551.55.9 : 551·503·99

A scheme for multi-level objective analysis of contour heights

B. V. SINGH and R. K. DATTA

Meteorological Office, New Delhi

ABSTRACT. The results of experiments performed with a scheme for objective analysis of the contour heights of five levels of the atmosphere are presented here. Whereas objective analysis so far produced in India related to only 500-mb level, in this study a scheme has been prepared for the analysis of 850, 700, 500, 300 and 200-mb levels simultaneously. The method known as 'successive corrections by weighted averages' has been employed. To maintain vertical consistency of analysis, thickness between two layers and the vector wind shear at the reporting stations have been taken into account. Use of satellite pictures and AIREP messages has been discussed.

1. Introduction

Objective analysis of meteorological elements or any other variable by machine, differs from conventional manual analysis in the sense that in the former, personal bias of the analyst is dispensed with. This is achieved by generating, by means of mathematical formulae, values of a variable at regularly spaced grid points from observations at irregularly spaced reporting stations.

Various methods of objective analysis in use at present by various meteorological services can be broadly classified into three categories, *viz.*, (*i*) fitting of a geometrical surface by polynomials, (*ii*) successive corrections by weighted averages, and (*iii*) optimal interpolation scheme.

In India, objective analysis of 500 mb (contour) chart was first prepared in 1969 at Northern Hemisphere Analysis Centre, New Delhi. This was done on IBM 1620 mark II computer. The results of that study have been discussed elsewhere (Datta *et al.* 1970). In that study we adopted 'successive corrections by weighted averages scheme', first proposed by Bergthorssen and Döös (1955) and later adopted by Cressman (1959). The method uses a first guess field over all the grid points and this is subsequently improved by using the current data of observing stations. The scheme presented in our earlier paper was put to real time test for a period of about two years. The results have been found to be encouraging.

At RMC, New Delhi, we are required to perform multi-level objective analysis. For any such analysis scheme to be successful, it is necessary that it should have built-in space and time continuity. Vertical consistency of reported observations is achieved by hydrostatic check and inter consistency between various stations at the same level is checked by horizontal check (Datta *et al.* 1970). Experience gained by manual analysis reveals that differential analysis scheme gives vertical consistency of analysis upto about 200 mb. Because of inaccuracies in the observed thickness above 200 mb, differential analysis scheme may not be very suitable. In the present study we have adopted a scheme very much similar to manual differential analysis scheme. The base layer for the analysis is taken as manually analysed sea level chart of the day, since the data coverage at this level is better than at upper levels. Important synoptic features over the sparse data regions in the upper air charts can thereby be properly built up. The analysis scheme is programmed to give the contour analysis for 850, 700, 500, 300 and 200-mb levels. Besides, provision has been kept to utilize so called 'bogus data'. The details of the scheme and the results of a few cases are discussed in the present paper.

2. Basic analysis procedure

2.1. The analysis procedure adopted in the study may be summarized as follows —

Suppose 0_k ($k=1, N$) is a set of N observations. Also for each grid point (i, j), preliminary guess for the variable being analysed (Z for contour height) is available, say Z^{s-1}, where $(s-1)$ indicates guess for the sth scan.

Let $\triangle Z_k^{s-1}$ ($k=1, N$) denote the difference between the observed value (Z_o) and its current estimate (Z_P) at the observation station (P). The current estimate at the point of observation is determined by using bilinear interpolation formula as discussed in the next section. The correction C^s, for the sth scan is evaluated by determining the contribution of each station which is within a distance R,
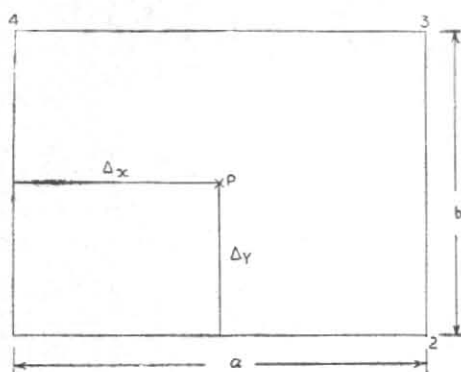
Fig. 1

