551. 575

Some studies on fog prediction at Dum Dum

K. C. MAJUMDAR

Regional Meteorological Centre, Calcutta

(*Received 2 August 1955*)

ABSTRACT. A quantitative analysis of the problem of fog prediction has been carried out by choosing the lowest visibility observed in the morning as the dependent factor in the multiple regression. A forecasting technique based on ideas analogous to statistical decision with a "fog-region" from the plot of points of factors like dew point depression and depth of inversion layer is found to be quite fruitful. The results of regression or "fog-region" when combined with synoptic criteria like High in Head Bay of Bengal in an objective manner are found to improve to a considerable extent. The possibility of the forecasting of fog with the help of mean dew point depression for fog days is also discussed.

1. Source of data

The data of February for the years 1952 to 1955 taken from the daily current weather registers, monthly meteorological registers and the radiosonde messages of the Calcutta Airport at Dum Dum have been used for the study. For synoptic situation weather charts of Alipore and *Indian Daily Weather Reports* published from Poona have been consulted.

2. Mean dew point depression

An analysis of variance of dew point depressions at 1800 GMT at Dum Dum for December to March for 5 years from 1949-1950 to 1954-1955 reveals that there is a significant month to month variation of the mean dew point depression, d (D.B. temperature—dew point temperature in °F) both for fog days as well as for all days in general, showing that mean of d for each month should be considered separately. It is also found by Fisher's t-test that mean of d for fog days are significantly lower from the general mean considering all days as shown in Table 1.

The mean for fog days may, therefore, be considered to be the most probable value of d that might lead to the formation of fog. Thus if on a day d is observed to be same as this mean and is further supported by favourable synoptic conditions, we may forecast fog with a greater degree of confidence.

3. A fog prediction formula

A regression equation has been worked out with the lowest visibility, v (in kilometres), attained during the period from midnight (0000 IST) till morning (0900 IST) of a day as the dependent variable and (*i*) dew point depression, d recorded at 1800 GMT, (*ii*) the depth of inversion layer, h (in ft) obtained from the 1500 GMT radiosonde data on the previous night, as independent variables.

The mean and standard deviation of the variables considered are respectively found to be 2·19 km and 1·78 for v, 6·63°F and 3·59 for d and 1076 ft and 580 for h. Although the correlation coefficient (C.C.) of v with d and h are only +0·234 and —0·231, it has been found that when logarithms of the variables are taken there is an improvement in their correlation coefficients. Thus C.C. between $\log v$ and $\log (0\cdot1+d)$ is 0·507 and that between $\log v$ and $\log (1+h)$ is —0·249. The multiple regression relation is then obtained with these three factors as shown in equation (1). It may be mentioned that factors like dew point temperature and surface wind speed at 1800 GMT were also tried but their C.Cs. with the lowest visibility were not encouraging being much less.

$$\log v = -0\cdot135+1\cdot683 \log (0\cdot1+d)- 0\cdot397 \log (1+h) \qquad (1)$$

This has a multiple correlation coefficient, R=0·607, which by F-test (F=23·57 with n_1=2 and n_2=81) though significant is, however, not sufficiently high.

TABLE 1

Mean dew point depression	Jan	Feb	Mar	Dec
For fog days	5·29	4·59	5·54	4·29
For all days	6·36	6·63	7·08	5·86
t	2·15†	3·21‡	1·97*	3·11‡
D.F.	180	175	184	169

*Significant at 5% level. †Significant at 1% level.
‡Significant at 0·1% level.

TABLE 2

	High (or Low)		No High (or Low)		Total
	(b)	(a)	(b)	(a)	
Fog	22 (8)	3 (7)	2 (3)	2 (11)	29
No fog	9 (23)	24 (20)	10 (9)	39 (30)	82
Total	31	27	12	41	111

Equation (1) can also be written as

$$v=0\cdot734(0\cdot1+d)^{1\cdot683}/(1+h)^{0\cdot397} \quad (2)$$

To get v readily from the equation, a nomogram drawn in Fig. 1 can be used. Whenever an observed point (d, h) falls on or below the line through $v=1\cdot0$ km (as indicated in the nomogram) fog may be expected, while if it falls on or below the line $v=2\cdot0$ km but above the line $v=1\cdot0$ km only mist may be forecasted.

4. Combination with synoptic situation

It is seen (Basu 1952) that a "High" in Head Bay of Bengal (or a High extended upto Head Bay) and less frequently a low to the west of long. 85°E on the 1200 GMT synoptic chart influence the fog formation at Dum Dum. The contingency chisquare, $\chi^2(=45\cdot0$ with 3 d.f.), and coefficient of contingency, C $(=0\cdot74)$ calculated from the data of Table 2 also establish that there is a definite dependence of fog on the position of Highs and Lows as stated.

