Variational method for objective analysis of scalar variable and its derivative

S G Narkhedkar* and S K Sinha**

Indian Institute of Tropical Meteorology, Dr. Homi Bhabha Road, Pashan, Pune 411 008, India.
** e-mail: narkhed@tropmet.res.in*
*** e-mail: sinha@tropmet.res.in*

In this study real time data have been used to compare the standard and triangle method by performing the objective analysis of mean sea level pressure. In the standard method, derivative fields are obtained from the grid point data using finite difference scheme whereas in the triangle method, a set of non-overlapping triangles are formed from the observations and the scalar and the spatial derivatives are computed directly at the centroid of each of the non-overlapping triangles. These scalars and their derivatives are then mapped to uniform grids by using the standard method. It has been found that objectively analysed scalar field obtained using standard method is superior to the scalar field derived by the triangle method, whereas the derivative fields produced by triangle method are superior to the derivative fields produced using standard method. A variational objective analysis scheme has been developed and an experiment has been carried out with depression case of June (11–15) 2004. It is found that the new scheme (variational) is able to extract the better parts of both triangle and standard methods. The results of this study will be useful in carrying out diagnostic calculations that involve derivative estimates.

1. Introduction

To produce the most accurate analysis is the ambition of meteorologists for nearly five decades. Interpolating scattered station data by hand requires experience and knowledge about the physics that governs the fields. Today, computers generally do this job, but the requirements remain the same. The work presented in this paper is about a variational objective analysis scheme which produces improved analysis of scalars and their derivatives. Objective analysis produces grid point data from observations which are irregularly distributed in space and time. This grid point data can be used as an initial guess for the numerical model. Secondly, the diagnosticians require grid point values to compute divergence, vorticity, frontogenesis, etc.

One of the most popular schemes of objective analysis is the Successive Correction (SC) scheme formulated by Cressman (1959). Weighting functions used in this scheme are empirical. Barnes' (1964, 1973) scheme which is also an SC scheme does not require any initial guess field. Both the schemes are simple and preferable for diagnostic studies. Optimum Interpolation (OI) is another popular scheme introduced in meteorology by Eliassen (1954) and Gandin (1963) in which weighting functions are determined on the basis of characteristic functions of the statistical structure of the given variable. But when the balance is needed between different fields, multivariate optimum interpolation scheme (Gandin 1963; Schlatter 1975; Bergman 1979; Dey and Morone 1985) or variational scheme (Sasaki 1958; Sinha *et al* 1998) may be required. Koch and Saleeby (2001), Ogura and Chen (1977), Koch *et al* (1983), Benjamin and Seaman (1985), Bussieres and Hogg (1989) and Mitra *et al* (2003) used SC scheme to analyse different meteorological parameters. Due to the poor quality of the derivatives produced by the above

Keywords. Standard method; variational method; triangle method; Delauney triangulation; Euler–Lagrange equation.

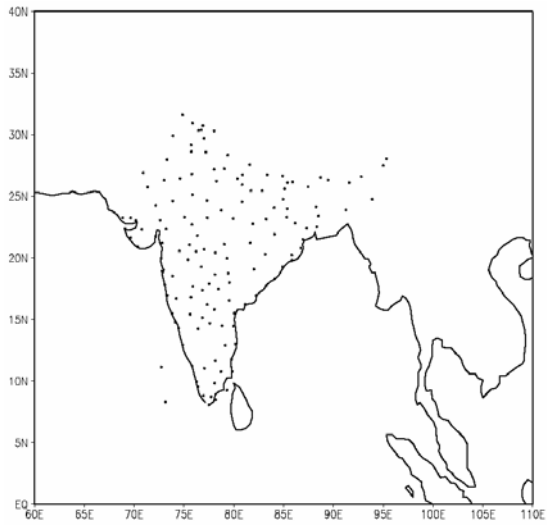


Figure 1. Scatter plot of the observations for a typical day.

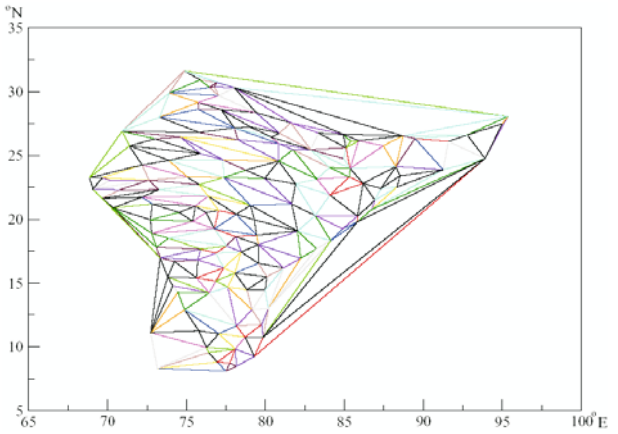


Figure 2. Triangles formed from the observations.

two schemes it is better to compute spatial derivatives directly from a set of observations.

Bellamy (1949) computed divergence and vorticity directly from the observations without using finite difference scheme. Ceselski and Sapp (1975) and Schaefer and Doswell (1979) computed spatial derivatives of wind at triangle centroids using line integrals. Zamora *et al* (1987) and Doswell and Caracena (1988) computed these spatial derivatives using linear vector point function method. Schaefer and Doswell (1979) computed derivatives directly from observations and then applied standard analysis scheme to obtain derivative fields at the grid points. They found that the grid point values of derivatives obtained using triangle centroid method are better than the grid point values of derivatives obtained using standard method. Spencer and Doswell (2001) also found the superiority of triangle method over standard method. All the above-mentioned authors considered only vector field and no scalar field has been analysed.

Endlich and Clark (1963) calculated the spatial derivatives of scalar quantities directly from observations using the triangle method assuming a linear variation of the scalar between reporting stations. In the triangle method values of scalars at the centroid are the mean of the three observations which constitute the triangle. But they have not carried out any objective analysis to interpolate the derivative fields which are at the centroid of the triangle. Barr *et al* (1971) used SC scheme to obtain grid point values by combining both the actual observations and the values at the centroid of the triangle. Recently, Spencer *et al* (2003) developed an improved objective analysis scheme (variational scheme) to analyse scalar field in which scalar field produced by standard method and derivative field produced by triangle method have been used as the input. They used the triangle method of Endlich and Clark (1963) to calculate spatial derivatives directly from analytically generated scalar variables. However, to the authors' knowledge no analysis has been made using the above scheme for real time data. In this paper, we have analysed mean sea level pressure (mslp) field (a scalar) using the technique of Spencer *et al* (2003) for a depression case of June 2004. Spencer *et al* (2003) used Barnes' scheme to produce scalar field whereas in this study scalar field has been derived using OI scheme. Section 2 contains a brief description of triangulation technique, OI technique and triangle methods of objective analysis. In section 3, data used to test the schemes has been discussed and summary of a depression which formed in June 2004 has been given. Performance of standard and triangle methods of objective analysis are discussed in section 4 and variational method of objective analysis is described in section 5. Finally, conclusions are drawn in section 6.

2. Triangulation and objective analysis schemes

2.1 *Delauney triangulation*

Triangulation involves creating from observational network a set of non-overlapping triangles, the vertices of the triangles are the input sample points. Among the number of triangulation algorithms Delauney triangulation is the most popular algorithm. It is closely related to Direchlet tessellation. A plane is divided into a number of polygonal regions by the tessellation. These regions are called tiles. In the interior of each tile there is one sample point. It is called generating point. All other points inside the tile are closer to the generating point than any other points. By connecting all generating points which share a common tile edge

Table 1. *Number of observations and corresponding number of triangles.*

Dates (June 2004)	11	12	13	14	15
Number of observations	142	141	140	133	132
Number of triangles formed	269	267	265	251	249
Number of triangles after deleting triangles having one of the angles $< 15^\circ$	193	188	186	180	184

Delauney triangulation has been created (Bourke 1989). Delauney triangulation has been discussed thoroughly by Ripley (1981).

Figure 1 shows the distribution of stations over the Indian region for a typical day having 142 observations. After the application of triangulation code nearly 269 triangles (figure 2) are obtained. Among the 269 triangles thus obtained, triangles having one of the angles less than 15° are deleted. Table 1 shows the number of observations and the corresponding number of triangles thus formed.

