Tellus (1979), **31**, 387–399

The development of the analysis error of the 500 mb height, 1946–1969

By JUHANI RINNE and KIM FRISK, *Department of Meteorology, University of Helsinki, Hallituskatu 11-13, SF-00100 Helsinki 10, Finland*

(Manuscript received May 29, 1978; in final form February 22, 1979)

ABSTRACT

The analysis error of the 500 mb height is estimated as a function of time. The parameters are the observation error, the forecasting error and the network of radiosonde stations from which the analysis centre received observations. With certain assumptions, the average analysis error north of 20° N was found to be 45 m, 22 m, 18 m and 18 m for certain periods in 1946, 1952, 1957 and 1969, respectively. The corresponding maximum errors are 108 m, 76 m, 73 m and 76 m. Areas with possible systematic errors introduced by forecasts and wind observations are determined.

1. Introduction

In a study dealing with a long series of 500 mb analyses, it proved necessary to estimate the analysis error as a function of time. The results may be of more general interest and are therefore reported in this paper. Hence, the aim of this paper is to describe the development of the analysis error in 1946–69. The analysis process is simulated using the method of optimum linear interpolation (Gandin, 1963). The parameters of the analysis process are the observation error of the height of the 500 mb surface, the observation error of the wind, the forecasting error and the station network.

K. Frisk is mainly responsible for Sections 2.3 and 3, and for the running of the programs; most of the remaining part of the paper was written by J. Rinne.

2. Basic formulae

In Gandin's method the analyzed value is found as a weighted mean of the observed anomalies so that the average difference between the true value and the analysis will be a minimum

$$E_0 = \overline{(z_0^{\text{true}} - z_0^{\text{anal}})^2}$$

$$= \overline{(z_0^{\text{true}} - \underbrace{\sum_{i=1}^{N} p_i z_i}_{\text{I}} - \underbrace{\sum_{i=N+1}^{2N} p_i \vartheta_i}_{\text{II}} - \underbrace{p_{2N+1} z_0^{\text{fore}}}_{\text{III}})^2} \quad (1)$$

The following symbols are used:

z_0^{true}	the true anomaly, i.e. the true departure of the height from the corresponding climatological value, at grid point 0
z_0^{anal}	the analyzed anomaly at grid point 0
z_0^{fore}	the forecast anomaly, i.e. the forecast departure of the height from the climatological value of the height, at grid point 0
z_i	the observed anomaly at station i
z	the anomaly of the height
ϑ	the anomalous wind (will be defined later); "anomalous" means that in the geostrophic approximation the wind is related to the height anomaly
N	the number of stations used in the analysis
p_i	the weights to be determined
E_0	the variance of the analysis error at grid point 0
I	the height group

0040-2826/79/050387-13$02.50/0

II the wind group
III the forecast group

The bar refers to averaging with respect to time.

The weights p_i are determined by setting

$$\frac{\partial E_0}{\partial p_i} = 0 \quad (i = 1, 2, \ldots, 2N + 1) \tag{2}$$

It is not necessary to know the actual values of the meteorological variables in order to estimate the variance of the analysis error (cf. Appendix). The covariance functions of these variables are the only parameters in (1). This is seen by solving eq. (2) and applying the result in (1).

Some of the covariance terms, scaled by m_{00}, are listed below:

1) $\overline{z_i z_j}/m_{00} = \overline{(z_i^{\text{true}} + \delta z_i^{\text{obs}})(z_j^{\text{true}} + \delta z_j^{\text{obs}})}/m_{00}$

$$= \begin{cases} R(r), & i \neq j \\ R(0) + \eta^z, & i = j \end{cases}$$

2) $\overline{z_0^{\text{true}} z_0^{\text{fore}}}/m_{00} = R(0)$

3) $\overline{z_0^{\text{fore}} z_0^{\text{fore}}}/m_{00} = R(0) + \eta^f$

where

m_{00} is the variance of the anomaly of the 500 mb height, $m_{00} = \overline{z_0^{\text{true}} z_0^{\text{true}}}$

δz_i^{obs} is the observation error of the height at station i

η^z is the relative height observation error $\eta^z = \overline{(\delta z_i^{\text{obs}})^2}/m_{00}$

η^f is the relative forecast error $\eta^f = \overline{(\delta z_i^{\text{fore}})^2}/m_{00}$

$R(r)$ is the correlation of the height anomaly at two points separated by a distance r

$R(0)$ is the relative variance of the true height anomaly, $R(0) = 1$. The relative variance of the observed anomaly will be $1 + \eta^z$

Certain assumptions usually made when applying Gandin's method are implicit in the present analysis system. We have assumed that

— the observation error is not correlated with the observation
— the observation errors of different stations are not correlated
— m_{00} does not vary across the study region (homogeneity)
— the climatological values are given exactly so that the observation error of the height anomaly equals the observation error of the height.

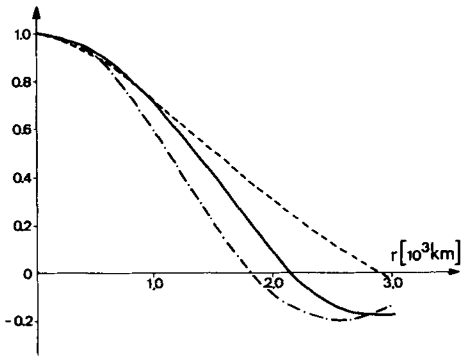

Fig. 1. Three different autocorrelation functions of the height anomaly (500 mb). Legend: ---- Gandin, 1963; —— Julian and Thiebaux, 1975; —·—·— Helminen, 1972.

