U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
NATIONAL METEOROLOGICAL CENTER

OFFICE NOTE 265

Optimum Interpolation: Basic Formulation and Characteristics

Ronald D. McPherson
Development Division

SEPTEMBER 1982

Lecture Prepared for the Stanstead Seminar
July 12-16, 1982

1. Introduction

Optimum interpolation is a technique for analyzing meteorological observations, transforming their information content into fields of meteorological variables. It is based on the statistical characteristics of the fields being analyzed. The earliest reference to the method is due to Eliassen (1954), although Gandin (1963) must receive a great deal of credit for the method's development and subsequent widespread use. Beginning in the early 1970's, it became apparent that the advent of remotely-sensed atmospheric data would greatly change the heretofore homogeneous character of the data base (largely radiosondes). Optimum interpolation offered a suitable framework for systematically treating data with different error characteristics. Rutherford (1972) and Schlatter (1975) began developing data assimilation systems based on optimum interpolation, and by 1979, the year of the Global Weather Experiment, such systems were in use at several research institutions and operational numerical weather prediction centers.

This paper presents a summary of the basic formulation of the method, within the very simple context of performing an analysis at one point using only three pieces of information. Some of the method's characteristics are then illustrated by means of a series of simple analysis problems.

2. Formulation

We assume for this development that we have one observation of geopotential (H_O), and one of wind which is treated in the form of eastward

(U_o) and northward (V_o) components, for a total of three pieces of information. The analyzed values of these parameters (H_a, U_a, V_a) at some point in the vicinity of the observation may be written as a "guess" value of the parameter (H_g, U_g, V_g), usually from a forecast, plus a correction formed by linear combinations involving the observations.

$$
\begin{aligned}
H_a &= H_g + a_1(H_o-H_g)_1 + a_2(U_o-U_g)_2 + a_3(V_o-V_g)_3 \\
U_a &= U_g + b_1(H_o-H_g)_1 + b_2(U_o-U_g)_2 + b_3(V_o-V_g)_3 \qquad (1) \\
V_a &= V_g + c_1(H_o-H_g)_1 + c_2(U_o-U_g)_2 + c_3(V_o-V_g)_3
\end{aligned}
$$

where (H_o-H_g), etc. are the differences, or "residuals", between the data at an observation location and the appropriate guess value interpolated to that location, and the a_i, b_i, c_i are coefficients which determine the weight each datum receives in the analysis, relative to the guess value. The analysis consists of determining the coefficients.

Note that observations of both geopotential and wind are used in the analyses of both variables: this type of analysis is said to be "multivariate". In contrast, an analysis which uses only observations of the variable being analyzed is said to be "univariate".

In order to calculate the coefficients (a_i, b_i, c_i), we begin by subtracting the true values (H_T, U_T, V_T) from both sides of eqns. (1), changing the signs, and introducing the following definitions:

$f_a = F_T-F_H$: true analysis error in the variable F(H, U, or V);

$f_g = F_T-F_g$: true guess error in F;

$F_o = F_T+\varepsilon_f$: ε_f = error in observation of the variable F.

We note that the "observed residuals" F_o-F_g, may be rewritten as

$$F_o-F_g = F_T-F_g+\varepsilon_f = f_g+\varepsilon_f$$

so that the observed residual may be expressed as the sum of the error in the guess plus the error in the observation. Eqns. (1) may then be recast in terms of these error quantities:

$$
\begin{aligned}
h_a &= h_g - [a_1(h_g+\varepsilon_h)_1 + a_2(u_g+\varepsilon_u)_2 + a_3(v_g+\varepsilon_v)_3] \\
u_a &= u_g - [b_1(h_g+\varepsilon_h)_1 + b_2(u_g+\varepsilon_u)_2 + b_3(v_g+\varepsilon_v)_3] \qquad (2) \\
v_a &= v_g - [c_1(h_g+\varepsilon_h)_1 + c_2(u_g+\varepsilon_u)_2 + c_3(v_g+\varepsilon_v)_3]
\end{aligned}
$$

