OBJECTIVE PROBABILITIES FOR FOG AT CANBERRA AIRPORT USING BAYESIAN AGGREGATION AND PRINCIPAL COMPONENTS

I. B. Mason

Regional Forecasting Centre, Bureau of Meteorology, Canberra

(Manuscript received January 1977; revised September 1977)

ABSTRACT

A statistical model for fog probability at Canberra Airport has been developed, based on local and pressure field data available at 3 pm on the preceding day. Bayesian aggregation of one-dimensional conditional probabilities using principal components enabled the effects of a large number of variables to be combined. Application of the model to independent data gave results comparable with subjective assessment of probabilities by forecasters on the same days. The model appears capable of further development.

INTRODUCTION

It is quite straightforward to find conditional probabilities for future weather events on one or two conditioning variables by extracting contingency tables. However, use of contingency tables to assess the combined effect of a large number of related variables is not feasible. The number of observations in some classes becomes too small, and multidimensional tables are unwieldy.

If the conditioning variables are mutually independent then any number of one-dimensional probabilities for each variable can be aggregated using Bayes' theorem (e.g., Winkler and Murphy 1971). Unfortunately, mutual independence is rare among meteorological variables. Use can be made, however, of the fact that normally distributed variables that also have coefficient of correlation zero are independent. In this note linear combinations of the meteorological variables are used as predictors, these linear combinations being constrained to have zero correlation. Values of the linear functions appear to be sufficiently close to normally distributed to make them for practical purposes independent.

The particular problem under consideration is quantitative assessment of the probability of fog at Canberra Airport in winter, or more precisely the probability of a visibility of 1000 m or less due to fog between 7 am and 12 noon (the most critical time for air movements). The assessments are required at 3 pm on the preceding day, in order to allow intending air travellers time to make alternative arrangements if necessary.

DATA

Two sets of data were used to develop the statistical model.

The first consisted of daily values of fourteen variables measured at Canberra Airport at 3 pm on 300 June days (1962-1971), together with an indicator variable for subsequent occurrence of fog. The particular variables chosen are given below, with reasons and some comments. The general criteria for selection were that values should be available at 3 pm, that they should be routinely measured meteorological variables, and that they had a plausible physical relationship with fog.
The second data set consisted of daily values of MSL pressure at 9 am at the fourteen grid points shown in Fig 1, for the same 300 June days as the first data set.

![Fig 1 Grid points for surface pressure](image)

The same data for a further 60 days (June 1973 and 1974) were extracted to use as independent test data.

All values were taken either from the register of aerodrome weather reports (Bureau of Meteorology Form A37) for Canberra Airport or from the Field Book (A9).

The predictand was an indicator variable (0/1) for occurrence of fog (visibility ≤ 1000 metres) between 7 am and 12 noon at Canberra Airport.

The definition of fog is intentionally rather narrow. Quite a number of days were entered as 'no fog', although 'fog patches' or 'fog at a distance' was noted in the register of aerodrome weather reports. The criterion that the horizontal visibility at the airport be 1000 m or less was rigidly applied. The reason for this rigidity lies in the broad aim of the project, which was to provide the most useful service possible; aircraft are delayed by fog at the airport, but not by fog patches or fog at a distance. It is desirable to attempt to forecast the operationally significant event, and to try to exclude occurrences of fog that do not cause problems. This policy probably has introduced some additional uncertainty into the predictions. It seems preferable that this uncertainty be incorporated in the objective predictions, rather than be left to be incorporated subjectively afterwards.

The following variables were selected as predictors:
1. $\bar{V}$: Average of surface wind speeds at 1 pm, 2 pm and 3 pm. An average was taken, rather than the single value at 3 pm, to smooth out short-period fluctuations. Surface wind speed should affect the probability of fog through its relationship with turbulence and vertical diffusion; stronger winds reduce the likelihood of fogs by increasing the depth of atmosphere through which the nocturnal radiational cooling is mixed. There is some evidence for a non-linear relationship between wind speed and fog at very low wind speeds, with the maximum probability of fog corresponding to a 3 pm wind around 5 to 6 kn. With lighter (but not dead calm) winds, and some other conditions being favourable, a water vapour pressure gradient towards the surface develops, resulting in downward transport of moisture which may sometimes be rapid enough to keep the relative humidity below 100 per cent despite rapid cooling of the air.

