Analysis Error as a Function of Observation Density for Satellite Temperature Soundings with Spatially Correlated Errors

KENNETH H. BERGMAN AND WILLIAM D. BONNER

National Oceanic and Atmospheric Administration, National Meteorological Center, Washington, D.C. 20233

(Manuscript received 29 January 1976, in revised form 21 June 1976)

ABSTRACT

A simple numerical experiment demonstrates that if the errors in satellite-derived temperatures are correlated spatially, the error of an optimum interpolation objective analysis using such temperature data is increased. Moreover, increasing the density of such observations beyond a threshold value (a spacing of about 400 km in the experiment) does not yield any significant improvement in analysis accuracy, in contrast to the case of observations with spatially uncorrelated errors.

1. Introduction

An important question in planning the observing system for the GARP Global Experiment is the horizontal resolution required for satellite-derived temperatures. Accuracy and resolution requirements for temperature measurements have been clearly established as ±1°C and 500 km (GARP, 1973). It seems unlikely at this time that the satellite-derived temperatures will be this accurate, and suggestions have been made that the failure to achieve 1°C accuracy can be compensated for by increasing the yield of satellite-derived temperatures. The purpose of this paper is to point out that our ability to do this may be severely limited by spatial correlations in the errors of the temperature retrievals.

In order to examine the effect of spatial error correlations on the usefulness of observational data, we have carried out a simple numerical experiment, similar in design to one by Alaka and Elvander (1972a, b). The significant departures from their experiment are that calculations are for temperatures rather than winds, that the analysis is considered to be in terms of deviations from a forecast rather than from some climatological state and, most important, that the errors in observations are not assumed to be randomly distributed in space. Systematic errors in the satellite temperature soundings may be expected to arise from the effects of large-scale cloud patterns on the radiance measurements.

2. The experiment

a. Analysis equations

Fig. 1 shows the basic design of the experiment. Twelve observations are arrayed in a rectangular grid of horizontal spacing h. Suppose an interpolated value of the meteorological variable is desired at the central point X. (This point might be a grid point for a numerical prediction model, for instance.) A statistically reliable way of doing this is by optimum interpolation, an application of linear regression theory to spatial interpolation of data.

Suppose temperature is the variable being analyzed. Following Gandin (1963), we let

\[ \hat{T} = \tilde{T} + \tilde{t}, \]

where \( \hat{T} \) is the observed value of the temperature at a given location, \( \tilde{T} \) is some “guess” value (provided by a forecast or climatology) of the temperature at the location, and \( \tilde{t} \) is the deviation of the observed temperature from the guess value. This may be re-expressed as

\[ \hat{T} = \tilde{T} + (t + \epsilon), \]

where \( t \) is the “true” temperature deviation at the location and \( \epsilon \) the observational error.

Then the “analyzed” value \( \hat{T}_a \) of the temperature field at an arbitrary location is given by a weighted linear combination of the deviational temperatures of the neighboring observations:

\[ \hat{T}_a = \tilde{T}_a + \sum_{i=1}^{n} c_i \tilde{t}_i = \tilde{T}_a + \sum_{i=1}^{n} c_i (t_i + \epsilon_i), \]

where \( \hat{T}_a \) is the guess value at the analysis location, the \( c_i \) are the weights to be assigned to the deviational temperature observations \( \tilde{t}_i \) and \( n \) is the number of observations used in the analysis. In the example of Fig. 1, \( n \) is 12 if all of the observations are used in the analysis for the central point.

In general, the right-hand side of (3) will not give the true value \( \hat{T}_a \) of the temperature at the analysis.
point; $\hat{T}_a$ will differ from $T_a$ by an amount called the analysis error $E$. It is generally not possible to eliminate this source of error exactly, no matter how the weights $c_i$ are chosen, but we can require that the statistical mean-square error of Eq. (3), i.e.,

$$E^2 = (T_a - \hat{T}_a)^2 = [T_a - \hat{T}_a - \sum_{i=1}^{n} c_i(t_i + e_i)]^2,$$  (4)

be a minimum for a large ensemble of interpolations and use this requirement to determine the weights.