2.1. *The autocorrelation of the height*

Three autocorrelation functions are shown in Fig. 1. The one presented in Julian and Thiebaux (1975) and in Thiebaux (1975) is

$$R(r) = [\alpha \cos(\omega r) + \beta][1 + (\lambda r)^2]^{\kappa} \tag{3}$$

They define m_{00} using the variance of the observed anomaly, whereas we have defined it using the true anomaly. Consequently, we have to take $\beta = 1 - \alpha$, because by definition we have $R(0) = 1$. We take $\alpha = 0.714$, $\beta = 0.286$, $\kappa = -0.34$, $\omega = 0.0009418$ km^{-1}, $\lambda = 0.0009418$ km^{-1}, r = distance in 1000 km. Unfortunately, there is an obvious printing error in Thiebaux (1975): the variable λ has been replaced by ω. This is why we have used $\lambda = 0.0009418$ km^{-1} instead of $\lambda = 0.0007055$ given in the paper referred to. As a result of this error, our autocorrelation function is slightly more dampened than that of Julian and Thiebaux. This kind of difference is small in comparison with the variation between the different versions of $R(r)$ in Fig. 1.

Helminen's (1972) graph (Fig. 1) has been computed from independent data and seems to support formula (3). This formula is the basis for the present computation. However, some comparisons will be made with the traditional formula given by Gandin:

$$R(r) = e^{-0.188 r^{1.30}} \cos(0.54\, r) \tag{4}$$

2.2. *The wind covariances*

The wind is resolved into two components respectively parallel and perpendicular to the line connecting the observation point to the grid point.

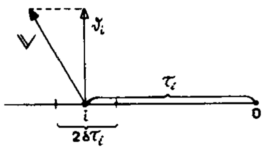

Fig. 2. Illustration of the derivation of the cross covariance between height and wind. τ_i is the distance of grid point 0 from station i, ϑ_i is the wind component used in computations and $\mathbb{V}$ is the wind observed at the station.

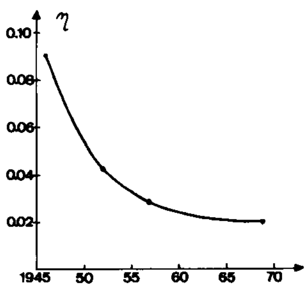

Fig. 3. The assumed relative observation error. The years 1946, 1952, 1957 and 1969 are indicated.

The correlation between the parallel component and the height at the grid point is poor (see e.g. Schlatter, 1975). Therefore only the transverse component (ϑ) will be utilized. As an example, the relative cross covariance $\overline{z_0^{\text{true}}\,\vartheta_i}/m_{00}$ will be derived in the following.

Using the geostrophic approximation we obtain (Fig. 2)

$$\frac{\overline{z_0^{\text{true}}\,\vartheta_i}}{m_{00}} = -\frac{g}{f_i}\,\frac{\overline{[z(\tau_i+\delta\tau_i)-z(\tau_i-\delta\tau_i)]z_0^{\text{true}}}}{2\delta\tau_i}\,\frac{1}{m_{00}}$$

We omit the local variation of the coriolis parameter and replace f_i by $f = f_0$ (value at the grid point). Then

$$\frac{\overline{z_0^{\text{true}}\,\vartheta_i}}{m_{00}} = -\frac{g}{f}\,\frac{R(\tau+\delta\tau_i)-R(\tau-\delta\tau_i)}{2\delta\tau_i}$$

Letting $\delta\tau_i \to 0$ results in

$$\frac{\overline{z_0^{\text{true}}\,\vartheta_i}}{m_{00}} = -\frac{g}{f}\left[\frac{\partial R}{\partial\tau}\right]_i$$

The other wind covariances are derived in the same manner (cf. Buell, 1972; Julian and Thiebaux, 1975). For instance,

$$\overline{\vartheta_i\,\vartheta_j}/m_{00} = -\left(\frac{g}{f}\right)^2\frac{\partial^2 R}{\partial r^2}, \quad \text{if } i \neq j$$

When $i = j$, we have an extra term due to the observation error of the wind:

$$\overline{\vartheta_i\,\vartheta_i}/m_{00} = -\left(\frac{g}{f}\right)^2\left[\frac{\partial^2 R}{\partial r^2}\right]_{r=0} + \eta^{\vartheta}\,\frac{\overline{\vartheta_i^{\text{true}}\,\vartheta_i^{\text{true}}}}{m_{00}}$$

where the relative wind error is

$$\eta^{\vartheta} = \overline{(\vartheta_i^{\text{obs}}-\vartheta_i^{\text{true}})^2}/\overline{\vartheta_i^{\text{true}}\,\vartheta_i^{\text{true}}}$$

Throughout this paper, $f(\varphi) = f(30°)$, if $\varphi < 30°$.

2.3. *The observation error*

Our estimations of the observation error (Fig. 3) are based on Harrison (1962), McInturff (1968), WMO (1971), Hooper (1975) and Vockeroth (1975). Furthermore, we assume (following Schlatter, 1975) that the relative error in the height and wind is equal, i.e. $\eta^z = \eta^{\vartheta}$. This assumption is a crude one. However, it makes the computations easier.

Thus, if the variance of the height anomaly ($m_{00} = \overline{z_0^{\text{true}}\,z_0^{\text{true}}}$) is 20,000 m², the standard deviations of the observation error associated with the height are estimated to be 45 m, 29 m, 24 m and 20 m, for the years 1945, 1952, 1957 and 1969, respectively. At 45° N the corresponding wind errors will be approximately 5 m/s, 3 m/s, 2.5 m/s and 2 m/s.