Optimum interpolation determines the coefficients (a_i, b_i, c_i) such that the mean-square analysis errors ($\overline{h_a^2}$, $\overline{u_a^2}$, $\overline{v_a^2}$) are minimized.

$$
\frac{\partial}{\partial a_i}(\overline{h_a{}^2}) = \frac{\partial}{\partial b_i}(\overline{u_a{}^2}) = \frac{\partial}{\partial c_i}(\overline{v_a{}^2}) = 0 \qquad (3)
$$

The minimization process results in sets of three linear equations, one set for each variable: for geopotential,

$$
\begin{aligned}
(\overline{h_1h_1}+ \varepsilon_h)a_1 + (\overline{h_1u_2})a_2 + (\overline{h_1v_3})a_3 &= \overline{h_1h_g} \\
(\overline{u_2h_1})a_1 + (\overline{u_2u_2}+\varepsilon_u)a_2 + (\overline{u_2v_3})a_3 &= \overline{u_2h_g} \qquad (4) \\
(\overline{v_3h_i})a_1 + (\overline{v_3u_2})a_2 + (\overline{v_3v_3}+\varepsilon_v)a_3 &= \overline{v_3h_g}.
\end{aligned}
$$

The parenthetical quantities are various covariances of guess, or forecast, errors. Note that the subscript g has been deleted from the forecast error at an observation location, but retained in the forecast error at the analysis point. For example, the term $\overline{h_1h_g}$ denotes the covariance between the forecast geopotential error at observation location #1 (h_1) and its counterpart at the analysis point (h_g). Quantities on the main diagonal such as $\overline{h_1h_1}$, etc., are recognized as the variances of the forecast errors σ_h^2, σ_u^2, σ_v^2. The off-diagonal quantities are seen to be cross-covariances between forecast errors of geopotential and those of wind.

It should be noted that terms such as $\overline{h_i \varepsilon_j}$ and $\overline{\varepsilon_h \varepsilon_v}$ have been assumed to vanish.

For the u- and v-components, similar sets of equations may be obtained, in which the a_i are replaced by b_i and c_i and the h_g in the covariances on the right-hand side is replaced by u_g and v_g, respectively. So long as the same set of observations is used for the analysis of geopotential and the wind components, the forecast error covariance matrix on the left-hand side of eqns. (4) is common to all three sets. The order of the three sets is determined by the number of observations used in the analysis.

If the forecast error covariance matrix and the right-hand side vectors can be specified, the matrix can be inverted and the three sets solved for the unknowns a_i, b_i, c_i. The analyzed values at the analysis point may then be calculated from eqns. (1). It may also be shown that the minimized analysis error variance may be determined from

$$
\begin{aligned}
h_a &= \sigma_h - a_1\overline{h_1 h_g} - a_2\overline{u_2 h_g} - a_3\overline{v_3 h_g} \\
u_a &= \sigma_u - b_1\overline{h_1 u_g} - b_2\overline{u_2 u_g} - b_3\overline{v_3 u_g} \\
v_a &= \sigma_v - a_1\overline{h_1 v_g} - c_2\overline{(u_2 v_g)} - c_3\overline{v_3 v_g}.
\end{aligned}
\tag{5}
$$

It is computationally convenient to model the forecast error covariance matrix by an analytic, differentiable function which approximates actual forecast error covariances. In the NMC system, the height-height error covariance is specified by

$$\overline{h_i h_j} = \{\sigma(h)_i \sigma(h)_j\}\{\exp[-K_H(s_i - s_j)^2]\}\{[1 + K_p \ln^2(p_i/p_j)]^{-1}\} \tag{6}$$

where $s_i - s_j$ is the horizontal separation between points (i) and (j), and $p_i - p_j$ is their vertical separation. This expression is in the form of a triple product of a variance, a horizontal correlation, and a vertical correlation. The forecast error cross-covariances in eqns. (4) are determined by assuming that the errors in geopotential, temperature, and wind are related hydrostatically and geostrophically. Under these assumptions, all forecast error covariances can be calculated by differentiation of eqn. (6). Table 1 gives the differential relationship between $\overline{h_i h_j}$ and all other convariances.

