ON SOME ASPECTS OF OBJECTIVE ANALYSIS OF HUMIDITY OVER INDIAN REGION BY THE OPTIMUM INTERPOLATION METHOD

S. K. Sinha, D. R. Talwalkar and *S. Rajamani*

Indian Institute of Tropical Meteorology, Pune-411005, India

Received June 5, 1986

ABSTRACT

On the basis of a three year record of Radiosonde observations over Indian region, the autocorrelation function and structure function of the humidity mixing ratio (r) were computed for 850 hPa level. These are necessary for the construction of a suitable objective analysis scheme for humidity over Indian region using optimum interpolation method. The statistics were derived for the monsoon period (June through September) for 850 hPa level.

In order to model the humidity correlation for Indian region, two types of curves were fitted: (i) $\mu(\rho)=a\exp(-b\rho)$, (ii) $\mu(\rho)=A\exp(-B\rho^2)$ where μ is the autocorrelation function — a function of distance ρ—between two observing stations. It was found that the best description of humidity correlation function was given by Eq. (1). The value of 'a' gives a quantitative impression of the observation error. Further, the mean random errors were computed from structure functions, the weighting factors for the observing stations with respect to each grid point were calculated, and objective analyses were made for the humidity mixing ratio.

I. INTRODUCTION

Objective analysis procedures for obtaining the meteorological parameters at specified grid points have gained wide importance for diagnostic as well as prognostic studies. A large number of analysis methods differing in interpolation schemes, algorithms reliability and applicability to various regions are in vogue. Cressman (1959) evolved a successive correction method in which the anomaly of a meteorological variable at a given grid point from the initial guess field is expressed as the linear combination of those anomalies at the surrounding observing stations, giving weights to the inverse of distances of the stations from the grid points. In this scheme the weighting functions are empirically determined. However, in another scheme known as optimum interpolation (O. I.) scheme, the climatological characteristics of the given parameters are used in determining the weighting functions (Gandin, 1963). The principle of optimum interpolation is to interpolate the observing station data to the grid points using suitable weights derived from climatological characteristics with the condition of minimum interpolation error.

On these lines, for tropical regions, Alaka and Elvander (1972), Rajamani et al. (1983) have made studies respectively over Caribbean region and Indian region. Among the various meteorological parameters, the geopotential heights, winds and temperature were considered for detailed study in order to develop suitable schemes of objective analysis. Typically, consideration of the moisture analysis in N. W. P. has been afterthought. In part, this attitude is probably due to the difficulty of the task. Atmospheric moisture is extremely non-homogeneous, streaky in the horizontal and highly stratified in the vertical. There may be relatively large areas of saturation or near saturation with adjacent areas of very low moisture content. Such strong gradients impose severe strains on objective analysis methodology. However, for short-period predictions, especially when dynamic forcing is weak, the initial moisture specification may be very important indeed. In the present study, the development of a suitable scheme for the analysis of humidity has been considered. Among the various humidity parameters viz. relative humidity, mixing ratio, logarithm of mixing ratio, dew point and dew point depression, we have chosen the mixing ratio for our analysis purpose because it is found that the deviation of its distribution from normality is not large, that it is independent of temperature, and that it is the variable generally used in forecasting models, although for precipitation forecasts the relative humidity is more useful.

In the present study, a scheme for objective analysis of mixing ratio (r) following this optimum interpolation method is developed and analyses are made at 850 hPa level for two situations, 1 July 1969 and 15 July 1969. The assessment of this objectively analysed field is made by comparing it with the subjectively analysed field of humidity and then computing the RMS errors.

