551. 509. 314: 509. 5: 547 .5

INVESTIGATION OF THE FORECASTING ERROR OF A SIMPLE BAROTROPIC MODEL WITH THF AID OF EMPIRICAL ORTHOGONAL FUNCTIONS

PART III

by

JUHANI RINNE

Department of Meteorology,
University of Helsinki

Abstract

In two independent samples 57 empirical (horizontal) orthogonal functions (e.o.f.'s) accounted for 96 per cent of the variance of the altitude of the 500 mb surface. The distribution of the residual in an independent sample partly follows that of the expected error of analysis. An analysis reconstructed with e.o.f.'s seems to be at least as correct as the original analysis. The method of correcting forecasts, developed in the earlier parts of the study, reduced the error variance by 21.6 per cent in routine 72-hour forecasts. It is estimated that, in optimum use, the method would reduce the error variance of 24, 48, 72 and 96-hour forecasts by 10, 20, 30 and 40 per cent, respectively. The method reduces the error all over the area and in all seasons of the year. An essential part of the method is a generalized smoother.

1. *Aim*

The aim of this paper is to verify the effectiveness of the method of correcting forecasts derived in Parts I and II [7 , 8].

2. *The method*

The method consists of the following stages:

A. Basic work

1. Compute the mean field $zm(i, j)$ and horizontal e.o.f.'s (empirical orthogonal functions) $f_\nu(i, j)$ from a sample of analyses $z(t, i, j)$, where i and j are grid point indices and t refers to time.

2. Compute the coefficients $D_\nu(t)$ and $C_\nu(t)$ from a sample of forecasts $z'(t, i, j)$ and from a sample of the corresponding verification analyses $z(t, i, j)$ by means of the equations

$$D_\nu(t) = \sum_{i,j} [z'(t, i, j) - zm(i, j)] f_\nu(i, j) dA(i, j) \tag{2.1}$$

and

$$C_\nu(t) = \sum_{i,j} [z(t, i, j) - zm(i, j)] f_\nu(i, j) dA(i, j) , \tag{2.2}$$

where $dA(i, j)$ is the element of area and $D_\nu(t)$ and $C_\nu(t)$ are the coefficients of the series

$$z'(t, i, j) \approx zm(i, j) + \sum_{\nu=1}^{N} D_\nu(t) f_\nu(i, j) \tag{2.3}$$

and

$$z(t, i, j) \approx zm(i, j) + \sum_{\nu=1}^{N} C_\nu(t) f_\nu(i, j) . \tag{2.4}$$

Find the best regression model

$$C_\nu(t) \approx \alpha_\nu + \sum_{\mu=1}^{N\mu} \beta_{\nu\mu} D_\mu (t) \qquad (\nu = 1, 2, \ldots . N) . \tag{2.5}$$

B. Application

3. Compute the coefficients D_ν of the actual forecast $z'(i, j)$ from

$$D_\nu = \sum_{i,j} [z'(i, j) - zm(i, j)] f_\nu(i, j) dA(i, j) . \tag{2.6}$$

Compute the new coefficients D''_ν from

$$D''_\nu = \alpha_\nu + \sum_{\mu=1}^{N\mu} \beta_{\nu\mu} D_\mu . \tag{2.7}$$

Find the final corrected forecast $z''(i, j)$

$$z''(i, j) = zm(i, j) + \sum_{\nu=1}^{N} D''_{\nu} f_{\nu}(i, j) . \tag{2.8}$$

3. *Analyses and forecasts*

Objective analyses this study is based upon were performed by using the Döös-Bergthorsson-Cressman method. The utilized data has been the monthly mean field, the preceding short-range forecast and the height and wind observations. The computations consist of successive scans with varying test and weight parameters.

During the first 36 (30) forecasting hours, when the initial analyses were valid at 00 GMT (12 GMT) the forecasting model was a baroclinic, quasigeostrophic, filtered three-parameter model. After these forecasting hours, the computation was continued with a

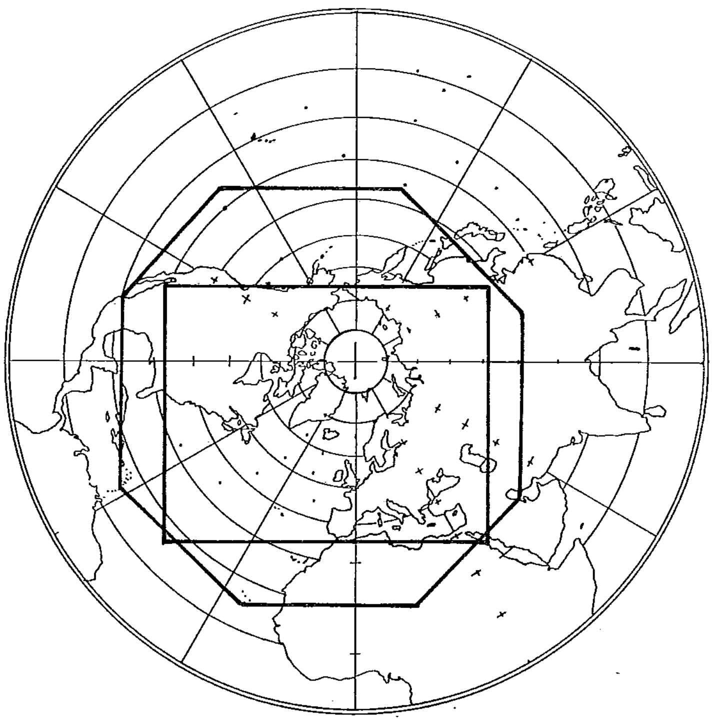


Fig. 1. The areas of the analysis and forecasting procedures (the larger area) and of the method of correction presented here in Part III (the smaller area).

divergent berotropic model. The area of the forecasts was the larger area depicted in Fig. 1. For details see Söderman [9].

