Optimum Interpolation of Radiosonde and Satellite-Derived Temperature Fields

By

Young Paul Yee

Department of Atmospheric Science
Colorado State University
Fort Collins, Colorado

Colorado State University

Department of Atmospheric Science

Paper No. 292

OPTIMUM INTERPOLATION OF RADIOSONDE AND SATELLITE-DERIVED TEMPERATURE FIELDS

Submitted by

Young Paul Yee

Atmospheric Science

In partial fulfillment of the requirements

for the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer, 1978

COLORADO STATE UNIVERSITY

Summer 1978

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION

BY Young Paul Yee

ENTITLED Optimum Interpolation of Radiosonde and Satellite-Derived Temperature Fields

BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF

Master of Science

Committee on Graduate Work

Adviser

ABSTRACT

An optimum interpolation scheme which interpolates temperature measurements to a specified mesoscale network of gridpoints is proposed. The analysis incorporates conventional synoptic-scale radiosonde observation at 00Z for thirty-one consecutive days in May 1976 and VTPR satellite-derived temperatures for a particular case study day. Structure functions were determined from the statistical nature of the 700 mb, 500 mb, and 300 mb temperature fields for this period. From these meteorological field samples, normalized auto-correlation and normalized mutual-correlation functions for the RAOB's and satellite-retrieved temperatures were calculated assuming general isotropic and homogeneous conditions over the region. Root-mean square instrumentation errors were estimated by fitting the unnormalized structure function to a linear regression curve to find the zero distance intercept. Applying the smoothed correlation functions, the mutual correlation functions, and the instrumentation errors to a simultaneous set of algebraic equations, the associated RAOB and satellite weighing factors for each gridpoint were computed. This study matches the vertical resolution of the radiosonde data to the horizontal resolution of the satellite radiance retrievals to produce a statistically better objective analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Figures Continued

List of Figures Continued

I. INTRODUCTION

Objective analysis of meteorological data is rapidly becoming a reality in light of the development of newer high speed, massive memory core computers. We can now look forward to an automated system which is capable of eliminating bad data points, interpolating between observations, smoothing out discrepancies within a network, and matching of different meteorological measurement systems. The end product can then be used in numerical weather prediction models where optimum interpolated values at specified grid points are needed. Another important application is the study of an optimum size network spacing for the measurement of singificant atmospheric systems and events. In the latter case, tradeoffs between the errors of interpolation as a function of grid-mesh spacing and the accuracy of measurement (instrumentation errors) must be weighed together. Optimum interpolation can tell us what the best station separation should be for measuring a particular atmospheric condition such as baroclinic systems, tropical disturbances, moisture fields, fronts, cyclonic activity, etc.

Panofsky (1949) was the first to consider the problem of ojective analysis by estimating the geopotential height fields using polynomial expansions. His scheme makes use of the statistical nature of the field in question whether it be pressure, geopotentials, temperature, or relative humidity. Since that time many objective analysis schemes have been proposed. Some methods such as Sasaki's (1958) matching technique which is an isoperimetric problem in the calculus of variations are too time-consuming and complicated to be operational.

Other methods work well in data dense regions but are inadequate in areas where radiosonde and aerological measurements are meager. In areas with good data coverage, one objective analysis scheme will perform just as well as another objective analysis scheme after allowing for the slight peculiarities within each technique. This is likened to the same peculiarities that will arise between two meteorologists with different backgrounds and experiences making a forecast from the same data. On the other hand, with bad data coverage many objective techniques fail to produce consistently reliable results. In the geopotential analysis done by the Russians (Gandin, L.S., 1963), noticeable changes between objective analysis schemes were observed in regions with sparse observational data. This was especially true when other sources of observations such as winds, ground measurements, climatological means, temperatures, etc. are assimilated into the interpolation of geopotentials. One of the major problems in objective analysis today is how to make maximum use of additional, available information.

The utilization of satellite radiance measurements provides an excellent source of real-time information. The data are received in a form which is convenient for rapid manipulations. Measurements are taken at regular grid intervals and their coverage is world-wide. For example, the NOAA-2 satellite is capable of computing over 14,000 temperature soundings a day and its scan-angle coverage is $\pm 7^{o}$ latitude on either side of the subsatellite point along its orbital path. Another advantage of a satellite is that it has a single calibrated instrument from which all its observations are taken; whereas, radiosondes

have internal biases in relation to each individual sounding and they will tend to drift away from the launch point. Looking at the radiosonde's point of view, the fine vertical temperature resolution of even one sounding cannot be easily duplicated by applying the inverse radiative transfer equation to the eight IR channels available on board the NOAA satellites. Except for the microwave region in the electromagnetic spectrum, satellites cannot measure below substantial cloud cover because the infrared radiation becomes attenuated by the water content in the air. Radiosonde soundings are not dependent upon the amount of clouds obscuring the atmosphere. From these considerations, the matching of satellite data with the conventional radiosonde observations should produce an optimum analysis of the temperature or moisture field, provided the proper weighing factors are calculated.

II. METHODS OF OBJECTIVE ANALYSIS

The Concept of objective interpolation is to compute meteorological values on a regular network of points from measured values at irregularly spaced observations as from recording ground stations or at irregularly timed observations as from satellites. Included in the analysis is the smoothing out of nonsystematic errors from the observations as well as the reducing of random roundoff errors. At the present time, three different methods of objective analysis are being used in various countries for the operational processing of meteorological information. The countries involved with the early development in this area are the United States, Sweden, and the Soviet Union. Panofsky (1949) and Gilchrist and Cressman (1954) devised one of the earliest methods of objective numerical weather analysis. Their method is essentially an interpolation scheme based upon fitting a second degree polynomial by the method of least-squares. In his early efforts, Panofsky applied his polynomial-fitting schemes to the calculations of meteorological fields of such elements as ageostrophic wind deviations and vertical velocities. He was also the first to investigate the possibility of matching different meteorological data fields. In the two cases he selected, Panofsky matched the pressure and wind fields by assuming that geostrophic conditions prevailed. At the time, all of his work was done without the help of computers. Several years later Gilchrist and Cressman followed Panofsky's original idea of polynomial representation of meteorological fields. In their study they made separate approximations for each gridpoint by using data within a 1000 x 1000 km square box. The constructed grid is shown in Figure 1 where a 19 x 19

point map is layed over the U.S. (gridpoint interval of about $.5^{o}$ latitude or 52.6 km). The results between the objective analysis and the subjective analysis versus the observed changes were compatible (Table 1).

Variations of the polynomial-fitting technique have been tried by the Joint Numerical Weather Prediction Unit (1957), Bushby and Huckle (1957), and Bykov, Kurbatkin and Gorelysheva (1964). Each of these modifications involved altering the weighting factors of the polynomials in some manner in order to minimize the mean square errors. Sometimes geostrophic wind data were used in conjunction with geopotential height fields. Currently, the polynomial approximation technique is being operationally applied in England and other countries.

Another objective method is the method of successive corrections as first proposed by the Bergthorsson and Doos (1955). Initially, a rough first approximation of some meteorological field is estimated from various sources such as numerical forecasts for that period, climatological data, or any observations from different levels or different parameters. The next step in the processing would be the revision of the initial field by comparison to the linear extrapolation of the current observations. The extrapolation values are weighted according to the distance between the observing stations and the grid point in consideration. Modifications have been since the advent of this method but the guiding principles are the same. Similarly Cressman (1959) and Haug (1959) have based their models upon that of Bergthorsson and Doos which essentially applies corrections to a first guess field. Today, the United States, Sweden, and Japan are the leading users of this technique.

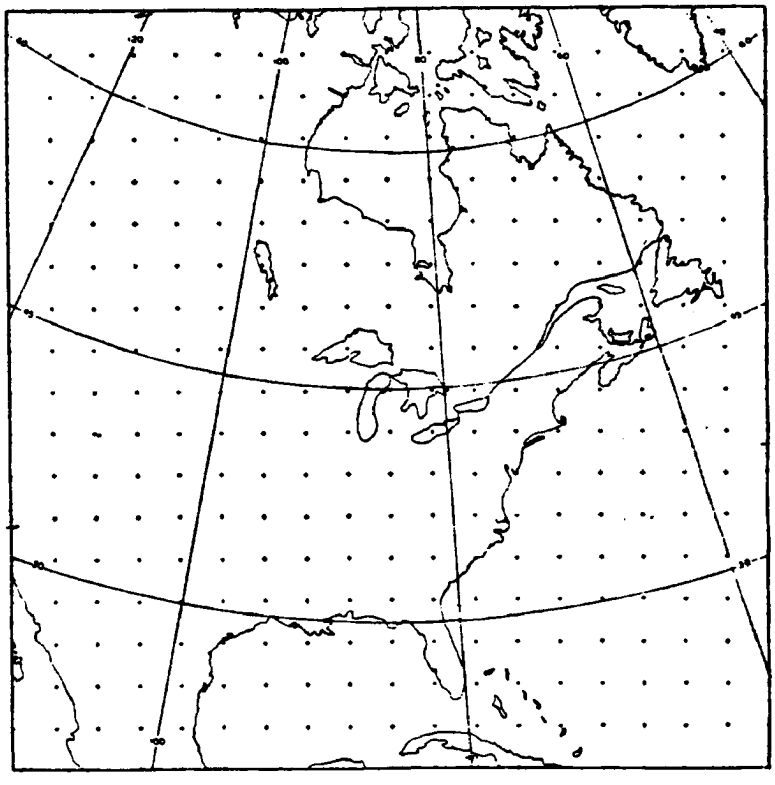


Figure 1. Outline of the 19 x 19 grid points used by Gilchrist and Cressman in their computational polynomial analysis.

	Correlation coefficient
Forecast change vs. observed change—subjective analysis	0.63
Forecast change vs. observed change—objective analysis	0.79
Forecast change—objective analysis vs. forecast change—subjective analysis	0.82

Table 1. Correlation coefficients for the subjective and objective determination of the observed change. (Gilcrist and Cressman, 1954)

The Russians are the main proponents of the optimum interpolation method whereby the value of an element at a grid point is determined from the values at stations with minimization of the root-mean-square errors. Because the previous method requires a large amount of initial data, the Russian meteorologists have dealt almost exclusively with optimum interpolation. In the early 1960's, Gandin (1963) became one of the leading spokesman in this particular field of study. In our preliminary case study, we preferred to use this method over the first two techniques because of certain faults that are inherently undesireable in the other two. The disadvantages of the polynomial approximation are: (1) It gives unsatisfactory results when applied to a sparse network. In such cases the amount of data used for the interpolation may be too close to the minimum data required to solve a multi-degree polynomial. For example, a second-degree polynomial would require a minimum of six initial meteorological quantities otherwise calculational instability develops in the equations. This instability begins with a minor variation in the initial data and soon leads to large variations in the interpolation weighting factors. Other deficiences in the polynomial approximation are described by Cressman (1957, 1959). (2) The polynomials are arbitrarily chosen so that their mathematical form may not correspond to the dynamic properties of the meteorological fields. (3) The size of the effective neighborhood around each grid point is arbitrarily chosen. (4) The ratio of the weighting factors such as for wind and heights can be altered during the course of the analysis.

In data dense areas, such as over the United States, the polynomial method rivals subjective methods as seen by the correlation coefficients in Table 1. However, this method was initially designed for areas of

relatively abundant data and will not be practical for a hemispheric region of the world. Although our observations will be taken over the U.S., our final intent is to show the feasibility of assimilating satellite information along with radiosonde measurements in data sparse regions over the globe or on scales smaller than conventional networks. To demonstrate the need for better coverage over certain parts of the Northern Hemisphere, Gubanova (1964) analyzed the geographical distribution of the mean absolute error $\sqrt{E}$ of the interpolation for the temperature field at 500 mb. The mean value of $\sqrt{E}$ (which will be derived later) was equal to 2.45^{0} C and the largest interpolation errors were observed over the Pacific Ocean and the north pole regions (see Figure 2). Miyakoda et. al. (1976) showed relatively large rms values of temperature difference for the Southern Hemisphere as compared to the Northern Hemisphere, particularly in winter. Again this is due to the scarcity of observational data there. Therefore, for these less dense data network areas it will be more advantages to consider the method of optimum interpolation. Eliassen (1954), Gandin (1963), and Kruger (1968) have also indicated their preference for the chosen method as being the most efficient one to minimize the error in their analysis.

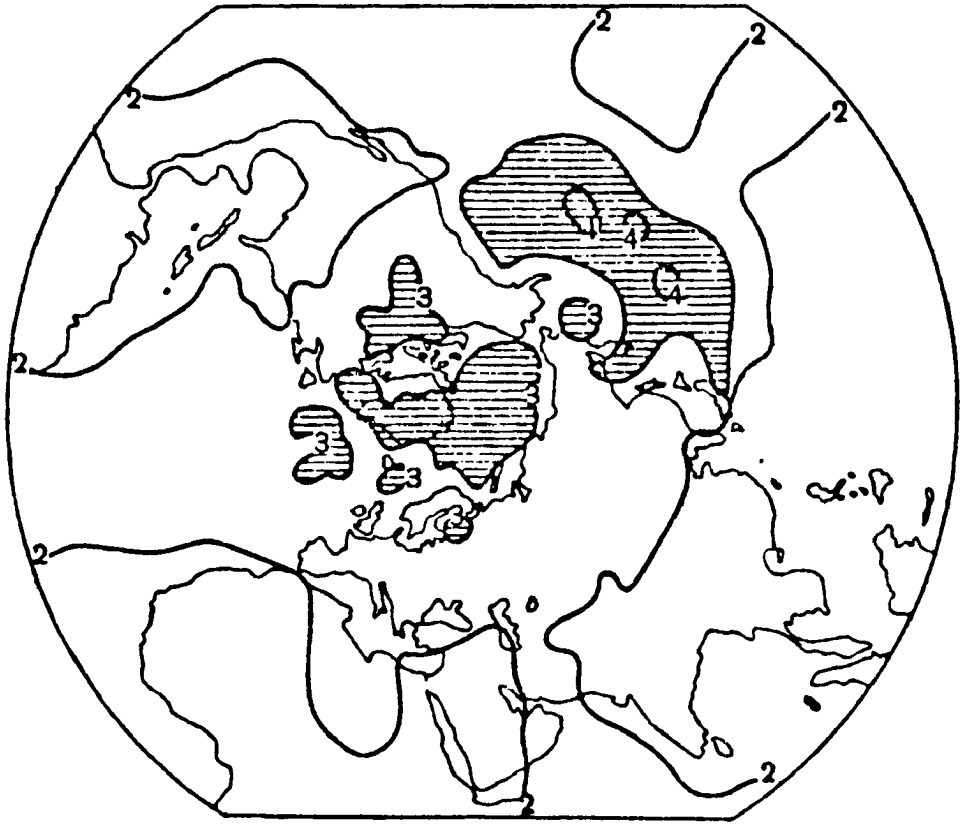


Figure 2. Global distribution of the mean absolute error for the 500 mb interpolation of temperature. (Taken from Gubanova, 1964)

III. OPTIMUM INTERPOLATION

Although optimum interpolation techniques have been known for over a decade by the Russians, only recently has there been a revived interest in this subject matter. The optimum objective analysis as first proposed by Gandin has been fully developed in its multivariate form by Thiebaux (1973) and by Gandin and Kagan (1974). Brief reviews pertaining to this subject can be found in Doos (1969), Gandin (1969), and Kruger (1968). Other various treatments of optimum interpolation are discussed by Bengtsson and Gustavsson (1971), Morel (1970), Rutherford (1972, 1973), Schlatter (1975), and Thiebaux (1975).

To begin our derivation we will consider a meteorological parameter, say temperature, T. The same treatment will hold for any other meteorological element we may wish to consider, for example mixing ratios or geopotentials. To determine T at a particular grid point 0, we can define a circular domain around this grid point with a radius R and note the number of observations within the domain (Figure 3). For example if we want to find the six closest radiosonde observations to a grid point, we will increase the radius of influence until six observations are found. This is the procedure that was taken in our case study. Once the six observations become available, we can represent the interpolated temperature value at the grid point as a linear combination of the observing points.

$$T'_0 = \sum_{i=1}^{n} p_i T'_i \qquad (1)$$

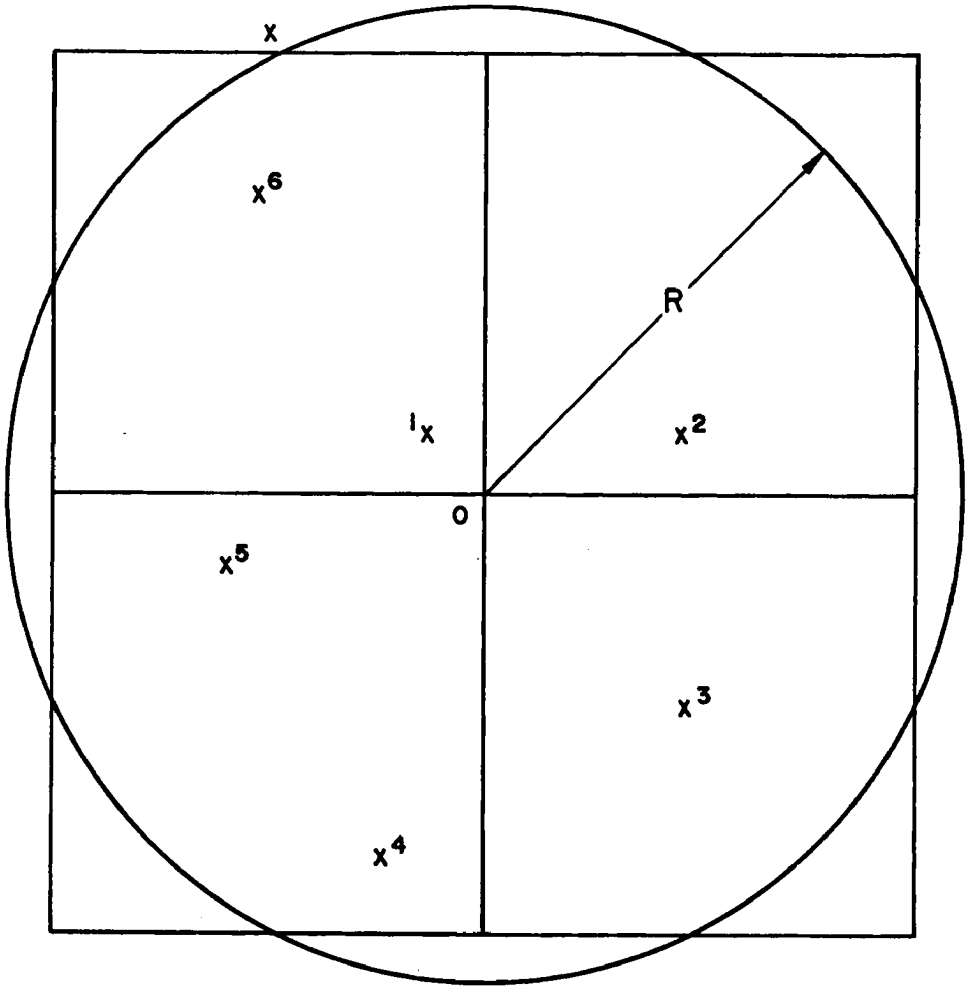


Figure 3. A circular area domain surrounding a grid point located at 0 in which observations are taken and weighted together.

where T_i' are the deviations of the temperatures from some climatological norm $\bar{T}$,

$$T_i' = T_i - \bar{T} . \qquad (2)$$

T_i is the value at the radiosonde station i and the primed variables denote deviations from the norm. The p_i's are weighting coefficients to be determined for each station i. The upper index n of the summation is the number of selected stations inside the domain of radius R and is taken to be six for our particular study. According to Gandin, the maximum number of stations surrounding a grid point should be less than 8, otherwise the solution will be unstable and the weighting factors will turn out to be unreasonable . The p_i's are found by requiring the mean square of the errors,

$$E = \overline{(T_0' - T_t')^2} \qquad (3)$$

be a minimum, where T_t' is the true temperature at the grid point and the bar denotes a statistical averaging. Substituting equation 1 into equation 3 we get

$$E = \sum_{i=1}^{n} \sum_{j=1}^{n} p_i p_j \overline{T_i' T_j'} - 2 \sum_{i=1}^{n} p_i \overline{T_i' T_t'} + \overline{T_t'^2} . \qquad (4)$$

To minimize E with respect to all p_k's the following condition must hold,

$$\frac{\partial E}{\partial p_k} = 0 \qquad (k = 1,\ldots,n) \qquad (5)$$

Thus we arrive at a set of simultaneous algebraic equations,

$$\sum_{j=1}^{n} p_j \overline{T_i' T_j'} = \overline{T_i' T_0'} \qquad \text{(for each } i = 1,\ldots,n) \quad (6)$$

By definition the covariances or correlation functions are the mean products of the deviations from the norms,

$$m_{ij} = \overline{T_i' T_j'} \qquad (7)$$

and in particular, the variance is

$$m_{ii} = \overline{T_i'^2} \qquad (8)$$

In later analysis it will be convenient to use normalized correlation functions as defined by

$$\mu_{ij} = \frac{m_{ij}}{(m_{ii} m_{jj})^{1/2}} \qquad (9)$$

i.e. the correlation functions divided by the square root of the product of the variances. The normalized correlation functions depend less on the averaging conditions than the corresponding unnormalized correlation function. In this way, various seasonal changes and annual biases can be factored out. Substituting equation 7 and then 9 into equation 6 and dividing by the product of the square root of the variances as defined in equation 8, we have another form of equation 6,

$$\sum_{j=1}^{n} \mu_{ij} p_j = \mu_{i0} \qquad \text{(for each } i = 1,\ldots,n) \quad (10)$$

The above equation holds true only if the measurements are perfect, that is without instrumental or random errors in the data. In this simplified form, the weighting factors are easily found once the functional forms of the normalized correlations are characterized.

Similar to the covariances are the structure functions defined as the mean-square difference between the temperature deviation at one station i and another station j,

$$b_{ij} = \overline{(T'_i - T'_j)^2} \tag{11}$$

If we carry through with the derivation, we find that the correlation function m_{ij} and the structure functions are related by the following equation,

$$b_{ij} = m_{ii} + m_{jj} - 2m_{ij} \tag{12}$$

Finally, if we want to account for the existence of random observational errors, the observed temperature deviations would be equal to the true values plus the observational errors ΔT at each point,

$$T'_i = T'_t + \Delta T'_i \tag{13}$$

A complete treatment on the influence of random observational errors on optimum interpolation can be found in Gandin (1963). The system of simultaneous equations equivalent to equation 10 now reduces down to

$$\sum_{j=1}^{n} \mu_{ij} p_j + \eta_i p_i = \mu_{i0} \qquad \text{(for each } i = 1,\ldots,n) \tag{14}$$

where η_i is called the measure of the random or observational errors. It is the ratio of the mean-square observational error in the data over the variance of the temperature,

$$\eta_i = \frac{\overline{\Delta T'^2}}{\overline{T'^2_i}} \tag{15}$$

The set of equations 14 simply implies that when solving for the matrix of weighting coefficients p_i's, the diagonal of the two-dimensional correlation matrix has the added term, η_i, as shown by the following matrix equation for six observations.

$$\begin{bmatrix} \mu_{11}+\eta_1 & \mu_{12} & \mu_{13} & \mu_{14} & \mu_{15} & \mu_{16} \\ \mu_{21} & \mu_{22}+\eta_2 & \mu_{23} & \mu_{24} & \mu_{25} & \mu_{26} \\ \mu_{31} & \mu_{32} & \mu_{33}+\eta_3 & \mu_{34} & \mu_{35} & \mu_{36} \\ \mu_{41} & \mu_{42} & \mu_{43} & \mu_{44}+\eta_4 & \mu_{45} & \mu_{46} \\ \mu_{51} & \mu_{52} & \mu_{53} & \mu_{54} & \mu_{55}+\eta_5 & \mu_{56} \\ \mu_{61} & \mu_{62} & \mu_{63} & \mu_{64} & \mu_{65} & \mu_{66}+\eta_6 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{bmatrix} = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \mu_{30} \\ \mu_{40} \\ \mu_{50} \\ \mu_{60} \end{bmatrix} \qquad (16)$$

When random errors are allowed for, the conditioning of the system of equations for determining the optimum weighting factors is improved. The larger η_i is, the smaller will be the off diagonal terms of the matrix of the coefficients in comparison with the diagonal terms. The simpliest case assumes that the observational accuracy is the same everywhere at all the points, which is usually true for radiosonde measurements (i.e. $\eta_i = \eta$ for all $i = 1,\ldots,n$). Obviously, we do not know exactly the value of the observational errors but knowing that errors do exist in the analysis can be compensated for by the method of optimum interpolation if the statistical properties of the errors are known.