2. $\Delta \bar{V}$: 24-hour change in $\bar{V}$ (defined above) to indicate synoptic-scale trends.

3. $\bar{U}$, $\bar{V}$: W-E and S-N components of the mean of surface winds at 1 pm, 2 pm and 3 pm.

4. $\bar{V}$: Canberra's geographic location suggests that moisture advection should be related to airstream direction.

5. $\bar{N}T$: Average amount of low cloud at noon and 3 pm. The average was taken, once again, to reduce short-period fluctuations. Cloud reduces the likelihood of fog by reducing the rate of cooling of the ground surface at night or even producing a slight warming, through downward long wave radiation. However, low cloud in the afternoon is often convective in origin, associated with an otherwise clear atmosphere, disappearing early in the evening, and may thus be associated with an increased probability of fog.

6. $N_m$: Amount of middle level cloud at 3 pm.

7. $\Delta N_m$: Change in amount of middle level cloud from 9 am to 3 pm. To detect situations in which middle level cloud is increasing or decreasing.

8. $T$: Dry-bulb temperature at 3 pm. Relatively high 3 pm temperatures in winter tend to be associated with clear skies, and thus should be associated with higher probabilities for fog.

9. $\Delta T$: 24-hour change in $T$ to indicate synoptic-scale trends.

10. $T_d$: Dew point at 3 pm. High dew points were expected to favour fog, and low dew points the reverse. It was in fact found that the higher dew points tended to be associated with wet, cloudy, windy conditions, which are not usually conducive to fog. Also, there is some reason to suspect that dew points below 0°C reduce the likelihood of fog even more than the lack of moisture would lead one to expect; it seems likely that lower dew points are associated with earlier formation of frost, which may accentuate the downward gradient of vapour pressure and hence the downward transport of atmospheric moisture by turbulent diffusion.

11. $\Delta T_d$: 24-hour change in 3 pm dew point.

12. $P$: 3 pm MSL pressure. Pressure is quite strongly correlated with fog occurrence, probably due to the relationship between pressure, cloud, and wind.
13. $\Delta P$ : 24-hour change in pressure to indicate synoptic-scale trends.

14. $R$ : Indicator variable (0/1) for occurrence of a trace or more of rain between 9 am and 3 pm. Soil moisture content affects the radiating efficiency of the ground surface, and the availability of moisture for fog formation.

Persistence was not included, although fog is sometimes regarded as having strong day-to-day persistence. A check of the data showed that persistence was a poor predictor for fog in the 10 months used in this study. Of 69 occurrences of fog, only 26 would have been correctly forecast using this strategy.

DEVELOPMENT OF THE MODEL

As stated in the introduction, Bayes' theorem can be used to aggregate probabilities, and reduces to a simple form if mutually independent variables are used. The following symbols are defined:

$F$ : the occurrence of an event $F$.

$\bar{F}$ : the event that $F$ did not occur.

$P(F)$: the a priori probability of $F$, i.e., the probability of $F$ in general, before anything is known of a particular occasion.

$A_i, i = 1, ..., n$ : $n$ variables to which $F$ is to be related. They may in general be either continuous variables or categorised. Later in this paper $A_i$ is used as the $i^{th}$ principal component.

$a_i$ : a value of the variable $A_i$.

$P(A_i = a_i|F)$ : the probability that $A_i$ has the value $a_i$, given that $F$ occurred.

$P(F|A_1 = a_1, A_2 = a_2, ..., A_n = a_n)$ : the probability of $F$, given that $A_i$ has the value $a_i$, etc.