2.4. *The forecasting errors*

The forecasting errors at adjacent points are normally correlated, so that little is lost by using only the forecast value of the grid point. The variance of the forecasting error at the grid point is estimated from

$$\overline{(z_0^{\text{fore}}-z_0^{\text{true}})^2} = 3\cdot\overline{(z_0^{\text{anal}}-z_0^{\text{true}})^2} + 1600\ \text{m}^2 \qquad (5)$$

The first term on the right-hand side represents inaccuracies in the initial conditions amplified threefold during, say, a 24-h forecast. If the initial analyses are accurate, the first term vanishes and the variance of the forecasting error is 1600 m².

The square root of this constant (40 m) thus estimates the standard deviation of the model error accumulated during 24 forecast hours. If $E_0 = \overline{(z_0^{\text{anal}} - z_0^{\text{true}})^2}$ equals, say, $(24\ \text{m})^2$, then the standard deviation of the forecasting error is estimated to be 59 m.

Assuming that $m_{00} = 20{,}000\ \text{m}^2$ we have from (5)

$$\eta^f = 3 \cdot E_0/m_{00} + 0.08 \qquad (6)$$

Hence E_0 must be known in order to estimate η^f, which is then used to compute E_0. This problem is solved by computing a preliminary field of E_0 using wind and height terms only. Then η^f can be estimated at every grid point and a final field of E_0 will be computed (cf. 4.3). Eq. 6 is associated with the 24-h forecasts. However, we can let it represent more generally all "predicted" values such as those extrapolated from the actual surface analysis or derived by experience.

Later objective analysis systems have had preliminary fields based on better forecasts. On the other hand, the earlier analysts, who worked without computers, used surface analysis and all their experience and knowledge. Therefore it is difficult to decide which analyses, the objective or the subjective, have been based on the better "prediction" methods. This is why the very same formula (6) has been applied during the whole period 1946–1969.

3. Station network

The number of observations received is smaller than the number of observation stations. According to Huovila (personal communication, see also WMO (1976), p. 8) it may even be possible that although most of the stations in a region work and send in their observations regularly, they are not received because of breaks in the telecommunication system. Generally, the breaks in the data flow depend on the transmitting centre, on the receiving centre and on the connections between them. Thus the analysis error is specific for each analyzing centre. In the present study, a specific analysis centre is chosen and its analysis error is computed for certain periods or moments: 1969, whole year; 1957, December; 1952, September; 1946, January 13 and January 21.

An exception is 1969, for which no analysis centre was chosen. In other words, it was assumed that the observations from the standard stations were regularly available everywhere. A "standard station" was defined as one listed by WMO (1968) and which observed the height every day at 00 GMT $\pm$ 2 h. The coordinates of the stations were checked against a network map (U.S. Weather Bureau, 1963). From the same map the coordinates of the stations in China (+ one station in North Vietnam) were picked up. The regularity of the observational work of these latter stations was checked from Daily Synoptic Weather Maps (U.S. Weather Bureau, 1967). Over the area given in Fig. 12, we found a total of 560 "standard stations".

In the case of December 1957, the analyzing centre is defined as the one that produced the Synoptic Weather Maps (U.S.W.B., 1957). A working station is now defined as one which regularly had plotted height observations (12 GMT $\pm$ 2 h) on Synoptic Weather Maps during December 1957. However, for stations south of 20° N the source map was that of January 1, 1959 (U.S.W.B., 1959). Plotting is defined as regular if the station has, on average, plotted values in three analyses out of four. For instance, ship K has been removed from the station list because it had no observations plotted for an 8-day period. On the other hand, there are two stations which were not found in WMO lists: 86° N 165° W and 83° N 165° W. The total number of stations accepted is 365 (Fig. 11).

In September 1952 there was a total of 265 stations which had regularly plotted height observations on Synoptic Weather Maps (U.S.W.B., 1952) at 15 GMT $\pm$ 3 h (Fig. 10).

In 1946 the variation in the amount of observed and transmitted data was large, and it was therefore impossible to collect a usable list of regular stations. Instead, 2 days were chosen to represent typical conditions. 130 stations had plotted height observations on the Historical Weather Map (Headquarters, AAF 1946) on January 13, 1946 at 04 GMT (Fig. 8). The other day chosen was January 21, 1946, with 143 accepted stations (Fig. 9). If we had chosen extreme days, the number of stations would have been either well below 100 or above 200.

4. Preliminary computations

The effect of the different terms of (1) is studied stepwise. The relative analysis error E_0/m_{00} is

Fig. 4. The relative analysis error when the analyses are based only on height observations (500 mb). The applied station network is that of 1969. Gandin's autocorrelation function is used. The minimum value is 0.01 (not indicated).

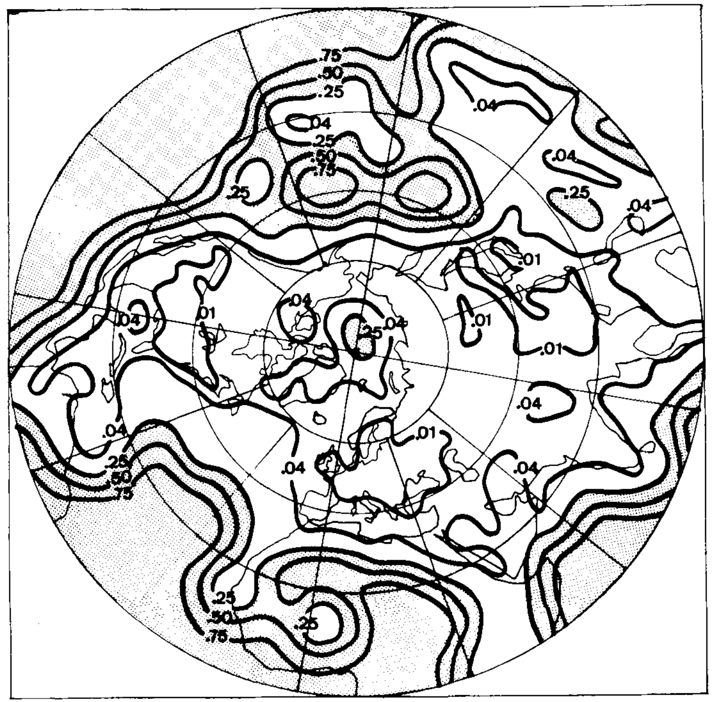

Fig. 5. The same as Fig. 4 but for the autocorrelation function of Julian and Thiebaux (1975). The minimum value is 0.004 (not indicated).