Using optimum interpolation techniques, Alaka and Elvander (1972) studied the trade-off characteristics between instrumental accuracy and network sampling resolution over the Carribbean area. The most practical aspect of the problem is to find the optimum density of

meteorological observations if the instrumental accuracy is already given. In our particular case, we have an estimate of the rms random errors from a radiosonde measurement and we have an estimate of the temperature profile accuracy derived from satellite radiance measurements. Table 2 shows some of the expected accuracies from satellite-based systems and constant level-balloons for the measurement of winds, temperature, pressure, and water vapor. According to Bessemoulin (1960) the probability that the error in the measured difference between two observations does not exceed a fraction k of the true difference can be represented by the next equation assuming a Gaussian shaped distribution of the meteorological variations.

$$P = 1 - \frac{2}{\pi} \tan^{-1} \frac{\lambda}{k} \{2(1 - \mu(\tau))\}^{-\frac{1}{2}} \qquad (17)$$

where P = probability that the measured error between two stations is not greater than a k^{th} of their true error.

λ = error index = $\frac{\sigma_\varepsilon}{\sigma}$ = $\frac{\text{rms error of observation}}{\text{standard deviation of the true variation of the measured element}}$

k = fractional part of the true error difference between stations.

$\mu(\tau)$ = correlation coefficient with respect to time or space.

As can be seen in Figure 4 for a given k, the probability increases with decreasing observational errors and increasing distance between two observations. As the distance between stations becomes shortened, the interpolation error becomes more sensitive to random errors of observation.

The set of equations 14 simply implies that when solving for the matrix of weighting coefficients p_i's, the diagonal of the two-dimensional correlation matrix has the added term, η_i, as shown by the following matrix equation for six observations.

$$\begin{bmatrix} \mu_{11}+\eta_1 & \mu_{12} & \mu_{13} & \mu_{14} & \mu_{15} & \mu_{16} \\ \mu_{21} & \mu_{22}+\eta_2 & \mu_{23} & \mu_{24} & \mu_{25} & \mu_{26} \\ \mu_{31} & \mu_{32} & \mu_{33}+\eta_3 & \mu_{34} & \mu_{35} & \mu_{36} \\ \mu_{41} & \mu_{42} & \mu_{43} & \mu_{44}+\eta_4 & \mu_{45} & \mu_{46} \\ \mu_{51} & \mu_{52} & \mu_{53} & \mu_{54} & \mu_{55}+\eta_5 & \mu_{56} \\ \mu_{61} & \mu_{62} & \mu_{63} & \mu_{64} & \mu_{65} & \mu_{66}+\eta_6 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{bmatrix} = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \mu_{30} \\ \mu_{40} \\ \mu_{50} \\ \mu_{60} \end{bmatrix} \qquad (16)$$

When random errors are allowed for, the conditioning of the system of equations for determining the optimum weighting factors is improved. The larger η_i is, the smaller will be the off diagonal terms of the matrix of the coefficients in comparison with the diagonal terms. The simpliest case assumes that the observational accuracy is the same everywhere at all the points, which is usually true for radiosonde measurements (i.e. $\eta_i = \eta$ for all $i = 1,\ldots,n$). Obviously, we do not know exactly the value of the observational errors but knowing that errors do exist in the analysis can be compensated for by the method of optimum interpolation if the statistical properties of the errors are known.

Using optimum interpolation techniques, Alaka and Elvander (1972) studied the trade-off characteristics between instrumental accuracy and network sampling resolution over the Carribbean area. The most practical aspect of the problem is to find the optimum density of

meteorological observations if the instrumental accuracy is already given. In our particular case, we have an estimate of the rms random errors from a radiosonde measurement and we have an estimate of the temperature profile accuracy derived from satellite radiance measurements. Table 2 shows some of the expected accuracies from satellite-based systems and constant level-balloons for the measurement of winds, temperature, pressure, and water vapor. According to Bessemoulin (1960) the probability that the error in the measured difference between two observations does not exceed a fraction k of the true difference can be represented by the next equation assuming a Gaussian shaped distribution of the meteorological variations.

$$P = 1 - \frac{2}{\pi} \tan^{-1} \frac{\lambda}{k} \{2(1 - \mu(\tau))\}^{-\frac{1}{2}} \qquad (17)$$

where P = probability that the measured error between two stations is not greater than a k^{th} of their true error.

λ = error index $= \frac{\sigma_\varepsilon}{\sigma} =$ rms error of observation / standard deviation of the true variation of the measured element

k = fractional part of the true error difference between stations.

$\mu(\tau)$ = correlation coefficient with respect to time or space.

As can be seen in Figure 4 for a given k, the probability increases with decreasing observational errors and increasing distance between two observations. As the distance between stations becomes shortened, the interpolation error becomes more sensitive to random errors of observation.

Atmospheric parameter		*Resolution*	*Accuracy*
Winds from constant-level balloons	V_b	——— 400 – 500 km – – – – less good -------- summer observations only	2 m/s
Winds from cloud displacement	V_c	not known yet, mainly day-time observations	3 m/s
Winds from buoys and automatic stations	V_a	not known yet, in tropics only	
Temperature from constant-level balloons	T_b	——— 400 – 500 km – – – – less good ------- summer observations only	0.5°C
Temperature from IR-soundings	T_I	quasi-global, good resolution if no clouds	2°C
Temperature from Microwave soundings	T_M	global, over ocean only, good resolution if no clouds	2°C
Sea surface temperature from high resolution IR radiometer	T_a	global in cloud free or scattered cloud areas Resolution 10's of km	1°C gradients to 0.5°C
Temperature from buoys and automatic stations	T_a	not known yet, in tropics only	
Pressure from constant level-balloons	z_b	——— 400 – 500 km – – – – less good -------- summer observations only	10 m (height of pressure level)
Pressure from buoys and automatic stations	p_a	not known yet, in tropics only	
Water vapour from IR-soundings	q_I	good horizontal resolution if no opaque cirrus	1 part in 10

——— = mid latitudes

– – – – = Tropics

-------- = high latitudes

Table 2. Accuracy comparisons of meteorological observations taken from space based systems. (From Bengtsson, L., 1975)

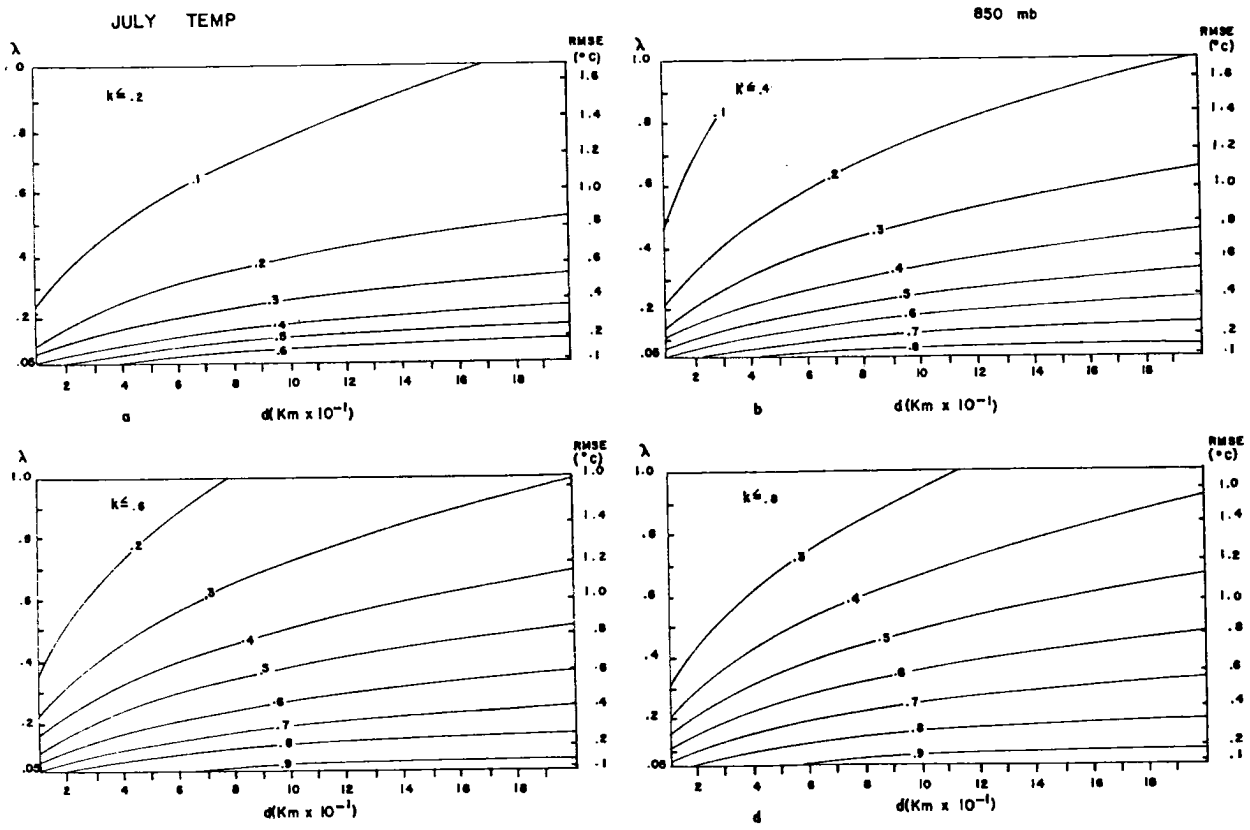


Figure 4. Lines of constant probabilities, such that the error in the measured difference between two temperature observations at 850 mb in July is equal to or less than a fraction k of the true difference, as a function of the rms random errors of the observations and the distance (in tens of kilometers) between them.

Optimum interpolation has many applications such as time projections of persistent meteorological features into forecasting models or space projection of an abundant data source into a region where there is a completely lack of data, a 'hole'. Other applications include optimum network studies and optimum matching of different sources of data. But in order to do this, the functional correlations must be found.

The next stage of the analysis will be to find an empirical formula for the representation of correlation functions as a function of distance only.

IV. STRUCTURE AND CORRELATION FUNCTIONS

The method of optimum interpolation requires a knowledge of the statistical nature of the meteorological field under study. This statistical knowledge is supplied by the structure functions and correlation functions. In practice, local homogeneity and isotropy is usually assumed for the correlation function μ_{ij}, meaning μ_{ij} is dependent only upon the distance between two observations $\mu_{ij} = \mu_{ij}(r)$ and not upon the position of the distance vector or the direction of the vector. Figure 5 demonstrates the idea of preserving isotropy and homogeneity.

Although most investigators of optimum interpolation (Buell 1972; Julian and Thiebaux 1975; Lacy 1973; Olevskays 1967; Ramanathan et. al. 1973; Rutherford 1973) have made use of the assumption of isotropy and homogeneity, Thiebaux has presented initial evidence indicating limitations of accuracy due to the isotropic assumption. Using the 500 mb winter height-field, Thiebaux found that the correlation values could be segregated on the basis of angle of station separation, generally into either a zonal or meridional direction. Comparing his two-dimensional anisotropic model to other standard isotropic models, Thiebaux claimed that anisotropy accounted for decreasing the sum of squared deviations to 0.44 of the sum of squared deviations for the best isotropic fit. Even though these results apply only to geopotential heights and not to temperature and mixing ratios as in our study, climatologically, there does exist a global zonal and/or meridional dependency with temperature. It is difficult to say how these global averages may affect a meteorological field within a small regional area but it must

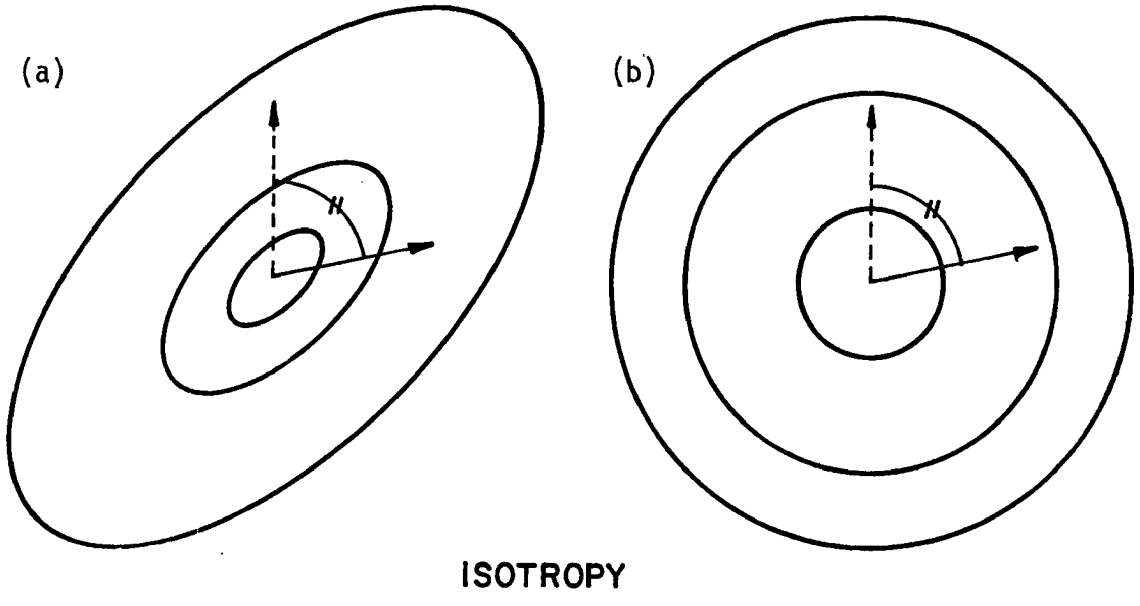
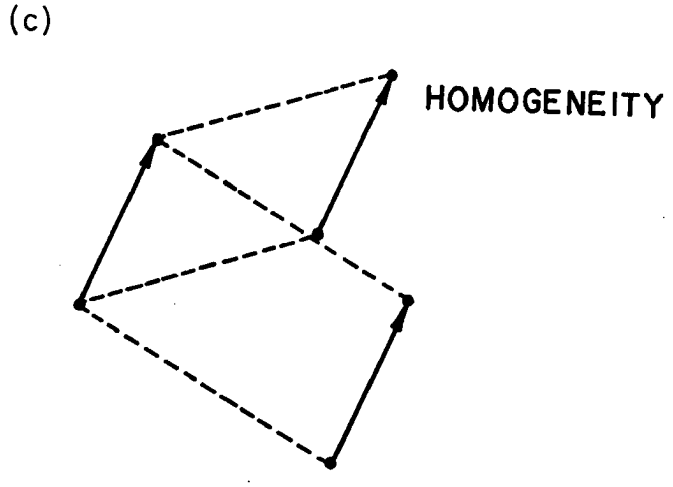


Figure 5. Illustrations of the concepts of isotropy and homogeneity. (a) This figure does not preserve isotropy in the field since the correlation values depend upon the direction between the distance vector of the two stations. (b) These concentric circles of correlation values preserves the idea of isotropy since correlations are dependent upon distance and not direction. (c) Homogeneity is preserved when the distance vector can be translated without affecting the value of the correlation function.



be bore in mind that they do exist. This suggests that parameterizing a field may improve the final analysis.

The functional forms of the correlations functions is not necessarily the same for temperature, wind, water vapor and geopotential fields. But one of the most widely used mathematical formula for the correlation function is

$$\mu = ae^{-bx^2} \tag{18}$$

where μ = correlation function

x = distance between observational locations

a,b = constants to be found by a least-squares fit.

Sometimes called the Gaussian correlation function by several authors, this exponential form has been applied to experiments dealing with the design of networks for precipitation measurements and interpolation of temperatures. Rutherford (1973, 1976) and Schlatter (1976) also use the Gaussian functional form in their optimum interpolation of synoptic systems.

Other correlation equations have been tried and tested for different cases and some of these forms are as follows:

$$\mu = \{b \cos(ax) + d\text{-}b\}\, e^{-cx^2} \tag{19}$$

$$\mu = \{b \cos(ax) + d\text{-}b\}\, \{1 + (cx)^2\}^{-\frac{1}{2}} \tag{20}$$

Equation 19 is called the "cosine-modified Gaussian" curve which gives a reasonable fit over a wide range of distances for an isotropic correlation array as shown in Figure 6 along with regular Gaussian

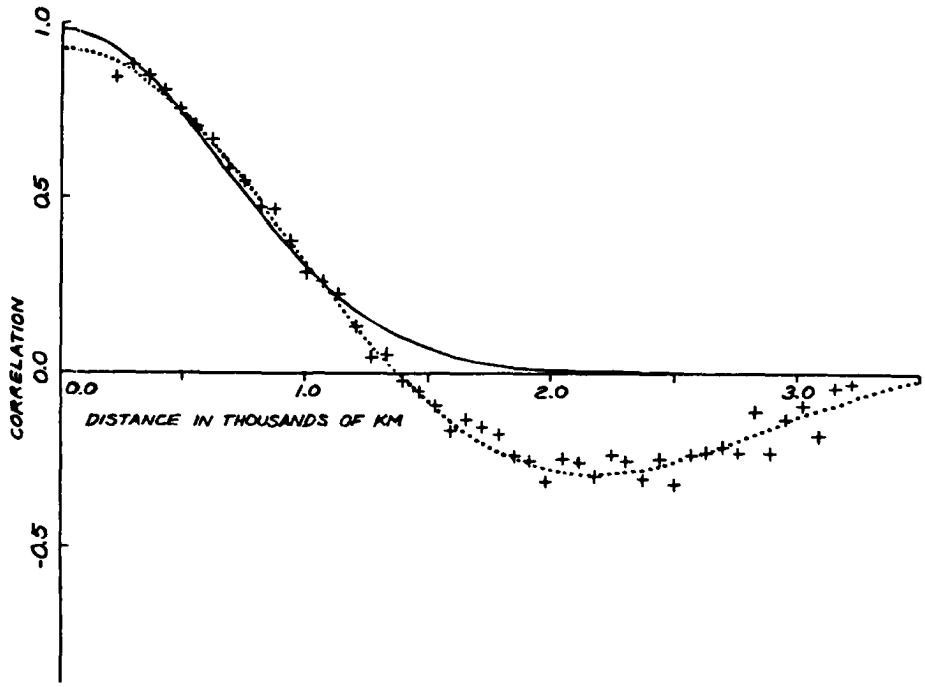


Figure 6. Distance-interval-averaged observed correlations of geopotentials with curves fitted to the correlation values, $\mu(x)=ae^{-cx^2}$ (solid curve) and $\mu(x)=(b\cos(ax) + d - b)e^{-cx^2}$ (dotted curve) where x is distance in thousands of km. (Julian and Thiebaux, 1975)

de^{-cx^2}	c = 1.1567 d = 0.9763
$\{b\cos(ax) + d - b\}e^{-cx^2}$	a = 1.2572 b = 0.7778 c = 0.1507 d = 0.9151
$\{b\cos(ax) + d - b\}\{1 + (cx)^2\}^{-\frac{1}{2}}$	a = 1.3800 b = 0.7380 c = 0.8480 d = 0.9800

curve. Julian and Thiebaux (1975) have tested the fit of the above equations to the observed correlations and power spectral behavior and they concluded that equation 20 performed best in modeling height field correlations. Its shape very closely resembles that of equation 19. Each functional representation will have some favorable advantages over the others depending upon the fields to be measured, special criteria to be satisfied, the area to be analyzed and the anisotropic or isotropic nature of the field.

V. MAY 1976 CASE STUDY OVER MID-WEST U.S.

To find the structure and correlation functions from radiosonde temperature and dew-point depression measurements, we read in data from twenty conventional radiosonde stations shown in Table 3. These stations were enclosed within the longitudial belts $90^{o}W$-$110^{o}W$ and the latitudial belts $30^{o}N$-$45^{o}N$ as shown in Figure 7. Thirty-one consecutive days of RAOB data were collected from May 1 to May 31, 1976 at 0000Z hour. The soundings were taken from three mandatory pressure levels in the atmosphere: 700 mb, 500 mb, and 300 mb.

After the dew-point depressions were read in, they were converted to mixing ratios by applying Teten's formula, for liquid water,

$$W_s = \frac{3.8}{P} \exp \frac{A(T^*-273)}{T^*-B} \qquad \begin{matrix} A = 17.27 \\ B = 36.0 \end{matrix} \qquad (21)$$

where W_s = mixing ratio

P = pressure in millibars

T^* = the dew-point temperature in ^{o}K.

The time-averaged norms for temperature and mixing ratios were then calculated for each of the twenty radiosonde sites (Table 4). The distances between radiosonde stations were computed according to their latitude-longitude location as shown by the distance matrix in Table 5. Since the distance vector is symmetic between two stations (i.e. $x_{ij} = x_{ji}$), the distance matrix will also be symmetric along the diagonal. It can be seen that the farthest station pair is 2015 km but there are no station pairs less than 250 km apart. The latter fact will make it difficult to interpolate the correlation functions for small distances.

	STATION	WMO STATION CODE	LATITUDE(^{o}N)	LONGITUDE(^{o}W)	ELEVATION(M)
1	JAN	72235	32.32	90.08	102
2	LCH	72240	30.12	93.22	5
3	GGG	72247	32.35	94.65	122
4	SEP	72260	32.22	98.18	402
5	MAF	72265	31.95	102.18	874
6	ELP	72270	31.80	106.40	1206
7	LIT	72340	34.50	92.16	172
8	UMN	72349	36.88	93.90	438
9	OKC	72353	35.40	97.60	397
10	AMA	72363	35.23	101.70	1099
11	ABQ	72365	35.05	106.62	1619
12	DDC	72451	37.77	99.97	790
13	TOP	72456	39.07	95.63	270
14	DEN	72469	39.75	104.87	1625
15	GJT	72476	39.12	108.53	1475
16	OMA	72553	41.37	96.02	406
17	LBF	72562	41.13	100.68	849
18	LND	72576	42.82	108.73	1700
19	HON	72654	44.38	98.22	393
20	RAP	72662	44.05	103.07	966

Table 3. Radiosonde stations with their corresponding WMO code, location, and elevation above sea level in meters.

