AD-A102 685 BEDFORD RESEARCH ASSOCIATES MA F/G 4/1

APPLICATION OF METHODS OF NUMERICAL ANALYSIS TO PHYSICAL AND EN--ETC(U)

OCT 80 R BOUCHER, T COSTELLO, P MEEHAN F19628-78-C-0241

UNCLASSIFIED AFGL-TR-80-0347 NL

1 OF 2

AD

A102685

AD A102685

LEVEL II

(12)

AFGL-TR-80-0347

APPLICATION OF METHODS OF NUMERICAL ANALYSIS TO PHYSICAL AND ENGINEERING DATA

R. Boucher
T. Costello
P. Meehan
J. Noonan

Bedford Research Associates
2 DeAngelo Drive
Bedford, Massachusetts 01730

Final Report
September 1978 - September 1980

October 15, 1980

Approved for public release; distribution unlimited

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	AFGL-TR-80-0347
2. GOVT ACCESSION NO.	AD A102685
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	APPLICATION OF METHODS OF NUMERICAL ANALYSIS TO PHYSICAL AND ENGINEERING DATA.
5. TYPE OF REPORT & PERIOD COVERED	Final Report Sept. 1978 - Sept. 1980
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s)	R. Boucher T. Costello P. Meehan J. Noonan
8. CONTRACT OR GRANT NUMBER(s)	F19628-78-C-0241
9. PERFORMING ORGANIZATION NAME AND ADDRESS	Bedford Research Associates 2 DeAngelo Drive Bedford, MA. 01730
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	62101F 9993XXXX
11. CONTROLLING OFFICE NAME AND ADDRESS	Air Force Geophysics Laboratory Hanscom AFB, Massachusetts 01731 Monitor/ John F. Kellaher/SUWA
12. REPORT DATE	October 15, 1980
13. NUMBER OF PAGES	121
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)

ATMOSPHERIC TURBULENCE, IONOSPHERIC CRITICAL FREQUENCY,
DMSP = Defense Meteorlogical Satellite Program
OPAQUE = Optical Atmospheric Quantities in Europe

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

The four major projects for which analysis and software development were performed on this contract are described in this report. These deal with Atmospheric Turbulence Modelling, Estimation of the Critical Frequency of the Ionosphere, Processing of Optimal Atmospheric Measurements, and Analysis of Data from the Defense Meteorological Satellite Program (DMSP).

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TABLE OF CONTENTS

ATMOSPHERIC TURBULENCE MODELLING IMPACTING STRATOSPHERIC ENVIRONMENT PROGRAM

OBJECTIVES:

The primary objectives of this study can be described as follows:

1) Develop methods of describing the constituents and thermal distributions of the normal stratosphere and calculate how these distributions are affected by Air Force flight operations and atmospheric transport coefficients.

2) Characterize stratospheric disturbances that can be produced from mountain waves and underlying thunderstorm activity.

3) Develop specialized numerical techniques to incorporate radiation transfer and couple photochemical processes with the heat balance of the stratosphere.

This task is the key to the stratospheric environment project wherein the results of the various measurements are used to establish a realistic model of the stratosphere. The model, in turn, is used to predict if environmental changes will occur as a result of AF flight operation. This task is essential to the preparation of environmental impact statements for the Air Force.

The major effort of this work on turbulence modelling is to improve stratospheric models for the prediction of Air Force systems operations in the stratosphere. In particular, the probability of occurrence of turbulence for a given location and season as a function of altitude is a required input to the tropospheric/stratospheric model used to predict effects of Air Force Systems emission pollutants in the stratosphere. Thus we are investigating the seasonal variability of the characteristic layers of turbulence in the troposphere and stratosphere and determining the seasonal variations in the natural levels of NO_x trace constituents which are also engine combustion by-products. This will result in refinements in the troposphere/stratosphere 1-D and 3-D chemical/radiation transport models which are used to predict environmental effects, such as ozone depletion, caused by Air Force Systems emissions in the troposphere and stratosphere.

Meterological data (temperature, winds, pressure and relative humidity) in the height range 0-60 km on magnetic tape is currently being restructured to provide efficient, easy access for processing with minimum computer time. A statistical approach is being used to determine the seasonal probability of occurrence of turbulence as a function of latitude, longitude and altitude through the use of the Richardson criteria for the presence or absence of turbulence in the atmosphere. NO_x trace constitutents measured by LKD will be compared with data from other experimental techniques to determine the extreme of seasonal variability in these constituents. The turbulence and NO_x seasonal results will be used to parameterize the important role of transport processes and provide fine adjustments in the tropospheric/stratospheric chemistry of the 1-D and 3-D models.

INTRODUCTION:

The stratospheric seasonal variations task under the Stratospheric Environment Project is currently studying the general aspects of turbulence in the troposphere and lower stratosphere. The data being used in this study has been provided by the Environmental Data Service National Climatic Center in Asheville, No. Carolina. The "rawinsonde data" consists of winds, temperature and pressure as a function of altitude obtained from all available stations (currently 144 stations) for the period 1948-1976. At the present time only the 1970-1976 data have been reformated for processing on the CDC 6600 computer system. This data is being used in calculating the Richardson number, a stability criterion for determining the presence or absence of atmospheric turbulence, and is given by the following relationship:

$$R_i = g/_T \left[\frac{\delta T}{\delta z} + \Gamma\right] \Big/ \left[\delta U/_{\delta z}\right]^2$$

where: g is the acceleration of gravity

$\delta T/_{\delta z}$ is the vertical temperature gradient.

Γ is the dry adiabatic lapse rate (10^{o} $K/_{km}$)

and $\left[\frac{\delta U}{\delta z}\right]^2 = \left[(\delta V_{x}/_{\delta z})^2 + (\delta V_{y}/_{\delta z})^2\right]$ is the square of the vertical shear of the horizontal winds.

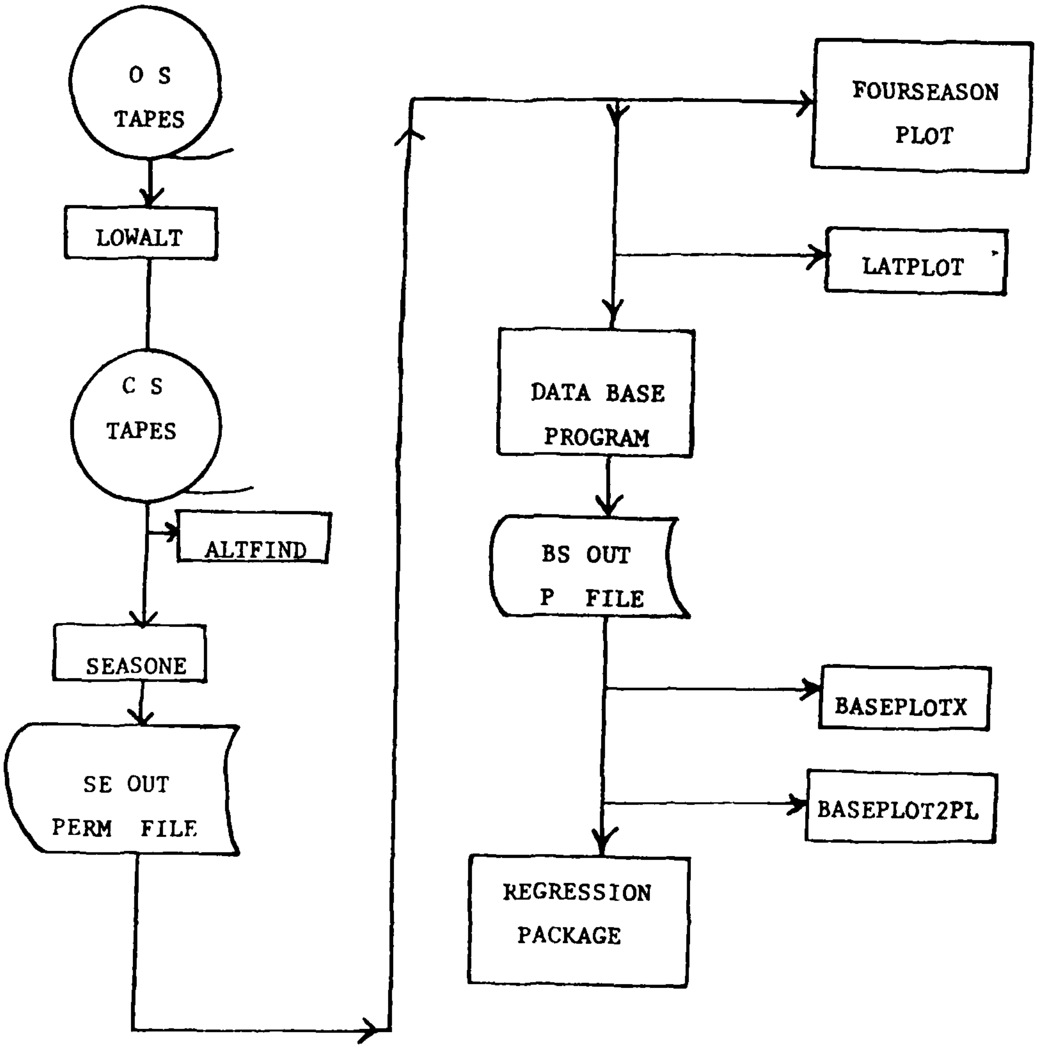
The rawinsonde system is not ideally suited to take measurements to provide thermodynamic data at the required sampling intervals for this application. The data being used in this study was originally intended to provide northern hemisphere pressure maps for a number of millibar pressure levels. In order to determine the Richardson number at equal height intervals, the temperature and wind component data points are fit using a Hermite interpolation algorithm. The temperature function along with the derivatives of the winds and temperature with respect to altitude are interpolated at 1/10 km intervals and used to calculate the Richardson number. The probability of occurrence of turbulence is obtained by defining a critical Richardson number (R_i = 1 to 1/4 according to current literature) and assuming turbulence present when the Richardson number is equal to or less than this value. The probability is then taken as the ratio of the total number of occurrences to the total number of measurements at a 1.0 km level over a season. These seasonal results are then averaged over one km and plotted as in the figure provided. Two techniques are used to establish the seasonal tropopause height. The seasonally averaged temperature gradient value is tested for a sustained change in sign which indicates the change in lapse rate from the troposphere to the stratosphere. Frequently, at high latitudes, the temperature remains fairly constant above the troposphere and through the stratosphere. Since a negative lapse rate may not develop, a second test is used which established the tropopause height as the maximum rate of change in the temperature gradient.

Initial results from stations at low, middle and high latitudes reveal interesting seasonal and latitudinal variation in the occurrence of turbulence as well as particular altitude levels which consistently reveal relatively high occurrences of turbulence. It is also interesting to note that turbulence rapidly decreases at heights which generally agree with those levels reported in models as the tropopause level and follows this pattern as the tropopause level decreases from its highest level in equatorial regions to its lowest levels at high latitudes. This particular result is substantiated in theory since the stratosphere region above the tropopause level is a dry stable region having a negative temperature lapse rate, and thus the degree of turbulent mixing and transport is greatly reduced from that occurring in the troposphere.

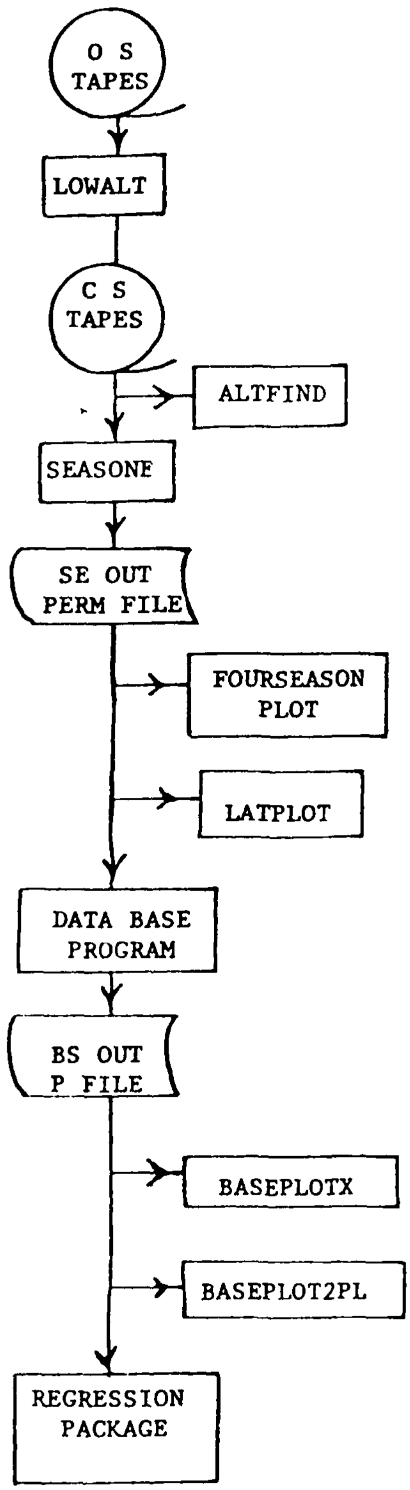
The resolution of the temperature and wind data (100m to a few kilometers) varies with the rise rate of the rawinsonde balloon and the criteria of selecting data points between mandatory levels (those standard millibar height levels recorded at all stations). The fit applied to the data to provide Richardson numbers at equal altitude intervals brings into question the validity of the Richardson number determination by possibly introducing effects not present in the thermodynamic structure of the atmosphere. The possibility of obtaining a limited amount of data at suitable sampling intervals to obtain some knowledge of the actual structure in the data has been investigated. Personnel at the Chatham, MA. rawinsonde station have cooperated in the initial investigation by demonstrating the present techniques in data acquisition during a visit to that site in September, 1978. They also provided at that time an analog strip chart recording of temperature along with the standard printout of thermodynamic data. Data points extracted from the temperature analog at the smallest possible interval indicate spectrum wave lengths of approximately a kilometer or longer.

TURBULENCE STUDY SOFTWARE

A description of the software used in this study is presented next. This is followed with a summary of statistical procedure which we have developed.



TURBULENCE STUDY FLOWCHART



FOURSEASON PLOT

The primary function of this plot is to use the Season E output and plot the Percent Occurrence of Turbulence (both RI $\leq$ .25 and RI $\leq$1.0) versus altitude. Each plot contains one year's data with all four seasons, and altitude 2 to 25 kilometers. Each plot presents the station name, number, longitude, latitude and year. Also incorporated into this program is the determination of the seasonal tropopause levels. This is accomplished and displayed using two approaches. Method one tests for a sustained sign change in the temperature gradient. Method two looks for the maximum change in slope of the temperature gradient. Output is on both pen plots and microfiche.

LATPLOT

This program uses data extracted from the Season E output. Created is a plot of achosen parameter versus latitude. Numerous stations arranged by latitude can be shown on one plot. The desired parameter with its standard deviation is plotted for a chosen altitude. Any or all seasons can be included on one plot. This plot is very good in showing relationships between station, and through latitudinal changes.

DATABASE

The data base program examines the Percent Occurrence of Richardson Number greater than -4.0 and less than or equal to 1.0. This program could also process any of the other parameters. Our analysis is done seasonally for years 1971 through 1975 and kilometer altitude levels 2-25. Calculations are performed by two methods, a five year average for each kilometer level and an average over the total 2-25 kilometer range per year. Values calculated are the mean, standard deviation, maximum, minimum and range. We also perform a two-way analysis of varirance. Results include the overall mean, variance component due to altitude bins, variance components due to yearly variations, total variance component, residual error and standard error. These are done for the entire 2-25 km range and also for altitude groups 2-7, 8-13, 14-19 and 20-25.

Output consists of for each season, the 24 levels of yearly means and the 24 levels of the 5 year means on permanent file storage.

BASEPLOTX

This plot uses the yearly data base output plus the 5 year average output. It created plots of Percent Occurrence of Turbulence (RI LE.1) versus altitude. A set of plots is created for each season with each of the 5 years and the 5 year average plotted per set. The sets are small and concise, ideal for comparing yearly and seasonal variations. Both pen and microfiche plots are available.

LOWALT

The Lowalt program selects, through card input, the station and year desired. The program searches the OS tape and extracts the raw data. The extracted records consist of a leader containing deck number, station number, year, month, day, hour and number of levels recorded. Each level contains the following parameters - pressure, altitude, temperature, relative humidity, wind speed and direction and level number. The data has an altitude range limit of 35 kilometers. Output is a 7 track CC tape containing the information. Approximately 15 years data recorded per tape.

ORIGINAL OS TAPES

The original OS tapes are a set of 40 9-track binary tapes. They contain Rawinsonde data from 144 stations throughout the Western Hemisphere. Testing occurred from 1970 to 1976.

ALTFIND

Altfind is a program created to examine the altitude structure of the Season E input. We take the validated input points from each launch and segment it into altitude levels of 2-7.5, 7.5-13.0, 13.0-20.0 and 20.0-25.0 km. We find the number of points and the average distance between them for each range. We then take a seasonal average of these values.

SEASON E

Season E reads the Lowalt generated 7 track CC tape. The data for each launch (2 per day) is verified and processed at each whole kilometer level by a spline interpolation. The desired turbulence parameters are calculated and averaged by season. The variables of major interest are the Richardson number and the percent occurrences of Richardson number less than .25 and less than 1.0. Output is a permanent file containing the 38 variables at each kilometer level, 1 to 35, for all four seasons and years 1970-1976. Another output option presents the daily raw input values and the daily calculated turbulence parameters.

BASEPLOT2PL

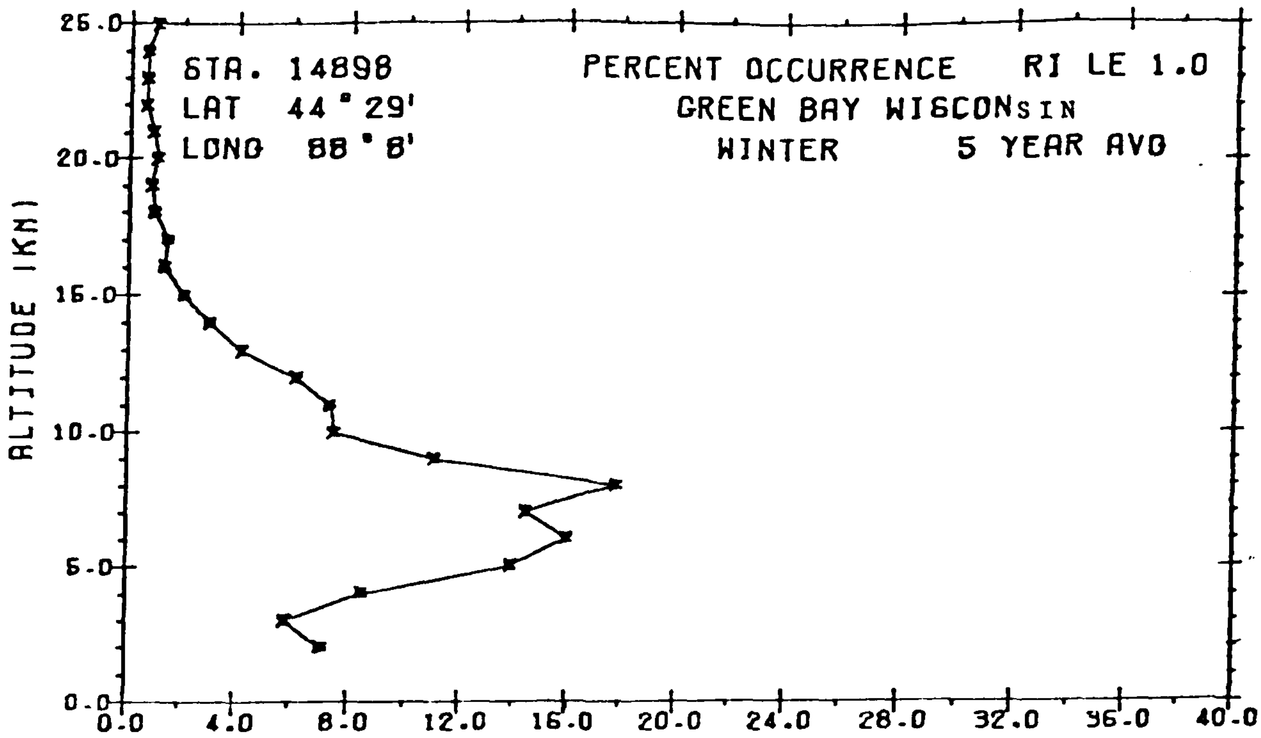
This plot uses the 5 year average data base output. Two plots for each season are created. Plot one shows the 5 year average per altitude data. The mean is plotted and the standard deviation, maximum and minimum are shown on the side. The second plot shows the yearly means taken over the total altitude range. The plot is good to examine yearly variations. The yearly standard deviation, maximum and minimum, are also shown.

The main parameter plotted has been the Percent Occurrence of Turbulence although many of the other variables have been examined through these plots. Output pen plots or microfiche.

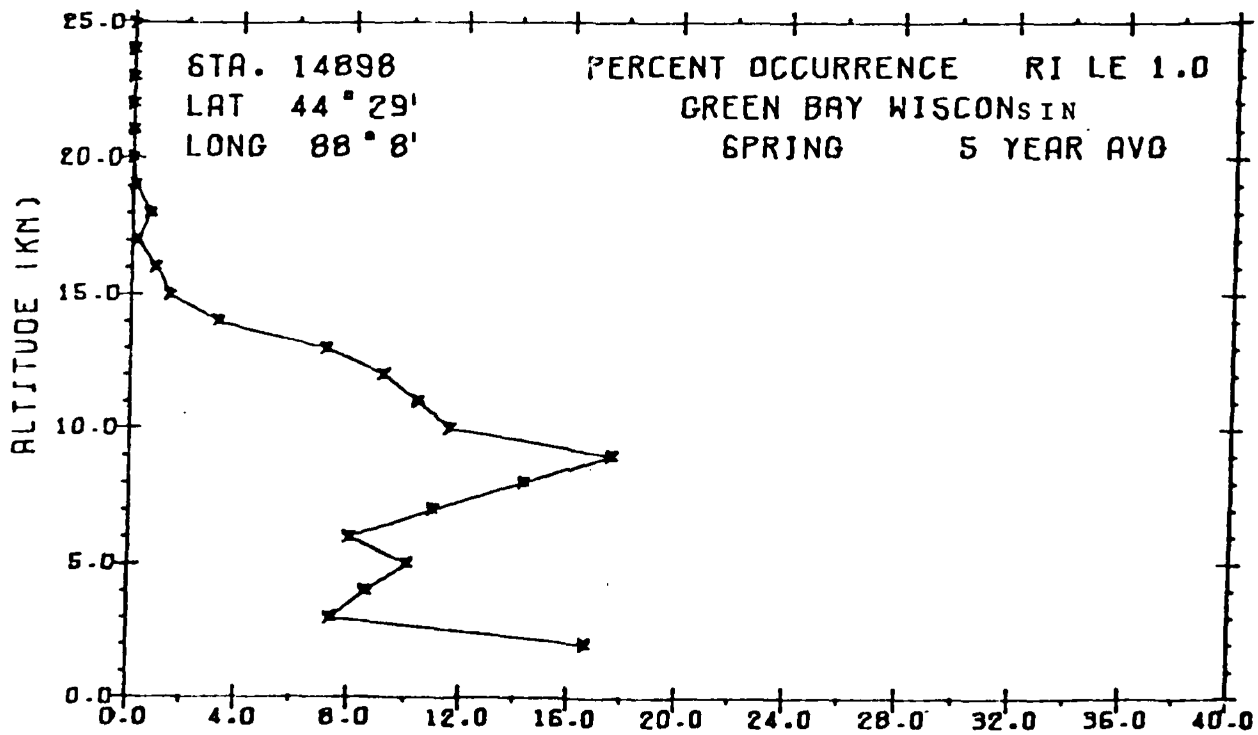
REGRESSION PACKAGE

The Regression Package is an extremely important element in the analysis. We use the data base output as input, and through this package perform a wide variety of statistical analysis. We examine the percent occurrence with parameters altitude, $altitude^2$, $altitude^3$, latitude, $latitude^2$, longitude, $longitude^2$, the four seasons and the five years. With this package we can examine the effects of each of these parameters on our model. Output includes means, deviation, correlation coefficients, corrected sums of squares and cross products, the sum of the residuals squared, multiple correlation coefficient, standard error of estimates, degrees of freedom, F-values and more. Scatter plots are also created.

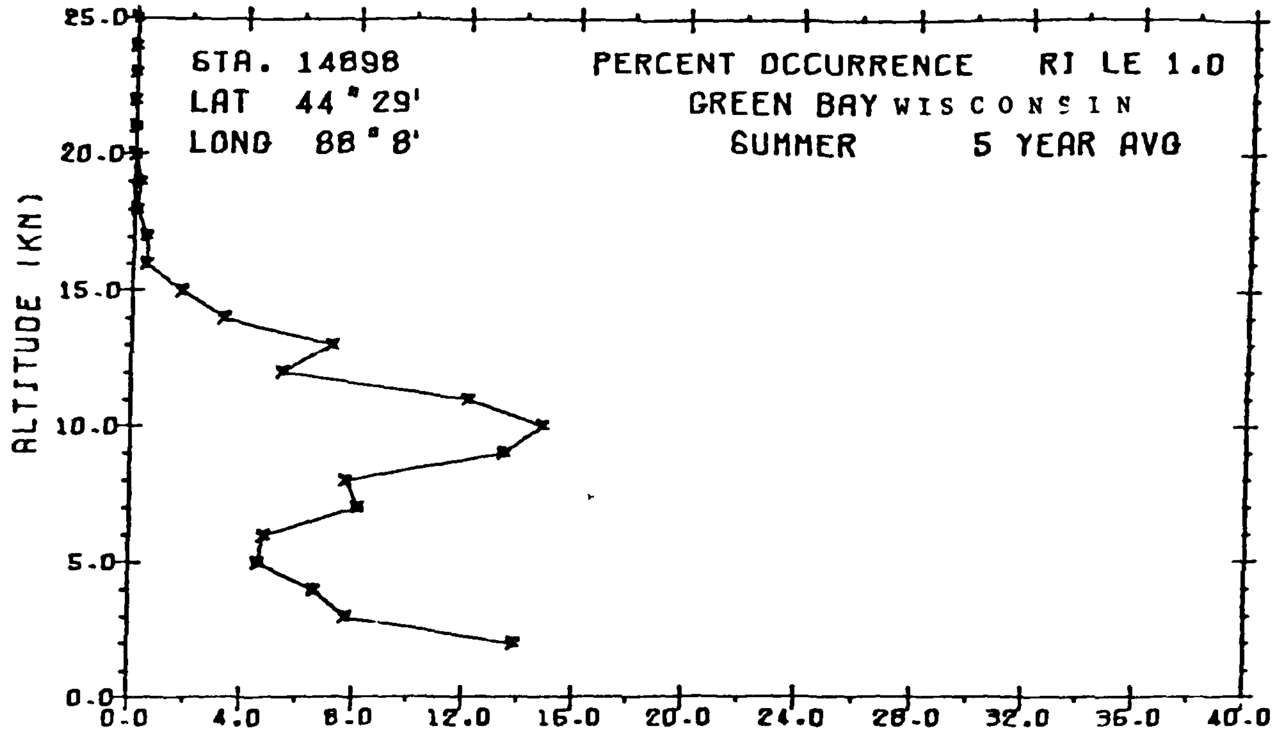
Typical percent occurrence plots follow.