Fig. 6. The relative analysis error when the analyses are based on height and wind observations. It is assumed that every station in the station list regularly observed both the height and the wind. The minimum value is 0.003.

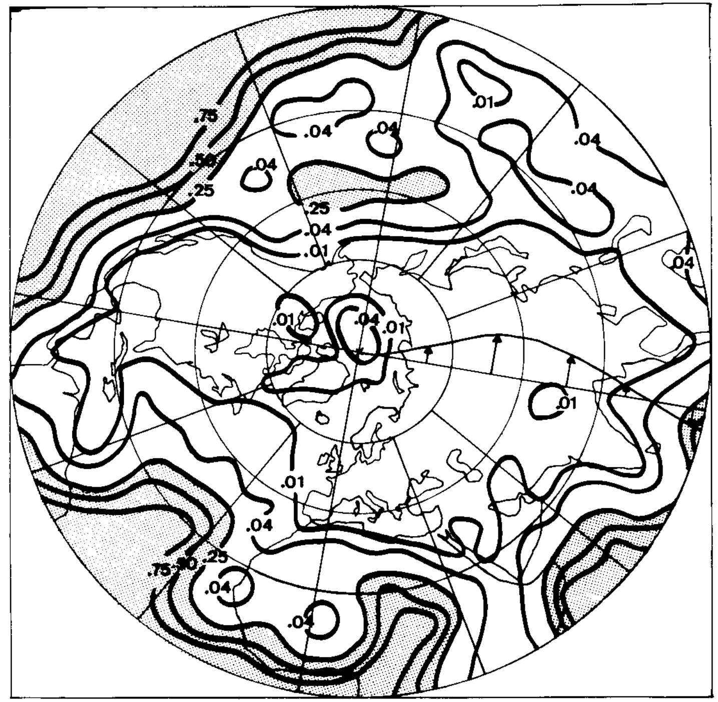

Fig. 7. The same as Fig. 6 but for analyses based on wind, height and forecast data. The minimum value is 0.002. Arrows indicate the assumed displacement of the analysis error during the forecasting process.

estimated using only the terms of group (I) in (1). In the next step both groups (I) and (II) have been used. These preliminary steps will be presented only for 1969. The third step is the final one, applying all the terms of I, II and III.

4.1. *Step 1: only height observations available*

The error of the analyses based only on the height observations is computed using the autocorrelation functions of both Gandin (Fig. 4) and Julian and Thiebaux (Fig. 5). In areas with a sparse network Gandin's function is more optimistic, i.e. the estimated analysis error is smaller. The largest differences are found at about 40° N, 170° E, where the estimate based on Gandin's function is 0.3 smaller than that in Fig. 5. Later on, only the autocorrelation of Julian and Thiebaux will be applied.

An increase in the relative observation error variance from 0.02 to 0.04 caused no marked changes. Most essential is the difference over a dense network; the increase in the observation error was followed by an increase in the minimum of the analysis error from 0.004 to 0.007, in terms of the relative variance.

The differences discussed above illustrate the accuracy of the present computations because these differences are due to the different basic assumptions.

4.2. *Step 2: height and wind observations available*

The effect of the wind observations has been computed by assuming that every station given in the station list has observed both the height and the wind. The result is given in Fig. 6. Over large areas the analysis error is now much less than half of the observation error ($0.5 \times 0.02 = 0.01$). Over the North Pole the introduction of the wind observations has changed the situation from intolerable to nearly acceptable. Large differences between Figs. 5 and 6 are found over the regions which are approximately limited by the isolines 0.04 and 0.75 in Fig. 6, the maximal differences being located between isolines 0.25 and 0.50. These latter isolines thus give the region where the wind observations most effectively reduce the relative analysis error. In this region the analyses with and without wind observations may be essentially different. Thus the region is the one where systematic errors due to the utilization of wind observations are most easily incorporated into the analyses.

4.3. *Step 3: effect of the forecasts*

In order to simulate the effect of forecasts, the error of the analyses based on wind and height (as in Fig. 6) is first advected along an assumed basic stream for 24 hours; the forecasting error is then computed from the advected analysing error using (6). This forecasting error is thus one of the "24-h" forecasts. If we had used "12-h" forecasts, the estimated forecasting error would have been smaller, and, as a consequence, the estimate of the analysis error would have been locally smaller. However, by advecting for 24 hours we let the small analysis error of the dense station net act widely over the regions of a sparse network. Thus the effect of the forecasts will be greatest where extra information is most necessary.

The assumed advection speed depends only on the latitude. The 24-h displacement of the advected analysis error is illustrated in Fig. 7. In comparison with Fig. 6, the main area of the influence of the forecasts in Fig. 7 seems to be that of $E_0/m_{00} > 0.25$. This is thus the region where systematic errors in the forecasts are most easily incorporated into the analyses. The relative analysis error over the North Pole is only a little larger in Fig. 6 than in Fig. 7. The forecasts have not affected the analyses because of the assumed weak advection in the north, i.e. there was only a minimal flow of information from known areas to unknown.

