

## Estimation of upper air relative humidity for objective analysis

P. L. KULKARNI, S. K. SINHA, D. R. TALWALKAR,  
S.G. NARKHEDKAR, S. NAIR and S. RAJAMANI

*Indian Institute of Tropical Meteorology, Pune*

*(Received 22 August 1989)*

सार—ऊपरी वायु सापेक्ष आर्द्रता और तत्सम्बन्धित सतह प्रेक्षण, जैसे मेघ आवरण, वर्तमान मौसम और सतह सापेक्ष आर्द्रता आदि के बीच एक बहु-समाश्रयण सम्बन्ध का विकास किया गया। कई सतह प्रेक्षण केन्द्रों में इस समाश्रयण सम्बन्ध का उपयोग करके, सतह सापेक्ष आर्द्रता, मेघ आवरण और वर्तमान मौसम को 4 से 9 जुलाई 1979 के छः दिनों के लिए 700 एच पी ए पर सापेक्ष आर्द्रता मूल्यों का आकलन किया गया। इन आकलित मूल्यों के संघात को निर्वारित करने के लिए, अधिकतम अन्तर्वेशन विधि का उपयोग करके विश्लेषणों के दो समूह बनाए गए, एक में आकलित आर एच मूल्य सम्मिलित थे और दूसरा मूल्यों को हटाकर था। आकलित आंकड़ों के साथ और बिना आकलित आंकड़ों के, विश्लेषण के दोनों समूह, प्रेक्षणों से मेल जुड़ते हैं यह परीक्षण करने के लिए सभी छः मामलों के लिए प्रेषित रेडियोसोण्डे आंकड़ों के साथ तुलना करते हुए आर एम एस त्रुटियों को संगणित किया गया। ई सी एम डब्ल्यू एफ—आर एच विश्लेषणों के साथ तुलना करते हुए आर एम एस त्रुटियों को भी संगणित किया गया। अन्त में, इन विश्लेषणों की उपग्रह मेघ चित्रों के साथ तुलना की गई और प्रेषित आंकड़ों के साथ निकट समझौता प्राप्त होता है।

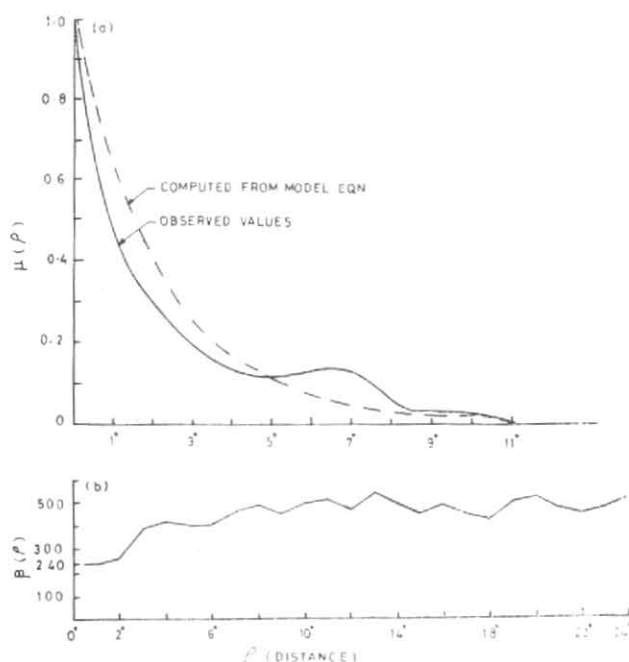
**ABSTRACT.** A multiple regression relation was developed between upper air relative humidity and the related surface observations, like cloud cover, present weather and surface relative humidity etc. Using this regression relation at a number of surface observing stations, the relative humidity values for upper air at 700 hPa were estimated from surface relative humidity, cloud cover and present weather for six days from 4 to 9 July 1979. In order to determine the impact of these estimated values, two sets of analyses were made using optimum interpolation scheme, one including estimated RH values and the second excluding them. RMS errors were calculated for all the six cases, comparing with the observed radiosonde data to examine how both sets of analyses with and without estimated data, fitted the observations. RMS errors were also computed comparing with ECMWF—RH Analyses. Lastly, these analyses were compared with satellite cloud pictures and found them in close agreement with the observed data.

