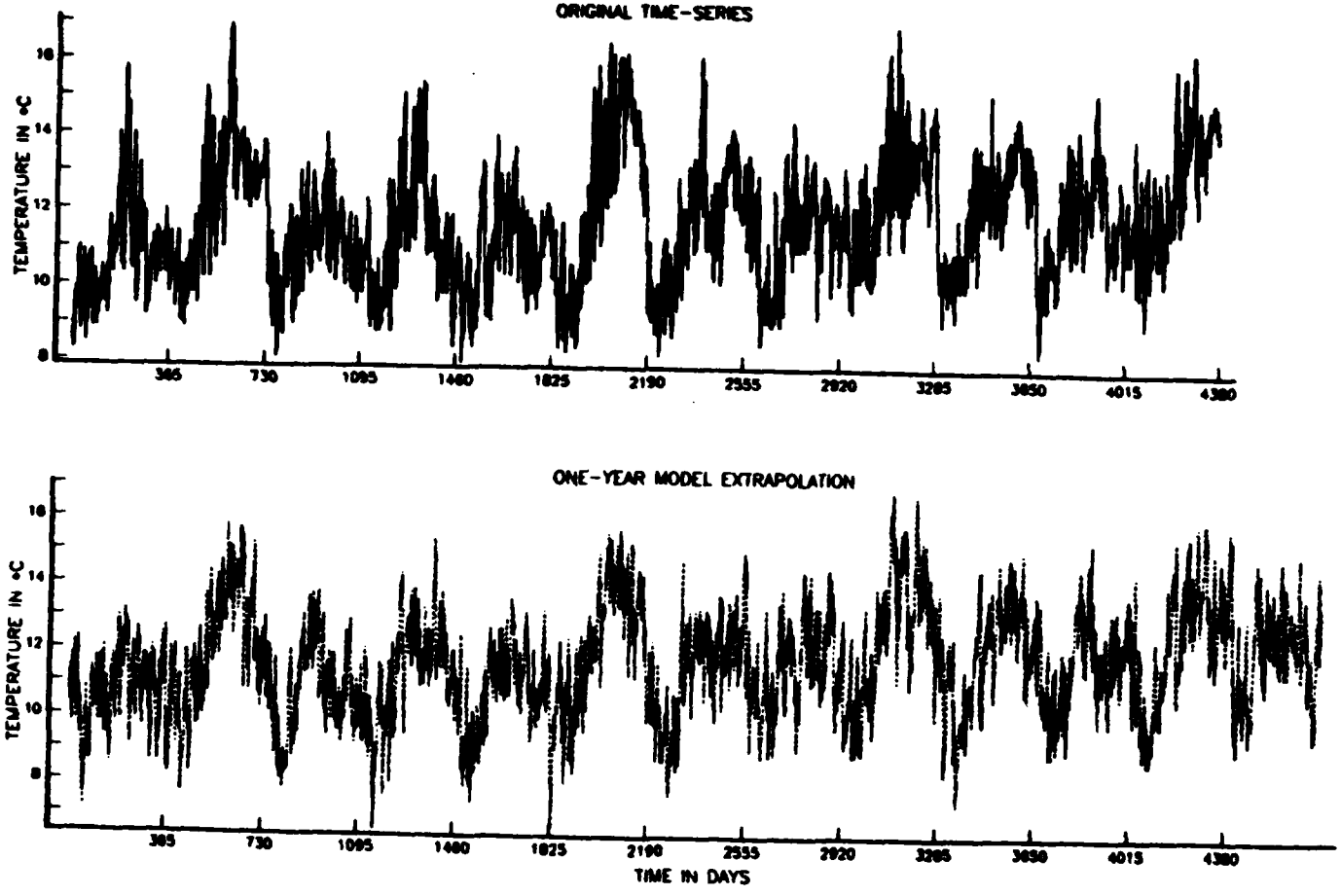


FIGURE 20

ONE-YEAR SIMULATED PROJECTION AT GRANITE CANYON

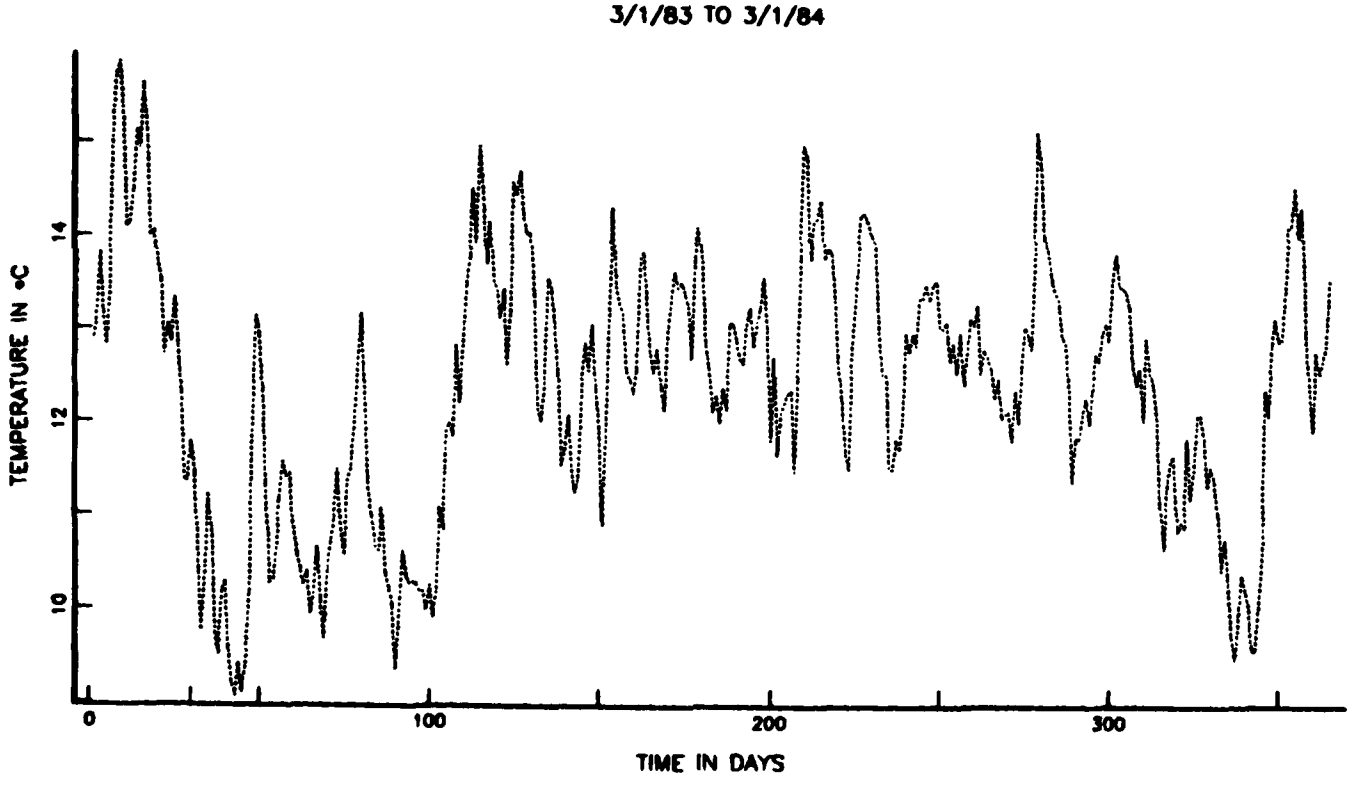


FIGURE 21

In Table 7 we give the results of the simulation for the parameters α and δ in the long term trend, for the serial correlations at lags one and two in the second residuals $r_2(t)$, for the associated coefficients a_1 and a_2 in the AR(2) model, and for the amplitudes A_1 and A_2 for the two dominant frequencies in s(t) with periods of one year and two years. These are also denoted A_1 and A_2 in Table 5. The nominal values are the values estimated from the real data which were used to generate the simulations.

Table 7

Simulation results - sampling every point

Parameter	Nominal value	Min	Max	Average	s.d.
α	10.9	10.78	10.95	10.91	0.045
δ	0.000374	0.000350	0.000391	0.000360	0.000011
$\rho(1)$	0.8511	0.8251	0.8543	0.8403	0.0074
$\rho(2)$	0.6973	0.6559	0.7067	0.6752	0.0141
a_1	0.9343	0.890	0.961	0.929	0.0187
a_2	-0.0975	-.136	-0.078	-0.106	0.0191
A_1 (1yr)	1.275	1.217	1.334	1.278	0.036
A_2 (2yr)	0.402	0.360	0.429	0.401	0.023

Since the estimated s.d. of the slope of the (linear) long term trend $m(t) = \alpha+\delta t$ is 0.000011, it is clear that the tilt in the data is statistically significant. Note that gross approximations give the s.d.'s of the estimates of $\rho(1)$ and $\rho(2)$ in a sample of size 4380 as $1/(4380)^{1/2} = 0.015$ which, given statistical fluctuations, are in reasonable agreement with the estimated s.d.'s of the serial correlations of 0.0074 and 0.0141. Note, however, that the estimates of $\rho(1)$ and $\rho(2)$ have averages below the nominal values. This reflects the known fact that estimates of serial correlations with high values are biased;