Finally we briefly compare the effects of the analysis errors and model errors in the "24-h" forecasts.

If the prediction error of a numerical model is given by (6), then the influence of the errors in the initial analysis dominate the influence of the model errors if $3\ E_0/m_{00} > 0.08$ or $E_0/m_{00} > 0.03$. Hence the isoline of 0.04 in Fig. 7 divides the area approximately into two different parts. When $E_0/m_{00} > 0.04$ the observational system should be developed, whereas over the areas of $E_0/m_{00} < 0.04$ an improvement in the prediction model is recommended.

5. The historical development of the analysis error

In the following, the error will be given only north of 20° N, because the present approximations are not sufficiently valid at southern latitudes. Absolute error values ($= \sqrt{E_0}$) will be presented

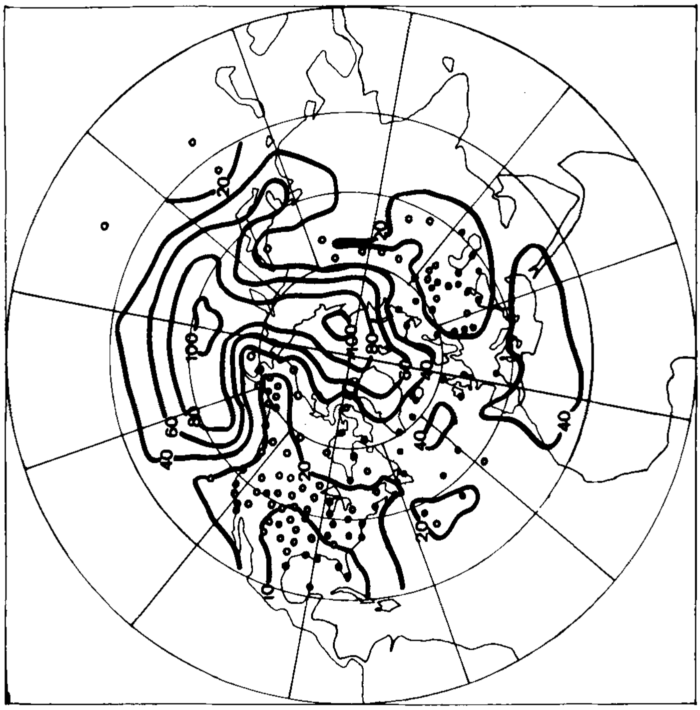

Fig. 9. The analysis error on Jan. 21, 1946.

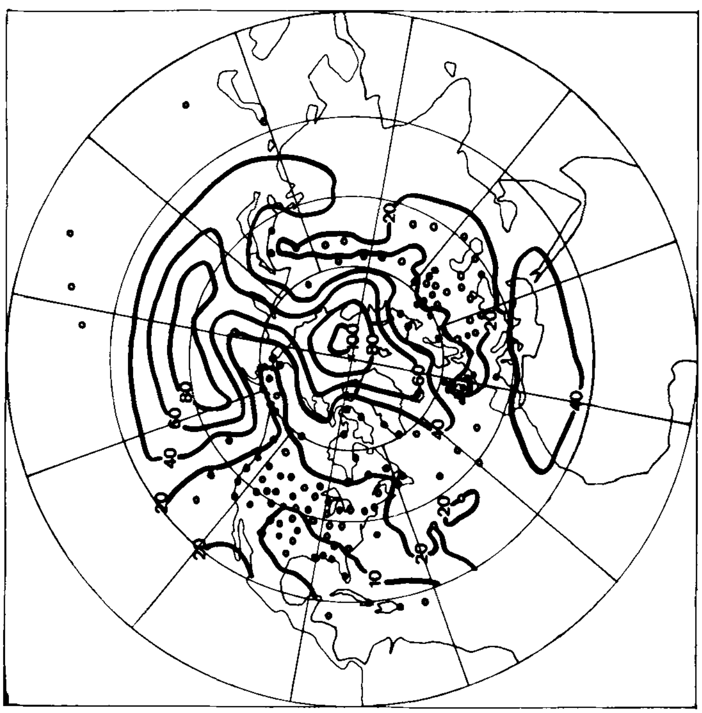

Fig. 8. The station network and the analysis error of Jan. 13, 1946. Unit: metre. In the south the metric values may be misleading.

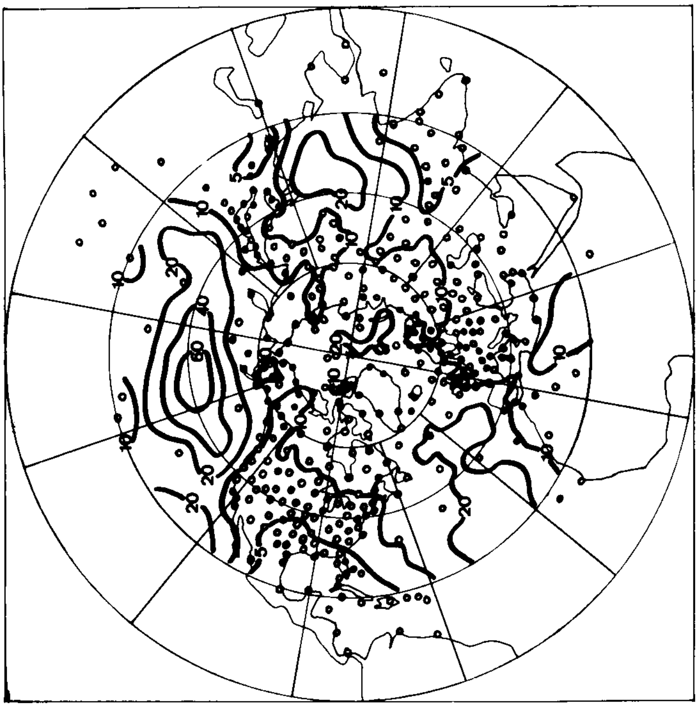

Fig. 11. The analysis error for Dec. 1957.