2.2 *Standard method of objective analysis*

Standard method of objective analysis discussed in this section is the univariate OI scheme (Gandin 1963) to which finite differencing scheme is applied to obtain grid point spatial derivatives. The analysis equation used in this scheme is given by

$$P_g^a = P_g^b + \sum_{i=1}^{N} w_i (P_i^o - P_i^b), \tag{1}$$

where P_g^a = analysed field at the grid point 'g', P_g^b = background field at the grid point 'g', P_i^o = observed field at location i, P_i^b = background field at location i, N = number of observations, and w_i = weights.

To derive weights, it is assumed that the observations and the background errors are uncorrelated. The expected analysis error variance derived from the above expression is minimized in relationship to the weights. The normalized expression for the weights is given by

$$\sum_{i=1}^{N} \left(\mu_{ij} + K_{ij}\lambda^2\right) w_i = \mu_{gj}, \quad j = 1, \ldots, N, \tag{2}$$

where μ_{ij} is the background error correlation between locations i and j; μ_{gj} is the background error correlation between locations g and j, K_{ij} is the kronecker delta and λ^2 is the normalized observational error. The expected analysis error variance is normalized and is given by

$$E^2 = 1 - \sum_{i=1}^{N} \mu_{ig} w_i. \tag{3}$$

The normalization is achieved using the background error covariances. For detailed derivation and discussion one can refer Rajamani *et al* (1983). The expression above provides the normalized expected analysis error variance for each analysis point (grid point). The two-dimensional structure of this quantity can be estimated if background error correlation structure is known. Ten years of mslp data for the monsoon months (June to September) are used to estimate the background error correlations. The background error correlations between two locations are given by

$$\mu_{ij}(d) = \frac{S_{ij}}{S_i S_j}, \tag{4}$$

where d is the distance between two locations (i, j), S_{ij} the covariance between mslp at locations i and j, S_i and S_j are the standard deviations of mslp field at i and j respectively.

The correlations involved in equation (2) are computed for every station with respect to every other station and are plotted against distance between two stations. As observed by Alaka and Elvander (1972), the scatter of points is partly due to anisotropy and non-homogeneity of the true correlations. Hence points within 1° segment are averaged to represent the mid-point of the segment. A gaussian function of the form

$$\mu(d) = A \exp(-Bd^2), \tag{5}$$

is fitted to the distance averaged correlations. The values of the regression constants 'A' and 'B' obtained are 0.793 and 0.003 respectively. Thiebaux (1976) has shown that the anisotropy of the correlation function is a significant source of errors to the interpolation scheme. However, the anisotropic nature of the correlation function is neglected in this work for ease of calculation. Since the derivative fields are more important

than the closeness of the analysis to the observations, derivative fields computed using the triangle method should be preferred. To estimate the derivatives inside the analysis domain central difference scheme is used and at the boundary one-way finite differencing scheme is used.

2.3 *Triangle method of objective analysis*

If the vertices of a triangle (the observing stations) have the co-ordinates (x_1, y_1), (x_2, y_2) and (x_3, y_3) then the co-ordinate of the centroid (x_c, y_c) is given by

$$x_c = \left(\frac{1}{3}\right)(x_1 + x_2 + x_3)$$

and

$$y_c = \left(\frac{1}{3}\right)(y_1 + y_2 + y_3). \tag{6}$$

Assuming that a scalar variable P varies linearly within each triangle created by the Delauney triangulation, then following Endlich and Clark (1963) and Spencer *et al* (2003), the value of the scalar at the centroid of the triangle is given by

$$P_c = \left(\frac{1}{3}\right)(P_1 + P_2 + P_3), \tag{7}$$

where P_1, P_2 and P_3 are the values of the scalar at the vertices of the triangle. In the neighbourhood of the triangle centroid, P may be expressed by a Taylor series expansion (retaining only first-order terms) and may be written as

$$P(x, y) = P_c + \frac{\partial P_c}{\partial x}(x - x_c) + \frac{\partial P_c}{\partial y}(y - y_c), \tag{8}$$

where $(\partial P_c/\partial x)$ and $(\partial P_c/\partial y)$ are the horizontal gradient components of P at the centroid. For each of the three observing stations which form a triangle, the linear expansion using equation (8) can be written as

$$P_1 = P_c + \frac{\partial P_c}{\partial x}(x_1 - x_c) + \frac{\partial P_c}{\partial y}(y_1 - y_c), \tag{9a}$$

$$P_2 = P_c + \frac{\partial P_c}{\partial x}(x_2 - x_c) + \frac{\partial P_c}{\partial y}(y_2 - y_c), \tag{9b}$$

$$P_3 = P_c + \frac{\partial P_c}{\partial x}(x_3 - x_c) + \frac{\partial P_c}{\partial y}(y_3 - y_c). \tag{9c}$$

From equations (9) (a and b) and (7) we get:

$$\frac{\partial P_c}{\partial x} = \frac{(P_1 - P_c)}{(x_1 - x_c)} + \left[\frac{(P_2 - P_c)(x_1 - x_c) - (P_1 - P_c)(x_2 - x_c)}{(y_1 - y_c)(x_2 - x_c) - (y_2 - y_c)(x_1 - x_c)}\right] \times \left(\frac{y_1 - y_c}{x_1 - x_c}\right) \tag{10}$$

and

$$\frac{\partial P_c}{\partial y} = \left[\frac{(P_2 - P_c)(x_1 - x_c) - (P_1 - P_c)(x_2 - x_c)}{(y_2 - y_c)(x_1 - x_c) - (y_1 - y_c)(x_2 - x_c)}\right]. \tag{11}$$

Thus the scalar P and its horizontal gradients at the centroid of each triangle are obtained from equations (7), (10) and (11). These scalar and gradient fields (which are located at the centroid of the triangles) are then interpolated to grid points using OI scheme.

The laplacian of the scalar field P is given by

$$\nabla^2 P = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} P), \tag{12}$$

where $\boldsymbol{\nabla} P$ is the scalar gradient at the grid points.

3. Data used and synoptic situation

The domain of the analysis is bounded by 65°E to 100°E and 5°N to 35°N covering 71 × 61 grid points in the longitude and latitude directions respectively, having the grid resolution of 50 km. The first guess values required for the objective analysis are taken from NCEP/NCAR reanalysis data. The station observations for different days have been collected from Indian Daily Weather Report (IDWR) published by India Meteorological Department (IMD), Pune. Broadly the synoptic situation prevailing over India and adjoining regions during 11 to 15 June 2004 has been described below. The information is based on the IDWR published by

IMD, Pune. The situation in the Indian Ocean was of the See-Saw type. Two deep depressions were formed at a time in Arabian Sea and Bay of Bengal (figure 3).

On 11 June 2004, the depression off south Konkan-Goa coasts intensified into Deep Depression (DD) and lay centred at 08:30 h IST within half a degree of lat. 17.5°N/long. 67.0°E. Under the influence of the trough of low on sea level over central Bay of Bengal a low was formed on 10 June and it concentrated into a depression on 11 June and lay centred within half a degree of lat. 15.5°N/long. 90.0°E. The depression over east-central Bay of Bengal moved slightly westwards and lay centred at 17:30 h IST of 11 June within half a degree of lat. 15.0°N/long. 88.5°E about 650 km southeast of Bhubaneshwar. Moving further in a northwesterly direction, it intensified into a DD and lay over lat. 17.5°N/long. 87.0°E on 12 June. The DD over east central Arabian Sea remained practically stationary and lay centred at 17:30 h IST of 11 June within half a degree of lat. 17.5°N/long. 66.5°E. It moved westwards and lay centred at 08:30 h of 12 June near lat. 17.5°N/long. 66.0°E. On 13 June the DD over west central Bay moved in northwesterly direction crossing the coast and lay close to Puri. The other DD over east central Arabian Sea moved in a northwesterly direction and weakened into depression and lay within half a degree of lat. 18.0°N/long. 85.0°E on 13 June. The DD of Bay of Bengal moved northwestwards and lay close to Jharsuguda on 14 June. The depression of Arabian Sea weakened into a well-marked low pressure area over the same area. On 15 June DD of Bay of Bengal weakened into a depression and subsequently into a well-marked low pressure area and lay over northwestern parts of Chattisgarh and adjoining east Madhya Pradesh. The well-marked low pressure area over east-central Arabian Sea weakened into a low pressure area over the same region.