II. BASIC EQUATIONS

As regard to the optimum interpolation method, some basic relations are given below. Let $f_i, \hat{f}_i, \bar{f}_i$ and $\hat{f}'_i$ be the true value, observed value, the initial guess value and the anomaly which is the deviation of the observed value from the initial guess value at the station i, then we have the following relations:

$$\hat{f}'_i = \hat{f}_i - \bar{f}_i, \qquad \hat{f}_i = f_i + \varepsilon_i, \tag{1}$$

where ε_i is the total error in the observed value. The analysed value or the interpolated value, f_o, at the grid point 'O', is given by

$$f_o = \bar{f}_o + f'_o,$$

where the anomaly f'_o is given as a linear combination of weighted anomalies at the stations surrounding the grid point:

$$\text{i.e.} \qquad f'_o = \sum_{i=1}^{n} \hat{f}'_i P_i + I_0, \tag{2}$$

where P_i is the weighting factor, and I_0 is the error in the interpolation. In optimum interpolation method, the mean square error $E = \overline{(I_0^2)}$ is made minimum. That

$$E = \overline{I_0^2} = \overline{\left[f'_o - \sum_{i=1}^{n} (f'_i + \varepsilon_i) P_i \right]^2} \tag{3}$$

is made minimum. This is achieved by making

$$\frac{\partial E}{\partial P_i} = 0, \tag{4}$$

we make two important assumptions. Firstly the total error ε_i is independent of the anomaly of the true value,

$$\text{i.e.} \quad \overline{\varepsilon_i f'_i} = 0 \tag{5}$$

and also independent of the total error ε_j at other stations

$$\text{i.e.} \qquad \overline{\varepsilon_i e_j} = \begin{cases} 0 & \text{when} \quad i \neq j, \\ \sigma^2_{\varepsilon i} & \text{when} \quad i = j. \end{cases} \tag{6}$$

Secondly, the variances are homogeneous and the covariances are both homogeneous and isotropic. With these assumption the condition (4) yields a set of n equations with n unknowns $P_i (i=1,2,\cdots n)$.

$$\sum_{j=1}^{n} \mu_{ij} P_j + \lambda^2 P_i = \mu_{0i} \qquad (i=1,2,\cdots n) \tag{7}$$

where

$$\mu_{ij} = \frac{\overline{f'_i f'_j}}{\sigma_0^2}, \qquad \mu_{0i} = \frac{\overline{f'_i f'_0}}{\sigma_0^2}$$

and

$$\lambda^2 = \frac{\sigma^2_{\varepsilon i}}{\sigma_0^2}, \tag{8}$$

μ_{ij} is the autocorrelation function for a pair of stations i and j, and μ_{0i} is the autocorrelation between the grid point 'O' and the station 'i', $\sigma^2_{\varepsilon i}$ is random error and σ_0^2 is the average variance for the region.

The random error $\sigma^2_{\varepsilon i}$ required for calculation of λ^2 is obtained from structure function. The true structure function

$$\beta(\rho) = \overline{(f'_i - f'_j)^2} \tag{9}$$

and the estimated structure function

$$\hat{\beta}(\rho) = \overline{(\hat{f}'_i - \hat{f}'_j)^2} \tag{10}$$

are related as follows (Alaka and Elvander, 1972):

$$\hat{\beta}(\rho) = \beta(\rho) + 2\sigma^2_{\varepsilon i}. \tag{11}$$

$\beta(\rho)$ becomes zero when ρ becomes zero but $\hat{\beta}(\rho)$ need not become zero at the same station. So when ρ becomes zero,

$$\hat{\beta}(0) = 2\sigma^2_{\varepsilon i}. \tag{12}$$

In other words, $2\sigma^2_{\varepsilon i}$ is estimated by fitting a curve to the computed structure function, $\hat{\beta}(\rho)$ plotted against distance, ρ, and extrapolating the curve until it intersects the axis of $\hat{\beta}(\rho)$ at $\rho = 0$.

Eq. (7) is solved by conjugate gradient method (Beckman, 1960) to obtain the values of P_i, which are in turn substituted in Eq. (2) to get the anomaly at a particular grid point (Rajamani et al., 1983).