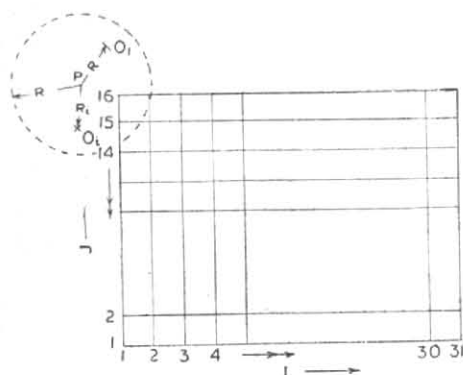Grid references for definition of interpolation formula
(Eq. 2·1)

Fig. 2

Grid references for current estimate for station outside the
grid area

the radius of influence. The correction to be made at the grid point (i,j) on scan s is :

$$C_{(i,j)}^{s} = \sum_{k=1}^{N} W_k \triangle Z_k^{s-1} \Big/ \sum_{k=1}^{N} W_k$$

where W is a distance dependent weighting function defined by $W = (R^2 - D^2)/(R^2 + D^2)$ for $D < R$

$$= 0 \quad \text{for } D \not> R$$

and D is the distance between the grid point (i,j) and the reporting station. Here the distance R is a characteristic distance beyond which corrections based on that observation are not taken into account. The new estimate at the grid point (i,j) is then calculated as follows —

$$Z_{(i,j)}^{s} = Z_{(i,j)}^{s-1} + C_{(i,j)}^{s}$$

2.2. *Current estimate at the point of observation*

The current estimate (Z_P) at the point of observation (P) must be determined before next scan of stations can be made to determine their contributions at the grid points. For computation of Z_P, the following cases have been considered.

2.2.1. *Station within the grid* — In Fig. 1, the reporting station is located at the point P and the four surrounding grid points are numbered 1 through 4. The estimate at the reporting station within the grid is determined from values at four surrounding points according to the formula given below—

$$Z_P = Z_1 + (Z_4 - Z_1)\frac{\triangle y}{b} + (Z_2 - Z_1)\frac{\triangle x}{a}$$

$$- (Z_2 - Z_3 + Z_4 - Z_1)\frac{\triangle x}{a} \cdot \frac{\triangle y}{b} \quad (2 \cdot 1)$$

where the subscripts on the right hand side of the equation refer to the grid points, a is the distance between the grid points along the x-axis and b is the corresponding distance along the y-axis.

2.2.2. *For station outside the grid* — To improve the analysis of boundary regions, reporting stations which are outside the boundary of the grid area have been included. The current estimate at the station P (Fig. 2) which is outside the grid is taken as distance weighted average of observations $0_1, 0_2, \ldots 0_i$, and the grid points $(1, 16), (1, 15)$ and $(2, 16)$ say which come within scan radius R of the reporting station P.

That is guess Z_P at P is :

$$Z_P = \sum_i W_i \, Z_i \Big/ \sum_i W_i$$

where $W_i = (R^2 - R_i^2)/(R^2 + R_i^2)$ and R_i is the distance of the ith observation/grid point from the reporting station P.

3. Height Analysis

In the case of height analysis, the calculated corrections are a function of the observed winds as well as the observed heights. Three different possibilities are considered; for each of these situations, the difference between the observed value and its estimate at the observation station is calculated as shown ahead.

(i) *For a station with height only*

$$(\triangle Z)_H = (Z_o - Z_P) \qquad (3.1)$$

where Z_o is the reported height and Z_P is preliminary guess value at the reporting station.

(ii) *For a station with both height and wind*

$$(\triangle Z)_T = Z_o + \frac{k.f}{m.g}(v\triangle x - u\triangle y) - Z_{(i,j)} \qquad (3.2)$$

where u and v are eastward and northward components of the reported wind, f is the coriolis parameter, and $\triangle x$, $\triangle y$ represent the distance between the reporting station and the grid point along the x and y-axes respectively and k is a constant to express the ratio of the geostrophic to the actual wind. In this study we have taken $k=0.80$. m is map scale factor.

(iii) *For a station with wind only*

$$(\triangle Z)_W = Z_P + \frac{k.f}{m.g}(v\triangle x - u\triangle y) - Z_{(i,j)} \qquad (3.3)$$

When all data within a radius R have been scanned, each reporting station yields a contribution to either $(\triangle Z)_H$, $(\triangle Z)_T$ or $(\triangle Z)_W$. The final correction $\triangle z$ that is applied at each grid point (i, j) to the first guess is given by :

$$\triangle z = \left[A \sum_K W_K (\triangle Z_H)_K + \sum_K W_K (\triangle Z_T)_K + \right.$$

$$\left. + \sum_K W_K (\triangle Z_W)_K \right] / [A N_H + N_W + N_T] \qquad (3.4)$$

where N_H, N_T and N_W are number of $(\triangle Z)_H$, $(\triangle Z)_T$ and $(\triangle Z)_W$ corrections. A is a constant factor to give weightage to the lateral gradient of the preliminary guess. In this study A is put equal to 0.3.