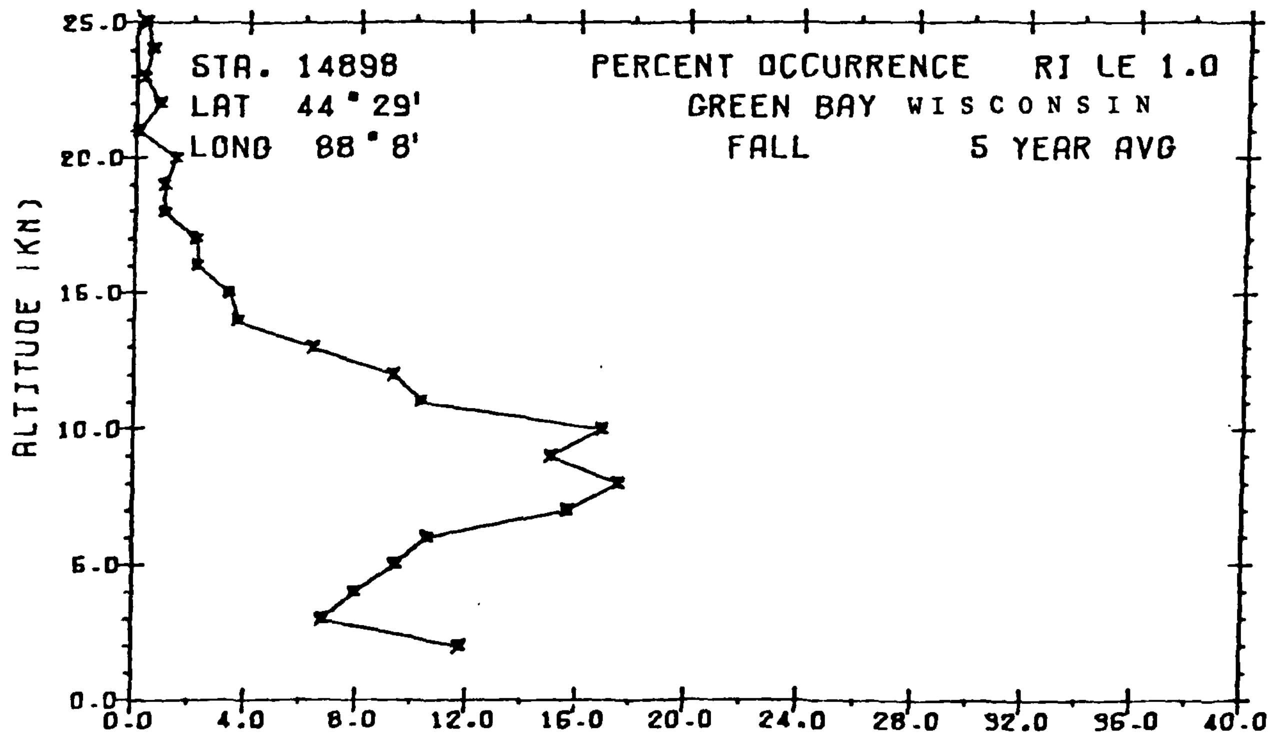


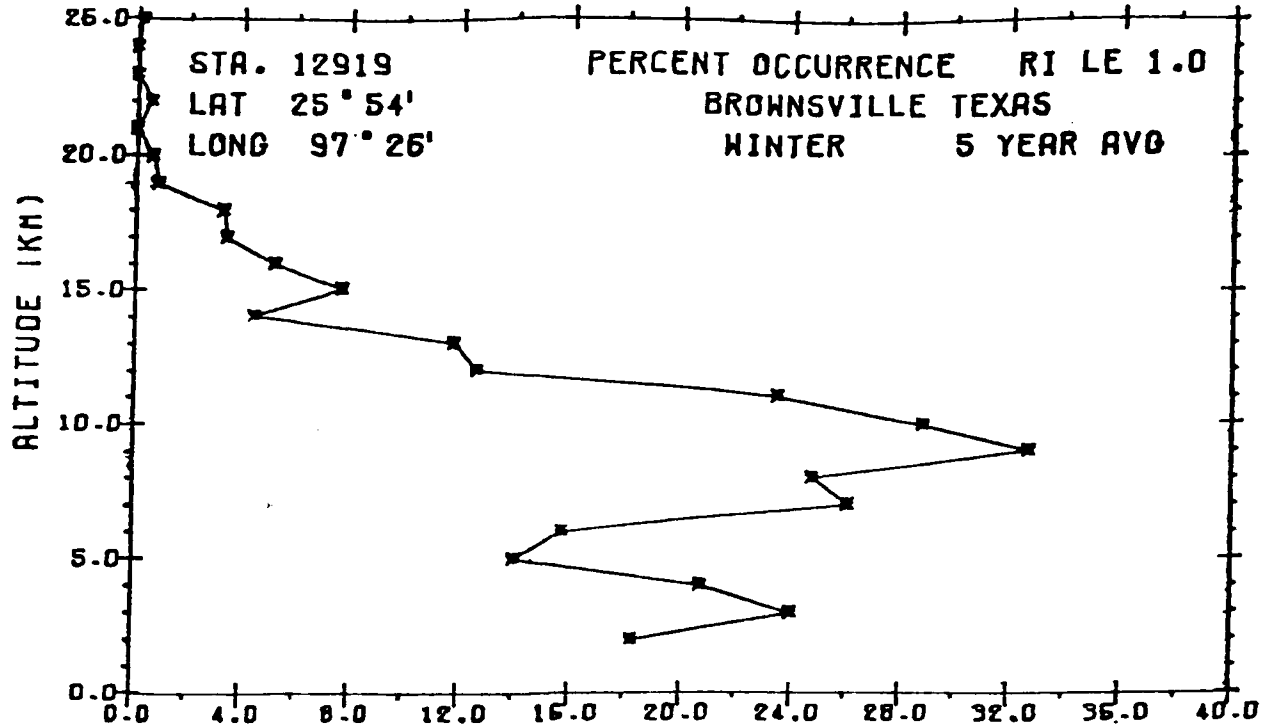
STA. 14898
LAT 44 ° 29'
LONG 88 ° 8'
PERCENT OCCURRENCE RI LE 1.0
GREEN BAY WISCONSIN
WINTER 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

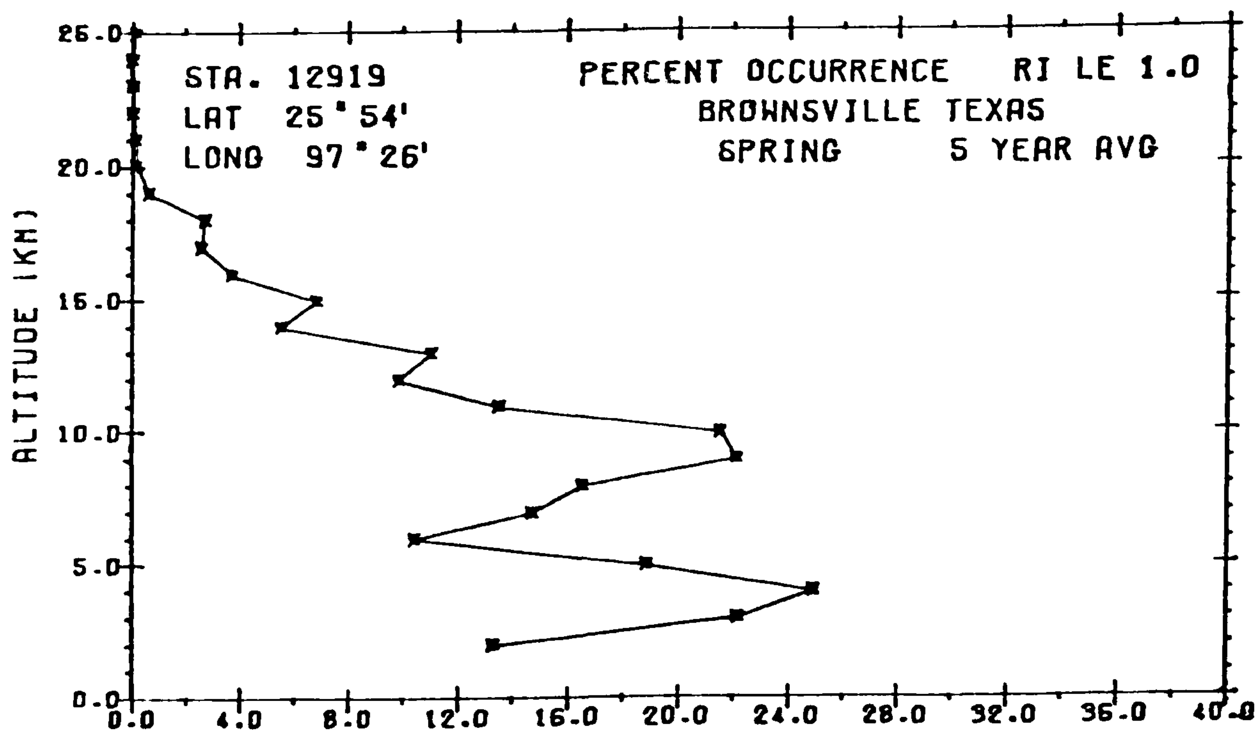


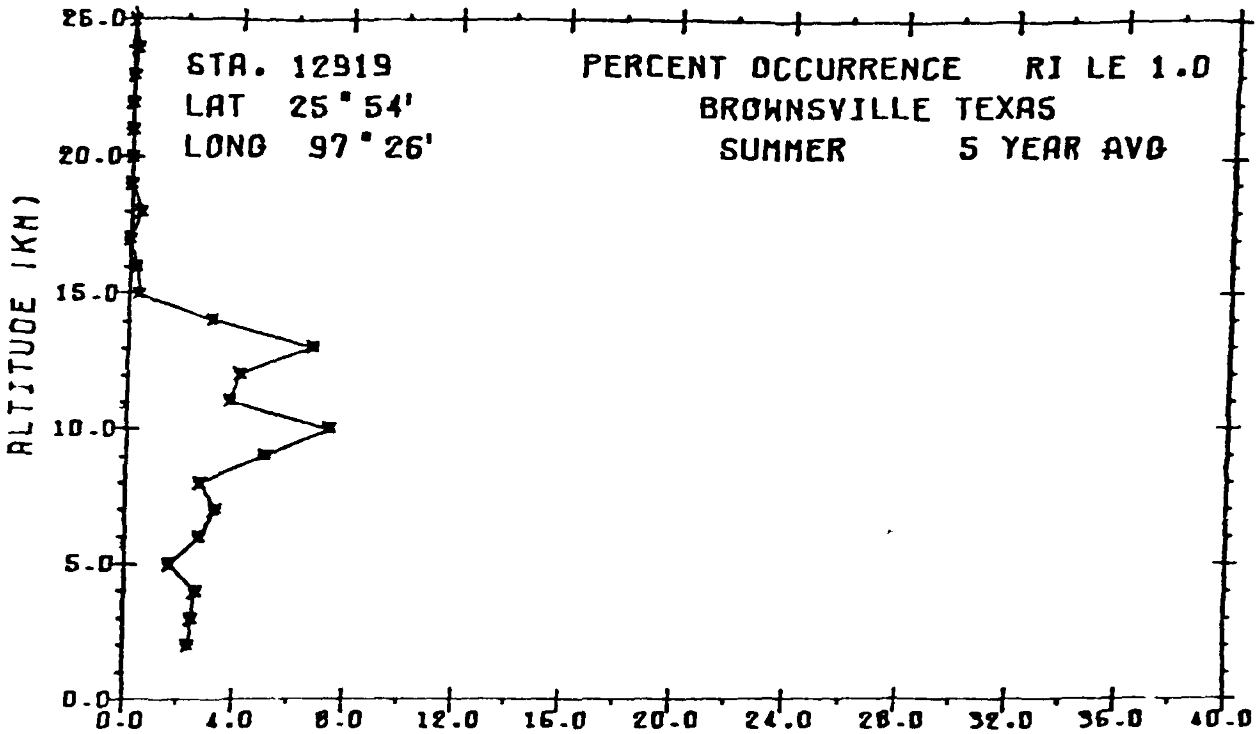
STA. 14898
LAT 44 ° 29'
LONG 88 ° 8'
PERCENT OCCURRENCE RI LE 1.0
GREEN BAY WISCONSIN
SPRING 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

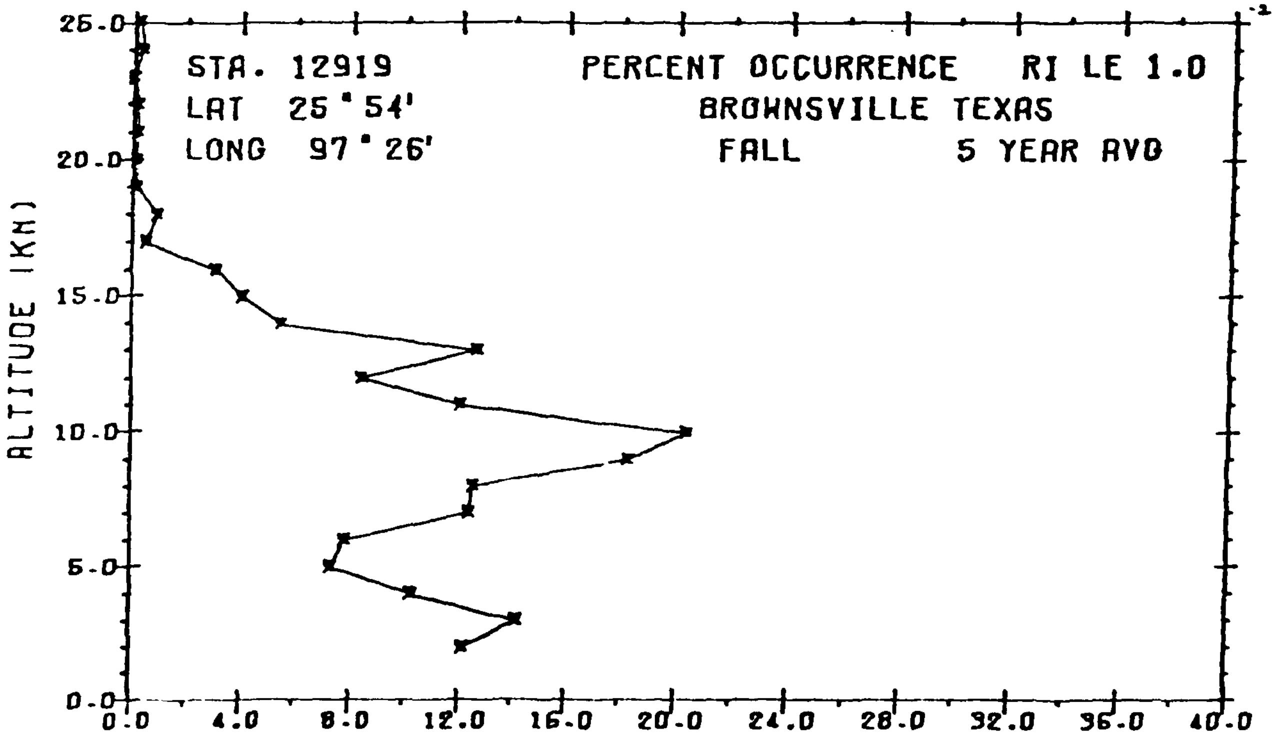

STA. 14898
LAT 44° 29'
LONG 88° 8'
PERCENT OCCURRENCE RI LE 1.0
GREEN BAY WISCONSIN
SUMMER 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

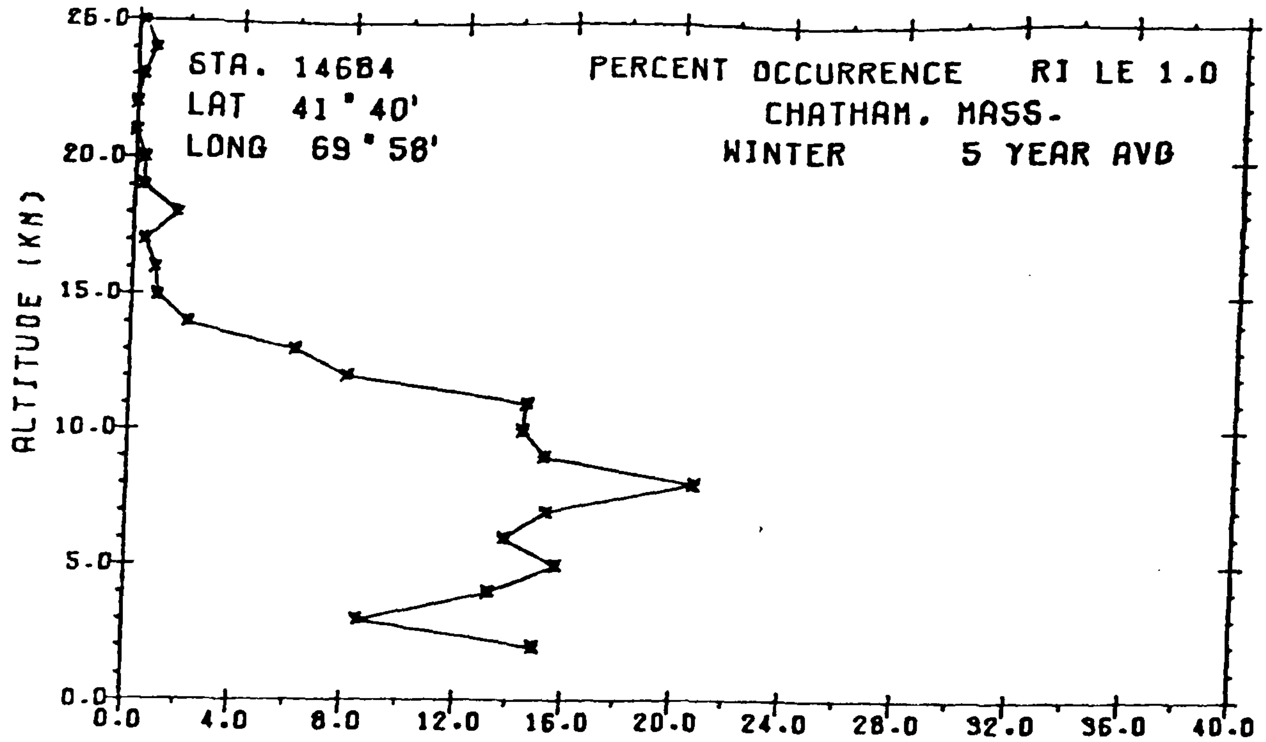

STA. 14898
LAT 44° 29'
LONG 88° 8'
PERCENT OCCURRENCE RI LE 1.0
GREEN BAY WISCONSIN
FALL 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

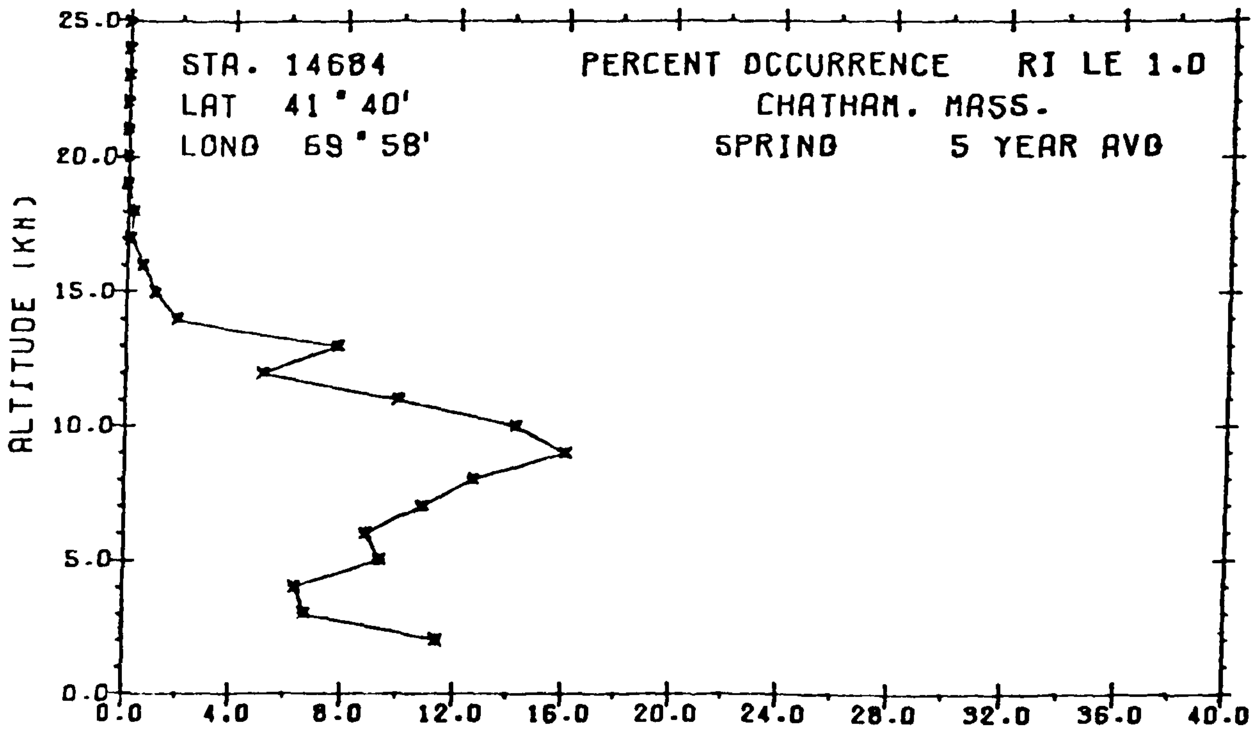

STA. 12919
LAT 25° 54'
LONG 97° 26'
PERCENT OCCURRENCE RI LE 1.0
BROWNSVILLE TEXAS
WINTER 5 YEAR AVG
ALTITUDE (KM)
0.0
5.0
10.0
15.0
20.0
25.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

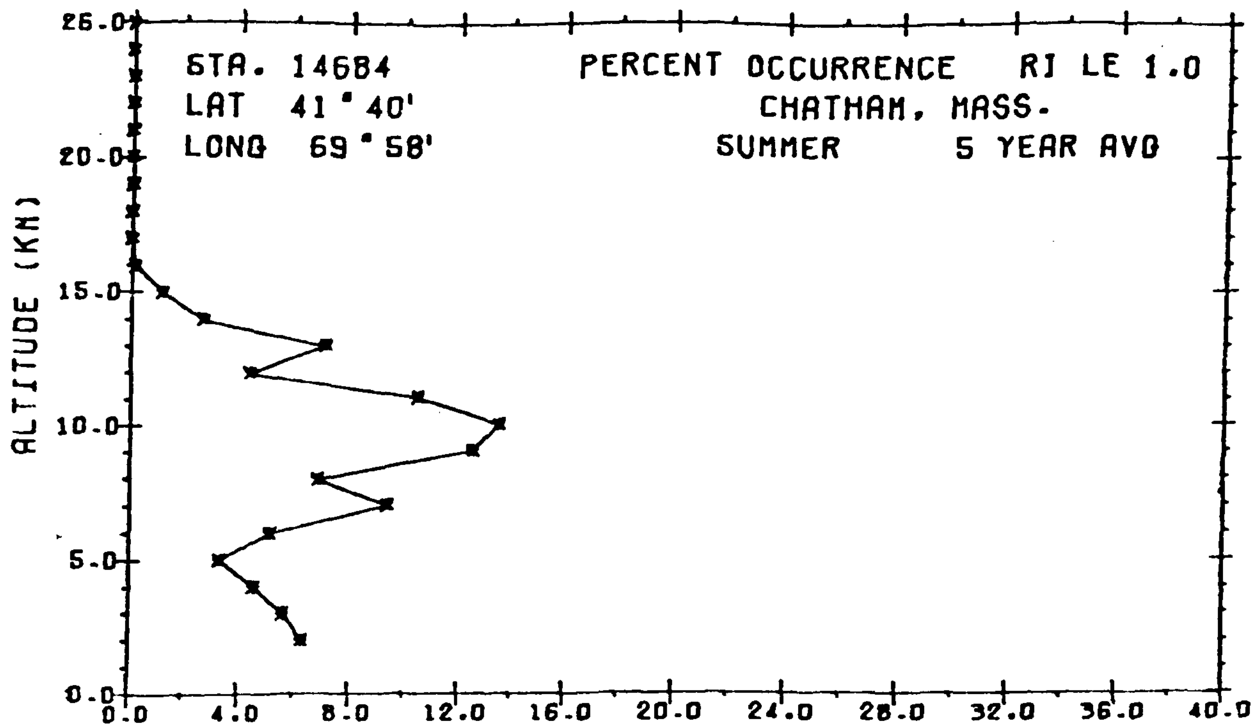

STA. 12919
LAT 25° 54'
LONG 97° 26'
PERCENT OCCURRENCE RI LE 1.0
BROWNSVILLE TEXAS
SPRING 5 YEAR AVG
ALTITUDE (KM)
0.0
5.0
10.0
15.0
20.0
25.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

