Preliminary Research in Objective Analysis

By D. H. JOHNSON, Central Forecasting Office, Dunstable

(Manuscript received November 13, 1956, revised March 8, 1957)

Abstract

Simple mathematical surfaces are used to interpolate numerically within two-dimensional networks of 500 mb contour heights and winds. The effect on the errors of interpolation of varying (i) the relative weights given to the measured contour heights and winds, (ii) the complexity of the mathematical surface, and (iii) the layout and dimensions of the network is examined. It is concluded that in areas where observations are reasonably dense, numerical analysis will be at least as satisfactory as the currently used subjective graphical method.

Introduction

The potential merits of objective analysis as a preliminary to numerical forecasting are well-recognised and objective wind interpolation may find application in other synoptic work.

Studies of objective analysis have been made in the United States (Panofsky 1949, Gilchrist and Cressman 1954) and in Sweden (Bergthorsson and Döös 1955). The results of the first phase of parallel work which is in progress at Dunstable, are given below.

Method of Approach

As a first step, attention has been confined to the basic problem of interpolating between a two-dimensional network of observations without regard to conditions at other levels or at other times, although it seems probable that fully successful schemes of objective analysis will eventually take into account continuity in time and consistency in the vertical.

One method of interpolating for the height of a contour surface at a given grid point is to represent the contour height over some area surrounding the grid point mathematically as a function of position, the function being chosen so as to give the appropriate values at the observing stations. By applying the geostrophic relation, the observed winds can also be taken into account.

Before the analysis can serve as a basis for a numerical forecast, it may be desirable to smooth the wind or contour fields so as to leave only those scales of disturbance which are to be forecast and which are consistent with the forecast equations. However, the proper degree of smoothing will depend upon the nature of the forecast and the particular mathematical model selected. So that the objective analysis should have general application, the principal aim in the work presently described has been to get accurate values for interpolated winds and heights. Smoothing to eliminate unwanted, though real, wavelengths has been taken to fall within the province of the forecast rather than that of the analysis.

Attempts were made to represent the contour height, H, as quadratic and cubic functions of suitable co-ordinates on the earth's surface. The co-ordinates chosen correspond to rectangular cartesian co-ordinates on a stereographic projection of the earth's surface from the south pole onto the tangent plane through the north pole. For the quadratic, H is expressible in the form

$$H = ax^2 + 2hxy + by^2 + 2gx + 2fy + c,$$

where a, h, b, g, f and c are constants determined from the observations. Because the observations contain errors, it is advisable not to fit them exactly but to introduce a degree of smoothing using the principle of least squares. A least-squares solution is obtained by choosing a, h, etc., so as to minimise the expression

$$\sum_i (H - H_i)^2 + T^2 \sum_i (\mathbf{V} - \mathbf{V}_i)^2, \qquad (1)$$

where the H_i are observed contour heights, the $\mathbf{V}_i$ are observed winds, $\mathbf{V}$ is the geostrophic wind, T decides the relative weight to be given to contour height and wind observations, and the summation is taken over i observing stations. Components of $\mathbf{V}$, V_x and V_y, are given by $-\frac{\beta g}{l}\frac{\partial H}{\partial y}$ and $\frac{\beta g}{l}\frac{\partial H}{\partial x}$ respectively, where g, l have their usual significance and β is a magnification factor arising from the choice of projection. Equation (1) is minimised when a, h, etc., satisfy the six simultaneous linear equations obtained by equating to zero the expressions given on partial differentiation of expression (1) by a, h, etc., in turn.