If the $A_i$ are mutually independent and $F$ is, as defined above, a dichotomous variable, then the following expression for Bayes' theorem can be derived (Winkler and Murphy 1971):

$$P(F|A_1 = a_1, A_2 = a_2, ..., A_n = a_n) = \frac{\Omega}{1 + \Omega}$$

where

$$\Omega = \frac{P(F) \cdot P(A_1 = a_1|F) \cdot P(A_2 = a_2|F) \cdot \ldots \cdot P(A_n = a_n|F)}{P(\bar{F}) \cdot P(A_1 = a_1|\bar{F}) \cdot P(A_2 = a_2|\bar{F}) \cdot \ldots \cdot P(A_n = a_n|\bar{F})}$$

Information about $F$ contained in the value of each of the $n$ 'predictors' $A_i$ taken separately is thus aggregated to yield a single conditional probability for $F$. The ratios $P(A_i = a_i|F)/P(A_i = a_i|\bar{F})$ are referred to in this note as conditional likelihood ratios.

$F$ is the event that a visibility of 1000 m or less due to fog was reported at Canberra Airport between 7 am and 12 noon in June. The conditioning variables $A_i$ are linear functions of the meteorological variables found by principal component analysis as described below.
It is not proposed to give a detailed exposition of the technique of principal component analysis here. It is well described in a number of textbooks and papers. The present writer found Sokal and Sneath (1973) useful. In a meteorological context Barry and Perry (1973), Stidd (1967), or Craddock and Flood (1969) among many others give discussions of the method. The particular program used was taken from Dixon (1964) and run on the Australian Bureau of Meteorology's IBM 360/65 computer.

The significance of the technique in the present study is that it replaced the original fourteen variables in each data set with fourteen new variables, each of which is a linear function of the original variables and each of which is uncorrelated with the others. The coefficients of the original variables in the new functions are the components of the eigenvectors of the correlation matrix. In passing it may be noted that the covariance matrix may give better results, for reasons discussed by Hartigan (1975), pages 60-3. The new variables found by principal component analysis generally provide an efficient description of the original data, in that only a few may account for most of the variance in the data. In the Canberra Airport data, for example, seven functions explained 80 per cent of the variance, and in the grid point pressure data three functions could be used to explain 98 per cent of the variance in June 9 am pressure fields. The most significant functions may be interpreted physically as 'weather patterns'. For example, the first component of the local data has the following form (coefficients have been rounded to one significant figure for clarity):

\[ A_1 = 0.5 \Delta V + 0.3 \Delta N + 0.5 u - 0.3 v + 0.09 \Delta T - 0.07 \Delta m - 0.1 \Delta N m - 0.2 T - 0.1 \Delta T - 0.05 \Delta d - 0.09 \Delta d - 0.4 \Delta p - 0.2 \Delta p + 0.04 r. \]

This component accounted for 19 per cent of the variance in the data. It can be seen that the variables with the strongest influence on the value of \( A_1 \) are wind-related variables and pressure. High values of \( A_1 \) will be found on days with high winds from the northwest quadrant and low pressures and low values of \( A_1 \) will correspond to days with light winds from the southeast and high pressures. It may be noted that the other variables have physically reasonable signs in this component; for example, high values of \( A_1 \) will tend to be associated with more low cloud than normal, low temperatures, and slightly increased rain. Low values of \( A_1 \) will tend to occur with higher temperatures, less cloud, and less rain. Thus the first eigenvector describes in broad terms two mutually exclusive types of winter day in Canberra - a day with fresh northwesterly winds and low pressures, and a day with light southeasterly winds and high pressures.

The coefficients of the first component of the MSL pressure fields are shown in Fig 2. High values of this component will occur with above normal pressures, whose spatial distribution is similar to that of the coefficients. Low values will occur when pressures are below normal with a distribution reversed in sign from that of Fig 1 (i.e., lowest values over southern New South Wales/central Victoria). This component accounted for 87 per cent of the variance in June pressure fields.

It is interesting to note that the correlation coefficient between the first component of the local data \((A_1, \text{above})\) and that of the pressure fields was -0.69. Thus high values of \( A_1 \) have a marked association with low values of the first pressure field component, and vice versa. This is physically reasonable.

Similar interpretations to the above can be made for most of the other components, although those corresponding to eigenvalues less than one have doubtful significance.