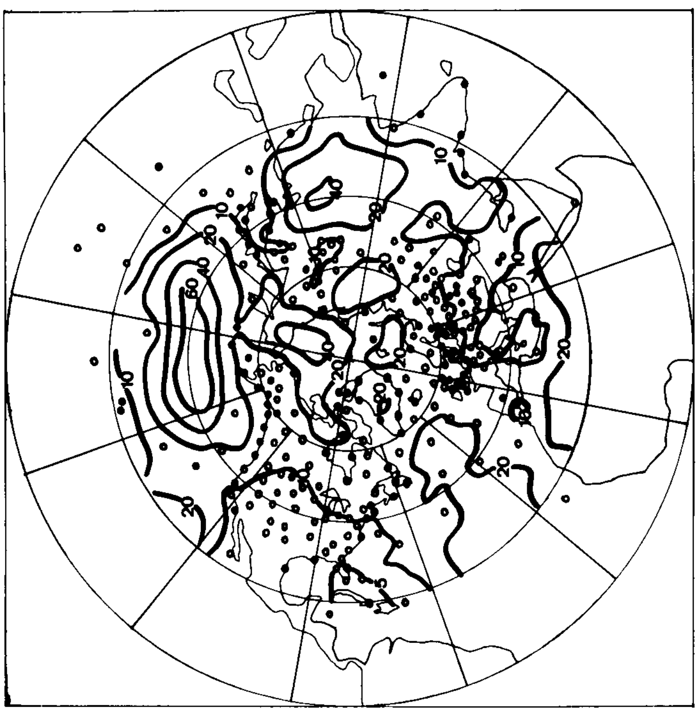

Fig. 10. The analysis error for Nov. 1952.

instead of relative ones (E_0/m_{00}). The absolute values are computed by multiplying the relative ones by m_{00}, the variance of the height anomaly. In the south, where m_{00} is small, the absolute values may be misleading. Nevertheless they will be used because the metric values may be easier to read. Quantity m_{00} will be a long-term (20 years) mean of $[z(i,t) - (zm(i) + C_1(t)_1(i))]^2$, where $z(i,t)$ is the height at grid point i, zm is its long-term mean and $C_1(t)f_1$ is the first empirical orthogonal function describing the variation associated with the annual wave. Thus the annual variation is not included in m_{00}.

In 1946 the analysis error was unacceptable nearly everywhere (Figs. 8 and 9). "Black holes" (error over 40 m) cover large areas. The station nets of January 13, 1946 and January 21, 1946 differ, especially over the Pacific, Europe and around Greenland and Iceland. The error patterns, however, are rather similar.

The rapid development from 1946 to 1952 is indicated by Fig. 10. The area covered by black holes has become smaller.

A clear improvement is observable from 1952 to 1957. Obviously the analysis error tended to decrease after 1952, though more slowly than before. Intensification of the observational work during the International Geophysical Year accelerated progress. Perhaps the influence was most marked at the end of the IGY. Our case, December 1957, is in the middle of the IGY (Fig. 11). The analysis error is below 20 m over large areas. The black hole close to the North Pole is removed thanks to the presence of some extra stations. Reports from China are still missing.

Developments after the IGY require a more detailed discussion. We have assumed, that the variance in the observation error of the additional (additional with respect to the radiosonde network) data sources can be estimated with the "24-h forecast" error variance. Thus we have omitted flight reports, for instance. In Synoptic Weather Maps, a specific plotting symbol of flight reports was given for the first time in January 1959. Thereafter some reports are usually plotted on the maps. The influence of the additional data sources should not be overestimated. One should first know how effectively the data were used. The observation error in the flight reports may have been large, the number of reports is often small and they often appear in areas where the error variance in Fig. 12 is smaller. The additional data obviously have had an effect and therefore our variance estimates of the analysis error in 1969 are too large. On the other hand, the criterion used in collecting the station list of 1969 differs from other cases. Therefore the number of stations may be too large and correspondingly the values in Fig. 12 too small.

Because of these inaccuracies in 1969 and because of the IGY, developments during the 60s cannot directly be derived from Figs. 11 and 12. The change over China is clear; the weather reports have become available. The differences at the North Pole are due to the extra IGY stations. Over the Pacific, the change in the analysis errors depends on changes in the activity of the weather ship, previously at 34° N, 164° E. The change over Africa is due to a decrease in the number of stations. The change may be larger than revealed by Figs. 11 and 12 because we have omitted stations observing wind only.

The following figures summarize the history of the analysis error. The table gives the minimum of the relative variance of the error, the maximum of the absolute error, and the standard deviation of the areal average of the absolute error, all computed over the region north of 20° N:

	$\min(E_0/m_{00})$	$\max(\sqrt{E_0})$	$\sqrt{\text{mean}(E_0)}$
1946	0.01	105–110 m	43–46 m
1952	0.006	76 m	22 m
1957	0.004	73 m	18 m
1969	0.0025	76 m	18 m

The first column illustrates changes over the densest network. The values can be compared with the assumed observation error, which is 0.02 for 1969. If this corresponds to a metric value of 20 m, then the minimum of 1969 (= 0.0025) corresponds to 7 m. Note that even the value for 1946 was acceptable.