Choice of weight

The weighting factor T, which occurs in expression (1), holds a balance between contour heights and winds and depends upon, amongst other things, the units of H and $\mathbf{V}$. The terms $\sum_i (H - H_i)^2$ and $T^2 \sum_i (\mathbf{V} - \mathbf{V}_i)^2$ are thereby permitted to attain the same order of magnitude; if these terms differed greatly, the least squares process would ignore the smaller of the two. In the computations described below it was convenient to work with $\mathbf{V}$ in ft/sec and H in ft. For the two terms of (1) to have equivalent dimensions, the units of T have been taken as seconds. The value of $\sum_i (H - H_i)^2$ depends on the random errors of observation of contour height and also upon a kind of topographical error which arises because quadratic surfaces comprise a restricted class of shapes; the true contour surface will not in general coincide with one of these shapes, although there may be a sufficient resemblance over a limited area. $\sum_i (\mathbf{V} - \mathbf{V}_i)^2$ contains ageostrophic effects as well as errors of observation and errors of a topographical character. The topographical errors will vary according to the function used to represent the contour surface, so T itself may require adjustment according to the function chosen.

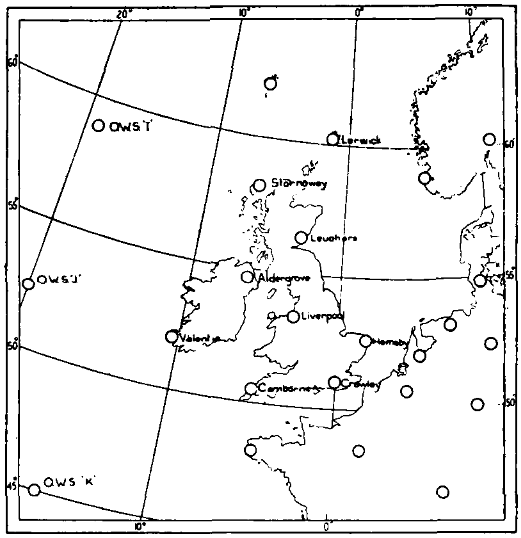

Fig. 1. Key to stations used in the computations.

Computation No. 1

To determine an optimum weighting factor a quadratic surface was fitted to 500 mb heights and winds from six British stations, Aldergrove, Camborne, Crawley, Hemsby, Leuchars and Valentia, for 100 occasions selected from the first nine months of 1954. Computations were made in each case for the ten integral values of T from 0 to 9. Cases were chosen when there were reports of contour height and wind from all six stations, but otherwise dates were picked at random. To test the value of the surfaces for interpolation, the implied Liverpool heights and geostrophic winds were compared with Liverpool upper air observations. Figure 1 shows the disposition of the stations used for these and subsequent calculations. The root-mean-square differences between observed and computed contour heights and winds at Liverpool for the 100

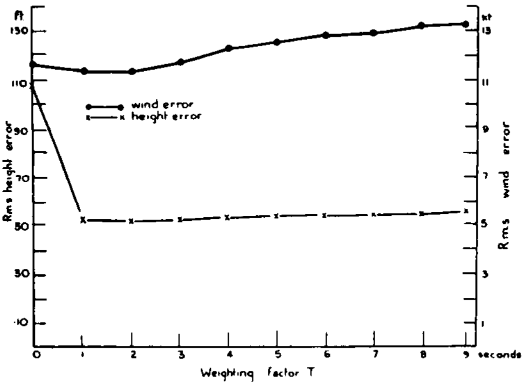

Fig. 2. Computation No. 1. Errors of representation at Liverpool.

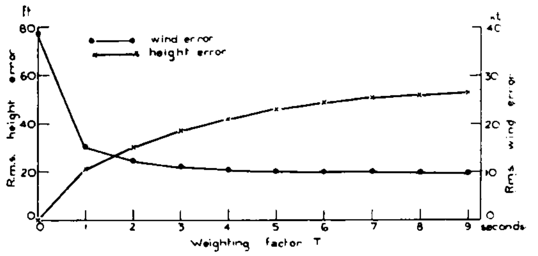

Fig. 3. Computation No. 1. Errors of representation at the six stations.

cases are given by Figure 2. Values for $T = 0$ were special since, if the winds are ignored, the six contour heights can be fitted exactly by a quadratic and there is no smoothing. The lowest interpolation errors occurred in the range $T = 1$ to $T = 2$, where the r.m.s. wind difference was 11.4 kt and the height difference was 52 ft.