4. Results and discussions

4.1 *Computations of error*

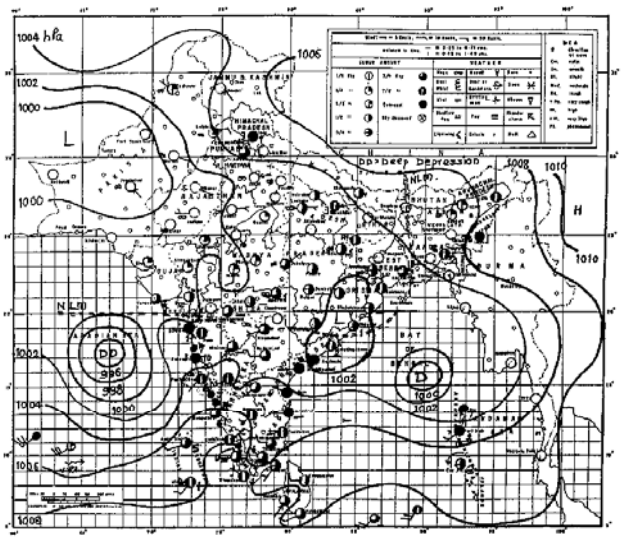
For quantitative determination of analysis errors, root mean square errors ($RMSE$) and correlation coefficients (CC) for scalar, gradient magnitude and laplacian fields are computed by comparing the standard and triangle analyses with National Center for Environmental Prediction (NCEP) analyses over land area only. Appendix A gives the formulae for the computations of mean square errors (MSE) and CC for scalar, gradient magnitude and laplacian.

4.2 *Comparison among analysis schemes*

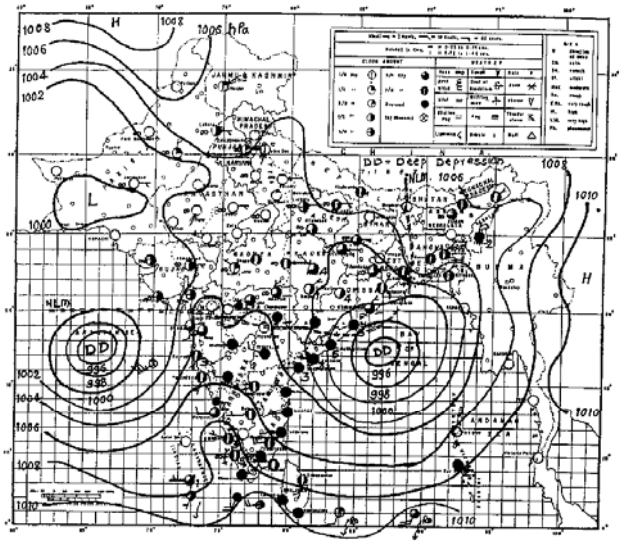
Spencer *et al* (2003) used three pass Barnes scheme (standard method) with analytically generated data to produce scalar, gradient magnitude and laplacian fields. In this study instead of Barnes scheme, OI scheme has been used to produce the above three fields. Before performing analysis with OI scheme using real time data we made analysis with OI scheme using analytically generated data (randomly located), to verify how the OI scheme performs as compared to Barnes scheme. Results are shown in figure 4(a–c). These fields compared well with Spencer's traditional analysis using $L = 20\Delta$, $\Delta = 20$ km. Results produced by OI scheme encouraged us to use real time data in producing the analyses.

The lows, depressions, deep depressions which formed over the Bay of Bengal and Arabian Sea during the period of study are well reflected in the scalar field produced by standard and triangle analysis schemes. Both these analysis schemes have their merits and demerits. The triangle analysis has been produced by using the newly generated centroid data points of the triangle making the region more dense (table 1). Since the quality of any analysis depends upon the data density, the triangle analysis which has more number of data points as compared to standard analysis has produced patterns which are slightly different from the standard analysis. From figures 5(a1–c1) and 6(a1–c1) we can see that the strength of the system has been depicted very well by both the schemes. But there are slight differences in depicting the position of the centre. Triangle analysis produces the centre of the system which is south of its position when compared with IMD's analysis. NCEP analyses for 10 to 15 June, 2004 are used as first guess field and also for comparison. Although analyses have been made for five days (11 to 15 June 2004), analyses of two days (13 and 14) are presented here. NCEP analyses of 13 and 14 June for scalar, gradient magnitude and laplacian are also produced here, figures 5(a1–c1) and 6(a1–c1) respectively. These analyses are used for computing $RMSE$ (square root of MSE) and CC. The $RMSE$ for a perfectly objective analysed field is zero, with larger $RMSE$ indicating decreasing accuracy of the analysed field. The CC is another measure of error between two analyses. It detects the similarities in the patterns between two analyses and is some times referred to as a pattern correlation. It is bounded by ± 1 and is not sensitive to bias in the analysis. Figures 5(a2–c2) and 6(a2–c2) are the analyses of scalar, gradient magnitude and laplacian for standard method of 13 and 14 June respectively. Similarly figures 5(a3–c3) and 6(a3–c3) show scalar, gradient magnitude and

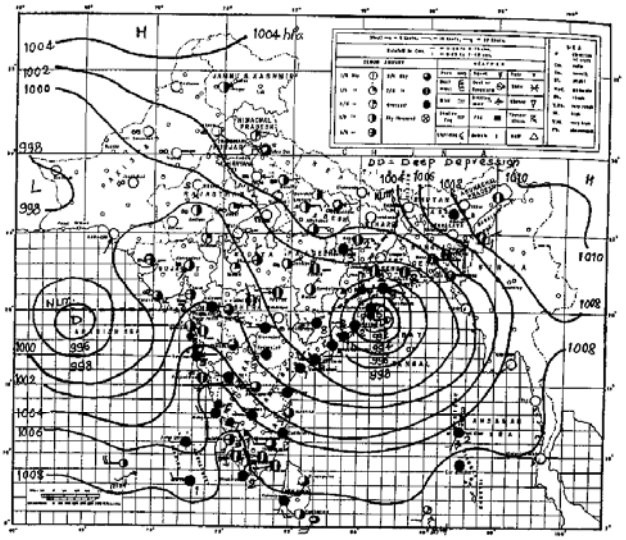
11 June 2004



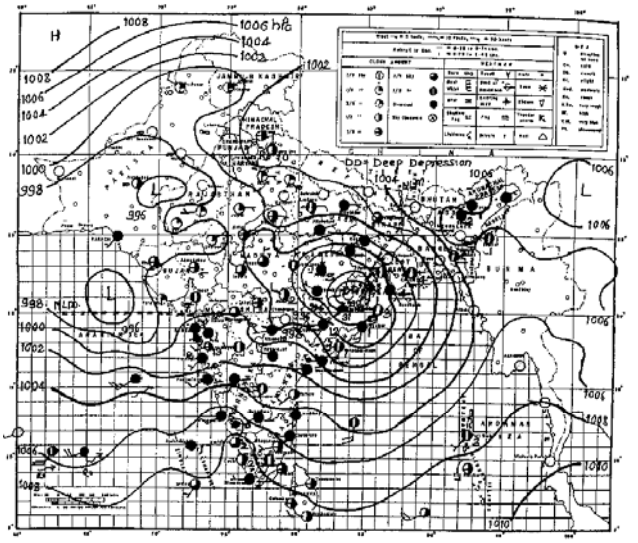
12 June 2004



13 June 2004



14 June 2004



15 June 2004

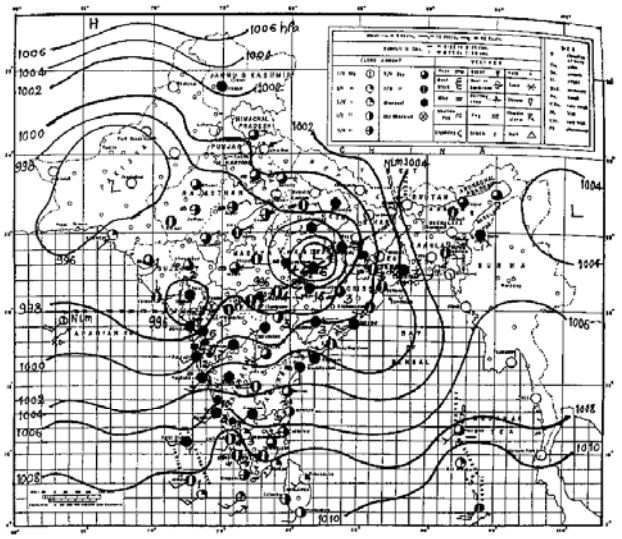


Figure 3. Weather maps at 08:30 h IST form IDWR, IMD, showing the synoptic situations prevailing over the Indian region during the period of study.