4. *Computation of the empirical orthogonal functions*

The number and accuracy of the e.o.f.'s used in the first two parts of this study were found to be too low. Therefore new e.o.f.'s were determined using a combination of all the samples of analyses described in Part II (*cf.* Part II, p. 216). These data consisted of 228 analyses of the altitude of the 500 mb surface in 1965—68, divided into 48 mutually almost non-correlated groups. These objective analyses were issued by the Swedish Meteorological and Hydrological Institute. This sample will be referred to below as sample A.

The mean field $zm(i\,,j)$ used throughout this paper was determined from sample A. When the mean field was computed, the analyses were weighted by $\frac{1}{48n}$, where n is the number of the analyses of the non-correlated group in question.

Details of the e.o.f.'s and of the computation will be described in a later paper.

The area in which the functions are orthogonal is the smaller area in Fig. 1. The grid points of the smaller (present work) and larger (routine forecast) areas coincide. The few grid points in the smaller area that lie outside of the larger area are non-essential.

The first 114 functions were determined from the 228 possible functions. Using the rule of the logarithmic linearity of noise eigenvalues[1]) presented by Craddock and Flood [4], the first 57 functions were found to be representative whereas the remaining functions can mostly be interpreted as noise.

5. *Determination of regression models*

The regression coefficients required for the correction method were computed from a sample of objective analyses and 72-hour forecasts issued by Finnish Meteorological Institute. The sample period was 4th February—31st July 1970 (about six months). The meteorological

---

[1]) estimated from sample B discussed in the following section

variable was the altitude of the 500 mb surface at 00 GMT and 12 GMT. The data were examined preliminarily computing the variance reduction due to the 57 e.o.f.'s. In four analyses (one forecast) the value of the residual variance was alarming. These data fields were then examined in more detail, found to be erroneous, and finally rejected from the sample. Some of the possible data fields were missing. Hence the final number of objective verification analyses (and of 72-hour forecasts) was 313. This sample will be referred to below as sample B.

The coefficients C_ν and D_ν in sample B were computed from Eqs. (2.1) and (2.2). The regression coefficients in Eq. (2.5) were computed by the stepwise regression method. The number N of single regression problems was 49 and the highest number N_μ of possible predictors was 114. Hence it was assumed that $\alpha_\nu = 0$ and $\beta_{\nu\mu} = 0$ if $\nu > 49$ or $\mu > 114$.

A large number of significant predictors[1]) within the meaning of regression analysis were found in each of the 49 single regression problems of model (2.5). As it is difficult to determine the significance of regression coefficients with the aid of dependent samples, the following models were formed:

Model 1.

This is the model given by regression analysis except that a few of the terms have been omitted. The hypothesis that a given regression coefficient $\beta_{\nu\mu}$ differed significantly from zero was t-tested: where $|t|$ was greater than 1.9, the regression coefficient was accepted, otherwise it was assumed that $\beta_{\nu\mu} = 0$.

Model 2.

As above, but with a tolerance limit $|t|$ greater than 4.9.

Model 3.

As above, but with a tolerance limit $|t|$ greater than 9.9. In this model, nearly all of the equations (2.5) were of the form

$$C_\nu(t) \approx \alpha_\nu + \beta_{\nu\nu} D_\nu(t) .$$

---

[1]) Actually the difference between the observed and forecasted amplitude, $C - D$, was the predictand. The resulting model was then converted into the form of Eq. (2.5).

Model 4.

Only the constants α_ν were retained in the model. Since $\beta_{\nu\nu} = 1$, only the mean field of the forecasts was changed (*cf.*Eq. (2.7)).

Model 5.

Also the coefficients α_ν were fixed: $\alpha_\nu = 0$. Hence the forecast remained unaltered except that components ν greater than 49 were omitted.

Of these models only the first and last were consistent. The first was consistent because it was close to the final regression model, the last because it was a pure elimination model (eliminating the terms $\nu > 49$) with no regression coefficients. The others were not true regression models but truncated regression models in which only some of the terms of the initial regression model were accepted. This disturbed the balance between the terms of the models. Model (4) was particularly unbalanced because the constant terms were not mean values of $C_\nu - D_\nu$ but had been selected from a complicated regression model. Truncated models were used instead of true ones to avoid subjectivism in the selection of models. The tolerance limits (1.9, 4.9, 9.9) were chosen to differ clearly from one another. The final construction of the models was performed purely mechanically by a computer program, so no effort was made to find the best t-value with the aid of the independent sample. The purpose of all this was to ensure the objectivity of the final test. No changes were made in the models except to the coefficients α_1 and β_{11} . Their signs during the first half-year were the opposite of those during the second half-year. This procedure can be justified theoretically (*cf.* Part II, Chap. 8). The effect of changing the signs was not tested.

6. *The independent sample*

The test sample was a sample of objective analyses and 72-hour forecasts issued by the Finnish Meteorological Institute between 1st September 1970 and 31st August 1971 (one year). This sample was not examined preliminarily, because the aim was to apply it as independently as possible. It will be referred to below as sample C.

Table 1. Variance values obtained with the aid of 57 empirical (horizontal) orthogonal functions. The mean yearly oscillation was estimated with the aid of the first e.o.f. The functions and the yearly mean field were determined from sample A. Unit: m^2. Variable: the height of the 500 mb surface.