The result of this experiment is comparable with that obtained by GILCHRIST and CRESSMANN (1954), who computed an objective analysis for a single occasion over a grid of 19×19 points covering most of North America and a part of the Atlantic. Their method was very similar to that described above: a quadratic surface was fitted to observations within a limited distance of a grid point to get an interpolated contour height. The interpolated heights were compared, not with observations, but with the mean of three subjective analyses. The minimum height "errors" occurred with a weighting factor corresponding to $T = 2$.

Figure 3 shows the variation with T of the r.m.s. height and wind differences when summed over the six stations to which surfaces were fitted and over the 100 cases. In the range $T = 1$ to $T = 2$ the r.m.s. height difference rose from 21 to 30 ft. whilst the r.m.s. wind difference fell from 15 to 12 kt (25 to 20 ft/sec). So the two terms of expression (1) did in practice attain the same order of magnitude: — at $T = 1$, $\sum_i (H - H_i)^2 = 400$ ft^2 while $T^2 \sum_i (\mathbf{V} - \mathbf{V}_i)^2 = 625$ ft^2 T^2 having units of sec^2; at $T = 2$, $\sum_i (H - H_i)^2 = 900$ ft^2 while $T^2 \sum_i (\mathbf{V} - \mathbf{V}_i)^2 = 1{,}600$ ft^2.

Computation No. 2

It was remarked above that the results obtained for $T = 0$ in Computation No. 1 were special since the contour heights at the six stations were exactly fitted. The computations were repeated for 40 of the 100 cases using observations from Stornoway and Lerwick in addition to the six previously used. In this case, even at $T = 0$, there were more pieces of information than were required to determine the co-efficients of the quadratic uniquely.

In Figure 4 the r.m.s. height and wind differences at Liverpool for the eight station computations are comperad with results for

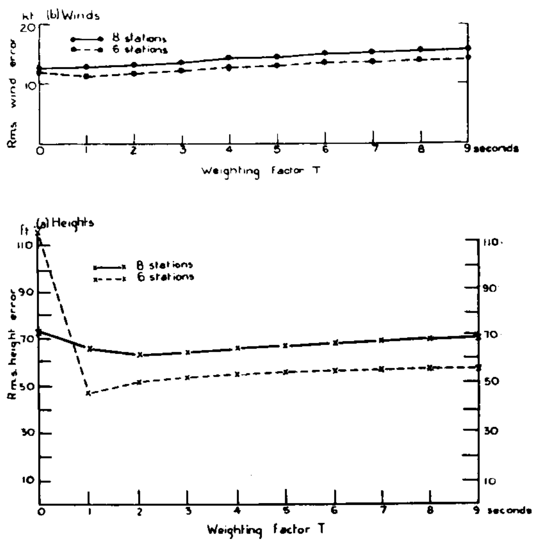

Fig. 4. Computation No. 2. Errors of representation at Liverpool.

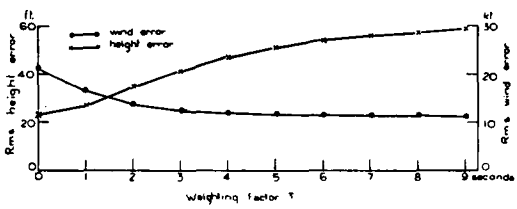

Fig. 5. Computation No. 2. Errors of representation at the eight stations.

the same 40 cases from Computation No. 1. Except for the heights at $T = 0$, the Computation No. 2 differences exceed those of Computation No. 1. The inclusion of the two extra heights reduced the r.m.s. Liverpool height difference considerably. The improvement here was due entirely to the beneficial effect of smoothing out errors of observation, etc., since results for $T = 1$ to $T = 9$ suggest that it was generally undesirable to extend the area over which the quadratic was fitted beyond that covered in Computation No. 1. As in Computation No. 1, values of T in the range from $T = 1$ to $T = 2$ appeared to be satisfactory from the point of view of minimising the errors of representation. Neither the wind nor the height errors were very sensitive to variations in T.

Figure 5 gives the standard errors of fitting at the eight stations. These may be compared with the corresponding values for Computation No. 1 in Figure 3. Computation No. 1 values were smaller except for the winds at $T = 0$.