The relationship with fog was also physically reasonable. High values of \( A_1 \) reduced the likelihood of fog on the following morning quite markedly, and low values increased the probability of fog. Similarly, high values for the first component of the pressure field data, which imply anticyclonic conditions in the region, increase the probability of fog, and low values reduce it.
A more detailed account of the patterns found in this data will be given in a separate note.

\[ P(A_i = a_i | F) \] and \[ P(A_i = a_i | \bar{F}) \] were found from conditional frequency distributions of each of the \( A_i \) with \( F \) and \( \bar{F} \) as follows. The days were separated into two groups, those followed by fog and those not followed by fog. Then for each variable \( A_i \) they were assigned to one of three classes within these groups using class boundaries of \( \pm 0.43075 \) for the value of \( A_i \). These boundaries would separate members of a standard normal distribution into three equal groups. Thus for each variable \( A_i \) the days of the data set are sorted into those with high positive values of \( A_i \), values around zero, and high negative values. The conditional likelihood ratios then have the form \[ P(a_i > 0.43075 | F)/P(a_i > 0.43075 | \bar{F}) \], \[ P(-0.43075 \leq a_i \leq 0.43075 | F)/P(-0.43075 \leq a_i \leq 0.43075 | \bar{F}) \], and \[ P(a_i < -0.43075 | F)/P(a_i < -0.43075 | \bar{F}) \]. They are found as the ratios of the corresponding relative frequencies within the fog and no-fog groups. Empirical relative frequencies were used as the probabilities, no assumption being made about any hypothetical underlying probability density except insofar as the assumption of mutual independence implies normality.

The two \( 14 \times 300 \) data matrices (14 variables on 300 days each) containing the local (Canberra Airport) and pressure field grid point data respectively were transformed using principal component analysis. Each of the two resulting transformed \( 14 \times 300 \) matrices contained daily values of fourteen linear functions of the original fourteen variables on the 300 days of the data set.

The two new \( 14 \times 300 \) matrices produced by principal component analysis were reduced to two smaller matrices, \( 7 \times 300 \) in the case of the airport data and \( 3 \times 300 \) for the pressure data, each of which accounts for most of the variance in its original data set, by deleting the rows of daily values of the least significant functions. These two matrices were then combined into a \( 10 \times 300 \) matrix, which was itself subjected to principal component analysis, providing a final
10 x 300 matrix of values of ten linear functions of both local data and grid
data pressure data. Five of these functions were sufficient to explain 73.5 per-
cent of the variance in the data, but all ten were used in the model.

If Bayes' theorem is to be applied in the form given above (Eqn 1) it is
necessary that the variables be mutually independent. Uncorrelated variables are
mutually independent if they are normally distributed. As the new variables are
linear functions of a large number (28) of basic meteorological variables it
seemed plausible on the basis of the central limit theorem (e.g., Heathcote
1971, page 177 et seq.) that they might have distributions at least approaching
normality. Accordingly, frequency distributions were extracted for the daily
values of each of the new functions and compared with the normal distribution
for the same mean and standard deviation, using chi-squared. Six of the ten
functions were definitely non-normal with significance levels for chi-squared less
than 0.01. However, since all ten distributions were more or less symmetrical and
uni-modal, it was decided to proceed with the analysis in the hope that departures
from independence would not introduce serious biases into the probabilities.
This did in fact seem to be the case.

The normalising transformations of Box and Cox (1964) could improve this
situation, possibly at the expense of re-introducing correlation between the
variables.

As a check, 'pair-wise' independence of the form \( P(A_i|A_j, F) = P(A_i|F) \) was
tested by forming 3 x 3 contingency tables between each pair of the new variables
for both \( F \) and \( F \) (i.e., fog and no-fog days). Class boundaries were ±0.43075,
the terciles of a standard normal distribution. From the resulting ninety con-
tingency tables five pairs of variables were found that could not be regarded as
independent at the 5 per cent level, and there was no pair significantly dependent
at 1 per cent. This was encouraging, but of course implies nothing about inde-
pendence of the form \( P(A_i/A_j, A_k, F) = P(A_i|F) \) or with more variables \( (A_1, A_3, \)
etc.) on the right hand side of the conditionality sign.

The assumption of mutual independence is a weak point in the analysis, and
can only be justified by appealing to the results.