Sample	Number of cases	Including the yearly oscillation				Not including the yearly oscillation			
		Total variance	Variance represented by e.o.f.'s	Residual	Effectiveness of the re-presentation (per cent)	Total variance	Variance represented by e.o.f.'s	Residual	Effectiveness of the re-presentation (per cent)
A	228	29,987	(29,619)	(368)	(98.8)	12,718	(12,350)	(368)	(97.1)
B	314	27,878	26,791	1,087	96.1	12,245	11,158	1,087	91.1
C	723	31,091	29,919	1,172	96.2	14,053	12,881	1,172	91.7
C:									
Sept. 70	60	23,835	22,938	897	96,2	10,433	9,536	897	91.4
Oct.	61	17,435	16,274	1,161	93.3	16,554	15,392	1,161	93.0
Nov.	60	25,138	23,932	1,205	95,2	18,485	17,280	1,205	93.5
Dec.	61	34,450	33,103	1,347	96.1	15,965	14,617	1,347	91.6
Jan. 71	60	53,450	52,009	1,441	97.3	20,921	19,480	1,441	93.1
Febr.	56	37,986	36,533	1,453	96.2	21,105	19,652	1,453	93.1
Mars	61	28,736	27,295	1,441	95.0	16,158	14,718	1,441	91.1
Apr.	60	17,499	16,125	1,374	92.2	15,027	13,652	1,374	90.9
May	62	12,215	11,002	1,214	90.1	9,382	8,168	1,214	87.1
June	60	33,772	32,851	921	97.3	10,273	9,352	921	91.0
Jul.	61	49,992	49,157	834	98.3	7,866	7,032	834	89.4
Aug.	61	39,433	38,637	797	98.0	7,197	6,400	797	88.9

7. *The residual of the e.o.f. analysis*

Some characteristic values of the 57 e.o.f.'s are given in Table 1. As the first function represents the yearly oscillation, the quantity $zm(i,j) + C_1(t)f_1(i,j)$ roughly represents the (daily) climatic conditions and hence the quantity

$$\sum_{i,j} \overline{\{z(t,i,j) - [zm(i,j) + C_1(t)f_1(i,j)]\}^2} dA(i,j)$$

approximately represents the variance of $z(t,i,j)$ from the climatological conditions (the values »not including the yearly oscillation» in Table 1). The variances »including the yearly oscillation» in Table 1 are those computed from the mean field $zm(i,j)$. The values determined from sample A are given in brackets as, in this case, sample A is a dependent sample.

Though the absolute variances and reductions of samples B and C differ, the relative reductions agree well. Hence it can be concluded that the variance represented by the 57 e.o.f.'s was 96% (of the total variance) including the yearly oscillation, and 91 . . . 92% not including it.

The area distribution of the variance of the residual was defined as

$$\overline{\{z(t,i,j) - [zm(i,j) + \sum_{\nu=1}^{57} C_\nu(t)f_\nu(i,j)]\}^2} . \qquad (7.1)$$

Fig. 2 shows the corresponding r.m.s. values. It can be expected that the distribution of the residual reflects that of the error of analyses. Before discussing the distribution of the residual, certain sources of error in our objective analyses will be mentioned:

1) The error of analyses is probably larger over regions lying at a great distance from the locality at which the analyses were made because these regions were not so important from the viewpoint of this central locality. Typical reasons for such errors are delays or disturbances in the transmission of data.

2) The density of the observational network is the most important factor affecting the error of analyses. Factors 3 and 4 below usually have an effect only in cases where the observational network is sparse.

3) In the analysis procedure a weighted mean of the preceding forecast and of the climatological conditions is used as the first guess. Hence, where the forecasts are bad, or

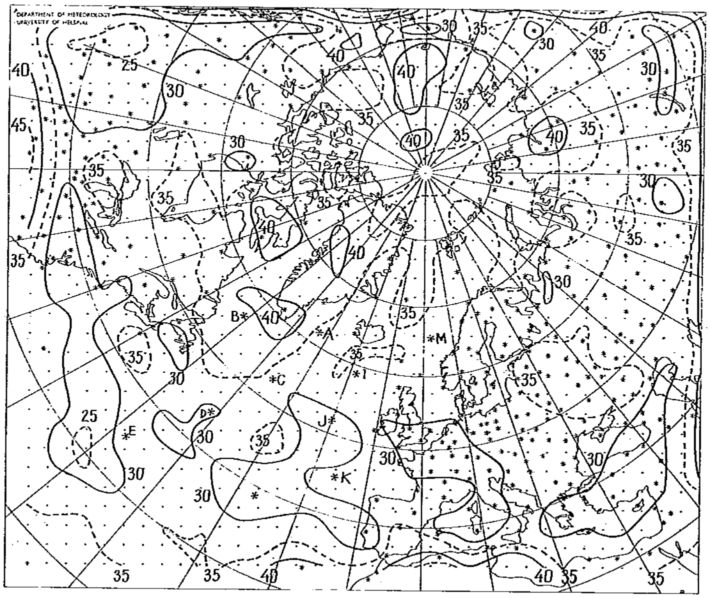


Fig. 2. The distribution of the r.m.s. residual (57 empirical orthogonal functions). Unit: m.
The observation stations used for the analyses studied in Part III are shown by asterisks.
The map is a polar stereographic projection, true at 60°N, where the grid interval is 300 kms.
The solid isolines denote full tens of metres, the broken lines units of five.

4) where deviations from the climatological mean are large, the error of analyses can be expected to be larger. It is possible that if climatological data were not used in the analyzing procedure, the error would still be related to the variance of the quantity to be analyzed. In Table 1, the residual term seems to depend on the variance »not including the yearly oscillation».

Turning back to Fig. 2, it can be seen that the distribution of the residual variance is fairly smooth. Nearly all the values lie between 30 m and 40 m, 35 m being a typical value. The minima are found where the synoptic activity is less intense and/or where the observational network is dense.

The field presented in Fig. 2 was analyzed before the stations were plotted on the map. This was done to emphasize the fact that the computation and analysis of the map were totally independent of the location of the observation points. The analysis in Fig. 2 is perhaps more detailed than it should have been in view of the data available. It was made so in order to bring the details out clearer. There is a belt of large residuals along the boundaries and a parallel belt of small residuals just inside them. All the analyses used in this study have been taken from a larger grid. Hence the large boundary values cannot have been due to boundary effects in the initial analyses, except at the corners. They may have been due to the paucity of data at the boundaries. Much of the capacity of the method of empirical orthogonal functions goes on minimizing residuals near the boundaries, since there are no data outside the boundary. This may result in fairly low values at points near the boundaries, where more information is available, which could be an explanation of the small residuals. The boundary effects, as such, seems to have been limited to the three outermost rows of points. The largest values, about 50 m, are found at the top in Fig. 2.