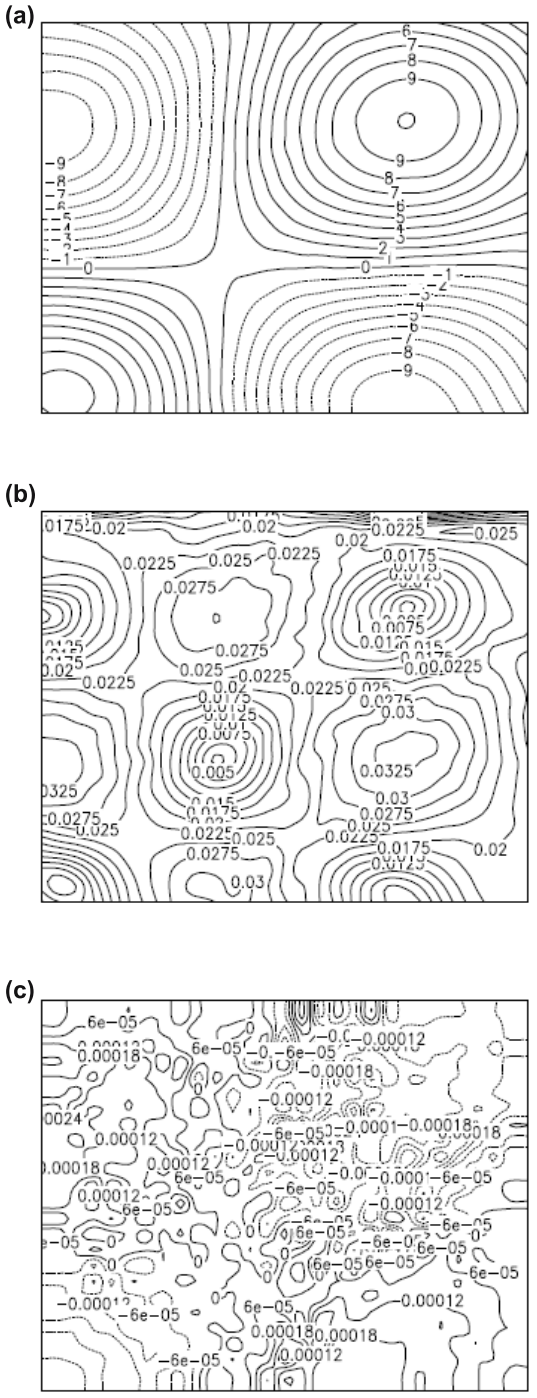


Figure 4. OI analyses of analytically generated data (randomly located): **(a)** scalar, **(b)** gradient magnitude and **(c)** laplacian. Dashed contour lines indicate negative values ($L = 25\Delta$, $\Delta = 100$ km).

laplacian analyses for triangle method of 13 and 14 June respectively. From figures 5(a2), 5(a3) and 6(a2), 6(a3) it is seen that for scalar objective analysis, analysis patterns are well represented by both the methods, standard and triangle (high correlation of 0.92). On comparing both the analyses (standard, triangle) with NCEP scalar analysis it is found that *RMSE* for standard method are less than triangle method and this is seen on most of the days except 14 June, figure 7(a). In order to see how the triangle method analyses derivative fields which are supposed to be more accurate than the standard method, let us compare figures 5(b2) with 5(b3) and 6(b2) with 6(b3) (gradient) and 5(c2) with 5(c3) and 6(c2) with 6(c3) (laplacian). Noisy gradient and laplacian fields are seen in the standard method of analysis, whereas gradient and laplacian fields produced by triangle method are comparatively smooth. Figure 7(a–f) shows the *RMSE* and *CC* for scalar, gradient and laplacian fields. Gradient and laplacian *RMSE* have been multiplied by 10^3 and 10^5 respectively before plotting. It is observed that triangle gradient analysis errors are less on most of the days except 11 June. The *CC* for triangle gradient analyses is lower than standard gradient analysis for 11 to 13 June and then has higher values on 14 and 15 June. From figures 5(c2) and 5(c3) (in the context of *RMSE* and *CC*) it is found that laplacian analyses produced by standard scheme are inferior to the laplacian analyses produced by triangle method. Due to higher *CC* of laplacian analysis (triangle method) better patterns of triangle analysis are observed. Figure 7(c, f) shows the *RMSE* and *CC* for the laplacian analyses. It can be seen from the figure 7(b, e) that in some cases, viz., on 11 June standard gradient *RMSE* are slightly less than that of triangle gradient, and also the *CC* for triangle gradient are less than standard gradient on 11, 12 and 13 June. Thus from the analysis errors (*RMSE* and *CC*) we cannot conclude the superiority or inferiority of one analysis over the other. Hence in order to produce an accurate scalar field and at the same time to get accurate analysis of gradient and laplacian fields we made an experiment based on variational analysis scheme developed by Spencer *et al* (2003), which produces scalar analysis resembling to standard scalar analysis and at the same time produces gradient and laplacian analysis which are similar to the gradient and laplacian fields produced by triangle method.

5. Variational objective analysis scheme

5.1 *Euler–Lagrange equation*

In regard to variational optimization of meteorological parameter a given measure of the 'distance' between the variational scalar analysis and the standard scalar analysis is minimized. The variational analysed field must at the same time satisfy some constraint. The constraint is that the difference between derivative of variational scalar and derivative of triangle scalar analysis is minimum. This is achieved by minimizing the cost function I,