The second column shows that, after 1952, the analysis error over the worst network has not changed. The analysis error is large and is comparable to the forecasting error for 1–2 day forecasts.

According to the third column, the average error decreased rapidly in the 40s. The rate of decrease then became slower and no change is observable after 1957.

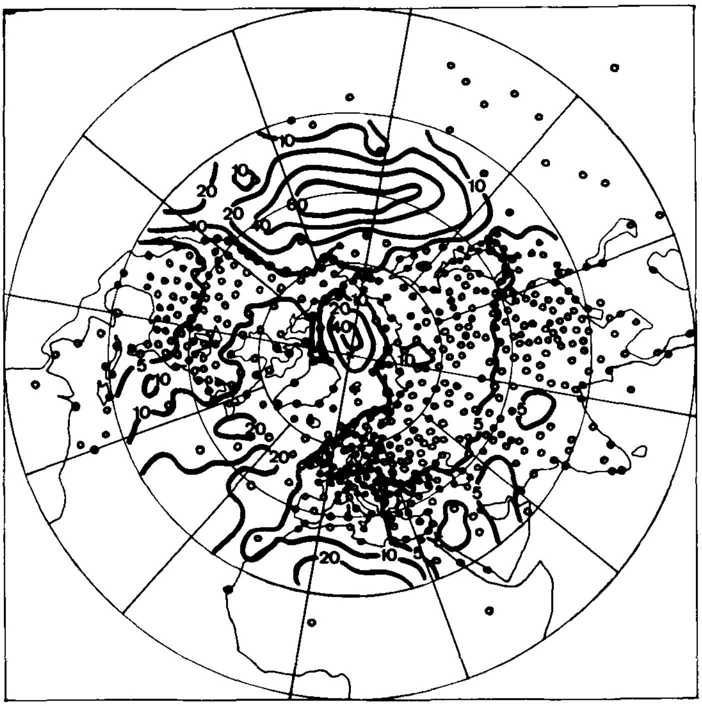

Fig. 12. The analysis error for 1969.

6. Critical comments

Some of the error sources of the present computations will be mentioned in the following. The assumed analysis method was that of optimum interpolation, which differs from other methods used in analysis centres, such as the manual method or the method of successive corrections.

It has been assumed that all the stations mentioned in the list observed both wind and height. Otherwise it would have been necessary to make separate lists for every day, because the group of stations observing the wind has varied. This is due to the use of the optical theodolite. Thus the influence of optical wind observations is overestimated. On the other hand, the assumed error (η^θ) of the earlier wind observations is large, which reduces the effect of the observations. We have omitted stations observing wind only, and this causes an overestimation of the error variance, especially over Africa and India. After all, the introduction of the wind into the computations is only approximative because of the assumption of $\eta^\theta = \eta^z$.

7. Comparison with results published by other authors

Bengtsson (1976, his Fig. 4.5) has presented a map showing the error variance of the 500 mb height. His station list and analysis method are more complete, but the computations, like the present ones, are based on height and wind radiosonde observations. Though there are differences in details, the locations of the maxima are similar. Our maximum over Africa is stronger. Obviously differences are due mainly to differences in the station network.

The estimation of the analysis error depends on the assumptions concerning the observation and forecasting errors. Our "forecasts" or "preliminary fields" have a relatively large error. They do not include all the extra information available. For this reason our absolute values may be too large.

Bengtsson gives values (his Table 5.4) of 5.3 m and 9.5 m in network cases of "conventional dense" and "like North Atlantic". In the latter case, flight reports are included. Our values (case of 1969) over Europe and the North Atlantic are around 7 m and 20 m. Larsen et al. (1977) have computed the analysis error of a network which is more complete than any ever used in routine analysis systems. The error variance in our cases should therefore not be smaller than theirs. Larsen et al. give standard errors of analysis of 15, 9 and 14 m for latitudes 80, 60 and 40° N, respectively. Our values (1969) are about 20, 10 and 25 m. Thus, though our absolute values possibly are too large, they generally are not much overlying.

8. Summary

From 1946 to 1957 there was, in the Northern Hemisphere, a rapid decrease in the variance of the analysis error of the 500 mb height. Thereafter the variance changed slightly, and even increased locally. However, the results of this paper are valid only if it can be assumed that the additional data (additional with respect to the radiosonde network) were of minor importance.

In the late 40s, the drop in the analysis error was due to improvements in instruments and telecommunications and to an increase in the number of stations. Later on, the last factor dominated. However, the decrease in the analysis error was concentrated in regions with a smaller error so that regions with a large analysis error have retained this feature. This is true for the Pacific, the North Pole, the South Atlantic and the oceans around India.

In the case of a sparse network, the wind observations and forecasts play a major role. Here the analysis error may include systematic errors due to systematic errors in forecasts and in the formula relating the wind to the height. In 1969, areas where the systematic prediction error may have had a stronger effect are given by $E_0/m_{00} > 0.25$ in Fig. 7. Areas of possible systematic errors due to to the application of wind data are given by $0.75 > E_0/m_{00} > 0.04$. In addition, it is estimated that, in order to improve forecasts, the analysis system should be developed when $E_0/m_{00} > 0.04$ and the prediction model when $E_0/m_{00} < 0.04$.

In order to get a more precise description of developments during the 60s (and 70s) an analyzing centre should first be fixed. It should then be verified how widely it did, in actual fact, apply modern analysis methods and how effectively it used data from new observations systems.