there is a technique for removing this kind of bias called jacknifing (Miller, 1974) and this could have been used. Again, the estimated precision of the estimate of A_1 is $0.036/1.275 \times 100\% = 2.8\%$, which is suitably low.

To investigate the problem of sampling rate, the simulations in Table 1 were run to produce estimates of α, δ, A_1 and A_2 based on every 2nd point in the generated series (Table 8) and every third point (Table 9).

Table 8 (every 2nd point)

Parameter	Nominal Value	Min	Max	Average	s.d.
α	10.9	10.79	10.95	10.91	0.042
δ	0.000374	0.000352	0.000387	0.000360	0.000010
A_1	1.275	1.215	1.330	1.277	0.035
A_2	0.402	0.362	0.429	0.400	0.023

Table 9 (every 3rd point)

Parameter	Nominal Value	Min	Max	Average	s.d.
α	10.9	10.78	10.95	10.91	0.049
δ	0.000374	0.000348	0.000393	0.000362	0.000012
A_1	1.275	1.213	1.342	1.279	0.037
A_2	0.402	0.369	0.429	0.403	0.022

First note that the estimated standard deviations in the last columns in Tables 7 and 8 are virtually identical. For independent data the s.d.'s in Table 8 would be approximately $\sqrt{2}$ times those in Table 7. The results therefore indicate that day-to-day correlation in the data is so high (local scales are so long) that sampling every day does not give any improvement to the sampling of long scale (one year and up) phenomena. Such phenomena are effectively summarized by the model through $s(t)$ and $m(t)$, exclusive of $\varepsilon(t)$.

In Table 9, results are given for estimation with sampling every third day. The standard deviations are beginning to increase, but not sufficiently to justify any finer sampling.