Computation No. 3

To examine whether a surface of greater complexity than the quadratic was of more value for interpolation, a cubic surface was fitted to the six stations used for Computation No. 1, in 38 cases. Computations were made for six values of T. $T = 1, 2^{\frac{1}{2}}, 2, 2^{\frac{8}{2}}, 4$ and 6.

The r.m.s. differences between observed and interpolated heights and winds at Liverpool are given in Figure 6. Neither the height nor the wind differences showed much variation over the range of T explored, but such variations as there were followed a different pattern from the quadratic results of Computation No. 1. The lowest Liverpool wind difference occurred at about $T = 4$ whereas the height

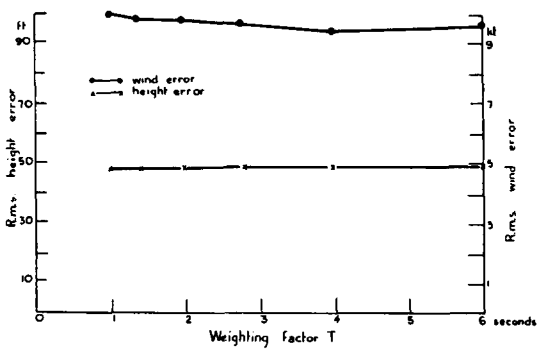

Fig. 6. Computation No. 3. Errors of representation at Liverpool.

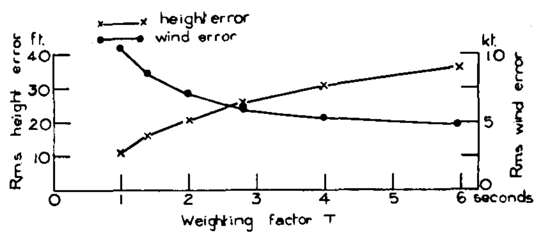

Fig. 7. Computation No. 3. Errors of representation at the six stations.

errors increased with T throughout the range examined. The cubic results are best compared with the quadratic six-station values given in Figure 4 which, with two exceptions, refer to the same occasions. The two exceptions do not materially affect the comparison. The cubic surfaces gave a slightly better fit than the quadratic at Liverpool.

Figure 7 gives the r.m.s. errors of representation at the six stations whose observation were used to obtain the cubics. The values are naturally lower than those obtained for the quadratic (figure 3) since the cubics have a greater degree of freedom.

The cubic results show how the choice of weighting factor T may depend upon the purpose for which the interpolation is required. A preliminary to the preparation of a numerical forecast is the evaluation of the Laplacian derivative of the initial contour field. Normally, the Laplacian is derived from the set of interpolated contour heights by means of the finite difference approximation. If the contour field is represented by some mathematical function, it is possible to obtain the Laplacian in a second way, namely by direct differentiation of the function itself. For this latter purpose it is desirable that the function should not just give an accurate interpolation for height at the grid point, but should also approximate closely to the true contour surface over a certain area surrounding the grid point. In Computation No. 3 the contour height interpolation error was lowest at $T = 1$ (Figure 6), but Figure 7 shows that at $T = 1$ the heights at the six stations were fitted with an error of only 10 ft. This closeness of fit was unrealistic since the standard error of measurement of 500 mb contour height at British stations is 30 ft. Futhermore the winds at the six stations were fitted at $T = 1$ with an error of about 10 kt, which is larger than the standard error, 3 kt, of British 500 mb wind observations. At $T = 4$, on the other hand, the r.m.s. height error at the six stations was 30 ft and the r.m.s. wind error was 5 kt. Thus, if a realistic fit is required over an area of 300 mi or so centred upon Liverpool, a weighting factor $T = 4$ is to be preferred to $T = 1$.