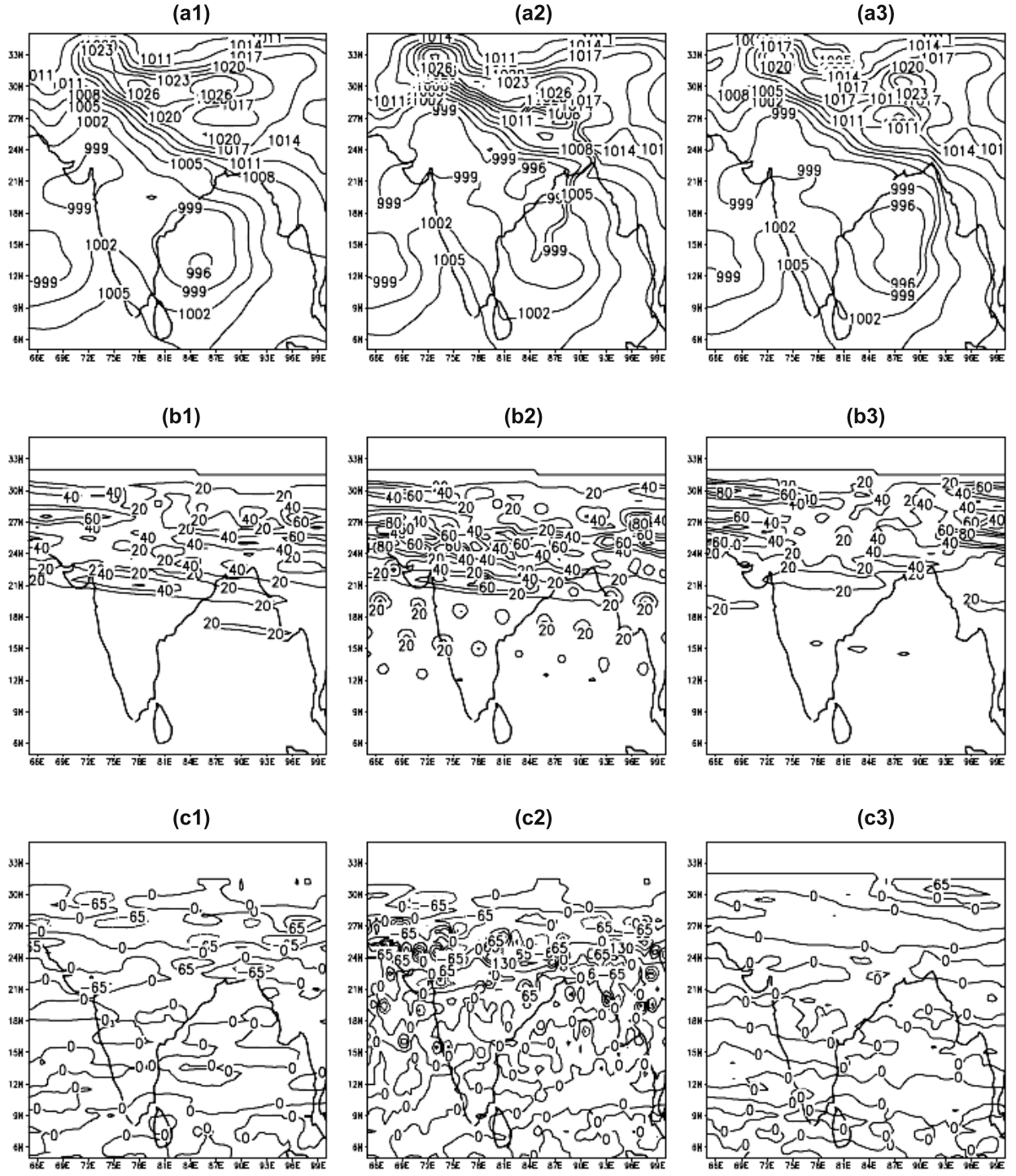


Figure 5. A 9-panel figure for 13 June 2004 showing, (a1)–(a3): scalar analyses for NCEP/NCAR, standard and triangle methods respectively. (b1)–(b3): Same as (a1)–(a3) but for gradient magnitude. (c1)–(c3): Same as (a1)–(a3) but for laplacian.

$$I = \iint \left[\alpha^2 (P^V - P^T)^2 + \beta^2 \left(\frac{\partial P^V}{\partial x} - \frac{\partial P^\Delta}{\partial x} \right)^2 + \beta^2 \left(\frac{\partial P^V}{\partial y} - \frac{\partial P^\Delta}{\partial y} \right)^2 \right] dxdy, \qquad (13)$$

where P^V is the variational scalar analysis, P^T is the standard scalar analysis, $(\partial P^\Delta/\partial x)$ and $(\partial P^\Delta/\partial y)$ are components of the gradient of the triangle scalar analysis along zonal and meridional directions respectively and α and β are the weighting functions. From the theory of calculus of

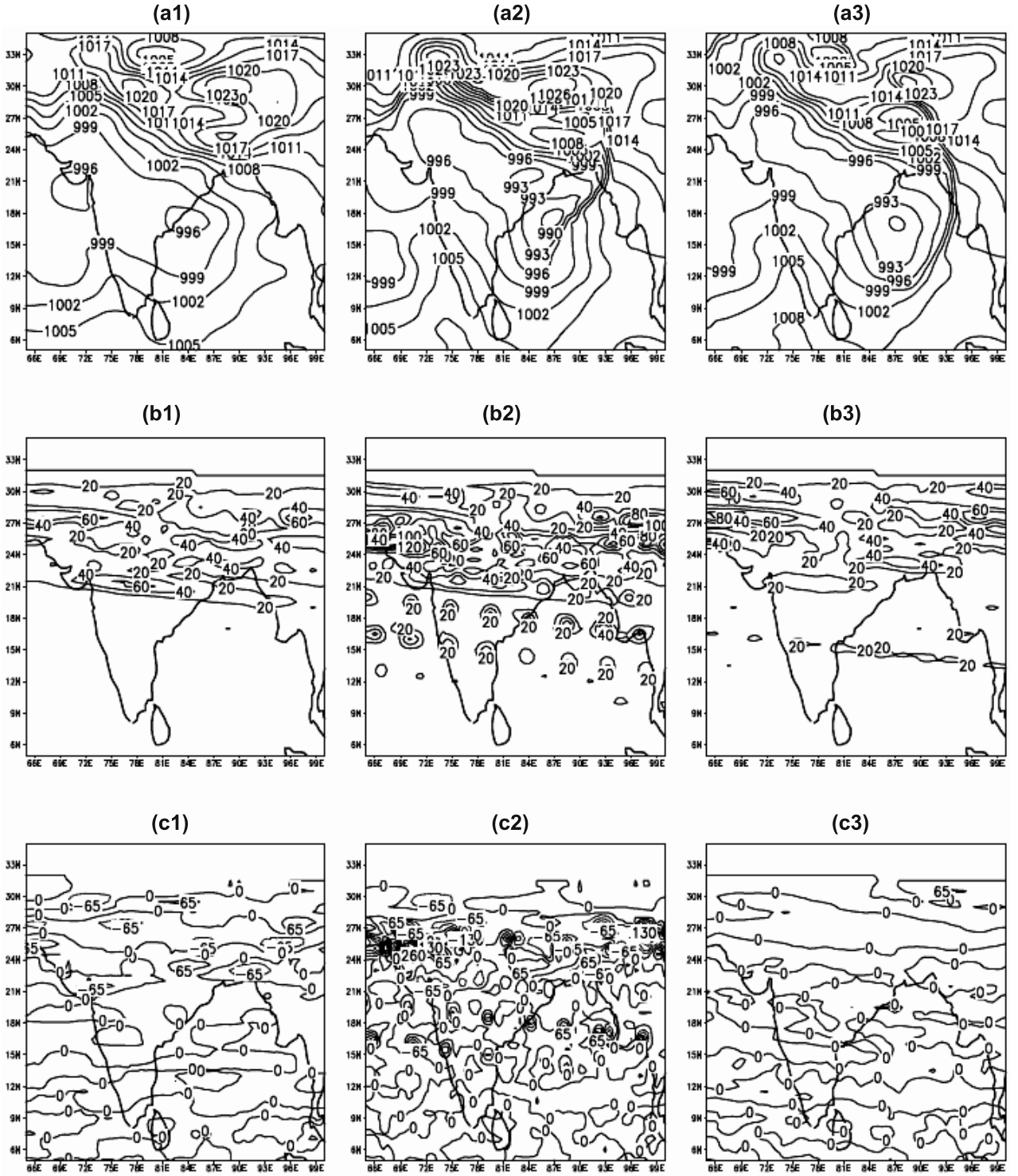


Figure 6. Same as figure 5 but for 14 June 2004.

variation, the minimization is achieved by making the first variance of the cost function I to zero, i.e.,

$$\delta I = 0.$$

The Euler–Lagrange equation corresponding to equation (13) is given by

$$(P^V - P^T) + \left(\frac{\beta}{\alpha}\right)^2 \times \left[\frac{\partial}{\partial x}\left(\frac{\partial P^\Delta}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P^\Delta}{\partial y}\right) - \nabla^2 P^V\right] = 0. \quad (14)$$

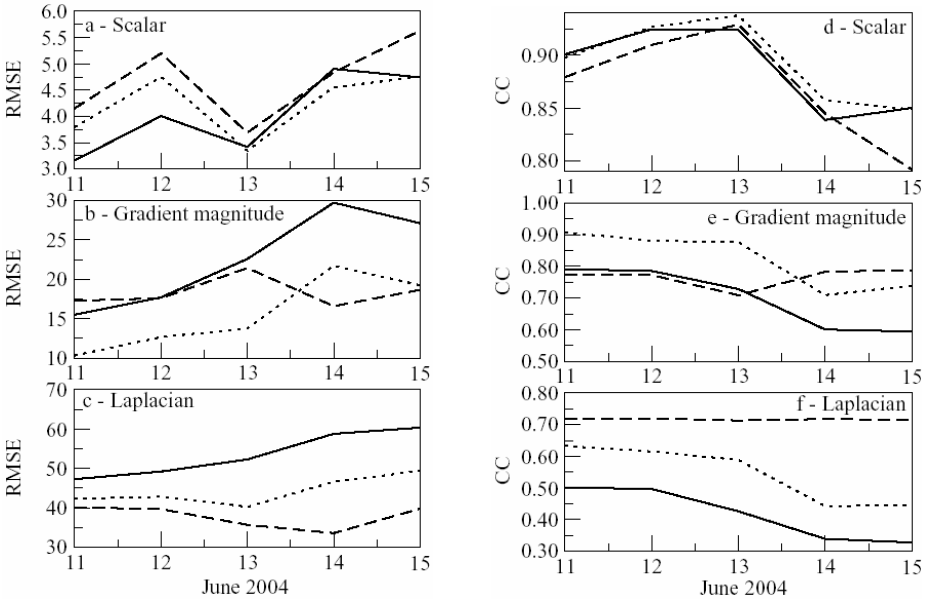


Figure 7. Plot for *RMSE* and *CC* for the standard (solid line), triangle (broken line) and variational (dotted line) for different days.