**Key Words**—Relative humidity, Multiple regression, Objective analysis, Optimum interpolation, Monsoon depression

### 1. Introduction

At present, forecast models are able to produce fairly accurate forecasts of various parameters including precipitation. In order to make precipitation forecasts, it is essential to specify the initial humidity field reasonably accurately. Also the humidity field influences the radiative processes of the model. However, the specification of the initial humidity field or in other words the analysis of humidity field at the grid points from the observing

stations has unique problems. It is more so over the Indian region because of the surrounding data sparse regions, like the Arabian Sea, the Bay of Bengal and the Indian ocean. Since some surface observations like cloud cover, present weather, surface humidity etc. are statistically related to upper air humidity, there have been a number of studies made to estimate the humidity at higher levels from surface observations in order to overcome the inadequate horizontal resolution of radiosonde network.



Figs. 1(a-b). (a) Autocorrelation function and (b) Structure function of relative humidity at 700 hPa

In this study, the technique of multiple linear regression is used to estimate the humidity at 700 hPa level to supplement the radiosonde observations with a view to improve the analysis. Since the relative humidity gives immediately the wet/dry conditions of air, this parameter is considered here.

## 2. Earlier work on estimation of relative humidity at higher levels from surface observations

Cressman (1959), Chisholm *et al.* (1968), Ball and Veigas (1968) estimated upper level dew point depression from surface reports and used them along with the radiosonde data for the analysis. Atkins (1970, 1974) in the study of analysis of relative humidity using anisotropic weighting functions, discussed the use of surface observations for the analysis at higher levels and concluded that surface mixing ratio fares better than the surface relative humidity. Using similar method, Jonas (1976) found that the relative humidity obtained from the surface observations were better than forecast values of British ten-level model.

In a similar pursuit on the estimation of the upper level relative humidity, Kastner (1974) has used vertical velocity obtained from quasi-geostrophic omega in a regression equation to obtain the first guess of relative humidity and subsequently its estimation at higher levels from surface observations of cloudiness, cloud type, present and past weather reports etc.

In the analysis technique used in the NMC operational model, Chu and Parrish (1977), radiosonde data, surface observations and bogus observations from satellite cloud photograph are used in the analysis of relative humidity.

Investigations of Rasmussen (1982) relating to the objective analysis of humidity for regional scale numerical weather prediction showed that the relative humidity

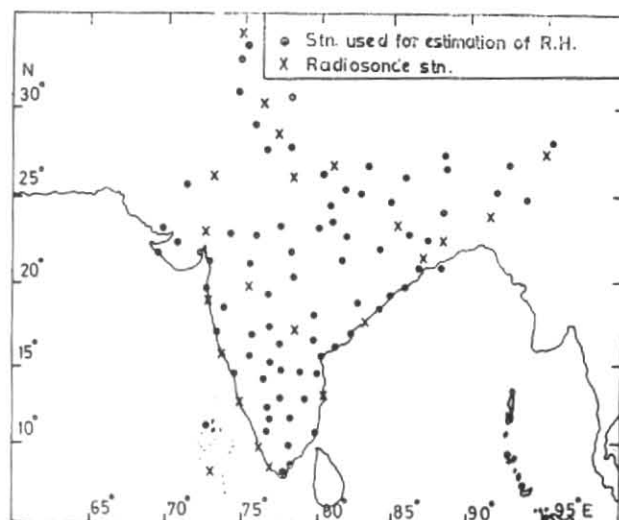


Fig. 2. Map showing radiosonde stations and stations where relative humidity were estimated

is a better analysis variable than the dew point depression and that the use of surface data and the cloud analysis data improve the analysis of upper air relative humidity. Vertical profiles of relative humidity from surface observations were estimated by Chu and Parrish (1977), Tibaldi (1982) and Illari (1985). Illari (1985) has studied the satellite observations of water vapour and concluded that, (i) the relative humidities inferred from the surface weather observations were too moist when compared with radiosonde data especially at lower levels and in partly cloudy conditions and (ii) the surface observations in areas of total or near total cloudiness could be used to compliment the satellite sensed humidities which are less reliable in the presence of overcast or near overcast conditions.

Dimego (1985) used satellite cloud imagery as supplementary information in several operational data assimilation experiments. Norquist (1986, 1988) used estimates of relative humidity from three different sources—satellite water vapour retrievals, surface weather observations and operational global cloud analysis and compared them with relative humidity from radiosonde reports. Surface observations and global cloud analysis found to be in sufficiently good agreement with radiosonde measurements and they have beneficial impact on analysis. Timchalk (1986) has shown that satellite bogus information can be used in regional analysis and forecasting system and that it compares well with conventional data to justify its use when no other data are available.

Begum *et al.* (1987) have used surface observations in the analysis of relative humidity over Indian region by relating the surface observations to one of the thirty five vertical profiles of the relative humidity and shown their positive impact on the analysis.