10. Oceanographic aspects of the data

10.1 Spring Transitions

Abrupt decreases in SST have been observed in the annual temperature cycle at Granite Canyon during the spring and appear to signal the onset of upwelling along the central California coast. These dramatic changes in temperature occur in at least 6 out of the 12 years considered and have been individually identified and tabulated previously in Table 1. From the table it is seen that these transitions typically occur in March and last for about a week. Also the associated temperatures drop from about 13 to about 10C. We have examined one spring transition in particular, the spring transition in 1980 starting on March 11th, which is shown in Figure 22. During this transition, temperatures dropped about 4C over a period of 7 days. As indicated previously in Figure 3, the entire water column experienced this transition down to at least 500 m. Alongshore, currents measured at the same location reversed from poleward to equatorward in direction at each of four depths between 113 and 510 m at the same time as the decreases in temperature occurred. The synoptic-scale winds at this location however did not become upwelling favorable until just after this event occurred. The latter observation suggests that this event may have been caused by wind forcing at another location. During a similar spring transition during 1981, Brink, Stuart, and Van Leer (1983) also concluded that local wind forcing was not responsible for the transition they observed off Pt. Conception during that year. Finally IR satellite imagery acquired a few days before and immediately following the event (not shown) indicated a change from isothermal surface conditions to conditions suggestive of intense upwelling in the vicinity of Pt. Sur.

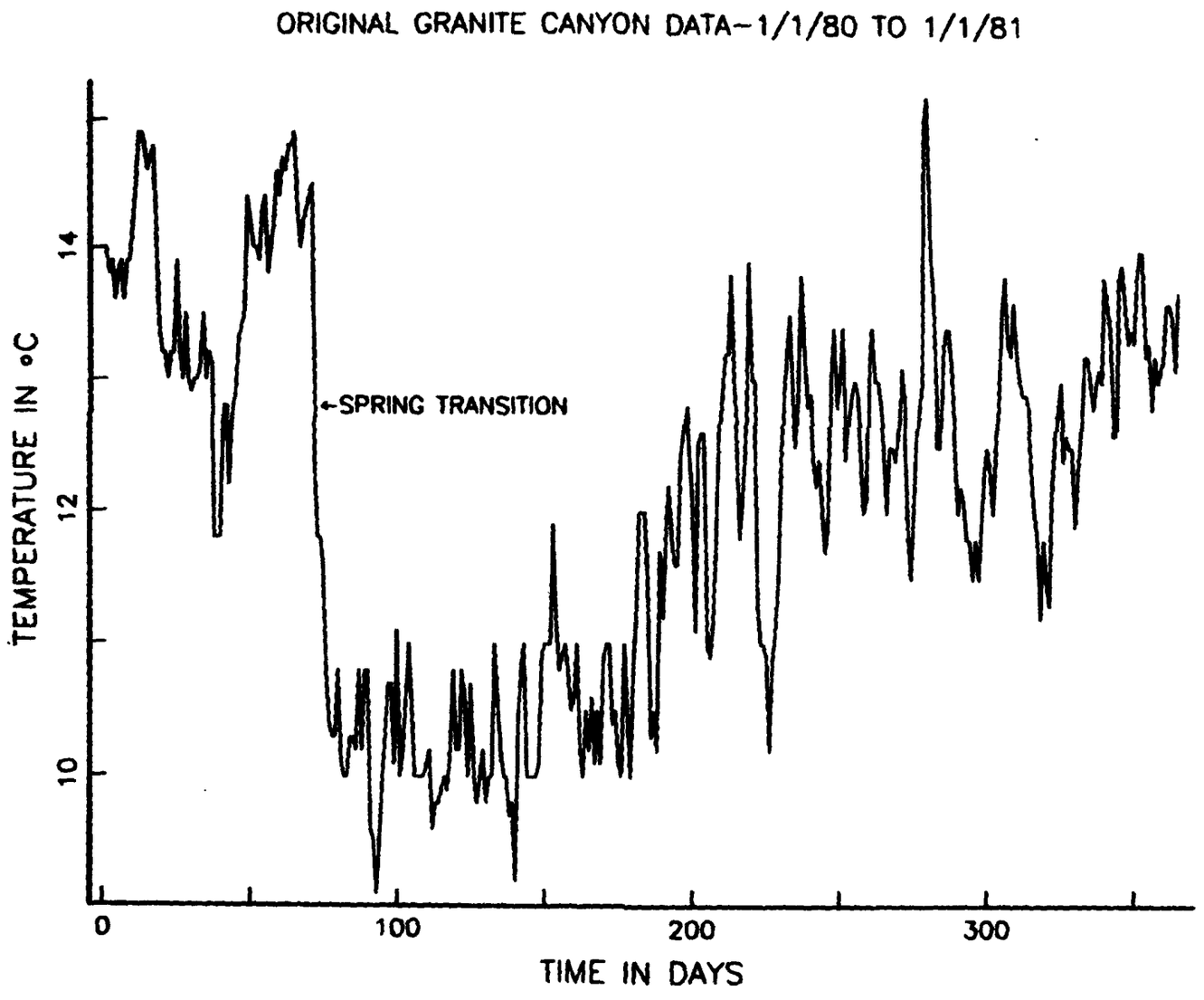

ORIGINAL GRANITE CANYON DATA–1/1/80 TO 1/1/81
TEMPERATURE IN °C
10
12
14
←SPRING TRANSITION
0
100
200
300
TIME IN DAYS

FIGURE 22

The relative times at which the 1980 spring temperature transition arrived at the four locations along the California coast for which we have data are seen in Figures 4, 5, and 6. This event is seen to arrive first at Diablo Canyon, second at Granite Canyon, third at the Farallons, and finally at Pacific Grove. Phase speeds for poleward propagating disturbances along the coasts of North and South America generally range between about 35 and 280 Km/day. In particular, Allen and Romea (1980) indicate that phase speeds of the order of 175 Km/day should be expected for barotropic continental shelf waves along the South American coast south of about 18°S. We take this value as representative along the North American coast as well, since the continental shelf-slope regions along both South and North America are generally similar (Shepard, 1973). The difference in arrival time between Diablo Canyon and Granite Canyon, which are about 175Km apart, is approximately a day, a time separation consistent with the propagation speed of 175 Km/day for barotropic continental shelf waves in this region. However, the separation in arrival times between Granite Canyon and the Farallons is considerably greater (~ 6 days) even though the distance between these points is again about 175 Km. The time separation in this case is not consistent with propagation of continental shelf waves since phase speed should in fact increase due to the slight increase in the Coriolis parameter and an increase in width of the continental shelf in this region. Consequently we conclude that these results are not necessarily consistent with shelf wave dynamics. We also note that the sea surface temperatures are only sampled once a day. Thus our ability to resolve the phase speeds of these waves is limited.