The result is a set of linear equations

$$\sum_{j=1}^{n} (t_i + e_i)(t_j + e_j)c_{ij} = \overline{t_i t_j}, \quad i = 1, 2, \ldots, n,$$  (5)

where $t_i = T_a - \hat{T}_a$, which may be solved for the weights $c_{ij}$, provided the statistical covariances of (5) are known or approximated.

For radiosonde temperatures, whose errors may safely be assumed to be random and mutually independent, the set of equations (5) simplifies to

$$\sum_{j=1}^{n} \overline{t_i t_j} + \overline{e_i e_j} = \overline{t_i t_j}, \quad i = 1, 2, \ldots, n.$$  (6)

Here $\overline{e_i^2}$ is simply the mean-square error of the $i$th observation. This is essentially the expression used for determining the weights in Alaka and Elvander's numerical experiment.

For satellite-derived temperatures with spatially correlated errors, $e_i e_j$ does not vanish when $i \neq j$; therefore, the set of equations (5) in this case becomes

$$\sum_{j=1}^{n} (t_i t_j + e_i e_j)c_{ij} = \overline{t_i t_j}, \quad i = 1, 2, \ldots, n,$$  (7)

where the covariances $\overline{e_i e_j}$ and $\overline{t_i t_j}$ have been assumed small compared to $e_i e_j$ and neglected.

The theory of optimum interpolation as used here assumes that the mean of the errors is zero at any particular analysis level. This may not be true of satellite-derived temperatures; however, it seems likely that statistical biases at individual levels could be removed or at least reduced in the temperature retrieval process. We are concerned here only with the effect of spatially correlated, unbiased measurement errors on errors in analysis.

It is convenient for computational purposes to normalize Eqs. (7) by dividing by the mean-square deviation $\overline{t_i^2}$. Eqs. (7) then take the form

$$\sum_{j=1}^{n} (\mu_{ij} + \rho_{ij} \overline{e_i e_j})c_{ij} = \mu_{ii}, \quad i = 1, 2, \ldots, n,$$  (8)

where

$$\mu_{ij} = \frac{\overline{t_i t_j}}{\overline{t_i^2}}.$$  (9a)

is the spatial correlation of the $i$th and $j$th observational temperature deviations,

$$\mu_{ii} = \frac{\overline{t_i^2}}{\overline{t_i^2}}$$  (9b)

is the spatial correlation of the $i$th temperature deviation with the temperature deviation being analyzed (at a grid point, usually),

$$\rho_{ij} = \frac{\overline{e_i e_j}}{(\overline{e_i^2} \overline{e_j^2})^{1/2}}$$  (9c)

is the spatial correlation of the observational errors, and

$$\sigma_{ii} = (\overline{e_i^2} / \overline{t_i^2})^{1/2}$$  (9d)

is the normalized standard deviation of observational error. Note that for unbiased errors the normalizing factor $\overline{t_i^2}$ is just the error variance of the "guess" temperature at the analysis point, and the $\mu_{ij}$ are in fact the spatial correlations of the error in the guess temperature field.

b. Spatial correlation of forecast errors

In the numerical experiment based on Fig. 1, the initial guess is assumed to be provided by a 12 h forecast. Spatial correlations of forecast errors have been published by Bengtsson and Gustavsson (1971) for 6, 12 and 24 h forecasts of the 500 mb geopotential height using a quasi-geostrophic model. Their spatial correlations are approximately independent of location or directional orientation; that is, they are approximately homogeneous and isotropic in character. Thus, the forecast error correlation is a function only of separation distance $s$ between the $i$th and $j$th points. Bengtsson and Gustavsson's experimentally
determined curve for 12 h forecast error correlations as a function of $s$ is given in Fig. 2. This curve is well approximated by the analytic function

$$
\mu_{ij}(s) = \exp(-k_\mu s^2),
$$

(10)

where $k_\mu = 1.56 \times 10^{-6}$ km$^{-2}$. Spatial correlations of 12 h forecast height errors by the NMC 6-layer primitive equations model have been evaluated by Peterson (1973) and appear to be in close agreement with (10). Correlation data on forecast temperature errors are not currently available; in this study we assume that (10) with the same value of $k_\mu$ also applies to the spatial correlation of temperature forecast errors.