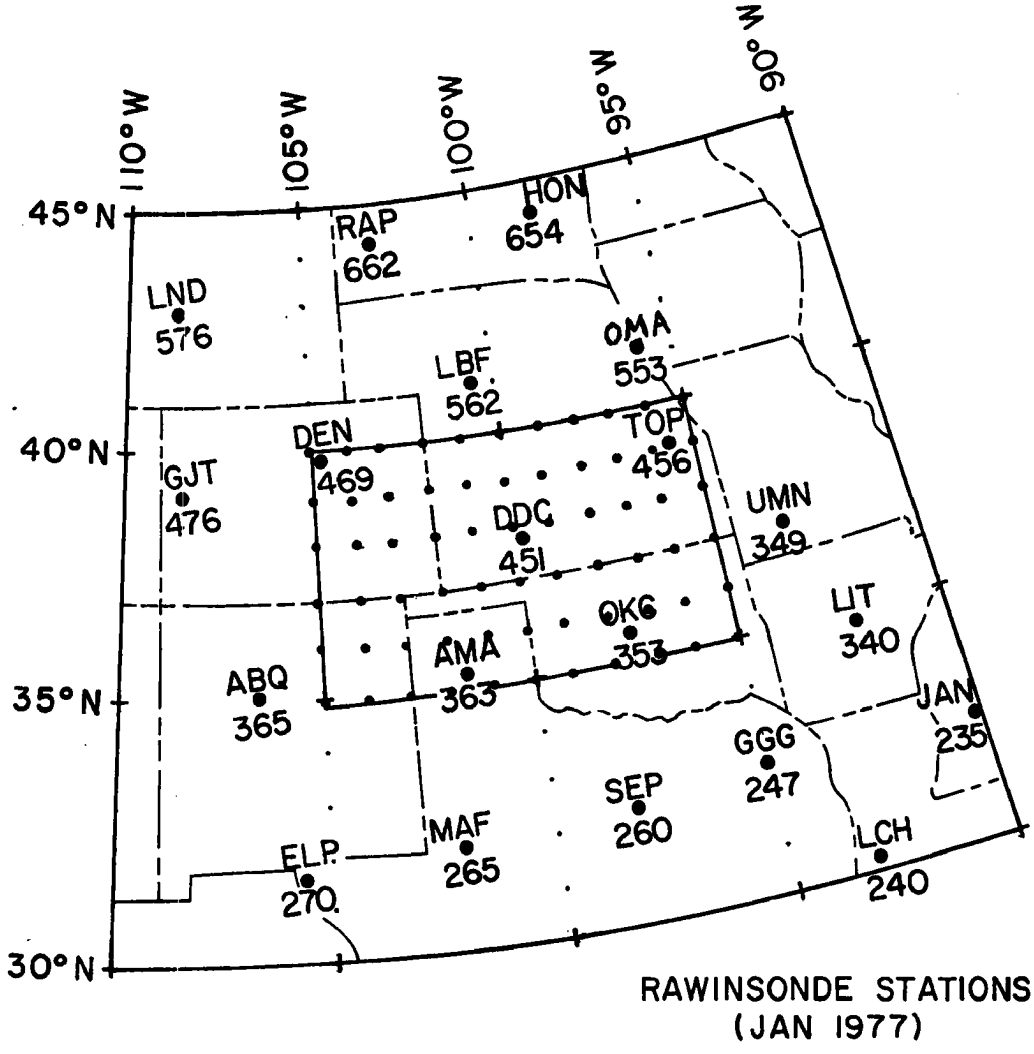


Figure 7. Geographical location of conventional radiosonde network within the latitudinal belts of 30°N to 45°N and the longitudinal belts of 90°W to 110°W. The boxed area between 35°N-40°N latitude and 95°W-105°W longitude will be the mesoscale region for optimum interpolation. The grid point intervals are in one degree latitude and longitude increments.

WMO STATION CODE		700 MB NORMS T(°C)	q(gm/kg)	500 MB NORMS T(°C)	q(gm/kg)	300 MB NORMS T(°C)
1	JAN	2.5	1.96	-13.9	.78	-40.9
2	LCH	4.5	2.43	-12.7	.84	-39.4
3	GGG	3.1	2.51	-13.9	.79	-41.1
4	SEP	4.9	2.79	-13.2	.82	-40.5
5	MAF	5.7	3.63	-13.1	.69	-41.1
6	ELP	6.6	2.34	-14.1	.56	-41.7
7	LIT	2.2	2.81	-15.0	.78	-41.8
8	UMN	1.7	2.56	-15.4	.59	-42.7
9	OKC	3.1	2.60	-15.0	.61	-42.8
10	AMA	4.6	3.69	-13.9	.60	-42.6
11	ABQ	7.2	2.43	-15.1	.95	-43.0
12	DDC	3.1	2.69	-15.2	.55	-43.2
13	TOP	1.8	2.32	-15.5	.66	-43.2
14	DEN	3.9	3.16	-16.0	1.02	-43.8
15	GJT	4.7	2.98	-15.8	.99	-44.2
16	OMA	1.5	1.98	-15.6	.59	-43.7
17	LBF	2.4	3.34	-15.4	.78	-43.9
18	LND	2.8	2.85	-16.9	.80	-44.9
19	HON	.3	1.89	-16.7	.72	-44.7
20	RAP	.8	3.12	-16.5	.76	-44.3

Table 4. The mean temperatures and mixing ratios (q(gm/kg)) during May 1976 over each of the twenty radiosonde stations.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	JAN	0	386	429	761	1139	1538	310	615	774	1121	1559	1085	982	1565	1828	1136	1360	2015	1516	1728
2	LCH	386	0	283	526	877	1270	497	754	715	975	1369	1053	1019	1507	1721	1275	1397	1978	1645	1777
3	GGG	429	283	0	332	710	1108	332	508	435	725	1147	773	752	1233	1461	1010	1114	1700	1372	1494
4	SEP	761	526	332	0	378	776	613	649	357	467	842	638	795	1031	1203	1035	1015	1501	1351	1382
5	MAF	1139	877	710	378	0	399	974	936	571	367	537	677	988	900	982	1182	1029	1339	1424	1347
6	ELP	1538	1270	1108	776	399	0	1358	1278	907	579	362	886	1266	894	836	1410	1156	1242	1570	1392
7	LIT	310	497	332	613	974	1358	0	308	505	874	1321	789	594	1268	1544	835	1050	1709	1215	1417
8	UMN	615	754	508	649	936	1278	308	0	371	724	1162	545	287	1008	1305	531	752	1427	910	1112
9	OKC	774	715	435	357	571	907	505	371	0	372	820	338	443	802	1051	678	691	1265	999	1069
10	AMA	1121	975	725	467	367	579	874	724	372	0	448	322	686	575	743	843	662	1039	1059	987
11	ABQ	1559	1369	1147	842	537	362	1321	1162	820	448	0	667	1072	545	483	1162	852	882	1261	1045
12	DDC	1085	1053	773	638	677	886	789	545	338	322	667	0	405	478	760	524	378	931	749	745
13	TOP	902	1019	752	795	988	1266	594	287	443	686	1072	405	0	797	1113	258	486	1176	628	830
14	DEN	1565	1507	1233	1031	900	894	1268	1008	802	575	545	478	797	0	322	769	386	469	752	580
15	GJT	1828	1721	1461	1209	982	836	1544	1305	1051	743	483	760	1113	322	0	1090	703	411	1035	711
16	OMA	1136	1275	1010	1035	1182	1410	835	531	678	843	1162	524	258	769	1090	0	390	1060	379	648
17	LBF	1360	1397	1114	1015	1029	1156	1050	752	691	662	852	378	486	386	703	390	0	691	413	379
18	LND	2015	1978	1700	1501	1339	1242	1709	1427	1265	1039	882	931	1176	469	411	1060	691	0	863	477
19	HON	1516	1645	1372	1351	1424	1570	1215	910	999	1059	1261	749	628	752	1035	379	413	863	0	388
20	RAP	1728	1777	1494	1382	1347	1392	1417	1112	1069	987	1045	745	830	580	711	648	379	477	388	0

Table 5. Distance matrix in km between all possible combinations of radiosonde stations' pair.

After the initializing part of the program and the checking for missing data (see Appendix C), the covariances or unnormalized correlation functions as defined by equation 7 are calculated and inserted into a 20 x 20 symmetrical array. The mean covariances for each station pair was taken over the thirty-one day period in May 1976. The unnormalized structure functions were calculated at this time and the results can be seen in the averaged structure functions per 50 km intervals for 700 mb, 500 mb, and 300 mb, where structure functions are plotted versus the distances between pairs of stations. The data points seem to fit a linear relationship for distances under 1500 km. Distances over about 1500 km begin to show lack of fit due mainly because the assumption of homogeneity becomes invalid. The scarcity of station pair distances over 1500 km is also likely to contribute to the random fluctuations in this range. The slope of the curve fit at 300 mb appears to be more defined and steeper than at 500 mb and 700 mb and this fact will help to lower the extrapolated mean-square error at this level. The ordinate in Figure 8 are in units of degrees-squared. The histogram corresponding to the number of correlation functions within each 50-km interval distance is given by Figure 9. Theoretically, it is possible to find two times the mean-square error in the data (i.e. $2\sigma^2$, where σ is sometimes called the average absolute error) by extrapolating the sample structure function curve to zero distance between observations. In practice, however, one cannot be certain of the accuracy of the functional curve representation at close distances unless there are a sufficient number of samples within that vicinity. Nevertheless, even a linear extrapolation can give us a reasonable estimate of the rms instrumental errors in the data sample.

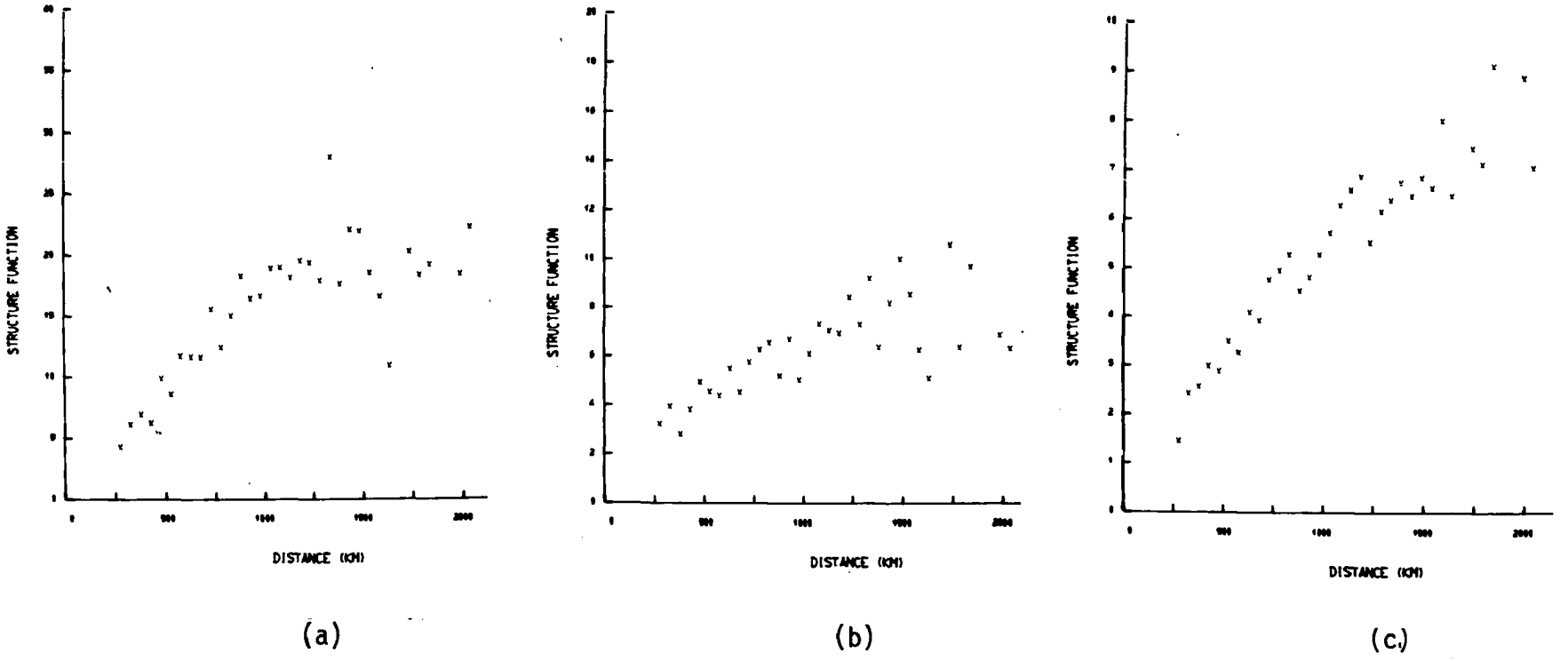


Figure 8. Unnormalized structure functions for RAOB temperatures at (a) 700 mb (b) 500 mb and (c) 300 mb. Samples are averaged over 50 km interval-distances. (Units of the structure function values are in degrees-squared.)

Figure 9. Histogram of the number of RAOB station pairs versus the incremental 50 km distances.

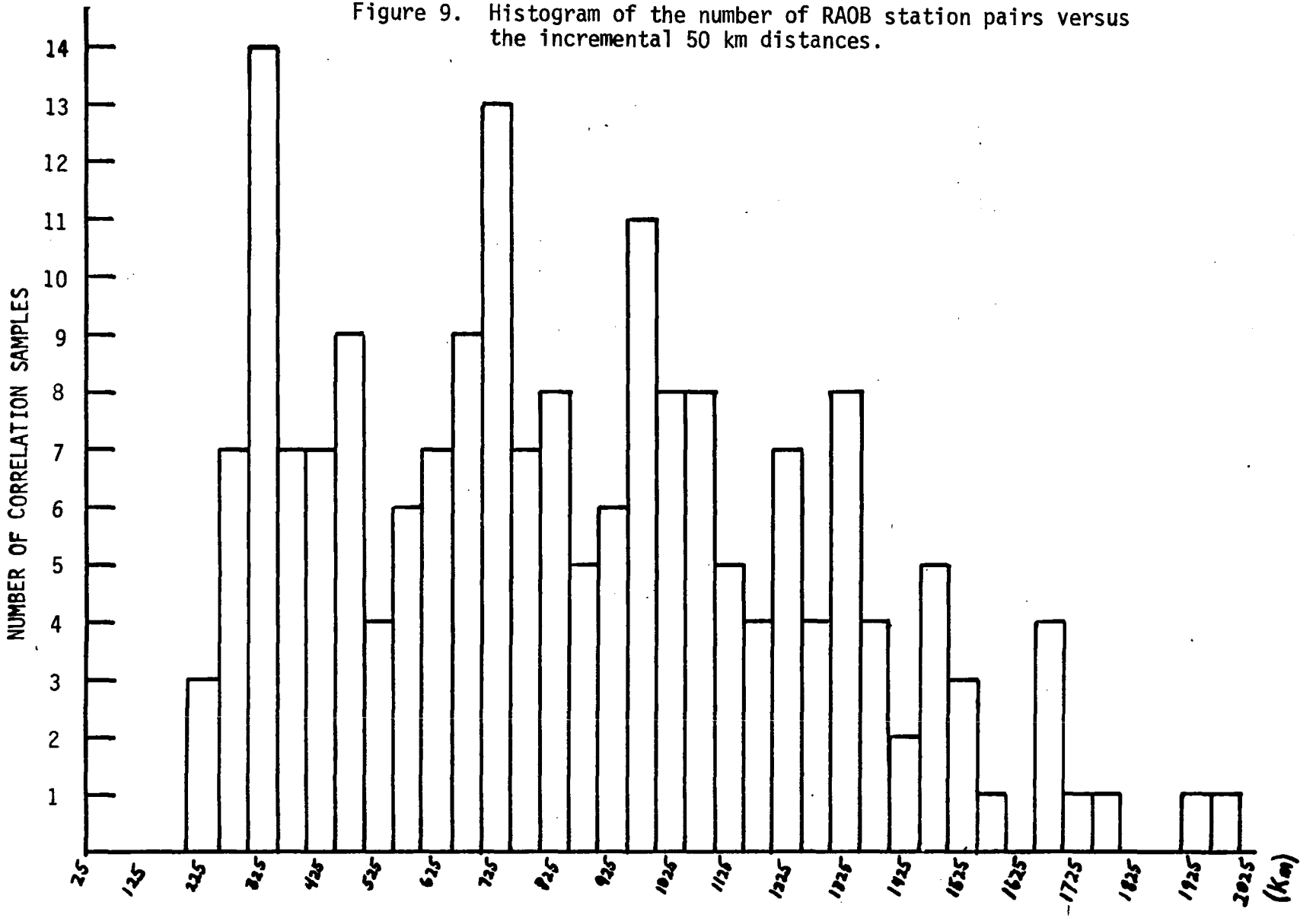


Table 6 summarized the different types of unnormalized structure function models used in curve fitting the sample temperature fields at 700 mb, 500 mb, and 300 mb. The coefficient of determination is the square of the correlation coefficient and gives a measure of the goodness of fit for each correlation value (μ) as a function of station separation (x). For each individual atmospheric layer the "goodness of fit" are rather similar relative to the models chosen. By extrapolating the linear regression and exponential curves to zero distance, an estimate of the mean-square errors can be found. For the 700 mb case, the root mean square error of the observations is $1.5^{o}C$ for the linear extrapolation and $1.6^{o}C$ for the exponential extrapolation, etc. The polynomial coefficients are erratic about the zero orgin so that extrapolation here is useless. The linear extrapolations of the rms errors tend to be slightly lower for the linear curve than for the corresponding exponential curve. We also notice that the rms errors decrease as we progress higher up in the atmosphere. From this additional information on the statistical nature of the errors, the measure of the random observational errors can be calculated by an equation equivalent to equation 15,

$$\eta_i = \frac{2\sigma^2}{\overline{T'^2}}$$

where σ = root-mean-square error in the observational data.

CURVE FITTING MODELS FOR THE UNNORMALIZED STRUCTURE FUNCTIONS
OF TEMPERATURES BY METHOD OF LEAST-SQUARES

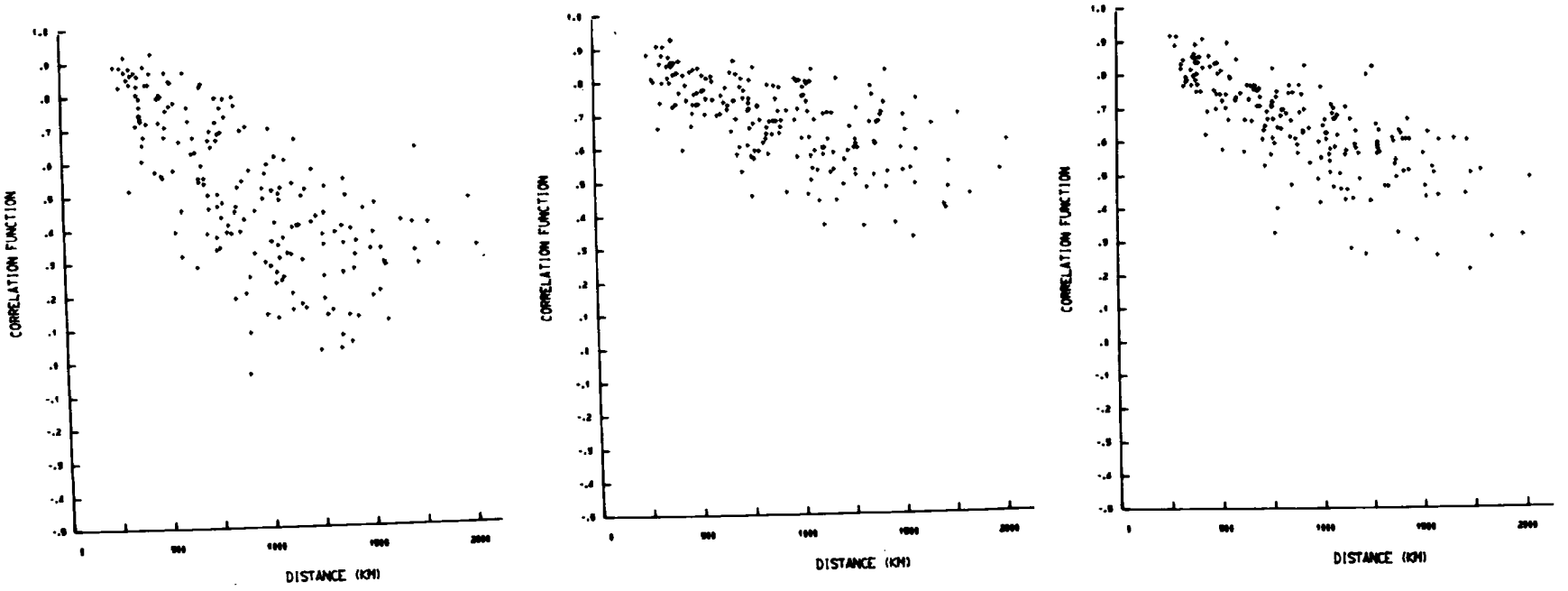
700 mb	Model	Parameter Values	Extrapolated rms Errors	Coefficient of Determination
Linear Regression	$\mu = a_0 + a_1 x$	$a_0 = 4.5281$ $a_1 = 1.1044 \times 10^{-2}$	1.5^o	.330
Exponential	$\mu = a \exp(bx)$	$a = 5.2598$ $b = 9.4360 \times 10^{-4}$	1.6^o	.394
Polynomial	$\mu = a_0 + a_1 x + a_2 x^2 + a_3 x^3$	$a_0 = -4.8931$ $a_1 = 3.5049 \times 10^{-2}$ $a_2 = -1.3649 \times 10^{-5}$ $a_3 = 8.8575 \times 10^{-10}$		.396
500 mb				
Linear Regression	$\mu = a_0 + a_1 x$	$a_0 = 2.5750$ $a_1 = 3.6685 \times 10^{-3}$	1.1^o	.285
Exponential	$\mu = a \exp(bx)$	$a = 2.9258$ $b = 6.5731 \times 10^{-4}$	1.2^o	.318
Polynomial	$\mu = a_0 + a_1 x + a_2 x^2 + a_3 x^3$	$a_0 = 1.5317$ $a_1 = 5.4882 \times 10^{-3}$ $a_2 = 2.3987 \times 10^{-7}$ $a_3 = -6.9704 \times 10^{-10}$		.303
300 mb				
Linear Regression	$\mu = a_0 + a_1 x$	$a_0 = 1.4737$ $a_1 = 3.8194 \times 10^{-3}$	$.9^o$	.600
Exponential	$\mu = a \exp(bx)$	$a = 2.0591$ $b = 8.6419 \times 10^{-4}$	1.0^o	.602
Polynomial	$\mu = a_0 + a_1 x + a_2 x^2 + a_3 x^3$	$a_0 = -5.3670 \times 10^{-1}$ $a_1 = 9.5905 \times 10^{-3}$ $a_2 = -4.2670 \times 10^{-6}$ $a_3 = 8.0383 \times 10^{-10}$		.626

Table 6. Curve fitting models for the unnormalized structure functions of RAOB temperatures by method of least-squares.

A. RADIOSONDE TEMPERATURE AND MOISTURE FIELDS

The structure functions were found primarily to estimate the rms errors of the RAOB measurements at each of the atmospheric pressure levels. The next step is to search for a representative correlation model for the associated normalized correlation functions. In our review of optimum interpolation models, the Gaussian type model was fitted adequately to the geopotentials heights but in our analysis of the temperature field this model did not fit better than the inverse power curve.

In the operational mode of data processing, the normalized correlation functions are averaged within a fixed distance interval regime. This interval cannot be so broad as to smooth out useful statistical relationships but it also cannot be so narrow as to retain the anisotropic scatter inherent in the data. A composite of all the data points used in our analysis is shown by the following scatter diagrams in Figure 10. The normalized correlation functions are plotted against the distances between different pairs of stations for each atmospheric layer. Overall, the best fit for these distance dependent correlations was the inverse power-fit curve charted in Table 7. According to the coefficient of determination we notice that the 300 mb level correlated better than the other two atmospheric levels. Both the anisotropic conditions in the covariance field and the noise in the covariance estimates will lead to a certain amount of scatter as seen in these diagrams. These are the functional forms of the normalized correlations which will be used to find the weighting factors as described in equations 10 and 14.



