J. Ind. Geophys. Union (2003)
Vol.7, No.2, pp.87-95

Cumulative Semivariogram Technique for Objective Analysis of height field over India and adjoining region

S.K.Sinha, S.G.Narkhedkar, D.R.Talwalkar and P.N.Mahajan
Indian Institute of Tropical Meteorology, Dr. Homi Bhabha Road, Pashan, Pune – 411008.

ABSTRACT
In this study objective analysis of height field using cumulative semivariogram (CSV) technique has been carried out for 850, 700 and 500 hPa levels. Normally in any objective analysis scheme we assume homogeneity of meteorological phenomena i.e. circular radius of influence, but in reality it is not so. Experimental CSV methods yield irregular radii of influence depending on the regional heterogeneous behaviour of the phenomenon. The experimental CSVs for daily height data over India and adjoining region have been computed. These experimental CSVs have been converted into experimental CSV weighting functions for the objective analysis. These experimental weights are compared with geometrical weighting functions, which are available in meteorological literature. It is found that none of the geometric weighting functions completely represent the regional variations. Unlike Optimum Interpolation (OI) in which at each grid point matrix inversions are carried out to determine the weights, there is no such matrix inversion in CSV technique. As such analysis using CSV scheme requires small computation time as compared to OI scheme without hampering the analysis.

INTRODUCTION

For objective analysis of meteorological variables a number of methods have been developed (Gustavsson 1981). The earlier computer aided analysis techniques (Panofsky 1949) handled this problem by fitting polynomials to the irregularly spaced synoptic observations. But this scheme is unstable in regions with non-uniform data coverage. Successive correction methods (Bergthorsson & Döös 1955; Cressman 1959), which have been used for the last four decades, are empirical in nature. Successive correction method has typical shortcomings: too much attention is given to the observations relative to first guess, observations occurring in areas of high observational density are given too much weights relative to observations in areas of low density. In this scheme, radius of influence is chosen as per one's experience and this value is decreased with successive scans and the resulting field of latest scan is taken as new approximation.

Eliassen (1954) and Gandin (1963) introduced optimum Interpolation (OI) procedure to meteorology. In OI scheme it is assumed that observations are spatially correlated as such observations that are close to each other are highly correlated. Hence as observations get farther apart regional dependence decreases. Details of OI scheme are discussed in later section.

In CSV scheme the radius of influence R is determined based on the observed data without any subjectivity. In the above scheme (successive correction) R was assumed constant without any directional variations. Hence, spatial anisotropy of observed fields is ignored. The objective of this paper is to analyse height field to produce gridded data from irregularly distributed sparse data within a region using CSV technique. This method provides information either regionally or along any preferred directions. Section Cumulative Semivariogram scheme gives detail of CSV technique.

OPTIMUM INTERPOLATION METHOD

Let the scalar field f denotes correlated variables such as height. Subscripts i and j refer to observation points and x refers to grid point. Superscripts o and p refer to observe and background field (predicted value) respectively. n is the number of observations, $\overline{f_i f_j}$ denotes the value of $\overline{\left(f_i^o - f_i^p\right)\left(f_j^o - f_j^p\right)}$, where bar represents average over number of cases. In this notation, OI scheme is given by following equations:

$$\sum_{j=1}^{n}\left(C_{ij} + \lambda^2\delta_{ij}\right)w_j = C_{xi} \qquad i = 1,....n \tag{1}$$

The w_j are the weights for OI correction. Analysed values at the grid point are given by

$$f_x^a = f_x^p + \sum_{j=1}^{n} w_j\left(f_j^o - f_j^p\right) \tag{2}$$

Where f_x^a, f_x^p represent analysed and predicted values at grid point x. $C_{ij}(= \overline{f_i f_j} / \sigma_o^2)$ is the spatial correlation coefficient, $\lambda^2 (= \sigma_{\varepsilon_i}^2 / \sigma_o^2)$ is normalized observational error variance, δ_{ij} is kroneker delta, $\sigma_{\varepsilon_i}^2$ is observational error variance and σ_o^2 is background error variance.

CUMULATIVE SEMIVARIOGRAM SCHEME

Sen (1989, 1997) proposed the CSV method, which is an alternate form of classical semivariogram technique of Matheron (1963). CSV is a graph, which shows the variation of successive