c. Spatial correlation of satellite temperature errors

A reliable determination of the spatial correlation of temperature errors from satellite soundings has not yet been made. However, for the purpose of this study, we obtained a tentative estimate of the error correlations as a function of separation distance from comparisons between temperature profiles derived from Nimbus 5 and three hand-drawn cross sections based on radiosonde data alone. Nimbus 5 soundings were derived by statistical regression of operational NMC temperature analyses against radiance measurements from the ITPR (Infrared Temperature Profile Radiometer), NEMS (Nimbus E Microwave Spectrometer) and SCR (Scanning Chopper Radiometer) instruments carried by Nimbus 5 (see Smith et al., 1974). All satellite temperature profiles used in this comparison were taken within 100 km of the plane of the cross section and within ±6 h of the radiosonde time. Cross sections were oriented along roughly north-south lines so that the satellite soundings used in each comparison were all nadir measurements from a single satellite orbit.

Differences were obtained between the satellite temperatures and the analyzed temperatures at nine mandatory levels. These differences were averaged at each level, and the mean difference or bias for that level was removed from each satellite temperature. Correlation coefficients were then computed for the remaining satellite “error” as a function of horizontal separation, stratified into categories of 250 km width. In order to increase the size of the statistical sample, coefficients for all levels and all three cross sections were combined to produce the three correlation values plotted in Fig. 3. The four points (including the correlation of 1 at 0 separation) fall along a smooth curve that is similar in shape to the curve for forecast errors (Fig. 2).

There are reservations about the interpretation of this profile, however. First, the statistical sample is small (see Fig. 3) and based only upon three synoptic situations. Second, the error calculations assume that the analysis of the radiosonde data defines the “true”
temperature field on some appropriate scale. Finally, it is likely that different correlation curves would be obtained with different retrieval methods. In particular, we would expect the correlations to fall off less rapidly with distance in systems which use long-term climatology or numerical forecasts to provide first guess profiles for "minimum information" retrievals (see Fritz et al., 1973).

With these reservations in mind, we decided to "model" the spatial error correlation $\rho_{ij}$ with the functional form

$$\rho_{ij} = \exp(-k_p r^{2})$$

and to perform three sets of computations with three different values of $k_p$:

1) $k_p = k_p$. This choice of $k_p$ assumes that satellite temperature errors and forecast errors have the same spatial correlation. This choice may be applicable to satellite temperature retrievals which rely heavily on forecast temperatures as a first guess.

2) $k_p = 8k_p$. This choice closely approximates the correlations obtained from the comparison of ITPR-rawinsonde analyses (Fig. 3), at least out to about 700 km separation.

3) $k_p = \infty$. This choice is equivalent to no spatial correlation of observational errors ($\rho_{ij} = 0$ when $i \neq j$) and is appropriate for rawinsonde temperature data.

d. Analysis error

Once the weights $c_i$ are determined, the mean-square analysis error is given by

$$\bar{E}^2 = I_a\left[1 - \sum_{i=1}^{n} \mu_i c_i\right].$$

Thus, the mean-square analysis error is a fraction of the error variance of the 12 h forecast. Although the observational error variance and the spatial correlations do not appear explicitly in this expression, their influence enters through the weights $c_i$ determined by solving Eqs. (8).