GANDIN ([5], p. 118, Fig. 25) has theoretically determined the distribution of r.m.s. error in an »optimum interpolation of the height of the 500 mb surface». The peaks and the minima of his results and of ours (Fig. 2) seem partly to coincide. In both, the most important peak over the Atlantic is located exactly half way between ships D and K, and there are roughly coinciding minima over Western Europe. Also coinciding are a Biscayan maximum and peaks west and east of Gibraltar. Similarly Gandin's results include peaks south and north-east of Iceland though these lie more to the north-east than ours. Ridges south-east and south-west of Novaja Zemlja seem to be present in both cases. The conditions over Novaja Zemlja are contradictory. The high values in our map are difficult to explain. The results over the continental areas are not directly comparable as the observational networks differed. Nevertheless, there is a distinct peak at 60°N, 80°E in both cases. In addition our minimum east of the dense network around the White Sea is weakly represented in Gandin's map. There is some evidence of larger values north of the Black Sea and north-west of the Caspian Sea in Gandin's results. Nevertheless, the intense maximum east of the Baltic Sea in our map is not satisfactory.

A less detailed hemispheric map has been published by BELOUSOV *et al.* ([2], p. 131, Fig. 3.9). Here, too, there are maxima above central Greenland and between the Bering Straits and the North Pole, but the peaks over the Baffin Islands and over the southernmost station in Greenland on our map seem to lie farther over the ocean on theirs. No Great Lakes peak can be seen in Belousov *et al*'s map.

There are strong contrasts along the east coast of North America in our results. For instance, the gridpoint S.S.E. of ship B has a value of 40 m (not shown in Fig. 2) and the minimum south of New Foundland is 29 m. In terms of variance, the residual has grown double over a distance of two grid intervals. The area in question seems to be subject to a maximum of forecasting error. Since, moreover, the synoptic activity is rather intense and the observational network sparse, the marked contrasts may be connected with the use of forecasts and climatological conditions as the preliminary analysis field.

Turning now to the absolute residual values, it is obvious that the values over the areas with dense networks are too large to be interpreted as an analysis error. For example, the isoline for 30 m over Central Europe agrees well with the corresponding isoline for 20 m in Gandin's map (*op.cit.*). The values associated with sparse networks are not comparable, as our analyses are based on more abundant data than Gandin's. However, the residual patterns over the sparse-network areas seem to reflect the distribution of the error of analysis. On the other hand, the residual patterns over the dense-network areas cannot always be identified as reflecting errors of analysis. Such a distribution of the residual might be explained as follows: series (2.4) should have less terms over the sparse-network areas, because the scale of the true phenomena observed is larger there. On the other hand, since more information is available in a dense network, the number of significant terms in Eqs. (2.4) should be larger over the dense-network areas. If this is true, the present case ($N = 57$) is a compromise. Decreasing the number of terms causes the error of the e.o.f. analysis to increase in the case of a dense network, whereas increasing the number would add insignificant terms to series (2.4) in the case of a sparse network. Further, as the present residual values are too large over continents, the residual includes rejected true information and the residual patterns may differ from the analysis error patterns. Over oceans the residual values are too small and the residual itself represents nothing but error. Since the error terms have been removed, the e.o.f. analysis over the oceans should be more accurate than the initial analysis. Over continents the most correct analysis lies somewhere between the initial analysis and the e.o.f. analysis. At all events, the e.o.f. analysis seems to be at least as usable as the initial analysis.

Figs. 3 show the mean residual field around the nine oceanic stations denoted by the large asterisks in Fig. 2. The values in the boundary zone ($3\frac{1}{2}$ grid intervals from the edges) have not been used. If there were less than three usable values in a row along a parallel of latitude for the station in question, these values were not used. If there were more data in the row, the missing data were estimated from the mean value for the row. Variations in the map scale have not been taken into account. The residual values increase on moving north. In the east-west direction the lowest relative values are found over the »mean» station, but the point (or line) of symmetry lies downstream of the station, at a distance of about 250 kms. This is obviously due to the use of forecast data in the preliminary field of the analysis. Since they are mean values, the differences in Fig. 3 are smooth. In individual cases, sharp differences can be found. For instance, the difference between the grid-point values in front of[1]) and behind ship I was 335 m^2, and the corresponding value for ship C was 400 m^2; the smaller values were 35 and 40%, respectively, of the larger values (those behind the ships). In experimenting with a rather dense network, MIYAKODA & TALAGRAND [6] found a value of 35% while BENGTSSON & GUSTAFSSON's result [3] with a network simulating the existing network corresponds approximately to 50%, when analysis errors in the cases »with» and »without» forecast information are compared to each other.

It should be pointed out that some of the details discussed above are not significant in a statistical sense. Nevertheless, they are illustrative.

8. The error reduction of the forecasts tested with an independent sample

For conformity's sake, the whole of the smaller area in Fig. 1 has been used as a verification area of the forecasts. Hence the variance of the forecasting error can be defined as

$$\overline{\sum_{i,j} [z(t\,,i\,,j) - z'(t\,,i\,,j)]^2 dA(i\,,j)}$$

This error quantity also includes the errors in the mean field of the forecasts.

---

[1]) *e.g.* eastward (downstream) from the ship.