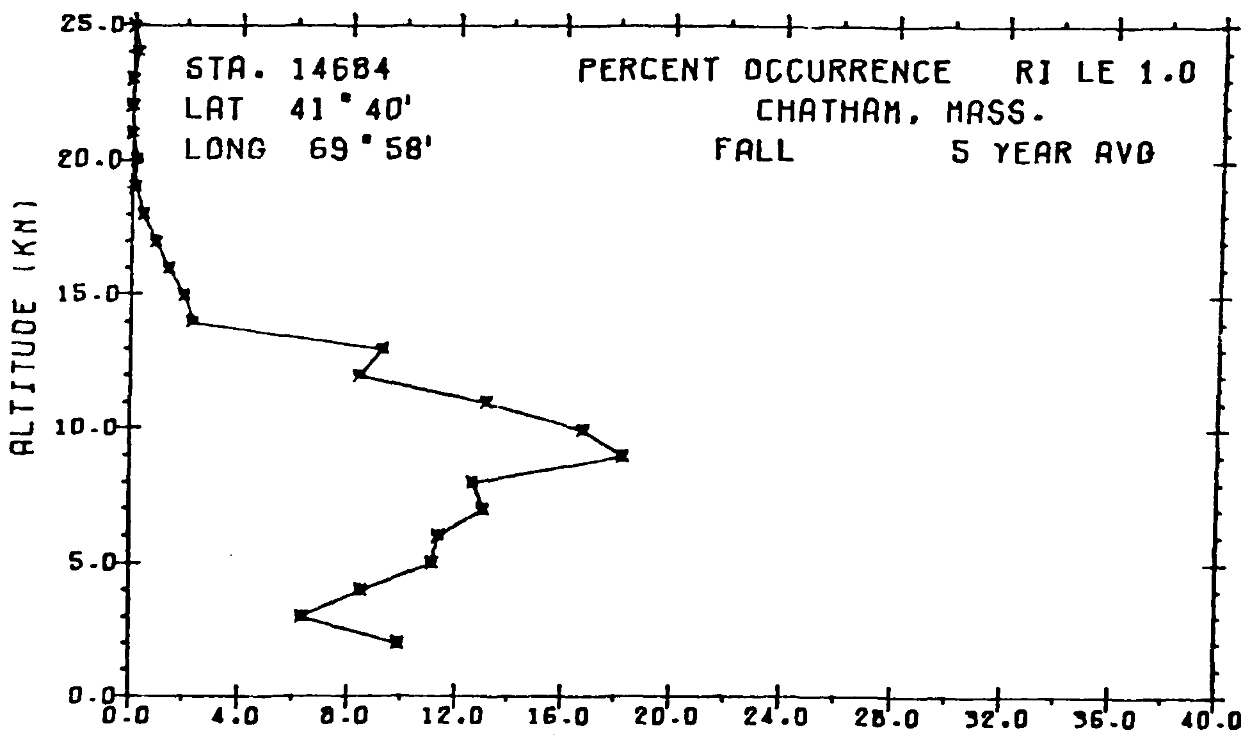

STA. 12919
LAT 25° 54'
LONG 97° 26'
PERCENT OCCURRENCE RI LE 1.0
BROWNSVILLE TEXAS
SUMMER 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0 4.0 8.0 12.0 16.0 20.0 24.0 28.0 32.0 36.0 40.0

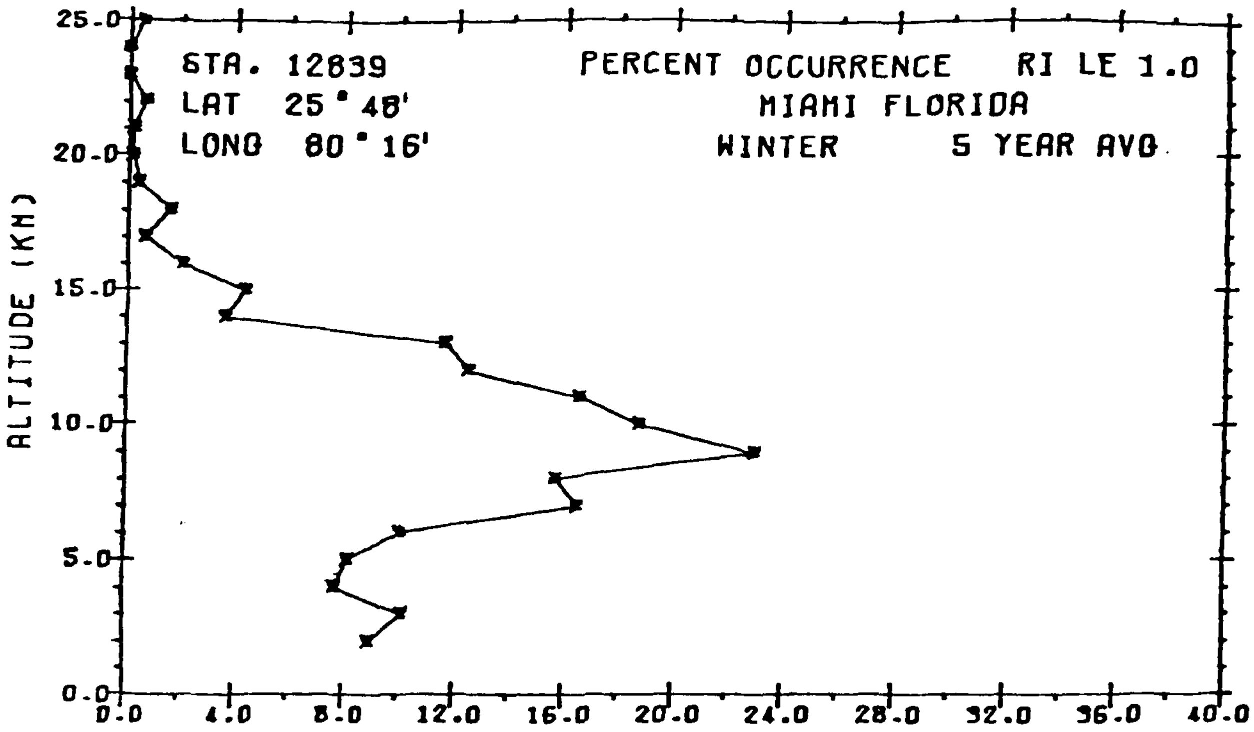

STA. 12919
LAT 25° 54'
LONG 97° 26'
PERCENT OCCURRENCE RI LE 1.0
BROWNSVILLE TEXAS
FALL 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0 4.0 8.0 12.0 16.0 20.0 24.0 28.0 32.0 36.0 40.0


STA. 14684
LAT 41 ° 40'
LONG 69 ° 58'
PERCENT OCCURRENCE RI LE 1.0
CHATHAM, MASS.
WINTER 5 YEAR AVG
ALTITUDE (KM)
0.0
5.0
10.0
15.0
20.0
25.0
0.0 4.0 8.0 12.0 16.0 20.0 24.0 28.0 32.0 36.0 40.0

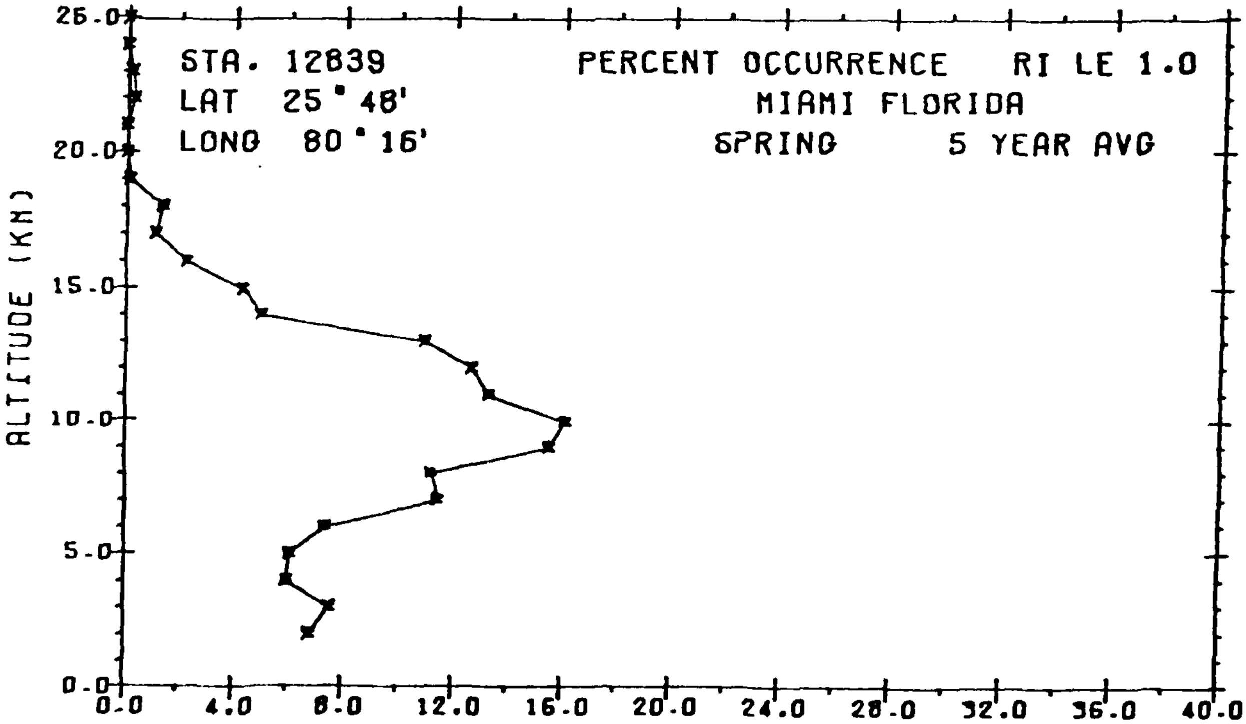

STA. 14684
LAT 41 ° 40'
LONG 69 ° 58'
PERCENT OCCURRENCE RI LE 1.0
CHATHAM, MASS.
SPRING 5 YEAR AVG
ALTITUDE (KM)
0.0
5.0
10.0
15.0
20.0
25.0
0.0 4.0 8.0 12.0 16.0 20.0 24.0 28.0 32.0 36.0 40.0


STA. 14684
LAT 41 ° 40'
LONG 69 ° 58'
PERCENT OCCURRENCE RI LE 1.0
CHATHAM, MASS.
SUMMER 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

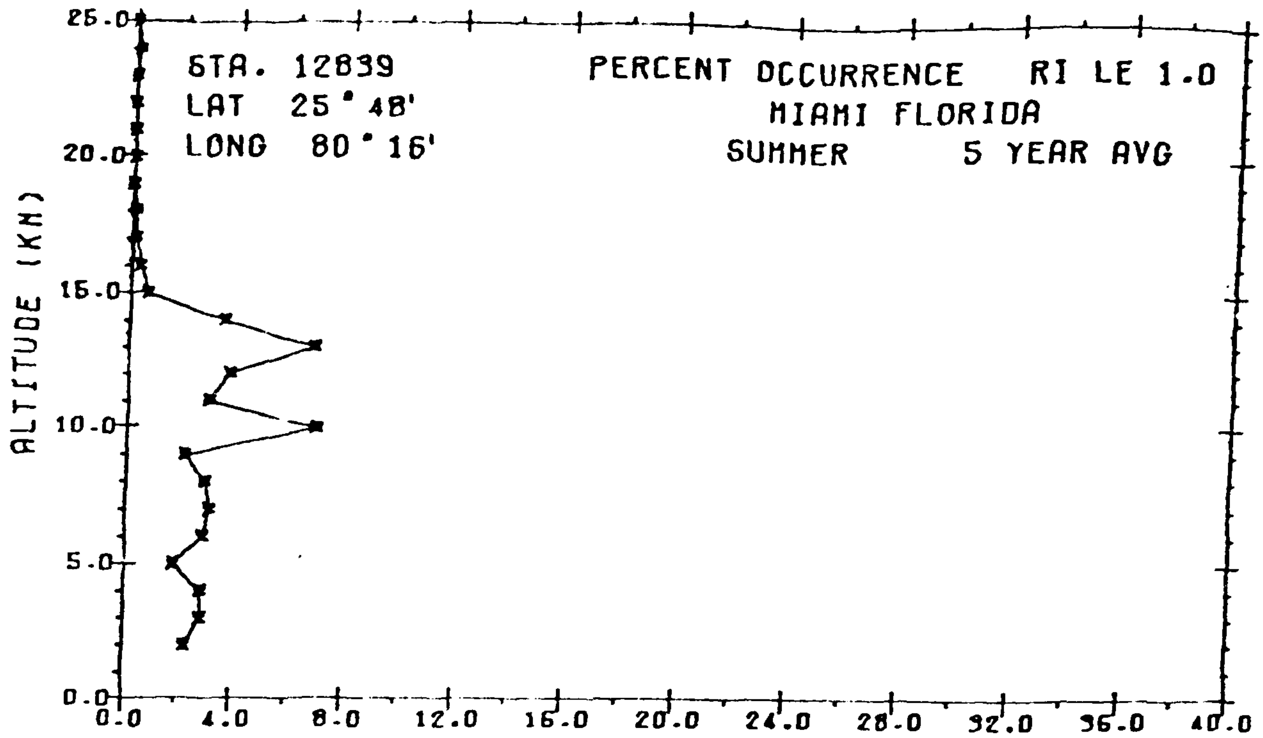

STA. 14684
LAT 41 ° 40'
LONG 69 ° 58'
PERCENT OCCURRENCE RI LE 1.0
CHATHAM, MASS.
FALL 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

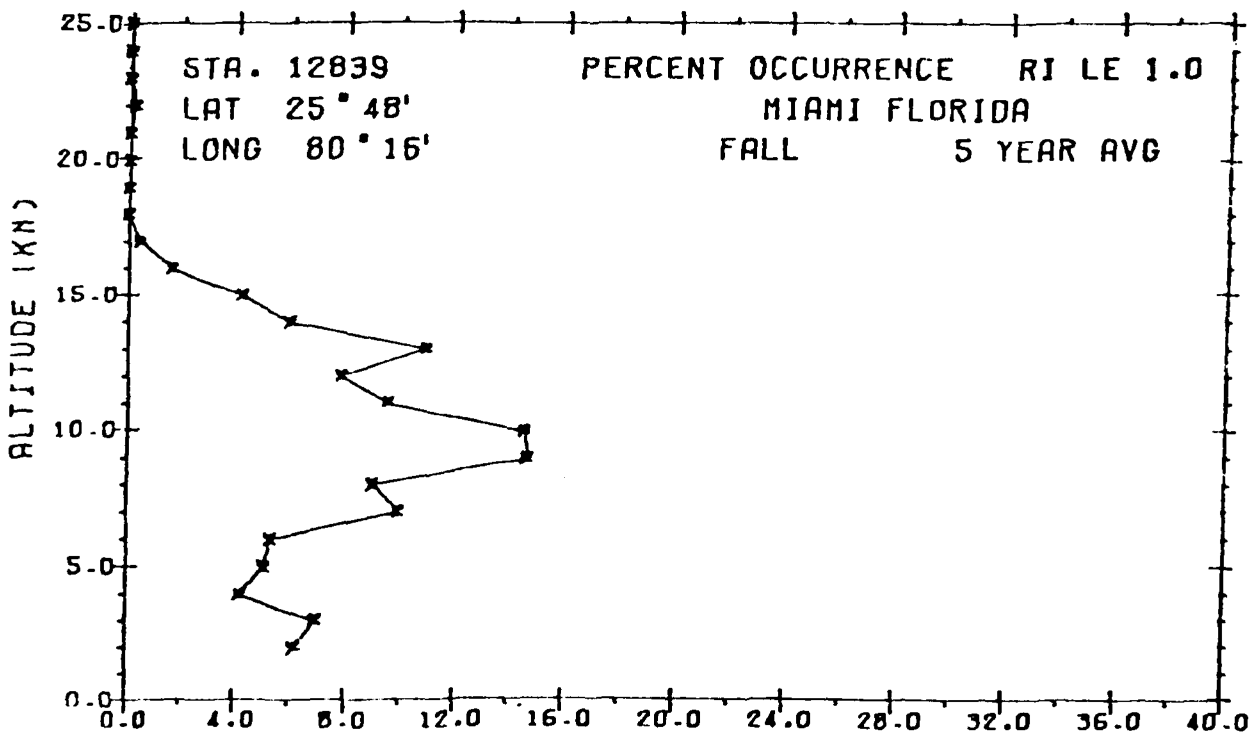


STA. 12839
LAT 25° 48'
LONG 80° 16'
PERCENT OCCURRENCE RI LE 1.0
MIAMI FLORIDA
WINTER 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0



STA. 12839
LAT 25° 48'
LONG 80° 16'
PERCENT OCCURRENCE RI LE 1.0
MIAMI FLORIDA
SPRING 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0


STA. 12839
LAT 25° 48'
LONG 80° 16'
PERCENT OCCURRENCE RI LE 1.0
MIAMI FLORIDA
SUMMER 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0


STA. 12839
LAT 25° 48'
LONG 80° 16'
PERCENT OCCURRENCE RI LE 1.0
MIAMI FLORIDA
FALL 5 YEAR AVG
ALTITUDE (KM)
25.0
20.0
15.0
10.0
5.0
0.0
0.0
4.0
8.0
12.0
16.0
20.0
24.0
28.0
32.0
36.0
40.0

TURBULENCE STUDY - STATISTICAL ANALYSIS

OBJECTIVES:

The statistical analysis plan to be described below has two major objectives:

1) To relate occurrence of turbulence to latitude, longitude, altitude and seasonal variation.

2) To identify areas of high turbulence.

I - SITE SELECTION

At present data are available on 144 sites (locations) for the years 1971-1976. From these sites we have selected a subset of 14 sites for initial analysis. This subset has been selected to include:

1) A wide range of latitude and longitude. The 14 sets are representative of sites ranging from 70° to 120° longitude (W) and 25° to 50° latitude (N). This will give the results broad applicability.

2) "Possibly" interesting variables such as land configurations (i.e. coastal, mountains, flat lands). This will allow us to investigate variables that may introduce variability into our results.

3) Stations which are geographically close. This will allow us to investigate the correlation between nearby sites.

4) Stations that also have data from 1948 to 1970. This will permit a more extensive time series analysis (if deemed useful).

The results obtained from the 14 selected sites should be representative of locations in the continental United States. When the analysis is completed on these sites data from other sites will be analyzed to act as a validation of the results obtained from the 14 original sites.

II - DATA FOR ANALYSIS

Dependent Variable - The major dependent variable is the proportion of days for a given time period (e.g., a season) with turbulence. Turbulence is determined according to whether the Richardson number is less than .25. At each site we have for a given balloon launch at most one Richardson value for each 1.0 k interval from ground level up to a height of 35km.

A typical situation is to have approximately two balloon launches per day providing two readings for up to 25 km. At this point the balloon usually bursts and measurements are no longer available.

Independent Variable - The independent variables include: site latitude, site longitude, altitude, average temperature, north-south wind speed, east-west wind speed, temperature variation (e.g., range of temperature over a season), density humidity, year, and season (i.e., appropriate time period).

III - ANALYSIS OF INDIVIDUAL SITES (FIRST ANALYSIS)

The first analysis is a site-by-site analysis. That is, each site is analyzed individually.

A - Summary Statistics

For a given time period (season) summary statistics on all variables will be computed. These include means, standard deviations, maximums/minimums and ranges for each altitude bin (each 1 kilometer is referred to as a bin).

In particular, we determined the number of altitude bins that can be analyzed with good statistical accuracy and precision. High altitude bins (say above 25 km) often did not have sufficient data for a good analysis. Based on the results we obtained from the 14 selected sites, we determined how many bins should be carried further in our analysis.

B - Altitude Analysis

For each altitude bin and time period (season) a two way analysis of variance will be performed using Time Period by year as the variables for classification. The objectives will be: (1) quantify yearly variation and (2) quantify time period variation.

C - Altitude by Time Period by Years Analysis

This analysis quantified variations due to altitude, variations due to years, variations due to time period and interactions among these three sources.

D - Regression Analysis

The regression model related proportion of turbulence to year, time period, altitude and weather variables such as temperature, density, humidity, wind direction, and interaction of these variables. Specifically the independent variables are: Year, Time Period (season or month), Altitude, Average Temperature, Temperature Variation, Wind Direction (North-South, East-West), Density, Humidity and interaction of these.

IV - COMBINATION OF SITES (SECOND ANALYSIS)

The second analysis consists of combining sites and analyzing them jointly. If deemed necessary the following analysis can be performed on subsets of sites.

A - Regression Analysis (altitude x latitude x longitude)

A regression relating proportion of turbulence to geographic location and altitude will be performed. This regression will also include year and time period (season). The model is

$$\text{(Proportion turbulence)} = b_o + b_1 \text{ (Altitude)} + b_2 \text{ (latitude)} + b_3 \text{ (longitude)} + b_4 \text{ (year)} + b_5 \text{ (time period)}.$$

If necessary the model will be expended to include interactions of the above variables.

B - Expanded Regression Analysis

The next analysis will further expand on the above regressions. Variables which quantify land configurations will be added. Also weather variables as discussed in III D will also be added.

V - LOGISTIC REGRESSION

The objective of analysis is to relate occurrence of turbulence to sets of variables. In particular we want to relate variables to the probability of turbulence. A useful model for this is

$$P = \frac{1}{1 + e^{\varepsilon\, b_i Xi + b_o}}$$

where P represents the probability of turbulence and $\Sigma\, b_i Xi + b_o$ represents a linear function relating variables $X_1, \ldots, X_n$ to the probability of P.

By use of the transformation

$$\ln \left(\frac{1-P}{P} \right)$$

we have

$$\ln \left(\frac{1-P}{P} \right) = b_o + \Sigma\, b_i X_i$$

and regression analysis is now possible using the transformed dependent variable. We will perform these logistic regressions in addition to the above regressions.

A - Individual Sites

A logistic regression will be performed using variables listed in III D.

B - Combination of Sites

Logistic regressions will be performed using variables described in IV A and also IV B.

Note: Analysis IV A, IV B, V B will allow us to relate the occurrence of turbulence to latitude, longitude, altitude and seasonal variation. Additionally these analyses will quantify importance of weather variables to turbulence. The functions developed in these analyses, especially the results of V B, will permit us to identify the profiles of variables that relate to higher turbulence - viz, we will be able to identify values of $X_1, \ldots, X_n$ given in V that produce high probabilities P given by the logistic regression.

Note: While the above analyses address mainly the problem of explaining past data, we will also investigate their ability to predict turbulence.

Note: As stated earlier the above will first be performed on the 14 selected sites. When this is completed other sites in the continental U.S. will be analyzed for validation. Then high latitude sites will be selected for an analysis similar to the one described above.

Rocketsonde Measurements - Radiosonde observations, or "RAOBS" as they are termed, are made to determine the pressure, temperature, and humidity from the surface to the point where the balloon bursts. The radiosonde consists of meteorological measuring elements coupled to a radio transmitter and assembled into a small lightweight box. The device is carried aloft by a balloon filled with hydrogren or helium gas. Included in the train is a small parachute to slow the descent of the instrument after the balloon bursts, thereby minimizing the danger of injury to life and property. As the balloon rises, measurements of pressure, temperature, and relative humidity are transmitted to a gound station where the data are recorded automatically. An observer then transcribes the information into a more commonly used form and plots it on various charts. Measurements of pressure are made in millibars, temperature in degrees Celsius, and moisture in percent of relative humidity. The Celsius temperature scale is the same as the centigrade scale and replaces that scale in all meteorlogical observations.

As the balloon rises, it is followed either visually by an observer at a theodolite, or electronically by radio-direction-finding equipment that tracks the transmitted signal. The balloon's drift away from the release point is plotted, and from this the direction and speed of the air movements are determined. The winds-aloft observation is termed a "RABAL" when the tracking is done visually, and a "RAWIN" when the tracking is done electronically. A combined rawin and raob is termed a "RAWINSONDE".

Rawinsonde System - The basic rawinsonde system consists of a balloon-borne radiosonde, a receiver and tracking unit, and a recorder. The most commonly used radiosonde transmits meteorlogical information consisting of pressure obtained from an aneroid cell, temperature, and relative humidity. The telemetered meteorological information from the tadiosonde is detected, amplified, and shaped by a receiver and the processed information is printed in graphic form on a strip chart recorder. An antenna array, which is an integral part of the receiving equipment, is used to track the radiosonde. Tracking is accomplished either manually or automatically. The elevation and azimuth angle information so obtained, in conjunction with height information computed from the meteorological data, is used to compute winds aloft.

Radar Wind - A target, suitably mathced to the frequency of the radar in use, is attached to the train of the radiosonde baloon. The radar will provide the slant range to the target by measurement of the time interval between the transmitted and reflected radar pulse. The azimuth indicator of the radar set will indicate the azimuth angle of the target. Slant range is used in conjunction with height data, computed from the radiosondes' meteorological data, to compute distance from observation point.

Radiosonde Instruments - The radiosonde is a baloon-borne, battery-powered instrument used together with the ground receiving equipment to delineate the vertical profile of the atmosphere. Pressure is measured by means of a baroswitch which employs an expanding aneroid pressure cell to move a contact arm across a commutator bar as the pressure decreases. Temperature is measured by a thermistor, the electrical resistance of the thermistor being a function of temperature. Relative humidity is measured by a hygristor, the electrical resistance of the hygristor being a function of relative humidity and, to some extent, temperature. As the radiosonde ascends, the thermistor and hygristor are switched sequentially into the modulator circuit by the baroswitch. The frequency of the received signal, therefore, is alternately a function of temperature and humidity, and may be any value from 10 to 200 Hz. In addition, fixed resistors are connected periodically into the modulator circuit by the baroswitch so that the frequency transmitted, when these are connected will show any change that may be occurring in the modulator characteristic due to variations of circuit parameters. Radiosondes are provided with the thermistor mounted externally on an outrigger which exposes it at some distance from the radiosonde case.