The effect of observational density on the resulting analysis error was simulated by varying the observational spacing $h$ in Fig. 1 from 100 to 1600 km. One might argue that, when $h$ is small, additional observations beyond the 12 shown in Fig. 1 should be included in order to correctly simulate increasing density. However, a trial run with $h=100$ km and 20 additional observations extending the grid of Fig. 1 resulted in very little difference in the analysis error, as the outer observations received virtually no weight in the analysis. Besides, any practical real-time analysis method must limit predictors to a relatively small number which is less than 12 in existing optimum interpolation analysis schemes.

For the limit $h=0$ (all observations coincide with the analysis point), the mean-square analysis error may be expressed as an explicit function of $\rho$ and $\sigma_e$, i.e.,

$$\bar{E}^2 = I_a\frac{\sigma_e^2[1+(n-1)\rho]}{n+\sigma_e^2[1+(n-1)\rho]}.$$  \hspace{1cm} (13)

Here $\rho$ denotes the limit value of the inter-observational error correlation, and it is assumed that all observations are of the same type; hence, $\sigma_{e1} = \sigma_{e2} = \sigma_e$. [Eq. (13) is derived in the Appendix.] When $\rho=0$, Eq. (13) reduces to

$$(\bar{E}^2/I_a)_{h=0,\rho=0} = \frac{\sigma_e^2}{n+\sigma_e^2}$$  \hspace{1cm} (14a)

an error expression appropriate to radiosonde data. On the other hand, if observational errors are
spatially correlated according to (11), then \( \rho \to 1 \) as \( h \to 0 \) and

\[
(\overline{E^2/i_0^2})_{h=0, \rho=1} = \frac{\sigma_a^2}{1+\sigma_e^2} \tag{14b}
\]

the form appropriate for satellite observations. It is evident that the former decreases without limit as \( n \) becomes larger, whereas the latter does not depend on the number of observations.

3. Results

The set of equations (8) with \( n=12 \) was solved numerically for the weights \( c_i \) using the iterative method of conjugate gradients (see, e.g., Beckman, 1960). It was assumed that all 12 observations are of the same type and hence have the same statistical error level. The solutions were obtained for normalized observational errors 0, 0.25, 0.50 and 1.00. (The normalized observational error is the ratio of the standard deviation of the observational error to that of the forecast error.) The correlation coefficients \( \rho_{ij} \) and \( \rho_{ij} \) were determined from the geometry of Fig. 1 and from (10) and (11). In (10), \( k_\mu \) was assigned the value \( 1.56 \times 10^8 \) km\(^{-2} \). In (11), \( k_\mu \) was assigned the values \( k_\mu \), \( 8k_\mu \) and \( \infty \). The observational spacing \( h \) was varied from 100 to 1600 km.

Once the weights \( c_i \) were obtained for a particular set of conditions, the normalized analysis error

\[
\sigma_a = \left( \overline{E^2/i_0^2} \right)^rac{1}{2}
\]

was obtained from (12). The results for \( \sigma_a \) as a function of \( h \) are plotted in Fig. 4 for \( \sigma_e = 0.25 \), Fig. 5 for \( \sigma_e = 0.50 \) and Fig. 6 for \( \sigma_e = 1.00 \). The limiting values of \( \sigma_a \) for \( h = 0 \) were obtained from (14a, b). In each of the figures, the dotted curve for \( \sigma_e = 0 \) (perfect ob-

![Fig. 5. As in Fig. 4 except for \( \sigma_e = 0 \) and 0.50.](Image)

![Fig. 6. As in Fig. 4 except for \( \sigma_e = 0 \) and 1.00.](Image)
servations) is shown for comparison. The other three curves are for the three values of \( k_p \) that are assumed. The solid curve for \( k_p = \infty \) is appropriate for conventional data with random independent errors, whereas the dashed curves apply to data with spatially correlated errors such as is assumed for satellite data.

The dotted curve indicates that portion of the analysis error which results from spatial interpolation of the observations to the analysis point; it is seen to be of negligible importance when \( h \) is less than about 400 km. For closely spaced observations, then, nearly all of the analysis error is a result of the observational errors. For widely spaced observations, on the other hand, the analysis error is primarily a result of the spatial interpolation.