(a) 5 JULY, 1979

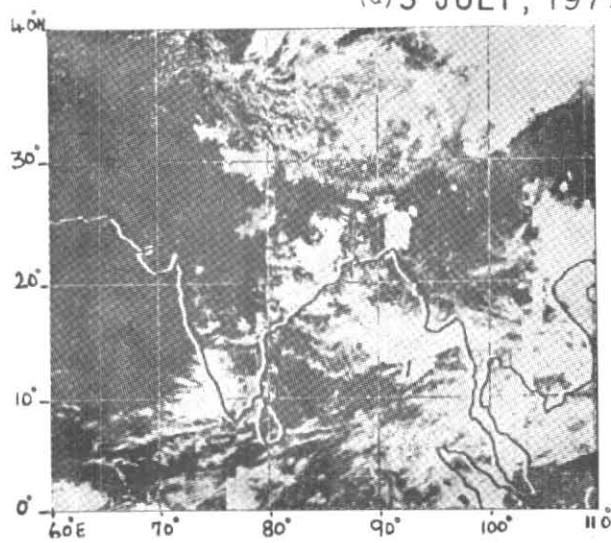


Fig. 3(a). Satellite cloud picture of 5 July 1979

(b) 7 JULY, 1979

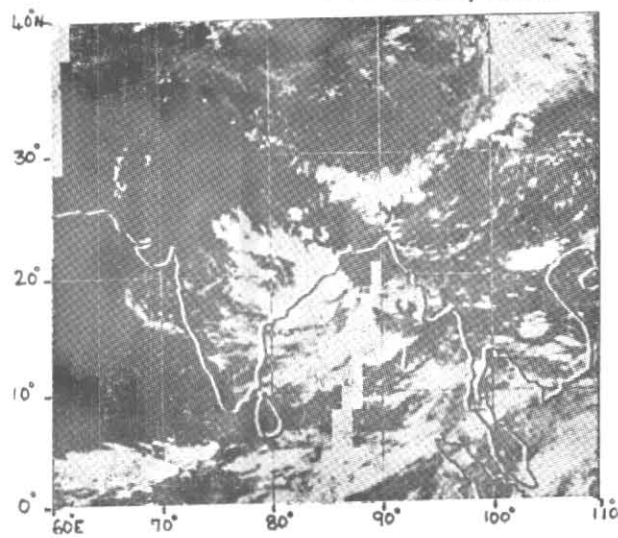


Fig. 3(b). Satellite cloud picture of 7 July 1979

(c) 9 JULY, 1979

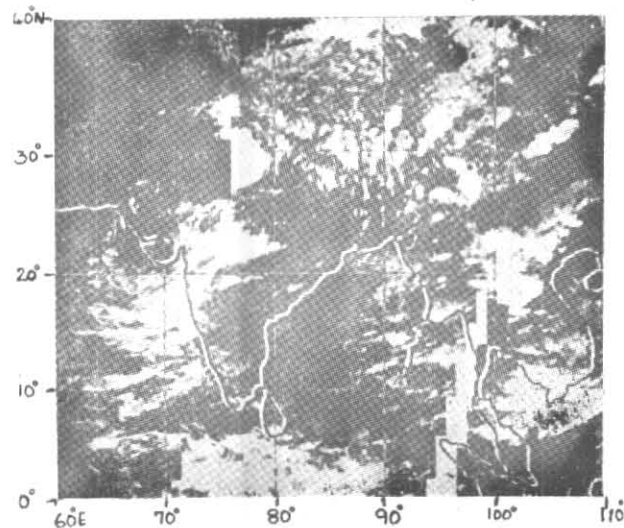
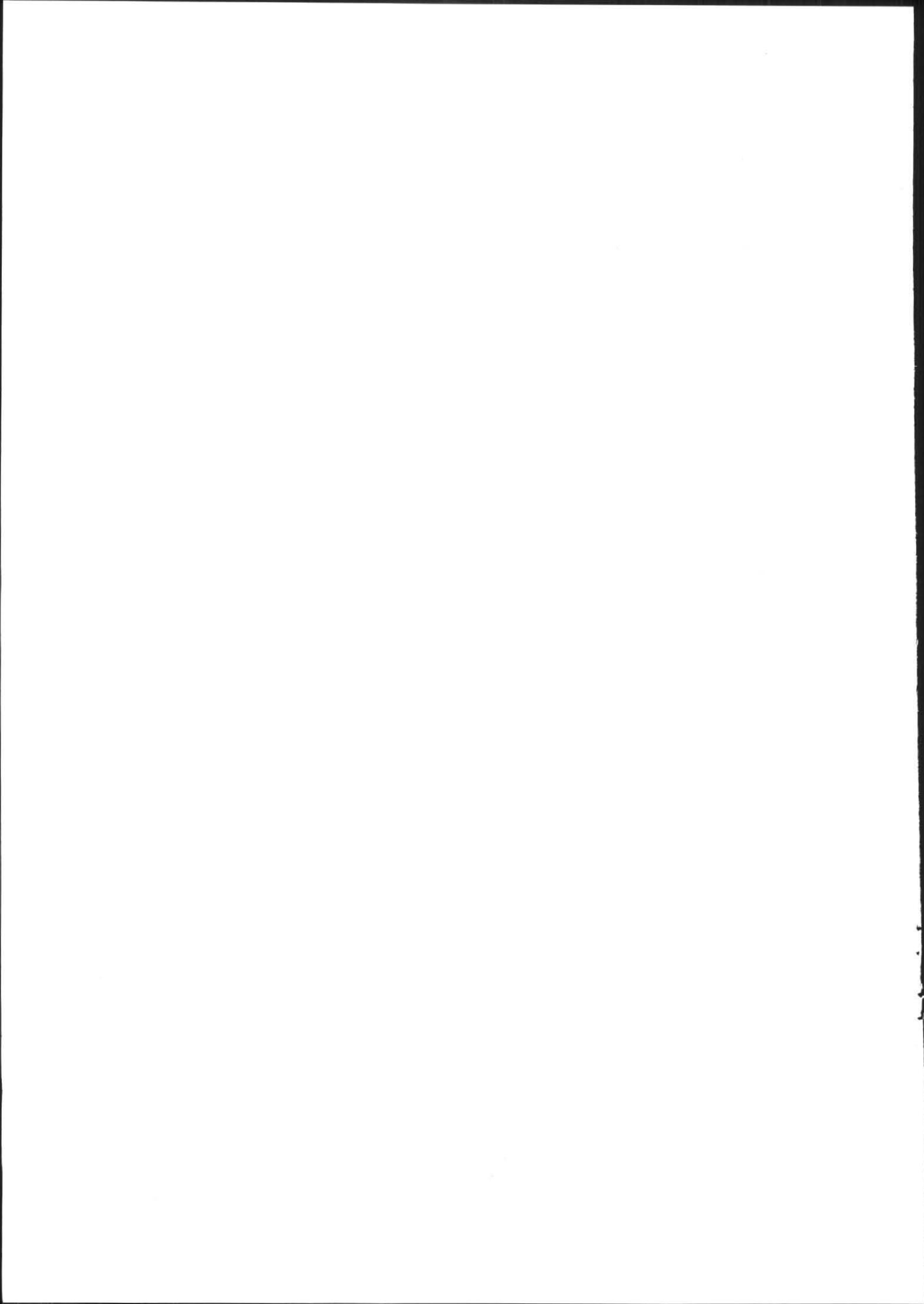


Fig. 3(c). Satellite cloud picture of 9 July 1979



In the present study, an attempt is made to use the surface observations in the analysis of upper air relative humidity especially over radiosonde sparse regions where these surface observations are available. Also this scheme could be made use as well over oceanic regions where ship observations of surface conditions would be available.

### 3. Methodology

The main objective of this study is to improve the initial specification of humidity field for the regional forecast model, in other words, to improve the analysis by developing a scheme in which it is possible to estimate relative humidity whenever radiosonde data are not available and add to the available radiosonde data for the analysis. Since, the upper air relative humidity is statistically related to the surface observations like relative humidity, cloud cover, etc, many studies have been made, as mentioned in earlier section, to estimate the upper air relative humidity from surface data by developing regression equation. In this study, the regression equation to estimate at 700 hPa level the relative humidity from some related surface observations was developed. These estimated relative humidity data were used to augment the radiosonde data and analyses were made.

In order to study the impact of the estimated data, analyses were also made without including estimated data and compared with earlier analyses for six cases by computing R.M.S. errors. Comparisons of these analyses with the cloud pictures were also made to examine which of these analyses depict the convective activity better.

#### 3.1. Optimum interpolation scheme of objective analysis

The analysis scheme used in this study is based on the optimum interpolation method first developed by Gandin (1963). The details of the scheme and the computational method are described in detail by Rajamani *et al.* (1983) for wind field and also by Sinha *et al.* (1987) for mixing ratio, so only a few important points of the scheme are given below.