4. Programming Technique

There are several approaches to the programming of this problem for an electronic computer. For example, the calculation of the corrections to be applied to the previous analysis at a grid point can be accomplished by scanning the grid points and considering all data within the region of influence of each grid point. Alternatively, one can scan the data array and determine the contribution of each piece of data to the corrections to be made at surrounding grid points. We have chosen the latter procedure. From the experience gained for analysis of single level, we have selected first scan of radius 5° Lat./Long. The second scan which is of 10° radius modifies only those grid points that lie beyond 5° but within 10° Lat./Long. These two scans taken together is termed as one iteration. It has been revealed from earlier study (Datta *et al.* 1970) that scan of 5 degrees radius brings out

small disturbances, such as minor troughs rather well while a larger scan of 10 or 15 degrees smoothens out small scale features and only retains the large scale features. After each iteration, number of stations whose reported value differed from its interpolated value by more than a pre-assigned limit, were marked out. And the iteration was repeated till the number of such stations was less than a pre-assigned number. Number of trials with single level showed that with three to four iterations, this convergence was achieved. Making use of this in multi-level analysis, we kept a provision for four iterations in all and this gave a fairly good objective analysis. Also the root mean square error (RMSE) between the reported observations and the interpolated values at the station from objective analysis was fairly low.

5. Maintenance of vertical consistency

From sea level analysis of the day, pressure values at the grid points are manually picked up. These values are converted to height of 1000 mb surface by using a mean value of conversion derived from Smithsonian Meteorological Tables. Latest available synoptic hour analysis (12 GMT in this case) is used as preliminary guess field for all the levels from 850 to 200 mb. Objective analysis of 850-mb level is done as analysis of a single layer discussed earlier. However vertical consistency with 1000 mb analysis is ensured by inclusion of 'bogus data'. Differential analysis scheme is adopted for 700 mb and above. The various steps followed for this are described below —

(1) Compute (1000-700) mb thickness field based on 1000 mb field and 700 mb guess field at the grid points. Call this thickness guess field for (1000-700) mb.

(2) Compute wind shear for (850-700) mb and thickness (1000-700) mb at the reporting stations based on current observations.

(3) Improve the thickness guess field at (1) by using wind shear and thickness for the reporting station (2) by the method of successive approximations. Call this actual thickness field for (1000-700) mb.

(4) Add 1000 mb field to (3). This gives first guess for 700 mb.

(5) Improve the guess at (4) using actual 700 mb observations by the method of successive approximations using iterations scheme. This gives us the final 700 mb analysis.

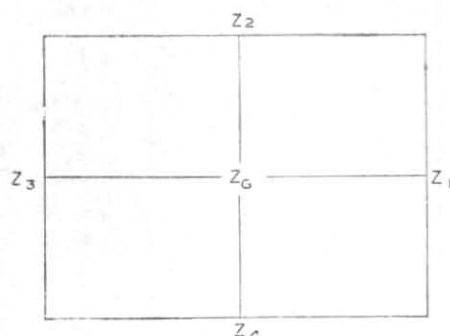This process is repeated for building up the 500 mb analysis by taking 700 mb analysis as base

Fig. 3

Grid references for detemining smoothened value at Z_G

layer and so on upto the final analysis of 200 mb. It may be noted that we did not adopt differential analysis scheme for building up 850 mb analysis for lack of height and wind observations for 1000 mb. Moreover for computation of wind shear between 1000 and 700 mb, we actually used winds of 850 and 700 mb.

6. Inclusion of bogus data

It is necessary to supplement observations in regions of data-hole, by use of so called 'bogus data'. By bogus data we mean all those asynoptic data obtained from APT pictures, AIREPS etc. Various studies on the satellite input for numerical analysis and prognosis (SINAP) have been conducted so as to exploit the use of satellite observed cloud pictures to improve the analysis over the oceanic area and areas of meagre data. Recently, one of the authors (Datta and Nagle 1971) has proposed the use of APT cloud pictures for the 500 mb analysis. In this study, however, we have tried to visually locate the centres of depressions/ cyclonic storms, location of western disturbances in the latest available APT pictures. Contour height for the centre is estimated from previous analysis and this estimated height and wind is used as 'bogus data' to improve objective analysis.