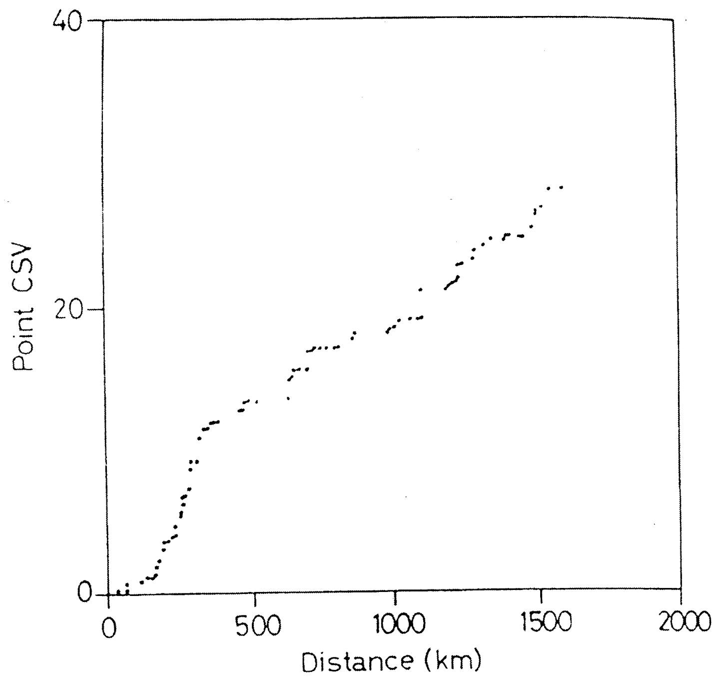

Figure 1. Sampled cumulative semivariogram functions against distance.

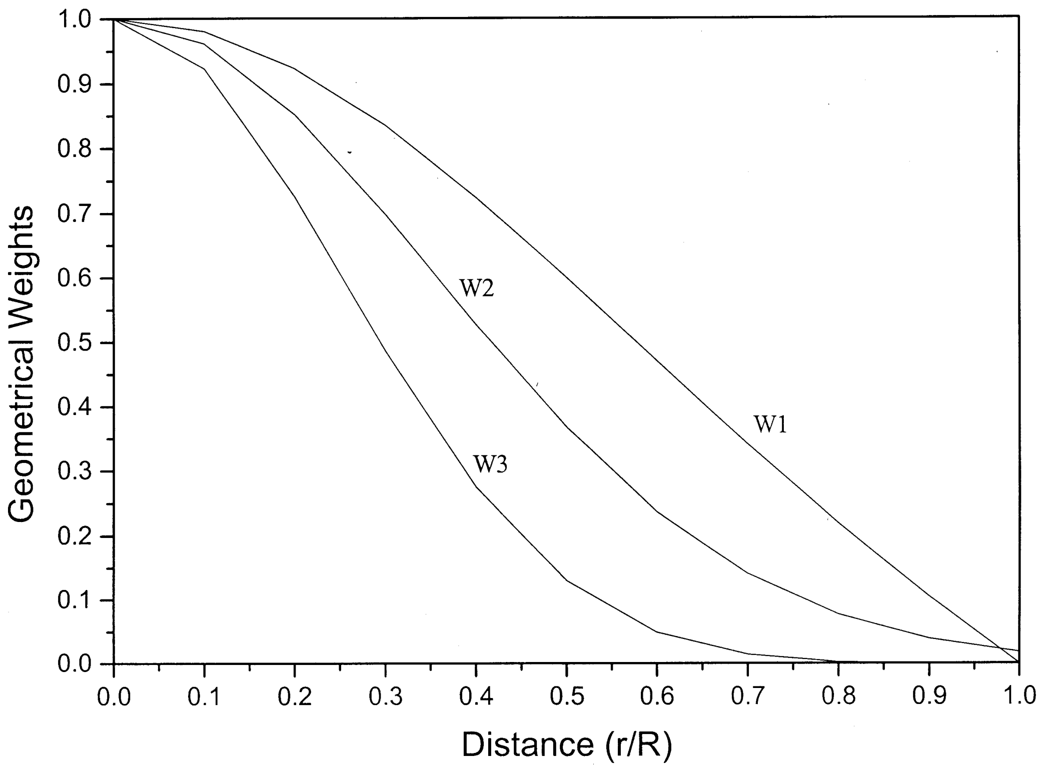

Figure 2. Geometrical weighting functions.

half squared difference summations with distance. This is an increasing sampled CSV function. Through CSV we get a measure of regional dependence such that two sites, which are close to each other, have higher correlations and hence smaller CSV values.

Computation of CSV

Let there be n number of observations. First the distances (h_i) between every possible pair of observations are computed (h_i, i = 1,....m). The number of possible pairs are m = n(n-1)/2. Next the corresponding half squared differences d (h_i) for the height field for each distance is obtained. For example if the height field has the values $Z_1(h_i)$ and $Z_2(h_i)$ at two distinct locations, then half squared difference is given by

$$d(h_i) = 0.5\,[\,Z_1(h_i)-Z_2(h_i)]^2. \qquad (3)$$

Next the successive summation of $d(h_i)$ is taken starting from the smallest distance to the largest in their order. These CSV values are plotted against the corresponding distance. The sample CSV function is shown in Fig.1. This sample CSV function is free of subjectivity as no previous selection of distance classes is made as done in OI schemes (Petersen & Truskey 1969).