The hygristor is enclosed within an airduct to shield the element from direct exposure to precipitation. The radiosonde contains an aneroid baroswitch for measuring pressure, a white-coated thermistor for temperature, and a carbon hygristor for relative humidity. Radiosondes of this type will be considered the basic radiosonde and will usually be identified by a serial number containing no suffix.

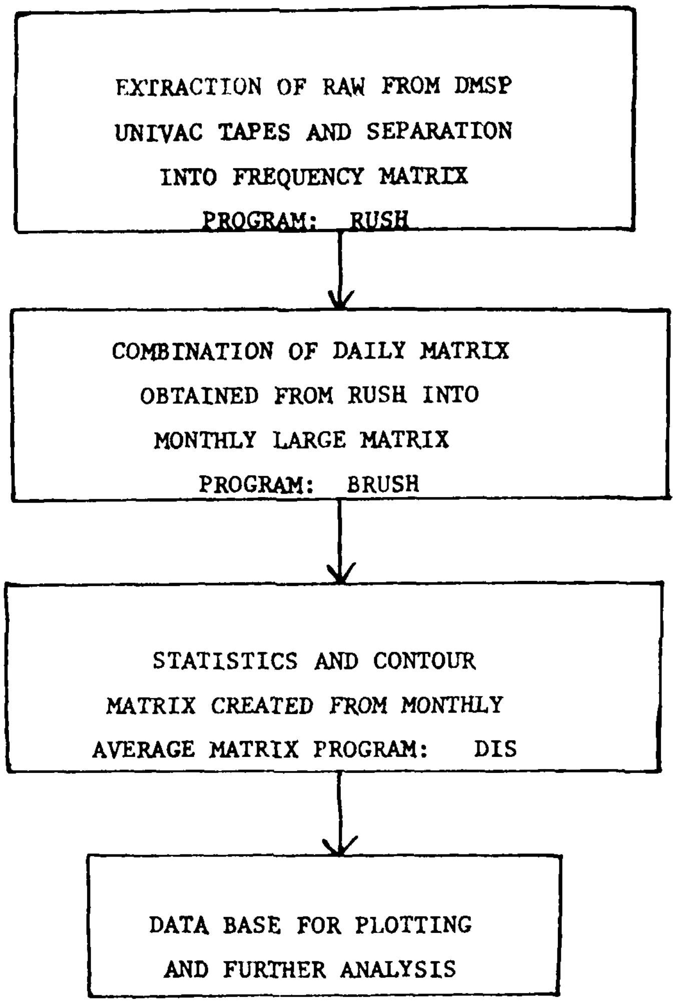
Defense Meteorological Satellite is equipped with a swept-frequency HF noise receiver. The receiver provides measurements of radio noise of terrestial origin every 100 KHz in the frequency range of 1.2 to 13.9 Mhz. Satellite is basically in a sun-synchronous morning or evening orbit. By analyzing data from successive morning or evening orbit, global maps of the noise intensity can be obtained at the satellite altitude. Work is being done in order to obtain world maps of noise intensity so as to assess the morphological behavior of the noise at satellite altitude as a function of frequency. Our studies have been concentrated on the frequencies of 1.5 to 13.5 Mhz in steps of 0.5 Mhz at both the dawn and dusk orbits. Monthly average noise maps are obtained in order to provide more representative maps of the average noise behavior.

The Air Force Global Weather Central processes several types of special sensor data including the data from the swept-frequency HF noise receiver. The DMSP tapes have been processed by GWC on the UNIVAC 1110 system, where the word length is 36 bits.

We next present a description of the overall software used in this study. Following this is a report deriving the optimal procedure for estimating the FoF2 of the Ionosphere.

IONOSPHERIC RADIO NOISE MEASUREMENT

SOFTWARE



SOFTWARE DESCRIPTION

PROGRAM: RUSH

RUSH unpacks data from one DMSP satellite tape. UNIVAC record is read and put into CDC words. All frequencies between 1.5 and 13.5 Mhz in steps of 0.5 Mhz are translated and loaded into frequency, latitude, and longitude matrix.

PROGRAM: BRUSH

Intermediate step before statistical analysis. Data files created by RUSH are combined until monthly matrix is complete. Up to five data sets will be combined with an existing data file.

PROGRAM: DIS

Program displays average count matrix for dawn and dusk and statistics about these matrices. Also printer contour is done to provide high and low noise peak for each frequency, at each latitude and longitude.

Data Base generated now used for further analysis such as estimates foF2 as described in this report.

I INTRODUCTION

This report describes an estimation procedure for determining the critical frequency of the ionospheric F2 layer. The estimate is obtained from measurements of radio noise obtained by a satellite orbiting above the F2 layer. A spectrum analyzer obtains noise power as a function of frequency in a band wide enough to include all possible values of the critical frequency.

The power measured at a particular frequency is a function of the radio energy being transmitted, and the transmittance of the F2 layer. The ionospheric parameters depend on the position of the satellite (particularly latitude), local time, and solar activity. The noise measured by the satellite receiver is considered to consist of two parts. The first is referred to as terrestrial noise, which passes through the F2 layer. This noise originates from such sources as radio transmitters located near the subsatellite point. The other component is background noise, which is not affected by the F2 layer, and can originate from within the satellite instrumentation itself or other sources located above the F2 layer.

The available data comes from a DMSP satellite with an on-board swept-frequency receiver. Counts are obtained for 128 frequency channels from 1.2 MHz to 13.9 MHz in increments of 100 KHz. One important aspect of the data is that there is coupling between adjacent channels, that is, a fraction of the power at one particular channel is also picked up by adjacent channels. This is most likely caused by the bandwidth of each channel filter being wider than the interchannel separation. While this produces some smoothing of the data, it also modifies the probability distribution of the data from what one would normally expect. Thus, it must be dealt with before the model of the data is constructed.

For estimation of the critical frequency foF2, the maximum likelihood (ML) technique is the method of choice. It is a well known and studied estimation method. For most applications it is well behaved and the optimum method. It is also systematic and intuitively meaningful since it can be derived directly from the model of the data.

II THEORY OF ESTIMATION METHOD

A. The Model

The effect of the foF2 layer on the radio noise spectrum received by the satellite can be modelled as a high pass filter. There are many candidate filters which can be used. For our purposes, it is desirable to use a filter with the least number of parameters, since each parameter must either be estimated or assigned a value. A suitable filter is the Butterworth filter of order n.

$$\left| H_n (f) \right|^2 = \frac{1}{1 + \left(\frac{fo}{f}\right)^{2n}} \qquad (1)$$

The parameter fo thus becomes the ionospheric critical frequency. The filter order n must also be chosen. Our procedure will be to choose several values for n and observe its effect on the estimation procedure.

The satellite receiver measures the power at discrete frequencies. A record of count data can be expressed as

$$S(f_k) = S_N(f_k) + \left| H_n(f_k) \right|^2 S_T(f_k) \qquad (2)$$
$$\text{for } k = 1,\ldots,N$$

where $S_N(f_k)$ is the noise power due to background and instrumentation noise, and $S_T(f_k)$ is the noise power due to radio signals of terrestrial origin which are affected by the F2 layer. However, S_N and S_T are random processes whose moments are unknown functions of frequency. In the absence of such information about these signals, we need to make certain assumptions;

namely, that these moments are constant with respect to frequency. In actuality, we need only consider one of the moments (such as the mean or variance) since these signals are chi-square distributed with all moments dependent on one of the moments. Thus, we obtain one additional parameter for S_N and one for S_T which must be estimated, along with the important quantity fo.

The problem can thus be expressed as a classical parameter estimation problem. The procedure is as follows:

1. Preprocess the data into a form acceptable by the estimation method, described in part B.
2. Estimate the critical frequency fo (and the S_N and S_T parameters as a by-product) by the maximum likelihood method (ML), described in part C.

Also in this chapter, part D describes and derives the Cramer-Rao bound for the estimated fo. Part E extends the method to include an arbitrary signal noise power function. Part F discusses a method for obtaining initial estimates of the S_N and S_T parameters which can then be used in the estimation procedure of part C.

B. Preprocessing of the Data

In the absence of interchannel coupling, each frequency channel would be a chi-square distributed random variable with two degrees of freedom. However, since each filter in the bank of filters has a finite, but unknown, bandwidth larger than the separation between channels, the received record is then expressed as the original power spectrum convolved with some "smearing function". Consequently, the degrees of freedom of each channel is increased.

The estimation method requires that the distribution of the data be of known form. Furthermore, computational efficiency is gained if the channels are uncorrelated among themselves. The solution is to remove the effect of coupling by filtering the record in the frequency domain by a digital filter:

$$X(f_k) = \sum_{m=-M}^{M} S(f_{k-m})\, h_d(m) \qquad k=1,2,\ldots,N$$

$X(f_k)$ is a nearly-Gaussian, zero mean, uncorrelated (and therefore independent) process. The filter $h_d(m)$ is a nonrecursive zero-phase (symmetric) digital filter, designed by the McClellan-Parks algorithm [1], of length $2M+1 = 41$ with an attenuation of 40 dB in the stopband. The problem encountered at the edges of the record is dealt with by extending the record at each end by its mirror image

$$S\left(f_{1-n}\right) = S\left(f_{1+n}\right) \text{ and } S\left(f_{N+n}\right) = S\left(f_{N-n}\right)$$
$$\text{for } n = 1,\ldots,M \text{ and } N = 128.$$

This processed record can then be expressed as

$$X(f_k) = N(f_k) + \left| H_n(f_k) \right|^2 T(f_k) \tag{3}$$

where $N(f_k)$ and $T(f_k)$ are independent Gaussian variables, with zero mean, but a variance which depends on the signal strength S_N and T_N. The variance of $X(f_k)$ is

$$\sigma_x^2(f_k) = \sigma_N^2 + \left| H_n(f_k) \right|^4 \sigma_T^2 \tag{4}$$

C. Maximum Likelihood Estimation

Parameter estimation via the maximum likelihood technique can be expressed intuitively as follows. We have a model, with parameters that can be chosen. Given an observation $\underline{X}$, what choice of parameters make the given observation most likely to have happened? In the context of the above model, the parameter vector is

$$\underline{\theta} = [\, f_o, \sigma_N^2, \sigma_T^2 \,]$$

and the observation vector is the preprocessed record

$$\underline{X} = [X(f_1),\ X(f_2),\ \ldots,\ X(f_N)] = [x_1, x_2, \ldots, x_N].$$

The (ML) estimation procedure involves finding the probability density of $\underline{X}$ as a function of the parameter vector $\underline{\theta}$, $p(\underline{X};\underline{\theta})$ which is then maximized with respect to $\underline{\theta}$. In practice, the likelihood function $l(\underline{\theta})$ is maximized, where

$$l(\underline{\theta}) = \log p(\underline{X};\underline{\theta}).$$

After preprocessing, the channels are independent zero-mean Gaussian variables. The probability density can be expressed simply as

$$p\,(\underline{X};\ f_o,\ \sigma_N^2, \sigma_T^2) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi}\,\sigma_x(f_k)} \exp\left[-\frac{1}{2}\frac{x_k^2}{\sigma_x^2(f_k)}\right] \quad (5)$$

where σ_x^2 is obtained from (4).

The likelihood function is

$$l\,(f_o,\sigma_N^2,\sigma_T^2) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\sum_{k=1}^{N}\left[\log\sigma_x^2(f_k) + \frac{X_k^2}{\sigma_x^2(f_k)}\right] \tag{6}$$

The first term can be ignored since it does not depend on the parameters. The ML estimate is obtained from

$$\text{maximize } l(\underline{\Theta}) = -\sum_{k=1}^{N}\left[\log\sigma_x^2(f_k) + \frac{X_k^2}{\sigma_x^2(f_k)}\right] \tag{7}$$

$$\text{where } \sigma_x^2(f_k) = \sigma_N^2 + \left[\frac{1}{1+\left(\frac{f_o}{f_k}\right)^{2n}}\right]^2 + \sigma_T^2$$

We can try to locate the maximum by taking the partial derivatives with respect to f_o, σ_N^2, σ_T^2 and eauate to zero. However, this yields a complex set of coupled nonlinear equations, and the solution of these equations does not guarantee a global maximum.

Instead, the solution is found numerically. This is a three-dimensional maximization, and some gradient directed algorithm would seem to be called for. However, $l(\underline{\Theta})$ is itself a random process, making its gradient too unreliable for such a sensitive algorithm. Instead, the following iterative approach is used:

1. Obtain initial estimates for σ_N^2 and σ_T^2. These can be rough guesses, or they can be obtained by the method in part F.

2. Keeping σ_N^2 and σ_T^2 constant, maximize l by a one-dimensional Fibonacci search on f_o.

3. Keeping σ_N^2 and f_o constant, do Fibonacci search on σ_T^2.
4. Keeping σ_T^2 and f_o constant, search for σ_N^2.
5. Go back to step 2 and repeat until convergence is achieved.

This method has been shown to be reliable with the given data, and converges in about three or four iterations, when initial estimates of σ_N^2 and σ_T^2 are provided by the method of part F.

D. Cramer-Rao Bound

Since the estimated f_o is a random variable, it would be useful to determine its mean and variance, and hence determine the accuracy of the estimator. For most nonlinear estimation problems, this is a difficult task. However, the theory does provide a quantity known as the Cramer-Rao bound, which is a lower bound on the variance of the estimate. This bound depends only on the model and the statistics of the data, and is independent of the actual estimation technique used. While it is not as useful as an upper bound, it is an inherent property of the data itself, and can tell us if our expectations concerning estimation accuracy are reasonable.

There are two formulas for calculating this bound [2]. In this case, the easier one is

$$\mathrm{Var}[f_o] \quad \mathrm{CRB} = \left[-E \left(\frac{\delta^2 \log p\left(\underline{X}; f_o\right)}{\delta f_o^2} \right) \right]^{-1} \tag{8}$$

We are interested only in the parameter f_o. Suppose σ_N^2 and σ_T^2 are constants. The likelihood function is

$$l(f_o) = \log p(\underline{X}\,;\, f_o, \sigma_N^2, \sigma_T^2) = -\frac{N}{2} \log 2\pi - \frac{1}{2} \sum_{k=1}^{N} \left[\log \sigma_x^2(f_k) + \frac{X_k}{\sigma_x^2} \right]$$

where

$$\sigma_x^{\,2}(f_k) = \sigma_N^{\,2} + g_n^{\,2}(f_k)\,\sigma_T^{\,2}$$

and define

$$g_n(f_k) = \left| H_n(f_k) \right|^{2} = \frac{1}{1+\left(\frac{f_o}{f_k}\right)^{2n}}$$

Then

$$\frac{1(f_o)}{f_o} = -\frac{1}{2}\sum_{k=1}^{N}\left(\frac{1}{\sigma_x^{\,2}(f_k)} - \frac{X_k^{\,2}}{\sigma_x^{\,4}(f_k)}\right)\sigma_T^{\,2}\;\frac{-4nf_o^{\,2n-1}}{f_k^{\,2n}}\;g_n^{\,3}(f_k)$$

$$\frac{2\;1(f_o)}{f_o} = -\sum_{k=1}^{N}\left\{\left(\frac{2X_k^{\,2}}{\sigma_x^{\,2}(f_k)} - 1\right)\frac{\sigma_T^{\,4}}{\sigma_x^{\,4}(f_k)}\;\frac{8n^2 f_o^{\,4n-2}}{f_k^{\,4n}}\;g_n^{\,6}(f_k) + \left(1-\frac{X_k^{\,2}}{\sigma_x^{\,2}(f_k)}\right)\frac{\sigma_T^{\,2}}{\sigma_x^{\,2}(f_k)}\;\frac{2n\,f_o^{\,2n-2}}{f_k^{\,2n}}\;g_n^{\,3}(f_k)\left[\frac{6\,f_o^{\,2n}}{f_k^{\,2n}} - 2n+1\right]\right\}$$

Taking the expectation over X, $E[X_k^{\,2}] = \sigma_x^{\,2}(f_k)$ results in

$$-E\left[\frac{{}^{2}1\,(f_o)}{f_o^{\,2}}\right] = \sum_{k=1}^{N}\frac{\sigma_T^{\,4}}{\sigma_x^{\,4}(f_k)}\;\frac{8n^2\,f_o^{\,4n-2}}{f_k^{\,4n}}\;g_n^{\,6}(f_k)$$

and the bound is

$$\mathrm{CRB} = \left[8n^2\sum_{k=1}^{N}\frac{\sigma_T^{\,4}}{\sigma_x^{\,4}(f_k)}\;\frac{f_o^{\,4n-2}}{\quad}\;g_n^{\,6}(f_k)\right]^{-1} \tag{9}$$

The bound depends on the parameters f_o, σ_N^2, σ_T^2 as well as on the filter order n. Sufficient accuracy can be obtained by using the estimated values of the parameters instead of the true values. Generally, the bound is lower for higher values of filter order n. It also decreases as σ_N^2 (background noise power) decreases.

Although there is no guarantee that an estimator meets the Cramer-Rao bound, some good properties of the ML estimator recommend its use. The ML estimator is consistent, i.e., as N (the number of observations) increases. the estimate converges to the true value. The ML estimator is also asymptotically efficient, i.e., as N increases, the variance tends toward the Cramer-Rao bound. Thus, the performance of a ML estimator can be improved by sampling the frequency spectrum more finely.

E. Extension to Arbitrary $S_T(f)$

In the previous sections, we assume that the variance of T(f) is constant over frequency. This assumption can be relaxed if information about the transmitted terrestrial radio noise is available. Then equation (3) can be modified as

$$X(f_k) = N(f_k) + \left| H_n\ (f_k) \right|^2 T_o\ (f_k)\ R(f_k)$$

where the variance of T_o dos not depend on frequency. $R(f_k)$ is a function which expresses the relative terrestrial noise power in the frequency band. It can be multiplied by any scale factor as convenient. The variance of X becomes

$$\sigma_x^2\ (f_k) = \sigma_n^2 + \left| H_n\ (f_k) \right|^2\ R^2\ (f_k)\ \sigma_T^2 \tag{10}$$

which can then be used in the maximum likelihood formulation of equation (7).

The Cramer-Rao bound becomes

$$\text{CRB} = g_n^{\,2} \sum_{k=1}^{N} \left[\frac{\sigma_T^{\,4}}{\sigma_X^{\,4}(f_k)} \quad \frac{f_o^{\,4n-2}}{f_k^{\,4n}} R^{\,4}(f_k)\, g_n^{\,6}(f_k) \right]^{-1} \tag{11}$$

Thus if the signal power being transmitted near the subsatellite point is known, it can be included in the above technique. In the absence of such information, the assumption $R(f_k)$ = constant must be made.

F. Initial Estimates of σN^2 and σT^2

Estimates of σ_N^2 and σ_T^2 are obtained by another maximum likelihood technique used in supervised learning [3]. Suppose the ionosphere behaves as an ideal filter. Then there are two cases

W1: Case 1. $S_N(f)$ only present $\sigma_X^2 = \sigma_N^2 = \sigma_1^2$

W2: Case 2. $S_N(f)$ and $S_T(f)$ present $\sigma_X^2 = \sigma_N^2 + \sigma_T^2 = 2^2$

The probability density of X_k can be described as the following mixture density

$$p(X_k;\underline{\theta}) = \sum_{j=1}^{2} p(X_k \mid W_j;\, \sigma_j^{\,2})\ P(Wj) \tag{12}$$

where $P(W_1)$ is the proportion of channels for which Case 1 applies, and $P(W_2)$ the proportion for which Case 2 applies, with $P(W_1) + P(W_2) = 1$, and the parameter vector is $\underline{\theta}$ is $(\sigma_1^{\,2}, \sigma_2^{\,2})$. The log likelihood function is

$$l(\underline{\theta}) = \sum_{k=1}^{N} \log p(X_k;\underline{\theta}) \tag{13}$$

the derivative with respect to $\sigma_i^{\,2}$ is

$$\frac{\partial 1}{\partial \sigma_i^{\;2}} = \sum_{k=1}^{N} \frac{1}{p(X_k;\underline{\theta})} \frac{\partial}{\partial \sigma_i^{2}} [p(X_k \;\; W_i \; ; \sigma_i^{2}) \; P(w_i)$$

Noting that $P(W_1 \; X_k \; ; \sigma_i^{\;2}) = \dfrac{p \; (X_k \; W_i \; , \;\; \sigma_i^{2}) \; P(W_i)}{p \;\; (X_k \; ; \; \underline{\theta})}$

we can obtain the useful form

$$\frac{\partial 1}{\partial \sigma_i^{\;2}} = \sum_{k=1}^{N} P(W_i \big/ X_k \; ; \;\; {}_k{}^{2}) \frac{\partial}{\partial \sigma_i^{\;2}} \log p(X_k \big/ W_i \; ; \sigma_i^{\;2}) = 0 \qquad (14)$$

Now since X_k is Gaussian and zero mean,

$$\log p(X_k \; W_i \; ; \;\; \sigma_i^{\;2}) = -\frac{1}{2} \log 2\pi - \frac{1}{2} \log \sigma_i^{\;2} - \frac{1}{2} \frac{X_k^{\;2}}{\sigma_i^{\;2}}$$

and then

$$\frac{\partial 1}{\partial \sigma_i^{\;2}} = -\frac{1}{2} \sum_{k=1}^{N} P \; (W_i \big/ X_k \; ; \sigma_i^{\;2}) \; (1 - \frac{X_k^{\;2}}{{}_i^{\;2}}) \; \frac{1}{\sigma_i 2} = 0 \qquad (15)$$

The above cannot be solved in closed form, but an iterative approach can be used:

$$\sigma_i^{\;2} \text{ (new)} = \frac{\sum_{k=1}^{N} P \; (W_i \; / X_k \; ; \sigma_i \; 2) \;\; X_k^{\;2}}{\sum_{k=1}^{N} P \; (W_i \big/ X_k \; ; \sigma_i^{\;2})} \qquad (16)$$

where

$$P \; (W_i \big/ X_k \; ; \sigma_i \; 2) = \frac{P \; (X_k \big/ W_i \; ; \sigma_i^{\;2}) \;\; P \; (W_i)}{\sum_{j=1}^{2} P \; (X_k \big/ W_j \; ; \sigma_j \; 2) \; P \; (W_j)} \qquad (17)$$

The $P(W_i)$ are also unknown. They are also found iteratively by

$$P(W_i)^{(new)} = \frac{1}{N} \sum_{K=1}^{N} P(W_i / X_k ; \sigma_i^2) \tag{18}$$

Equations (16-18) provide the method used to obtain initial estimates of the σ_i^2, and therefore σ_N^2. Typically this algorithm is run for about ten iterations, and the results then used in the estimation method of Part C.

REFERENCES

[1] J. H. McClellan, T. W. Parks, L. R. Rabiner "A Computer Program for Designing FIR Linear Phase Digital Filters" IEEE Trans. on Audio and Electroacoustics, Vol AU-21, No. 6, 506-526 - December 1973.

[2] H. L. Van Trees <u>Detection, Estimation and Modulation Theory, Part 1</u>. John Wiley and Sons, Inc., New York, 1968 Chapter 2.4.2

[3] R. O. Duda, P. E. Hart <u>Pattern Classification and Scene Analysis</u> John Wiley & Sons, Inc., New York 1973 pp 192-201

OPAQUE ---- A MEASUREMENT PROGRAM ON OPTICAL ATMOSPHERIC QUANTITIES IN EUROPE

PROGRAM PURPOSE AND OBJECTIVE

The ultimate objective for the research program is to develop a data base of atmospheric optical and infrared (IR) parameters which affect military systems. Until the present time, efforts to measure atmospheric optical parameters have been on a purely national basis. The time has now come to bring those efforts together and to consolidate the results, thus providing a more universally acceptable data base. Possible applications for the resulting data, would be in the optimization of optical or IR vision systems by taking due account of the influence of the atmosphere on the system capability. For example, one can immediately envisage the data being used in the optimization of Night Recon, Weapon Guidance and Laser Systems. The results from the measurement program would also be useful for specifying the atmospheric optical and IR (infrared) properties for operational (system deployment) purposes.

It will be possible to derive from the measured data, distribution of the frequency of occurrence for various parameters such as illumination levels, transmission, contrast reduction, as a function of geographical location, time, wave-length region, altitude, etc. Attempts are also being made to derive from the statistics, material relationships between the atmospheric optical parameters themselves as well as with the general meteorological condtions, which would be useful for prediction purposes.

Electromagnetic radiation from the ultraviolet through the visible and infrared spectrum, is affected in its propagation through the earth's atmosphere by absorption and scattering from air molecules and aerosol or haze particles and by the refractive effects of atmospheric density changes. Therefore, any optical system whose function it is to look through the atmosphere is affected in its "seeing" capabilities by these atmospheric effects. Optical systems are being used in such military functions as reconnaissance, target detection, recognition, and acquisition; also, in ranging and target designation, in weapons guidance and in optical communications, just to mention some examples. Optical systems use a wide range of wavelengths from the photopic response of the human eye through narrow and broadband systems in the visible and infrared (IR) and quasimonochromatic laser systems.

The spectral band width of such optical systems generally has to be a compromise between the desire to keep it as narrow as possible for rejection of undesirable background, and the need to make it wider to be able to receive sufficient radiation energy on the detector. Active laser systems, on the other hand, have created now the possibility for extremely narrow path band systems.

Atmospheric optical properties affect systems in different ways. In all cases atmospheric extinction reduces the optical signal, originating from the object scene, along its path through the atmosphere to the receiver. This atmospheric extinction is due to the loss of light energy in the direct beam caused by molecular and aerosol absorption and also due to light scattered out of the direction of the beam by molecules and aerosols.

Some systems, such as reconnaissance systems or contrast seekers for weapons guidance, depend on the detection of brightness differences, or contrast, between object scene elements; for instance, a target and its surrounding background. The inherent contrast between two different elements in an object scene is reduced by the atmosphere because of light scattered along the path into the direction of the receiver. This scattered light is called air light or path radiance. If the intensity of path radiance becomes of the same magnitude as the radiance from the object scene, the contrast begins to disappear and different scene elements can no longer be distinguished.