III. DATA

For this study, daily 00z, 850 hPa dew point temperatures of three years (1968, 1969 and 1971) from radiosonde stations over Indian region for the monsoon period (June through September) were utilized. From these dew point temperatures, the mixing ratios were computed.

IV. COMPUTATIONS AND DISCUSSION OF RESULTS

1. *Structure Functions*

Using Eq. (10), the structure function $\hat{\beta}(\rho)$ for mixing ratio was computed for 850 hPa level and plotted against distance ρ (Fig. 1). The mean random error $\sigma^2_{\varepsilon_i}$ at 850 hPa level for mixing ratio was obtained from Eq. (12) by extrapolating the structure function curve to zero distance. This random error is given in Table 1.

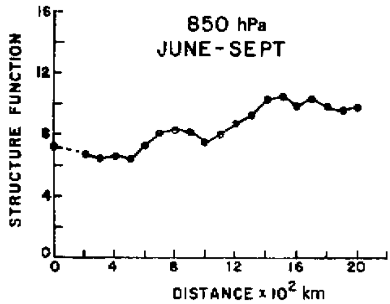


Fig. 1. Structure function of mixing ratio at 850 hPa over Indian region.

Table 1. Estimate of Random Error

Variable	Period	Level (hPa)	Extrapolated value of structure function at distance $\rho=0$	$\sigma^2_{\varepsilon_i}$	Random error (g/kg)
Mixing ratio(r)	June—Sept.	850	7.3	3.65	1.91

2. *Autocorrelation Function*

The covariances $\overline{f'_i f'_j}$ were computed for every station with respect to every other station in the region. They were normalized by dividing with σ^2_0 before plotting them against distance separating the two given stations. The computed autocorrelations, plotted against distance ρ, showed a certain amount of scatter, partly as a result of anisotropy and non-homogeneity of the true autocorrelations. So this scatter of points was divided into 4 segments with middle points located at a distance $d=$ 2, 3, 4, ... 20 degrees, and then we determined the values of $\mu(d)$ in each interval by means of the following equation:

$$\mu(d)=\sum_{i=1}^{N_T}\mu(\rho_i)\frac{N_i\left[\frac{1}{2}+\frac{1}{2}\cos\frac{\pi(\rho_i-d)}{100}\right]}{\sum_{j=1}^{N_T}N_j\left[\frac{1}{2}+\frac{1}{2}\cos\frac{\pi(\rho_i-d)}{100}\right]}, \tag{13}$$

where N_T is the total number of covariance values contributing to μ, and N_i is the number of points within 4° segment.

Two empirical curves represented by Eqs. (14) and (15) were considered for fitting the autocorrelation values obtained earlier using Eq. (13).

$$\mu(\rho)=a\exp(-b\rho), \tag{14}$$

$$\mu(\rho)=A\exp(-B\rho^2), \tag{15}$$

where a,b,A and B are constants. The values of the constants a,b,A and B were evaluated and are given in Table 2. These forms were selected because they insure that the fitted autocorrelation functions are positive definite. Of course, this is difficult to test, but the conditions may be replaced (essentially) by the request that the two-dimensional Fourier transforms (Hankel transforms) must not be negative for any value of the wave number. The Eq. (14) was used for further computations because it fits the data better than the function of Eq. (15). In other words, the sum of the squares of differences between observed and fitted ordinates is minimum. Fig. 2 shows the autocorrelation function thus obtained.

Table 2. Values of Constants of the Empirical Curves Fitted for Autocorrelation Functions

Functions	Period	Level (hPa)	Constants a	b	A	B
$\mu=a\exp(-b\rho)$	Jun.—Sept.	850	0.492	0.166		
$\mu=A\exp(-B\rho^2)$	Jun.—Sept.	850			0.248	0.008

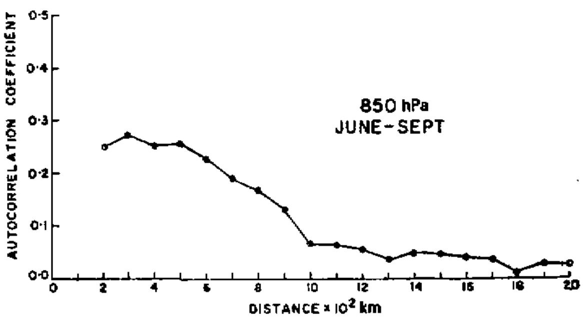


Fig. 2. Autocorrelation function of mixing ratio at 850 hPa over Indian region.