Geometrical weighting function

A number of geometrical weighting functions available in meteorological literature that are used by different researchers are given below. Fig.2 shows the various dimensionless geometrical weighting functions. These are leveled as

(a) Ratio Model (W1) : $W = [(R^2 - r^2)/(R^2 + r^2)],\ r < R$
$= 0,\ r \geq R$

(b) Exponential Model (W2) : $W = \exp[-4\,(r/R)^2]$,

(c) Power Model (W3) : $W = [(R^2 - r^2)/(R^2 + r^2)]^4,\ r < R$
$= 0,\ r \geq R$

Where r is the distance between the grid point and the observation. R is radius of influence and is determined subjectively from personal experience. These geometrical weighting functions cannot reflect morphology i.e. regional variability of the phenomena. Our main aim is to determine the event dependent weighting function by using experimental CSVs.

DATA AND SYNOPTIC SITUATION

Daily height data for 10 years (1976-1985) of July month, 12 GMT for the levels 850, 700 and 500 hPa at all the available RS stations over India are used for this study. Analyses for two different synoptic situations (4-8 July 1979 and 26-30 July 1991, 12 GMT) have been carried out with both the schemes (CSV and OI). A region bounded by 41.250°E and 108.750°E longitude and 1.875°N to 39.375°N latitude with a grid resolution of 1.875° is considered for the first case and in the second case region bounded by 40°E to 120°E and 25°S to 40°N with a grid resolution of 2.5° is considered for analysis. FGGE analyses for 3-7 July 1979 and NCMRWF analyses for 25-29 July 1991 are used as first guess fields.

July 1979 depression

On 4 July a low was formed over the region 20°N, 90°E, it moved slowly at the beginning and intensified into a depression on 7 July with its center near 20°N and 88°E. By 8 July it moved westnorthwestward and crossed north Orrisa coast.

July 1991 depression

A cyclonic circulation was observed on 25 July, which extended upto mid-troposphere levels over North Bay and adjoining central Bay. Under this influence a low was formed over northwest Bay on 26 July and became a depression on 27 July with its center lying about 250 km southeast of Calcutta. It further intensified into a deep depression with its center at south of southwest Calcutta. It crossed Paradip coast on 29 July. It moved in westerly direction and on 30 July it was over Madhya Pradesh.

COMPUTATIONS AND RESULTS

In both these schemes instead of raw observations the deviations (anomalies) of the observations from some background field are analysed for computation purpose. The linear combination of the weighted anomalies are added to the background field to get the analysed values in case of OI (Eq.2) scheme and linear combination of weighted average of observed anomalies i.e.

$$\sum_{j=1}^{n} w_j \left(f_j^{o} - f_j^{p}\right) / \sum_{j=1}^{n} w_j \qquad (4)$$

are added to the background field to get the analysed values in case of CSV. Fig.3 shows the CSV values against the distance for the three levels. It is observed that CSV values are slowly increasing and reach asymptotic values at large distances viz. 3245 kms for 850, 700 hPa levels and 3072 kms for 500 hPa level. This means there is no regional effect of one station on the other after these distances. These distances correspond to R (radius of influence). CSVs intercept the X (distance) axis approximately at 230 km. In other words this is the smallest distance between the stations. The largest CSV value at large distance occurs at 850-hPa level whereas the smallest CSV value occurs at 500-hPa level.

After extracting the maximum CSV and corresponding distance values from Fig.3 the graphs are made dimensionless by dividing the computed CSVs and corresponding distances by their maximum values. The dimensionless CSV values now vary between 0 to 1. The experimental CSV weights are obtained by subtracting the dimensionless CSV values at each distance from 1. These experimental CSV weighting functions for three