Table 2 gives frequencies of fog, High (Head Bay) or Low west of 85°E with respect to the point (d, h) lying (a) above $v=1\cdot0$ km and (b) on or below $v=1\cdot0$ km in Fig. 1 which is the graphical representation of regression equation (1). Frequencies derived on the assumption that there is no association between fog and the synoptic criteria are shown in brackets. The substantial differences between some of the expected and actual frequencies confirm that the factors are related.

From Table 2 it is seen that if we forecast fog in all cases with such a High (or Low) only we shall be correct in 43 per cent cases, while skill score* (Brier and Allen 1951) of such a forecast is 33 per cent. On the other hand, taking the cases under (b) of Table 2, the percentage accuracy and the skill score of the forecast based on the regression equation alone are respectively 56 and 51 per cent. Now if we predict fog when there is a favourable synoptic situation mentioned above together with the point (d, h) lying on or below the line $v=1\cdot0$ km of Fig. 1, we find from data under (b) against High (or Low) of Table 2 that the percentage accuracy and the skill score of the forecast will be respectively 71 and 63 per cent which show a great improvement of the fog forecast by a combined effect of regression and synoptic condition over that by only one of them.

5. A fog prediction diagram using d and h

Fig. 2 gives a scatter diagram of sample points $p(d, h)$ plotted with d as abscissa and h as ordinate. The bounded region in this figure, say a region R (d,h), is similar to the critical region in the theory of statistical decision (Blackwell and Girshick 1954). Here this may be termed as "fog region". Now if an observed point p (d, h) falls within

*Skill Score$=F-E/T-E$ where F is the number of correct forecasts of fog out of total, T and $E=(C_1 R_1+C_2 R_2)/T$, C_i and R_i being respectively the i^{th} column total and i^{th} row total$(i=1,2)$ of the 2×2 contingency table

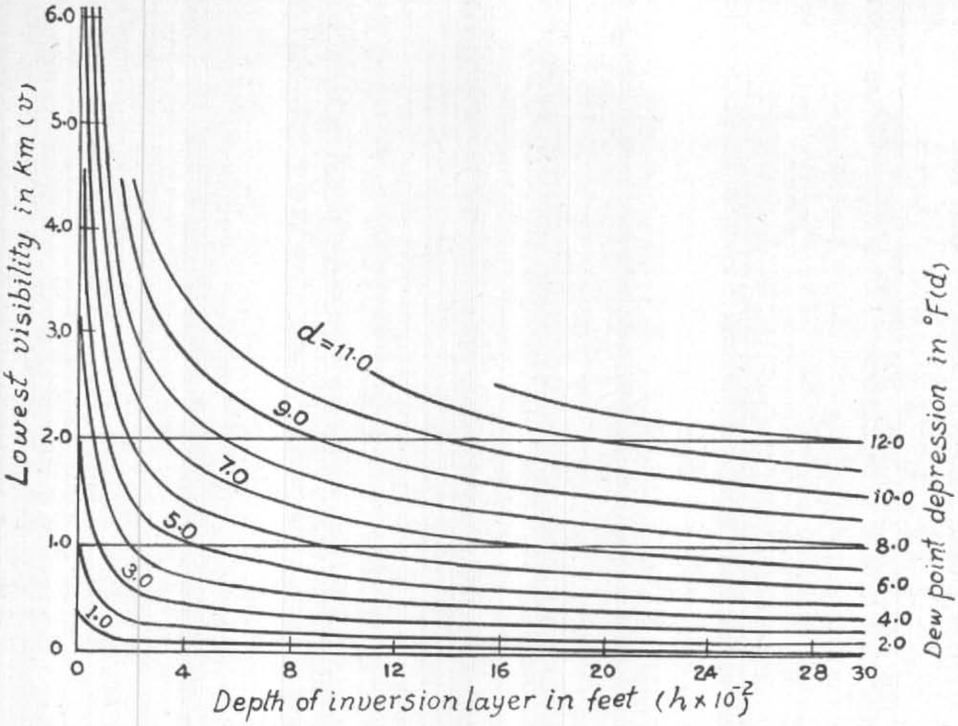

Fig. 1. Curves showing the expected lowest visibility in February at Dum Dum for different values of depth of inversion layer and dew point depression (d in °F) recorded on previous night

the region R (d, h) fog may be forecasted, while fog will not be forecasted if $p(d, h)$ lies outside R (d, h). In the case considered, forecasts of "fog" and "no fog" will go wrong in $\alpha = 40$ per cent and $\beta=6$ per cent cases respectively as seen from Table 3. α and β are equivalent to first and second kinds of error and $p(d, h)$ is analogous to the decision function. The region $R(d, h)$ is drawn so as to minimise the second kind of erorr β, because a larger value of β would mean more chance and risk of plane crashes or similar other disasters due to wrong forecast of "no fog". The percentage accuracy and the skill score of the fog forecast calculated on the basis of this "fog region" are respectively 60 and 57 per cent as obtained from the data of Table 3 which gives the frequencies of fog, High (Head Bay) or Low west of 85°E with respect to the point $p(d, h)$ lying within or outside the fog-region, $R(d, h)$ (*vide* Fig. 2), with expected frequencies given in brackets.