In this scheme the weights for the observing stations are obtained from autocorrelation of the given parameter, viz., relative humidity, (denoted by "f") over Indian region by solving a set of equations given in Eqn. (1).

$$\sum_{j=1}^n \mu_{ij} P_j + \lambda^2 P_i = \mu_{oi} \quad i = 1, 2, \dots, n \quad (1)$$

$$\text{where, } \mu_{ij} = \frac{f_i' f_j'}{\sigma_0^2}; \mu_{oi} = \frac{f_o' f_i'}{\sigma_0^2};$$

$$\lambda^2 = \frac{\sigma_{\epsilon_i}^2}{\sigma_0^2} \text{ and } \sigma_0^2 = \sigma_i^2 - \sigma_{\epsilon_i}^2$$

Here,  $\sigma_i^2$  is the average variance and  $\sigma_{\epsilon_i}^2$  is the total mean random error in the observation of relative humidity.

$P_i$  is the weighting factor for the observing station,  $i$ , and  $f_i' f_j'$ ,  $f_o' f_i'$  are the covariances of the observed

anomalies of "f" between the stations  $i$  and  $j$  and the grid point "0" and the station  $i$  respectively.

Figs. 1(a & b) depict both the autocorrelation function and structure function of relative humidity at 700 hPa for the July month. The autocorrelation,  $\mu(\rho)$ , which is a function of distance is modelled into the following equation.

$$\mu(\rho) = C_1 e^{-C_2 \rho} \quad (2)$$

The mean random error in the observations  $\sigma_{\epsilon_i}^2$  was computed by extrapolating the structure function curve to zero distance [Fig. 1(b)].  $C_1$  and  $C_2$  the constants in the Eqn. (2) are computed by the least square method.

#### 3.2. Estimation of upper air relative humidity from regression relation

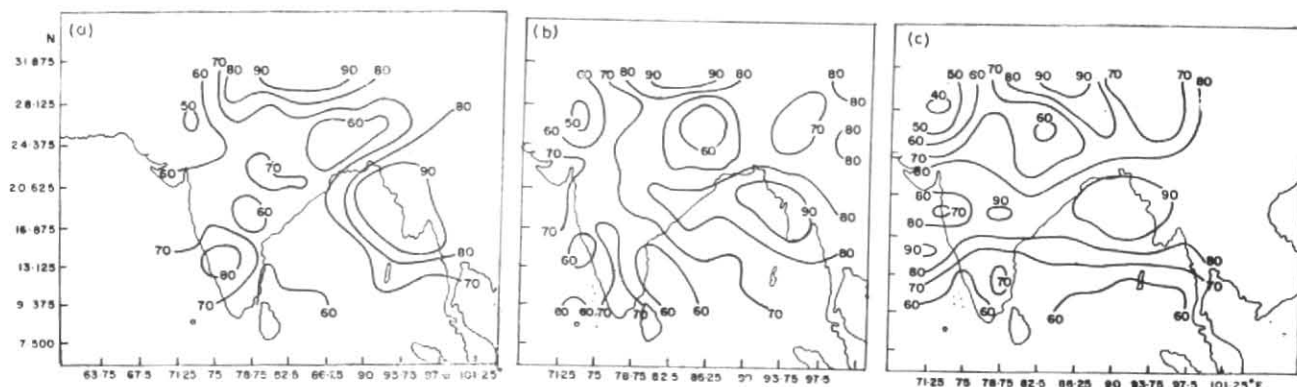
To supplement the inadequate radiosonde observations of relative humidity, some investigators have developed techniques for using the surface observations to estimate the upper air humidity (Chisholm *et al.* 1968, Atkins 1974, Jonas 1976, Chu and Parrish 1977, Rasmussen 1982). Chisholm used a decision tree approach to combine surface observations which gave reliable estimate of upper level humidity and remaining cases he analysed by a technique of REEP, i.e., regression estimation by event probabilities.

A decision tree approach can be used when surface reports consist of qualitative observations (e.g., present weather) and quantitative observations (e.g., humidity) but drawback of this method is that the estimated humidity is not a continuous function of input data. To overcome the drawback of decision tree approach, Rasmussen (1982) used a multiple linear regression technique. Following Rasmussen we have also in this study used this technique for estimation of relative humidity at 700 hPa.

The dependent variable is relative humidity at 700 hPa and independent variables are surface humidity, total cloud amount and present weather.

#### 3.3. Use of dummy variables in regression

Present weather is reported in codes (00 to 99). It is very difficult to use the codes as such in regression equation because the coded numbers do not vary systematically with the variation of relative humidity in the atmosphere. So, following Chisholm *et al.* (1968) and Rasmussen (1982), the present weather was categorised into three groups (such groups are called dummy variables). Chisholm *et al.* (1968) and Rasmussen (1982) have used this technique for dew point spread and low cloud type respectively. In the present study we have classified the present weather codes into three categories as dry, showery and wet. These three categories were assigned the values 1, 2 and 3 respectively. In other words, the values 1, 2 or 3 were used in the computations while developing the prediction equation when present weather was respectively either dry, showery or wet.