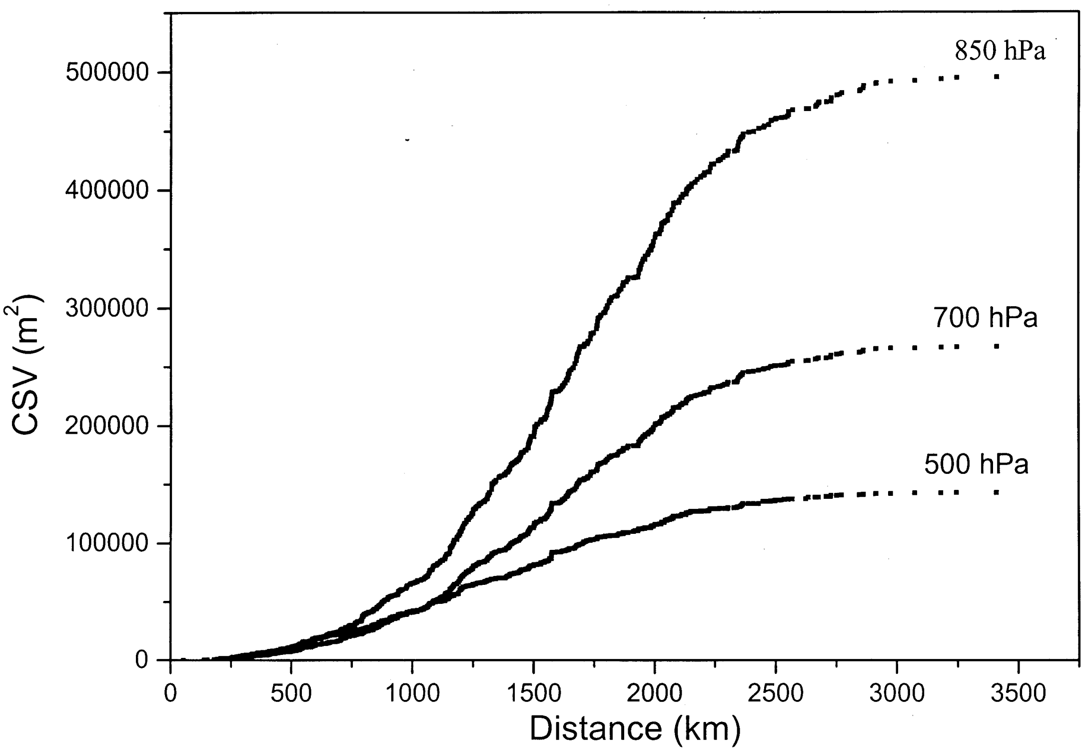

Figure 3. CSV against distance for 850, 700 and 500 hPa levels.

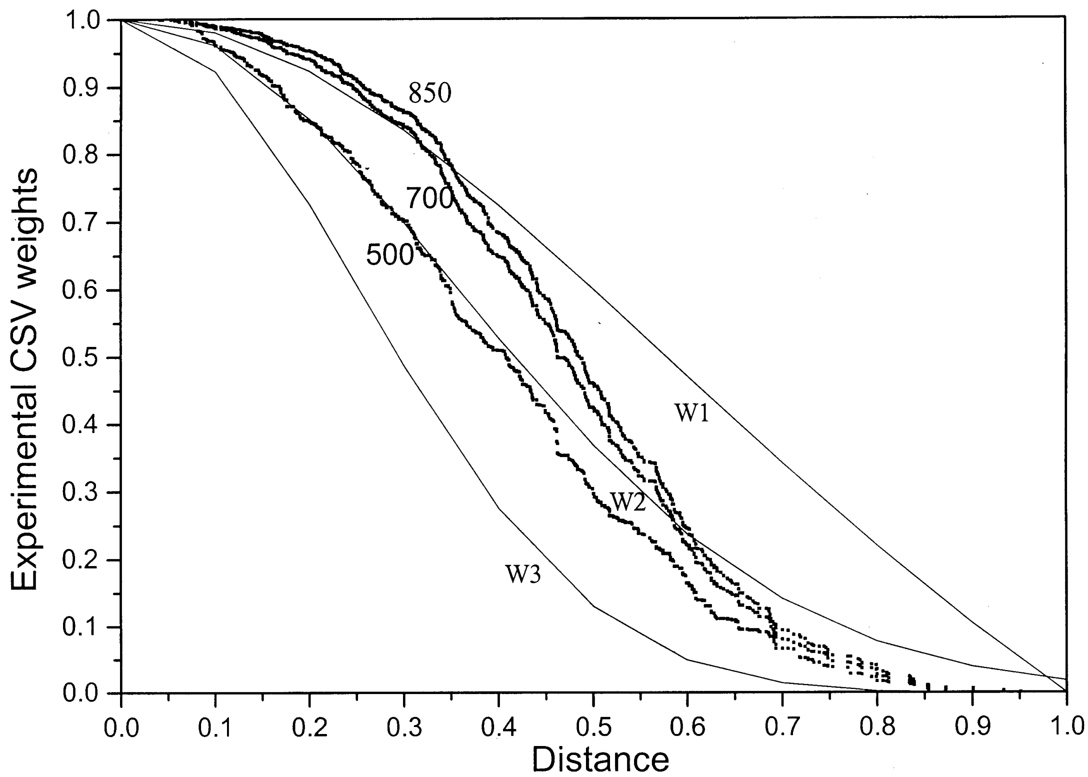

Figure 4. Experimental CSV weights for 850, 700 and 500 hPa levels.