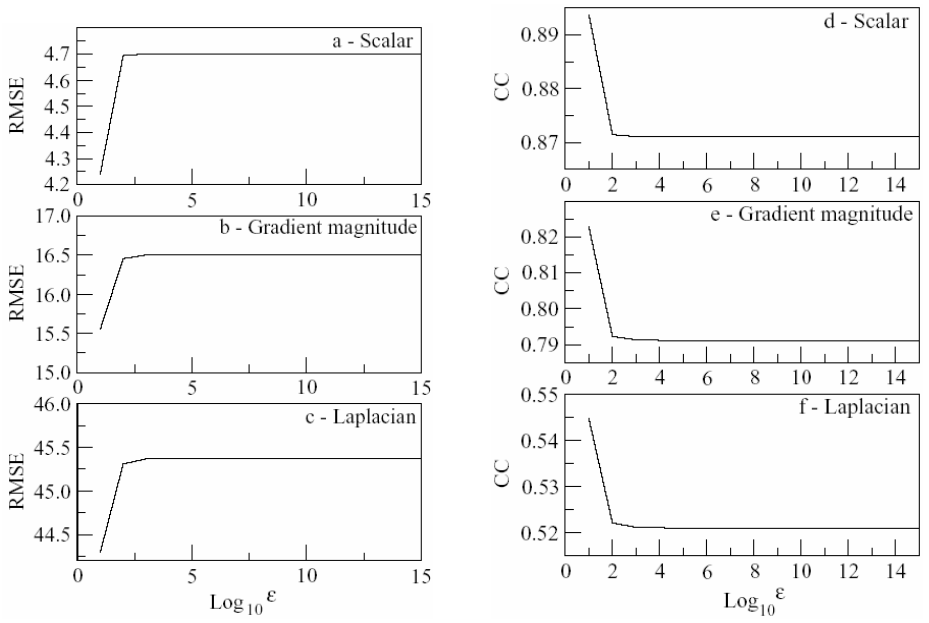


Figure 8. Plot for *RMSE* and *CC* of variational method analyses as a function of the logarithm of the weighting factor ε for the scalar, gradient and laplacian. Five days of *RMSE* are averaged to create these curves.

The complete derivation of the Euler–Lagrange equation is given in Appendix B. Equation (14) can be further written in the following form:

$$\nabla^2 P^V - \frac{P^V}{\varepsilon} = \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} P^{\Delta}\right) - \frac{P^T}{\varepsilon}, \qquad (15)$$

where $\varepsilon = (\beta/\alpha)^2$. It is seen from equation (15) that ε plays an important role in producing improved analysis using the variational method. As the value of ε increases $\nabla^2 P^V$ approaches $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla} P^{\Delta})$ and virtually for large value of ε, $\nabla^2 P^V$ equals $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} P^{\Delta})$ where $\boldsymbol{\nabla} P^{\Delta} = \hat{i}(\partial P^{\Delta}/\partial x) + \hat{j}(\partial P^{\Delta}/\partial y)$ and $\hat{i}$ and $\hat{j}$ are the unit vectors in zonal and meridional directions respectively. Equation (15) is an elliptic partial differential equation and is numerically solved for P^V using the relaxation technique.

As already stated the value of the weighting function ε should be chosen very carefully such

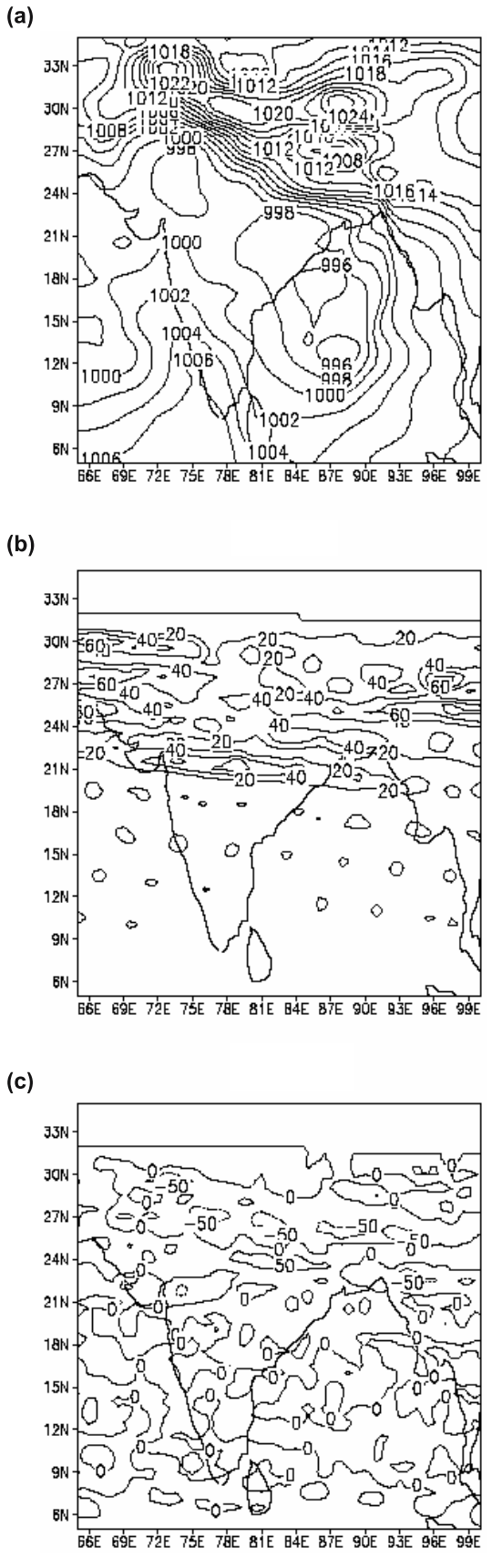


Figure 9. Variational analysis for 13 June 2004, **(a)** scalar, **(b)** gradient and **(c)** laplacian.