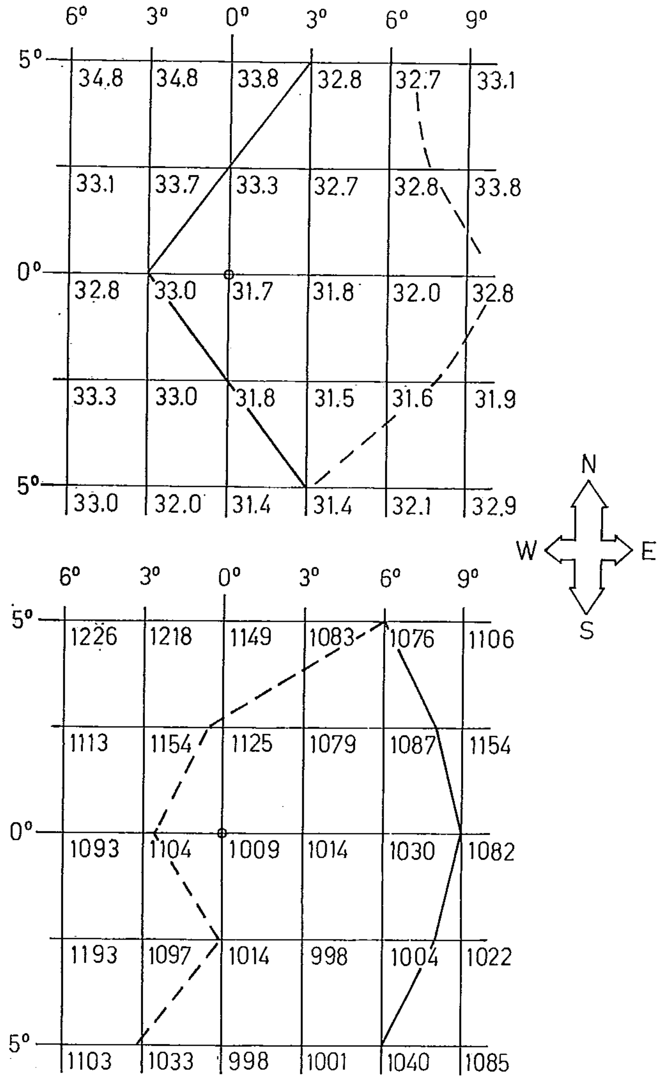


Fig. 3. The residual field (57 eigenvectors) around a mean oceanic station computed as the mean of the residual fields around nine stations (large asterisks in Fig. 2.). Distances are given in degrees of longitude and latitude from the station. The quantities and their units are the mean residual in m (above) and the mean variance in m² (below). Solid lines were drawn arbitrarily. Then points with values equalling the values determined by the solid lines in the east-west direction were analyzed (broken lines). As a result, the point of symmetry seems to lie about 3° east of the station.

Table 2. Error reduction in an independent sample of routine 72-hour 500 mb forecasts.

Month	Number of cases	Initial error variance	Reduction of error variance in m^2					Reduction of error variance as percentages				
			Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
Sept. 70	54	7,546	−1,699	− 344	314	43	162	−22.5	− 4.6	4.2	0.6	2.1
Oct.	60	10,972	− 801	− 321	1,421	107	259	− 7.3	− 2.9	13.0	1.0	2.4
Nov.	60	14,538	1,091	2,342	3,598	287	375	7.5	16.1	24.7	2.0	2.6
Dec.	60	16,708	4,035	5,202	5,752	850	410	24.2	31.1	34.4	5.1	2.5
Jan. 71	58	23,585	2,546	5,101	6,146	850	555	10.8	21.6	26.1	3.6	2.4
Febr.	56	15,321	− 288	693	1,809	335	309	− 1.9	4.5	11.8	2.2	2.0
Mars	60	15,782	1,854	2,692	3,471	436	383	11.7	17.1	22.0	2.8	2.4
Apr.	60	12,360	740	1,946	2,760	412	381	6.0	15.7	22.3	3.3	3.1
May	62	9,406	1,302	2,399	2,820	498	289	13.8	25.5	30.0	5.3	3.1
June	60	6,799	450	1,128	1,397	155	191	6.6	16.6	20.5	2.3	2.8
July	60	5,254	− 607	150	578	−28	126	−11.6	2.9	11.0	−0.5	2.4
Aug.	60	6,067	− 63	383	820	143	256	− 1.0	6.3	13.5	2.4	4.2
Mean	710	12,007	736	1,797	2,588	342	307	6.1	15.0	21.6	2.8	2.6

In principle the method used to reduce the forecasting error was one of the sum of least squares. Thus it can be expected that forecasts with large errors would be improved while those with small errors would deteriorate. This means that errors in the winter forecasts would be decreased most effectively.

Table 2 gives monthly error reduction values. The simplest model (model 5) in which series (2.3) is only truncated at $N = 49$, reduced the error variance by 2.6%, as could be expected. The reduction would have been larger if the series had been truncated at $N = 10 \ldots 20$. This was tried out with sample B, and the result was an error reduction of 6.5%.

The regression model including only the constants a_ν (model 4) was not better than the simplest model. This was obviously due to the inconsistencies in the constants. The constants, too, seem to have been valid only for the winter months; for summer, this model is clearly inefficient.

The most complex model (model 1), including the significant regression coefficients ($|t| > 1.9$) has obviously been exaggerated. The error reduction was 6.1%. In the dependent sample B, it was more than 50%!

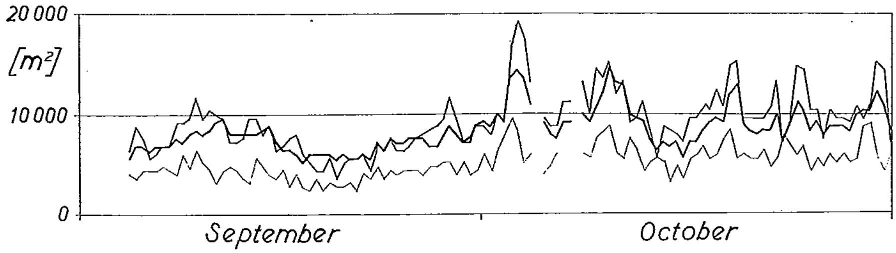
The model including all the very significant terms (model 2) has been a satisfactory one. When it is compared to the best model (3), however, the additional terms of the more complex model (2) seem to be of no use. These terms seem to be more useful during the winter months.