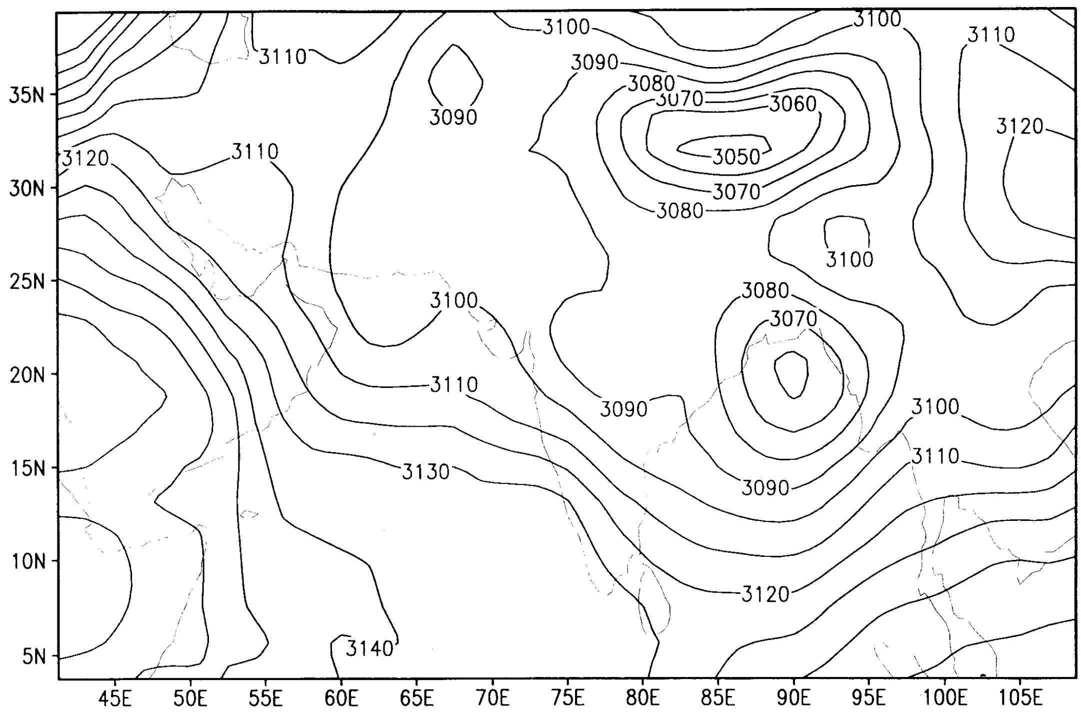

Figure 5. Objective analysis of 7 July 1979, 12 GMT of height field at 700 hPa level using CSV scheme.

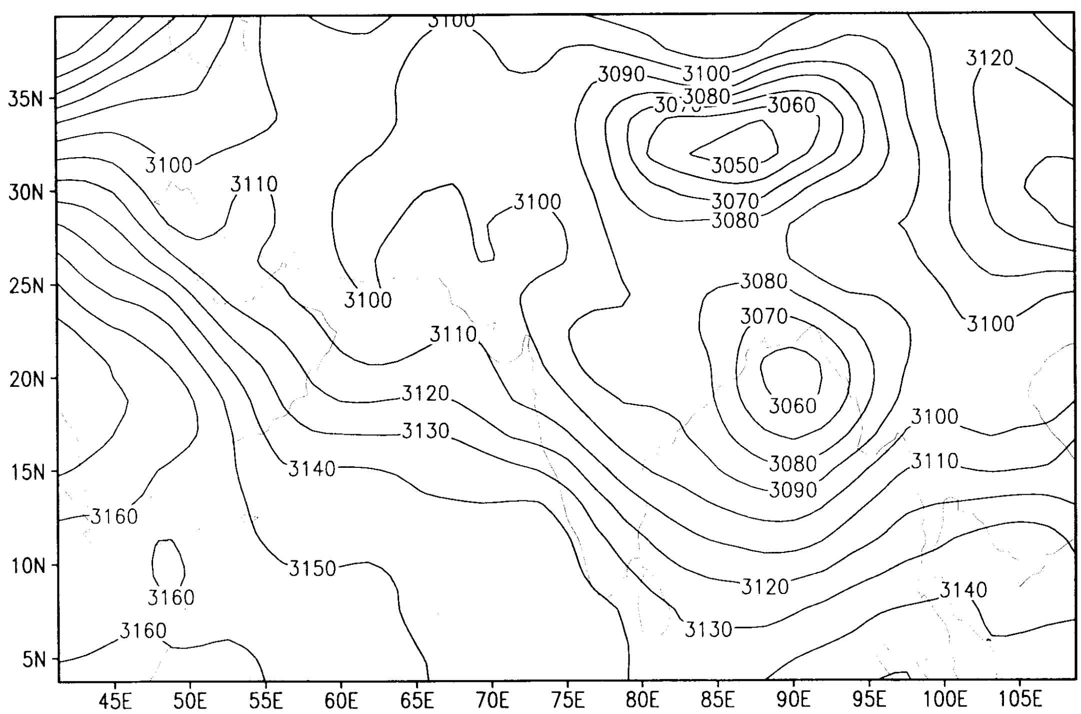

Figure 6. Same as Fig. 5 but for OI scheme.