According to MURRAY (1954) the r.m.s. difference between observed winds and geostrophic winds measured from the C.F.O. working charts is 14 kt at 500 mb over the British Isles. It is interesting to note that the wind values interpolated from the cubic surfaces when compared with independent observed winds gave a r.m.s. difference of less than 10 kt. Murray shows that the 14 kt can be broken down into the following components: 9 kt due to the ageostrophic wind; 4 kt due to turbulent wind fluctuations; 3 kt due to errors in the observed wind; 9 kt due to errors inherent in the graphical technique. The error of representation of the wind field given by the true contour pattern will contain the first three of these four components and should therefore have a root-mean-square value of about 10 kt. On this basis it appears that the 5 kt error of representation of the six stations, given by the cubics for a weighting factor $T = 4$, is lower than is justified. However, in many techniques of numerical forecasting the contour height field is used essentially as a stream function and a closer fitting of the winds than seems proper on purely geostrophic grounds may give a better approximation to such a stream function than is otherwise obtained.

PANOFSKY (1949) fitted a few cubic polynomials to the wind and contour fields over an area of the order of 10^6 square mi, which is about four times greater than the area used for Computation No. 1. Winds were weighted by a factor corresponding to $T = 2$. Visual comparison of the objective analyses with corresponding subjective analyses showed that the two methods gave similar contour patterns, but the objective analyses were somewhat smoother. Panofsky also computed the Laplacian derivatives of the subjective and objective contour fields. Here the agreement was "not good". The last result may have been anticipated, since the Laplacian of a cubic is a linear function and experience of vorticity fields computed from subjective analyses suggests that the true vorticity fields are very far from linear; in fact they are usually more complex than the contour patterns themselves. The implication of comparative simplicity of the vorticity field may be a considerable objection to the use of polynomials in representing contour patterns over large areas.

Various networks

Computation No. 4

The six-station network, centred upon Liverpool, was chosen somewhat arbitrarily as suitable for use in the above interpolation experiments. To examine the effect on the errors of representation of varying the distribution and quantity of data, quadratic surfaces were fitted to sixteen selections of data from the eight stations used in Computation No. 2. The computations were made with a weighting factor $T = 2^{\frac{1}{2}}$ for each of 40 occasions. The r.m.s. errors of representation at Liverpool and at stations used in the computations are given for each selection in Table 1, which also contains, for comparison, values previously obtained for corresponding occasions in the six-station and eight-station calculations (Computations Nos. 1 and 2).

The first six rows of Table 1 (series B, C, D, E, F and six-station) give information regarding the relative values of contour heights and winds. Comparing the Liverpool interpolation errors, series D, E and F, which contained more heights than winds, improved markedly upon series B and C in which winds preponderated.

Series M, K and I contained wind and contour height observations from five, four and three stations respectively, a compromise being made in the choice of stations between proximity to Liverpool and regularity of distribution. Each of these arrangements have errors of interpolation similar to those of the six-station data, except that the three-station arrangement (series I) gave a slightly better fit to the Liverpool wind: Liverpool is not far from the centroid of the triangle formed by three stations, Aldergrove, Hemsby and Camborne. When the size of the triangle was increased (series J and H), the interpolation errors were also increased.

Most of the above selections were chosen so as to have a reasonably symmetric distribution with respect to Liverpool, although comparison between series D and E shows how the disposition of the two winds materially affected the contour height errors. In series N, O and P the data were not symmetrically arranged about Liverpool, in fact for series O and P the problem was one of extrapolation rather than interpolation. In these last three cases the errors of representation were comparatively large.

Oceanic networks

The foregoing experiments gave promising results, but the British network of upper air observations is one of the most satisfactory for accuracy and density of observations. The worst networks are found over the oceans. A difficulty which arises in experimenting with a weather ship network is that the reservation of one observation for verification purposes alone seriously depletes the network. The alternative to checking the objective interpolations against observations is to compare them with values read from a conventionally analysed chart. This latter practice may be unsound however, particularly for experiments in which T is varied, since the conventional analysis is subjective and itself depends upon the analyst's opinion regarding the relative weights to be given to wind and height observations. To avoid these difficulties the five stations, weather ships "I", "J" and "K", Crawley and Stornoway were chosen as being equivalent to a mid-ocean network. Valentia observations were used to test the accuracy of the interpolated values.