(a) (b) (c)

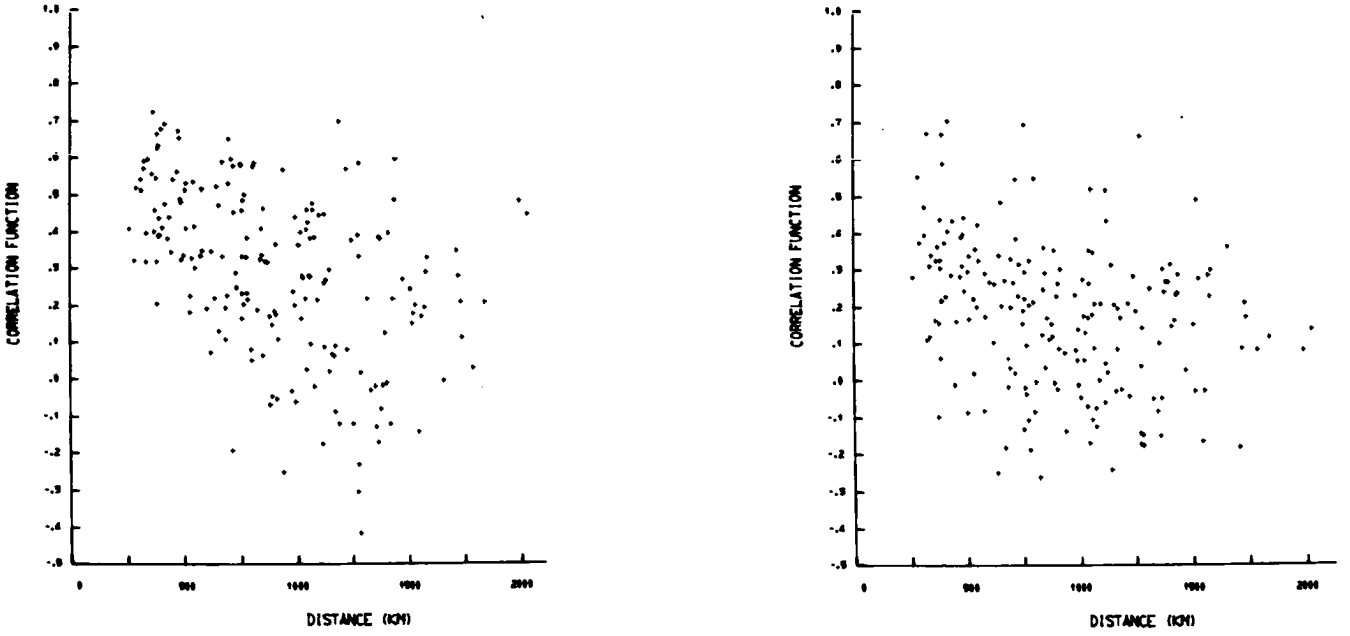
Figure 10. Scatter diagrams of the normalized RAOB temperature correlation functions versus distances for (a) 700 mb, (b) 500 mb, and (c) 300 mb.

MODELS FOR THE NORMALIZED CORRELATION FUNCTIONS
OF TEMPERATURES BY METHOD OF LEAST-SQUARES

700 mb	Model	Parameter Values	Measure of the Random Errors (η)	Coefficient of Determination
Inverse Power Fit	$\mu = 1 - ax^b$	$a = 9.726 \times 10^{-4}$ $b = 9.085 \times 10^{-1}$	.1565	$r^2 = .594$
500 mb				
Inverse Power Fit	$\mu = 1 - ax^b$	$a = 5.274 \times 10^{-3}$ $b = 5.996 \times 10^{-1}$	.1368	$r^2 = .468$
300 mb				
Inverse Power Fit	$\mu = 1 - ax^b$	$a = 2.730 \times 10^{-3}$ $b = 7.138 \times 10^{-1}$	.1177	$r^2 = .616$

Table 7. Model for the normalized correlation function of RAOB temperature by the method of least-squares.

An identical analysis for the structure functions of the 700 mb and 500 mb moisture fields was undertaken and the scatter diagrams are shown in Figure 11. It becomes apparent that even at distances of 250 km the amount of correlation scatter is large. This implies that the average radiosonde grid spacing of approximately 250 km cannot resolve the moisture fields with any reasonable degree of accuracy. Perhaps the dew-point temperature were too dry for the 700 mb and 500 mb layers during this less than 50 percent of normal precipation for May 1976 in a larger portion of the analyzed region. A composite of many seasons should produce a statistically better average of the precipitation and thus better correlation between stations. Any attempts to curve fit the mixing ratio structure function or correlation functions were abandoned. Much interest has been devoted to the objective analysis of humidity whereby precipitation predictions can be anticipated. But when using radiosonde soundings it is a known fact that the transmitted moisture-content information is subject to considerable lag errors. These results lead us to the conclusion that satellite radiance measurements of the total vertical precipitable water at grid-point spacing of about 70 km may provide better correlations than the standard radiosonde type network of stations.



(a) (b)

Figure 11. Scatter diagrams of the normalized mixing ratio functions at (a) 700 mb and (b) 500 mb from RAOB measurements.

B. SATELLITE-DERIVED TEMPERATURE FIELD

Radiance measurements from the NOAA polar-orbiting satellites were converted to vertical temperature profiles (see Appendix B) along the paths shown in Figure 12. From these satellite-derived temperatures for 700 mb, 500 mb, and 300 mb, the norms (see Table 8) were found by taking an area mean over each 50-km interval distance between scan spots. The areal mean-temperature value was chosen instead of the time-averaged norm over a particular location as done in the RAOB analysis because the satellite's sensing path will descend over different earth locations. Since it is unlikely that a region will be perfectly homogeneous and isotropic for one day, the areal norms of fourteen clear-column radiance days in May and June 1976 were averaged over time. On the other hand, RAOB observations are taken at fixed locations so that a climatological or time-averaged mean temperature is more convenient. This discrepancy emphasizes the importance of combining a satellite's areal coverage with the fixed but detailed vertical structure measured by radiosondes.

The normalized correlation functions for the satellite-derived temperatures at the three specified atmospheric levels (Figure 13-15) are taken from work done by Hillger (1977). The solid and dotted lines represent two different cloud correction schemes used in the iterative inverse radiative transfer program. The temperatures from satellite radiances correlate very well for distance separations below 1000 km in all levels as compared with the equivalent radiosonde correlation functions. From Table 9 the coefficient of determination

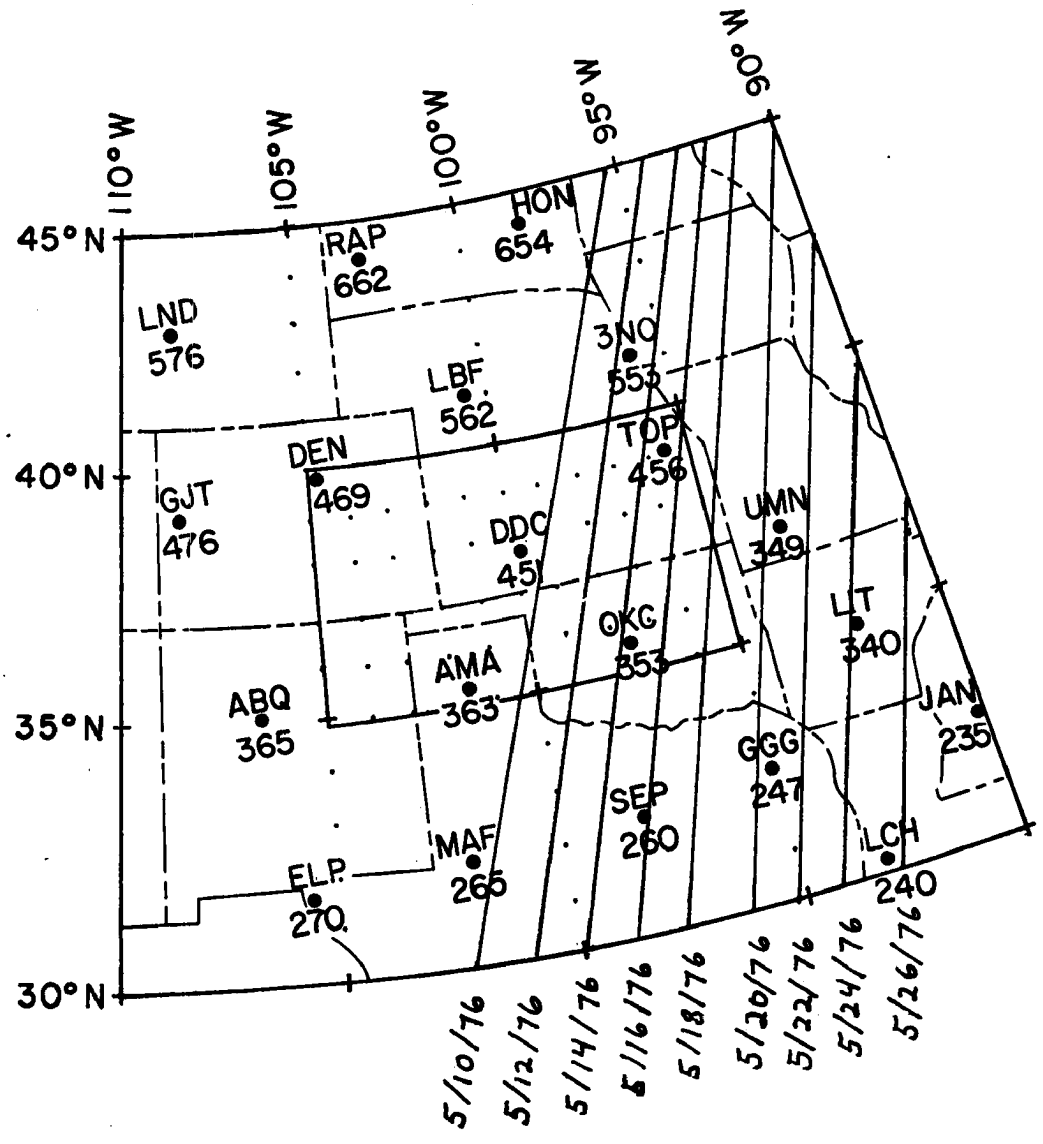


Figure 12. Subsatellite radiance measuring paths from the NOAA-2 satellite during May 1976.

DISTANCE	NUMBER OF SAMPLE PAIRS	700 MB SAT. NORMS (^{o}C)	500 MB SAT. NORMS (^{o}C)	300 MB SAT. NORMS (^{o}C)
75	2856	5.3	-13.6	-40.5
125	3098	5.5	-13.4	-40.5
175	4553	5.4	-13.5	-40.5
225	6035	5.3	-13.5	-40.5
275	5426	5.2	-13.5	-40.5
325	5529	5.4	-13.3	-40.4
375	6998	5.2	-13.4	-40.5
425	5952	5.1	-13.4	-40.4
475	5560	5.3	-13.4	-40.4
525	6358	5.1	-13.4	-40.4
575	5260	5.1	-13.4	-40.4
625	4624	5.2	-13.3	-40.4
675	5056	5.0	-13.5	-40.4
725	4033	5.0	-13.5	-40.4
775	3205	5.1	-13.4	-40.4
825	3079	4.9	-13.6	-40.4
875	2153	4.8	-13.6	-40.4
925	1466	5.0	-13.6	-40.4
975	973	4.8	-13.6	-40.3
1025	501	4.9	-13.7	-40.4

Table 8. Areal distance-averaged norms of satellite-derived temperatures over 50-km intervals at 700 mb, 500 mb, and 300 mb. (distances in km)

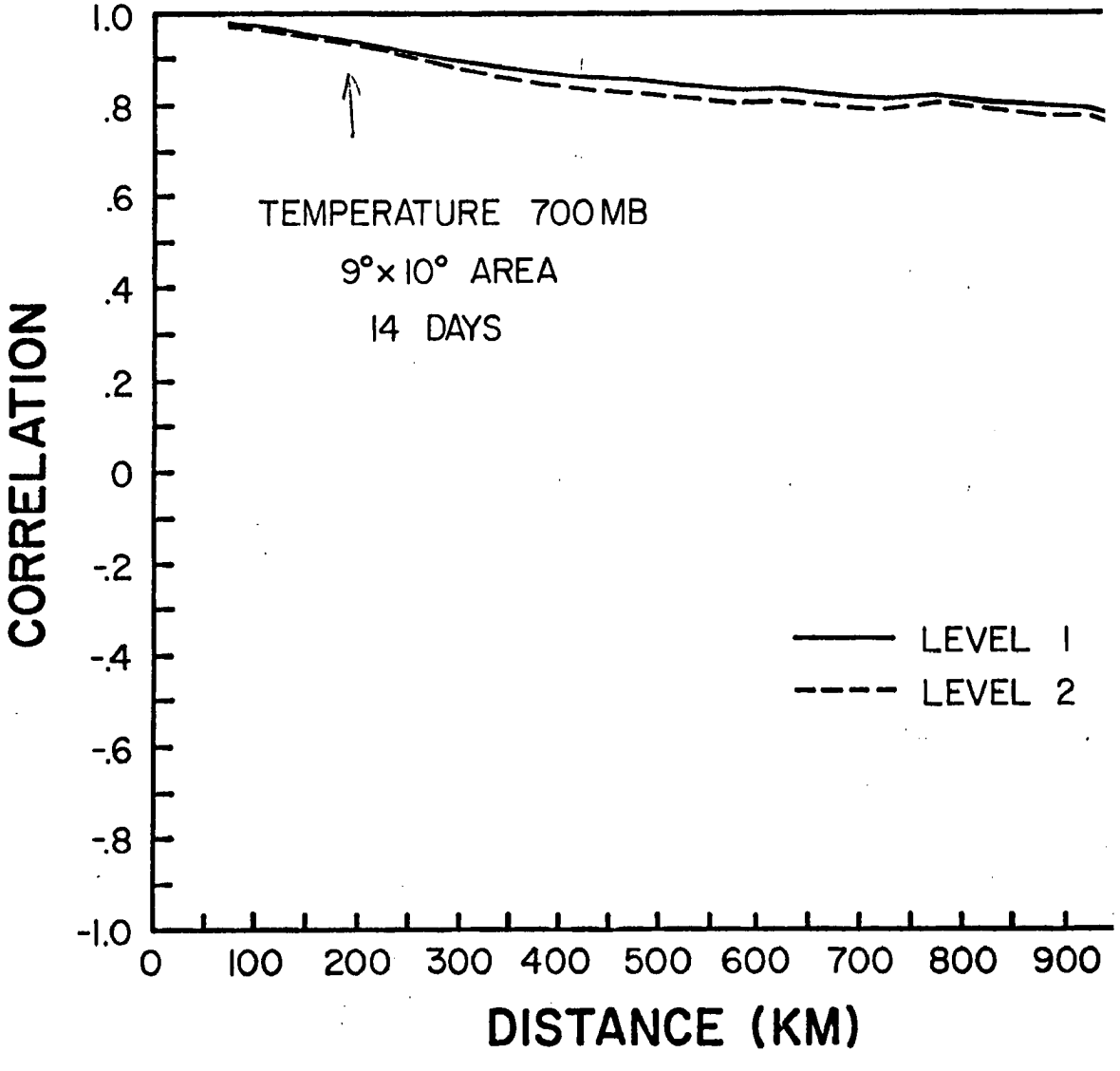


Figure 13. Normalized correlation functions of satellite-derived temperatures at 700 mb. Level 1 and level 2 are two different sets of cloud corrections used in the solution of the inverse radiative transfer equation. (From Hillger, 1977)

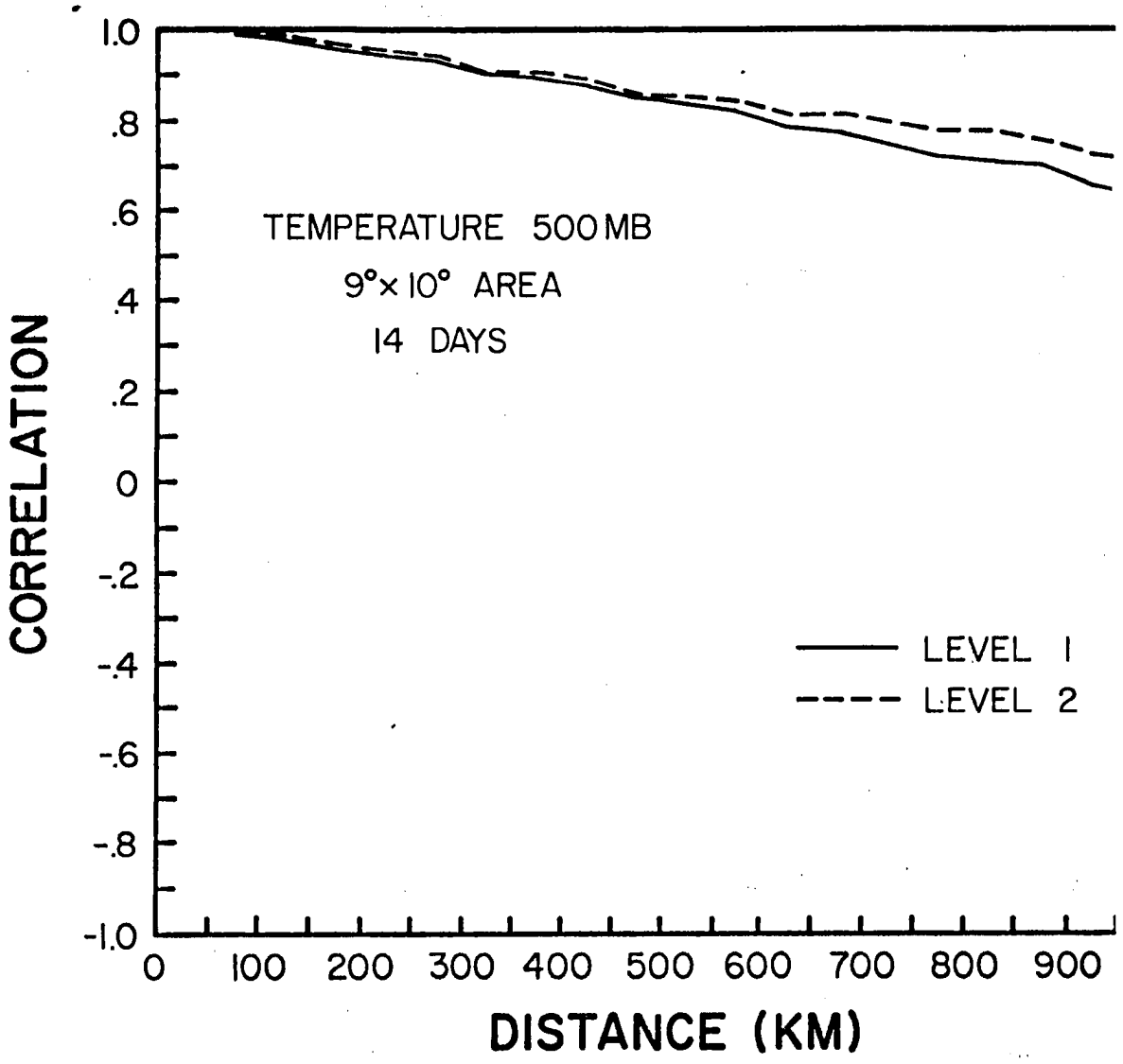


Figure 14. Normalized correlation functions of satellite-derived temperatures at 500 mb. Level 1 and level 2 are described in figure 13. (From Hillger, 1977)

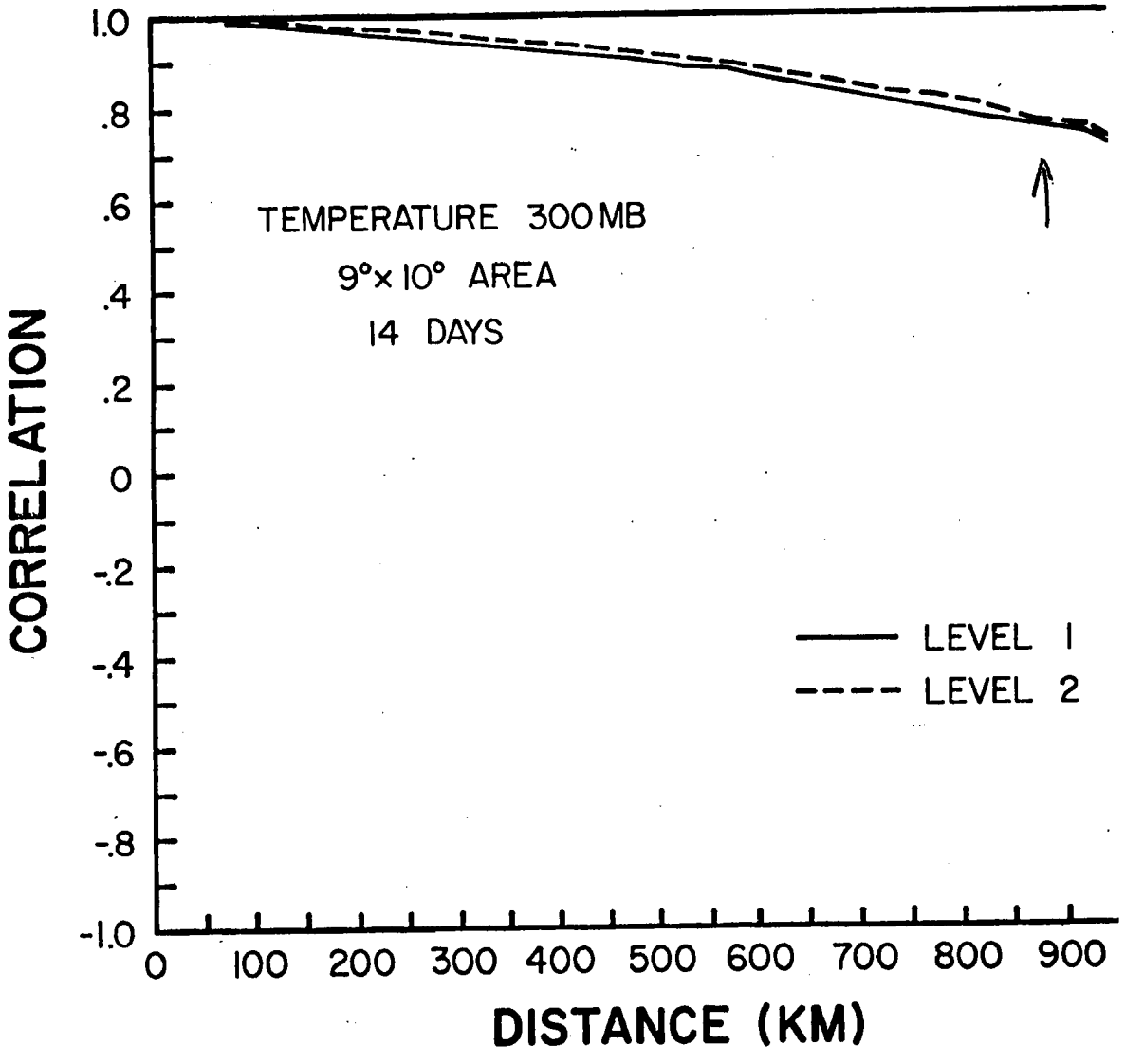


Figure 15. Normalized correlation functions of satellite-derived temperatures at 300 mb. Level 1 and level 2 are described in figure 13. (From Hillger, 1977)

show that these correlation values fit a linear relationship to a high degree.

700 mb	Model	Parameter Values	Measure of the Random Errors (ξ)	Coefficient of Determination
Linear Regression	$\mu = a_o + a_1 x$	$a_o = 0.9642$ $a_1 = -2.331 \times 10^{-4}$	0.592	$r^2 = .910$
500 mb				
Linear Regression	$\mu = a_o + a_1 x$	$a_o = 1.0075$ $a_1 = -2.982 \times 10^{-4}$	0.942	$r^2 = .985$
300 mb				
Linear Regression	$\mu = a_o + a_1 x$	$a_o = 1.0507$ $a_1 = -3.223 \times 10^{-4}$	2.577	$r^2 = .949$

Table 9. Models of the normalized satellite-derived temperature correlations by the method of least-squares.

C. MUTUAL CORRELATION FUNCTIONS BETWEEN RADIOSONDE MEASURED AND SATELLITE-DERIVED TEMPERATURES

As pointed out by Yates (1974) there are major differences in the way temperatures are measured by radiosondes and by satellite. Radiosondes take point measurements as selected levels of the atmosphere. Satellite-derived temperatures are indirectly calculated by the inverse radiative transfer equation which has no unique solution. When converting radiances at various frequency channels into temperature profiles, we are taking volume averaged measurements. As in the case of the NOAA-2 satellite, the approximate area of an averaged temperature over a layer is a 70 km x 70 km scan spot at the subsatellite point. From this, it is not surprising that satellite-radiosonde comparisons prove to be inadequate.