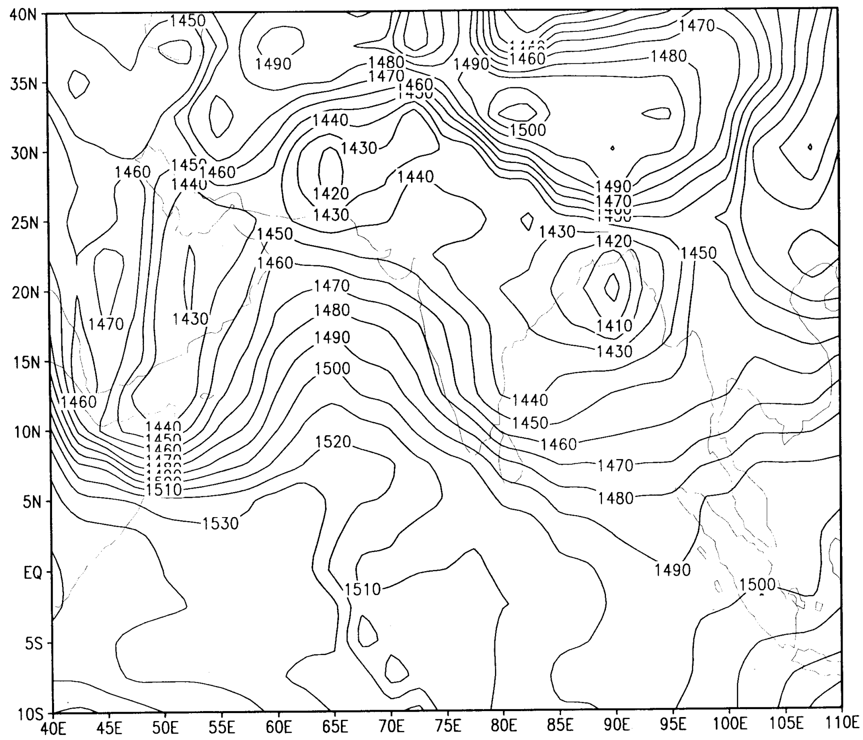

Figure 7. Same as Fig. 5 but for 29 July 1991 at 850 hPa level.

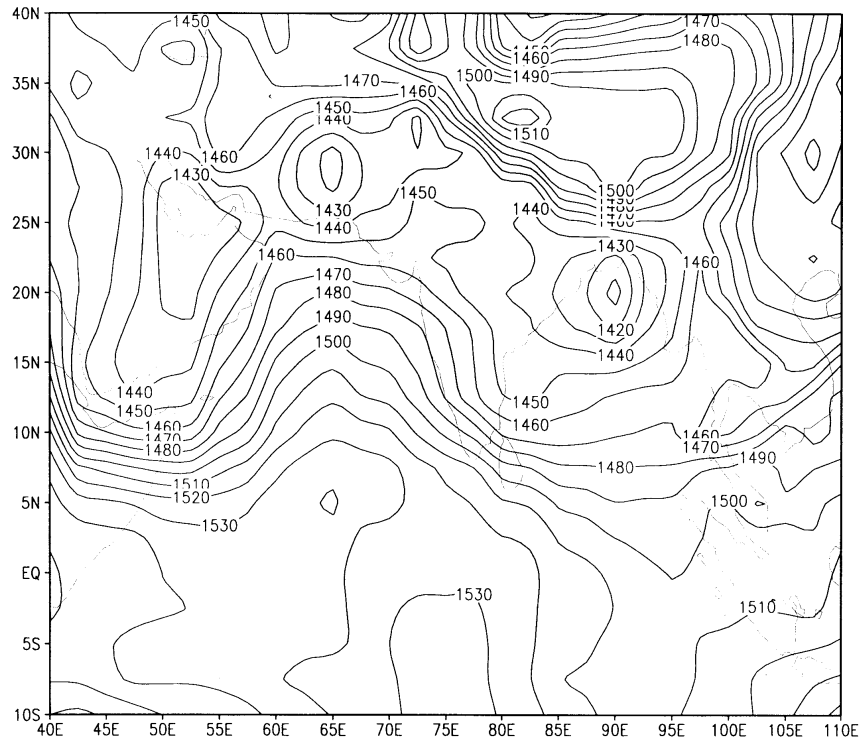

Figure 8. Same as Fig. 7 but for OI scheme.