TABLE 3

	High (or Low)		No High(or Low)		Total
	$p(d,h)$ within $R(d,h)$	$p(d,h)$ outside $R(d,h)$	$p(d, h)$ within $R(d,h)$	$p(d, h)$ outside $R(d,h)$	
Fog	23 (8)	2 (7)	2 (3)	2 (11)	29
No fog	9 (24)	25 (20)	8 (7)	40 (31)	82
Total	32	27	10	42	111

Now if we combine these results of "fog region" with the synoptic conditions considered in the previous section the percentage accuracy and skill score of the fog forecast considerably improve to 72 and 63 per cent. It is seen that these results of fog prediction are quite comparable with those obtained with regression formula in combination with synoptic situation.

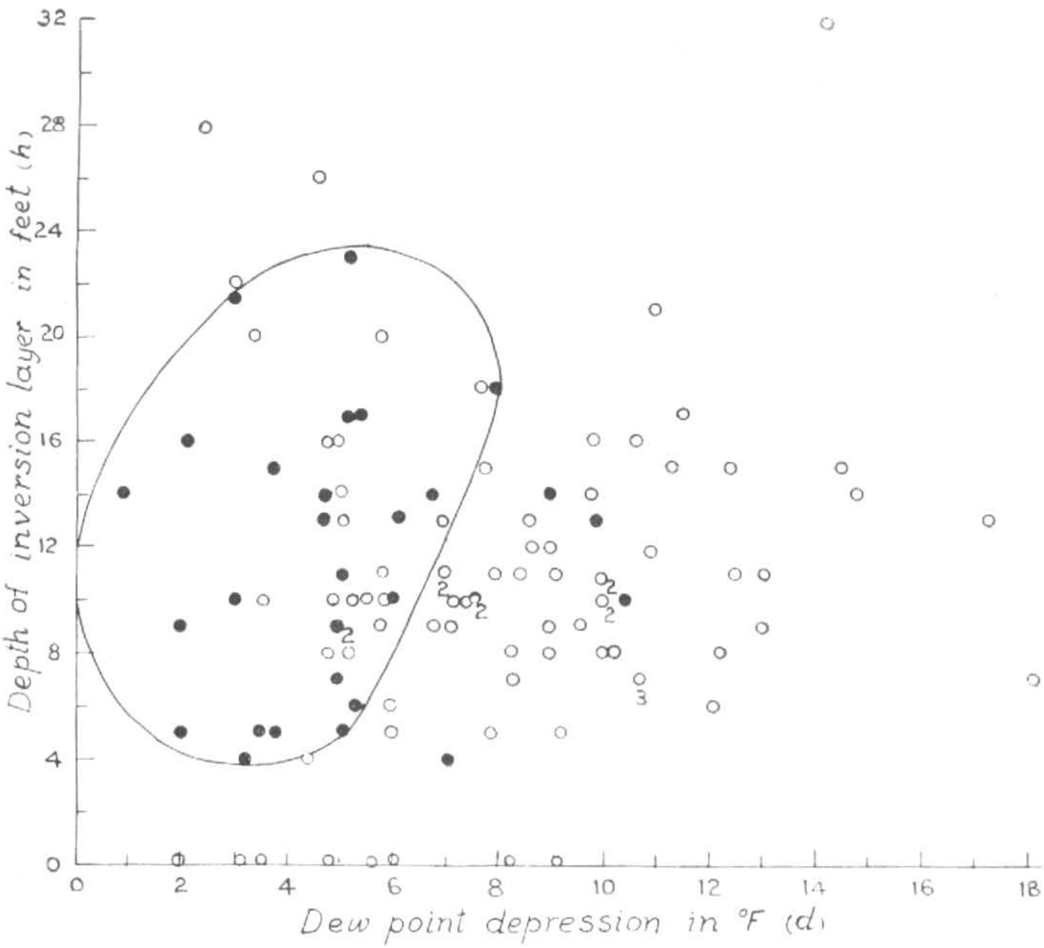

Fig. 2. A fog prediction diagram with fog region (bounded area) for February at Dum Dum

(Filled and open circles indicate points $p(d, h)$ corresponding to fog and no fog respectively. The numerals below the circles indicate the number of overlapping points)

6. Concluding remarks

The results of the studies made in the foregoing sections clearly indicate the usefulness of the various statistical methods in fog prediction. The improvement of the results of regression and "fog-region" when combined with synoptic situation in an objective manner is very encouraging. Such studies if carried out extensively with good many years' data would be all the more fruitful.

REFERENCES

Basu, S. C.	1952	*Indian J. Met. Geophys.*, **3**, 4, p. 281.
Blackwell, D. and Girshick, M. A.	1954	*Theory of Games and Statistical Decision*, pp. 81.
Brier, G.W. and Allen, R.A.	1951	*Compendium of Meteorology*, pp. 844, 846.