Table 1. Covariance of the row variable with the column variable in terms of the geopotential autocovariance $\overline{hh}$, assuming that height, temperature, and wind residuals are related through the geostrophic and hydrostatic equations.

	h	t	u	v
h	$\overline{hh}$	$-\frac{g}{R}\frac{\partial \overline{hh}}{\partial z}$	$-\frac{g}{f}\frac{\partial \overline{hh}}{\partial y}$	$\frac{g}{f}\frac{\partial \overline{hh}}{\partial x}$
t	$-\frac{g}{R}\frac{\partial \overline{hh}}{\partial \zeta}$	$\left(\frac{g}{R}\right)^2 \frac{\partial^2 \overline{hh}}{\partial \zeta \partial z}$	$\frac{g^2}{fR}\frac{\partial^2 \overline{hh}}{\partial \zeta \partial y}$	$-\frac{g^2}{fR}\frac{\partial^2 \overline{hh}}{\partial \zeta \partial x}$
u	$-\frac{g}{f}\frac{\partial \overline{hh}}{\partial \eta}$	$\frac{g^2}{fR}\frac{\partial^2 \overline{hh}}{\partial \eta \partial z}$	$\frac{g^2}{f^2}\frac{\partial^2 \overline{hh}}{\partial \eta \partial y}$	$-\frac{g^2}{f^2}\frac{\partial^2 \overline{hh}}{\partial \eta \partial x}$
v	$\frac{g}{f}\frac{\partial \overline{hh}}{\partial \xi}$	$-\frac{g^2}{fR}\frac{\partial^2 \overline{hh}}{\partial \xi \partial z}$	$-\frac{g^2}{f^2}\frac{\partial^2 \overline{hh}}{\partial \xi \partial y}$	$\frac{g^2}{f^2}\frac{\partial^2 \overline{hh}}{\partial \xi \partial x}$

The main diagonal terms of eqns. (4) contain forecast error and observational error variances which must be specified. With respect to the former, it should be noted that the assumption of a hydrostatic and geostrophic relationship between forecast errors implies a relationship between the variances of the errors as well. It can be shown that

$$\sigma_u^2 = \sigma_h^2 (2K_H g^2/f^2), \tag{7}$$

and

$$\sigma_t^2 = \sigma_h^2 (2K_p g^2/R^2), \tag{8}$$

where σ_t is the forecast temperature error variance.

In an analysis/forecast data assimilation cycle, it is necessary to predict the forecast geopotential error variance σ_h^2 valid at the next analysis time, usually a few hours in advance. For the NMC 6h cycle, this is accomplished by assuming that σ_h at the next analysis time (denoted by superscript n+1) is related to the analysis error variance at time n:

$$\sigma_h^{n+1} = [(\overline{h_a^2})^{\frac{1}{2}}]^n + D \tag{9}$$

where D is an estimate of the forecast error growth rate determined from verification statistics. In the NMC system, D is a function of variable and pressure level, but not horizontal position.

The observational error variances ($\overline{\varepsilon_h^2}$, $\overline{\varepsilon_u^2}$, $\overline{\varepsilon_v^2}$) in the covariance matrix must be pre-specified. Typically, this is done by classes of observations; that is, radiosondes are assigned one error variance, satellite data another, aircraft another, etc. Current values in use at NMC were adopted from the European Centre for Medium-Range Weather Forecasting and may be found in Bengtsson (1981).