The next step was extraction as described above of ten sets of 'conditional
likelihood ratios', and the a priori likelihood ratio \( P(F)/P(\bar{F}) \).

The resulting values are given in Table 1.

A value greater than unity for the conditional likelihood ratio implies that
the probability of fog is higher than its a priori ('climatological') level, and a
value less than unity that the probability is reduced below climatology. For
example, suppose that after substitution of values of the meteorological variables
in the expression for \( A_i \) it is found that the value of \( A_1 \) is greater than 0.43075. From
Table 1 it can be seen that

\[
P(A_1 > 0.43075|F)\]
\[
P(A_1 > 0.43075|\bar{F})\]

\[
\frac{P(F)\ P(A_1 > 0.43075|F)}{P(\bar{F})\ P(A_1 > 0.43075|\bar{F})} = \frac{0.299 \times 0.46300}{0.138} = 0.138 \]

and \( P(F|A_1 > 0.43075) = \frac{\Omega}{1 + \Omega} = 0.121. \)
Table 1  Conditional likelihood ratios (see text for explanation). The value of the class boundary is truncated to 0.43 in the headings to this table. The value actually used was 0.43075.

| Principle component | $P(a_1 < -0.43|F)$ | $P(-0.43 < a_1 < 0.43|F)$ | $P(a_1 > 0.43|F)$ |
|---------------------|---------------------|-----------------------------|---------------------|
| $A_1$               | 1.36120             | 1.58280                     | 0.46300             |
| $A_2$               | 1.58036             | 0.93974                     | 0.48402             |
| $A_3$               | 0.89535             | 1.79791                     | 0.62542             |
| $A_4$               | 1.04054             | 1.08999                     | 0.84874             |
| $A_5$               | 0.39855             | 1.48792                     | 1.20522             |
| $A_6$               | 1.27312             | 0.88378                     | 0.86957             |
| $A_7$               | 0.61241             | 1.37347                     | 1.03735             |
| $A_8$               | 0.51949             | 1.19964                     | 1.07383             |
| $A_9$               | 0.71739             | 1.10096                     | 0.92065             |
| $A_{10}$            | 2.86957             | 0.96095                     | 0.41848             |

The a priori likelihood ratio, $P(F)/P(\bar{F})$, was 0.299.

Thus the information that the value of $A_1$ is greater than 0.43075 has reduced the probability of fog from its climatological value of 0.23 to 0.12. The effects of $A_2$, $A_3$, etc., on the probability of fog can be incorporated in the same way.

From the $10 \times 300$ matrix of daily values of each of the new eigenvector coefficients a value of

$$\Omega = \frac{P(F) \cdot P(A_1|F)}{P(F) \cdot P(\bar{A}_1|\bar{F})} \frac{P(A_{10}|F)}{P(\bar{A}_{10}|\bar{F})}$$

was computed for each day, and hence daily values of $P(F|A_1, A_2, \ldots A_{10}) = \ldots 5$

The resulting 300 probabilities were compared with subsequent occurrence of fog by rounding each probability to one significant figure (i.e., 0.0, 0.1, ..., 0.9, 1.0) and finding the relative frequency of fog for each probability. The results are shown in Table 2.

The independent data (June 1973 and 1974) were treated in the same way. Daily probabilities for fog were found using the principal components and conditional likelihood ratios derived from the original (1962-1971) data, and compared with the occurrence of fog as described above. Table 3 shows the relationship between probability and relative frequency for the independent data.

Mean values of the Brier probability score (Brier 1950) were computed. The Brier score is a measure of accuracy, and has the form, in Murphy and Epstein's formulation (1967):

$$PS = 1 - \frac{K}{2p_j} + \frac{K}{\sum p_i^2} \quad (0 \leq p_i \leq 1; \sum p_i = 1; \quad i = 1, 2, \ldots, K) \quad \ldots 6$$
where \( K \) is the number of mutually exclusive and exhaustive categories for the predictor (in this case \( K = 2 \)), \( p_i \), \( i = 1, K \) are the forecast probabilities and \( p_j \) is the probability forecast for the category that occurred. PS has range \([0, 2]\) and lower scores are better. It is a 'strictly proper' score in the sense of Winkler and Murphy (1968).