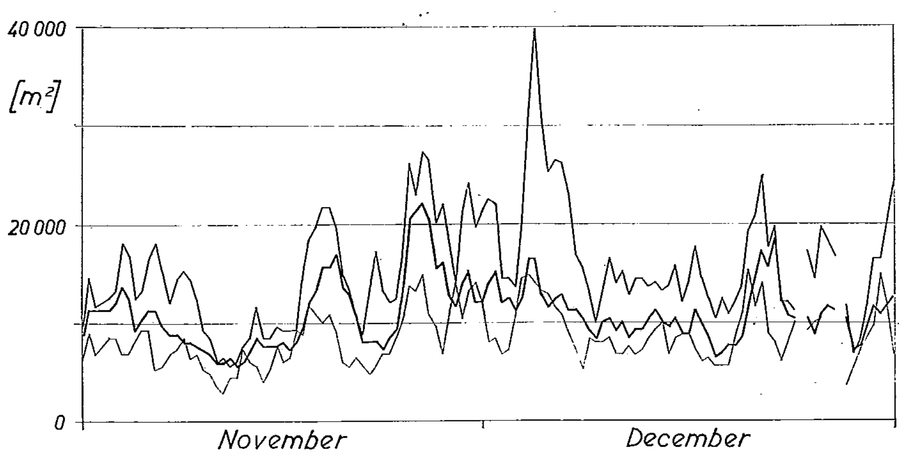
The best model was the regression model (3). It resulted in a remarkable reduction of error especially for the winter and spring months. The yearly error variances of the initial and corrected forecasts were 12007 m^2 and 9419 m^2, respectively. The error variance of the conventional 48-hour forecasts was 6825 m^2. Hence the corrected 72-hour forecasts have the same error variance as »60-hour» forecasts would have had.

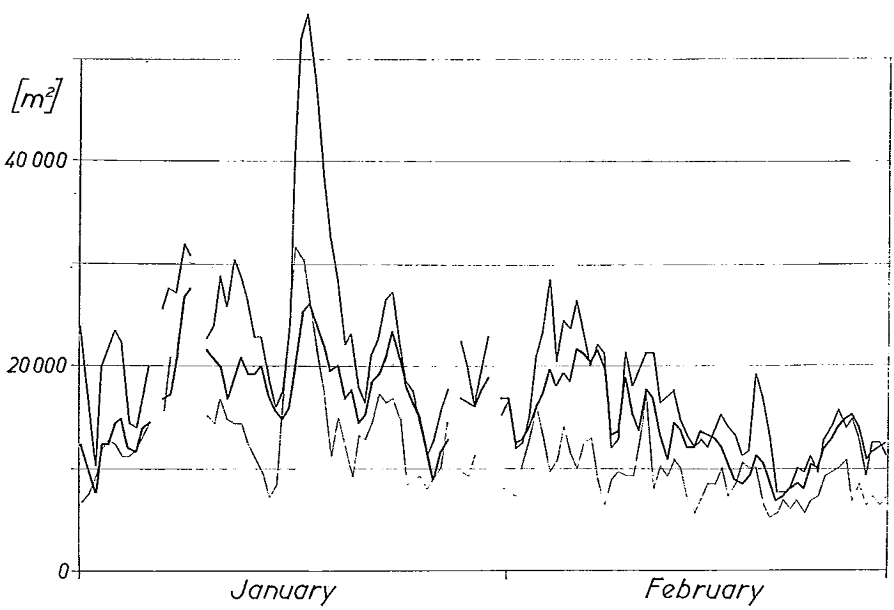
In some cases the error of the corrected 72-hour forecasts was even smaller than that of the corresponding 48-hour forecasts (Fig. 4, Dec., April). However, there were some unfavourable periods — such as September and October. In some cases the error of a bad forecast was reduced strongly (beginning of December), in others large errors remained (end of November).

The distribution of the yearly r.m.s. error,

$$\overline{[z[t , i , j) - z'(t , i , j)]^2}$$


20 000
[m²]
10 000
0
September
October


40 000
[m²]
20 000
0
November
December


[m²]
40 000
20 000
0
January
February

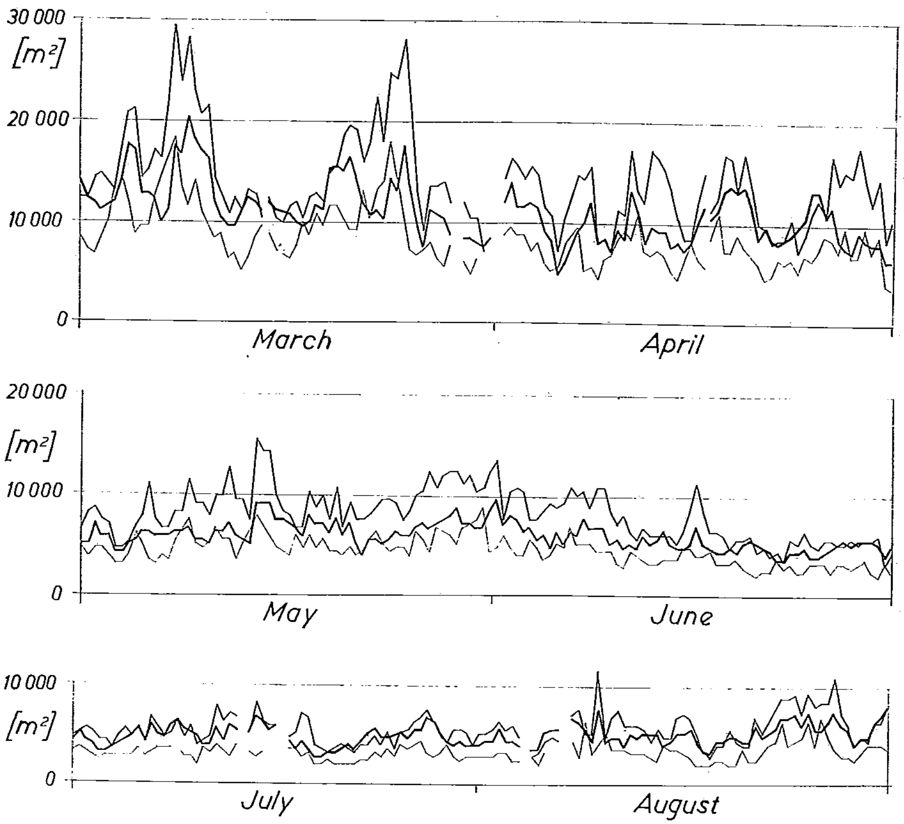


Fig. 4. Error variances of 48-hour (lower thin line) and 72-hour (upper thin line) routine 500 mb forecasts and of corrected 72-hour forecasts (thick line). The latter seems to fall between the routine forecasts. Independent sample.

is shown in Fig. 5 separately for the initial and corrected forecasts. Once again, initially small errors were increased and initially large errors decreased. Owing to errors in the verification analyses, the increase of the error in the southernmost latitudes is not necessarily real; the corrected forecasts may even be better than the initial ones.