Figs. 4(a-c). Analysis II of relative humidity at 700 mb including estimated data of: (a) 5 July '79, (b) 7 July '79 and (c) 9 July '79

The following table shows the use of dummy variables to replace the code numbers of the variable (WW) in regression.

Present weather code (India Met. Dep.)	Category	Dummy variable
0 to 12, 28, 30 to 37 and 40 to 49	dry	1
13 to 19, 25 to 27, 29 and 80 to 99	showery	2
20 to 24, 38, 39 and 50 to 79	wet	3

The regression relation between the upper air relative humidity at 700 hPa level and the surface observations was obtained using the daily data for four July months (1976 to 1979) of 19 radiosonde stations over the Indian region. The daily data of only 16 stations, observations numbering around 1100, were used in deriving the coefficients of the regression relations Eqn. (3). The daily data of remaining three stations were kept as an independent data set for testing of the regression relations. The prediction equation obtained for 700 hPa relative humidity is of the following form:

$$RH_{700} = 48.9 + 3.88 \times WW + 1.29 \times N + 0.15 \times RH_s \quad (3)$$

WW has a value 1, 2 and 3 depending on whether the present weather is dry, showery or wet.  $N$  is the cloud cover in octas and  $RH_s$  is surface relative humidity.

The coefficients in Eqn. (3) were tested for significance and found to be highly significant for the three predictors. The multiple correlation coefficients ( $R$ ) was found to be .953 and so the coefficient of multiple determination ( $R^2$ ) indicates that about 91% of the variance of 700 hPa relative humidity has been explained by this relation. The stations over which the relative humidity values were estimated are shown in Fig. 2.

The regression relation as mentioned earlier was tested with an independent data set, *i.e.*, the data of three stations which were not used in deriving regression relation. This test showed that the errors in the estimation of RH were less than 10% for 63% of the cases and between 10-20% for 24% of the cases and 21-30% in 13% of the cases. The root mean square error of estimation was found to be 15%.

The regression relation was also tested with another set of independent data. The data of all radiosonde stations for the period from 1 to 10 July 1984 were used. The total number were about 250 observations which reported all necessary parameters. In this case, the errors in the estimation of RH were less than 10% for 43% of the cases, between 11 to 20% for 39% of the cases and 21% and above for 18% of the cases. The R.M.S. error for the estimated values was found to be 15.6% considering all estimations.

#### 4. Analysis and discussion of results

##### 4.1. Analysis of relative humidity

Using the univariate optimum interpolation scheme, the analyses of relative humidity at 700 hPa level were made for six days from 4 to 9 July '79 over the region from 9.375°N to 30°N and from 71.250°E to 99.375°E with 1.875 deg. lat/long grid. ECMWF analyses of the corresponding previous days were used as initial guess field. First the analysis (Analysis I) were made only with radiosonde data as input and then analyses (Analysis II) were made with radiosonde data as well as the relative humidity at 700 hPa, estimated using the regression equation from surface observations. This experiment was carried out using climatology also as initial guess in the analysis scheme.

During the period 4 to 9 July 1979 considered for the analysis, a low had formed over the head Bay of Bengal, it intensified into depression on 7 July and crossed the Indian coast on 8 July. This depression had westnorth-westerly movement and its position on 9 July was 25°N and 77°E (approx.) consequent of the movement of the depression, the rainfall pattern changed and the region of maximum rainfall moved westwards as can be seen from the cloud pictures given in the FSU Report No. 83-7 by Krishnamurti *et al.* (1983) Figs. 3(a-c). The cloud pictures also showed intensive convective activity in the west coast of southern Peninsular India. In the subsequent sections it would be examined whether the relative humidity analysis I & II depict these features.

##### 4.2. Discussion on the impact of estimated relative humidity

The objective of this study is to improve the analysis scheme by making provisions to estimate the relative humidity wherever radiosonde data are not available

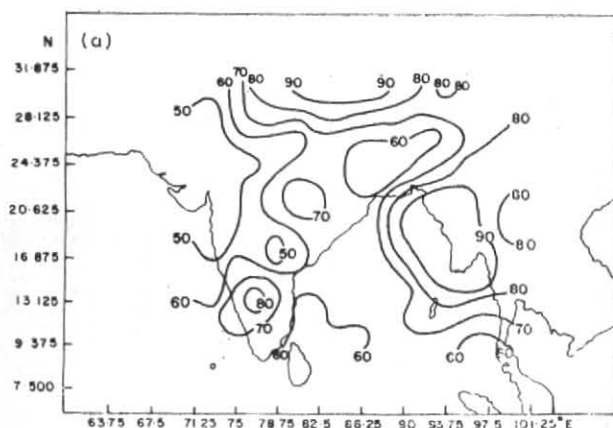


Fig. 5(a). Analysis I of relative humidity (only radiosonde data included) 5 July '79