levels are shown in Fig.4. In this figure the geometrical weighting functions, which are shown in Fig.2, are also drawn. At 850-hPa level the experimental CSV weighting function is closer to ratio model up to about 0.45 dimensionless distance and thereafter it is closer to exponential model. At 700 hPa level similar pattern is observed. At 500-hPa level it is slightly different as compared to other two levels. Initially CSV weighting function is closer to ratio model up to 0.08 dimensionless distance and thereafter it follows the exponential model up to 0.65 dimensionless distance and finally it is closer to power model. This shows that any single geometrical weighting model cannot be considered for the entire meteorological phenomenon. Once the weighting functions have been determined objective analyses for above period and for three levels have been carried out using Eq. 4. The analyses using CSV and OI are carried out for 850, 700 and 500 hPa levels for 4-8 July 1979 but we present here the analyses for only 700 hPa of 7 July 1979, Fig.5 and Fig.6 respectively. It is observed that the major features such as low pressure are well captured without difficulty in both the schemes. The lowest value surrounding the center of depression for 700 hPa level is 3060 m for CSV as well as for OI. Experiments using above two schemes (CSV and OI) are also carried out for the 26–30 July 1991 at 850, 700 and 500 hPa levels. However, we present here the analyses for the 29 July 1991, 850 hPa, Figs. 7 and 8. Centres of the depressions are well depicted in both the schemes in the second case also. OI scheme utilizes a gaussian function e.g. a.exp(-b.S_{ij}^2) to model the observed correlations, which are obtained from the above, mentioned 10 years of height data for July month. S_{ij} is the distance between the two locations i and j. The values of the constants a and b involved in the gaussian function and the values of random error (noise to signal ratio) for different levels are given in Tables 1 and 2. C_{ij} (correlation between two observations) and C_{xi} (correlation between grid point and observation point), which are required for the solution of Eq.1, are computed using this gaussian function. These systems of simultaneous equations are then solved to determine weights. It is observed that both the sets of analyses (CSV and OI) are in well agreement with each other for both the situations (1979 and 1991). Root mean square (RMS) errors compared with FGGE analysis for three levels and for all the days (1979) and for both the schemes are shown in Table 3. RMS errors for the 1991 case are given in Table 4. It is found that the RMS errors using CSV scheme are comparatively less than the OI scheme on most of the days for all the levels. RMS errors graph for only 850-hPa level is shown in Fig. 9. As already mentioned in earlier section, the CSV analyses do not

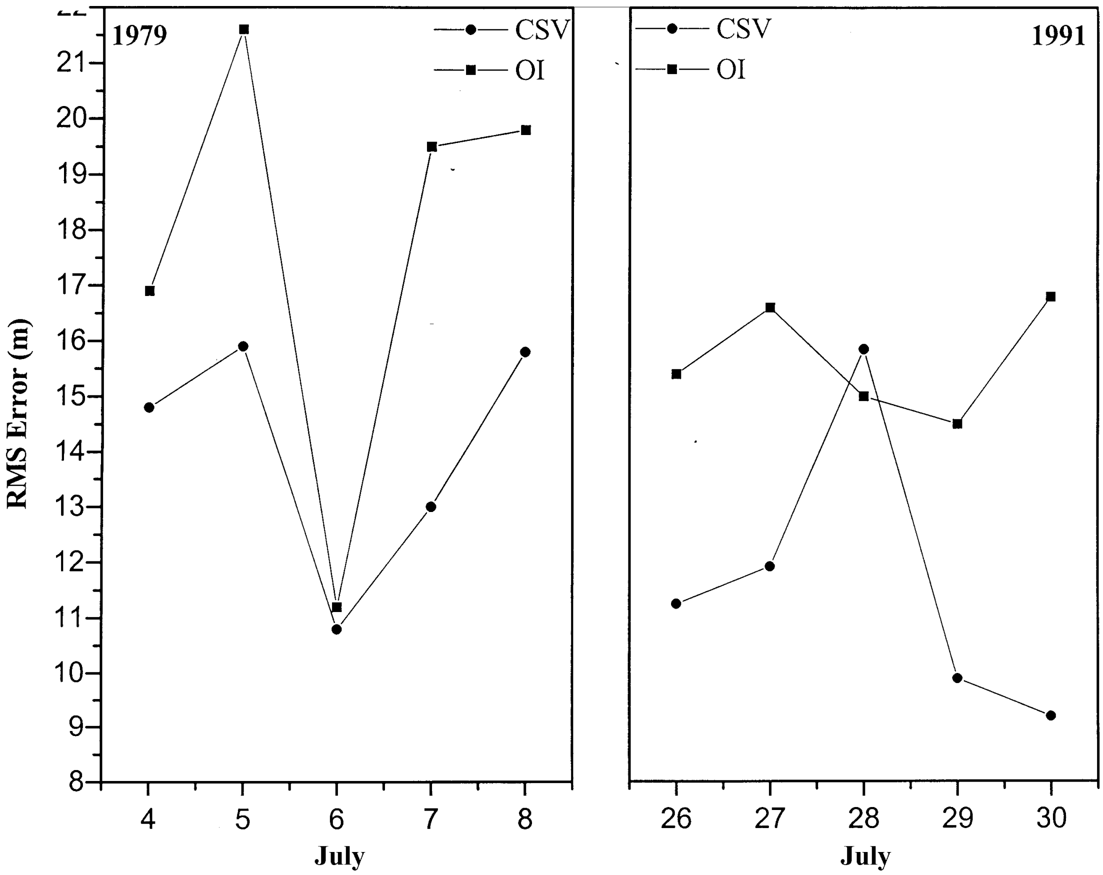

Figure 9. Graph showing Root Mean Square errors for both the schemes for different days for July 1979 and 1991 cases at 850 hPa level.