3. Characteristics

To illustrate some of the method's characteristics, we first consider a problem in which we wish to perform an univariate analysis of geopotential using only two geopotential observations. We will assume for all but the last of these examples that the observations are arranged as in the schematic:

that is, each is colinear with the analysis point and separated from it by a distance δ_1 or δ_2. The analysis equation may be written as

$$h_a - h_g = \Delta h = a_1(h_1+\varepsilon_1) + a_2(h_2+\varepsilon_2) \tag{10}$$

Minimization of the mean-square analysis error leads to

$$\begin{aligned}(\overline{h_1h_1} + \overline{\varepsilon_1^2})a_1 + (\overline{h_1h_2} + \overline{\varepsilon_1\varepsilon_2})a_2 &= \overline{h_1h_g}\\ (\overline{h_2h_1} + \overline{\varepsilon_2\varepsilon_1})a_1 + (\overline{h_2h_2} + \overline{\varepsilon_2^2})a_2 &= \overline{h_2h_g}\end{aligned} \tag{11}$$

Note that we have left open the possibility of correlated observational errors through the presence of the term $\overline{\varepsilon_1\varepsilon_2}$.

For the covariance model, we assume that observations and analysis point are at the same level, so that

$$\overline{h_ih_j} = \sigma_h^2\, e^{-k\delta_{ij}^2}, \tag{12}$$

where δ_{ij} is the distance between (i) and (j). The elements of the forecast error covariance matrix are

$$\overline{h_1h_1} = \overline{h_2h_2} = \sigma_h^2;\ \overline{h_1h_2} = \overline{h_2h_1} = \sigma_h^2\, e^{-k(\delta_1 + \delta_2)^2}, \tag{13}$$

and the right-hand side is

$$\overline{h_1h_g} = \overline{h_2h_g} = \sigma_h^2\, e^{-k\delta_{1,2}^2}. \tag{14}$$

With eqns. (13) and (14) the solution of eqns. (11) are

$$a_1 = \frac{(\sigma^2 e^{-k\delta_1^2})(\sigma^2+\overline{\varepsilon_2^2}) - (\sigma^2 e^{-k\delta_2^2})[\sigma^2 e^{-k(\delta_1+\delta_2)^2} + \overline{\varepsilon_1\varepsilon_2}]}{(\sigma^2+\overline{\varepsilon_1^2})(\sigma^2+\overline{\varepsilon_2^2}) - [\sigma^2 e^{-k(\delta_1+\delta_2)^2} + \overline{\varepsilon_1\varepsilon_2}]^2} \tag{15}$$

and

$$a_2 = \frac{(\sigma^2 e^{-k\delta_2^2})(\sigma^2+\overline{\varepsilon_1^2}) - (\sigma^2 e^{-k\delta_1^2})[\sigma^2 e^{-k(\delta_1+\delta_2)^2} + \overline{\varepsilon_1\varepsilon_2}]}{(\sigma^2+\overline{\varepsilon_1^2})(\sigma^2+\overline{\varepsilon_2^2}) - [\sigma^2 e^{-k(\delta_1+\delta_2)^2} + \overline{\varepsilon_1\varepsilon_2}]^2}, \tag{16}$$

where σ_h^2 has become σ^2 for convenience.

We now consider a series of examples in which (15) and (16) can be simplified to illustrate some characteristics of optimum interpolation.

A. Random observational error: $\overline{\varepsilon_1\varepsilon_2} = 0$

Datum #1 at analysis point: $\delta_1 = 0$

Datum #2 far removed: $\delta_2 \to \infty$

Both reports of same type: $\overline{\varepsilon_1{}^2} = \overline{\varepsilon_2{}^2} = \overline{\varepsilon^2}$

The solutions become

$$a_1 = \frac{\sigma^2(\sigma^2+\overline{\varepsilon^2})}{(\sigma^2+\overline{\varepsilon^2})^2} = \frac{1}{1+\overline{\varepsilon^2}/\sigma^2} \; ; \; a_2 = 0$$

Thus, for a single observation located at the analysis point, we see that if the datum is error-free, $(\overline{\varepsilon^2}=0)$ it receives a weight of unity; that is, the analysis exactly reflects the datum. The effect of imperfect data is to reduce the influence of the data in the analysis, in proportion to the ratio of the observational error variance to the forecast error variance.