From Fig. 5 it is obvious that there was no essential boundary error in the initial forecasts, of the kind found in Parts I and II. The forecasting model is now as correct as possible. Thus it can be concluded that the method of correction has sharply reduced the forecasting error, and that this reduction has occurred all over in space and in time. It must be remembered, however, that the only measure of error considered has been the variance of error.

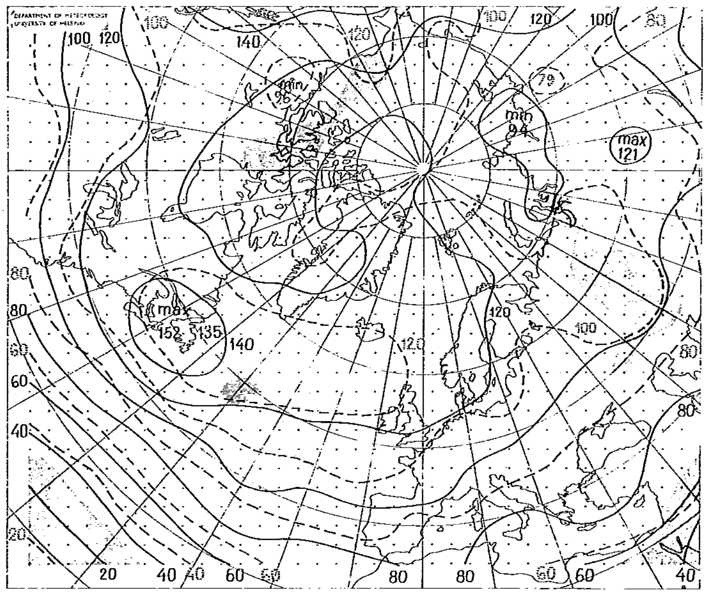


Fig. 5. Distribution of the r.m.s. error of the routine (solid line) and corrected (broken line) 72-hour 500 mb forecsts. The areas in which the method of correction has increased the error are shaded. Unit: m.

9. *Experiments made since the independent test*

Regression model (3) is based on a small sample. The e.o.f.'s were determined from a small sample. The mean field correction was unsuccesful. Hence the value obtained for the error reduction is an underestimate of the true efficiency of the method.

In Part II, it was found to be difficult to depict mean fields with the aid of e.o.f.'s designed to depict anomalies. Optimum use of the method of correction therefore requires a determination of the monthly mean error fields, so that these can be eliminated from the forecasts before applying the method of correction. In sample B, the error variance due to errors in the mean field (averaged over the whole sampling period) was 4.6% in the case of 72 hour forecasts. Thus an additional reduction of 5% can be expected when monthly mean error fields are applied.

How much have the inconsistencies of the regression model affected the results? To find out this, instead of truncating a more complex model, we derived a pure regression model from sample B using the same test values as for the truncated model 3. The resulting reduction of the error was 21.6%. Thus the inconsistency of the model did not essentially affect the results. The best model, including only the terms for which $|t| > 8$ and the constants α_ν resulted in 23% reduction of the error in sample C. The error reduction achieved with this model in the dependent sample B was not much greater (29%). The model

$$C_\nu \approx \alpha_\nu + \beta_{\nu\nu} D_\nu \, , \nu = 1 \, , 2 \, , \ldots \ldots \, 57 \, , \tag{9.1}$$

where only the significant coefficients α_ν and β_ν determined from sample C (a large sample) were accepted, resulted in an error reduction of 29% in sample C. With an independent sample, therefore, this model would have had an error reduction capacity of about 23%. If the correction of the mean field, the yearly variation of the regression coefficients and the small size of the basic sample A were taken into account, the error reduction might have been about 30%. The same model (9.1) had a reduction capacity of 19%, 29% and 37% for 48, 72 and 96-hour forecasts. Hence it can be estimated, that in optimum use, the mean yearly efficiency of model (9.1) could be about 10, 20, 30 and 40% for 24, 48, 72 and 96-hour forecasts.

It was very roughly estimated with the aid of the computed regression coefficients that, in model (9.1),

$$\beta_{\nu\nu} \sim e^{-0.22(T-1)-0.009\nu} \, , (\nu > 1) \tag{9.2}$$

where T is the forecasting period in days, and ν the component number. When the $\beta_{\nu\nu}$ coefficients were obtained from Eq. (9.2), model (9.1) gave reduction values of 18%, 29% and 37% for the 48, 72 and 96-hour forecasts. Thus the differences between the true regression coefficients and coefficients (9.2) do not seem to have affected the results. In this test the three parameters of Eq. (9.2) produced 56 + 56 + 56 coefficients, which were applied to 721 + 710 + 708 forecasts. A test was performed with $e^{-0.22T-0.009\nu}$ owing to a programming error. The resulting reduction values were 12%, 27% and 36%. These can be compared to the values 18%, 29% and 37% from Eq. (9.2) or to the values 19%, 29% and 37% with the true regression coefficients given above. The conclusion is that the method is not very sensitive to variations in the $\beta_{\nu\nu}$ coefficients. Though the tests were made with a dependent sample it is certain that the three-parameter Eq. (9.2) is significant and it accounts for a large part of the efficiency of the method of correction.

10. What does the method of correction achieve?

The method of correction

1) eliminates small-scale components with high indexes. This is the same as smoothing in the conventional sense, except that only non-predictable small-scale components are eliminated.