The existence of errors of the measuring instrument makes the value of a less than one, but still the observed correlations would exactly fit the function $a\exp(-b\rho)$. The deviation of a from one is not only caused by the effect of instrumental errors, but also by the combined effect of instrumental errors and the deviations of the correlation behaviour at a small distance between the true atmospheric field and the autocorrelation model (Maanen, 1981). ρ is a function of four variables (x_i, y_i), (x_j, y_j). The inconsistency in using Eq. (14) is that the humidity field is differentiable with respect to x, whereas the autocorrelation $\mu(\rho)$ is not differentiable to x_i at any point. The problem arises during the analysis of height field. Since the humidity gradient is not involved in humidity analysis, the consideration of fitting with the observed correlations was given preference.

3. *Weighting Functions and Objective Analysis*

For our experiment we have chosen the region from 8° to 30°N and from 70°E to 94°E as the region of our analysis. The pattern of the autocorrelation of all stations with Nagpur as central station (Fig. 3) suggests that the autocorrelation is nearly circular and thus direction is independent. Therefore, we can assume that the correlation function is isotropic. We also assume that the correlation function is homogeneous; hence, autocorrelation function $\mu(\rho)$ for any given distance, ρ, could be computed by using the functional relation (14), knowing the values of the constants a and b. Thus, using Eq. (14), μ_{ij} (for the pair of stations i and j) and μ_{oi} (for the station i and the grid point 'O') were obtained. σ^2_{ri} was computed from structure function curve in Fig. 1 as mentioned earlier, and subsequently λ by using Eq. (8). Thus a set of equations Eq. (7) was formed and solved by the conjugate gradient method to obtain the weighting functions P_i's of the stations ($i = 1, 2, \ldots n$) with respect to each of the grid points in the region considered. The anomaly at each of the grid points was computed as a linear combination of the weighted anomalies at the surrounding stations and, by adding it to the initial guess at that grid point, the analysed value was obtained. The grid size of 2° was considered for our analysis. The analyses were made for 1 July and 10 to 17 July 1969 for 850 hPa level.

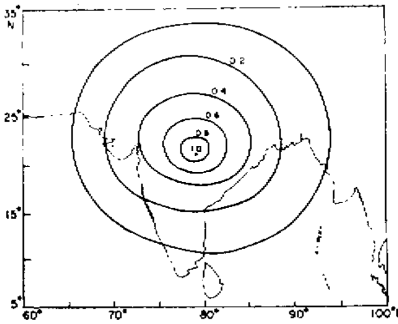
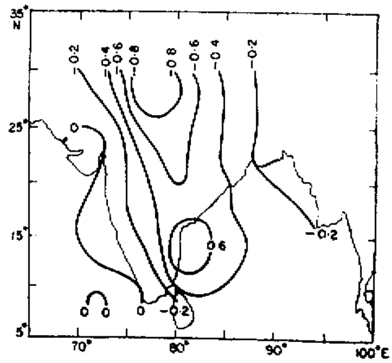


Fig. 3. Autocorrelation of mixing ratio at 850 hPa with respect to Nagpur.