Computation No. 5

Quadratic surfaces were fitted to observations from the five stations for 100 occasions chosen from the year 1954. Computation were made for T-values of 1, 2 and 4. The errors of representation appear in Table 2. The interpolation errors considerably exceeded those obtained in previous computations.

Computation No. 6

Computation No. 5 was repeated, replacing the quadratic surfaces by cubics. Table 2 contains the results. The Valentia errors are slightly smaller for the cubics than for the quadratics.

Assessment

The experiments described above show that in a close network of observations, a reasonably accurate objective interpolation for contour height and wind can be made. Comparisons with the errors of representation of the conventional contour analysis suggest that in regions of good data coverage, objective analyses will be at least as satisfactory as conventional analyses for numerical prediction. The ocean network computations suggest that only in limited regions will an analysis which takes no account of continuity in time and in the vertical, give very good results. There is no difficulty in principle in introducing an element of time continuity into the analysis for, as GILCHRIST and CRESSMANN (1949) point out, forecast values of height and wind can be taken into account in the least squares process on the same basis as observations. The weight given to a forecast value would naturally be much less than the weight given to an observation.[1]

The object of this series of computations was to derive interpolated heights and winds for a single grid point. Before extending the technique to get interpolations over a network of grid points one is led to consider whether it is better to represent the contour field over the whole network by a single function or whether a fresh function should be derived for each grid point. GILCHRIST and CRESSMANN

[1] A number of analyses have now been made by F. H. BUSHBY, for the North Atlantic using this technique. These will be published shortly.

(1949) remark that the latter alternative agrees with synoptic practice in considering data only from the immediate vicinity of the grid point in question. For the former alternative difficulties may arise due to the dependence of the expression for the contour height at a given grid point on data from distant regions. If a single function is used for representation over a wide area, the degree of complexity of the implied geostrophic vorticity field may be over-restricted, although this objection would be less valid for harmonic functions for example, than for polynomials. On balance it seems that the method which treats each grid point independently is likely to prove the more flexible approach.

Conclusions

1) In a close network of reliable observations, objective analyses are capable of representing the 500 mb winds with a standard error of about 10 kt. This figure compares favourably with the errors of representation, about 14 kt., of routine analyses for the same area.

2) In more open networks, such as occur over the Atlantic Ocean, provided that both winds and contour heights are available, winds can be interpolated with a standard error of about 16 or 17 kt. Errors of the conventional analysis are not known for such a region, but they are unlikely to be less than 16 or 17 kt, judging from the data available for the land area.

3) The optimum weight to be given to wind observations as compared with contour heights varies according to the nature of the interpolation. When interpolating for wind or contour height using quadratic surfaces the weighting factor is best chosen so that a discrepancy of 30 ft in contour height receives the same weight as a discrepancy of 10 to 15 kt in wind. Using cubic surfaces, a similar weight gives good results for contour height interpolations but the optimum weight for wind interpolation is obtained by equating a discrepancy of 30 ft in contour height to a discrepancy of 5 kt in wind.

4) The errors of interpolation are not very sensitive either to small variations in the relative weights given to contour heights and winds or to small variations in the number and disposition of the observations.

Acknowledgements

This paper is published by permission of the Director of the Meteorological Office. The least-squares solutions were calculated electronically using the Ferranti Mk. 1 machine at Manchester University. The computations were coded by F. H. Bushby and Vera M. Huckle. Benefit of discussion with J. S. Sawyer, G. A. Corby and F. H. Bushby is gratefully acknowledged.

REFERENCES

Bergthorsson, P., and Döös, B., 1955: Numerical weather-map analysis. *Tellus*, *7*, p. 329.

Gilchrist, B., and Cressman, G. P., 1954. An experiment in objective analysis. *Tellus*, *6*, p. 309.

Murray, R., 1954: On the accuracy of contour charts in forecasting upper winds. *Met. Office, London, Professional Notes*, 7, No. 110.

Panofsky, H. A., 1949: Objective weather-map analysis. *J. of Meteorology*, *6*, p. 386.