10.2 Annual cycles

A mean annual temperature cycle has been calculated for the Granite Canyon data, realizing from our previous statistical calculations that a mean value may not be a truly representative measure of central tendency in our data. Points

corresponding to the average temperatures and a smoothed version of that average (with a window of 25 points) are both presented in Figure 23. According to List and Koh (1976), and Reid, Roden and Wyllie (1958), a well-defined annual cycle in SST might not be expected off Granite Canyon, due to warming during the winter from the Davidson Current and cooling during the spring and summer from coastal upwelling. To the contrary, Figure 23 shows a well-defined annual cycle at Granite Canyon, with a range of about 3C. However the annual range of variability in temperature is reduced at Granite Canyon compared to the range of variation usually found at this latitude. This is shown in the insert in Figure 23 where deep ocean SSTs (Robinson, 1976) are compared to the Granite Canyon SSTs. The effect of averaging the data has been in part to "smear" the occurrence of the spring transition over a two-month period from about mid-February to mid-April. Maximum temperatures are seen to occur during mid-October, about a month later in the year than might be expected at this latitude. Obviously the phase relationships between such factors as the onset and cessation of the Davidson Current, and the onset and cessation of coastal upwelling determine in part the detailed structure of the annual temperature cycle off Pt. Sur in any given year.

Skogsberg (1936) originally specified three oceanic seasons for Monterey Bay, an upwelling season from about February to July, an oceanic season from about July to November, and a Davidson Current season from about November to February. According to the smoothed average annual cycle presented in Figure 23, at least two "seasons" can be identified, an upwelling season and a non-upwelling season. Based on our data the distinction between an oceanic season and a Davidson Current season is not clear. We note however that Skogsberg used spatially distributed temperature data as well as chemical and biological observations to construct his seasonal description. Also his