2) corrects errors in mean field of forecasts. As we have seen above this correction is not an optimum one, and not very efficient.

3) corrects errors in forecasted amplitudes. It was found that the variance of forecasted amplitudes is larger than that of observed ones. Hence the forecasted amplitudes should be multiplied by a factor less than one. The correction is optimum and it may vary for different components. In the case of 72-hour forecasts, for instance, $\beta_{(\nu=45)} = 0.59$ but $\beta_{(\nu=44)} = 0.29$. Thus the filtering power is doubled in the latter case even though the components are adjacent.

4) smoothes amplitudes. This corresponds to smoothing in the conventional sense only in the case of high indexes. Large-scale components are also modified — and in an optimum way. This smoothing of amplitudes seems to be the most efficient part of the method of correction.

What does the correction method fail to achieve? The statement (9.1) does not include terms of the form

$$C_\nu \sim \alpha_\nu + \beta_{\nu\nu} D_\nu + \beta_{\nu\mu} D_\mu \, , \, \mu \neq \nu \, .$$

In other words, there is no mechanism of the following type: If a component μ is in a given phase, then the error of the amplitude D_ν is of a given kind. Significant correlations of this kind were found in sample B but they were applied without success to the independent sample C. Such correlations could perhaps be discovered in a more detailed study if variations with time were taken into account.

11. Points of contact with methods presented by other authors

The present methods corrects the forecasted amplitudes in a similar way under similar conditions. YUDIN *et al.*'s [10] method corrects the error field of a forecast by taking into account error fields observed earlier under analogous conditions. The analogies are determined with the aid of the most important amplitudes of the empirical orthogonal functions, using data from both analyses and forecasts. So there is some connection between the two methods.

BAUER [1] has studied statistical forecasting with the aid of empirical orthogonal functions. In the present case, dynamic forecasts were corrected statistically. Both methods seem to include some smoothing, which is also applied to large-scale waves. The extremely long-range forecasts in both cases tend to resemble a climatological mean.

12. Summary

(Analyses are discussed in addition to forecasts. So the statement made in Chapter 1 is not exactly true.)

The variance represented by the first 57 e.o.f.'s determined from a sample of 228 analyses of the altitude of the 500 mb surface, was 96% and 91—92% of the total variance, with and without the yearly oscillation, respectively. In the dependent sample the corresponding values were 99% and 97%. The distribution of the residual resembled that of the error of analysis, especially over regions with sparse observational networks. Over areas with dense networks the residual values seemed to be too large. The impression is that the residual field was the result of a compromise: additional terms in the e.o.f. analysis might have decreased and increased the error of the analysis over sparse and dense network regions, respectively. It seems unlikely that the number of the terms (57) was too large; it could even have been made larger. The effect of the climatological and forecast data on the analysis error is demonstrated.

The efficiency of the method of forecast correction derived in Parts I and II of this study has been verified with independent test data. The test forecasts were daily forecasts issued under routine conditions. In addition to eliminating non-predictable terms, the method discussed in this paper includes correction of forecasted amplitudes with the aid of a regression model. The best regression model turned out to be a simple one operating mainly as a smoother. In this case, however, the term »smoothing» should not be taken in the conventional sense. The yearly average of the variance reduction of the forecasting error of 72-hour forecasts was 21.6%. It has been estimated that, in optimum use, the correction method could reduce the forecasting error of 24, 48, 72 and 96-hour forecasts by 10%, 20%, 30% and 40% in terms of error variance. The method seems to reduce the error all over the area and in all seasons of the year.

Acknowledgements: My thanks are due to Miss Anneli Pollari for typing the manuscript and to Miss Päivi Leppiaho for drafting the graphs.

This study was started in collaboration with Miss Riitta Järvinen. She made a hard work in performing a large part of the computations associated with sample B. For practical reasons Mrs. Riitta Long could not continue the work that she had so succesfully begun.

REFERENCES

1. BAUER, K. G., 1971: *Linear prediction of a multivariate time series applied to atmospheric scalar fields.* University of Wisconsin, 180 pp.
2. BELOUSOV, S. L., GANDIN, L. S., and S. A. MASHKOVICH, 1968: Computer processing of meteorological data. *Gidrometeorologicheskoe Izdatel'stvo,* Leningrad 1968. Translated from Russian into English, Israel Program for Scientific Translations, Jerusalem, 1971, 210 pp.
3. BENGTSSON, L. and N. GUSTAVSSON, 1971: An experiment in the assimilation of data in dynamical analysis, *Tellus* **23**, 328—336.
4. CRADDOCK, J. M. and C. R. FLOOD, 1969: Eigenvectors for representing the 500 mb geopotential surface over the Northern Hemisphere. *Quart. J. roy. Meteor. Soc.,* **95**, 576—593.
5. GANDIN, L. S., 1963: Objective analysis of meteorological fields. *Gidrometeorologischeskoe Izdatel'stvo,* Leningrad, 1963. Translated from Russian into English, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.
6. MIYAKODA, K. and O. TALAGRAND, 1971: The assimilation of data in dynamical analysis. Part I. *Tellus* **23**, 310—317.
7. RINNE, J., 1971: Investigation of the forecasting error of a simple barotropic model with the aid of empirical orthogonal functions. Part I. *Geophysica,* **11**, 185—213.
8. —»— 1971: Investigation of the forecasting error of a simple barotropic model with the aid of empirical orthogonal functions. Part II. *Ibid.,* 215—239.
9. SÖDERMAN, D., 1972: Numerical weather prediction in Finland during 1970. *Techn. Rep.* 2, *Finnish Meteorological Institute,* Helsinki, 21 pp.
10. YUDIN, M. I. and G. D. KUDASHKIN, 1968: Consideration of group errors of numerical forecasts as a means of improving their accuracy. *Proceedings of the WMO/IUGG Symposium on Numerical Weather Prediction in Tokyo,* II/107—117, Japan Meteorological Agency, Tokyo.