The figures clearly indicate that the analysis error is less for independent observations (\( k_p = \infty \)) than for observations with spatially correlated errors. Of these, the case with the higher degree of correlation (\( k_p = k_\mu \)) has the larger analysis error. For both correlated cases, the analysis error diminishes with decreasing separation only until some limiting value is reached. This value occurs with a station separation of about 500 km for \( k_p = k_\mu \) and 400 km for \( k_p = 8k_\mu \). (The increase in \( \sigma_\varepsilon \) as \( h \) approaches zero, to the value given by (14b) in the \( k_p = 8k_\mu \) case is a consequence of assuming perfect correlation of observational errors in the limit \( h = 0 \) [see Eq. (14b)].) These results suggest that there is little information to be gained from increasing the yield of satellite observations beyond the point where the average station separation is less than 400–500 km.

In Figs. 7, 8 and 9, the normalized analysis error \( \sigma_\varepsilon \) is shown as a function of both the normalized
observational error $\sigma_a$ and the observational spacing $h$ for \( k_a = \infty \), \( 8k_a \) and \( k_a \) respectively. Where the isopleths of $\sigma_a$ are approximately vertical, the analysis error is primarily dependent on the spacing of observations. Where the isopleths are approximately horizontal, analysis error is a function of the errors in observation and cannot be reduced by increasing the density of observations. The figures show clearly that, for both cases with spatially correlated errors, there is no increase in analysis accuracy associated with an increase in observation density to a spacing that is less than about 400 km. For values of $h < 200$ km (Fig. 8) or 300 km (Fig. 9), analysis errors are at least twice as large as the corresponding errors for the uncorrelated case.

The weights $c_i$ obtained by solving (8) are of secondary interest here. In a real analysis situation, the analyzed value is given by (3) once the weights have been determined. In this experiment, the observed values $l_i$ have been left unspecified, as here we are interested in the analysis error rather than the analyzed value itself. Because of the symmetry of the observational array (Fig. 1), the four inner observations all receive the same weight, as do the eight outer ones. Generally, the inner observations receive the bulk of the weight, the outer ones comparatively little. When the observational error level $\sigma_a$ is high or when $h$ is large, the weights $c_i$ tend to decrease in magnitude, effectively giving greater importance to the guess value $T_a$ in the final analysis.

4. Summary and conclusions

A simple numerical experiment has been performed which extends the study of Alaka and Elvander (1972a,b) to include the effects of spatially correlated errors. Results indicate that this spatial error correlation reduces the information content of point observations compared to that for data with random errors. They suggest that increasing the density of observations beyond a certain threshold (a spacing of about 400 km in the experiment) will yield little or no improvement in temperature analyses produced from satellite soundings. This statement assumes that the sounding errors are spatially correlated over distances of 500 km or more (see Fig. 3). Although the assumption appears reasonable and has been used in simulation studies (Baumhefner and Julian, 1972), such correlations have not yet been determined with the desired precision. Only a rough approximation based on Nimbus 5 data is given here. The problem is made difficult by the fact that operational satellite soundings are produced only over the oceans where it is impossible to define a "true" temperature or height field from radiosonde data.

APPENDIX

Derivation of Limiting Value of Mean-Square Analysis Error

Eq. (8) states that the observational weights $c_i$ are obtained by solving the set of linear equations

$$\sum_{j=1}^{n} (\mu_{ij} + \rho_{ij}\sigma_i\sigma_j)c_j = \mu_{ai}, \quad i = 1, 2, \ldots, n,$$

where the symbols have the same meaning as in the text. Separating the "diagonal" term, where $j = i$, from the other $(n-1)$ terms on the left side leads to

$$\sum_{j \neq i} [(\mu_{ij} + \rho_{ij}\sigma_i\sigma_j)c_j] + (\mu_{ii} + \sigma_i^2)c_i = \mu_{ai},$$