Bruce et al (1977) measured the variability of RAOB temperatures as a function of interstation separation distance for eleven layers of the atmosphere over White Sands Missile Range. The averaged radiosonde errors for all distances (<160 km) at 700 mb was about .6^{o}K and at 500 mb it was about .65^{o}K and at 300 mb it was approximately .7^{o}K. Their results for comparing the satellite's area averaged temperature with that from the RAOB's are shown by Figure 16 for three different situations. The first two error profiles are for radiosonde measurements made within a satellite's area of observation (about 100 km). The expected difference is around 1^{o} to 1.5^{o}K. The last profile exemplifies a satellite-centered and radiosonde observation separation of 200 km. The temperature error for 200 km is between 1.5^{o}K and 2.0^{o}K. If there is a constant bias between the RAOB's and satellite of over 2^{o}K, we must

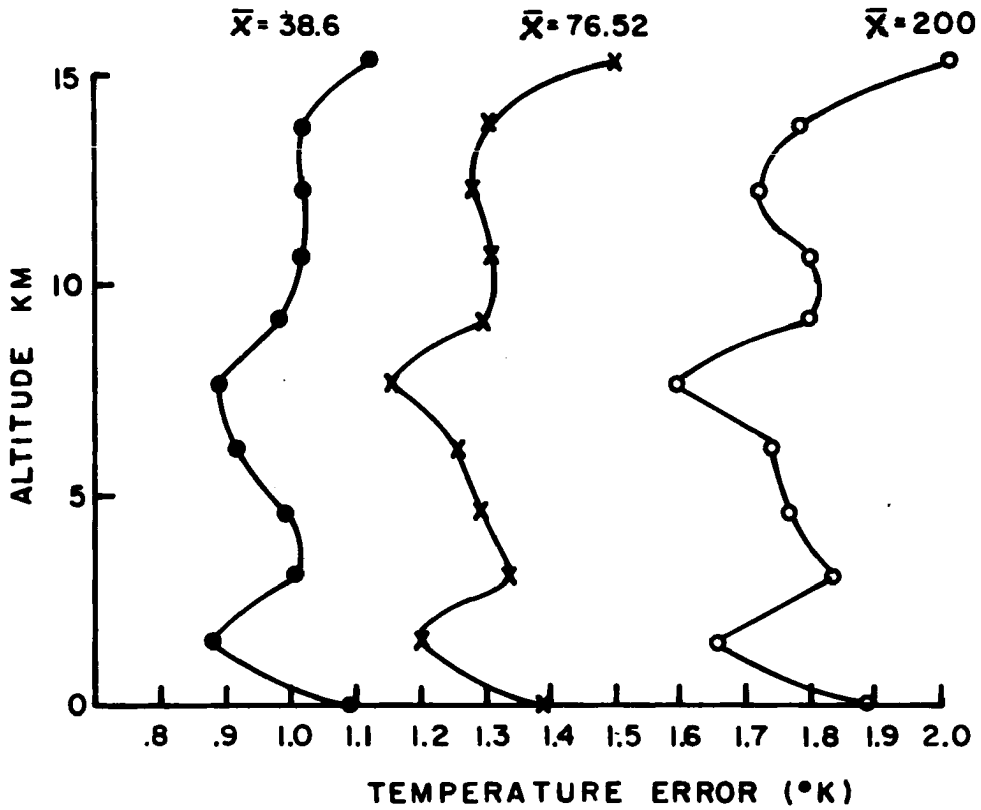


Figure 16. Expected temperature comparison difference between a satellite centered observation and a radiosonde measurement for three mean distances, $\bar{x}$ in km. (Taken from Bruce etal, 1977)

attribute this error to the various techniques involved in solving for temperature profiles from the raw radiances.

The mutual normalized correlation functions necessary for the final stage of the objective analysis (i.e. the matching of information from both satellites and radiosondes) were taken from one test day on May 18, 1976. Satellite passes on this date were at 1600Z and the local afternoon radiosonde measurements were at 0000Z. As in the previous correlation functions, the mutual correlations were distance-averaged over 50-km intervals. A histogram of the distribution versus each distance interval is given by Figure 17. The associated mutual correlation graphs are plotted in Figure 18 for the 700 mb to 500 mb, and 300 mb levels, respectively. At 700 mb the mutual correlations drop off very rapidly over 500 km averaged distances. Information content was sparse after 1500 km so that these averaged distance interval correlations were omitted from the curve fitting routines. At 500 mb the mutual correlation was high for small distances but after 500 km random fluctuations seem to prevail. At 300 mb the linear correlation model resembled that at 700 mb.

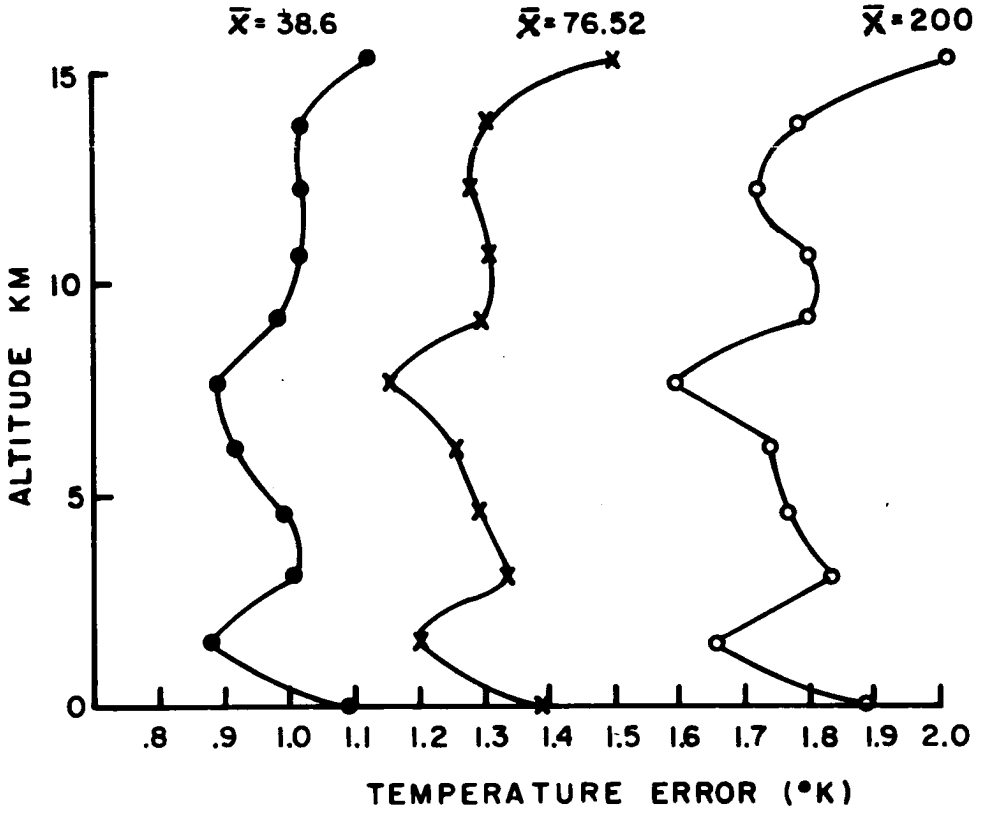


Figure 16. Expected temperature comparison difference between a satellite centered observation and a radiosonde measurement for three mean distances, $\bar{x}$ in km. (Taken from Bruce etal, 1977)

attribute this error to the various techniques involved in solving for temperature profiles from the raw radiances.

The mutual normalized correlation functions necessary for the final stage of the objective analysis (i.e. the matching of information from both satellites and radiosondes) were taken from one test day on May 18, 1976. Satellite passes on this date were at 1600Z and the local afternoon radiosonde measurements were at 0000Z. As in the previous correlation functions, the mutual correlations were distance-averaged over 50-km intervals. A histogram of the distribution versus each distance interval is given by Figure 17. The associated mutual correlation graphs are plotted in Figure 18 for the 700 mb to 500 mb, and 300 mb levels, respectively. At 700 mb the mutual correlations drop off very rapidly over 500 km averaged distances. Information content was sparse after 1500 km so that these averaged distance interval correlations were omitted from the curve fitting routines. At 500 mb the mutual correlation was high for small distances but after 500 km random fluctuations seem to prevail. At 300 mb the linear correlation model resembled that at 700 mb.

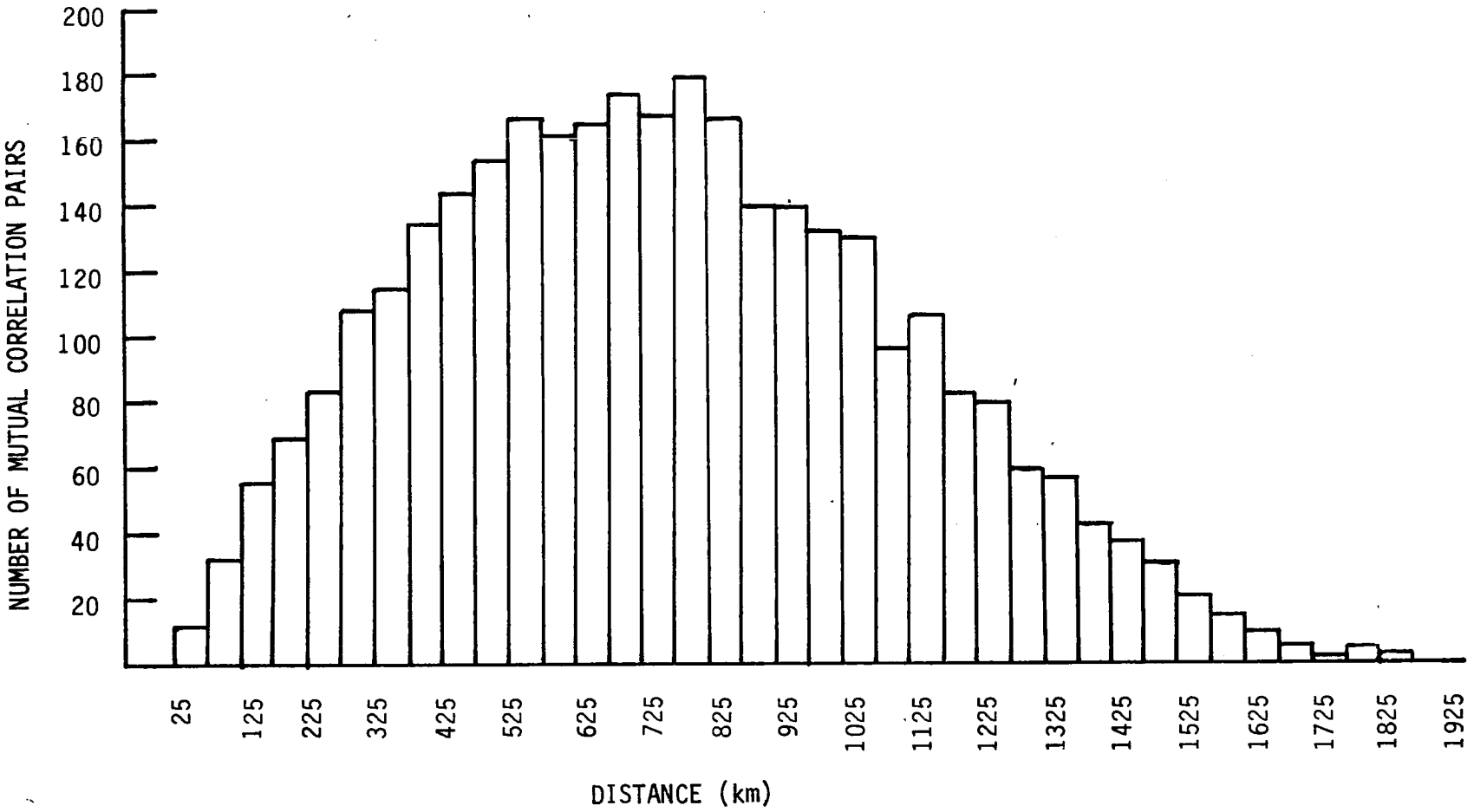
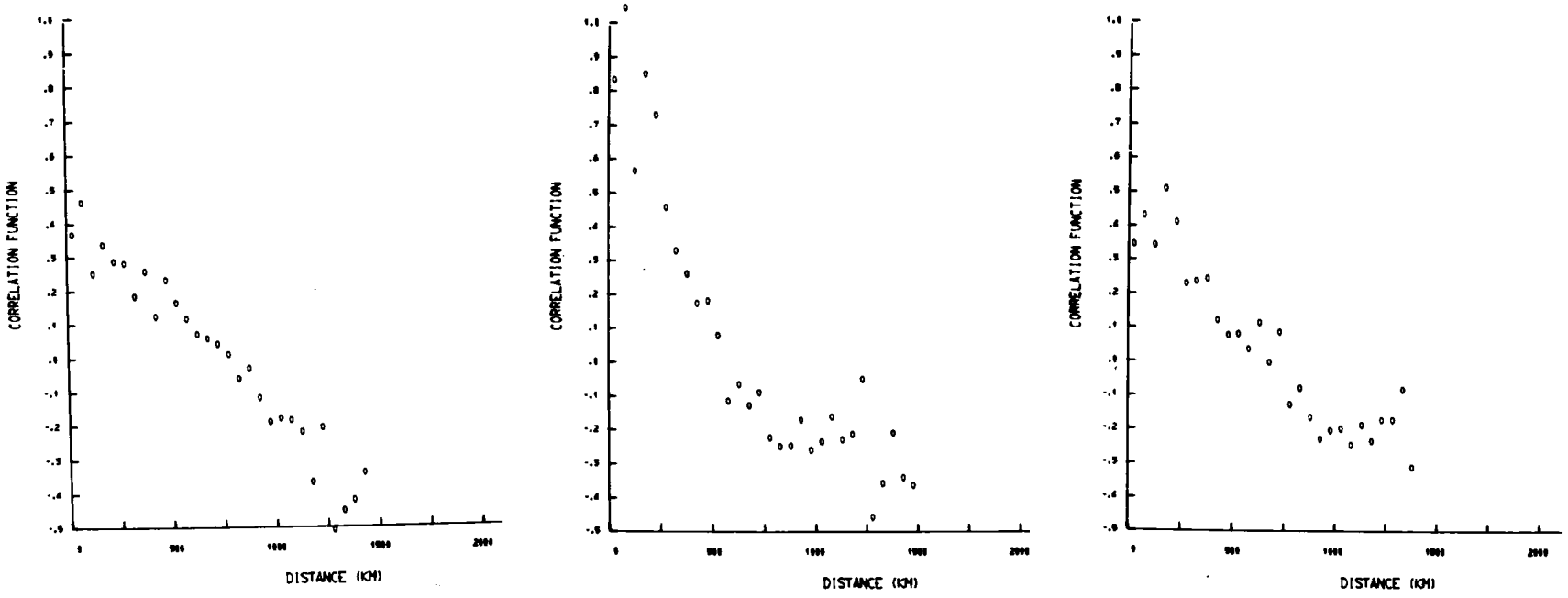


Figure 17. Histogram of the number density of mutual correlation pairs versus the distance-averaged interval (50 km).



(a) (b) (c)

Figure 18. Normalized mutual correlation functions averaged over 50-km intervals for (a) 700 mb, (b) 500 mb, and (c) 300 mb.

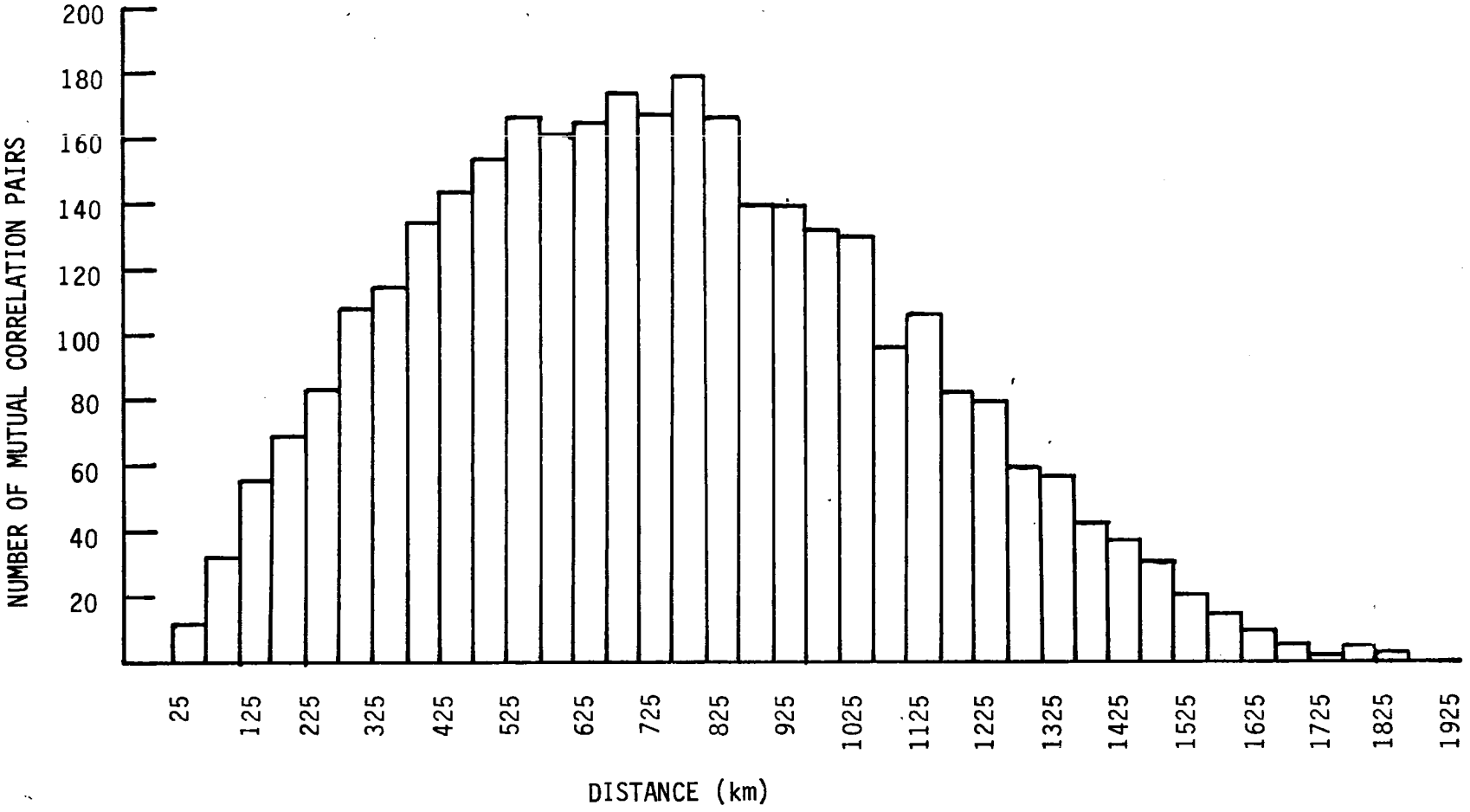


Figure 17. Histogram of the number density of mutual correlation pairs versus the distance-averaged interval (50 km).

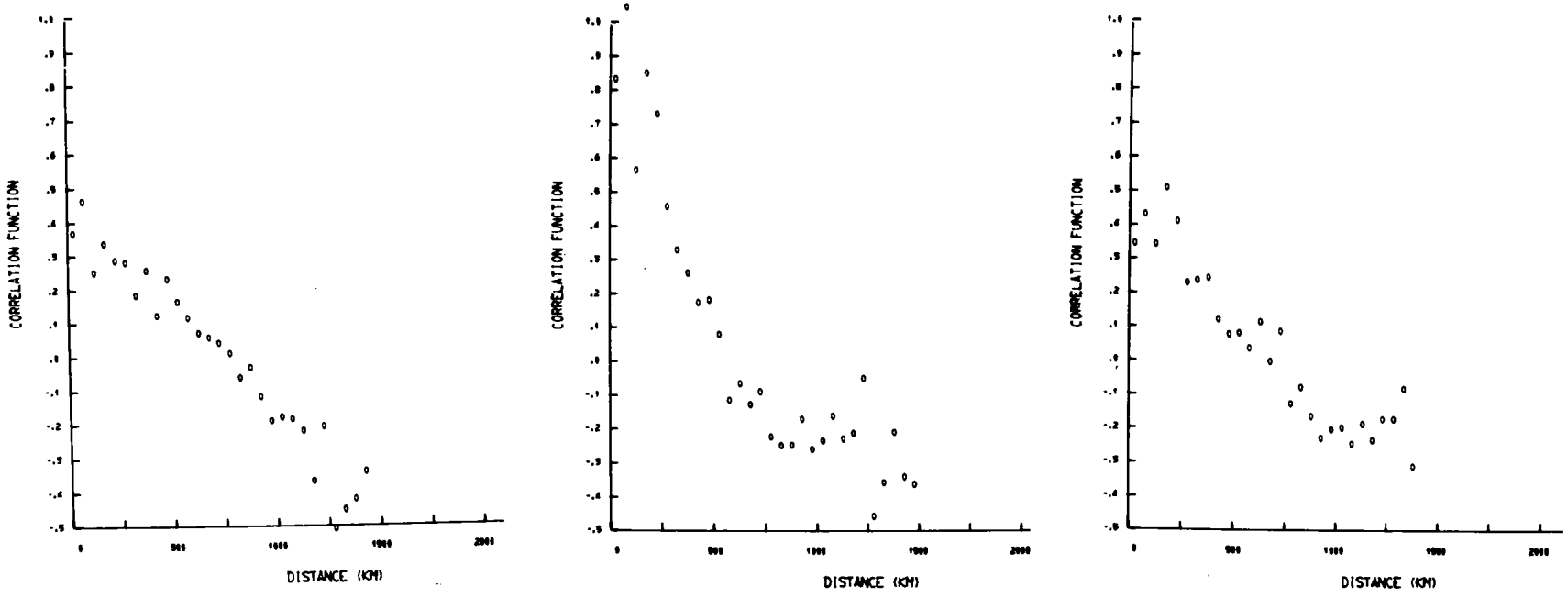


Figure 18. Normalized mutual correlation functions averaged over 50-km intervals for (a) 700 mb, (b) 500 mb, and (c) 300 mb.

VI. OPTIMUM MATCHING OF RAOB AND SATELLITE TEMPERATURE DATA

After an empirical formula for the normalized correlation functions has been found for both the radiosonde data and the satellite data, the next step will be the matching of these two meteorological fields. In this way the merging of the two data sets will provide the best analysis of the temperatures and mixing ratios over a particular region. This is an important phase in objective analysis since it allows a more accurate representation of a meteorological element based upon certain relationships with other meteorological elements.

Previous investigators have tried to use additional information to analyze the geopotentials of isobaric surfaces. Gilchrist and Cressman (1954) using their polynomial method matched the observed geopotentials with the wind fields assuming geostrophic conditions. Briefly, the geopotential at a certain grid point (Φ_0) is expressed as a linear combination of the observed geopotential (Φ_1) and of the wind components (u_i and v_j) in the general vicinity of the grid point:

$$\Phi_0 = \sum_{i=1}^{n} b_i \Phi_i + \sum_{j=1}^{m} (c_j u_j + d_j v_j) \qquad (22)$$

Once they are determined, the weighting factors b_i, c_j, d_j will not be dependent upon the observational data.

A different approach to the matching of meteorological fields was done by Sasaki (1958) who employed calculus of variation techniques in two case studies for his objective analysis. One case is based upon the assumption of quasi-geostrophic and thermal wind conditions. The other case assumes barotropic flow which evokes the conditions of non-

divergence and of the balance equation. The short comings in this method are the relatively complicated procedures for finding a solution and the arbitrariness of the weighting factors. Differing from these schemes, optimum interpolation makes use of information concerning the statistical nature of two or more or two different measurements of the same element by minimizing the mean-square interpolation error. Thus, when a matching of meteorological fields is desired, it also requires a minimization of the mean-square error between the true value and a linear combination of two interrelated fields.