B. Random observational errors: $\overline{\varepsilon_1\varepsilon_2} = 0$

Datum #1 at analysis point: $\delta_1 = 0$

Datum #2 δ away: $\delta_2 = \delta$

Both of same type: $\overline{\varepsilon_1{}^2} = \overline{\varepsilon_2{}^2} = \overline{\varepsilon^2}$

The solutions are

$$a_1 = 1-\overline{\varepsilon^2}\left[\frac{\sigma^2+\overline{\varepsilon^2}}{(\sigma^2+\overline{\varepsilon^2})^2-(\sigma^2 e^{-k\delta^2})^2}\right]$$

$$a_2 = \overline{\varepsilon^2}\left[\frac{\sigma^2 e^{-k\delta^2}}{(\sigma^2+\overline{\varepsilon^2})^2-(\sigma^2 e^{-k\delta^2})^2}\right].$$

From this it may be seen that even if another observation is nearby, a perfect observation ($\overline{\varepsilon^2} = 0$) located at the analysis point will be reflected exactly and the others will receive no weight. For imperfect data, however, both observations receive some weight.

C. Same circumstances as in (A) except that datum #1 is a distance from the analysis point: $\delta_1 = \delta,\ \delta_2 \to \infty$.

$$a_1 = \frac{e^{-k\delta^2}}{1+\overline{\varepsilon^2}/\sigma^2}\ ,\ a_2 = 0$$

Thus the influence of a single imperfect datum is given by a the assumed forecast error correlation function, reduced by the error variance ratio $\overline{\varepsilon^2}/\sigma^2$.

D. Random observational errors: $\overline{\varepsilon_1\varepsilon_2} = 0$

Both observations equidistant from analysis point, but far enough from each other to be uncorrelated: $\overline{h_1h_2} = \overline{h_2h_1} = 0$

Each datum has different error characteristics: $\overline{\varepsilon_1{}^2} \neq \overline{\varepsilon_2{}^2}$

$$a_1 = \frac{e^{-k\delta^2}}{1+\overline{\varepsilon_1{}^2}/\sigma^2}\ ;\ a_2 = \frac{e^{-k\delta^2}}{1+\overline{\varepsilon_2{}^2}/\sigma^2}$$

If $\overline{\varepsilon_1{}^2} < \overline{\varepsilon_2{}^2}$, $a_1 > a_2$; better quality data receives more weight in the analysis.

E. Same circumstances as (D), except that $\overline{h_1h_2},\ \overline{h_2h_1} \neq 0$, but the data are of the same type: $\overline{\varepsilon_1{}^2} = \overline{\varepsilon_2{}^2} = \overline{\varepsilon^2}$:

$$a_1 = a_2 = \frac{e^{-k\delta^2}}{1+\overline{\varepsilon^2}/\sigma^2+e^{-4k\delta^2}}$$

The two observations receive the same weight, but less than in (D); thus the effect of interobservational correlation is to reduce the influence of each datum in a univariate analysis.

F. Observations and analysis point are at the vertices of an equilateral triangle: $\delta_1 = \delta_2 = \delta_{12} = \delta$

Both data of same type: $\overline{e_1^{\,2}} = \overline{\varepsilon_2^{\,2}} = \overline{\varepsilon^2}$

Correlated observational errors: $\overline{\varepsilon_1\varepsilon_2} \neq 0$

$$a_1 = a_2 = \frac{e^{-k\delta^2}}{1+\overline{\varepsilon^2}/\sigma^2+e^{-k\delta^2}+\overline{\varepsilon_1\varepsilon_2}/\sigma^2}$$

Thus data with positively-correlated observational errors receive less weight than those with random errors, in univariate analysis. This is not true in multivariate analysis.