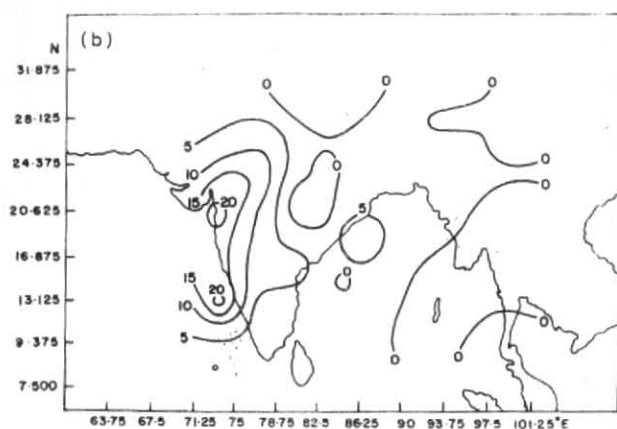


Fig. 5(b). Difference in analyses (Analysis II — Analysis I) 5 July '79

and including them in the analysis. The next step is to assess the impact of the estimated relative humidity on the analysis. In order to do this, the analysis I and II which were made respectively without and with including the estimated relative humidity, were examined to determine whether analysis II are closer to reality. This could be done in different ways. One way is to examine whether analysis II fit the radiosonde observations of relative humidity better; another way is to examine the statistical measure of the analysis difference with respect to a standard analysis, and yet another way is to examine how the analysis depict the physical features like cloudiness etc.

The R.M.S. errors were calculated by interpolating at observing points the relative humidity value from the analysis I or II and were compared with actual observations (Table 1). This shows that on all days except on 4 and 6 July, the R.M.S. errors of analysis II are smaller than those of analysis I suggesting that when estimated relative humidity values were included, the analysis II fits better with observations. Of course the decrease in R.M.S. errors is only marginal.

Table 1 also shows the bias computed as the mean of the error in the analysis, for the two sets of analyses. Except on 4 July in the case of analysis II when the error was positive, on all other days for both sets of analyses,

TABLE 1

Bias towards observations and R.M.S. errors compared with (i) ECMWF analysis and (ii) radiosonde observations

Date (July)	Bias towards observations	R.M.S. errors in % (compared with ECMWF analysis)	R.M.S. errors in % (compared with observations)
<b>Radiosonde and estimated data</b>			
4	5.09	10.49	10.53
5	-1.30	9.42	7.14
6	-0.64	9.60	11.55
7	-0.25	7.95	8.27
8	-0.36	11.57	9.22
9	-4.15	10.65	8.15
<b>Only radiosonde data</b>			
4	-1.27	11.78	9.13
5	-3.00	8.72	7.58
6	-3.64	10.55	9.92
7	-0.58	8.67	8.49
8	-1.71	12.09	9.28
9	-2.46	12.30	8.27

the mean errors were negative. This implies that on all cases (except one) the analyses are drier than the actual observation.

The R.M.S. errors computed by comparing the standard analysis from ECMWF showed that except on 5 July, on other days the R.M.S. errors of analysis II are smaller than those of analysis I. This suggests that analysis II has been closer to ECMWF analyses.

Next, comparison of the analyses I and II are made with reference to the satellite cloud pictures over the Indian region. Both analyses showed that the region of cloudiness agreed with the region of maximum relative humidity. Especially, the isopleths of 80% agreed with the cloudiness due to the monsoon depression on all days. To illustrate this, the satellite cloud pictures and the analysis II for a limited number of days, viz., 5, 7 and 9 July were given in Figs. 3 and 4 respectively. Similar features are also seen in the analysis I (However, these figures are not included here). The convective activities over the west coast of southern Peninsular India on 5 and 9 July were represented by higher values of relative humidity in both analyses, whereas on 7 July, the cloud patch was not seen over there. This has been reflected in the RH analyses with only 60% isopleth being there (Figs. 3 and 4).

In general, the analyses I are drier than those of analyses II implying that impact of the estimated values is essentially to make the analyses more moist. Fig. 5(a) shows analysis I of 5 July which may be compared with Fig. 4(a) analysis II of 5 July to examine the differences between them. Fig. 5(b) shows that the difference in analyses (Analysis II—Analysis I) has been more to the west of 80°E with maximum difference up to 24% over the west coast of India.