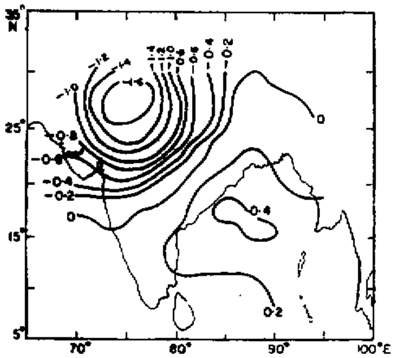
Experiments were conducted for five different cases. Initially, all observations which were within 30° from the grid points were used for determining the weighting functions. In the subsequent cases, observing stations within 12.5°, 10°, 7.5° and 5° from the grid point were taken into consideration for determining the weighting functions, and subsequently the anomalies at the grid points. The computed anomalies at the grid points for the two dates 1 July and 15 July 1969 are given in Fig. 4 (a—e) and 5 (a—e) respectively. Fig. 6 (a) and Fig. 6 (b) depict the objective and subjective analyses of mixing ratio for the dates 1 July and 15 July 1969. For examining the objectively analysed field quantitatively, RMS errors were computed by comparing the objective analyses with the subjective analyses for nine days. The RMS errors for different scan lengths for 1 July and 15 July 1969 are shown in Fig. 7 and, for 1 July and 10 July to 17 July 1969 are given in Table 3.

4. *Analysis of* 1 *July and* 15 *July* 1969

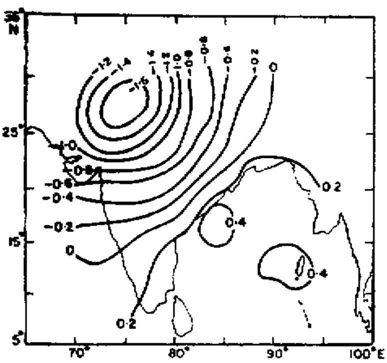
The analyses were made, considering stations within five different scan lengths viz. 30°, 12.5°, 10°, 7.5° and 5°, in order to determine the scan length which would give better analysis, and this would depend on the wavelength of the system which frequents the region. On examination of the analyses of moisture field (mixing ratio) made for 1 July 1969 for different scan lengths, it is seen in general that the computed anomalies at the grid points are rather small and consequently the analysed field is similar to the initial guess field. The RMS errors computed for the different cases show that the RMS error is minimum for 10°, the value being 2.64 g/kg.



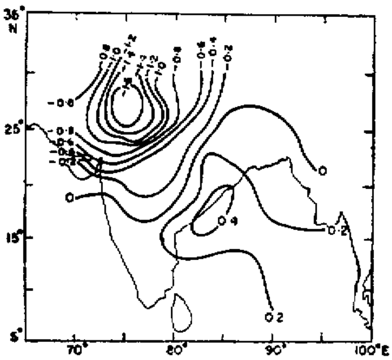
(a)



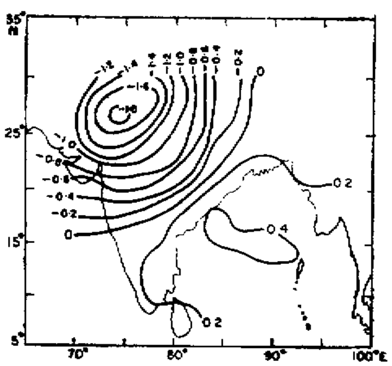
(d)



(b)



(e)



(c)

Fig. 4. (a) Anomalies at grid points computed using weights of the O. I. scheme for July, 1969 for scan length 30°;
(b) Same as (a) for scan length 12.5°;
(c) Same as (a) for scan length 10.0°;
(d) Same as (a) for scan length 7.5°;
(e) Sameas (a) for scan length 5.0°.