Imaging systems, looking over very long paths, require very high angular resolution. In such systems, the resolution is also reduced by atmospheric turbulence and resulting optical scintillation phenomena along the path. The picture becomes blurred and sharp edges become fuzzy.

Many optical military systems use active illumination systems, in particular lasers (for example, target designator systems, range finders). In these kinds of systems, one is concerned with the possibility that the laser beam might be detected from off-axis scattering by atmospheric aerosols and molecules.

Many studies have been conducted to determine the magnitude of these various atmospheric effects on military systems and operations. For example, in the employment of the F4/MAVERICK/EQ missile, cloud ceiling is a determining parameter for aircraft attack profiles; visibility affects the lock-on range for the target-background contrast sensor of the Maverick missile.

In another area, this data can be used to estimate the scattered light flux on an aircraft borne sensor originating from the scattering of a laser beam which is not directly hitting the aircraft.

The OPAQUE program is operating a number of stations covering the areas of Scandinavia, Central, Western and Southern Europe including some island or shipboard stations.

The choice for the location of these sites was based first on military tactical considerations; secondly on their representativeness with respect to the geographical environment and atmospheric meterological condtions; and thirdly on availability of logistical support.

All data is being reduced to a commonly agreed upon data format compatible with computerized data analysis. All the data is to be reduced to an agreed format on a six monthly basis, and then collected in a central data bank from which the participating nations will be able to obtain consolidated data tapes. Distribution of the consolidated data will be at an agreed classification level. Each member nation, or group of member nations may then perform their own analysis. Published reports will then be submitted to a central secretary, to provide a common series of OPAQUE reports for distribution to the participating nations and to others as may be agreed jointly. It is also desirable that a six monthly review be made of the data analysis with the aim of avoiding unnecessary duplication.

PHYSCIAL BACKGROUND

Every optical of IR system which is looking through the atmosphere, is affected by atmospheric light scattering, absorption, turbulence, atmospheric refraction, and nonlinear effects. The important parameters for any given line of sight for the atmospheric transmission, the atmospheric path radiance, and the turbulence factor.

These quantities are functions of the illumination conditions, the atmospheric properties themselves, and also the average terrain-background albedo. These factors affect the amount of radiation energy transmitted from a source along an atmospheric path to a receiver, the reduction of image contrast, and the distribution (spatial and temporal) of energy in the transmitted light beam.

For example, under homogeneous conditions the apparent luminance of target and background at range R can be written as

$$L_R = L_o \ e^{-\sigma R} + L_H(1 - e^{-\sigma R})$$

and

$$L'_R = L'_o \ e^{-\sigma R} + L_H(1 - e^{-\sigma R})$$

defining the contrast at zero range as

$$C_o = \frac{L_o - L'_o}{L_o + L'_o}$$

and the apparent contrast at range R as

$$C_R = \frac{L_R - L'_R}{L_R + L'_R}$$

it can be shown that

$$C_R = \frac{C_o}{1 + \frac{2L_H}{L_o + L'_o}\left(e^{+\sigma R} - 1\right)} \tag{1}$$

Alternatively, using the diffuse reflectance of the target and background A and A', Eq. (1) can be written as

$$C_R = \frac{C_o}{1 + \frac{2\pi\Omega_o L_H}{(A + A')E_v}\left(e^{\sigma R} - 1\right)} \tag{2}$$

The following definitions are relevant to the previous equations:

(1)	C_o	= Contrast at zero range
(2)	L, L'	= Luminance of target and background [Cd M^{-2}]
(3)	T	= Transmission = $e^{-\sigma R}$
(4)	o	= Attenuation coefficient [M^{-1}]
(5)	R	= Range (m)
(6)	E_v	= Illuminance indident on a vertical plane [Lux]
(7)	A, A'	= Diffuse reflectance of target and background

(8) L_H = Horizon luminance [Cd M^{-2}]

(9) Ω_o = 1 sr

(10) Lp = Path Luminance (Cd M^{-2}) per range R

From a study of these parameters and Eqs. (1) and (2), it is possible to specify a minimum number of parameters which must be measured in any program before worthwhile results can be obtained. Section 3 gives what are considered to be the minimum set which it is worthwhile to measure, including IR transmission.

ϕ_o is the off-axis angle (0^o would be the direct beam). Such signals could be detected, and in response an aircraft could take appropriate counter-actions.

The effect of illumination levels on the ability of an observer to detect targets of various sizes has gained interest and lately with the development of image intensifier systems and Low Light Level Television (LLLTV). These systems have extended passive visual seeing from twilight down to nighttime illumination levels.

The aerosol scattering and absorption properties are described in the Mie theory. This theory applies to spherical particles and a few other simple particle configurations; however, comparisons with experimental results show good enough agreement to justify for most practical purposes the application of the Mie theory to natural aerosols. The scattering and absorption coefficients for aerosol particles between sizes r_1 and r_2 and distribution N(r) are

$$S_{\lambda a} = \int_{r_1}^{r_2} r^2 \pi \cdot Q_s \left(\frac{r}{\lambda}, m_\lambda\right) \cdot N(r)\, dr$$

and

$$k_{\lambda a} = \int_{r_1}^{r_2} r^2 \pi \cdot Q_a \left(\frac{r}{\lambda}, m_\lambda \right) \cdot N(r)\, dr$$

Q_s and Q_a are functions of the ratio of particle size to wavelength, and the aerosol particle refractive index m; they are derived from the Mie theory and are called efficiency factors for scattering and absorption, respectively. The refractive index of aerosol particles is a real number of completely transparent substances (no absorption); for absorbing substances it is a complex number,

$$m = n - in'$$

where n is the real part of the refractive index and n' the imaginary part. The quantity $4\pi n'/\lambda$ is the absorption coefficient of the aerosol substance; its dimension is $(cm)^{-1}$.

These equations define all the aerosol absorption and scattering quantities which are needed to calculate the transmission of a light beam through the atmosphere.

The Raleigh and Mie theory also give relationships for the intensity of radiation scattered out of a light beam into different directions. For molecular Rayleigh scattering, the angular scattering intensities are given by

$$I_\lambda(\varphi) = I_{\lambda,o} \frac{3}{16\pi} \cdot S_{\lambda m} (1 + \cos^2 \varphi) \quad .$$

For spherical particles with sizes of the same order of magnitude or larger than the wavelength of the incident light, the Mie theory gives for the angular scattering intensitities

$$I_\lambda(\varphi) = I_{\lambda o} \cdot \frac{\lambda^2}{4\pi^2} \frac{i_1 + i_2}{2}$$

In order to predict the atmospheric optical/IR effects on the performance of a given optical system, two things are needed: (a) algorithms to calculate the specific quantity required for the physical condtions under which a system functions (for example, the transmittance for a specified spectral band and along a given atmospheric slant path), and (b) a description of the atmospheric properties which are required as input to (a) (for example the aerosol and molecular extinction coefficients as a function of wavelength, and the vertical profile of these parameters). The development of the algorithms is primarily a theoretical effort; the atmospheric properties must be derived from field measurements.

ATMOSPHERIC DATA BASIS

The basic atmospheric quantities which are needed to calculate atmospheric optical effects are: the distribution of atmospheric molecules and aerosols and their absorption and scattering properties, and the turbulance structure constant C_n1 For most of these parameters, models describing their spatial distribution and in some cases temporal variations have been developed.

Models for the distribution of atmospheric molecular density, temperature and water vapor concentration are defined in the "Standard Atmospheres" and are based on long series of measurements. The data base for other minor gaseous components are less extensive. Laboratory and field measurements exist of optical properties of the atmospheric molecular compnents, and have been compiled for optical calculations. Data on over 100,000 absorption lines from different molecular species have been compiled for wavelengths from the visible into the far infrared.

Whereas most of the molecular absorption is caused by two or three molecular species only (water vapor, carbon dioxide and ozone), the extinction due to atmospheric particulates or aerosol particles is much more complex. The concentration, size distribution, and composition of aerosols is extremely variable, spatially and in time; the development of representative models for the aerosol properties, therefore, is in a more preliminary state. Aerosol models have been developed for continental environments, such as rural or urban areas, and for maritime regions. These models describe also the variation of aerosol properties with altitude from the surface up to 100 km altitude. The vicissitude even within these models, however, makes their applicability to a given real world situation very difficult.

With the exception of those parameters, which are being collected routinely as part of the standard meterological observations (that is, temperature, humidity, surface visibility), the present data base for atmospheric optical modeling has no statistical significance. It describes only "general average" conditions.

ALGORITHMS FOR OPTICAL CALCULATIONS

Model calculations can be only as good as the input data for them. It can be stated at the outset, that in general the theoretical concepts and the algorithms for computing optical propagation characteristics are superior to the accuracy of the input data. This is certainly correct for transmission, emission, and single scattering computations, and even for most multiple scattering computations.

Algorithms for predicting scattering intensities are needed primarily for contrast reduction and sky radiance calculations. Path radiance P is composed of light which has been scattered several times along its path through the atmosphere. Depending on how optically thick the atmosphere is, light which has been scattered up to ten times or more may contribute significantly to the radiation intensities of concern. No rigorous analytical solution has been presented as of this date for the radiation transfer problem involving multiple scattering in a real, aerosol-containing atmosphere.

NEED FOR EXTENDED ATMOSPHERIC OPTICAL DATA BASE

The military community is confronted with two different types of problems in predicting the atmospheric optical effects on the performance of a given optical system. The first one is associated with the question: "What is the probability that a given system will be able to function successfully in a certain environment?" This question is being asked by the systems designer to derive the appropriate systems specifications. It is also being asked by the systems analyst to develop the proper operational procedures and deployment plans for a system.

The second type of requirement is for a specific forecast for the atmospheric optical environment at the time and place of deployment of the system.

The first problem is one of developing an atmospheric optical climatology, and the second one requires a capability to forecase atmospheric optical conditions from the available standard meteorological observations and the general weather forecast. At present, however, one has in many cases only a qualitative understanding of the physical relationships between the meteorological factors and the environment on one side, and the atmospheric optical properties on the other. A quantitiative definition of these relationships requires also a reliable and comprehensive data base as a source for empirical correlations.

While such data are useful for system design planning, they must be used cautiously in operational planning. The data will be misused if employed to estimate the chances that a single or small number of missions will find suitable conditions of cloud cover and surface visibility. If the data are used in relation to large number of missions, then more conficence can be attached to them.

These data can be of use to the operational weather forecaster, however, by providing him useful relationships for shaping his forecasts. Since the type of air mass is so important in influencing visibility and cloud cover, a refinement of the data by air mass type of synoptic situation would be even more useful to the forecaster.

The atmospheric environment is of a quite decisive influence on many electro-optical systems. In most cases, the atmospheric medium constitutes a limiting factor for the propagation of electro-magnetic signals. In some other cases, there are the optical properties of this medium which render the systems infeasible. Thus to support electro-optics system planners, designers and users with probability of occurrence information on earth surface propagation conditions, it is desirable to have statistics of specific propagation parameters available.

The natural aerosol is a basic constituent of the atmosphere, having considerable impact generally on optical propagation properties. Its basic parameters are composition, concentration, and distribution.

Though considerable effort has been addressed in the past by many research workers to measuring and understanding the characteristics of natural aerosols, the state-of-the-art must be considered still yet as quite preliminary. In fact there are neither comprehensive statistics available containing sufficient information for purposes mentioned above, nor are there unique models which can be used for operational purposes free of ambiguities.

The reason for this lies in the complexity of aerosol composition, the vast mechanisms of its formation, its complex dynamics, and in its complex way of ageing. One other problem in the past has been the state-of-the-art of data collecting devices which did not easily permit measuring aerosol data over larger size intervals with satisfactory resolution.

The situation is improving now with the advent of new aerosol particle counters.

Typically, aerosol size distribution measurements generate a huge stream of original data. These must be channeled down, mainly for two reasons. First, because of economics all redundancy in the data should be removed to the largest extent possible. Secondly, there is virtually no way to sensibly work with, analyze, or use in later processing applications this bulk of original data.

Generally, in a situation like this the benefit of a model consists in providing some kind of mathematical expression through which the data could be greatly compressed. However, there is no such single expression capable of describing size distributions. Junge's formula is referred to very widely in the literature. It is an empirical model in the sense that observations show that it tends to be followed, but not that it must be followed. Yet, for the applications metnioned above, Junge's formula does not seem to be adequate enough. Other distribution functions are the standard and the modified gamma distribution or the normal and the lognormal distributions. Gamma, normal and lognormal distributions have proven to be quite versatile. Generally they are used for typical background aerosols. Sums of normal and lognormal distributions are used to fit distributions originating from a single source through a single mechanism. These few examples were cited to show that very often the choice of a specific mathematical expression seems to be justified only by the purpose of use. A highly accurate method for fitting aerosol data is available through the use of spline functions. However, this approach normally yields a large number of coefficients which do not correspond to sphysically meaningful parameters.

2. AEROSOL SIZE DISTRIBUTION MEASUREMENTS

2.1 Size Distribution Function

Assume an aerosol is contained in a unit probing volume. The basic method for determining its size distribution function is to count the total number, N(D) of aerosol particles (per unit volume) whose diameter values fall below a certain value, D. The function N(D) is a cumulative function with values between zero and N, the total number of aerosol particles contained within the unit volume. Generally, the number of aerosol particles in a given volume is very large, and their sizes may be assumed to vary continuously. Thus, N(D) approaches a continuous, smooth function as N increases. It is then possible to derive the size distribution function, $n(D)$, by

$$n(D) = \frac{dN(D)}{dD} \tag{2.1}$$

or in integral representation,

$$N(D) = \int_0^D n(D')\, dD' \tag{2.2}$$

In practice the determination of $N(D)$ leads to serious problems since it is, for experimental reasons, possible only to measure particles with sizes above a certain, device-dependent, lower limit, say D_L. Thus $N(D)$ is determined up to a constant $N(D_L)$. Since $n(D)$ is a derivative function, this ambiguity is removed.

The aerosol counters employed in this work provide data on the number of particles (per unit volume) for some specified diameter intervals $D_i \leq D \leq D_{i+1}$, with $i = 1,2,\ldots,M$. If N_i denotes the number of particles for the i-th channel, then because of (2.2) one has

$$N_i = \int_{D_i}^{D_{i+1}} n(D')\, d\, D' \tag{2.3}$$

Accordingly, N_i may be used directly to evaluate an estimate for the size distribution function, n, at the point D_i that is

$$n(D_i) \approx n_i := \frac{N_i}{D_{i+1} - D_i} \; ; \; i = 1,2,\ldots,M. \tag{2.4}$$

3. DATA FIT OBJECTIVES

If data are going to be fitted by some algorithm, two basically different cases are to be distinguished. In the one case, the data result from a process the physical nature of which is understood in the sense that a physical model or theory is available. In this case, data fitting permits the determination of some free model parameters through adjusting the model predictions to the set of measured data. In the other case, there is no physical model at hand to predict the experimental findings, and data fits in this case are primarily another way of data presentation. If the fit algorithm is chosen carefully and appropriately adapted to the problem, a few important objectives may be met simultaneously:

A. Data representation:

The insertion of a smooth fit curve into the plot of data points is a useful way to assist the understanding of the general trends underlying the data set by inspection.

B. Error smoothing:

The smooth fit curve is likely to smooth out high frequency contributions within the set of data due to random error. This is especially true if a least-squares-fit algorithm is applied.

C. Data reduction:

Through data fitting, one can generally reduce the large amount of data to a few functional parameters only. Frome these, it is possible to retrieve the original data (in the accuracy desired), as well as to calculate new interpolated data.

D. Empirical modelling:

Generally, it is much easier to correlate any typical condition of the process under study with a typical pattern of the reduced data set, than with the original samples.

The Least Squares Method

Measured data with values v_i ($i = 1,2,\ldots,M$) are assumed to be a function of a single variable, t. The experimental result is a list of M data pairs (t_i,v_i). The measuring points, t_i, need not be spaced regularly. In the least-squares method, a function F from an appropriate collection $\mathcal{F}$ is determined such that

$$\sum_{i=1}^{M} (F(t_i) - v_i)^2 = \min. \tag{4.1}$$

For ease of computation, the function space $\mathcal{F}$ is chosen as to be formed by all linear combinations of some basis functions, f_k. That is, any $F \in \mathcal{F}$ can be written as

$$F(t) = \sum_{k=0}^{K} b_k f_k (t) . \tag{4.2}$$

The set of f_k must be linearly independent. Though desirable in many cases, the f_k need not necessarily obey some orthogonality condition. In most applications, the function space $\mathcal{F}$ is deliberately restricted by requiring

$$K < M. \tag{4.3}$$

The free amplitudes, b_k, in (4.2) are determined through (4.1) as solutions of the Gaussian normal equations

$$A \,.\, B = C, \tag{4.4}$$

where A is a positive definite, square symmetric matrix with elements

$$a_{kl} = \sum_{i=1}^{M} f_k (t_i) f_l (t_i), \; (k, l, = 1,\ldots K), \tag{4.5}$$

and C is a column vector with elements

$$c_k = \sum_{i=1}^{M} f_k(t_i) \cdot v_i, \quad (k - 1,\ldots K). \tag{4.6}$$

The vector B of elements b_k is the solution.

The problem encountered in determining solutions of Eq. (4.4) is that the matrix A usually tends to be ill-conditioned. As a result, numerical instabilities arise. The problem is reduced somewhat if an appropriate function system, f_k, is used for expansion. Furthermore, it is advisable not to apply the genral Gaussian inversion technique to deduce the solution, B, rather than make use of Cholesky's method, a numerically stable matrix decomposition technique which fully exploits the symmetry and positive definiteness of matrix A.

4.2 Data Transformation

Size distributions of natural aerosols tend to drop off in magnitude over several decades within the diameter range from .1 to 30 m. Therefore, the fit usually is applied not to the original data n_i but rather to their logarithms

$$v_i = \log n_i, \quad (i=1,\ldots,M). \tag{4.7}$$

It should be noted that his is, in fact, a substantial step in the data analysis. Indeed, if the fitted curves are going to be employed in any optical propagation calculation, taking the logarithms is a necessary approach. It has been shown that the results of such calculations depend to a very large extent on the number of small aerosol particles as well as on that of larger particles, although the latter usually are much less populated. For the purpose of optical calculations, the "goodness" of fit should be uniform over the entire spectrum of particle diameters considered regardless of their magnitude in concentration. This is quite naturally achieved in using the logarithmic scale since then the least-squares fit results in an overall minimized relative error for the original data.

Typically for the aerosol counters employed in this type of work, the widths of their different size channels generally increase with increasing diameter values, D_i.

4.3 Chebyshev Polynomials

The success of the fitting procedure depends critically on the set of basis functions, f_k, to be used for expansion of the data points. In the present work it is suggested that for fitting natural aerosol size distributions, the use of Chebyshev polynomials is highly favorable. In this case, formula 4.2 is called a Chebyshev expansion and is of the form

$$F(t) = b_o T_o + b_1 T_1 + \ldots + b_k T_k. \tag{4.10}$$

This saves computing time considerably, and enables an easy check for convergence. However, it is then more difficult to interpret the results of the fit in a straight forward and general way.

Chebyshev polynomials are well known from many textbooks on approximation theory and smoothing. There, one usually may also find an outline of their special properties which makes them a rather unique set of basis functions for use in numberical fit algorithms. The first few members of this set are given by (10)

$$\begin{aligned}
T_o(t): &\quad = 1 \\
T_1(t): &\quad = t \\
T_2(t): &\quad = -1 \quad +2t^2 \\
T_3(t): &\quad = -3t \quad + 4t^3 \\
T_4(t): &\quad = 1 \quad - 8t^2 \quad + 8t^4 \\
T_5(t): &\quad = 5t \quad -20t^3 \quad +16t^5 \quad .
\end{aligned} \tag{4.11}$$

The first polynomials T_o to T_7 are shown graphically in Figure 2 for the interval -1 t 1. From this figure it may be seen that

$$\left| T_k(t) \quad 1 \; ; \; |t| \leq 1. \qquad (4.12)$$

The definitions (4.11) indicate that the polynomials T_k are built up from either even or odd power functions in t. The amplitudes of the power functions obviously tend to be very large in magnitude, and it is because of their alternate signs that the T_k stay within the limits (4.12). The numerical evaluation of series like this may generally be subject to serious computational errors for computers having limited word lengths. A convenient way out of this problem is to use the three-terms recurrence relation

$$T_{k+2}\ (t) = 2t \quad . \quad T_{k+1}\ (t) - T_k(t) \quad , \ k=0,1,2,... \qquad (4.13)$$

in conjunction with the basic definition (4.11) for T_o and T_1.

Some Examples

Though the examples chosen are rather simple, they are of general interest in the study of aerosols. They are:

- The Junge size distribution
- A unimodal lognormal size distribution
- The haze L model

The Junge size distribution is of the form

$$n\ (D): = \frac{dN}{dC} = C' \quad D^{-(\beta+1)} \ ,$$

where the prime is to indicate that, since n is a function of diameter rather than of radius, the concentration C differs in value from those to be found in conventional tabulations.

On a log-log-scale, n(D) plots as a straight line. In this case, the Chebyshev fit results in a linear function

$$F(t) = \sum_{k=0}^{1} b_k T_k(t)$$

$$= b_o + b_1 t . \qquad (4.15)$$

The interpretation of the two amplitudes b_o and b_1 is most conveniently done in the double logarithmic plot. It may be seen that b_o mainly determines the absolute concentration while b_1 determines the slope.

Junge distribution and Chebyshev approximations

The Junge parameters, C' and B can be expressed uniquely in terms of the amplitudes, b_o and b_1.

$$\log C' = b_o - b_1 \left(\frac{\log (D_H \cdot D_L)}{\log (D_H / D_L)} \right) \tag{4.16}$$

$$B + 1 = \frac{2}{\log (D_H / D_L)} \cdot b_1 .$$

From this it is clear that the amplitudes b_k are dependent on the low and high ends, D_L and D_h respectively, of the measured diameter interval. Thus, when comparing Chebyshev amplitudes for measurements taken with different devices, the corresponding values for D_L and D_H must be taken into consideration.

In case of Junge distributions, we can avoid this unpleasant device-dependence. However, the problem is much more involved for more general distributions containing many more terms. We hope to resolve this problem in future studies.

The next example is that of a unimodal lognormal distribution

$$n(D) = \frac{2N'}{D. \log \delta \sqrt{2\pi}} \exp - \frac{1}{2} \cdot \frac{\log \frac{D}{D_m}}{\log \delta g}$$

Again, the prime indicates that in the current notation the parameter N' is different in value from that in the literature where n has been considered to be a function of the particle radius. The example illustrated is an average size distribution for maritime aerosols, and taken here at the experimental meshpoints of the present work. The median number diameter accordingly is $D_m = 0.52\mu m$; the logarithmic standard deviation is $\delta_g = 4.8$.

One can see that the Chebyshev fit of a unimodal lognormal distribution is quite excellent. Moreover, three coefficients are sufficient to specify this fit, i.e., b_o, b_1, and b_2 as one would expect since a lognormal distribution is represented in the double logarithmic scale by a parabola of order 2.

The physical parameters N',D_m and δ_g can be expressed uniquely in terms of the amplitudes b_o, b_1, and b_2.

The last example we consider is that of Deirmendjian's Haze L model distribution

$$n(D) = a' D^{\alpha} \exp \left(-b' \cdot D^{\gamma}\right) \tag{4.18}$$

with the following parameter values

$$a' = 2.4878 \quad 10^{6} \; [\mu m^{-\alpha-1}] \quad , \quad \alpha = 2$$

$$b' = 10.6905 \quad [\mu m^{\gamma}] \quad , \quad \delta = 5$$

Unlike the two other examples, the distribution does not in the log-log-scale represent a simple polynomial. Thus, the Chebyshev fit in this case is really an approximation.

The few examples presented have been chosen to demonstrate the main advantages and shortcomings of fit time algorithms. The fit was shown to be versatile in the sense that various size distributions can be treated although they are of quite different nature. All of them are basically related to the aerosol field. Moreover, unlike in cases of orthogonal polynomials, the fixed basis of Chebyshev interpretation of the fitted data in terms of a reduced data set, namely the resulting Chebyshev amplitudes. A remaining difficulty lies in the fact that the resulting amplitudes are dependent on the range of particle diameters covered in the fit, and therefore to some extent are device-dependent.

OPAQUE DATA BASE

I. DATA BASE

A. Various data base definitions have been examined and implemented for test purposes. Questions in this area revolve around two points of concern:

1. Data Structure - A relatively stable logical structure has evolved since the data content is well defined and understood.
2. Data Retrieval - (Repeating group time assignments, element designation) - A good deal of experimentation and consideration has been given to the method of storing data i.e. the representation of a logical entry (HR., DAY, MO.). This will play a key role in determining data base size and retrieval and update capability. Hour as the logical entry will allow data history storage, while day as logical entry requires data sets to represent hours of data.

B. Data Base Loading - Two methods of initial data base loading have been considered:

1. Natual language "load" capability - natural language provides a rudimentary capability to mass load a data base with machine stored data. This capability has been examined and tested. Although successful, the method is cumbersome requiring a two-step process.
2. PLI Capability - PLI provides more powerful tools for data base communication i.e. use of Host Fortran Processing) and appears to be the method which will be chosen for data base loading. Preliminary Fortran coding has been completed and PLI statement requirements examined, but no PLI tests have been performed.