G. We now consider the effect of wind data on the mass analysis (i.e., multivariate analysis). We assume one height observation and one wind observation, colocated a distance δ from the analysis point:

The analysis equation for geopotential is

$$h_a - h_g = \Delta h = a_1(v_1+\varepsilon_v) + a_2(h_2+\varepsilon_h) \tag{17}$$

and the obtain after minimizing the mean-square analysis error

$$\begin{aligned} (\overline{v_1v_1}+\overline{\varepsilon_v^{\,2}})a_1 + (\overline{v_1h_2})\ a_2 &= \overline{v_1h_g} \\ (\overline{h_2v_1})\ a_1 + (\overline{h_2h_2}+\overline{\varepsilon_h^{\,2}})\ a_2 &= \overline{h_2h_g} \end{aligned} \tag{18}$$

Our covariance model is

$$\overline{h_ih_j} = \sigma^2 e^{-k\delta_{ij}^2};\ \overline{h_iv_j} = \frac{g}{f}\frac{\partial}{\partial x}(\overline{h_ih_j}) \tag{19}$$

so that the elements of the error covariance matrix become

$$\left.\begin{aligned} &\overline{h_2h_2} = \sigma_h^{\,2};\ \overline{v_1v_1} = \sigma_v^{\,2} = \sigma_h^{\,2}\ (2kg^2/f^2) \\ &\overline{v_1h_2} = \overline{h_2v_1} = 0 \\ &v_1h_g = -\frac{2k|\delta|g}{f}\sigma_h^{\,2}e^{-k\delta^2};\ \overline{h_2h_g} = \sigma_h^{\,2}e^{-k\delta^2}. \end{aligned}\right\} \tag{20}$$

The solutions for a_1 and a_2 are

$$a_1 = \frac{-\left(\frac{f\delta}{g}\right)\sigma_v^{\,2} e^{-k\delta^2}}{\sigma_v^{\,2}+\overline{\varepsilon_v^{\,2}}} \; ; \; a_2 = \frac{\sigma_h^{\,2} e^{-k\delta^2}}{\sigma_h^{\,2}+\overline{\varepsilon_h^{\,2}}} \; . \tag{21}$$

Because a_1 and a_2 have different dimensions, direct comparison as in the univariate case is not possible. We instead examine the separate contributions of the wind and height observations to the reduction of analysis error variance,

$$\overline{h_a^{\,2}} = \sigma_h^{\,2} - a_1\overline{v_1 h_g} - a_2\overline{h_2 h_g} \tag{22}$$

For the case of perfect data ($\overline{\varepsilon_r^{\,2}} = \overline{\varepsilon_h^{\,2}} = 0$),

$$h_a^{\,2} = \sigma_h^{\,2}[1-e^{2k\delta^2}(1+2k\delta^2)] \tag{23}$$

The factor $e^{-2k\delta^2}$ represents the reduction of error variance due only to the height observation; its product with $(1+2k\delta^2)$ represents the reduction due to both height and wind together. Eqn. (23) has been evaluated numerically for several values of the separation distance δ and the results are presented in Table 2. Even at 8° latitude separate, the reduction of error variance with both wind and height data is over 50%, while with a height report alone, the reduction is only 20%. The effect of the wind observation on the height analysis is clearly to improve the accuracy of the analysis.

H. We now examine the influence of mass data on the wind analysis. The configuration in example (A) is assumed. The analysis equation is

$$V_a - v_g = \Delta v = a_1(h_1+\varepsilon_h)+a_2(h_2+\varepsilon_h) \, . \tag{24}$$

Analogous to eqns. (4), there results from the minimization

$$\begin{aligned}(\overline{h_1h_1}+\overline{\varepsilon_h^{\,2}})a_1 + (\overline{h_1h_2}+\overline{\varepsilon_1\varepsilon_2})a_2 &= \overline{h_1v_g}\\ (\overline{h_2h_1}+\overline{\varepsilon_2\varepsilon_1})a_1 + (\overline{h_2h_2}+\overline{\varepsilon_h^{\,2}})a_2 &= \overline{h_2v_g}\end{aligned} \tag{25}$$