7. Scale of analysis and smoothing

It is well known that the rectangular grid cannot resolve wavelengths whose dimensions are less than two grid units. It is, therefore, necessary to smooth out amplitudes arising from disturbances whose wavelengths are in this size range. The averaging formula 3·4 applied to the correction field, acts as a smoother but in practice it is found that short wavelengths may still be present in the final analysis. Consequently, an explicit filtering procedure has been employed to set a lower limit to the wavelengths present in the final analysis.

Schuman (1957) has discussed the relative power transfer functions of the 5 and 9-point two dimensional filters. He has shown that the 9-point filter is about twice as efficient as the five-point filter in eliminating small scale disturbances. We tried Schuman five and nine-point and five-point smoothers defined by 7·1 and used by Canadian Met. Services. A large number of cases of single level analysis were tried with these three smoothers. It was found that Schuman 9-point smoother is rather strong and sweeps out some important features as well. We found that five-point mild smoother (7·1), was better compared to 9-point smoother. for use after each alternate iteration for the reason that we are building up upper level contour charts from lower level charts which had already been smoothened.

$$Z_G' = \tfrac{1}{8}[Z_1 + Z_2 + Z_3 + Z_4 + 4Z_G] \qquad (7\cdot1)$$

where the subscripts refer to the grid points as shown in Fig. 3.

8. Discussion of results

A number of trial runs were made to study the working and test the accuracy of the analysis scheme presented in this paper. The analysis was confined to an area between 25·0 to 136·1°E and 0 to 48·4°N. The area was divided into (31×16) grid poin's with a grid length of 381 km. The chart used was Mercator projection, scale 1 : 20,000,000 true at Lat. 22°30′. Machine analysis was compared with conventional analysis for all the cases. It was found that the two analyses compared well. For the economy of space, we present a case of analysis for a situation on 00 GMT of 10 April 1972. The conventional analysis for the day is also presented for comparison.

Figs. 4(A) to 4(E) show conventional analysis for 00 GMT for 850, 700, 560, 300 and 200-mb levels and Figs. 5(A) to 5(E) show the corresponding machine analysis. For presentation of vertical consistency and to present a three-dimensional view, the charts have been shown in one block, with meridians tilting towards west and not vertical as in a Mercator projection. However, the location and orientation of the systems have been preserved as in the original projection.

8.1. *850 mb chart*: *Figs. 4(A) and 5(A)* — The chief synoptic features over the Indian region were a cyclonic storm in the Bay of Bengal, an extended low pressure system centred between Delhi and Srinagar and a low pressure area over Trivandrum and adjoining regions.
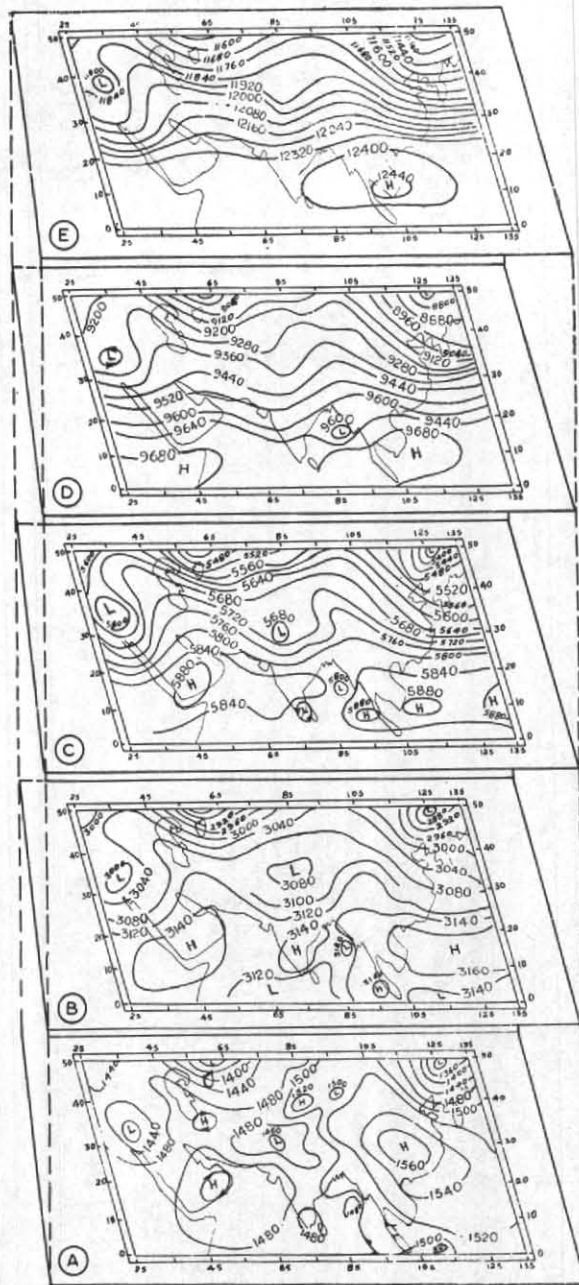
Since, no data was available from the cyclonic