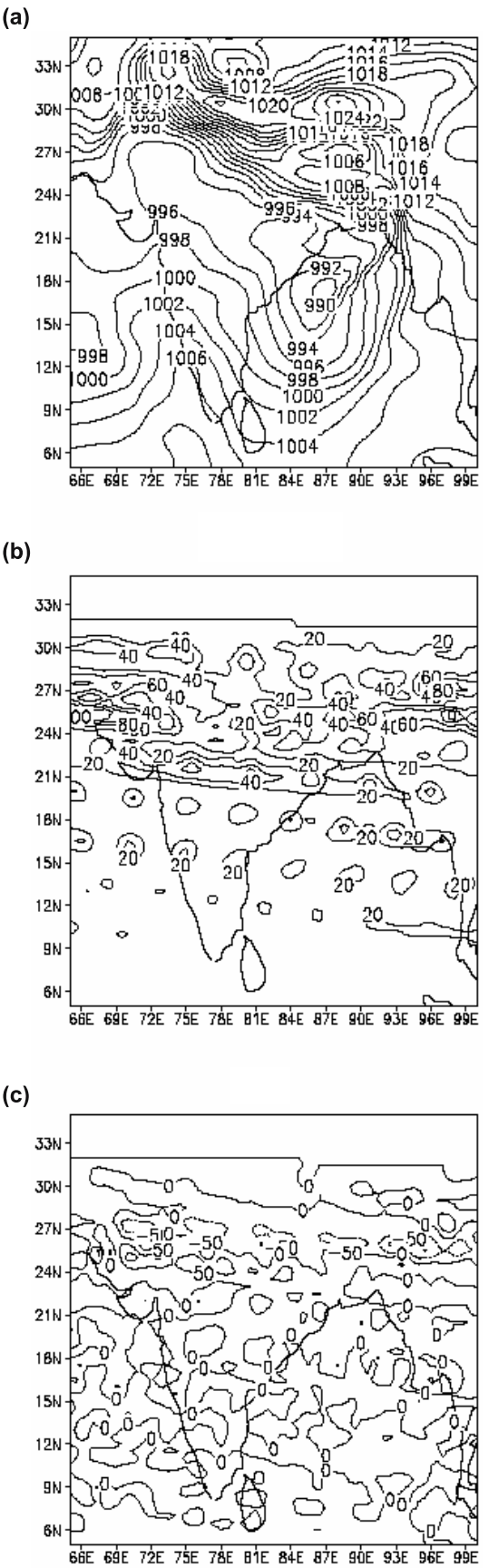


Figure 10. Same as figure 9 but for 14 June 2004.

that accurate values of both derivative and scalar field are obtained. For this purpose, the *RMSE* and *CC* are computed by comparing the variational analyses with the NCEP analyses for different values of ε ranging from 10 to 10^{15} ($P^V, \boldsymbol{\nabla} P^V, \nabla^2 P^V$), figures 8(a–f). It is found that *RMSE* in all the three cases are minimum for $\varepsilon = 10$ and are maximum at $\varepsilon = 10^2$ and then remain constant for $\varepsilon > 10^2$. A similar situation is seen in the case of *CC*. *CC* are maximum for $\varepsilon = 10$ and then sharply decreased to a minimum at $\varepsilon = 10^3$ and thereafter remain constant for $\varepsilon > 10^3$.

5.2 *Results*

The right hand side of equation (15) is the combination of standard and triangle methods to produce the variational analysis. It is seen that variational analysis uses the best aspects of both (standard and triangle) schemes. In this study we have used OI scheme as standard scheme and also the triangle analysis is generated by using OI scheme. OI scheme requires first guess values which have been taken from NCEP/NCAR reanalysis data. It is obvious that the variational analysis produced will certainly contain the information from first guess particularly over data sparse oceanic region. Variational analyses of 13 and 14 June are shown in figures 9 and 10 respectively. Comparing these figures with the figures 5(a2–c2) and 5(a3–c3) it is seen that variational scalar analyses produced patterns which are similar to those of standard scalar analysis and at the same time variational gradient and laplacian analyses match those of the triangle method. From figure 7(a) it is found that *RMSE* for variational scalar analysis for 13 and 14 June are slightly higher than standard analysis but less than triangle analysis. On 13, 14 and 15 June, variational *RMSE* are closer to standard analysis error as compared to triangle analysis. Computations of *CC* also support the above observations. *CC*s of variational derivative analysis are higher than those of standard derivative analysis. The variational scalar analyses of 13 and 14 June 2004 (figures 9a and 10a) match the standard scalar analyses of 13 and 14 June 2004 (figures 5(a2) and 6(a2)) but with few differences. The variational analysis has slightly higher *RMSE* relative to the standard analysis but better than the triangle scalar analysis. Further, we calculated the skill score (SS) which gives the percentage improvement of one analysis over the reference analysis (NCEP). It is expressed as:

$$SS = \left[1.0 - \frac{\sum_{j=1}^{n} MSE(j)}{\sum_{j=1}^{n} MSE_{\text{ref}}(j)}\right] \times 100\%, \quad (16)$$

where *MSE* is the mean square error of any particular analysis (standard or variational analysis) and MSE_{ref} is the mean square of the reference analysis when compared with observations. $\sum$ is the sum of n individual analysis. It was found that *SS* for the variational analysis (36%) is higher than the *SS* for the standard analysis (25%). From the higher value of *SS* it can be concluded that variational scheme produces analysis which is superior to standard analysis so far as the scalar field is concerned. Thus a relatively accurate smooth analysis can be produced by variational analysis scheme in which standard scalar analysis and triangle derivative analysis are used as input. Hence the purpose of developing variational analysis scheme is achieved to some extent.

6. Conclusions

The work described in this paper is based on the work of Spencer *et al* (2003). They applied this technique to analyse analytically generated data. In this study a real time data test has been carried out. Three different analysis schemes (standard, triangle and variational) for the analysis of scalar variable and its derivatives have been discussed here. The standard method of objective analysis is OI scheme which maps the mslp field (scalar) to grid points. This analysis is found to be superior to the triangle method, which interpolates triangle centroid scalar variable to grid points. The value of the scalar at the centroid of the triangle is obtained by taking the arithmetic mean of observed value at the three stations comprising each triangle. As mentioned in section 2, on a particular day the number of observations is 142 over the Indian region. From these observations, 269 triangles have been formed and the gradients have been computed at the centroid of each triangle. In fact 196 triangles are obtained after deleting those triangles whose one angle is less than 15°. Thus there is an increase of 38% of more data for the triangle method as compared to standard method. Thus the increased number of observations reduces the gap (increased data density) and therefore produces better analysis as compared to the analysis over large data void region. The triangle method computes derivatives directly from the observations and then interpolates these derivatives to the grid points using OI scheme. The triangle method of estimating derivative field is superior to the standard method in which finite difference scheme is applied to the grid point scalar field to obtain derivative field.

In order to produce superior scalar and superior derivative fields simultaneously, variational analysis scheme has been developed to analyse the

mslp field over Indian and adjoining regions. This scheme combines the better part of both the triangle method and the standard method. Variational scheme uses simultaneously, the scalar field from standard analysis and derivative estimate from triangle method to produce better analysis. The variational scheme is developed simply to minimize the difference between the variational scalar analysis and the standard scalar analysis, while at the same time minimizing the difference between the variational gradient analysis and the triangle gradient analysis. To determine the appropriate values of the weighting function ε involved in equation (15) analysis with variational scheme was carried out for different values of ε ranging from 10 to 10^{15} and computing *RMSE* and *CC*s for scalar and derivative fields comparing with NCEP analyses.

Acknowledgements

The authors are grateful to the Director, Indian Institute of Tropical Meteorology, Pune, India for his interest and for providing the necessary facilities to carryout this work. Thanks are due to the Head, Forecasting Research Division for his encouragement. We are extremely grateful to Dr. Phillip L. Spencer of NOAA/NSSL, USA for his help during different stages of this study and for running the triangulation code of Dr. Paul Bourke. Thanks are also due to the India Meteorological Department for the daily pressure data and NCEP/NCAR for providing the reanalysis datasets. The authors thank the two anonymous reviewers for their constructive comments that led to improvements in the final manuscript.

Appendix A

Determination of error and correlation

Let P^g and P^{ncp} be the analyses of standard scalar variable and the NCEP scalar analysis respectively, then mean square error (*MSE*) is given by:

$$MSE = \frac{\sum_{i,j} (P^g - P^{ncp})^2}{N_g}. \tag{A1}$$

As the gradient is a vector quantity, *MSEV* is computed using the following formula:

$$MSEV = \frac{\sum_{i,j} \left\langle (P_x^g - P_x^{ncp})^2 + (P_y^g - P_y^{ncp})^2 \right\rangle}{N_g}. \tag{A2}$$

The *MSE* of the laplacian (*MSEP*) is calculated as:

$$MSEP = \frac{\sum_{i,j} \left\langle \left\{ (P_x^g)_x + (P_y^g)_y \right\} - \nabla^2 P^{ncp} \right\rangle^2}{N_g}, \tag{A3}$$

where N_g is the number of grid points, P_x^g, P_y^g and $(P_x^g)_x$, $(P_y^g)_y$ are the first and second derivatives of the scalar P^g in zonal and meridional directions respectively.