RESULTS

Table 2 shows the relation between probability and relative frequency of fog when the model was applied to the data from which it was derived. The fourth line ('significance') gives the probability of observing the corresponding relative frequency on the assumption that the probabilities are unbiased and that the subset of days corresponding to each probability can be regarded as a homogeneous set of Bernoulli trials. It can be seen that none of the relative frequencies would require rejection, at the 5 per cent level of confidence, of the hypothesis that the probabilities were correctly assessed, and that differences between probability and relative frequency can be attributed to sampling error.

The mean value of the Brier score for this set of probabilities was 0.2934. A standard of comparison is provided by fifty subjective assessments of fog probability made during June and July 1976 by forecasters in the Canberra Regional Forecasting Centre in identical form to those discussed above, for which the score was 0.2744 (recall that lower scores indicate more accurate probabilities).

Standard deviation of daily values of the Brier score for the subjective forecasts was 0.3, implying a standard deviation for a mean of 50 daily values of 0.3/\( \sqrt{50} \) = 0.04. The difference between the mean scores for objective and subjective forecasts was 0.019, less than one standard deviation. Thus the difference cannot be regarded as significant.

Table 2 Comparison of computed probabilities and relative frequencies of fog in developmental data.

<table>
<thead>
<tr>
<th>Probability classification</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of days when computed probability was in the class</td>
<td>56</td>
<td>69</td>
<td>67</td>
<td>28</td>
<td>33</td>
<td>21</td>
<td>14</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Number of days when fog was observed</td>
<td>3</td>
<td>7</td>
<td>13</td>
<td>5</td>
<td>15</td>
<td>10</td>
<td>9</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Observed relative frequency of fog when computed probability was in the class</td>
<td>0.054</td>
<td>0.101</td>
<td>0.194</td>
<td>0.179</td>
<td>0.455</td>
<td>0.476</td>
<td>0.643</td>
<td>0.500</td>
<td>1.000</td>
<td>-</td>
<td>\textbf{1.000}</td>
</tr>
<tr>
<td>'Significance'</td>
<td>0.11*</td>
<td>0.16</td>
<td>0.12</td>
<td>0.06</td>
<td>0.11</td>
<td>0.17</td>
<td>0.21</td>
<td>0.10</td>
<td>0.80</td>
<td>0.90</td>
<td>-</td>
</tr>
</tbody>
</table>

* Using \( p = 0.025 \).

Use of persistence as the forecast strategy gave a score of 0.573, much less accurate than either of the sets of probability forecasts.

Table 3 shows the results of application of the model to sixty days of independent data for June 1973 and 1974. The correspondence between probability and relative frequency is evidently much weaker than in Table 1.
Table 3  Comparison of computed probabilities and fog relative frequencies - independent data, June 1973 and 1974.

<table>
<thead>
<tr>
<th>Probability classification</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of days when computed probability was in the class</td>
<td>8</td>
<td>21</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Number of days when fog was observed</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Observed relative frequency of fog when computed probability was in the class</td>
<td>0.125</td>
<td>0.143</td>
<td>0.250</td>
<td>0.0</td>
<td>0.571</td>
<td>0.200</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>'Significance'</td>
<td>0.17*</td>
<td>0.20</td>
<td>0.29</td>
<td>0.17</td>
<td>0.19</td>
<td>0.16</td>
<td>0.03</td>
<td>0.09</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* Using $p = 0.025$.

Due to the small size of the sample, conclusions about bias in the probabilities cannot be drawn with any great certainty except in the case of probabilities assessed as 0.6. Of the four occasions on which this probability was produced by the model none was followed by fog. The probability of getting no 'hits' in a sample of four from a binomial population with $P = 0.6$ is only 0.026. All the other relative frequencies were well above the 5 per cent level of significance. The Brier score for the forecasts as they stand is 0.3723, indicating a lower accuracy than Table 2 as would be expected, and also lower than the accuracy of the subjective probabilities in 1976. It is interesting, however, to compare the probabilities computed for these independent days with subjective fog risk forecasts prepared at the same time (1500) for the same event (and issued to the public) on forty of the same days, from 20 to 30 June 1973 and most of June 1974. The subjective forecasts were given as one of four categories of risk, intended to correspond with numerical probabilities as follows:

- **Negligible risk**  \( Pr(\text{fog}) < 0.05 \)
- **Slight risk**  \( 0.05 \leq Pr(\text{fog}) \leq 0.30 \)
- **Moderate risk**  \( 0.30 < Pr(\text{fog}) \leq 0.70 \)
- **High risk**  \( Pr(\text{fog}) > 0.70 \)

(The exact values of probability for the boundaries of these risk categories were of course somewhat arbitrarily chosen. In particular, it may be considered that a probability of 0.3 is too low to warrant calling it a moderate risk. However, the climatological probability of fog is 0.23, and it is arguable that a probability of 0.3 is far enough above normal to make 'moderate' a reasonable term.) Taking the mid-point of the range as the forecast probability the Brier score for these forecasts was 0.4054. Categorising the objective probabilities for the same forty days in the same way the score was 0.3812. The difference is not statistically significant, but this result does show at least that the objective model produces probabilities that are of comparable accuracy with subjective assessments of risk.

It is also interesting to note that subjective and objective risks agree quite well with each other. Table 4 shows the numbers of days on which each subjective risk corresponded with each objective risk categorised as described above.
Table 4 Comparison between subjective and objectively derived risk categories for fog at Canberra Airport, 40 days, 20 June to 29 June 1973 and 1 June to 30 June 1974. N = negligible risk; SL = slight; M = moderate; H = high. Note that this is a subset of the data described in Table 3.

<table>
<thead>
<tr>
<th>Subjective</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>S</td>
<td>3</td>
</tr>
<tr>
<td>M</td>
<td>0</td>
</tr>
<tr>
<td>H</td>
<td>0</td>
</tr>
</tbody>
</table>

On exactly half the days both techniques produced the same risk, and on only two days were the risks widely different (two categories apart). The single occasion in the forty days on which both gave a high risk was not followed by fog between 0700 and 1200. There was fog earlier in the night, which lifted to stratus at about 0500. This illustrates the difficulty of the task.

Table 4 shows that the two sets of risk assessments agree with each other reasonably well, but does not indicate how well either agreed with the subsequent observations. Tables 5 and 6 show the relationship between risk and observation for the subjective and objective risks respectively.

Table 5 Verification of *subjective* assessments of fog risk for 40 days from 20 to 29 June 1973 and 1 to 30 June 1974.

<table>
<thead>
<tr>
<th>Assessed risk</th>
<th>Observed</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No fog</td>
<td>Fog</td>
</tr>
<tr>
<td>N</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>M</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>H</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 6 Verification of *objective* assessments of fog risk for 40 days from 20 to 29 June 1973 and 1 to 30 June 1974.

<table>
<thead>
<tr>
<th>Assessed risk</th>
<th>Observed</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No fog</td>
<td>Fog</td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>M</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
The objective technique was much more conservative than the subjective, in that the extreme categories were selected less frequently in the former case. Whether or not this is a drawback depends on the nature of the operation using the forecasts. Two examples may clarify this point.

First, consider an operation with a weather-sensitive component for which the cost of taking precautions against adverse weather, \( C \), is $700, and the loss if no precautions are taken and adverse weather occurs, \( L \), is $1000. It can be shown (e.g., in Thompson and Brier 1955) that the expected value of the cost to the operation is minimised if the decision to protect is taken on the following criterion, where \( p \) is the probability of adverse weather:

\[
\text{protect, if } p > \frac{C}{L} \\
\text{do not protect, if } p < \frac{C}{L} \\
\text{(either, if } p = \frac{C}{L}).
\]

In the present case protection would be provided when a high risk was assessed, but not otherwise. Referring to Table 5, the subjective risk was high on nine occasions, implying a total cost of protection of $6300. There were five occasions on which fog followed a risk of \( N \), \( L \), or \( M \), giving a cost of $5000, and thus a total cost to the operation of $11300 over these forty days.