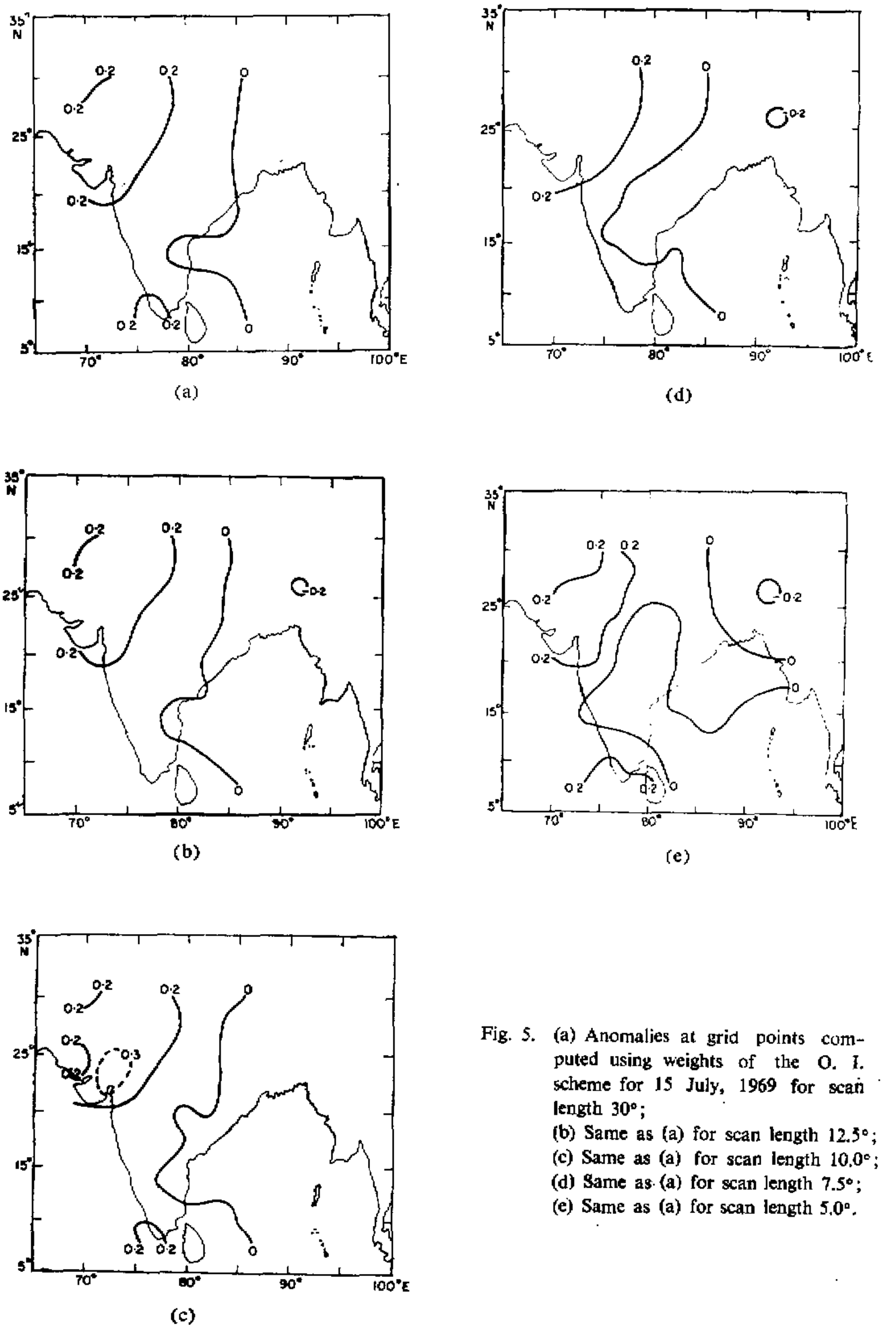
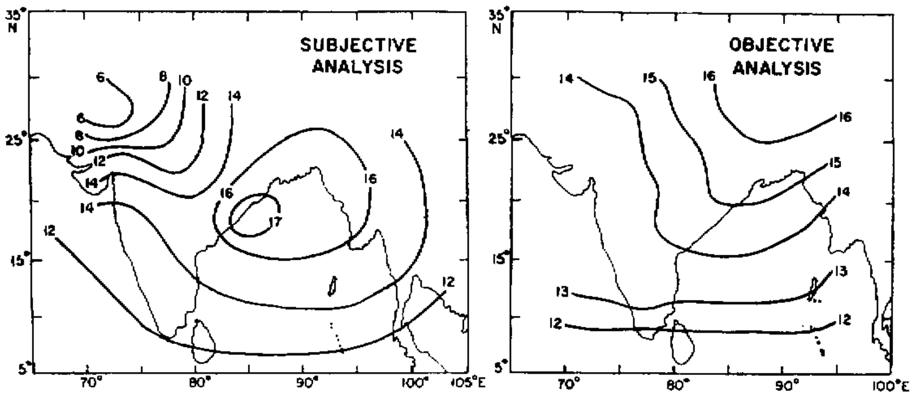
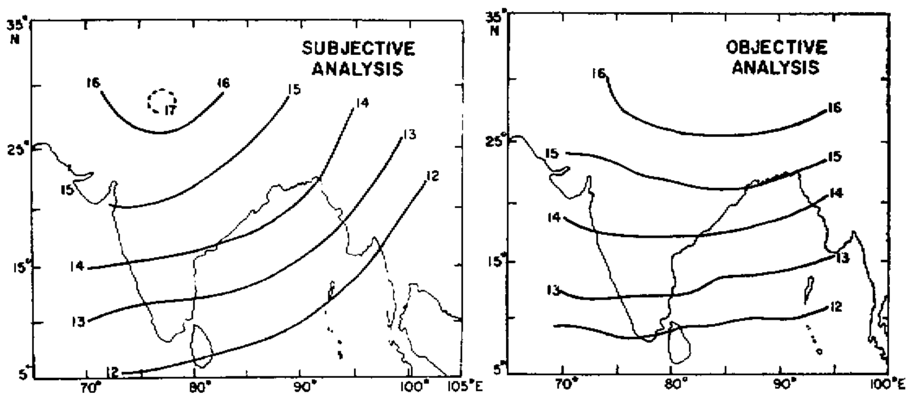