Consider two meteorological elements T^s and T^r where T^s is the derived temperature from satellite radiances and T^r is the directly measured temperature from radiosonde observations. These two temperature values are not exact and therefore, they must contain a certain amount of instrumental errors designated by ΔT^s and ΔT^r. Let

ΔT^r = Error in determining radiosonde temperatures.

ΔT^s = Error in determining satellite temperatures.

We assume that the elements T^r and T^s are correlated to each other by a functional relationship

$$F_1(T^r) = F_2(T^s),$$

where $F_1(T^r)$ and $F_2(T^s)$ are the function elements to be matched. In the more specific case, we are concerned only with the simple relationship $T^r=T^s$ provided we are able to find the correct correlation coefficient. In passing on, it is only necessary to match one element field, say T^s, to the field T^r since the solution matrix is perfectly symmetrical with respect to the two fields, T^r and T^s.

Since we will be fixing the number of data points influencing any particular grid point in our network, we shall derive our field matching using six RAOB data points from the closest observing stations. As for the satellite observations, we shall use only the four nearest satellite-derived temperatures that fall within a radius of 100 km over each individual grid point.

Let's derive the interpolated element perturbation for one represented grid point location. We assume that the values of T^r at six points (denoted by T^s) are known. From these data points we want to find the value of T^r at the grid point (denoted by subscript 0). The grid point can be represented as a linear combination of the T_i^r's and T_k^s's

$$T_0^r = \sum_{i=1}^{6} a_i T_i^r + b_1 T_1^s \tag{23}$$

where the a_i's and b_k's can be found by minimizing the mean-square error

$$E = \overline{\{T_0^r - (\sum_{i=1}^{6} a_i T_i^r + b_1 T_1^s)\}^2} \tag{24}$$

or

$$E = \overline{(T_0^r)^2} - 2\sum_{i=1}^{6} a_i \overline{T_0^r T_i^r} - 2b_1 \overline{T_0^r T_1^s} + \sum_{i=1}^{6}\sum_{j=1}^{6} a_i a_j \overline{T_i^r T_j^r}$$

$$+ b_1^2 \overline{(T_1^s)^2} + 2\sum_{i=1}^{6} a_i b_1 \overline{T_i^r T_1^s} \tag{25}$$

and substituting equations 13, we get

$$E = \overline{(T_0^r)^2} - 2\sum_{i=1}^{6} a_i \overline{T_0^r(T_i^r + \Delta T_i^r)} - 2b_1 \overline{T_0^r(T_1^s + \Delta T_1^s)}$$

$$+ \sum_{i=1}^{6}\sum_{j=1}^{6} a_i a_j \overline{(T_i^r + \Delta T_i^r)(T_j^r + \Delta T_j^r)} + b_1^2 \overline{(T_1^s + \Delta T_1^s)^2}$$

$$+ 2\sum_{i=1}^{6} a_i b_1 \overline{(T_i^r + \Delta T_i^r)(T_1^s + \Delta T_1^s)}. \qquad (26)$$

where T_i^r or T_k^s is the error in $\bar{T}_i^r$ or T_k^s, respectively.

Assuming all the errors are completely random, we can evoke the following relationships.

$$\overline{(\Delta T_i^r)(\Delta T_j^s)} = \overline{(\Delta T_i^r)(T_k^s)} = \overline{(\Delta T_k^s)(T_i^r)} = \overline{(\Delta T_k^s)(T_i^s)} = 0 \qquad (27)$$

$$\overline{(\Delta T_i^r)(\Delta T_j^r)} = 0 \qquad (i \neq j) \qquad (28)$$

$$(\Delta T_k^s)(\Delta T_1^s) = 0 \qquad (k \neq 1) \qquad (29)$$

$$(\Delta T_i^r)(\Delta T_k^s) = 0 \qquad (30)$$

Applying equations (27-30) to equation 26, the mean square error is

$$E = \overline{(T_0^r)^2} - 2\sum_{i=1}^{6} a_i \overline{T_0^r T_i^r} - 2b_1 \overline{T_0^r T_1^s} + \sum_{i=1}^{6}\sum_{j=1}^{6} a_i a_j \overline{T_i^r T_j^r}$$

$$+ \sum_{i=1}^{6} a_i^2 \sigma_i^2 + b_1^2 \overline{(T_1^s)^2} + b_1^2 \tau_1^2 + 2\sum_{i=1}^{6} a_i b_1 \overline{T_i^r T_1^s} \qquad (31)$$

where $\sigma_i^2 = (\Delta T_i^r)^2$ and $\tau_1^2 = (\Delta T_1^s)^2$ (32)

are the mean-square error. In all our analysis the values we are dealing with are deviation from the norm rather than the element value. As discussed earlier most meteorological fields can be considered to be somewhat homogeneous and isotropic. If not, the second order moments (variances) between any two points are equal, i.e.

$$\overline{(T_i^r)^2} = \overline{(T_j^r)^2} = \overline{(T^r)^2} \tag{33}$$

$$\overline{(T_i^s)^2} = \overline{(T_j^s)^2} = \overline{(T^s)^2} \tag{34}$$

and the covariances will depend only on the distance apart two points.

For convenience, we will calculate the normalized autocorrelation function for T^r as

$$\mu_{ij} = \frac{\overline{T_i^r T_j^r}\ \text{(covariance)}}{\overline{(T^r)^2}\ \text{(variance)}} \tag{35}$$

and the normalized autocorrelation function for T^s as

$$\nu_{ii} = \frac{\overline{T_i^s T_i^s}}{\overline{(T^s)^2}} \tag{36}$$

Since only one satellite observation will be used per grid point interpolation, $\nu_{ii} = \nu_{ll} = 1$. The normalized mutual correlation function for T^r and T^s will be represented as

$$r_{il} = \frac{\overline{T_i^r T_l^s}}{\{\overline{(T^r)^2}\,\overline{(T^s)^2}\}^{\frac{1}{2}}} \tag{37}$$

and the normalized mean-square errors or measures of error T^r and T^s as

$$\eta_i = \frac{\sigma_i^2}{\overline{(T^r)^2}} \qquad \text{and} \qquad \xi_l = \frac{\tau_l^2}{\overline{(T^s)^2}} \ . \tag{38}$$

The measure of error for matching is

$$\varepsilon = \frac{E}{\overline{(T^r)^2}} \tag{39}$$

The ratio of the variances of T^r and T^s is

$$\lambda^2 = \frac{\overline{(T^s)^2}}{\overline{(T^r)^2}} \tag{40}$$

With these new definitions, equation 39 can be written as

$$\varepsilon = 1 - 2\sum_{i=1}^{6} a_i \mu_{0i} - 2 b_1 r_{01} + \sum_{i=1}^{6}\sum_{j=1}^{6} a_i a_j \mu_{ij} + \sum_{i=1}^{6} a_i^2 \eta_i$$

$$+ \lambda^2 b_1^2 \nu_{11} + \lambda^2 b_1^2 \xi_1 + 2\lambda \sum_{i=1}^{6} a_i b_1 r_{i1} \quad . \tag{41}$$

To solve for the unknown coefficients a_i's and b_k's we must minimize ε by finding the partial derivatives of ε with respect to a_i and b_k.

$$\frac{\partial \varepsilon}{\partial a_i} = 0 \qquad \frac{\partial \varepsilon}{\partial b_k} = 0 \tag{42}$$

By doing this, we get the following set of seven linear equations:

$$\sum_{j=1}^{6} \mu_{ij} a_j + r_{i1} b_1 + \eta_i a_i = \mu_{0i} \qquad (i = 1,\ldots,6) \tag{43}$$

$$\sum_{i=1}^{6} r_{i1} a_i + \lambda \nu_{11} b_1 + \lambda \xi_1 b_1 = r_{01} \tag{44}$$

Also the measure of error for matching can be simplified as

$$\varepsilon = 1 - \sum_{i=1}^{6} a_i \mu_{0i} - \lambda b_1 r_{01} \quad . \tag{45}$$

VII. RESULTS AND DISCUSSION

The preliminary tests were run without estimating the instrumental errors from the statistical information and without inserting these errors into the final analysis. The weighting factors for the idealized case assuming perfect measurements fluctuated excessively about zero producing some unrealistically large negative factors. According to theory if the meteorological elements were measured absolutely without error, any additional data from other true measuring stations will always make the objective analysis better. But since radiosondes and satellites are not perfect-measuring instruments we must expect instability in the solution of simultaneous equations (eq. 11). In lieu of these results, we included the instrumental errors along the diagonal of the normalized correlation matrix. The resulting weighting factors were more evenly distributed for each grid point and the small negative values did not contribute significantly to the interpolated temperatures.

For comparison purposes, we computed the optimum interpolation temperatures for the observed radiosonde data on May 18, 1976 using the previous calculated weighting facotrs p_i's. May 18th was conveniently picked such that the NOAA-2 satellite was able to take measurements over the same defined network region. If the satellites radiometer failed to make a measurement over a grid point, the computer program will automatically interpolate the temperature from the six closest radiosonde stations available, as in the previous analysis. That is, it will ignore the last weighting factor of equation 23 pertaining to any satellite information over that grid point and calculate only the six original weighting factors from the RAOB correlation functions.

The situation in the Northern Hemisphere for May 1976 was generally chaotic. The exception was the mean 700 mb ridge over western Northern America together with the trough to its east (Figure 19). During this period in our area of study, the surface temperature departures from a 30-year mean were $1^{o}C$ to $3^{o}C$ below the average. This would suggest that our computed climatological norm temperatures for this month are not entirely representative of the true homogeneous and isotropic nature existing within this region. Therefore, more seasonal data is required as input into the optimum interpolation model.

The synoptic conditions for the case study day (May 18, 1976) is shown by the following daily weather map (Figure 20). A cold front extending from the eastern side of Nevada to the western portions of the Dakotas is proceeding eastward into a high pressure area over Kansas and Missouri. There is ridging over the Mid-west and a trough over the Great Lakes at 500 mb.

In Table 10 the correlation coefficient of the mesoscale radiosonde interpolated temperatures versus the combined satellite-radiosonde analysis for test day May 18, 1976 are quite high. The covariance of the 300 mb level is small compared to the other atmospheric levels implying a very stable regime where the maximum temperature gradient across the network is only about $2^{o}C$ from ABQ to TOP. Here the differences between the two analysis were minute, emphasizing the satellite's sensitivity at 300 mb (Figure 21 and 22). The fact that both types of interpolation at 300 mb were nearly identical suggests that the satellite data at this atmospheric level contributes little additional information. This is partly due to the relatively stable temperatures at 300 mb and

to the measuring instruments' limiting rms errors. There is the possibility that measuring the upper level temperatures, which are free from excessive clouds, can be made exclusively by satellite. At 500 mb the temperature gradients for the RAOB-only case are stronger than the corresponding match RAOB-Satellite case (Figure 23 and 24). Looking at Figures 25 and 26 for the 700 mb level the opposite is true. The satellite measurements made the 700 mb temperature gradients greater than the RAOB-only case. In the latter two analysis, the satellite-derived temperatures helped to bring out the very small mesoscale features within the field of study. This is a prime example of using satellite data to bring out the minute small scale changes in the 2-dimensional horizontal area. The satellite temperatures tend to raise the mesoscale field by $1^{o}C$ whenever the VTPR scanner passed over a grid point. Nevertheless, the differences are well below the expected rms errors of either the radiosonde or satellite-determined temperatures.

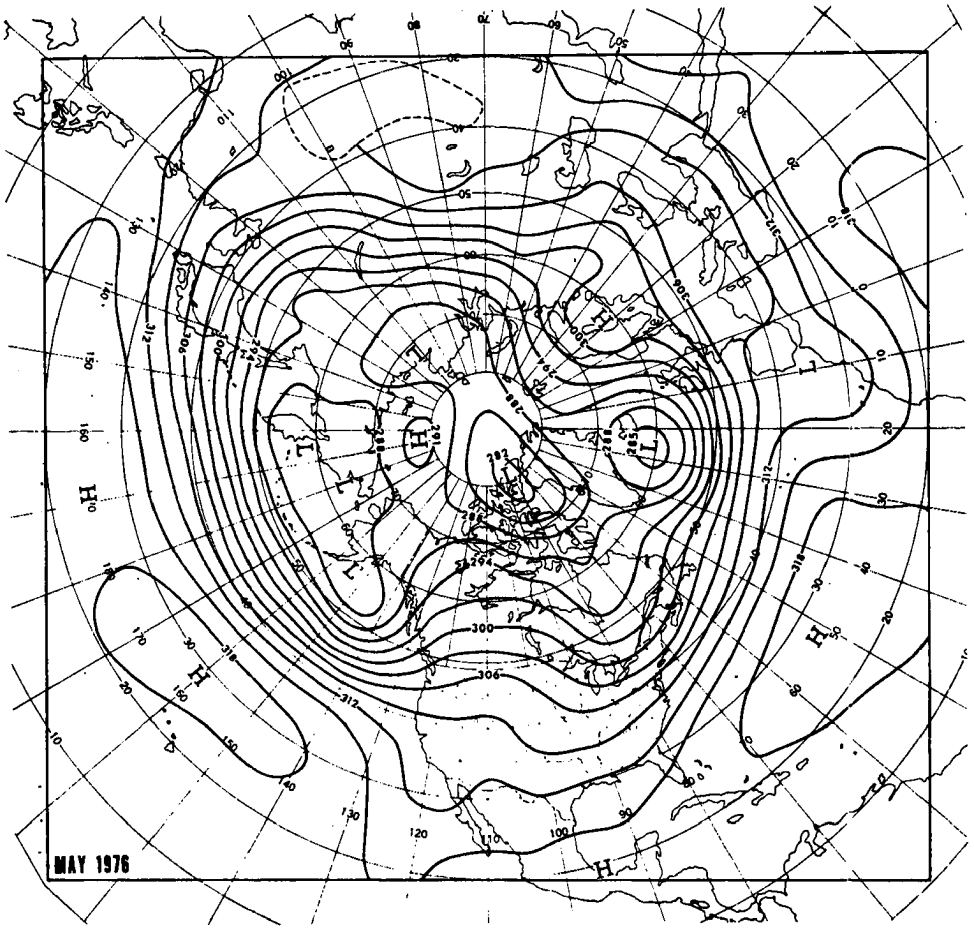


Figure 19. Mean 700 mb height contours (dam) for May 1976.

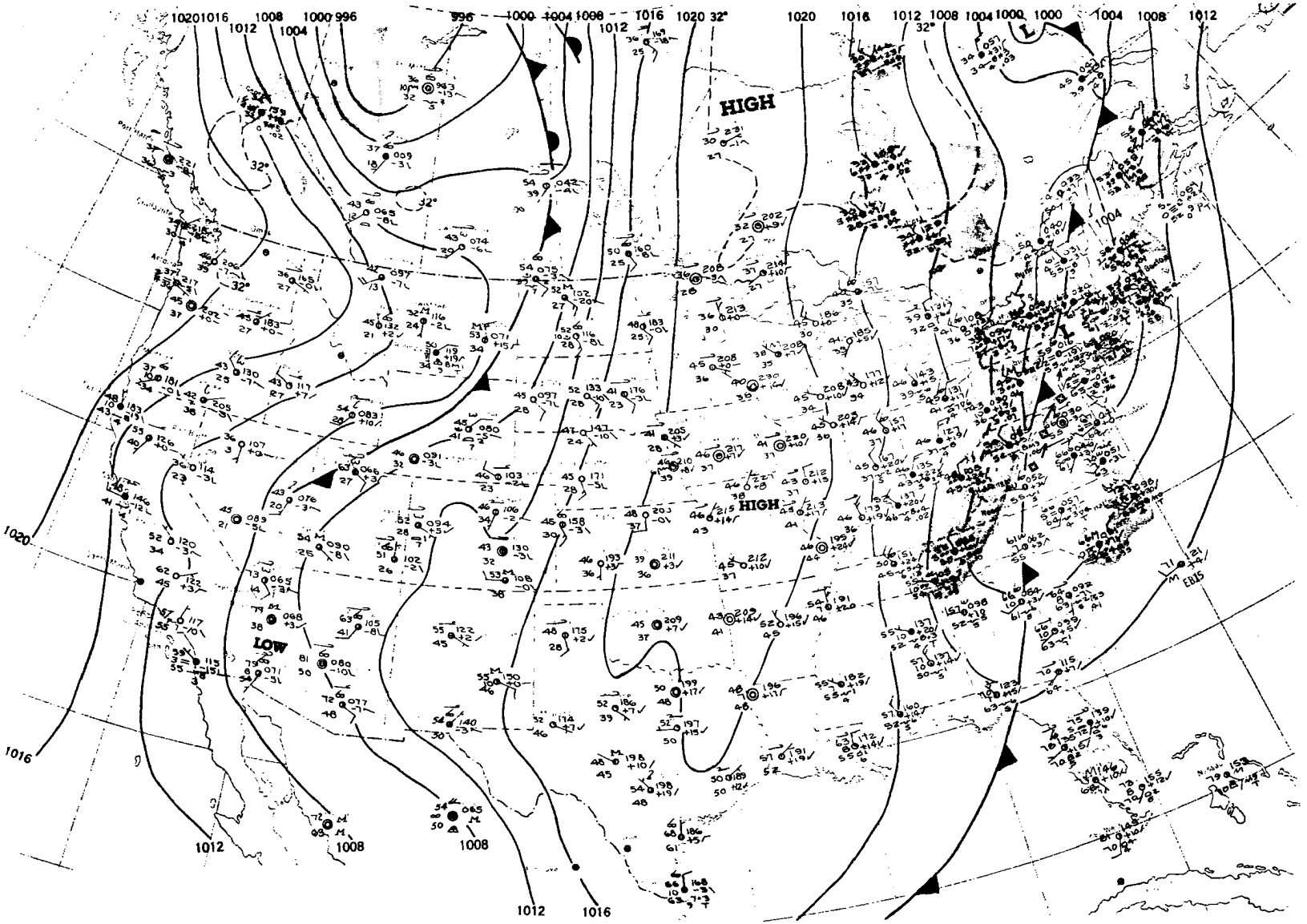


Figure 20. Surface weather map for test day, May 18, 1976 at 7: a.m. E.S.T.

	Number of Matching Grid Points	Covariance	Correlation Coefficient
700 mb	49	5.19	.980
500 mb	49	2.44	.997
300 mb	49	.39	.992

Table 10. Correlation between mesoscale RAOB interpolated temperatures and satellite-RAOB interpolated temperatures for 18 May, 1976.

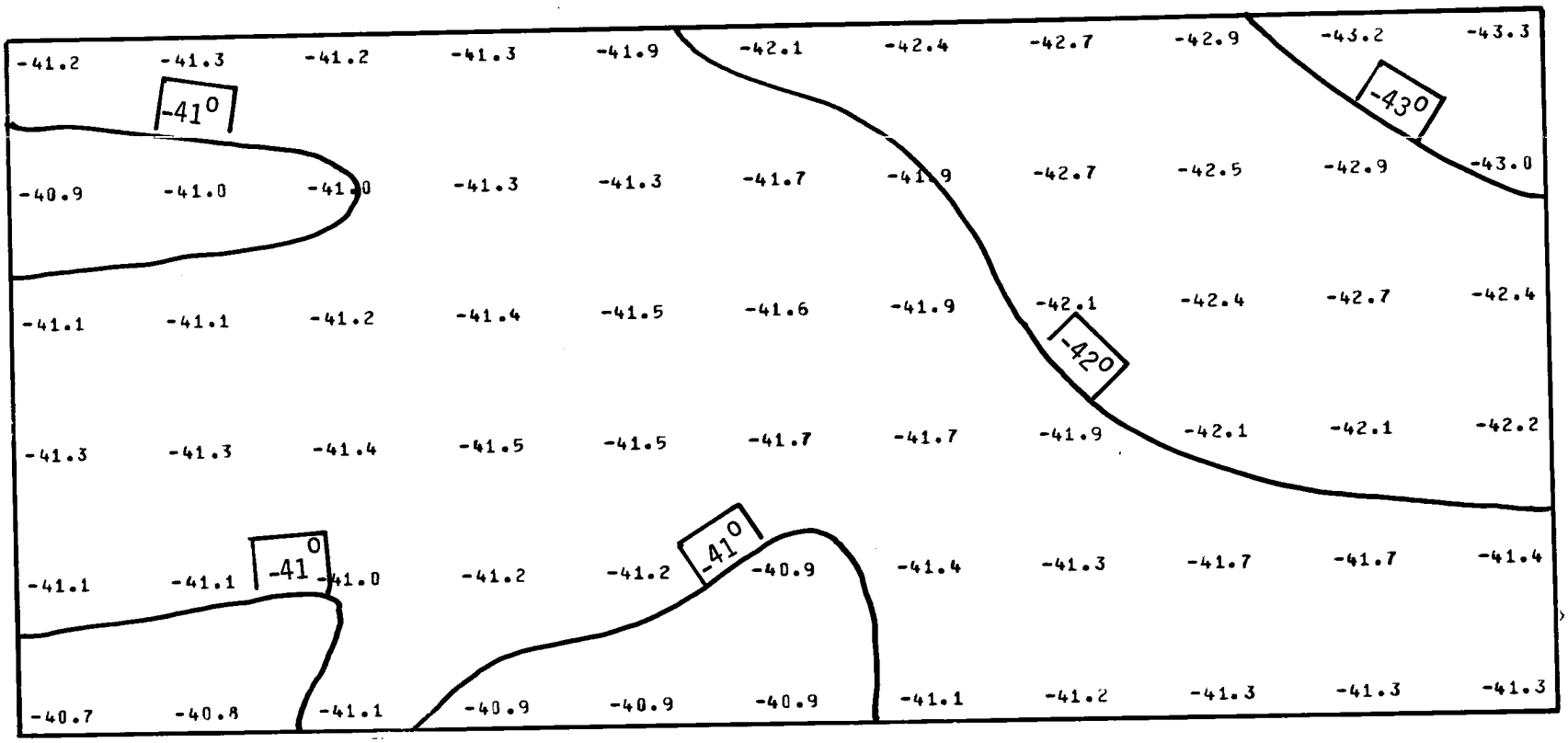


Figure 21. Interpolated radiosonde temperature values (^{o}C) for each latitude-longitude grid point in the 40^{o}-35^{o} Lat. and 105^{o}-95^{o} Long. network for 300 mb.