From the covariance model in example (G), we may calculate the elements of the error covariance matrix

$$\left.\begin{aligned}&\overline{h_1h_1} = \overline{h_2h_2} = \sigma_h^{\,2}\\ &h_1h_2 = h_2h_1 = \sigma_h^{\,2}e^{-k(2\delta)^2}, \text{ with } |\delta_1| = |\delta_2|\\ &\qquad\qquad\qquad\qquad\qquad\quad \delta_1 = X(0)-X(1)\\ &\qquad\qquad\qquad\qquad\qquad\quad \delta_2 = X(2)-X(0)\\ &h_1v_g = -h_2v_g = \frac{-2k|\delta|g}{f}\sigma_h^{\,2}e^{-k\delta^2}\ .\end{aligned}\right\} \tag{26}$$

The solutions for the coefficients are

$$a_1 = -a_2 = \frac{-2k|\delta|g}{f}\left[\frac{e^{-k\delta^2}}{1+R^2(1-\rho)-e^{-4k\delta^2}}\right] \tag{27}$$

where $R^2 = \overline{\varepsilon_h^{\,2}}/\sigma_h^{\,2}$, $R^2\rho = \overline{\varepsilon_1\varepsilon_2}/\sigma_h^{\,2}$. We may write the analysis error variance as

$$v_a^{\,2} = \sigma_v^{\,2}\left[1 - \frac{4k\delta^2e^{-2k\delta^2}}{1+R^2(1-\rho)-e^{-4k\delta^2}}\right] \tag{28}$$

We may see from eqn. (27) that, in contrast to the univariate analysis in example (F), the effect of correlated observational errors ($\rho \neq o$) is to increase the weight such an observation receives, and to increase the reduction of analysis error variance. Indeed, data with perfectly correlated errors ($\rho=1$) are as good as error-free data in this example. It may also be noted that the effect of interobservational correlation is to improve the analysis in this case.

4. Summary

By means of a series of simple analysis problems, we have illustrated some of the behavior characteristics of optimum interpolation. Among them are:

- o Observations are weighted in proportion to the ratio of the accuracy of the data to the accuracy of the forecast in the vicinity of the observation: the more accurate the data relative to the forecast the more weight it receives in the analysis;
- o All else being equal, observations with smaller error variances receive more weight than those with larger ones;
- o The effect on non-independence of observations is recognized by reducing the weight such observations receive in univariate analyses;
- o In the analysis of geopotential, a wind observation added to a geopotential observation is more beneficial than a geopotential obseration by itself.
- o Observations with random errors receive more weight in a univariate analysis than those with correlated errors; but it is better to have correlated than random errors in an analysis where observations of one variable are being used to analyze another variable related to the gradient of the first.

References

Bengtsson, L., 1981: "Data analysis, initialization, and data assimilation." Proceedings, International Conference on Preliminary FGGE Data Analysis and Results. Bergan, Norway, 23-27 June 1980, 162-185.

Eliassen, A., 1954: "Provisional report on calculation of spatial covariance and autocorrelation of the pressure field." Rapport No. 5, Videnskaps-Akademie Institute for Vaer-Ogklinaforskning, Oslo, Norway. Reprinted in Dynamic Meteorology: Data Assimilation Methods, L. Bengtsson, M. Ghil, and E. Kallen, eds. Applied Mathematical Sciences, V36, Springer-Verlag, New York, 1981.

Gandin, L., 1963: "Objective analysis of meteorological fields". Gidrometeorologicheskoe Isdatel Stvo, Leningrad. Translated from Russian, Israel Program for Scientific Translations, Jerusalem (1965), 242.

Rutherford, I., 1973: "Experiments in the updating of P. E. forecasts with real wind and geopotential data." Preprint, Third Conference on Probability and Statistics in Meteorology, Boston, Amer. Meteor. Soc., 198-201.

Schlatter, T., 1975: "Some experiments with a multivariate statistical objective analysis scheme." Mon. Wea. Rev., 103, 246-257.