C. Data Base Maintenance - The following areas are of interest:

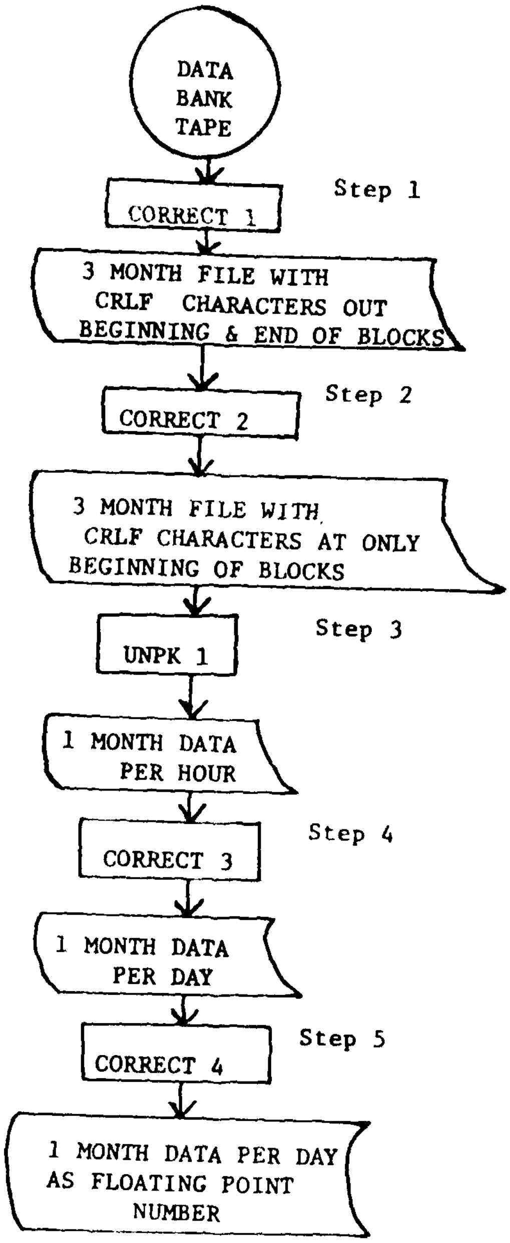
1. Retrieval - plans have been made to incorporate PLI statements into existing interactive graphics software. No testing has occurred but preliminary coding has been completed.

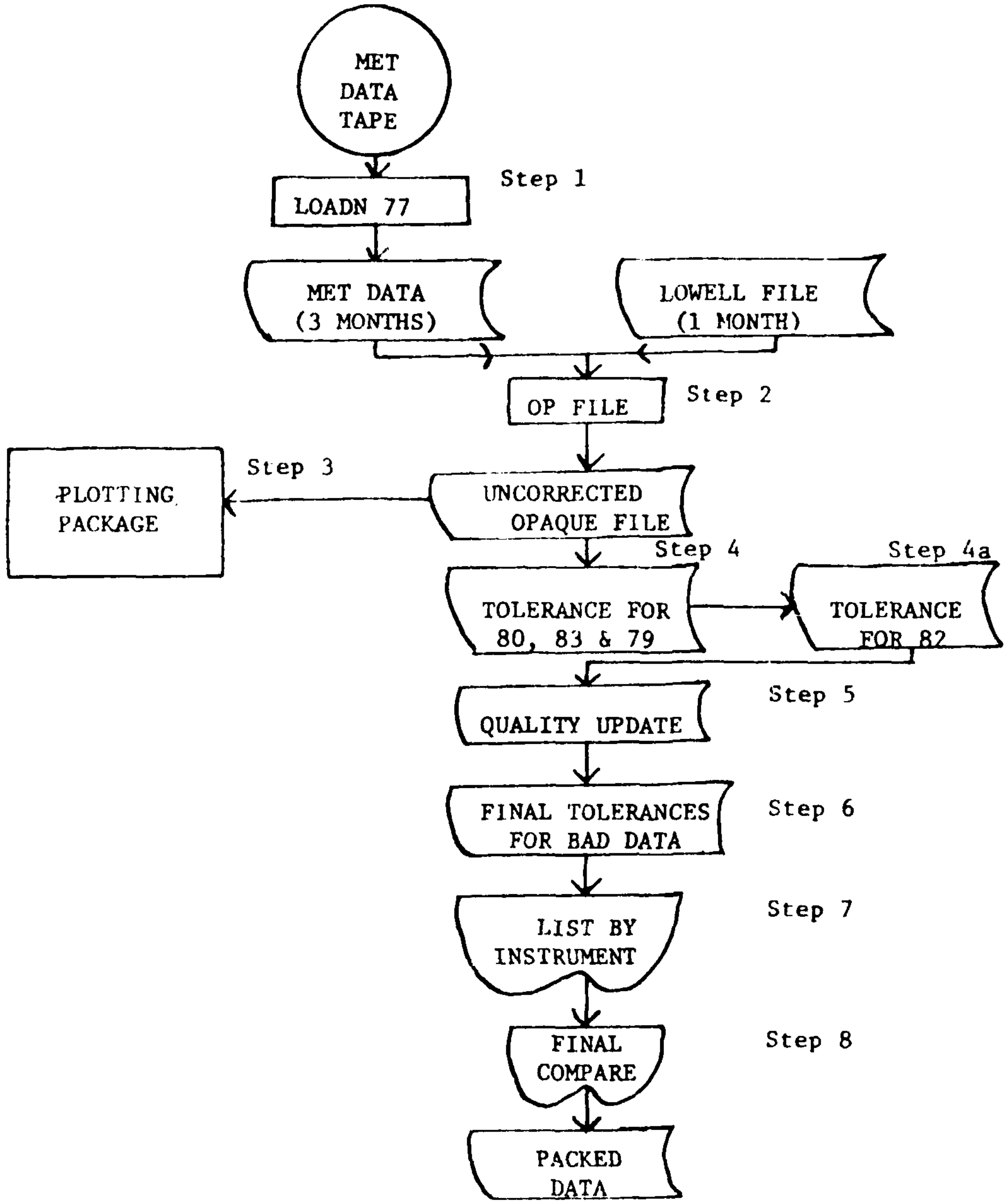
2. Update - a procedure has been defined to allow interactive editing of OPAQUE data resident in the data base.
3. DBA Functions - a procedure to provide for transaction log, automatic backup, recovery, etc. has been defined.

II. INTERACTIVE PLOTTING SOFTWARE

A. Interactive plotting software currently exists and will provide single plots of any data word. No provision exists to test quality value or data out of range, but log or linear plots are available, automatic or manual Y-Axis scaling is possible, labels are provided, and titles may be entered by the user.

B. A second version of (A) above exists which incorporates those features absent from (A) but no testing has occurred. This system will be split to form a log and a linear program.


DATA
BANK
TAPE
Step 1
CORRECT 1
3 MONTH FILE WITH
CRLF CHARACTERS OUT
BEGINNING & END OF BLOCKS
Step 2
CORRECT 2
3 MONTH FILE WITH
CRLF CHARACTERS AT ONLY
BEGINNING OF BLOCKS
Step 3
UNPK 1
1 MONTH DATA
PER HOUR
Step 4
CORRECT 3
1 MONTH DATA
PER DAY
Step 5
CORRECT 4
1 MONTH DATA PER DAY
AS FLOATING POINT
NUMBER


MET DATA TAPE
Step 1
LOADN 77
MET DATA (3 MONTHS)
LOWELL FILE (1 MONTH)
OP FILE
Step 2
PLOTTING PACKAGE
Step 3
UNCORRECTED OPAQUE FILE
Step 4
Step 4a
TOLERANCE FOR 80, 83 & 79
TOLERANCE FOR 82
QUALITY UPDATE
Step 5
FINAL TOLERANCES FOR BAD DATA
Step 6
LIST BY INSTRUMENT
Step 7
FINAL COMPARE
Step 8
PACKED DATA

ARRAY SPECIFICATION - DIRECT ACCESS FILE - DATA BASE SOFTWARE

TYPE: INTEGER

DIMENSION: 140

CONTENT BY WORD:

ARRAY WORD #	DATA	OPAQUE WORD #	FORMAT
1	Station No.	1	12
2	Date	2	16
3	Time	3	14
4	Duration of Meas. Cyc.	4	12
5	Comment Code 1	5	13
6	" " 2	6	13
7	" " 3	7	13
8	" " 4	8	13
9	" " 5	9	13
10	Scattering Value	10-1	11
11	Filter Value	10-2	11
12	Humidity Value	10-3	11
13	MRI Photopic, Beg. Val.	11	14
14	" " , Fin. Val.	12	14
15	" " , Max. Val.	13	14
16	" " , Min. Val.	14	14
17	" " , No. Meas.	15	13
18	Eltro Transmissometer, Beg. Val.	16	14
19	" " , Fin. Val.	17	14
20	" " , Max. Val.	18	14
21	" " , Min. Val.	19	14
22	" " , No. Meas.	20	13
23	Horizontal Luxmeter, Beg. Val.	21	14
24	" " , Fin. Val.	22	14
25	" " , Max. Val.	23	14
26	" " , Min. Val.	24	14
27	" " , No. Meas.	25	13
28	Vertical Luxmeter, North Value	26	14
29	" " , East Val.	27	14
30	" " , West Val.	28	14
31	" " , South Val.	29	14
32	Night Path Luminance, Beg. Val.	30	14
33	" " " , Fin. Val.	31	14
34	" " " , Max. Val.	32	14
35	" " " , Min. Val.	33	14
36	" " " , No. Meas.	34	13

ARRAY WORD #	DATA	OPAQUE WORD #	FORMAT
37	Variance Path Pct. Meter, Beg. Val.	35	14
38	" " " " , Fin. Val.	36	14
39	" " " " , Max. Val.	37	14
40	" " " " , Min. Val.	38	14
41	" " " " , No Meas.	39	13
42	" " " " , South Val.	40	14
43	" " " " , West Val.	41	14
44	" " " " , North Val.	42	14
45	Filtered Epplet Data, = .945	43	14
46	" " " , = .4	44	14
47	" " " , = .87	45	14
48	" " " , = 1.06	46	14
49	" " " , = .75	47	14
50	" " " , = .55	48	14
51	" " " , Photopic	49	14
52	" " " , + .3-3.51	50	14
53	" " " , Direct	51	14
54	" " " . = .3-3.52	52	14
55	Barnes Trans., 3-5	53	14
56	" " , 8-12	54	14
57	" " , 8-13	55	14
58	" " , Open or 4	56	14
59	" " , 3-5 Fin	57	14
60	Cloud Cover	68	11
61	Wind Dir. at 10^M	69	12
62	Wind Speed at 10^M	70	12
63	Wind Dir. at 2^M	71	12
64	Wind Speed at 2^M	72	12
65	Pressure	73	13
66	Temperature	74	13
67	Dew Point Temperature	75	13
68	Rain Rare	76	13
69	General Ground State	77	11
70	Total Rain Past Hour	85	13
71	Contel and Add. Meas. Code	64-1	11
72	Aerosol Ins. Code	58-1	11
73	Aerosol Date Code	58-2	11
74	Max. Channel Used	58-4	11
75	Aerosol Rel. Code	58-3	11
76	Add. Extinction Val.	64-3,4,5	13
77	Add. Extinction Val. Rel. Code	65-1	11
78	CO2 Concentration Val.	65-2,3,4	13
79	" " Rel. Code	65-5	11

ARRAY WORD #	DATA	OPAQUE WORD #	FORMAT
123	Barnes Transmissometer Qual.,55	84-1	11
124	" " " ,56	84-2	11
125	" " " ,57	84-3	11
126	" " " ,57	84-4	11
127	" " " ,59	84-5	11
128	Aerosol Channel 1 Data	58-5,59-1	12
129	" " 2 "	59-2,3	12
130	" " 3 "	59-4,5	12
131	" " 4 "	60-1,2	12
132	" " 5 "	60-3,4	12
133	" " 6 "	60-5,61-1	12
134	" " 7 "	61-2,3	12
135	" " 8 "	61-4,5	12
136	" " 9 "	62-1,2	12
137	" " 10 "	62-3,4	12
138	" " 11 "	62-5,63-1	12
139	" " 12 "	63-2,3	12
140	" " 13 "	63-4,5	12

A) The following terminology has been used in producing this document:

a) Data Bank - OPAQUE Random Access Disk Files.
b) Data Base - OPAQUE Data Managed by S2K DBMS.

B) Hyphenated data bank word numbers reflect digit positions within that word, with position "1" always the most significant digit.

C) All data base component names have been chosen to be Fortran compatible (length of name, default variable type) for ease of use with PLI software.

D) Data base component names follow one of two formats as follows:

a) Non-quality data
   i) Letter "I" (or "N" for "no. meas." data item) - indicating Fortran integer type variable.

   ii) Data item code - it should be noted that when partial data bank words are used, an attempt has been made, in general, to indicate this by including the digit positions in the component name.

b) Quality data
   i) Letter "I" - indicating Fortran integer type variable.

   ii) Letter "Q" - indicating quality data.

   iii) Data base component number - indicating which data quality figure is for.

E) Data bank word 64-2 (data item F-2) is not used and is set to zero in the data bank - this has no bearing on the OPAQUE data base.

DATA BANK RECORD NO.	DATA BANK COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION	PICTURE	DECODING ALGORITHM
	C1	ISTN	-	STATION NO.	9(2)	4
2	C2	IDATE	-	DATE	9(6)	YYMMDD
3	C3	ITIME	-	TIME	9(4)	HHMM
4	C4	IDUR	-	DURATION OF MEAS. CYC.	9(2)	MM
-	C5	ISTAT	-	UPDATE STATUS	9(1)	5
to 9	C11	ICOMNT	-	COMMENT CODE	9(3)	6
10-1	C21	ISCAT	-	SCATTERING	9(1)	7
10-2	C22	IFILT	-	FILTER	9(1)	8
10-3	C23	IHUM	-	HUMIDITY	9(1)	9
11	C101	ISSBEG	$S_S BEG$	MRI PHOTPIC, BEG. VAL.	9(4)	1
12	C102	ISSFIN	$S_S FIN$	" " , FIN. VAL.	9(4)	1
13	C103	ISSMAX	$S_S MAX$	" " , MAX. VAL.	9(4)	1
14	C104	ISSMIN	$S_S MIN$	" " , MIN. VAL.	9(4)	1
15	C105	NVSS	N_V	" " , NO. MEAS.	9(3)	--
16	C201	IEGBEG	$E_G BEG$	ELTRO TRANSMISSOMETER, BEG. VAL.	9(4)	1
17	C202	IEGFIN	$E_G FIN$	" " , FIN. VAL.	9(4)	1
18	C203	IEGMAX	$E_G MAX$	" " , MAX. VAL.	9(4)	1
19	C204	IEGMIN	$E_G MIN$	" " , MIN. VAL.	9(4)	1
20	C205	NVEG	N_V	" " , NO. MEAS.	9(3)	--
21	C301	IELBEG	$E_L BEG$	HORIZONTAL LUXMETER, BEG. VAL.	9(4)	1
22	C302	IELFIN	$E_L FIN$	" " , FIN. VAL.	9(4)	1
23	C303	IELMAX	$E_L MAX$	" " , MAX. VAL.	9(4)	1
24	C304	IELMIN	$E_L MIN$	" " , MIN. VAL.	9(4)	1
25	C305	NVEL	N_V	" " , NO. MEAS.	9(3)	--

DATA BANK RECORD NO.	DATA BASE COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION	PICTURE	DECODING ALGORITHM
26	C401	IEVN	E_V^N(north)	VERTICAL LUXMETER, NORTH VAL.	9(4)	1
27	C402	IEVE	E_V^E(east)	" " EAST VAL.	9(4)	1
28	C403	IEVS	E_V^S(south)	" " SOUTH VAL.	9(4)	1
29	C404	IEVW	E_V^W(west)	" " WEST VAL.	9(4)	1
30	C501	ILPBEG	L_P^{NT} BEG	NIGHT PATH LUMINANCE, BEG. VAL.	9(4)	2
31	C502	ILPFIN	L_P^{NT} FIN	" " " , FIN. VAL.	9(4)	2
32	C503	ILPMAX	L_P^{NT} MAX	" " " , MAX. VAL.	9(4)	2
33	C%04	ILPMIN	L_P^{NT} MIN	" " " , MIN. VAL.	9(4)	2
34	C505	NVLP	N_V	" " " , NO. MEAS.	9(3)	--
35	C601	IFPBEG	F_P^E BEG	VARIABLE PATH PCT.METER, BEG.VAL	9(4)	2
36	C602	IFPFIN	F_P^E FIN	" " " " , FIN. "	9(4)	2
37	C603	IFPMAX	F_P^E MAX	" " " " , MAX. "	9(4)	2
38	C604	IFPMIN	F_P^E MIN	" " " " , MIN. "	9(4)	2
39	C605	NVFP	N_V	" " " " ,NO.MEAS.	9(3)	--
40	C606	IFPS	F_P^S	" " " "SOUTH VAL.	9(4)	2
41	C607	IFPW	F_P^W	" " " "WEST VAL.	9(4)	2
42	C608	IFPN	F_P^N	" " " "NORTH VAL.	9(4)	2

DATA BANK RECORD NO.	DATA BASE COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION	PICTURE	DECODING ALGORITHM
43	C701	IE01	E_0^1	FILTERED EPPLEY DATA, λ = .946	9(4)	1
44	C702	IE02	E_0^2	" " " λ = .4	9(4)	1
45	C703	IE03	E_0^3	" " " λ = .87	9(4)	1
46	C704	IE04	E_0^4	" " " λ =1.06	9(4)	1
47	C705	IE05	E_0^5	" " " λ = .75	9(4)	1
48	C706	IE06	E_0^6	" " " λ = .55	9(4)	1
49	C707	IE07	E_0^7	" " " PHOTOPIC	9(4)	1
50	C708	IE08	E_0^8	" " " λ = .3to3.5	9(4)	1
51	C709	IE09	E_0^9	" " " DIRECT	9(4)	1
52	C710	IE10	E_0^{10}	" " " λ = .3to3.5	9(4)	1
53	C801	IT1	T_1	BARNES TRANSMISSOMETER, 3-5μ BEG	9(4)	1
54	C802	IT2	T_2	" " 8-12 μ	9(4)	1
55	C803	IT3	T_3	" " 8-13 μ	9(4)	1
56	C804	ITX	T_x	" " OPEN OR 4μ	9(4)	1
57	C805	IT8	T_8	" " 3-5μ FIN	9(4)	1
58-1	C1001	IX1	X	AEROSOL INSTRUMENT CODE	9(1)	10
58-2	C1002	IX2	X	AEROSOL DATA CODE	9(1)	11
58-3	C1004	IX3	X	AEROSOL RELIABILITY CODE	9(1)	22
58-4	C1003	IX4	X	MAX. CHANNEL USED	9(1)	12

DATA BANK RECORD NO.	DATA BASE COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION	PICTURE	DECODING ALGORITHM
58-5,59-1	C1101	IX5A1	X,A	AEROSOL DATA, CHANNEL 1	9(2)	13
59-2,3	C1102	IA23	A	" " " 2	9(2)	13
59-4,5	C1103	IA45	A	" " " 3	9(2)	13
60-1,2	C1104	IB12	B	" " " 4	9(2)	13
60-3,4	C1105	IB34	B	" " " 5	9(2)	13
60-5,61-1	C1106	IB5C1	B,C	" " " 6	9(2)	13
61-2,3	C1107	IC23	C	" " " 7	9(2)	13
61-4,5	C1108	IC45	C	" " " 8	9(2)	13
62-1,2	C1109	ID12	D	" " " 9	9(2)	13
62-2,3	C1110	ID23	D	" " " 10	9(2)	13
62-5,63-1	C1111	ID4E1	D,E	" " " 11	9(2)	13
63-2,3	C1112	IE23	E	" " " 12	9(2)	13
63-4,5	C1113	IE45	E	" " " 13	9(2)	13
64-1	C31	IFI	F	CONTEL & ADD. MEAS. CODE	9(1)	14
64-3,4,5	C41	IF345	F	ADD. EXTINCTION VALUE	9(3)	?
65-1	C42	IGI	G	ADD. EXTINCTION RELIABILITY	9(1)	?
65-2,3,4	C51	IG234	G	CO_2 CONCENTRATION VALUE	9(3)	?
65-5	C52	IG5	G	CO_2 CONCENTRATION RELIABILITY	9(1)	?
66-1,2,3	C61	IH123	H	LASER TURBULENCE VALUE	9(3)	?

DATA BANK RECORD NO.	DATA BASE COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION	PICTURE	DECODING ALGORITHM
66-4	C62	IH4	H	LASER TURBULENCE RELIABILITY	9(1)	?
66-6,67-1	C71	IH511	H,I	BARNES INDICATOR VALUE -1	9(2)	?
67-2	C72	II2	I	" " RELIABILITY -1	9(1)	?
67-3,4	C81	II34	I	" " VALUE -2	9(2)	?
67-5	C82	II5	I	" " RELIABILITY -2	9(1)	?
68	C901	IN	N	CLOUD COVER	9(1)	15
69	C902	IDD	DD	WIND DIRECTION AT 10 M	9(2)	16
70	C903	IFF	FF	WIND SPEED AT 10 M	9(2)	17
71	C904	ID2D2	D_2D_2	WIND DIRECTION AT 2 M	9(2)	16
72	C905	IF2F2	F_2F_2	WIND SPEED AT 2 M	9(2)	17
73	C906	IPPP	PPP	PRESSURE	9(3)	18
74	C907	ITTT	TTT	TEMPERATURE	9(3)	19
75	C908	ITDTD	$T_DT_DT_D$	DEW POINT TEMPERATURE	9(3)	19
75	C908	ITDTD	$H_RH_RH_R$	RELATIVE HUMIDITY	9(3)	20
76	C909	IRRATE	RRR-1	RAIN RATE	9(3)	21
77	C910	IE	E	GENERAL GROUND STATE	9(1)	3
78-1	C2101	IQ101	Q	MRI PHOTOPIC QUALITY, 11	9(1)	22
78-2	C2102	IQ102	Q	" " " 12	9(1)	22
78-3	C2103	IQ103	Q	" " " 13	9(1)	22
78-4	C2104	IQ104	Q	" " " 14	9(1)	22

DATA BANK RECORD NO.	DATA BASE COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION	PICTURE	DECODING ALGORITHM
79-1	C2201	IQ201	Q	ELTRO TRANSMISSOMETER QUALITY	16 9(1)	22
79-2	C2202	IQ202	Q	" " "	17 9(1)	22
79-3	C2203	IQ203	Q	" " "	18 9(1)	22
79-4	C2204	IQ204	Q	" " "	19 9(1)	22
80-1	C2301	IQ301	Q	HORIZONTAL LUXMETER QUALITY	21 9(1)	22
80-2	C2302	IQ302	Q	" " "	22 9(1)	22
80-3	C2303	IQ303	Q	" " "	23 9(1)	22
80-4	C2304	IQ304	Q	" " "	24 9(1)	22
80-5	C2401	IQ401	Q	VERTICAL LUXMETER QUALITY	26 9(1)	22
80-6	C2402	IQ402	Q	" " "	27 9(1)	22
80-7	C2403	IQ403	Q	" " "	28 9(1)	22
80-8	C2404	IQ404	Q	" " "	29 9(1)	22
81-1	C2501	IQ501	Q	NIGHT PATH LUMINENCE QUALITY	30 9(1)	22
81-2	C2502	IQ502	Q	" " " "	31 9(1)	22
81-3	C2503	IQ503	Q	" " " "	32 9(1)	22
81-4	C2504	IQ504	Q	" " " "	33 9(1)	22
82-1	C2601	IQ601	Q	VARIABLE PATH FCT. METER QUALITY	35 9(1)	22
82-2	C2602	IQ602	Q	" " " " "	36 9(1)	22
82-3	C2603	IQ603	Q	" " " " "	37 9(1)	22

DATA BANK RECORD NO.	DATA BASE COMPONENT NO.	DATA BASE COMPONENT NAME	DATA ITEM	DATA DESCRIPTION		PICTURE	DECODING ALGORITHM
82-4	C2604	IQ604	Q	VARIABLE PATH FCT. METER QUALITY	38	9(1)	22
82-5	C2605	IQ605	Q	" " " " "	40	9(1)	22
82-6	C2606	IQ606	Q	" " " " "	41	9(1)	22
82-7	C2607	IQ607	Q	" " " " "	42	9(1)	22
83-1	C2701	IQ701	Q	FILTERED EPPLEY DATA QUALITY	43	9(1)	22
83-2	C2702	IQ702	Q	" " " "	44	9(1)	22
83-3	C2703	IQ703	Q	" " " "	45	9(1)	22
83-4	C2704	IQ704	Q	" " " "	46	9(1)	22
83-5	C2705	IQ705	Q	" " " "	47	9(1)	22
83-6	C2706	IQ706	Q	" " " "	48	9(1)	22
83-7	C2707	IQ707	Q	" " " "	49	9(1)	22
83-8	C2708	IQ708	Q	" " " "	50	9(1)	22
83-9	C2709	IQ709	Q	" " " "	51	9(1)	22
83-10	C2710	IQ710	Q	" " " "	52	9(1)	22
84-1	C2801	IQ801	Q	BARNES TRANSMISSOMETER QUALITY	53	9(1)	22
84-2	C2802	IQ802	Q	" " "	54	9(1)	22
84-3	C2803	IQ803	Q	' " "	55	9(1)	22
84-4	C2804	IQ804	Q	" " "	56	9(1)	22
84-5	C2805	IQ805	Q	" " "	57	9(1)	22

FOOTNOTES:

1) $I_1I_2I_3I_4$ = Float $(I_1I_2I_3)$) * (10. ** Float $(I_4 + 6)$)

2) $I_1I_2I_3I_4$ = Float $(I_1I_2I_3)$) * (10. ** Float $(I_4 + 8)$)

3) See Appendix 1 for codes and their significance.

4) I_1I_2 = a) I_1 = Station location code (see appendix 2)

b) I_2 = Issue level of tape

i) $1 \leq I_2 \leq 4$ = Preliminary data

ii) $5 \leq I_2 \leq 9$ = Reviewed data

5) See Appendix 3 for code

6) See Appendix 4 for code

7) Scattering Code: = 1 - Denotes scattering being presented

= 2 - Denotes extinction being presented

= 3 - Denotes both being presented

8) Filter Code: = 4 - 3.4 - 5.5 m

= 5 - Open (2-14 m)

= 6 - 4 calibration filter

= 7 - To be decided

9) Humidity Code: = 1 - Denotes direct humidity value to .1%

10) See Appendix 4 for code

11) See Appendix 5 for code

12) Maximum Channel: = 0 - No data

= (Max. channel - 4) - Data available (N.B. - data is present, at least 5 channels are used).