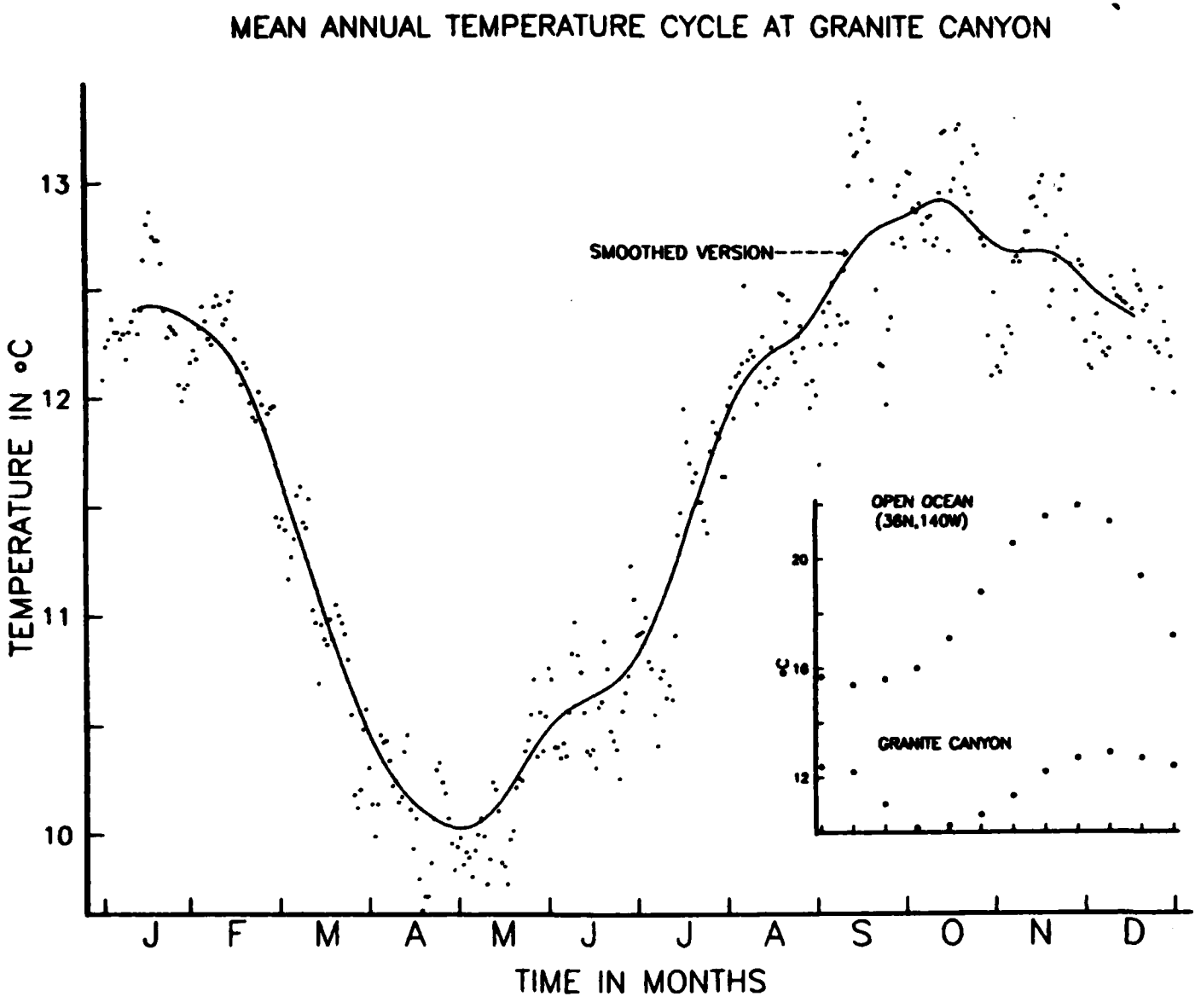

MEAN ANNUAL TEMPERATURE CYCLE AT GRANITE CANYON
SMOOTHED VERSION
TEMPERATURE IN °C
13
12
11
10
J
F
M
A
M
J
J
A
S
O
N
D
TIME IN MONTHS
OPEN OCEAN
(36N,140W)
20
°C
16
GRANITE CANYON
12

FIGURE 23

seasonal model was derived for Monterey Bay and as a result it may not be valid to extrapolate his model to regions outside the Bay.

10.3 El Niño events

In addition to the seasonal variation in temperature at Granite Canyon, inter-annual variability is also evident, as mentioned previously in Section 3. The large increases in temperature seen during the winters of 1972, 1976, and 1983 correspond to major El Niño warming events that took place during those years. These events have also been observed off central California through biological measurements and measurements of sea level by Chelton, Bernal and McGowan (1982). The increase in temperature seen during the latter part of 1979 is not associated with an El Niño warming event (Enfield, 1983).

Of particular interest is the present major El Niño warming event. This warming was shown to extend at least 500 Km off the coast of central California during February 1983 (Linder and Breaker, 1983). The effects of the present warm anomaly are clearly evident in Figure 24 which compares the 3/1/82 to 3/1/83 time period (i.e., the solid curve) to the 12 year average data (i.e., the dashed curve). Early in 1982 warming off Granite Canyon appeared to be intermittent. However, by late 1982 warming at this location had become well-established. As of February 1983, the warm anomaly at Granite Canyon had reached almost 3C. A smoothed version of Figure 24 is shown in Figure 25 and helps clarify the present departure in SST from the 12-year mean at this location.

10.4 Interpretation of higher frequency variations

From an oceanographic viewpoint, the periodogram shown previously in Figure 9 is somewhat surprising for its lack of any major spectral components beyond 100 cycles. For example, spectra of winds along the California coast