9. Appendix

9.1. *Expansion of formulae (1) and (2)*

Expression (1) can be written as

$$E_0 = \overline{z_0^{\text{true}} z_0^{\text{true}}} - 2\left[\sum_{i=1}^{N} p_i \overline{z_i z_0^{\text{true}}} + \sum_{i=N+1}^{2N} p_i \overline{\vartheta_i z_0^{\text{true}}} + p_{2N+1} \overline{z^{\text{fore}} z^{\text{true}}}\right] + \overline{\left[\sum_{i=1}^{N} p_i z_i + \sum_{i=N+1}^{2N} p_i \vartheta_i\right]^2} + p_{2N+1} p_{2N+1} \overline{z^{\text{fore}} z^{\text{fore}}} + 2\left[\sum_{i=1}^{N} p_i p_{2N+1} \overline{z_i z^{\text{fore}}} + \sum_{i=N+1}^{2N} p_i p_{2N+1} \overline{\vartheta_i z^{\text{fore}}}\right] \quad \text{A(1)}$$

The third last term in this expression may be expanded as

$$\sum_{i=1}^{N}\sum_{j=1}^{N} p_i p_j \overline{z_i z_j} + 2\sum_{i=1}^{N}\sum_{j=N+1}^{2N} p_i p_j \overline{z_i \vartheta_j} + \sum_{i=N+1}^{2N}\sum_{j=N+1}^{2N} p_i p_j \overline{\vartheta_i \vartheta_j} \quad \text{A(2)}$$

Thus error variance E_0 is given in terms of covariances and variances of height and wind anomalies. Weights p_i are solved from equation group (2). When $i \le N$, for instance, we have

$$\overline{z_0^{\text{true}} z_l} - \sum_{i=1}^{N} p_i \overline{z_i z_l} - \sum_{i=N+1}^{2N} p_i \overline{\vartheta_i z_l} - p_{2N+1} \overline{z^{\text{fore}} z_l} = 0 \quad \text{A(3)}$$

Hence the weights are obtained as functions of the covariance terms.

10. Acknowledgements

This study was supported by the Academy of Finland. Pirkko and Timo Saarikivi assisted in preparing the manuscript.

REFERENCES

Bengtsson, L. 1976. Initial data and some practical aspects of weather forecasting. In *Weather forecasting and weather forecasts. Models, systems, and users* (Ed. A. Murphy and D. Williamson), NCAR, Colorado, Vol. 1, 254–419.

Buell, C. E. 1972. Correlation functions for wind and geopotential on isobaric surfaces. *J. Appl. Meteor. 11*, 51–59.

Gandin, L. S. 1963. *Objective analysis of meteorological fields, Leningrad*, Gidrometeoizdat.

Harrison, D. N. 1962. The errors of the Meteorological Office radiosonde Mark 2B. Met. Off. Paper 15, London.

Headquarters, Air Weather Service, AAF 1946. Nothern hemisphere historical weather maps, sea level and 500 millibars.

Helminen, J. 1972. Gandinin analyysimenetelmä (in Finnish), Report No. 2, Dep. of Met. Univ. of Helsinki.

Hooper, A. H. 1975. Upper air sounding studies, Vol. I, Studies on radiosonde performance. WMO Technical Note 140.

Julian, P. R. and Thiebaux, H. J. 1975. On some properties of correlation functions used in optimum interpolation schemes. *Mon. Wea. Rev. 103*, 605–616.

Larsen, G., Little, C., Lorenc, A. and Rutherford, I. 1977. Analysis error calculations for the FGGE, International Report 11, Research Department, European Centre for Medium Range Weather Forecasts, Bracknell, England. (Permission to quote obtained.)

McInturff, R. M. and Finger, F. G. 1968. The compatibility of radiosonde data at stratospheric levels over the northern hemisphere, ESSA Techn. Mem. WBTM DATAC 2.

Schlatter, T. W. 1975. Some experiments with a multivariate statistical objective analysis scheme. *Mon. Wea. Rev. 103*, 246–257.

Thiebaux, H. J. 1975. Experiments with correlation representations for objective analysis. *Mon. Wea. Rev. 103*, 617–627.

U.S. Weather Bureau, 1952, 1959, 1967. Daily series synoptic weather maps, northern hemisphere sea level and 500-millibar charts.

U.S. Weather Bureau, 1957. IGY world weather maps, Part I, Northern hemisphere, sea level and 500-millibar charts.

U.S. Weather Bureau, 1963. Northern hemisphere RAOB and Rawinsonde network, January 1962, Asheville (map).

WMO, 1953, 1960, 1968, 1973 Oct. Weather reporting, observing stations or weather reports, stations, codes and transmissions, Vol. a, No. 9 TP 4, Geneva.

WMO, 1971. Guide to meteorological instruments and observing practices, 4th ed., Geneva.

WMO, 1976. Report of the eighth session of the tropical experiment board. GARP Special Report 20, Geneva.

Vockeroth, R. E. 1975. Upper air sounding studies. Vol. II, Manual computation of radiowind. WMO Technical Note 140.

РАЗВИТИЕ ОШИБКИ АНАЛИЗА ВЫСОТЫ ПОВЕРХНОСТИ 500 МБ ЗА 1946-1969 ГОДЫ

Ошибка анализа высоты поверхности 500 мб оценивается как функция времени. Параметрами являются ошибки наблюдений, ошибки прогноза и сеть радиозондовых станций, с которых центр анализа получает данные наблюдений. При определенных предположениях средняя ошибка анализа к северу от 20°с.ш. найдена равной 45 м, 22 м, 18 м и 18 м для определенных периодов 1946, 1952, 1957 и 1969 годов, соответственно. Соответствующие максимальные ошибки равны 108 м, 76 м, 73 м и 76 м. Определены области с возможными систематическими ошибками, производимыми при прогнозе и при наблюдениях ветра.