Table 1. Values of the constants for Gaussian Function: $a.exp(-b.S_{ij}^2)$

Levels	850 hPa	700 hPa	500 hPa
a	0.6939	0.4881	0.2372
b	0.0023	0.0027	0.0046

Table 2. Random Errors for different levels

Levels	850 hPa	700 hPa	500 hPa
λ^2 = noise/signal	0.1666	0.6313	2.3527

Table 3. Root Mean Square errors (meter) for different levels and days as compared with FGGE analysis

Days	850 hPa		700 hPa		500 hPa	
	CSV	OI	CSV	OI	CSV	OI
04.07.79	14.8	16.9	12.5	17.0	18.1	17.5
05.07.79	15.9	21.6	15.3	17.8	20.0	19.8
06.07.79	10.8	11.2	09.5	10.9	13.7	14.5
07.07.79	13.0	19.5	13.0	16.0	16.0	16.8
08.07.79	15.8	19.8	18.2	19.5	21.2	21.2

Table 4. Root Mean Square errors (meter) for different levels

Days	850 hPa		700 hPa		500 hPa	
	CSV	OI	CSV	OI	CSV	OI
26.07.91	11.3	15.4	16.7	14.0	16.9	14.5
27.07.91	11.9	16.6	15.5	14.0	13.1	15.0
28.07.91	15.9	15.0	11.8	14.2	13.9	13.3
29.07.91	09.9	14.5	12.5	12.6	15.6	14.2
30.07.91	09.2	16.8	12.8	14.3	13.5	15.6

require any matrix solution whereas in OI scheme at each grid point the matrix inversions are carried out to determine the weighting functions. Hence CPU time for CSV analysis is largely reduced as compared to OI. Although we have made analyses for three levels and five days in each situation (4-8 July 1979 and 26-30 July 1991) we have presented the results of 7 July 1979 and 29 July 1991 only as the system reached maximum intensity on these two days.

CONCLUDING REMARKS

1) In CSV scheme the weighting function and radius of influence are computed based on the data values at each site hence no subjectivity is involved in CSV scheme.
2) In CSV scheme there is no successive scans and grid point values are directly estimated from the surrounding observed values within the radius of influence.
3) Experimental CSV provides circular radius of influences only in the case of homogeneous meteorological phenomena, but irregular radii of influence in the case of heterogeneous phenomena. Whereas conventional objective analysis technique assumes circular radii of influence irrespective of whether the phenomenon is homogeneous or not.
4) To represent regional variability a combination of three geometrical models should be considered.
5) There are no matrix solutions in CSV technique, hence it requires a small computation time and as such it will be very much useful for operational forecast.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. G.B. Pant, Director, Indian Institute of Tropical Meteorology, Pune for his constant encouragement and Dr. S.S. Singh, Head, F.R.D. for his interest in this study. Thanks are also due to N.D.C., IMD, Pune for providing the data.

REFERENCES

Bergthorsson, P. & Döös, B.R., 1955. Numerical weather map analysis, Tellus, 7, 329-340.

Cressman, G.P., 1959. An operational objective analysis system, Mon. Wea. Rev., 87, 367-374.

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

Gandin, L.S., 1963. The objective analysis of meteorological fields, Israel program for Scientific Translations, Jerusalam, 242 pp.

Gustavsson, N., 1981. A review of methods of objective analysis, Data Assimilation Methods (Ed. L. Bengtsson, M. Ghil, E. Kallen), Springer-Verlag, 17-76.

Matheron, G., 1963. Principles of Geostatistics, Economic Geol., 58, 1246-1266.

Panofsky, H.A., 1949. Objective weather map analysis, J. Meteor., 6, 386-392.

Petersen, D.P. & Truskey, T.N., 1969. A study of objective analysis techniques for meteorological fields, Final Report, Contract No. EE-163 (69), University of New Mexico, Albuquerque, 155 pp.

Sen, Z., 1989. Cumulative semivariogram models of regionalised variables, Math. Geol., 21, 891-903.

Sen, Z., 1997. Objective analysis by Cumulative semivariogram techniques and its application in Turkey, J. Appl. Meteor., 36, 1712-1724.

(Accepted 2003 March 5. Received 2003 February 6; in original form 2002 September 5)