COMPARISON OF SEA-SURFACE TEMPERATURE AT GRANITE CANYON
BETWEEN PAST YEAR AND 12-YEAR MEAN

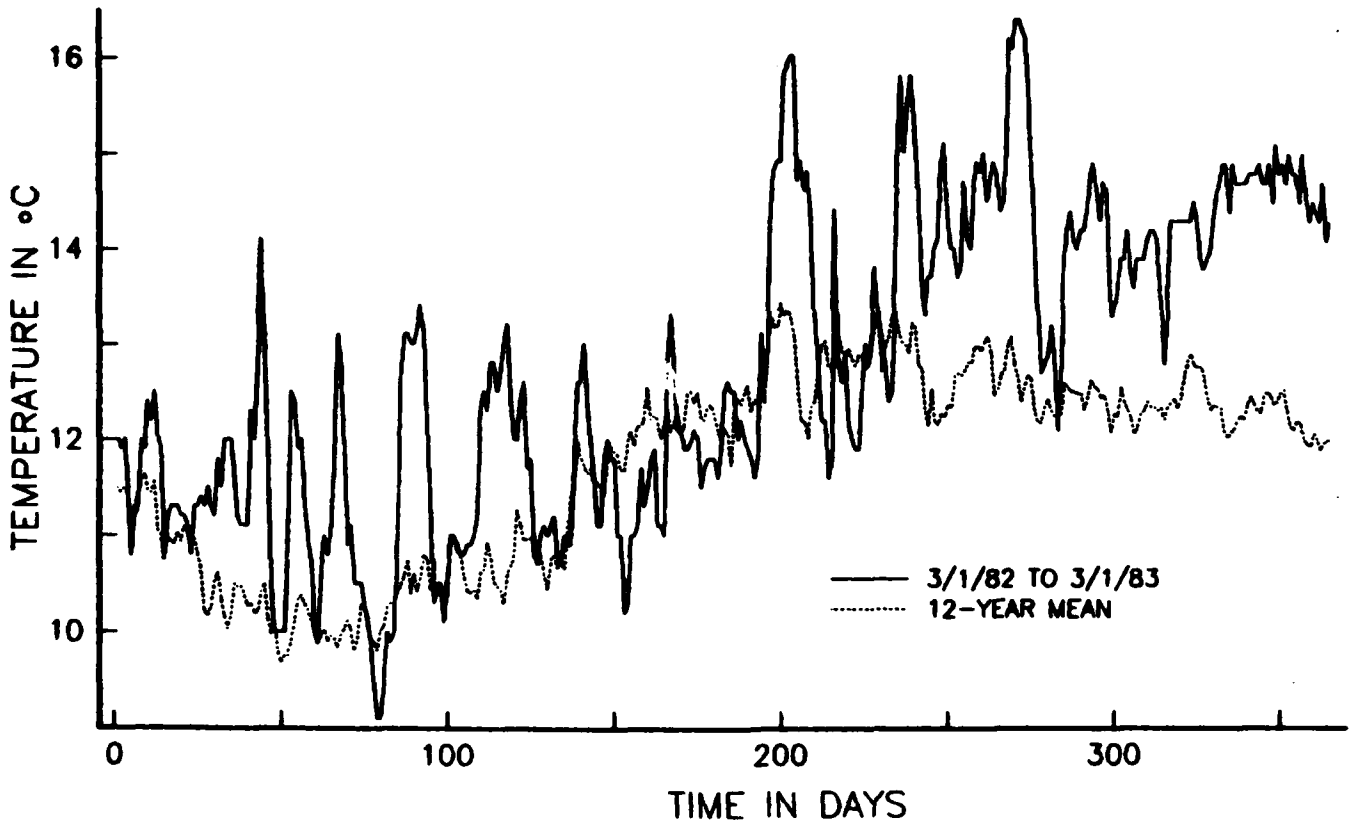


FIGURE 24

GRANITE CANYON SEA–SURFACE TEMPERATURE COMPARISON
SMOOTHED

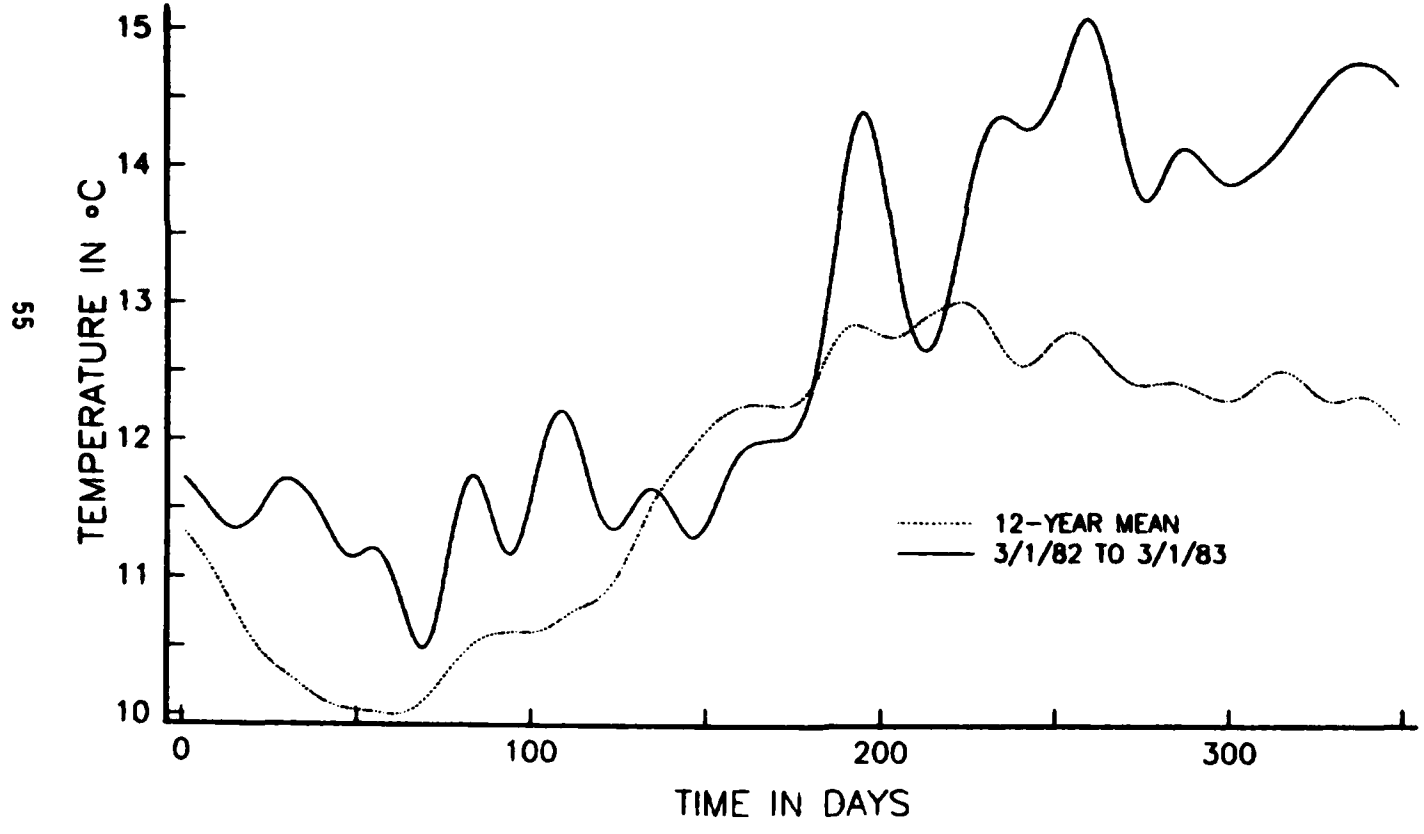


FIGURE 25

often show higher amplitudes at periods between 3 and 10 days (Van Patten, 1981). There are three possible explanations for this apparent lack of consistency between wind and temperature spectra. First, the periodic effects in the winds may be real, but continual stirring of the near-surface environment where observations of ocean temperature are made, may inhibit these periodic influences on SSTs. Alternatively, higher amplitudes in the coastal wind spectra may be caused by correlation in the wind data rather than by underlying periodicities. As such, a periodic-like response in SST to the local winds should not necessarily be expected. A third possibility of course is that the spectrum of the wind field at this particular location has no discrete components to start with. Data to check this third possibility unfortunately do not exist over the full 12 year period but there is evidence from other locations (Hugus, 1982) that 3 to 10 day periodicities do not show up consistently in wind observations.

Originally the possibility of aliasing by the semiduirnal and/or diurnal tides was a concern. Analysis of similar observations at other locations along the California coast by List and Koh (1976) indicated that tidal aliasing (or a natural periodicity corresponding to the spring/neap tides) by the semidiurnal tide at 14.78 days could be important. To the contrary we do not find evidence that the alias of either tidal component (148 cycles for the diurnal tide and 296 cycles for the semidiurnal tide) appears in the calculated periodogram.