Fig. 4

Conventional analysis for (A) 850, (B) 700, (C) 500, (D) 300 and (E) 200 mb for 00 GMT of 10 April 1972

storm area, bogus data on the basis of previous 12 GMT analysis and latest APT picture was utilised. It can be seen that in the machine analysis, the cyclone has been produced remarkably well as regards position but central value is different due to the approximate nature of the 'bogus data' used. A slightly lower value of 'bogus data' would have produced two closed contours around the storm as in the case of conventional analysis.

The low pressure area between Delhi and Srinagar drawn in the conventional analysis has not been brought out due to its size which is of a single grid dis'ance and as such gets smoothed out by the application of smoother. The other 'low' over Tibet area brought out in the machine analysis is a left-over from the 12 GMT guess field which cannot be modified due to non-availability of observations over the area.
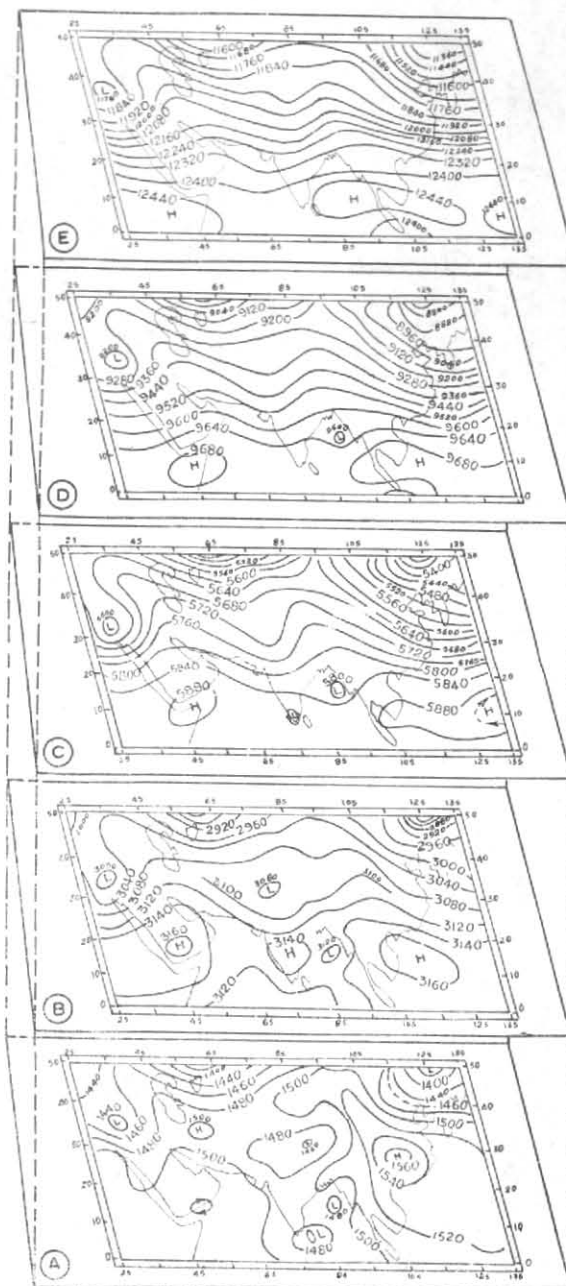
Fig. 5

Machine analysis for (A) 850, (B) 700, (C) 500, (D) 300 and (E) 200 mb for 00 GMT of 10 April 1972

The 'low' near Trivandrum is fairly well brought out in the machine analysis, to an accuracy of about 1° Long.

Over the extra-Indian areas, the features like low pressure areas east of Mediterranean Sea, northeast of Caspian Sea and over Japan and anticyclones over Arabia, China etc brought out

by machine analysis are in agreement with their positions in the conventional analysis.