The correlation coefficient is given by:

$$CC = \frac{\sum_{i,j} \left\{ (A_{i,j} - \bar{A})(T_{i,j} - \bar{T}) \right\}}{\left[\sum_{i,j} (A_{i,j} - \bar{A})^2 \sum_{i,j} (T_{i,j} - \bar{T})^2 \right]^{1/2}}, \tag{A4}$$

where $A_{i,j}$ is the analysed value, $\bar{A}$ is the average value of the analysis, $T_{i,j}$ is grid point NCEP value and $\bar{T}$ is average value of NCEP analysis.

Appendix B

Derivation of Euler–Lagrange equation

In the case of two independent variables x and y, the cost function I can be written as:

$$I\left\{P^V(x,y)\right\} = \iint_S F dx dy, \tag{B1}$$

where

$$F = \alpha^2 (P^V - P^T)^2 + \beta^2 \left(\frac{\partial P^V}{\partial x} - \frac{\partial P^\Delta}{\partial x} \right)^2 + \beta^2 \left(\frac{\partial P^V}{\partial y} - \frac{\partial P^\Delta}{\partial y} \right)^2. \tag{B2}$$

S is the domain on the (x, y) plane where P^V is continuous and has continuous derivatives up to second order and P^V is the prescribed values on the boundary of the domain.

(B2) can be further written as:

$$F = \alpha^2 (P^V - P^T)^2 + \beta^2 \left(P_x^V - \frac{\partial P^\Delta}{\partial x} \right)^2 + \beta^2 \left(P_y^V - \frac{\partial P^\Delta}{\partial y} \right)^2, \tag{B3}$$

where

$$P_x^V = \frac{\partial P^V}{\partial x} \quad \text{and} \quad P_y^V = \frac{\partial P^V}{\partial y}.$$

The problem is now to determine the value of P^V. For this we require I to be minimum.

Based on the theory of calculus of variation,

$$\delta I = 0. \tag{B4}$$

The necessary condition for the vanishing of δI is called the Euler–Lagrange equation and is given as:

$$\frac{\partial F}{\partial P^V} - \frac{\partial}{\partial x}\left(\frac{\partial F}{\partial P_x^V}\right) - \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial P_y^V}\right) = 0. \tag{B5}$$

Thus from equations (B3) and (B5) we have:

$$2\alpha^2(P^V - P^T) - 2\beta^2\left\{P_{xx}^V - \frac{\partial}{\partial x}\left(\frac{\partial P^\Delta}{\partial x}\right)\right\} - 2\beta^2\left\{P_{yy}^V - \frac{\partial}{\partial y}\left(\frac{\partial P^\Delta}{\partial y}\right)\right\} = 0,$$

$$\alpha^2(P^V - P^T) - \beta^2(P_{xx}^V + P_{yy}^V) + \beta^2\left\{\frac{\partial}{\partial x}\left(\frac{\partial P^\Delta}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P^\Delta}{\partial y}\right)\right\} = 0,$$

$$(P^V - P^T) - \left(\frac{\beta}{\alpha}\right)^2 \nabla^2 P^V + \left(\frac{\beta}{\alpha}\right)^2 \left(\hat{i}\frac{\partial}{\partial x} + \hat{j}\frac{\partial}{\partial y}\right) \cdot \left(\hat{i}\frac{\partial P^\Delta}{\partial x} + \hat{j}\frac{\partial P^\Delta}{\partial y}\right) = 0,$$

$$(P^V - P^T) - \varepsilon\nabla^2 P^V + \varepsilon\left\{\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}P^V)\right\} = 0,$$

where $\varepsilon = (\beta/\alpha)^2$.

Thus finally we have:

$$\nabla^2 P^V - \frac{P^V}{\varepsilon} = \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}P^V) - \frac{P^T}{\varepsilon}. \tag{B6}$$

(B6) is the required equation.

References

Alaka M A and Elvander R C 1972 Optimum Interpolation from observations of mixed quality; *Mon. Weather Rev.* **100** 612–624.

Barnes S L 1964 A technique for maximizing details in numerical weather map analysis; *J. Appl. Meteor.* **3** 396–409.

Barnes S L 1973 Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK, 41 pp. [NTIS COM-73-10781].

Barr S, Widger Jr. W K, Miller I A and Stanton R 1971 Objective subsynoptic upper level analysis; *J. Appl. Meteor.* **10** 410–417.

Bellamy J C 1949 Objective calculation of divergence, vertical velocity and vorticity; *Bull. Amer. Meteor. Soc.* **30** 45–49.

Benjamin S G and Seaman N L 1985 A simple scheme for objective analysis in curved flow; *Mon. Weather Rev.* **113** 1184–1198.

Bergman K H 1979 Multivariate analysis of temperature and winds using optimum interpolation; *Mon. Weather Rev.* **107** 1423–1444.

Bourke P 1989 An algorithm for interpolating irregularly-spaced data with application in terrain modeling; presented at Pan Pacific Computer Conference, Beijing, China, January 1989.

Bussieres N and Hogg W 1989 The objective analysis of daily rainfall by distance weighting schemes on a mesoscale grid; *Atmos.-Ocean* **27** 521–541.

Ceselski B F and Sapp L L 1975 Objective wind field analysis using line integral; *Mon. Weather Rev.* **103** 89–100.

Cressman G 1959 An operational objective analysis system; *Mon. Weather Rev.* **87** 367–374.

Dey C H and Morone L L 1985 Evolution and performance of the National Meteorological center global data assimilation system: Jan.–Dec., 1982; *Mon. Weather Rev.* **113** 304–318.

Doswell C A III and Caracena F 1988 Derivative estimation from marginally sampled vector point functions; *J. Atmos. Sci.* **45** 242–253.

Eliassen A 1954 Provisional report on spatial covariance and autocorrelations of the pressure field; *Inst. Weather and Climate Research Academy of Science*, Oslo, Rep. No. 5.

Endlich R M and Clark J R 1963 Objective computation of some meteorological quantities; *J. Appl. Meteor.* **2** 66–81.

Gandin L S 1963 The objective analysis of meteorological fields; Israel program for scientific translations, Jerusalam, 240 pp.

Koch S E, desJardins M and Kocin P J 1983 An interactive Barnes objective map analysis scheme for use with satellite and conventional data; *J. Climate Appl. Meteor.* **22** 1487–1503.

Koch S E and Saleeby S 2001 An automated system for the analysis of gravity waves and other mesoscale phenomena; *Wea. Forecasting* **16** 661–679.

Mitra A K, Dasgupta M, Singh S V and Krishnamurti T N 2003 Daily rainfall for the Indian monsoon region from merged satellite and rain gauge values: Large-scale analysis from real-time data; *J. Hydrometeorology* **4** 769–781.

Ogura Y and Chen Y L 1977 A life history of an intense mesoscale convective storm in Oklahoma; *J. Atmos. Sci.* **34** 1458–1476.

Rajamani S, Talwalkar D R, Ray S P and Upasani P U 1983 Objective analysis of wind field over Indian region by Optimum Interpolation method; *Mausam* **34** 43–50.

Ripley B D 1981 Spatial statistics; John Wiley and Sons, 252 pp.

Sasaki Y 1958 An objective analysis based on the variational method; *J. Meteor. Soc. Japan* **36** 77–81.

Schaefer J T and Doswell C A III 1979 On the interpolation of a vector field; *Mon. Weather Rev.* **107** 458–476.

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

Sinha S K, Narkhedkar S G and Rajamani S 1998 Application of Sasaki's numerical variational technique to the analysis of height and wind fields over Indian region; *Mausam* **49** 1–10.

Spencer P L and Doswell C A III 2001 A quantitative comparison between standard and line integral methods of derivative estimation; *Mon. Weather Rev.* **129** 2538–2554.

Spencer P L, Stensrud D J and Fritsch J M 2003 A method for improved analyses of scalars and their derivatives; *Mon. Weather Rev.* **131** 2555–2576.

Thiebaux H J 1976 Anisotropic correlation function for objective analysis; *Mon. Weather Rev.* **104** 994–1002.

Zamora R J, Shapiro M A and Dowell C A III 1987 The diagnosis of upper tropospheric divergence and ageostrophic wind using profiler wind observations; *Mon. Weather Rev.* **115** 871–884.

MS received 19 February 2008; revised 14 May 2008; accepted 23 May 2008