In examining the periodogram of second residuals, particularly at high frequencies, it is of interest to briefly consider the likely influence of turbulence on these results. Turbulence production undoubtedly occurs from the breaking of incoming surface waves and perhaps from local drift currents as well. In this regard a portion of the periodogram of second residuals has been

smoothed and replotted on log coordinates in Figure 26. The uniform decrease in spectral amplitude with increasing frequency is reminiscent of the spectral decay often associated with the inertial subrange from homogeneous turbulence theory (Batchelor, 1953). Between 200 (~ 2 days) and 2000 cycles (~22 days), the spectral slope is essentially constant at -9/5, close to the value of -5/3 predicted by classical turbulence theory (Phillips, 1977). According to this theory temperature is assumed to be a conservative quantity, an assumption which could be quite unrealistic at the sea-surface. In our case time replaces distance which is normally taken as the independent variable, and if it is assumed that these coordinate frames are linearly related, then the spectral slopes observed here are directly comparable with those usually derived in oceanic turbulence calculations as a function of distance. In turbulence models where temperature is not assumed to be conservative, spectral slopes considerably greater than -5/3 are predicted (Panchev, 1980). Measurements of the spatial surface temperature spectrum by Saunders (1972) and Holladay and O'Brien (1975) for example, yielded spectral slopes of -2.2 and -3 respectively, suggesting that temperature should not be treated as a conservative quantity at least under certain conditions.

In conclusion, our earlier analyses suggest that the (likely) influence of turbulence in our data corresponds to the autoregressive component in our statistical decomposition since the spectrum of the fourth residuals is essentially flat. A more detailed analysis of our data would include a comparison of the spectra for various orders of autoregression with spectra predicted from various theoretical models of ocean turbulence.

PARTIAL SPECTRUM OF SECOND RESIDUALS

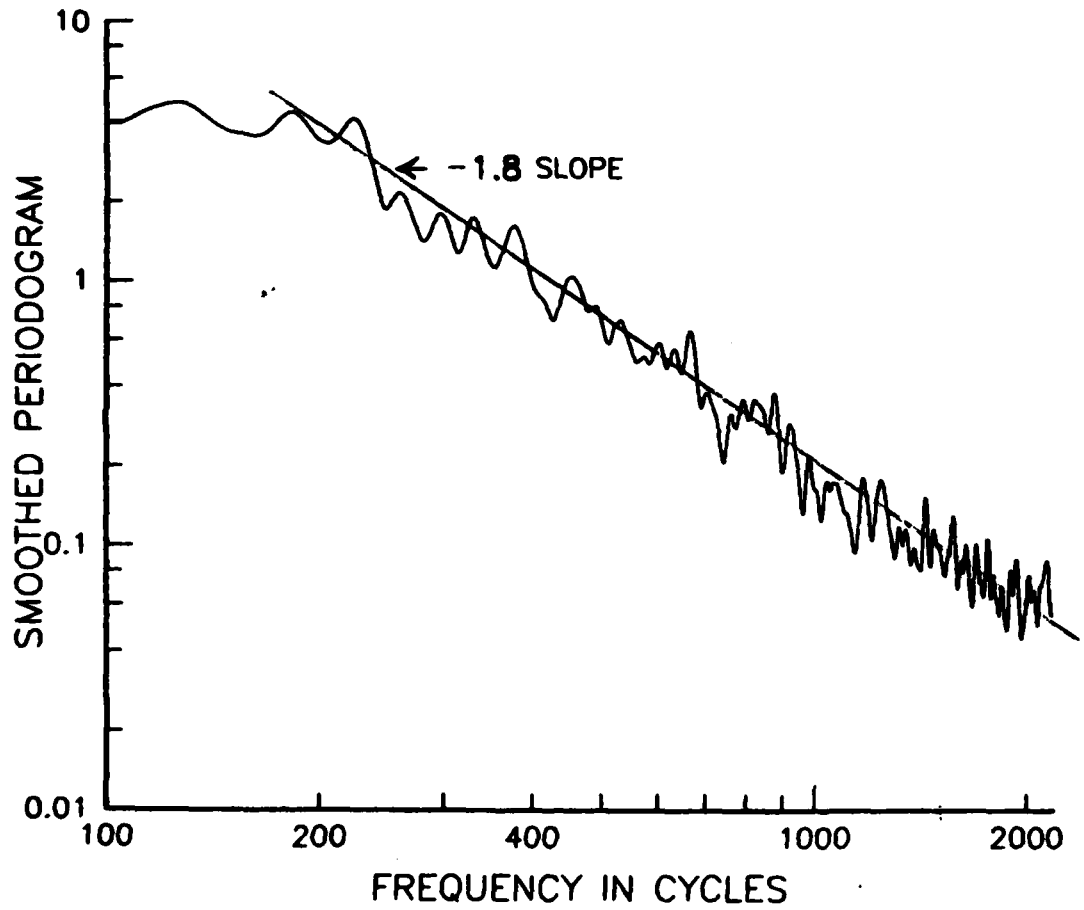


FIGURE 26

ACKNOWLEDGEMENTS: The work of P.A.W. Lewis and E.J. Orav was supported by the Office of Naval Research under Grant NR-042 and by the Naval Postgraduate School Foundation. The work of L. Breaker was supported by the Oceanography Department of the Naval Postgraduate School. The authors would particularly like to acknowledge Mr. Earl Ebert and the staff at the Granite Canyon Marine Culture Laboratory for their dedication in acquiring the data used in this report. Mr. Vern Wyman from PG&E is also thanked for furnishing the temperature data from Diablo Canyon. The subsurface temperature data were provided by J.B. Wickham and S.P. Tucker of NPS. The authors are indebted to C.N.K. Mooers for guidance and careful reading of the manuscript.