8.2. *700 mb chart: Figs. 4(B) and 5(B)* — Important synoptic features over Indian and extra-Indian regions are in good agreement in the two analyses. However small anticyclone of value 3160 gpm near Singapore in the conventional analysis has not

appeared in the machine analysis. This is probably because conventional analyst has given full weightage to a single observation of 3163 gpm surrounded by three observations of values 3152 gpm, and drawn a closed anticyclone. However in the machine analysis, surrounding observations of lower values have weakened the 'high', resulting in its disappearence.

It may, however, be mentioned that ordinarily analysis is not done over the Himalayas and Tibetan Plateau for 850 and 700 mb charts for reasons of orography. We have, however, shown the analyses in these two charts for the reason that we did not program our scheme to skip the analysis over this area.

8.3. *500 mb chart : Figs. 4(C) and 5(C)* — The only apparent difference between the two analysis is the delineation of the 5840 contour. In conventional analysis the low in the Bay of Bengal has been shown as a part of easterly flow, whereas the machine analysis has shown it embedded in westerlies. Separate low over Delhi in the conventional analysis has appeared as a deep trough in the machine analysis. All other features are in good agreement in the two analyses.

8.4. *300 mb chart : Figs. 4(D) and 5(D)* — Significant features in the conventional analysis are the cyclonic storm, the low over east of Mediterranean with ridge east of it and a trough in the westerlies along 77°E which have been brought out remarkably well in the machine analysis.

8.5. *200 mb chart : Figs. 4(E) and 5(E)* — The two analyses are almost identical. A special mention may be made of the steep gradient between 25-37°N along 105-136°E, with maximum concentration of contours at 27°N along 120-130°E. This is incidentally the region of the core of westerly jet stream with a maximum wind speed of 140 kt. Jet stream, however, has not been drawn in the conventional analysis in order to retain the clarity of the diagram.

9. Verification of analysis

For verification of machine analysis root mean square error (RMSE) of the reported observations and the one interpolated at the reporting station from four grid points surrounding the station, for all the stations within the grid area have been worked out.

Let Z_A be the actual and Z_I be the interpolated height at the station. Then,

$$\text{RMSE} = \sum_{i=1}^{n} (\Sigma E_i)^2 / n$$

TABLE 1

Root mean square error (metres) between the observed heights and those interpolated on the objective analysis maps for various level

Date	Levels (mb)				
	850	700	500	300	200
11 Feb 1972	5·9	8·6	11·8	18·5	18·2
4 Oct 1971	7·9	11·2	17·1	26·9	31·4
9 Aug 1971	14·8	9·1	14·1	27·3	38·1
10 Apr 1972	9·0	7·9	12·3	20·0	26·7
Average	9·4	9·2	13·8	23·2	28·6

where $E_i = (Z_A - Z_I)_i$ for the ith station and n is the total number of height reporting stations within the grid area. RMSE values for all the levels for 4 cases have been listed in Table 1. It will be seen that RMSE for 850, 700, 500, 300 and 200 mb are of the order of 9, 9, 14, 23 and 28 gpm respectively.

Gandin and Lugina (1969) compared the accuracy of 500 mb objective analyses produced in various countries for different regions. For our area (Region 3) objective analysis of 500-mb level performed by different countries have RMSE values varying from 30 gpm (USA) to 65 gpm (Belgium). It will be seen that for the corresponding level, RMSE of 14 gpm in our case is very low.

10. Conclusions

The results of this study suggest that the scheme of differential analysis for multi-level machine analysis is quite statisfactory and the final output maintains vertical consistency at least upto 200-mb level. The scheme can be put to operational tests for further modifications and developments.

REFERENCES

Bergthorssen, P. and Döös, B.	1955	*Tellus*, **7**, 3, pp. 329-340.
Cressman, G. P.	1959	*Month. Weath. Rev.*, **87**, 1, p. 368.
Datta, R. K., Chhabra, B. M. and Singh, B. V.	1970	*Indian J. Met. Geophys.*, **21**, 3, pp. 437-442.
Datta, R. K. and Roland, E. Nagle	1971	*Ibid.*, **22**, 2, pp. 197-202.
Gandin, L. S. and Lugina, K. M.	1969	Comparison of the accuracy of objective analysis produced in various countries, Proc. WMO/IUGG Symp. on NWP in Tokyo, Nov. 26-Dec. 4, 1968.
Schuman, F. G.	1957	*Mon. Weath. Rev.*, **85**, 11, pp. 357-361.