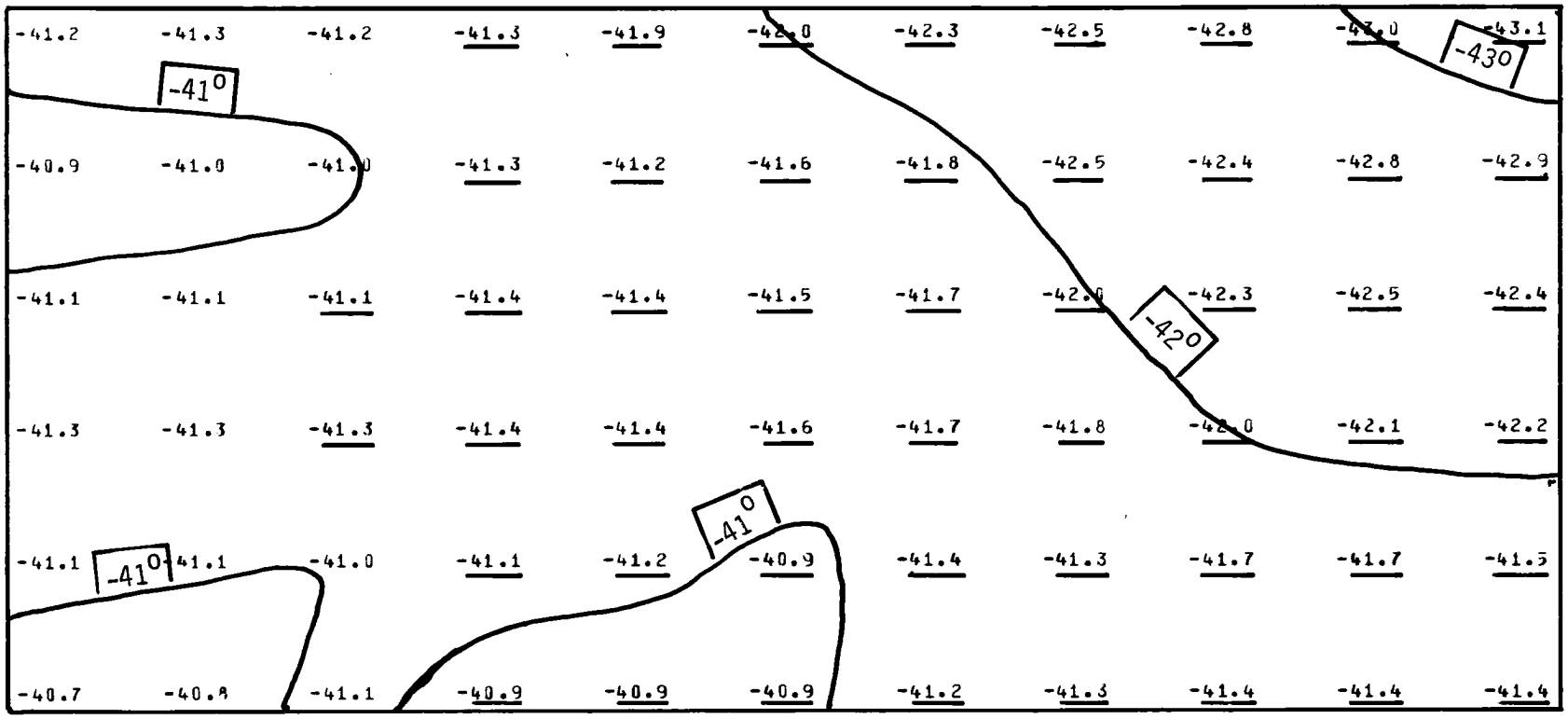


Figure 22. Interpolated radiosonde-satellite derived (^{o}C) for each latitude-longitude grid point in the 40^{o}-35^{o} Lat. and 105^{o}-95^{o} Long. network for 300 mb. (Under lined temperature represent grid points in which the NOAA-2 radiometer was able to make scan-spot radiance measurements.)

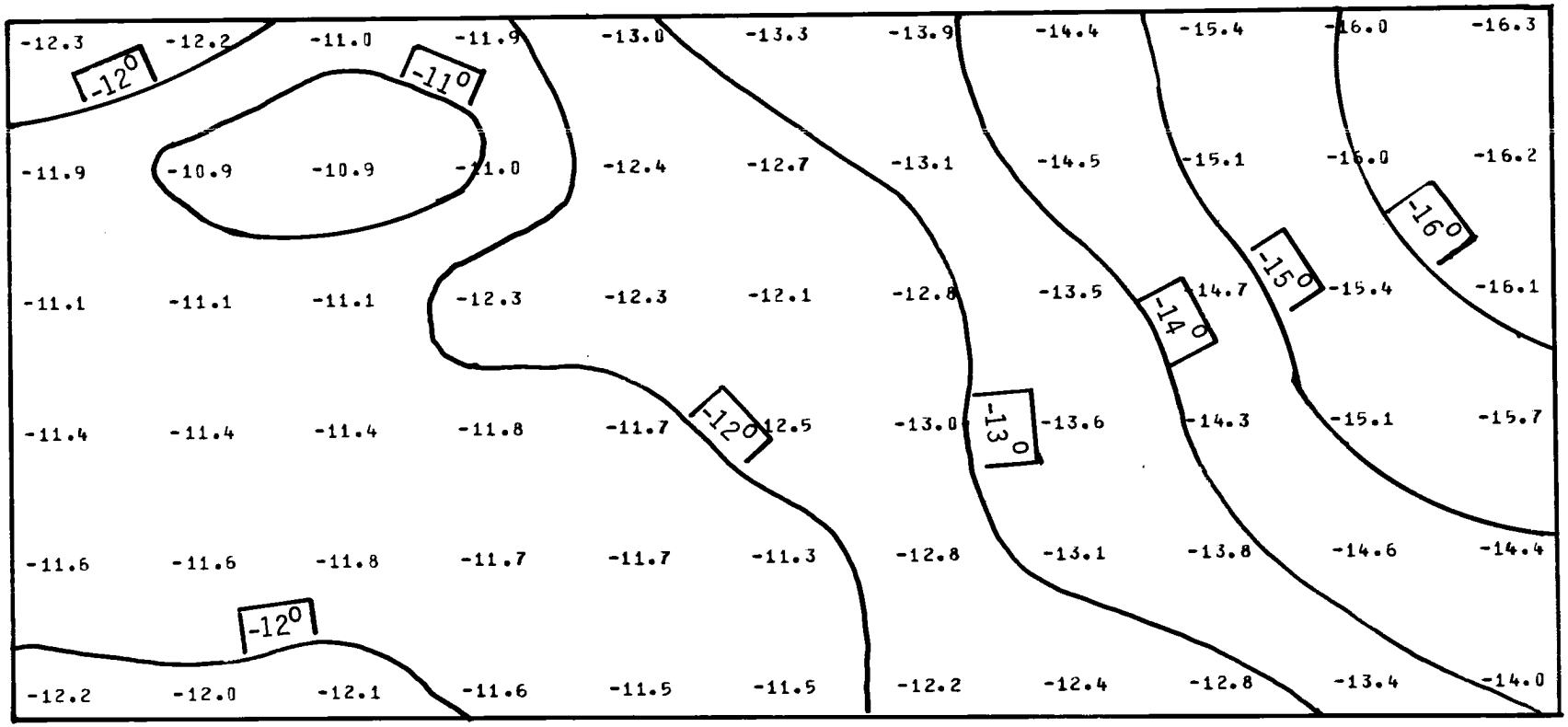


Figure 23. Interpolated radiosonde temperature values (^{o}C) for each latitude-longitude grid point in the 40^{o}-35^{o} Lat. and 105^{o}-95^{o} Long. network for 500 mb.

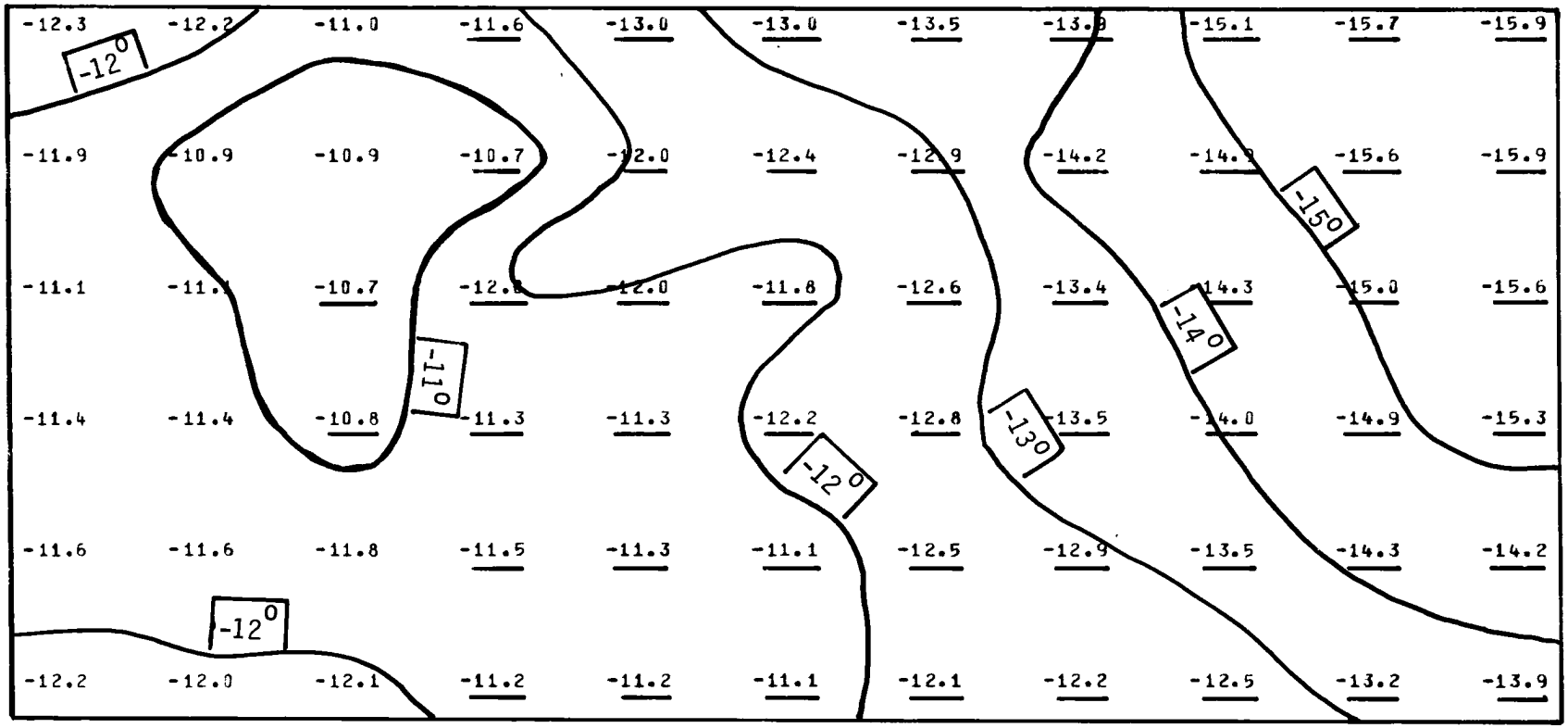


Figure 24. Interpolated radiosonde-satellite derived (C) for each latitude-longitude grid point in the 40^{o}-35^{o} Lat. and 105^{o}-95^{o} Long. network for 500 mb. (Underlined temperature represent grid points in which the NOAA-2 radiometer was able to make scan-spot radiance measurements.)

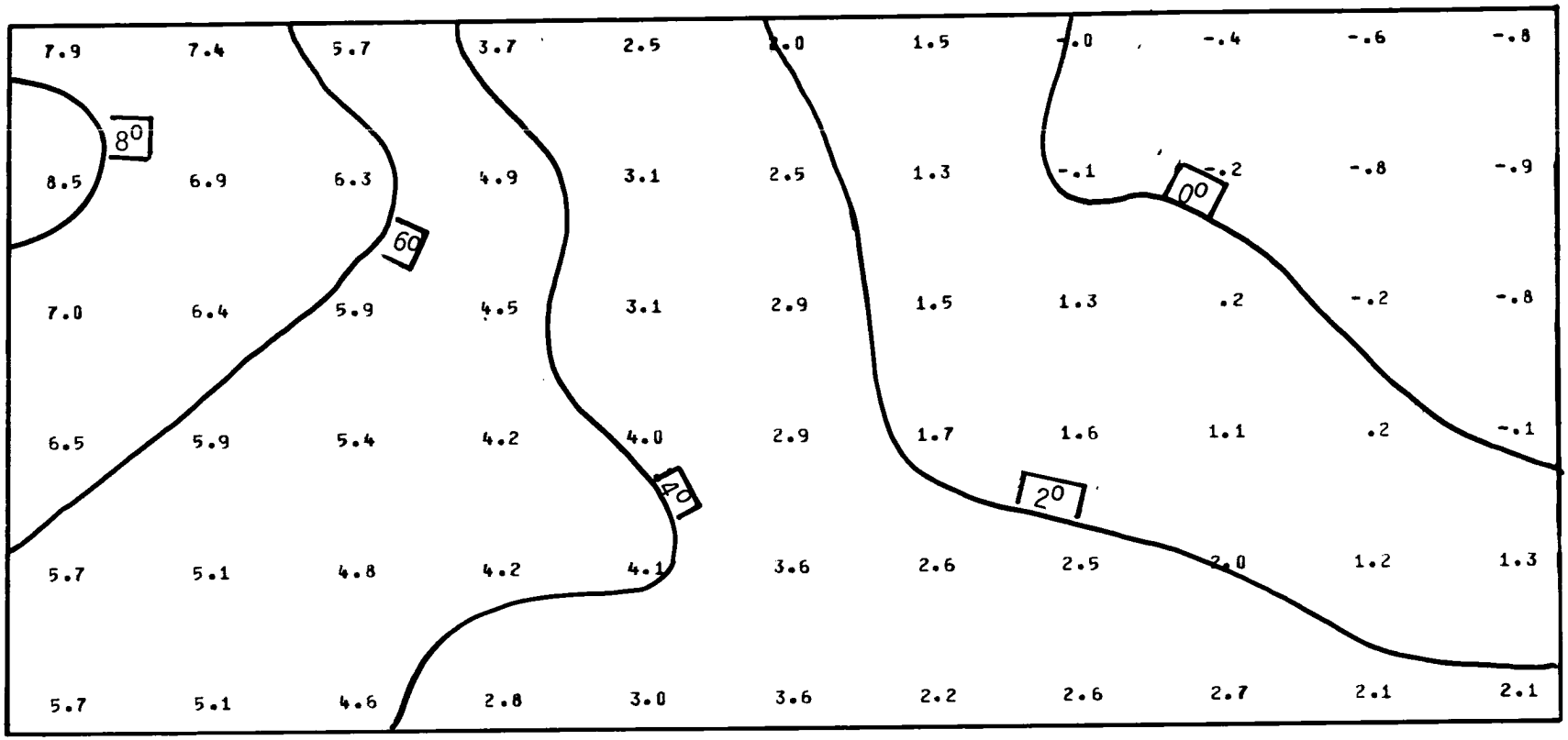


Figure 25. Interpolated radiosonde temperature values (^{o}C) for each latitude-longitude grid point in the 40^{o}-35^{o} Lat. and 105^{o}-95^{o} Long. network for 700 mb.

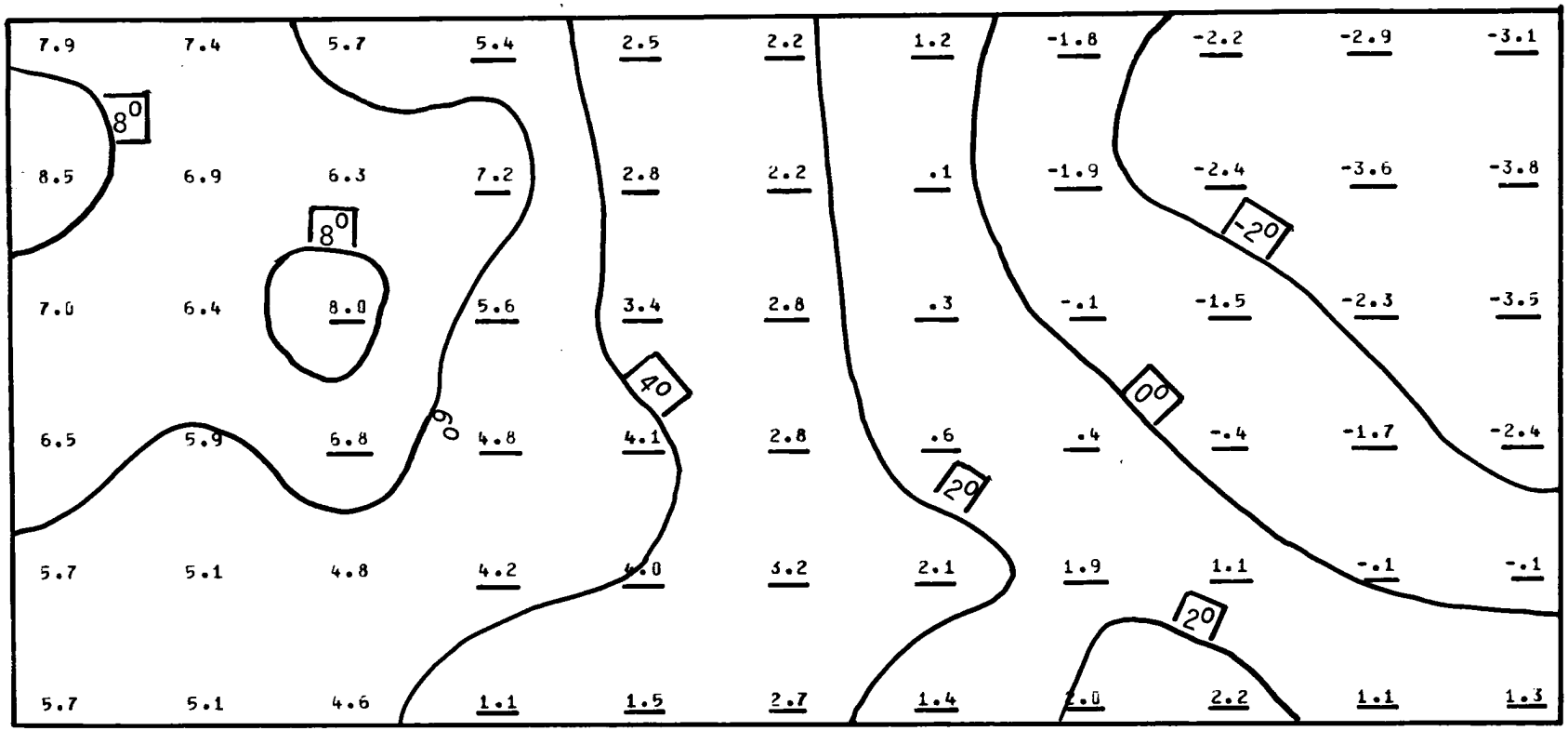


Figure 26. Interpolated radiosonde-satellite derived (°C) for each latitude-longitude grid point in the 40°-35° Lat. and 105°-95° Long. network for 700 mb. (Underlined temperature represent grid points in which the NOAA-2 radiometer was able to make scan-spot radiance measurements.)

VIII. CONCLUSION

In these studies it should be pointed out that the statistical nature of a meteorological element for an area of interest is compiled and retained in the structure and correlation functions and in the climatological norms of that region. The principle may be likened to a typical forecaster who must be accustomed to the mesoscale peculiarities within an area. He will apply his past experience as well as his new learned experience to predict the weather. In the same way, optimum interpolation will assimilate new information as it comes in and thus standardize its prediction or analysis in a systematic and methodical manner. For the moment the question is how to properly weigh one piece of information in relation to another. The weighting factors will change in accordance with the meteorological elements being measured, the instrumentation errors, the random observational errors, the mutual and auto-correlation functions, the number of climatological sample, and finally the isotropic nature of the field. Resolving all these factors for a particular region should result in an accurate, efficient, and objective analysis of any meteorological element.

In our test day of May 18, 1976 we noted peculiarities between the RAOB-only analysis and the combined RAOB-Satellite analysis. At 300 mb the interpolated temperature field were almost identical. At 700 mb and 500 mb the temperature gradients and the general form of the isentropic contours were similar at least within a few degrees celsius of each other. The addition of the satellite data over a grid point influenced certain mesoscale areas of about 100 km x 100 km to the same degree that a radiosonde observation launched 300 km away

would have affected the grid point temperature. Although the temperature gradients of the non-satellite analysis were smoother and weaker than the satellite analysis, it is difficult to judge the validity of the analyses since simultaneous verification at the specific grid points were not made. The main point in that satellite measurements can be extremely helpful in data scarce regions where standard radiosonde measurements are more than 300 km away. Analysis of this kind demonstrate the need for better data assimilation in order to maximize the capabilities of all data collecting systems such as those on board air-borne balloons, aircrafts, naval ships, and satellites. Optimum interpolation provides the statistical relationships between data sets necessary for proper handling of various meteorological inputs.

IX. SUGGESTIONS FOR FUTURE RESEARCH

This study demonstrated the feasibility of applying optimum interpolation techniques in objective analysis. In Chapter VI the special treatment of matching the radiosonde data with the satellite-derived temperature data proved to be very encouraging. Although the objective procedure described in this report is complete in itself, there remains several areas of study which will require future research. The following list of suggestions are bases for further investigations.

(a) Collecting meteorological data over many years or decades. The data sample used in our analysis is small and therefore may not be valid for subsequent year's months unless the anomalies observed on May 1976 are fortuitously repeated. The range of our analysis is valid only for one month. In order to make a complete and thorough analysis, meteorological data must be taken and compiled over many years perhaps decades. Collecting past climatological information with updating every month for instance will provide a firm base for future operational objective analysis. Improvement of correlation models can be accomplished by compiling statistical data over different seasons and over many years.

(b) Modeling of the normalized structure and correlation functions. An important phase of objective optimum interpolation is the proper representation of the statistical parameters via the structure and correlation functions. Empirical correlation models as formulated by previous investigators must be tested and verified for various mesoscale as well as synoptic scale distances. It may be that certain models fit

well for a close network of stations (say <200 km) and other models fit well for larger station distances (say >1000 km). Still, much work needs to be done in this area of study.

(c) Inserting anisotropic and inhomogeneous conditions in the structure and correlation functions. Obviously most geographical areas of the world will not be perfectly isotropic and homogeneous. It may turn out that a particular empirical form of the correlation function works best for station pairs situated in a zonal or meridional position. In this case two different representations of the correlations must be used depending upon the stations' orientation. Other features such as mountainous terrain, deserts, vegetation may also be characterized in somewhat the same manner.

(d) Monitoring erroneous data input or mistakes in recording the measured information. Dynamic meteorological controls must be established to screen off and eliminate gross inconsistencies in the initial data. Time consistencies (i.e. a comparison of recent data to data from preceding periods) as well as horizontal and vertical data consistencies (i.e. comparison with data at surrounding points) must be checked. Belousov (1968) discusses one method of static control for geopotentials and temperature.

(e) Characterizing moisture and precipitation fields by the use of satellite and radiosonde data. An attempt to analysis the mixing ratio via the dew point temperatures was made but the results were inconclusive. More data is needed for a thorough study. Use of the H_2O radiance channel on the VTPR satellite instrument may provide additional information.

(f) Matching of other sources of meteorological data such as wind

velocity components. Other sources can be aircraft and naval measurements or constant balloon data.

(g) Assimilating time and space into a 4-dimensional optimum interpolation analysis. The techniques applied in this paper are somewhat a 3-dimensional analysis of the mesoscale horizontal area for three atmospheric pressure levels over the grid network. The minimization of the mean-square errors is also applicable to the time interpolation of meteorological data. An outline of this approach can be found in Bengtsson (1975) as part of the GARP program.

(h) Computing the drift factor in making standard radiosonde balloon measurements. The distances between observations plays an important role in finding the relationships between station pairs. If radiosondes or any measuring instrument drifts with time, this must be evaluated since distance errors of 100 km are not uncommon. For mesoscale measurements in the upper atmosphere drift must be computed.

(i) Finding the optimum network grid size for distinguishing fronts, cyclonic activity, mesoscale features. Depending on the meteorological parameter to be interpolated, the spatial arrangement of the observations can be optimized to detect and measure a particular dynamic feature.

(j) Accounting for missing data. If one of the meteorological observations in missing, either a search for new data in the same vicinity is initiated or the weighting factors minus the missing data must be recomputed. This should not be a major obstacle after the missing data is discovered and the proper correlation functions are known.

REFERENCES

Alaka, M.A. and R.C. Evander, 1972: Matching of observational accuracy and sampling resolution in meteorological data acquisition experiments. Journal of Applied Meteorology, vol. 11, pp.567-576.

Belousov, S.L., L.S. Gandin, and S.A. Mashkovich, 1968: Computer processing of Meteorological Data. Gidrometeorologicheskoe Izdatel'stvo, Leningrad.

Bengtsson, L., 1975, 4-Dimensional Assimilation of Meteorological Observations, WMO GARP, Garp Publication Series No. 15, January.

Bengtsson, L. and N. Gustavsson, 1971: An experiment in the assimilation of data in dynamical analysis. Tellus, 23, 328-336.

Bergthorssen, P. and B. Doos, 1955: Numerical Weather Map Analysis, Tellus, vol. 7, no. 3, Aug., pp. 329-340.

Bessemoulin, J., 1960: Rapport preliminaire du groupe de travail de la Commission de Meteorologie Synoptique sur les reseaux. Note Tech. No. 30, WMO, Geneva, 91 pp.