The graphs were produced by an experimental APL package GRAFST2 which the Naval Postgraduate School is using under a test agreement with IBM Watson Research Center, Yorktown Heights, N.Y. We are grateful to Dr. P.D. Welch and Dr. P. Heidelberger for making GRAFST2 available to us.

Finally we are indebted to Ellen Saunders for good typing and good cheer.

References

Allen, J.S., and Romea, R.D. (1980). On coastal trapped waves at low latitudes in a stratified ocean. J. Fluid Mech., 98,3, 555-585.

Anscombe, F. (1981). Computing in statistical science through APL. New York: Springer-Verlag

Batchelor, G.K. (1953). The theory of homogeneous turbulence. Cambridge: Cambridge University Press.

Bolin, R.L. and Abbott, D.P. (1963). Studies on the marine climate and phytoplankton of the central coastal area of California, 1954-1960. CalCOFI Rep., 9, 23-45

Bretschneider, D.E., and McLain, D.P. (1983). Sea level variations at Monterey, California. NOAA Tech. Rpt. NMFS SSRF-761.

Brink, K.H., Stuart, D.W. and VanLeer, J.C. (1983). Observations of the coastal upwelling region near 34°30'N off California: spring 1981. WHOI Contribution No. 5319.

Chatfield, C. (1980). The analysis of time series: an Introduction. 2nd Edition New York: Chapman and Hall.

Chelton, D.B., Bernal, P.A. and McGowan, J.A. (1982). Large-scale interannual physical and biological interaction in the California Current. Jour. Mar. Res. 40(4), 1095-1125.

Cox, D.R. and Lewis, P.A.W. (1966). The Statistical Analysis of Series of Events. New York: Chapman Hall

Enfield, D. (1983). Personal Communication.

Fisher, R.A. (1929). Tests of significance in harmonic analysis. Proc. Roy. Soc. Ser A, 125, 54-59

Gilmer, L. and Rose, A.J. (1976). APL An interactive approach. 2nd Edition., Rev. printing. New York: Wiley

Holladay, C.G., and O'Brien, J.J. (1975). Mesoscale variability of sea surface temperature. Jour. Phys. Oceanogr., 5,(5), 761-772.

Hugus, D.K., (1982). Extension of some models for positive-valued time series. Ph.D. Thesis, Naval Postgraduate School, Monterey, California.

Huyer, A., Sobey, E.J., and Smith, R.L. (1979). The spring transition in currents over the Oregon continental shelf. Jour. Geophys. Res. 84(011) 6995-7011.

Jenkins, G.M. and Watts, D.G. (1969). Spectral analysis and its applications. San Francisco: Holden-Day

Linder, D., and Breaker, L.C. (1983). Warm sea-surface temperatures along the California coast. Submitted to the Coastal Oceanography and Climatology News for publication.

List, E.J. and Koh, R.C.Y. (1976). Variations in coastal temperatures on the Southern and Central California coast. J. Geophys. Res. 81(12): 1971-1979.

Panchev, S. (1980). Spectral structure of horizontal oceanic turbulence - semi-empirical models, 43-55. in Marine Turbulence, edited by J.C.J. Nihoul. New York: Elsevier.

Phillips, O.M. (1977). The dynamics of the upper ocean, 2nd edition, Cambridge: Cambridge University Press.

Reid, J.L., Jr., Roden, G.I., and Wyllie, J.G. (1958). Studies of the California Current System. CalCOFI REP., 1 July 1956-1 January 1958, 28-56.

Robinson, M.K. (1957). Sea temperature in the Gulf of Alaska and in the Northeast Pacific Ocean, 1941-1952. Bull. Scripps Inst. Oceanogr. 7(1),

Robinson, M.K. (1976). Atlas of North Pacific Ocean monthly mean temperatures and mean salinities at the surface layer. Ref. Pub. 2, Naval Oceanographic Office, Washington, D.C.

Roden, G.I. (1963). On statistical estimation of monthly extreme sea-surface temperatures along the west coast of the United States. J. Mar. Res., 21, 172-190

Saunders, P.M. (1972). Space and time variability of temperature in the upper ocean. Deep Sea Res., 19, 467-480.

Schuster, A. (1898). On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. Terr. Mag. Atmos. Elect., 3, 13-41

SIO Reference 81-30,(1980). Surface water temperatures at shore stations. Data report, UCSIO.

Shepard, F.P. (1963). Submarine geology, 2nd edition. New York: Harper and Row

Skogsberg, T. (1936) Hydrography of Monterey Bay, California thermal conditions, 1929-1931. Am. Phil. Soc. Trans., 29.

Skogsberg T., and Phelps, A. (1946) Hydrography of Monterey Bay, California thermal conditions, Part II (1934-1937). Proc. Amer. Philos. Soc. 90(5), 350-386

Stewart, H.B., Jr., Zetler, B.D. and Taylor, C.B. (1958). Recent increases in coastal water temperature and sea level - California to Alaska. Tech. Bull. No. 3, U.S Cst. Geodet. Survey.

Tsiao, G.C. and Box, G.E.P. (1981). Modelling multiple time series with applications. J. Amer. Statist. Assoc., 76, 302-816.

Van Patten, B. (1980). Long-term spectra of wind and temperature in the vacinity of the marine inversion. Second Conference on Coastal Meterology, Los Angeles.

DISTRIBUTION LIST

	NO. OF COPIES
Library, Code 0142 Naval Postgraduate School Monterey, CA 93940	4
Dean of Research Code 012A Naval Postgraduate School Monterey, CA 93940	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93940	1
Professor P. A. W. Lewis Code 55Lw Naval Postgraduate School Monterey, CA 93940	250
Chief of Naval Research Arlington, VA 22217	1


DA
FILM