13) I_1I_2 Float (I_1I_2) (See Appendix 7 for further information)

14) See Appendix 7 for code

15) $I_1 \leq 8 \rightarrow .125 * \text{Float}(I_1)$; $I_1 = 9$ No measurement

16) I_1I_2 Integer value indicating the nearest 10^o value from true North as follows:

-5^o to $5^o \rightarrow 36$

5^o to $15^o \rightarrow 01$

.

.

.

.

345^o to $355^o \rightarrow 35$

Special values: I_1I_2 = 00 Calm

= 99 No observation or failure

17) I_1I_2 Integer value indicating wind speed in M/sec.
Special value: I_1I_2 = 99 No observation or failure.

18) $I_1I_2I_3$ Integer value of pressure in MBAR.
e.g. $I_1I_2I_3$ = 024 1024 MBAR
$I_1I_2I_3$ = 973 .973 MBAR
(N.B. Pressure range is approximately 950-1050)

19) $I_1I_2I_3$ $I_1I_2I_3 < 500$ $0.1 * \text{Float}(I_1I_2I_3)$

$I_1I_2I_3 > 500$ $-0.1 * \text{Float}(I_1I_2I_3 - 500)$

$I_1I_2I_3$ 999 Failure code

20) $I_1I_2I_3$ $0.1 * \text{Float}(I_1I_2I_3)$

21) $I_1I_2I_3$ $0.1 * \text{Float}(I_1I_2I_3)$

Mean rain rate value in MM/HR. The recommended sampling period is 4 minutes. Alternative periods must be noted in the comments.

Special value: $I_1I_2I_3$ = 999 Failure code

22) See Appendix 8 for codes

23) $I_1I_2I_3$ Total rain in the past hr. in MM
Special value: $I_1I_2I_3$ = 999 No observation or failure.

APPENDIX 1 - GROUND STATE CODES

0 = Dry

1 = Wet

2 = Flooded

3 = Frozen

4 = Snow patches

5 = Snow cover or hail cover

6 = Ice or frozen snow cover

7 = Melting snow

8 = Situation outside OPAQUE scheme; see comment list

9 = No observation

APPENDIX 2 - STATION CODES

1 = Canada/Denmark

2 = France

3 = Germany

4 = Italy

5 = Netherlands

6 = UK

7 = USA/Germany

APPENDIX 3 - UPDATE STATUS

NOTE - This item does not appear in the original OPAQUE data bank, but has been included in the data base for convenience. It is meant to serve as an indicator of the status of any given logical entry (hour) in the data base. The following convention will be used:

0 = Raw data from random access data file

1 =

2 =

3 =

4 =

5 =

6 =

7 =

8 =

9 =

APPENDIX 4 - AEROSOL INSTRUMENT CODE

0 = No data collected
1 = Active probe (A)
2 = Classical probe (C)
3 = (A) and (C)
4 = Royco (R)
5 = (A) and (R)
6 = (C) and (R)
7 = (A) and (C) and (R)
8 = Other - see comments

APPENDIX 5 - AEROSOL DATA CODE

0 = Data not reviewed for features
1 = Reasonable approximation
2 = Statistically unreliable
3 = Statistically unreliable at larger sizes - low counts
4 = Single peak distribution
5 = Multiple peak exists

APPENDIX 6 - AEROSOL DATA DERIVATION

The two digits of stored aerosol data are determined by the algorithm:

$$I_1 I_2 = \mathrm{IFIX}(10.*(\log(\frac{d}{d \log \delta}(N)) + 4))$$

where the following constraint applies:

$$0 \leq (\log(\frac{d}{d \log \delta}(N) + 4) \leq 100$$

and where: δ = Particle diameter

N = Particle count/BIN

In 13 channels as indicated below:

0.200	-	0.283
0.283	-	0.400
0.400	-	0.566
0.566	-	0.800
0.800	-	1.131
1.131	-	1.600
1.600	-	2.263
2.263	-	3.200
3.200	-	4.525
4.525	-	6.400
6.400	-	9.051
9.051	-	12.800
12.800	-	18.100

APPENDIX 7 - CONTEL AND ADDITIONAL MEASUREMENT CODE

0 = No measurement

1 = Contel

2 = Sky camera

3 = Contel and sky camera

4 = Additional measurements only as in comments

5 = Contel and additional measurements

6 = Sky camera and additional measurements

7 = Contel and sky camera and additional measurements.

APPENDIX 8 - RELIABILITY CODES

First group, no data available

0 No measurement, as the instrument is not (yet) working (min. period a week, for shorter time use 1).

1 No measurement in this measurement series or may be in the neighboring series (max. some weeks, for longer time use 0)
or
signal outside the range covered by the instrument or signal destroyed by other influence (e.g. by looking direct in the direction of the sun).

Second group, the relative error of the data is below 50% (Barnes 15% abs.), i.e. not good for relationships but useful for statistics.

2 Quality below normal standards is assumed, as the data has not been checked.

3 Quality below normal standard, caused by measurement circumstances (as sun facing the instrument or fluctuating light) or known by detailed knowledge of the specific instrument (e.g. noise at low signal levels), or proven by bad calibration results.

Third group, the relative error of the data is below 15% (Barnes 5% abs.), i.e. useful for all normal purposes. All data are checked for bad circumstances as e.g. sun problems.

4 Quality normal, assumed by detailed knowledge, even though it is some time since calibrations have been performed or proven by frequent, agreeing calibrations.

Fourth group, the relative error of the data is below 5% (Barnes 2% abs.), i.e. they are especially useful for demanding functional correlation relationships.

5 Quality above normal standard, based on knowledge of instrumental performance and information contained in the basic data set (e.g. absence of veiling glare in the path luminence meters from sun close to optical axis, as derived from the sun position and a threshold angle equivalent to a 5% effect with respect to effects of artificial light sources (see 6).

6 Quality above normal standard, as 5, but also supported, for specific interpretations, by information in addition to that contained within the basic data set:

a) in case of Epply measurements: without any clouds obscuring the sun,

b) in case of Barnes measurements: during rain whether indicated or not by rain rate meters,

c) in case of extinction coefficient measurements: with values at homogeneous conditions,

d) in case of path luminance during night: without any occasional artificial light sources affecting the measurement.

REDUCTION AND ANALYSIS OF DATA FROM THE DEFENSE METEOROLOGICAL SATELLITE PROGRAM

The DMSP program is the successor to the Air Force Defense Systems Application program for performing meteorological satellite imagery and to the Data Acquisition and Processing Program. DMSP is a total program involving sensors, data communications and ground processing equipment. Some awareness of all of these systems is necessary to understand and apply DMSP data. The purposes of the DMSP are threefold:

1) provide globally recorded visual and infrared cloud cover and other specialized environmental data,
2) provide real time direct readout of local area environmental data to mobile receiving terminals throughtout the world, and
3) continue advancement of environmental satellite technology.

Basic interpretation techniques are valid for both conventional satellite data and DMSP data. But, DMSP data have several unique characteristics in addition to a very sharp resolution capability.

The primary sensors used to obtain DMSP data are contained in an aerospace vehicle electronics package which is mounted on the spacecraft so that, with nominal attitude control, the sensors are always oriented toward the earth. These sensors are scanning radiometers which respond to amounts of radiation within specific spectral ranges. Two types of data are obtained: visual data, which is a measure of reflected solar radiation; and infrared data, which is a measure of emitted earth and cloud radiation.

The spectral interval used for obtaining visual data is 0.4-1.1 micrometers, i.e. the sensor reacts as there was a negative blue filter blocking most of the response in the blue end of the range. This minimizes the blue light backscatter of the atmosphere. The response at the other end of the spectral range is a result of the type of sensor used. Its natural response provides sensing into the near infrared, but radiation sensed is still primarily reflected solar radiation.

A thermister bolometer is used to sense infrared radiation in the 8-13 micrometer range. This range is used primarily because it contains the peak radiations emitted by the earth and its atmosphere. The broad spectral range was selected so that the field of view of the detector could be narrowed to obtain good spacial resolution while still maintaining a good signal to noise ratio at the detector.

The scanning mirror in the high resolution, infrared mode (HR/MI) rotates at 1/3 the speed of the VHR/WHR mirror, so only 1/6 the number of scan lines are produced. The HR/MI scan rate produces a 2 nautical mile resolution of visual data and a 2.4 nm resolution of infrared data. Obtaining proper spacial resolution requires that the VHR and WHR detectors be smaller than their HR and MI detectors. Consequently, the VHR data are available only during daytime when there is sufficient reflected solar illumination of the earth scene to provide adequate S/N ratio at the detector. For infrared data there is little difference between the total amount of radiation emitted from the earth during day or night. However, a greater S/N ratio problem exists for this data than for other detectors. The face on the spacecraft is earth oriented and therefore the sensors on the right are earth oriented and therefore the sensors on the right are earth oriented. The top of the spacecraft is always oriented toward the sun. In this orbit there is a possibility, near the poleward portion of each revolution, of direct sunlight impunging on the sensors.

Some zero resolution, special purpose, radiometer measures the amount of sunlight incident on the spacecraft and, with the earth oriented sensor system, measures the solar illumination of the earth scene. This output in conjunction with a programmable gain memory unit, controls the signal level output for the visual data are normalized on the spacecraft for solar illumination; this is an extremely important and unique feature in data production capabilities of DMSP.

There are four supplementary sensors which go on the DMSP spacecraft. Normally each package contains two supplementary sensors, the exact complement differing according to the time period the orbit is planned for, the condition of the sensors on older spacecraft, etc. One of these, supplementary sensor J (SSJ) is an electron spectrograph with one fixed channel and one stepping channel. The channels detect energetic electrons over ranges of energies associated with visible aurora. The fixed channel is 6 KeV and the stepping channel cycles through eight energy thresholds: 54, 98, 219, 600, 1400, 3540, 8200 and 1970 eV. The data sample is taken approximately every second and the field of view is 3 degrees by 12 degrees. SSJ has been improved; the second generation is called SSJ/2. It consists of a single stepping channel with six energy ranges with nominal energy steps 0.3, 0.68, 1.6, 3.5, 7.9, and 18 KeV. The sampling rate is 0.0922 seconds per energy step and the field of view is a 30 degree anti-earth cone.

The J/3 is such an instrument which is an electron spectrometer designed for SAMSO. The instrument consists of two electrostatic analyzers which measure the spectrum of precipitating electrons that are responsible for the aurora. The data from J/3 will be used to determine the location of the boundaries of the auroral oval and will be used to correlate electron fluxes with auroral observations. During the design of the instrument, a computer program was written to trace the particle orbits from the entrance slits to the detectors in an effort to determine the response and detection efficiencies of the instrument as closely as possible. Tests indicate the calibration is better than was anticipated. An additional task that the data will be used for is to correlate electron precipitation with substorm phenomena at synchronous altitudes along the same magnetic field line and for predicting total electron content of the ionosphere.

Four types of data are produced by the spacecraft's primary sensors; however, it is not possible for a tactical readout site to receive all of them. The DMSP transmitter operates at 512 kilobits per second (kbs). The normal operating mode is for VHR transmission on the ascending portion of an orbit and WHR on the descending orbit. Both HR and MI data are transmitted to all receiving sites. If the spacecraft is placed in an early mording ascending (early evening descending) orbit where there is an earth terminator problem in the visual scene, the VHR data are transmitted on the ascending portion of the orbit near the equator where there is sufficient imagery. When the spacecraft progresses poleward far enough so that the edge of the VHR scan line enters the terminator, the transmission is switched to WHR.

a. Visual Data The two types of visual data, VHR and HR, complement each other for several meteorological applications. The VHR is designed to examine weather systems in detail such that not only is the overall synoptic pattern readily identifiable, but the details of the cloud patterns and their structure are observed. Since HR data has a lower data volume rate than VHR, it is preferred for computer processing.

b. Infrared Data The two types of infrared data are MI and WHR. The infrared sensors are co-located with visual sensors. Common optics insure that the visual and infrared scene are viewed simultaneously and in the same perspective. This arrangement allows direct comparison of the visual and infrared data, and permits the meteorologist to envision the three dimensional aspects of the cloud scene generated by atmospheric motions.

The MI and WHR sensors were developed to serve several different purposes for the meteorological analyst. MI data, with its 2.4 nm spatial resolution at subpoint, is used as the primary infrared data for comparison with all visual data. It is also the more accurate infrared data for use in quantitative thermal determinations.

AD-A102 685 BEDFORD RESEARCH ASSOCIATES MA F/G 4/1

APPLICATION OF METHODS OF NUMERICAL ANALYSIS TO PHYSICAL AND EN--ETC(U)

OCT 80 R BOUCHER, T COSTELLO, P MEEHAN F19628-78-C-0241

UNCLASSIFIED AFGL-TR-80-0347 NL

2 OF 2

AD
A102685

END

DATE
FILMED

9-81

DTIC

Data The DMSP sensors were designed to provide high quality data. The system approach was taken to ensure that the resolution inherent in the basic sensor electronic signal would be faithfully preserved through amplification, recording, transmission, and eventual display on the ground. The sensors were engineered to provide subpoint spatial resolution of 1/3 NM for VHR and WHR data, 2 NM for HR data, and 2.4 NM for MI data.

All these data are rectified by the Signal Processor No. 1 but the rectification process does not improve spatial resolution.

As a general rule of thumb, VHR imagery displayed in a 1:15 million scale has a spatial resolution of one-third to one-half nm within the first inch of data either side of subpoint. From one inch and two and one-half inches either side of the subpoint, the spatial resolution degrades from one-half to one nm. From the two and one-half inch point to data edge, the spatial resolution further degrades to two nm. These degradations are due to foreshortening.

Spatial resolution is more difficult to quantitatively determine for infrared sensors than it is for visual sensors. In the infrared, the spatial resolving power of the sensor can be masked by the thermal accuracy of the sensing system.

The error between actual temperature and effective radiating temperature was observed to be related to the moisture distribution. The regression equation describing it is:

$$\text{Error} \quad (T_a - \text{ERT}) = 35.97 + 30.98 \text{ Log } (\bar{\bar{w}}/\cos\beta)$$

where: $\bar{\bar{w}}/\text{Cos}\beta$ represents a line of sight measurement through the moisture layers.

A study on absolute thermal accuracy was performed. Radiosonde data was used from within six hours of a satellite pass. A correlation was derived between radiosonde "observed" cloud top temperatures and the effective radiating temperature indicated by MI data.

It is apparent that atmospheric attenuation factors are different for each site depending on its location and season. Many sites have a continuing evaluation program so that accurate temperatures can be obtained for mission support. To assist those who do not have access to site peculiar calibration data, theoretical atmospheric attenuation curves for a standard atmosphere have been developed for the MI sensor.

$$ERT_{(corr)} = ERT + 0.0125\ (ERT - 220) + 0.00156\ (ERT - 220)^2$$

Orbit Characteristics The ability to use DMSP data is greatly enhanced by unique data location procedures. Data location, or gridding of data, at the very best, is a complex science. An understanding is necessary of the sensing mechanism, spacecraft orbit, and display equipment. The scanners are essentially looking at a curved surface from a fixed point in space. The resulting effect is termed foreshortening. Not only are these data foreshortened because of "looking" over the curvature of the earth, but the spacecraft travels in an orbit such that the longitude and latitude lines are curved and their orientation changes depending on the subpoint latitude.

The nominal orbit for the DMSP spacecraft is a 450 nm circular orbit with an inclination angle of 98.7 (exactly 98.747). This means the orbital plane is inclined 98.7 to the equatorial plane where the spacecraft crosses the equator northbound (ascending node). The inclination angle was selected to insure that the nominal 450 nm circular orbit is sun-synchronous (maintains a relatively constant relationship to the sun so that the ascending node remains at a constant solar time). The combination of a 450 nm circular orbit and 98.7 inclination angle means that the satellite's orbital plane rotates slowly around the earth at the same rate and direction that the earth rotates around the sun. At the most poleward positions of the spacecraft, its motion is tangent to the parallels of latitude, adding to the complexity of a grid's latitude/longitude configuration.

The nodal period of this sun-synchronous orbit is 101.56 minutes. During the time of one revolution, the earth revolves under the orbit to the east and the orbital plane precesses slightly to the east. The result is that each nodal crossing is approximately 25.4 west of the previous crossing. Approximately every five days there is a 15 revolution day; therefore, a space-craft peculiar cycle is established which causes both daily data coverage and local time of data variations.

Once DMSP data are acquired, only the longitude and nodal crossing and the elapsed time since the nodal crossing are needed in order to accurately locate these data.

At AFGWC a continuous sinusoidal correction is applied to the imagery. A tactical site usually sets an altitude setting which is assumed to be valid for the few minutes the spacecraft is within readout range of the site.

Data is received from various satellites as they orbit the earth. This data is archived onto tape by the World Data Center A for Solar-Terrestrial Physics in Boulder, Colorado. For the past eighteen months and on a continuing basis they have been and are sending tapes containing the data collected from two satellites F2 and F4. Various other satellites will be sent up in the future and their data in turn will be processed in the same manner as F2 and F4, (though at a higher data rate).

One magnetic tape from Boulder contains approximately ten days worth of data. If there are two satellites (as at present) on a tape each tape contains five calendar days worth of data. After considerable effort on our part to make the existing processing programs as efficient and compact as possible, it still takes approximately 2000 CP seconds computer time to set up the present data base and produce the microfiche flux plotting for five calendar days from one satellite. This data base consists of one tape per 15 calendar days of packed channel counts for one satellite and one tape per month of flux data. Already under this present system approximately two years worth of data of one satellite and six months of two satellites have been processed and we now have a data base of approximately eighty tapes. Some of the tapes are labeled and some not. We have found several tapes mislabeled and others with data missing. Also this processing will expand to accommodate three, four or more satellites sent up in the future.

The preceeding describes the software required in setting up a CDC data base. This consists of only a part of the total project. On a continuous day to day basis, we receive several requests to run several of the existing application programs which keep a programmer quite busy. In addition it involves the running of the statistical survey program whose input is all of the data base flux tapes (approximately 40 tapes).

Through meticulous effort put into packing the data and by use of high density tapes, we feel we can put about six to eight months worth of data onto one tape. The old system would require 18 to 24 tapes. The production runs will be reduced to a fraction of what they are now (less than 500 seconds). We intend to set up a partly interactive system whereby an operator at an Intercom-terminal will type in the date of data desired, the time intervals, and application program desired and a batch job would be generated to run the request. After logging in the operator would be prompted by clear and straightforward messages. Any of the countless day-to-day requests could be generated by the scientists or initiators themselves even if they have little or no programming experience. We would then have more time to develop new software which could be very easily added into the existing software.

REDUCTION AND ANALYSIS OF HIGH ENERGY ELECTRON DATA

DMSP PROGRAM

1. Normal production of microfiche and archived output tapes.

2. Production of particle spectrum and paper plots for publication.

3. Dump of the edge of the auroral oval.

4. Plar maps of auroral oval.

5. Replotting microfiche with abnormal problems.

6. Testing method for evaluating equatorial edge of oval.

7. Checking particle data against data from existing DMSP tapes.

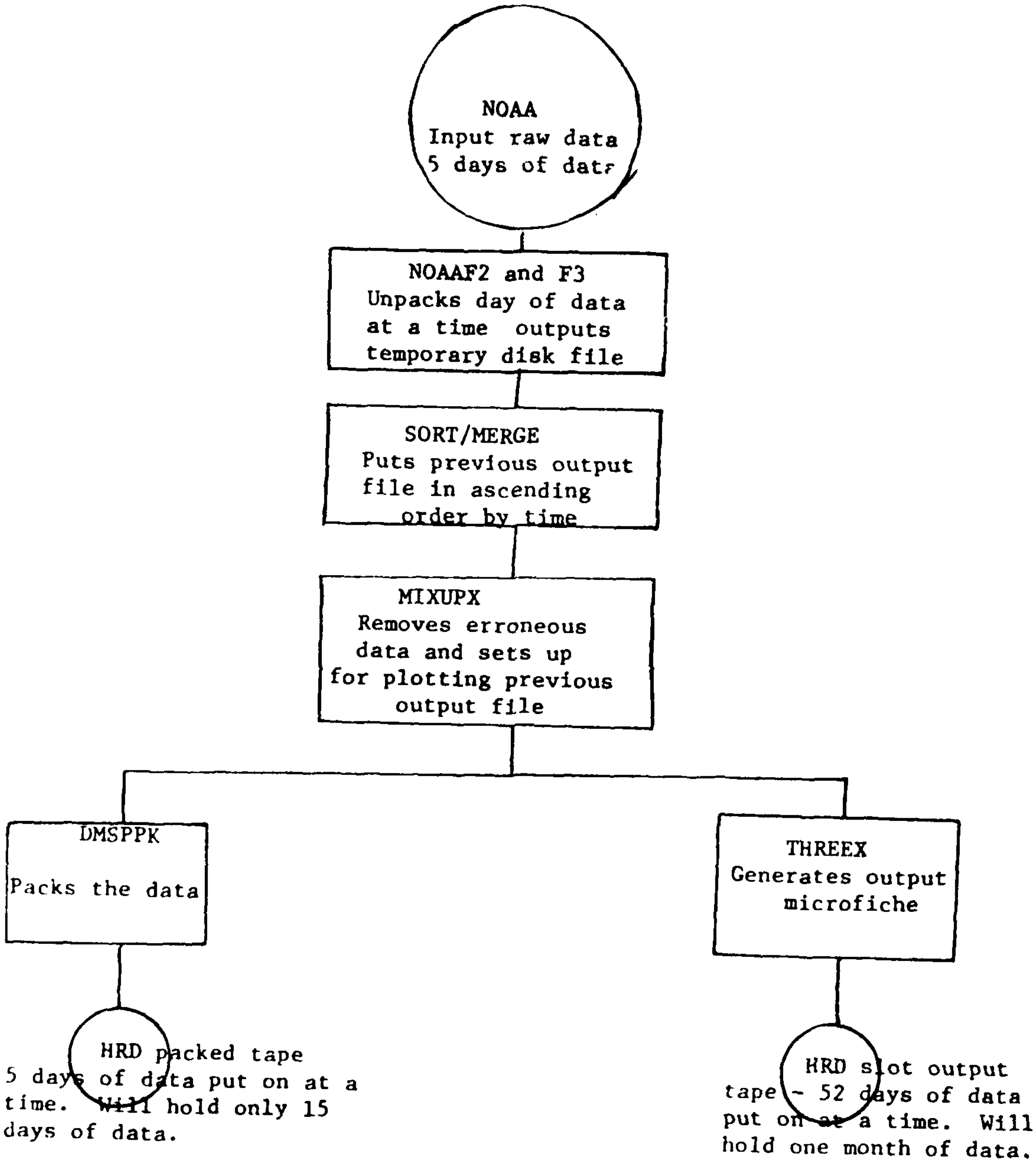
SUMMARY

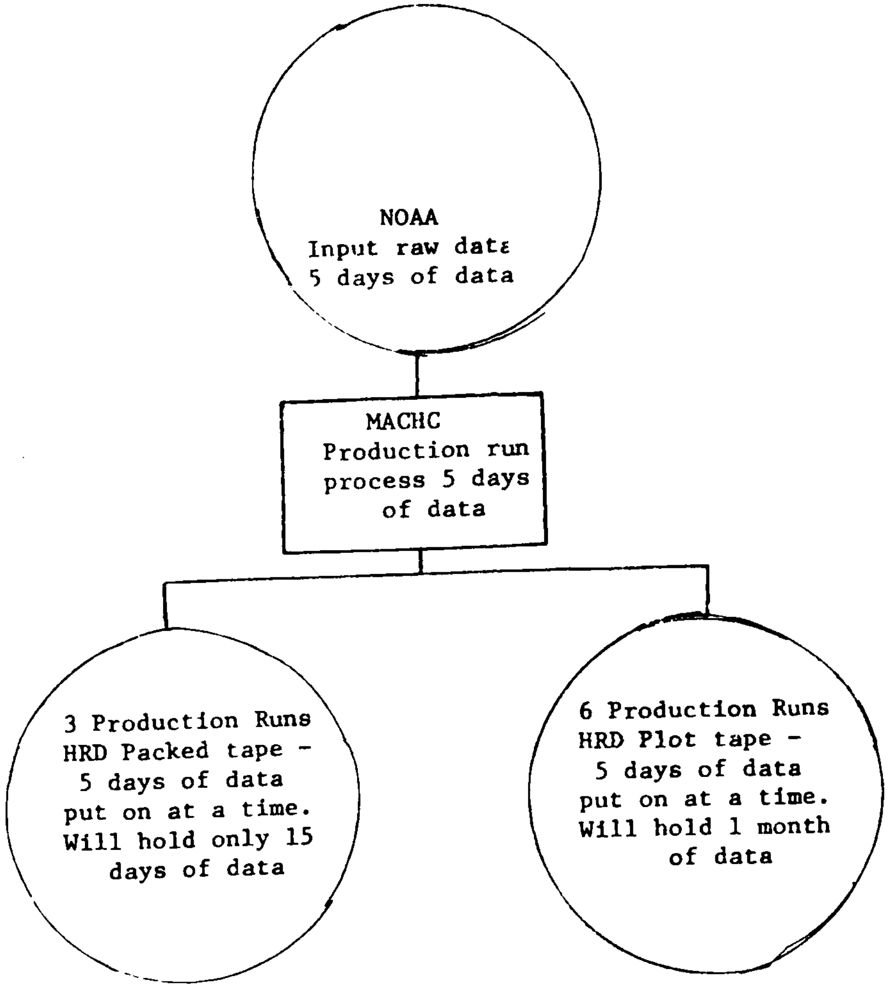
1. Tape Output

Files are accumulated on tapes up to six production runs on one output tape. Thus, one production run can affect others. Sometimes jobs abort through mis-keypunching, parity, tape breaking, system problems, etc. and thus garbage data and half files are put on tape. The further runs have to read data on tape before adding data and runs are not independent. The ideal is to get consistently one run a night for each calendar months processing. Large chunks of time of say 5, 6 or 8 hours are not too beneficial as it now stands.