Fig. 5. (a) Anomalies at grid points computed using weights of the O. I. scheme for 15 July, 1969 for scan length 30°;
(b) Same as (a) for scan length 12.5°;
(c) Same as (a) for scan length 10.0°;
(d) Same as (a) for scan length 7.5°;
(e) Same as (a) for scan length 5.0°.



(a)



(b)

Fig. 6. (a) Subjective analysis and objective analysis of mixing ratio by the O. I. scheme for 1 July, 1969;
(b) Same as (a) for 15 July, 1969.

Table 3. RMS Errors Obtained by Comparing the Observed Humidity with the Objectively Analysed Humidity

Scan Length	RMS Errors in g/kg 850 hPa Level									Average RMS Errors for 10 July to 17 July, 1969
	1.7.69	10.7.69	11.7.69	12.7.69	13.7.69	14.7.69	15.7.69	16.7.69	17.7.69	
30.0°	3.3221	0.9857	1.0400	1.2085	1.5162	1.0431	0.7174	0.9819	0.7861	1.0349
12.5°	2.6516	0.9859	1.0475	1.2198	1.5086	1.0321	0.7082	0.9758	0.7967	1.0343
10.0°	2.6480	0.9809	1.0530	1.2252	1.4991	1.0252	0.7119	0.9703	0.7974	1.0337
7.5°	2.6841	0.9741	1.0688	1.2359	1.4875	1.0289	0.7148	0.9714	0.8033	1.0356
5.0°	2.7918	0.9766	1.0976	1.2618	1.4897	1.0584	0.7480	0.9877	0.8135	1.0542

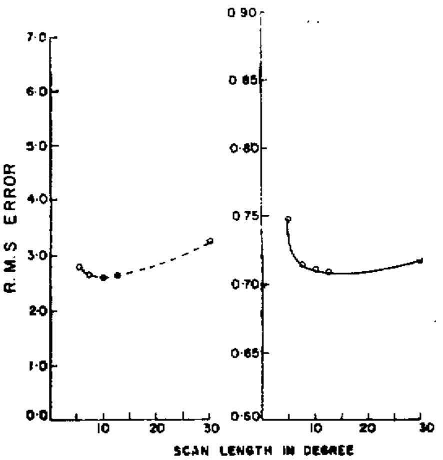


Fig. 7. Root Mean Square Errors in g/kg in the analyses for 1 July, 1969 (dashed curve) and 15 July, 1969 (full curve).