$$i = 1, 2, \ldots, n, \quad (A1)$$

where $\sigma_i^2$ is the normalized error variance of the $i$th observation. (Note that $\rho_{ii}$ is necessarily unity;
observational errors are perfectly correlated with themselves.) We assume that all observations are of the same type and have the same statistical error, i.e.,

$$\sigma_{ij} = \sigma_{ai} = \sigma_e.$$  

In the limiting case where all \( n \) observations coincide with the analysis point, the correlations \( \mu_{ij} \) between pairs of observations and \( \mu_{ai} \) between observations and analysis point become unity; thus,

$$\sum_{j \neq i} [(1 + \rho_{ij}\sigma^2_e)c_i] + (1 + \sigma^2_e)c_i = 1, \quad i = 1, 2, \ldots n. \quad (A2)$$

If in this limiting case the error correlation between pairs of observations is also statistically uniform, \( \rho_{ij} = \rho \), then

$$(1 + \rho\sigma^2_e) \sum_{j \neq i} c_j + (1 + \sigma^2_e)c_i = 1, \quad i = 1, 2, \ldots n. \quad (A3)$$

There are three possible cases to consider:

I. \( \sigma_e = 0 \) and \( \rho \neq 1 \)

In this case each of the \( n \) equations (A3) has \( (n-1) \) coefficients \( (1 + \rho\sigma^2_e) \) and one coefficient \( (1 + \sigma^2_e) \). Each equation is identical except that the weights \( c_i \) are interchanged. This is possible only if all the \( c_i \) are equal and given by

$$c^* = \frac{1}{(n-1)(1 + \rho\sigma^2_e) + (1 + \sigma^2_e)} = \frac{1}{\sigma^2_e[1 + (n-1)\rho] + n}. \quad (A4)$$

Then, from Eq. (12),

$$\bar{E}^2 = t_n^2(1 - \sum_{i=1}^{n} c_i) = t_n^2(1 - nc^*) = t_n^2 \frac{\sigma^2_e[1 + (n-1)\rho]}{n + \sigma^2_e[1 + (n-1)\rho]}. \quad (A5)$$

This expression is the same as Eq. (13) of the text.

II. \( \sigma_e \neq 0 \) but \( \rho = 1 \)

In this case (A3) reduces to a set of \( n \) identical equations

$$\sum_{j=1}^{n} c_j = \frac{1}{1 + \sigma^2_e}, \quad \sigma_e = \frac{n}{n + 1}, \quad \sigma^2_e = \frac{1}{n-1}. \quad (A6)$$

so that even though the individual weights are undetermined, their sum is known. Then again from Eq. (12),

$$\bar{E}^2 = t_n^2(1 - \sum_{i=1}^{n} c_i) = t_n^2 \frac{\sigma^2_e}{1 + \sigma^2_e}. \quad (A7)$$

This expression is in agreement with (A5) when \( \rho = 1 \).

III. \( \sigma_e = 0 \)

In this trivial case (A3) reduces to

$$\sum_{j=1}^{n} c_j = 1. \quad (A8)$$

Eq. (12) then gives

$$\bar{E}^2 = 0, \quad (A9)$$

the expected result when \( \sigma_e = 0 \), and one which is consistent with (A5). Thus, the analysis error for all possible cases is correctly given by (A5) and hence by Eq. (13) of the text.

Acknowledgments. Radiosonde cross sections used in this study were constructed by Mr. Paul Lemar and Mr. Robert Van Haaren of the Data Assimilation Branch, NMC. Nimbus 5 soundings were derived for us by Dr. C. Hayden of NESS. We would like to thank Dr. L. Bengtsson of the European Center for Medium Range Weather Forecasting and Dr. M. Alaka of Techniques Development Laboratory, NWS, NOAA, for helpful discussions on the application of optimum interpolation techniques to problems of this type.

REFERENCES