2. Disk Output - Achieved to Tape

This method is more appropriate if we are getting discrete chunks of say 5, 6, or 8 hours of computer time. Runs are completely independent. Output files are catalogued on disk. However this involves the extra step of achieving disk files to tape then purging the disk file. About four production runs could be put on a disk pack. To make this feasible we would need another disk pack.


NOAA
Input raw data
5 days of data
NOAAF2 and F3
Unpacks day of data
at a time outputs
temporary disk file
SORT/MERGE
Puts previous output
file in ascending
order by time
MIXUPX
Removes erroneous
data and sets up
for plotting previous
output file
DMSPPK
Packs the data
THREEX
Generates output
microfiche
HRD packed tape
5 days of data put on at a
time. Will hold only 15
days of data.
HRD slot output
tape - 52 days of data
put on at a time. Will
hold one month of data.


NOAA
Input raw data
5 days of data
MACHC
Production run
process 5 days
of data
3 Production Runs
HRD Packed tape -
5 days of data
put on at a time.
Will hold only 15
days of data
6 Production Runs
HRD Plot tape -
5 days of data
put on at a time.
Will hold 1 month
of data

Typical DMSP Activities

- Generate high latitude maps of the electron precipitation patterns based upon Kp.

- Using the SSF/3 data determine the relationship between Kp, the equatorward boundary of the auroral oval, and magnetic local time. This requires the determination of the equatorward boundary for several thousand orbits of SSJ/3 data.

- The project is attempting to find a way to rapidly determine the equatorward edge of the oval and to specify its position. The work already done on the large statistical program has shown that a fully automated determination of the boundary is not possible. Hence, some interactive capability perhaps using the Tektronix graphics is required.

- Determine an algorithm to be used in real time at AWS for giving the approximate position of the equatorward boundary. This project requires the study of a number of auroral passes of SSJ/3 and come up with several possible algorithm.

- Over the long run AWS would like to see the global statistical study redone based on different energy bins. Now, in the global study the integral flux, energy flux and average energy are the three quantities studied. For the follow on study, they would like to see the same quantities studied but to have them calculated separately for the energy ranges causing the F & E regions in the ionosphere respectively. In addition, they would like to see the global maps combined with a simple code to actually predict the E region profile.

Finally, AWS would like AFGL to produce an auroral activity index using the equatorward boundary of the oval. This would require that for every pass of the SSJ/3 data the boundary be determined.

Many of these tasks have been completed. The statistical program is finished. The development of the plotting programs to display the data must still be completed.

The algorithm to find the equatorward boundary has also been successfully tested. This work is being done to support the Air Weather Service. Since Air Weather Service plans to use the algorithm, the programming had to be fully documented. This will be incorporated in an in-house report.

The printout program is also working. The test we use to determine where to printout the data is not as accurate as the algorithm for determining the equatorward boundary. This program was modified to use the algorithm to determine the equatorward edge and to printout the data for 1 minute before and after this point.

Additional work falls into two categories. The first category includes projects to modify and extend the global statistics work. The second category includes projects to analyze the boundary determinations we have made using the SSJ/3 data.

The first project in extending the global statistics work is to change the binning size. At present the bins are one half hour wide in magnetic local time by 2 degrees in magnetic latitude. This means that as a function of increasing latitude, the area of the bins decrease and the total number of points averaged in each bin is not the same. To partially eliminate this problem, the bins should vary in latitudinal size as follows:

50-60 degrees	1/2 degree magnetic latitude x 1/2 hr MLT
60-70 degrees	1 degree magnetic latitude x 1/2 hr MLT
70-80 degrees	2 degrees magnetic latitude x 1/2 hr MLT
80-90 degrees	2.5 degrees magnetic latitude x 1/2 hr MLT

The averaging interval should also be decreased from 10 seconds to 4 seconds. This will require generating a tape of 4 second averages. The program to ignore bad data will have to be modified to run off such a tape. We will wish to redo the statistical study using two years of data with the new binning and the new averaging intervals.

The second new project involving global statistics involves binning the data according to the orientation of the Interplanetary Magnetic Field. We have 6 months of data on the IMF. The data exist as 1 hour averages available on tape and listing. Because of the crude time averages of the IMF we have, only the original average sizes and bin sizes need to be used. The programming will involve the capability to determine for each ten second average the orientation of the IMF for a period 1 hour before the time of the ten second average. As a first division the data will be binned as follows:

$B_y > 0$ and $B_z > 0$
$B_y > 0$ and $B_z < 0$
$B_y < 0$ and $B_z > 0$
$B_y < 0$ and $B_z < 0$
$B_z > 0$ irrespective of B_y
$B_z < 0$ irrespective of B_y

Finer division according to the actual magnitude of B_z and B_y will be an additional step after this initial work is completed.

As a parallel job we have been determining the equatorward boundary of the auroral zone from the survey microfiche. Approximately 5000 to 6000 boundaries have been calculated and typed onto computer cards. Each card includes the boundaries on each side of the oval for each pass over the north or south pole (magnetic latitude and magnetic local time). They include also the day and year of each crossing, the approximate time in the day in Universal Time seconds, and the Kp value for the three hour interval in which measurement was made.

We anticipated first analyzing the manner in which the equatorward boundary varies in magnetic latitude and local time as a function of the geomagnetic activity index Kp. Specifically, we designed programs with the capability to read each card, determine in what hour of magnetic local time the boundaries occur and for each whole hour in magnetic local time to bin the magnetic latitude of the boundary according to each level in Kp. We have calculated approximately twnety boundaries for each level of Kp (0, 0+, 1-.........etc.) in each hour of local time to which the satellite orvit gave access. For each level of Kp in a magnetic local time zone the program calculates the average and standard deviation of the magnetic latitude of the boundaries. The program also calculates using all of the boundary determinations in a local time zone, the least squares straight line fit to the data and the linear correlation coefficient.

Secondly, we anticipated analyzing the manner in which the morning equatorward boundary varies as a function of the evening equatorward boundary. Each computer card has an evening and morning equatorward value for a pass over the north or south pole. What we require are programs that will read the cards and, for specific magnetic local time bins on both the evening and morning sides of the oval, will plot the magnetic latitude of the evening side boundary versus the magnetic latitude of the morning side boundary. For example, we may need to look at those orbits that pass through the magnetic local time zone 1800 to 1900 on the evening side and 0500 to 0600 on the morning side. For each orbit, so identified, we wish to plot the latitude of evening boundary vs the latitude of the morning boundary. For each pair of local times selected all the points should be plotted and the data least square fitted and the linear correlation coefficient calculated.

The last project is a study of the manner in which the equatorward boundary varies as a function of the orientation of the IMF. As in the other statistical study using the IMF, this project requires the ability to read the cards of the IMF. For each one hour magnetic local time zone, the program must determine the manner in which the equatorward boundary varies as a function of the B_y and B_z components of the IMF according to the following separations.

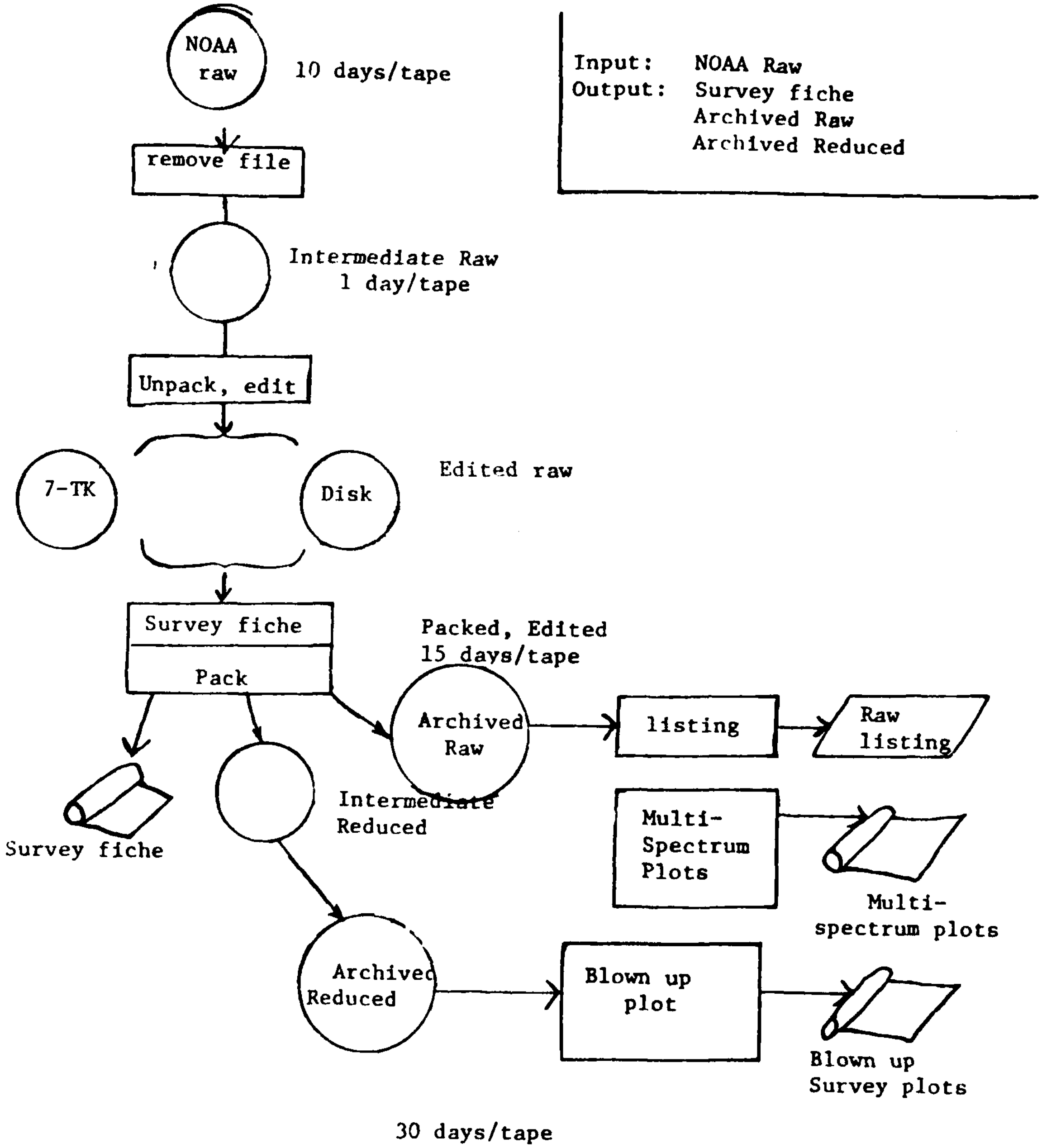
a. B_y in 1 gamma bins from +15 gamma to -15 gamma for $B_z > 0$,

b. B_y in 1 gamma bins from +15 gamma to -15 gamma for $B_z < 0$.

c. B_z in 1 gamma bins from +10 gamma to -10 gamma for all B_y.

For each bin in B_y or B_z, the mean and standard deviation should be calculated and using all the data points the least square fit and linear correlation coefficient should be calculated. As in the other study with IMF, the 1 hour average should be used that preceeds by 1 hour the time at which a given boundary determination was made. The study also might be later expanded to include separations according to the solar wind plasma velocity.

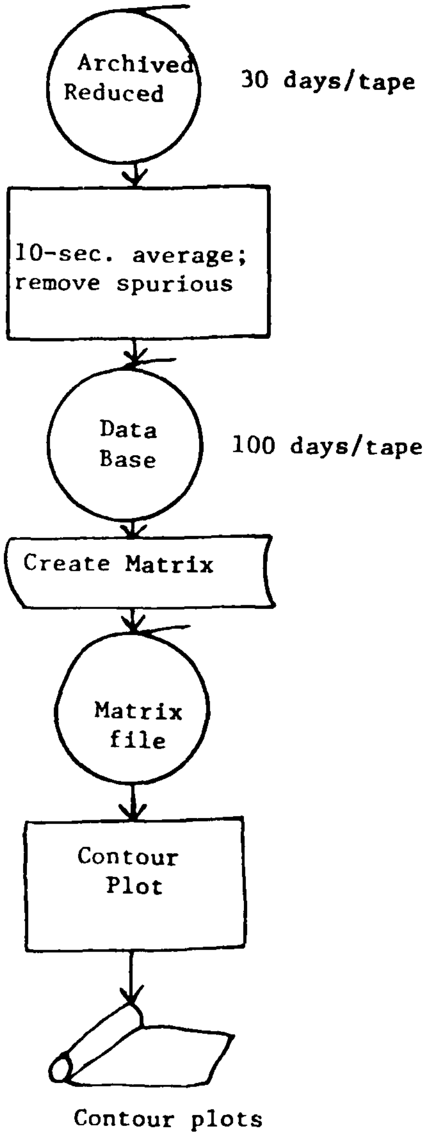
DMSP J/3 Processing: Normal Production and Associated Tasks



DMSP J/3 Processing: Additional tasks

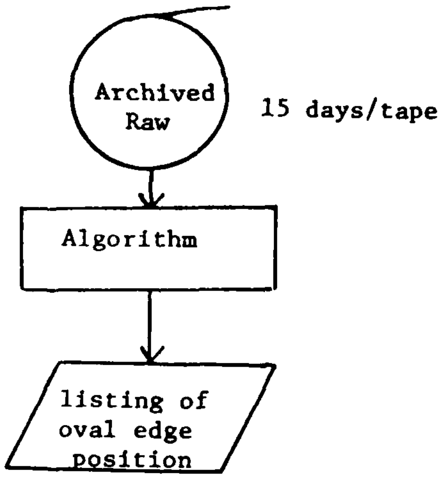
Priority 1

Electron Precipitation high latitude maps:



Priority 2

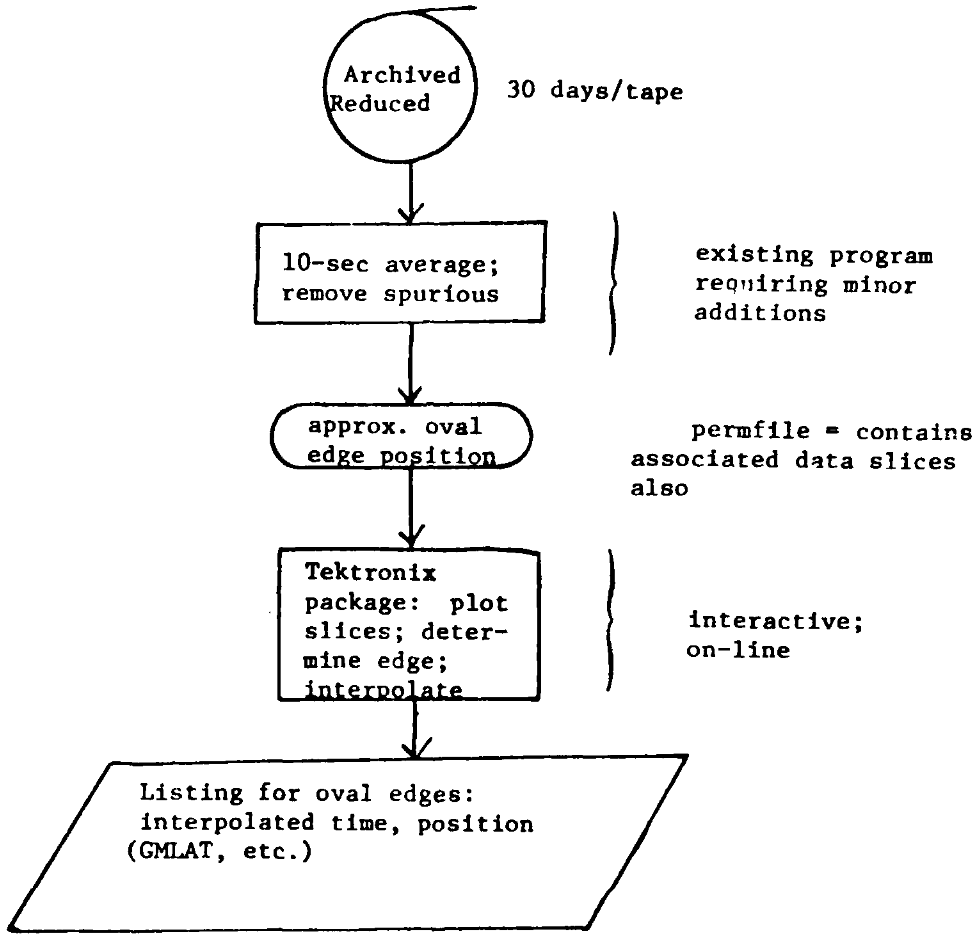
Real time algorithm:



DMSP J/3 Processing: Additional Tasks (continued)

Priority 3

Auroral Oval Determination



DMSP Stacked Data Tape (DSDT) Format

I. General Format

The DSDT tape contains a distillation of the original data tapes written by the Digital Data Formatting System (DDFS). Bad data has been eliminated, and the data has been separated out and calibrated.

The DSDT tapes are written at 556 BPI, in odd parity, on a 7 track drive.

The DSDT tapes contain a series of records making up on giant file containing all of the data from a single pass. (This file may span 2 tapes). Each record contains one data line. Three types of records (data lines) appear: V Data, I Data, and T Data. T Data being calibrated I Data. Generally, but not necessarily, due to noise problems in the original data which may cause loss of one or more data lines, 6 V records will appear for each I and T record pair. (A T record always follows each I record).

Each record consists of an identifier portion and a data portion. The identifier portion contains twnety identifier parameters. The first parameter is a record (data) type discriminator, 4 BCD characters, VVVV for V Data, IIII for I Data, and TTTT for T Data. (Each BCD character - 6 BITS).

The remaining nineteen identifier parameters are 24 BIT binary integers (4 tape characters each).

The data values are either 6 BIT digital count values (V Data and I Data) or 24 BIT integers (T Data) as described below.

II. V Data T = 63
I = 31
V = 65

Identifier Parameters

1. VVVV
2. V line number - counts V lines written on DSDT tape.
3. Pass number
4. DMSP satellite number
5. Data day
6. Data month
7. Data year
8. First line time, tenths of seconds after ascending node
9. Current line time, tenths of seconds after ascending node (counting up from first line time described above).
10. DZ - zero resolution sensor output (0-63).
11. DR - dark reference value (0-63).
12. PSGC - preset gain code (0-63) (not yet used).
13. Bad sample count - number of samples set to zero due to problems with the data.
14. Subpoint sample number (not yet used).
15. Satellite height, N. miles (not yet used).

16-20, Not yet used.

Following the 20th identifier parameter are 3400 V Data values, each a 6 BIT binary integer (tape character). A zero corresponds to black, 63 to white. The data includes the complete earth scan.

III. I Data

Identifier Parameters

1. IIII
2. I line number - counts I lines written on DSDT tapes.

3-9. Same as V Format

10. Thermistor output count (0-63).
11. Calibration source IR count (0-63).
12. Free space reference count (0-63).
13. 31 OK reference count (0-63).
14. 21 OK reference count (0-63).
15. First horizon sample number.
16. Second horizon sample number.
17. DM data word - not yet used.
18. Satellite height, N. miles - not yet used.

19-20. Not used.

Following the 20th identifier parameter are 700 T Data values. Each is a 6 BIT binary integer (tape character). 0 represents the warmest value and 63 the coldest. The data includes the complete earth scan.

IV. T Data

Identifier Parameters

1. TTTT

2-20. Same as I Data

Following the 20th identifier parameter are 700 calibrated temperature values corresponding exactly to the 700 I Data values in the preceeding I Data record. Each temperature value is a 24 BIT binary integer (4 tape characters). The values are in tenths of a degree celsius. Negative values follow the convention by which 77777776_8 is $^{-}1$, 77777775_8 is $^{-}2$, etc., and 00000000_8 or 77777777_8 is 0.

DATA FORMATTER OUTPUT FORMATS FOR DMSP BLOCK 5C PRIMARY DATA

I. FORMATS

1. Data is arranged in 36 binary bit words on seven track tapes with odd parity and have a density of 800 bits per inch.

2. The Data Formatter output, or (DF) data tapes are generated in an IBM-compatible format. Each DF tape contains a header record, N number of data records, and an End-of-File, followed by another header record. When data runs require two DF tapes, the first tape ends with two End-of-Files, and the second data tape eands with an End-of-File followed by the header record. The header record contains 27 words and provides unique identification for each data run.

3. Block 5C primary data can be in either of two different formats: Analog or Digital. Column 7 of the header record contains the control that designates either Analog or Digital data: An "A" in this column designates Analog: a Blank designates Digital.

a. In the "Analog Mode", an analog sub-carrier is demodulated and digitized by the DF;

b. In the "Digital Mode", the analog subcarrier is processed by a Block V Formatter (BVF) at the Command Readout Station (CRS) and sent to AFGWC via an American Satellite Communications link as Digital Data, where the DF then formats the data for processing. The analog video base bands may be digitized by either the DF in Analog Mode or a BVF in Digital Mode.

II. ANALOG MODE

1. In the Analog Mode, H/I data is formatted into 3840 (decimal) word records for processing. The output formats for this Analog Mode data are shown in Figures 5 and 6. It should be noted that the 3840 word record length provides for an integral number of H and I scan lines in each record (no partial scan lines). The first word found in each record will be a

"frame/line counter" and the last word will always be an "assembled time-word".

2. Each scan line is provided with a unique frame/line counter to designate the beginning of the scan line. As part of this frame/line counter, the most significant bit provides an indicator flag for the type of data to follow. A '1' in this position indicates <u>visible</u> data (H or V): a '0' indicates <u>infrared</u> data (I or W).

FRAME/LINE COUNTER	CALIBRATION DATA	II/I DATA	CALIBRATION DATA	BLANKS	TIME WORDS

"ANALOG" SCAN LINE FORMAT

3. The calibration slots are digitized and depicted by several calibration words. These calibration words are repeats of the data in the calibration slots; thus, if the time code slot contained a '1' bit, the time code words would contain 77 (octal). In the "Digital Mode" these calibration slots have been assembled into a single calibration word for the several calibration parameters.

4. The H and I data may be in line pairs; that is, a visible scan line followed by an infrared scan line. Or the data may be H only or I only, but the same record length is still used. The most significant bit in the frame/line counter is the key to identifying the data.

5. At the end of each scan line there is an "assembled spacecraft time-word". This time-word is updated onboard the spacecraft once every two minutes real-time. Therefore, this time-word (in the H and I data) will change approximately once every 214 scan lines. However, if an error occurs during the assembly of a time-word, the previous value will be held and the erroneous time-word discarded until the next two-minute update occurs. Thus, it is possible to have more than 214 scan lines between time-words (a multiple of 214), however, in most cases, there will be

about 214 scan lines. In the V/W mode, the time-word changes approximately every 1284 scan lines.

6. The V/W data is similar to the H/I, except the record length is 3222 words (decimal), and the data is collected so that blocks of data are processed on the tape. Due to the high density of the V/W data and transfer rate constraints of the tape control unit, V/W data is processed in one of two different modes: Only every other scan line is processed; or only one scan line in four is processed (as designated by the DF set-up card, Column 10.

It is possible for a switch from visible to IR data, or IR to visible data, to occur within any given data run (watch for the significant bit in the Frame/Line Counter).

III. DIGITAL MODE

1. The digital data is processed in a similar manner, except the calibration information has been placed in one word at the end of the data stream. Thus, the total length of each scan line has been reduced by removing the calibration overhead. The exact record lengths for processing the digital H/I and V/W data have not been selected; however, if a digital tape header record is examined, Columns 2 through 6 contain the number of words in each record will be adjusted to provide an integral number of scan lines per record.

FRAME/LINE COUNTER	DATA	CALIBRATION WORD	TIME WORD

"DIGITAL" SCAN LINE FORMAT

IV. THE DATA

1. An individual data sample consists of raw radiances as detected by the onboar sensor in six-bit binary representation of the radiance value, where zero is dark and 77 (octal) is maximum brightness in visible data. The IR data sample is similar; however, may seem reversed in a depiction. The zero is 210 degrees Kelvin and 77 (octal is 310 degrees Kelvin.

V. UNIVAC TAPE FORMATS

1. All records are separated by an inner record gap referred to as an End-Of-Record (EOR) mark. The EOR is a s/4 inch blank gap on the tape. When an I/O request is made to a tape unit, the entire data record (EOR to EOR) is read/written by the tape unit. If the user requests only 1 word of a 2000 word record, 1 word will be transferred to his input buffer but the tape will advance to the beginning of the next record, i.e. the remaining 1999 words of the record will be dropped.

2. All files on a tape are separated by End-Of-File (EOF) marks. The EOF is a hardware mark written on the tape by the tape drive. All tape displays use at a minimum 2 files. The first file contains descriptive information and data records. The second file usually contains an End-Of-Display sentinel. On occasion, it may contain a continuation sentinel which indicates that the display is continued on another tape. The data file contains integral numbers of scan lines which are used to build a picture a line at a time. The number of scan lines per record is constant for any given display, but may vary from display to display because of the different map scales employed.

VI. DATA FORMATTER TAPES

Tapes produced by the Harris Corporation Data Formatter (DF) contain the following information:

1. <u>Record 1 (DF Header)</u> The first record contains 28 words in column binary format. A breakdown of column binary is given in attachment 1. The 8 column DF card input from which the 28 word record is derived is given in attachment 2. This card is read by the DF in column binary and written to record number 1 of the tape. The processing parameters specified are used to direct both the data formatter and subsequent computer processing on a larger mainframe. Columns 2-6 of the card specify the size of data records desired. This entry is a 5 digit octal number.

2. <u>Record 2 through N (DATA NNNNN WORDS)</u> All subsequent records on the DF tape contain NNNNN data words where NNNNN is the record size as specified on the DF card. Each record after the header record up to and including the End-Of-File mark (EOF) contains a fixed number of words. The size of each record is specified by columns 2-6 of the DF card. In addition, each record of NNNNN words contains an integral number of scan lines. Video and IR data lines, when both are present, are interleaved.

Note that video lines are distinguished from IR lines by a frame line counter which has bit 35 (leftmost bit) set.

ATE

LMED

8