Fig. 4(a) shows that, when the scan length is 30°, the computed anomalies at all grid points are negative on 1 July 1969. This has been caused by larger negative anomalies of the observing stations in the north-western region which is dry due to the non-arrival of the monsoon current by 1 July 1969. However, by 15 July 1969 when the monsoon has well established over the northwest region of India and the synoptic conditions would be closer to the mean July conditions, the anomalies were not large. This brings out the fact that the closer the initial guess to the actual conditions, the better would be the analysis.

When the scan length is 12.5° or less, the stations having large negative anomalies (for the date 1 July 1969) affect nearby grid points, and so the corrections are negative only in those regions. In the case of the analysis for 15 July 1969, the station anomalies are smaller and hence the variation in the scan length does not affect the analyses so much and the RMS errors do not vary much, as can be seen from Fig. 7. As a result, the analysis obtained for different scan lengths are similar. The small values in RMS errors compared to the case of 1 July indicate that the analyses are closer to the subjective analysis for the case of 10 July to 17 July 1969. The average RMS errors for 10 July to 17 July 1969 for different scan lengths are given in Table 3. It is observed that the minimum average RMS error is 1.0337 for the 10° scan length.

V. CONCLUSIONS

The model curve for the autocorrelation function was determined. The distribution of autocorrelation function with respect to a central station Nagpur suggests that the autocorrelation function decreases uniformly in all directions or, in other words, it is more or less isotropic.

The experiment made with different scan lengths shows that only those observing stations within certain distance should be used for analysis and this is particularly true when the synoptic situation is rather changing than well set as in the case of 15 July 1969.

The autocorrelation function and structure function of the humidity mixing ratio (r)

were computed on the basis of a three-year record of radiosonde observations over Indian region and the value of the random error obtained was 1.91 g/kg. On further increase of data from three to five years, it is found that the value of the random error is 1.95 and it is also found that there is no appreciable change in the value of the autocorrelation function with the addition of two years of data. However, in order to see that the values of the autocorrelation function and structure function obtained are stable irrespective of further increase of data, the computations have to be made with ten-years data and examined.

The authors are thankful to Shri D. R. Sikka and Dr. S. S. Singh, Indian Institute of Tropical Meteorology, Pune for their interest in this study. They also wish to thank Shri A. S. Gade for the preparation of the diagrams and Smt. C. Bardhan for typing the manuscript.

REFERENCES

Alaka, M. A. and Elvander, R. C. (1972), Optimum interpolation from observations of mixed quality, *Mon. Weath. Rev.*, **100**: 612—624.

Alaka, M. A. and Elvander, R. C. (1972), Matching of observational accuracy and sampling resolution in meteorological data acquisition experiments, *J. Appl. Met.*, **11**: 567—577.

Beckman, F. S. (1960), Mathematical Methods for Digital Computers, Vol. 1, A. Ralston and H. S. Wilf, Eds., Wiley, 62—72.

Cressman, G. P. (1959), An operational objective analysis system, *Mon. Weath. Rev.*, **87**: 367—374.

Gandin, L. S. (1963), Objective analysis of meteorological fields, Gidrometeorologicheskoe Izdatel'stvo, Leningrad, USSR, 286.

Maanen, J. V. (1981), Objective analysis of humidity by the optimum interpolation method. *Tellus*, **33**: 113—122.

Rajamani, S. et al. (1983), Objective analysis of wind field over Indian region by optimum interpolation method, *Mausam*, **34**: 43—50.