The satellite cloud picture corresponding to 5 July shows that there are three cloudy regions one over the Bay of Bengal extending up to Orissa, and the second at the foot hills of the Himalayas and the third over the southern parts of the west coast of Peninsular India. Of course in the both analysis I and II there are three maxima corresponding to the region of maximum cloudiness. However, analysis II on all days, have greater magnitude particularly over the Peninsular region. From this it can be inferred that the analysis II depicts the cloudy region over west coast of Peninsular India better than analysis I. However, both analysis were similar at the other two cloudy regions.

To sum up the examination of the analyses I and II and the corresponding satellite cloud pictures shows that the addition of estimated relative humidity to radiosonde observations improves the analysis only marginally. However, no analysis was deteriorated by the addition of estimated data. This strongly favours the use of estimated data wherever, radiosonde data are not available. Also, over the oceanic region, if reliable ship observations are available, they could be used indirectly in the analysis.

##### 5. Concluding remarks

The regression equation for estimating relative humidity values at 700 mb level from surface observation has been developed using daily data of four July months. This regression equation was verified with an independent data set and the estimated values were found to be within 15% of the observed values. When the estimated relative humidity at 700 mb were used as supplementary data to radiosonde data the resultant analysis had R.M.S. errors marginally less than those when estimated values were not used.

When comparison was carried out of the satellite cloud pictures for the days of analysis with both the analysis I and II made without and with estimated data, the following inferences were made. Firstly, both the analysis of relative humidity have depicted fairly well

the cloudy region due to the depression as well as the convective region in the west coast of the southern Peninsula. Secondly, the analysis II made including the estimated relative humidity values depict better the convective region in the west coast of India. This suggests that over radiosonde sparse regions as well as over oceanic regions (where ships data are available) relative humidity can be estimated from surface observations and used to supplement the radiosonde data in order to improve the upper air relative humidity analysis.

##### Acknowledgements

The authors are thankful to Shri. D. R. Sikka for his interest in this study and to Dr. S. S. Singh and Shri P. P. Kamte for going through the manuscript and making very useful comments. They also wish to thank Shri K. D. Barner for typing the manuscript.

##### References

- Atkins, M.J., 1970, *Met. Mag.*, **99**, pp. 98-110.  
 Atkins, M.J., 1974, *Tellus*, **24**, pp. 663-667.  
 Ball, J.T. and Veigas, K.W., 1968, *J. appl. met.*, **7**, pp. 620-625.  
 Begum, Z.N., Bansal, R.K. and Datta, R.K., 1987, *Mausam*, **38**, 3, pp. 341-344.  
 Chisholm, D.A., Ball, J.T., Veigas, K.W. and Luty, P.W., 1968, *J. appl. met.*, **7**, pp. 613-619.  
 Chu, R. and Parish, D., 1977, NMC Office Note No. 149, p. 29.  
 Cressman, G.P., 1959, *Mon. Weath. Rev.*, **87**, pp. 367-374.  
 Dimego, G.J., 1985, Proc. of workshop on high resolution analysis. ECMWF, pp. 213-258.  
 Gandin, L.S., 1963, *Gidrometeorologicheskoc*, Izdatel, Stov, Lenin-grad, USSR, p. 286.  
 Illari, L., 1985, Proc. of workshop on high resolution analysis, ECMWF, pp. 41-68.  
 Jonas, P.R., 1976, *Met. Mag.*, **105**, pp. 44-56.  
 Kastner, A., 1974, *Arch. Met. Geoph. Biokl. Ser A*, **23**, pp. 137-148.  
 Krishnamurti, T.N., Steven Cocks, Richard Pasch and Simon Low-Nam, 1983, FSU Report No. 83-7, May 1983.  
 Nerquist, D.C., 1986, AFGL-TR-86-0194 Air Force Geophysics Laboratory.  
 Nerquist, D.C., 1988, *Mon. Weath. Rev.*, **116**, 2, pp. 452-470.  
 Rajamani, S., Talwalkar, D.R., Ray, S.P. and Upasani, P.U., 1983, *Mausam*, **34**, 1, pp. 43-50.  
 Rasmussen, R.G., 1982, Ph. D. Thesis, Univ. of Drexel, U.S.A.  
 Sinha, S.K., Talwalkar, D.R. and Rajamani, S., 1987, *Advances in atmospheric sciences*, **4**, 3, pp. 332-342.  
 Tibaldi, S., 1982, *Rivista Di Meteorologia, Aeronautica*, **42**, pp. 309-328.  
 Timchalk, A., 1986, NOAA Tech. Rep. NESDIS 24, Wash. DC, p. 60.