Now, referring to Table 6, the objectively derived risk was high just once, resulting in a cost of protection of $700. There were nine occasions when fog occurred after lower risks had been given resulting in a cost of $9000 and total cost of $9700, $1600 less than if the subjective forecasts had been used. It may be noted that a still cheaper strategy would have been to ignore both sets of forecasts and take no precautions, saving another $700 over the objective forecasts. For an operation with \( C/L \) as high as 0.7 both systems had an unacceptably high 'false alarm' rate.

Now, secondly, consider an operation with \( C = $300 \) and \( L = $1000 \), and suppose that protection will be provided when the risk is assessed as \( M \) or \( H \), but not for \( N \) or \( S \). By similar reasoning to that shown above, the total cost to the operation if subjective forecasts were used would be \( 14 \times \frac{300}{1000} + 4 \times \frac{1000}{300} = 8200 \), and for the objective forecasts \( 16 \times \frac{300}{1000} + 4 \times \frac{1000}{300} = 8800 \). Clearly for this operation the subjective assessments of risk were preferable to the objective, and both were cheaper than neglecting precautions altogether. Other applications of this method of evaluation of probability forecasts can be found in the paper by Thompson and Brier, and in Mason (1975, 1976, 1977). The Brier probability score gives a reasonable ranking of forecasts with respect to average value over the whole set of users, and the values of this score given above show that the objective forecasts were slightly, although not significantly, better.

**FURTHER DEVELOPMENT**

It seems likely that the model could be improved by seeking class boundaries that maximise the discrimination of the data for fog, rather than the somewhat arbitrarily selected ±0.43075.

An alternative to the use of this 'contingency table' approach would be to model the conditional probabilities \( P(A_1|F) \) and \( P(A_1|F) \) using probability density functions, then form the conditional likelihood ratios, or to fit some analytic function of \( a_1 \) to the conditional likelihood ratios.

The component analysis performed in this study used the correlation matrix, implying that values of the original variables had all been 'standardised' to have a variance of one. This may be undesirable (see Hartigan (1975) for a full discussion of the reasons) and any future work should use the covariance matrix.
Application of the normalising transformations of Box and Cox (1964), either before or after the principal component analysis, should be investigated. This could be expected to reduce the dependence between the variables, and should thus improve the validity of the probabilities.

It is not claimed that the meteorological variables selected in this paper are the best possible. The choice of variables was based only on physical considerations. Use of screening regression on a large range of variables, or on combinations of variables, should make it possible to get a more effective set of predictors.

CONCLUSIONS AND COMMENTS

A statistical model for fog probability at Canberra Airport has been developed using Bayesian aggregation of one-dimensional conditional probabilities based on principal components.

Application of the model to its own developmental data produces probabilities that are sufficiently close to observed relative frequency for the difference to be regarded as sampling error. Accuracy as indicated by the Brier probability score is comparable with that of subjective assessments of fog probability made by forecasters in June and July 1976.

On independent data there is significant bias in probabilities given as 0.6. Accuracy is lower than in the developmental data, but is still at least as good as that shown by a set of subjective fog risk assessments made by forecasters on some of the same days.

It has been shown that the method described in this paper can provide valid objective estimates of the probability of future weather events.

The initial development is somewhat complex, and requires significant computer power; however, once the eigenvectors and conditional likelihood ratios have been found, routine application is possible on quite a modest machine. Any of the current HP98 series would be more than adequate.

The applicability of this method is quite general, and is not necessarily confined to dichotomous predictands. It could be used to provide valid objective probability forecasts for any parameter for which enough developmental data are available.

ACKNOWLEDGMENTS

Ms Margaret Bell's assistance in the Bureau of Meteorology's ADP Branch was indispensable to this project. Dr Neil Macdonald of the Geography Department at the Australian National University introduced the writer to principal component analysis and provided considerable useful information about the technique. John Sullivan extracted the raw data in the Australian Capital Territory Regional Office of the Bureau, and Mrs Sonja Purkiss typed the manuscript.

REFERENCES


Dixon, W.J. 1964. *BMED - Biomedical Computer Programs*. (School of Medicine, University of California: Los Angeles.)