Bruce, R.E., L.D. Duncan and J.H. Pierluissi, 1977: Experimental study of the relationship between radiosonde temperatures and satellite-derived temperatures, Monthly Weather Review, vol. 105, 493-496.

Buell, C.E., 1972: Correlation functions for wind and geopotential on isobaric surfaces, Journal of Applied Meteorology, 11, 51-59.

Bushby, F.H. and V.M. Huckle, 1957: Objective Analysis in Numerical Forecasting. Quarterly Journal of the Royal Meteorological Society, vol. 83, no. 356.

Bykov, V.V., G.P. Kurbatkin, and I.V. Gorelysheva, 1964: Building a Multilevel Model for Numerical Analysis of Aerological Data. Trudy MMTs, no. 2.

Chahine, M.T., 1970: Inverse problems in radiative transfer: Determination of atmospheric parameters, Journal of Atmospheric Sciences, Vol. 27, No. 6, pp. 960-967.

Cressman, G.P., 1957: An Objective Analysis Study, Joint Numerical Weather Prediction Unit, Technical Memoranda, no. 12.

Cressman, G.P., 1959: An Operational Objective Analysis System, Monthly Weather Review, vol. 87, no. 10, Oct.

velocity components. Other sources can be aircraft and naval measurements or constant balloon data.

(g) Assimilating time and space into a 4-dimensional optimum interpolation analysis. The techniques applied in this paper are somewhat a 3-dimensional analysis of the mesoscale horizontal area for three atmospheric pressure levels over the grid network. The minimization of the mean-square errors is also applicable to the time interpolation of meteorological data. An outline of this approach can be found in Bengtsson (1975) as part of the GARP program.

(h) Computing the drift factor in making standard radiosonde balloon measurements. The distances between observations plays an important role in finding the relationships between station pairs. If radiosondes or any measuring instrument drifts with time, this must be evaluated since distance errors of 100 km are not uncommon. For mesoscale measurements in the upper atmosphere drift must be computed.

(i) Finding the optimum network grid size for distinguishing fronts, cyclonic activity, mesoscale features. Depending on the meteorological parameter to be interpolated, the spatial arrangement of the observations can be optimized to detect and measure a particular dynamic feature.

(j) Accounting for missing data. If one of the meteorological observations in missing, either a search for new data in the same vicinity is initiated or the weighting factors minus the missing data must be recomputed. This should not be a major obstacle after the missing data is discovered and the proper correlation functions are known.

REFERENCES

Alaka, M.A. and R.C. Evander, 1972: Matching of observational accuracy and sampling resolution in meteorological data acquisition experiments. Journal of Applied Meteorology, vol. 11, pp.567-576.

Belousov, S.L., L.S. Gandin, and S.A. Mashkovich, 1968: Computer processing of Meteorological Data. Gidrometeorologicheskoe Izdatel'stvo, Leningrad.

Bengtsson, L., 1975, 4-Dimensional Assimilation of Meteorological Observations, WMO GARP, Garp Publication Series No. 15, January.

Bengtsson, L. and N. Gustavsson, 1971: An experiment in the assimilation of data in dynamical analysis. Tellus, 23, 328-336.

Bergthorssen, P. and B. Doos, 1955: Numerical Weather Map Analysis, Tellus, vol. 7, no. 3, Aug., pp. 329-340.

Bessemoulin, J., 1960: Rapport preliminaire du groupe de travail de la Commission de Meteorologie Synoptique sur les reseaux. Note Tech. No. 30, WMO, Geneva, 91 pp.

Bruce, R.E., L.D. Duncan and J.H. Pierluissi, 1977: Experimental study of the relationship between radiosonde temperatures and satellite-derived temperatures, Monthly Weather Review, vol. 105, 493-496.

Buell, C.E., 1972: Correlation functions for wind and geopotential on isobaric surfaces, Journal of Applied Meteorology, 11, 51-59.

Bushby, F.H. and V.M. Huckle, 1957: Objective Analysis in Numerical Forecasting. Quarterly Journal of the Royal Meteorological Society, vol. 83, no. 356.

Bykov, V.V., G.P. Kurbatkin, and I.V. Gorelysheva, 1964: Building a Multilevel Model for Numerical Analysis of Aerological Data. Trudy MMTs, no. 2.

Chahine, M.T., 1970: Inverse problems in radiative transfer: Determination of atmospheric parameters, Journal of Atmospheric Sciences, Vol. 27, No. 6, pp. 960-967.

Cressman, G.P., 1957: An Objective Analysis Study, Joint Numerical Weather Prediction Unit, Technical Memoranda, no. 12.

Cressman, G.P., 1959: An Operational Objective Analysis System, Monthly Weather Review, vol. 87, no. 10, Oct.

Doos, B.R., 1969: Numerical analysis of meteorological data. Lectures on Numerical Short Range Weather Prediction, WMO Regional Training Seminar, Moscow, 17 Nov.-14 Dec. 1965, Hydrometeoizdat, Leningrad, pp. 678-706.

Eliassen, A., 1954: Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Videnskaps-Akademiets Institutt for Vaer-og Klimaforskning, Report No. 5, 11 pp.

Falbel, G. and D. Zink, 1971: A sequiential filter radiometer for sensing the vertical temperature profile of the earth's atmosphere, Proceedings, IERE Conference on Infrared Techniques, (University of Reading), pp. 231-255.

Gandin, L.S., 1963: Objective Analysis of Meteorological Fields. Gidrometeorologicheskoe Izdatel'stvo Leningrad, 1960. Translated from Russian into English. Israel Program for Scientific Translations, Jerusalem, 1965, pp. 242.

Gandin, L,S., 1969: Objective analysis. Lectures on Numerical Short Range Weather Prediction, WMO Regional Training Seminar, Moscow, 17 Nov.-14 Dec. 1965, Hydrometeoizdat, Leningrad, pp. 633-677.

Gandin, L.S. and R.L. Kagan, 1974: Construction of a system for objective analysis of heterogeneous data based on the method of optimum interpolation and optimum agreement. Meteorology and Hydrology, Washington, D.C., 5, 1-10.

Gilchrist, B. and G.P. Cressman, 1954: An Experiment in Objective Analysis, Tellus, vol. 6, no. 4, Nov., pp. 309-318.

Gubanova, S.I., 1964: Objective Control of Initial Data at Stations of the Northern Hemisphere by the Method of Optimum Interpolation, Izv. An Uzbssr, ser. fiz-mat., no. 4.

Haug, O., 1959: A Method for Numerical Weather May Analysis, Scientific Report, no. 5, Det Norske Meteorogiske Institutt, 10 pp.

Hillger, D.W. and T.H. Vonder Haar, 1976: Mesoscale Temperature and Moisture Fields from Satellite Infrared Soundings, Department of Atmospheric Science, Colorado State University, Paper No. 249.

Hillger, D.W. and T.H. Vonder Haar, 1977: An analysis of satellite infrared soundings at the mesoscale using statistical structure and correlation functions. Submitted to Journal of Atmospheric Science.

Julian, P.R. and H.J. Thiebaux, 1975: On some properties of correlation functions used in optimum interpolation schemes. Monthly Weather Review, 103, 605-616.

Kruger, H.B. 1968: General and special approaches to the problem of objective analysis of meteorological variables. Quarterly Journal of the Royal Meteorological Society, 95, 21-39.

Lacy, C.H., 1973: Objective analysis using modeled space-time covariances; An evaluation, ECOM5514 R&D Tech. Report, Fort Monmouth, N.J.

Miyakoda, K., L. Umscheid, D.H. Lee, J. Sirutis, R. Lusen, and F. Pratte, 1976: The Near-real-time global, 4-Dimensional Analysis Experiment during the GATE Period, Part I, Journal of Atmospheric Science, Vol. 33, No. 4, pp. 561-591.

Morel, P., 1970: Space and time meteorological data analysis and initialization. Lecture No. 5 in Dynamic Meteorology: Lectures Delivered at the Summer School of Space Physics of the Centre National D'Etudes Spatiales, Lannion, France, 7 Aug-12 Sept., D. Reidel.

Olevskays, S.M., 1967: Space-time structure of the geopotential field H500, Atmospheric and Oceanic Physics, 2, 751-753.

Panofsky, H.A., 1949: Objective Weather Map Analysis, Journal of Meteorology, vol. 6, No. 6, Dec., pp. 386-392.

Ramanthan, Y., P. Kulkarni, and D.R. Sikka, 1973: On a study of winter season wind structure at 500 mb in the Indian region for use in objective analysis of the wind field, Journal of Applied Meteorology, 12, 977-983.

Rutherford, I.D., 1972: Data assimilation by statistical interpolation of forecast error fields. Journal of Atmospheric Science, 29, 809-815.

Rutherford, I.D., 1973: Experiments on the updating of P.E. forecasts with real wind and geopotential data. Preprints Third Conference of Probability and Statistics in Atmospheric Science, Boulder, Colo., American Meteorological Society, 198-201.

Rutherford, I.D., 1976: An operational three-dimensional multivariate statistical objective analysis scheme. GARP Report No. 11 (Procceedings of the JOC Study Group Conference on Four-Dimensional Data Assimilation, Paris, 17-21 Nov. 1975), 98-121.

Sasaki, Y., 1958: An objective analysis based on the variational method, Journal of the Meteorological Society of Japan, vol. 36, No. 3, 77-88.

Schlatter, T.W., 1975: Some experiments with a multivariate statistical objective analysis scheme, Monthly Weather Review, vol. 103, pp. 246-257.

Schlatter, T.W., G.W. Branstator, and L.G. Thiel, 1976: Testing a global multivariate statistical objective analysis scheme with observed data. Monthly Weather Review, 104.

Staff Members, Joint Numerical Weather Prediction Unit. One year of Operational Numerical Weather Predicting. Bulletin of the American Meteorological Society, vol. 38, No. 5, 1957.

Thiebaux, H.J., 1973: Maximally stable estimation of meteorological parameters at grid points. Journal of Atmospheric Science, 30, 1710-1714.

Thiebaux, H.J., 1975: Experiments with correlation representations for objective analysis. Monthly Weather Review, 103, 617-627.

Yates, H.W., 1974: Limitations and prospects for atmospheric soundings, Procceedings of Photo-Optical Instrument Engineering, 51, 1-20.

ACKNOWLEDGEMENTS

The text of this paper was originally the thesis presented in partial fulfillment of the requirements for the Degree of Master of Science for Young Paul Yee, Summer 1978. The author would like to express his deep appreciation to his advisor, Dr. Thomas H. Vonder Haar, for his invaluable advice and his aid in initiating this report's investigation of objective analysis. The manuscript was expertly typed by Mary Lou Yee who devoted much of her time and effort into this project. Also special thanks go to Duane Barnhart in reducing and photographing the many figures and tables in the text, the secretaries at the Atmospheric Science Department who were very helpful at the time, and ERC of Colorado State University for the drafting portion of the thesis.

The remaining members of the committee (Dr. William Cotton and Dr. Howard Frisinger) provided useful suggestions and recommendations. Throughout this research, much guidance was provided by my collaborating colleagues in room 301 of the Atmospheric Science Department, CSU. Lastly, we would like to acknowledge the constant cooperation of the National Center for Atmospheric Research during the countless hours of computer programming. This research was supported by National Science Foundation (NSF) Grant no. ATM 76-21307 and ATM 78-00530.

APPENDIX A

LIST OF SYMBOLS

APPENDIX A

LIST OF SYMBOLS

A,B	Constants in applying Teten's mixing ratio formula for liquid water.
$B\{f_j,T'(p_j)\}$	Planck's blackbody value as a function of f_j and $T'(p_j)$, a guess temperature profile.
$B\{f_j,T(p_j)\}$	Modified Planck's blackbody radiance value as a function of f_j corresponding to a new modified temperature.
a,b,c,d	Constants in the empirical forms of the correlation function.
a_i,b_k,c_j,d_j	Weighting factors for a linear combination of observed meteorological elements.
b_{ij}	Struction function.
b_1	Weighting factor of the satellite-derived temperature within 70 km radius of the grid point.
E	Mean square of the errors between calculated temperature and true temperature.
f_j	Sounding frequencies of the eight VTPR channels.
$I(f_j,\bar{p})$	Measured outgoing radiance at a particular sounding frequency f_j and mean atmospheric pressure $\bar{p}$.
$I'(f_j,\bar{p})$	Computed radiance for f_j and mean atmospheric pressure level $\bar{p}$.
k	Fractional part of the true error difference between two stations.
m_{ij}	Covariance or correlation function.
m_{ii}	Variance.
P	Probability that the measured error between two stations is not greater than a k^{th} of their true error.
p_i	Weighting factors for a linear combination of observed temperatures.
p	Pressure in millibars.

r_{il}	Normalized mutual correlation function for T^r and T^s.
T	Temperature at a location.
$\overline{T}$	Climatological temperature.
T'_i	Temperature deviation from the norm.
T'_t	True temperature deviation from the norm.
$\Delta T'_i$	Observational temperature error.
T^r	Radiosonde measured temperature.
T^s	Satellite-derived temperature.
$\Delta T^r, \Delta T^s$	Temperature errors in determining T^r and T^s.
T^r_0	Interpolated temperature at a particular grid point 0.
T^*	Dew point temperature in degrees Kelvin.
u_j, v_j	Zonal and meridional wind components, respectively.
w_s	Mixing ratio.
x	Distance between observational locations.
ε	Measure of error for matching meteorological data.
ξ_l	Normalized mean-square errors or measures of errors for T^s.
η_i	Measure of the random or observational errors.
λ^2	Ratio of the variance of T^r over the variance of T^s.
μ_{ij}	Normalized correlation function.
σ	Averaged absolute error or root-mean-square error in the observational data.
$\tau(f,p)$	Transmitance at frequency f and pressure p.
ν_{ll}	Normalized autocorrelation function for T^s.
Φ_0	Geopotential height at a grid point.
Φ_i	Observed geopotentials.

APPENDIX B

VTPR

APPENDIX B

VTPR

Satellite-derived temperatures were taken by a Vertical Temperature Profile Radiometer (VTPR) and descriptions of the VTPR instrument can be found in Fabel and Zink (1971). The iterative inverse radiative transfer program is outlined in Hilger (1976) and is patterned after Chahines (1970) iterative relaxation equation,

$$\frac{I(f_j,\bar{p})}{I'(f_j,\bar{p})} \simeq \frac{B\{f_j,T(p_j)\}}{B\{f_j,T'(p_j)\}}$$

where $I(f_j,\bar{p})$ = measured outgoing radiance at a particular sounding frequency and pressure level.

$I'(f_j,\bar{p})$ = computed radiance for frequency f_j and mean pressure level $\bar{p}$.

$B\{f_j,T'(p_j)\}$ = Planck's Blackbody value as a function of frequency f_j and a guess temperature.

$B\{f_j,T(p_j)\}$ = Modified Planck's Blackbody value as a function of frequency corresponding to a new temperature.

The temperature profile to the n^{th} interaction $T^{(n)}(p_j)$ becomes a solution to the radiative transfer equation,

$$I(f,\bar{p}) = B\{f,T(p_o)\}\tau(f,p_o) + \int_{\ln p_o}^{\ln \bar{p}} B\{f,T(p)\} \frac{\partial \tau}{\partial(\ln p)} d(\ln p)$$

when the residual radiances between the computed and measured values approach zero. This final solution is not unique and the fact that there exist random errors in the satellite radiance RMS errors of $\pm 1^{o}C$ in the iterated temperature profile retrievals.

In operation, the VTPR scans across the earth's surface over angles of $\pm 32^{o}$ from the nadir on each side of the orbit ground track

and generates 23 vertical temperature profiles per scan line, or 14,000 profiles per orbit. The data is processed 24 hours a day and it provides continuous real-time global coverage at altitudes of up to 60 kilometers. This program has never been realized before in the history of meteorology.

Although satellites will supply an enormous bank of new information to the users, the iterative inverse transfer solution is still highly dependent upon a good initial temperature profile guess. Also, cloud cover will obscure many areas of interest from the VTPR. In leiu of these obstacles, the VTPR must be aided by radiosonde soundings situated at key locations. Optimum interpolation can make use of both types of information and at the same time determine the optimum combination of satellite temperature or moisture data to balloon measurements. For example, in relatively flat and very isotropic regions such as over deserts and oceans, one radiosonde sounding may be all that is needed to provide a temperature guess profile for numerous satellite temperature sounding in these regions. Or in areas where cloudiness is persistent, optimum interpolation of radiosondes is all that is necessary.

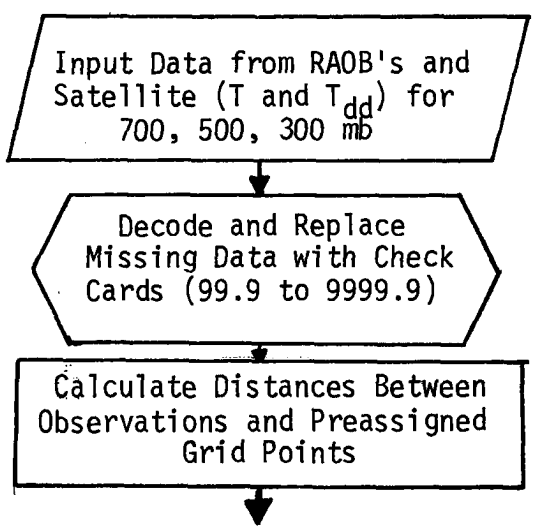
APPENDIX C

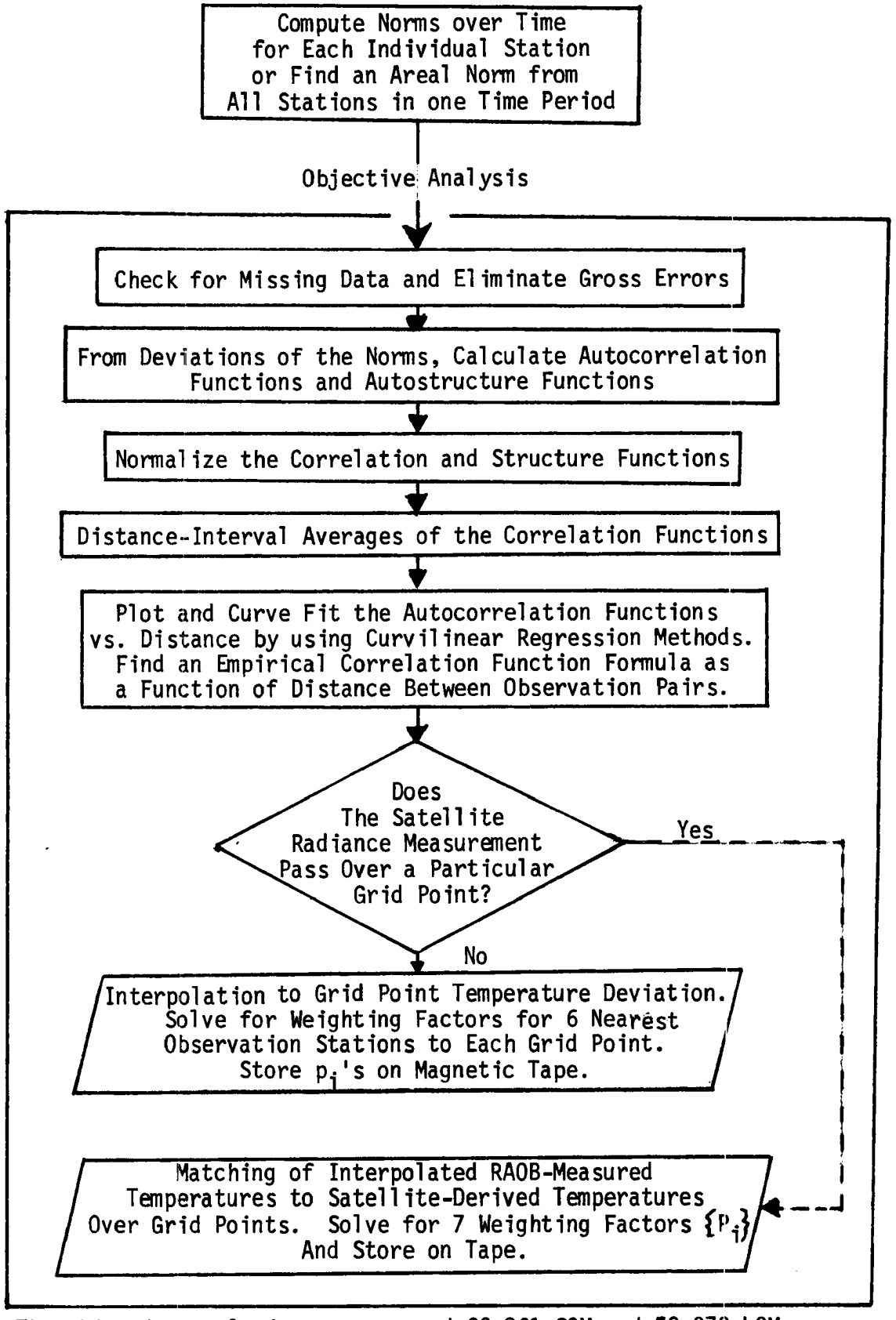
PROGRAM FLOW CHART

APPENDIX C

PROGRAM FLOW CHART

NCAR's (National Center for Atmospheric Research) Control Data 7600 computer facilities at Boulder, Colorado were used to run the optimum interpolation program. The 7600 has twelve PPU's (Peripheral Processing Units), each with 4,096 12 bit words of storage and a 275 user cycle time. 53,000 SCM (small core memory) are available to the user. Our program utilities the Plotting and Solution of Linear Systems subroutines in the NCAR Software Support Library (NSSL). For the symmetric positive definite linear system of equations (matrix equation 16), the Cholesky decomposition technique was called and a iterative improvement was made on the calculated solution vector $\{p_i\}$. For the sets of equations (matrix equation 43 and 44) which includes satellite-derived temperature functions, the Gaussian elimination technique was used. Time on the CDC 7600 to solve a 30 x 30 matrix is about 9.5 milliseconds. A description of the data processing is given by the following flow chart.





The objective analysis program used 22,801 SCM and 32,272 LCM (large core memory)and the total CPU time was 4.532 sec.

BIBLIOGRAPHIC DATA SHEET	1. Report No. CSU-ATSP-292	2.	3. Recipient's Accession No.
4. Title and Subtitle Optimum Interpolation of Radiosonde and Satellite-Derived Temperature Fields			5. Report Date August, 1978 6.
7. Author(s) Young Paul Yee			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Atmospheric Science Department Colorado State University Fort Collins, CO 80523			10. Project/Task/Work Unit No. 11. Contract/Grant No. ATM 76-21307 and ATM 78-00530
12. Sponsoring Organization Name and Address National Science Foundation Atmospheric Research Section Washington, D.C. 20550			13. Type of Report & Period Covered 14.
15. Supplementary Notes			

16. Abstracts

An optimum interpolation scheme which interpolates temperature measurements to a specified mesoscale network of gridpoints is proposed. The analysis incorporates conventional synoptic-scale radiosonde observation at 00Z for thirty-one consecutive days in May 1976 and VTPR satellite-derived temperatures for a particular case study day. Structure functions were determined from the statistical nature of the 700 mb, 500 mb, and 300 mb temperature fields for this period. From these meteorological field samples, normalized auto-correlation and normalized mutual-correlation functions for the RAOB's and satellite-retrieved temperatures were calculated assuming general isotropic and homogeneous conditions over the region. Root-mean square instrumentation errors were estimated by fitting the unnormalized structure function to a linear regression curve to find the zero distance intercept. Applying the smoothed correlation functions, the mutual correlation functions, and the instrumentation errors to a simultaneous set of algebraic equations, the associated RAOB and satellite weighing factors for each gridpoint were computed. This study matches the vertical resolution of the radiosonde data to the horizontal resolution of the satellite radiance retrievals to produce a statistically better objective analysis.

17a. Key Words and Document Analysis.

Optimum interpolation
Mesoscale network
Structure functions

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages 88
	20. Security Class (This Page) UNCLASSIFIED	22. Price