On the Economic Value of Probability of Precipitation Forecasts in Canada

AMBURY STUART

Atmospheric Environment Service, Downsview, Ontario, Canada

(Manuscript received 28 April 1981, in final form 4 January 1982)

ABSTRACT

Canadian weather offices have recently begun to receive probability of precipitation information estimated from numerical weather prediction guidance, but this information is not given in the public forecast. This paper uses a well-known procedure to estimate the increased economic value of these probability forecasts over equivalent categorical forecasts for Toronto. The cost-loss matrix given by Epstein (1969) is generalized and applied in economic value calculations of a Probability of Precipitation Amount (POPA) procedure for the same Toronto data set. The increased economic value of these probability forecasts is a substantial fraction of the value of categorical forecasts and appears to justify the expense involved in a marketing/education program necessary for the public to understand these forecasts.

1. Introduction

The format of Canadian public forecasts differs from the format of American forecasts in that the former do not provide a numerical estimate of the probability of precipitation. In the United States, probability of precipitation (POP) forecasts have been given since 1965, and several authors including Thompson (1962), and Murphy (1977) have demonstrated the potential for improvements in the economic value of weather information if these forecasts are provided. Those opposed to the implementation of POP forecasts point to studies by Rogell (1972) which show that most members of the general public do not understand these forecasts with the result that these potential economic advantages are not realized. Recent Canadian surveys (Stuchbury and Lapczak, 1979) confirm the results of Rogell for Toronto. However, Murphy et al. (1980) have shown that these misunderstandings relate to the events of concern rather than the probabilities themselves. Furthermore, the general public in the United States prefers numerical probabilities to verbal expressions of uncertainty by a ratio of 7 to 3 (see M.S.I. Services, 1981).

This controversy underlines the fact that the economic value of weather forecasts depends upon many factors including accuracy, forecast format, and the ability of users to understand and apply the information to weather sensitive activities. Most national weather services, including the Atmospheric Environment Service (AES) of Canada, concentrate their activities on improving forecast accuracy and time-line. Substantial funding has been received in the past and is continually being sought for these activities. AES argues that the return on this investment of public funds is realized, at least in part, through the enhanced economic value of weather forecasts. Recently, however, the importance of the consumer has been identified through the recognition of the need for AES to give more attention to marketing its products and services. The effort to enhance the economic value of weather forecasts must include both product improvements and user-oriented activities.

This paper is, in part, an examination of the economic value of weather forecasts that may be attributed to the forecasts' accuracy and/or to a user's knowledge of the concepts necessary to apply probability forecasts to weather-sensitive activities. Forecast users who do not have this knowledge are presumed to follow the categorical version of the forecasts; the economic advantage of this advice is attributed to forecast accuracy. If forecasts are unavailable, knowledgeable users are presumed to follow climatological tables; the economic advantage of this information is attributed to user skills. Probability forecasts have economic value that is attributed both to forecast accuracy and user skills.

Sanders (1973, 1979) and Cook and Smith (1977) have shown that improvements in forecast accuracy are now occurring only very slowly. It is possible that the public might be better served if part of the funding now directed to improving forecast accuracy is diverted to a public education program. If probability forecasts were better understood, then the economic value of weather information would be increased without any increase in forecast accuracy. To deter-
mine whether or not this expenditure is prudent, it is first necessary to estimate the marginal economic utility of probability forecasts. AES has developed an objective scheme for predicting precipitation probability using only numerical weather prediction output. This paper is the first attempt to estimate the economic value of these forecasts.

The Probability of Precipitation Amount (POPA) scheme used here is described by Agnew (1979). Using only numerical model guidance, four probability values are generated indicating the likelihood of: 1) no measurable precipitation; 2) light precipitation (less than 2 mm of equivalent water in a 12-h period); 3) moderate precipitation (more than 2 mm but less than 10 mm equivalent water in 12 h); or 4) heavy precipitation (more than 10 mm of equivalent water in 12 h). These values can be reduced to the simpler probability of precipitation occurrence (POP) format by noting that the first value is the probability of no precipitation, while the sum of the last three values, is the probability of precipitation. Agnew's method was recently applied to Numerical Weather Prediction model output from the Canadian Meteorological Centre (CMC) and provided 24 and 48 h POPA forecasts for Toronto during the period 1 December 1979–30 April 1980.

The development of an N×N cost-loss matrix to define weather sensitivity is described in Section 2. In Section 3, the economic value calculation procedure is outlined; the results and conclusions are given in Section 4.

2. The user sensitivity matrix

This section extends the work of Epstein (1969) through the development of a more general cost-loss matrix. Consider a user of public forecasts who incurs different losses $L_i$ for each of $N$ mutually exclusive and jointly exhaustive weather states. For each of these states, there is assumed to be a unique protective action, which, if taken, protects the user from any loss should this, or a less severe state, occur. Each action is administered at some cost $c_i$. Both the cost of protection and the unprotected loss increase with increasing weather severity. The economic advantage of protecting against the occurrence of weather state $i$, given that this state is correctly predicted, is

$$L_i - c_i = L_i(1 - a_i),$$

(1)

where $a_i$ is the cost-loss ratio for state $i$. If one chooses protective action $i$ and a more severe weather state $j > i$ occurs, it is assumed that the economic advantage of this protection is the same as would be the case if state $i$ had occurred. In summary, the expense $E_{ij}$ of taking protective action $i$ with the subsequent occurrence of weather state $j$ is

$$E_{ij} = a_i L_i,$$  

$i > j$

$$E_{ij} = L_i - L_i(1 - a_i),$$  

$i < j$

(2)

The least severe weather state ($j = 1$) means "no adverse weather," or in this paper, "no precipitation." The associated protective action is "take no action" which implies that both $L_1$ and $c_i$ are equal to zero.

Consider now the case of a user who always chooses the same protective action. Over a period of time, a distribution of weather states, $n_i$, occurs. The expenses for such a user are $X_i$ where

$$X_i = \sum_{j=1}^{N} E_{ij} n_j = a_i L_i \sum_{j=1}^{N} n_j + \sum_{j=i+1}^{N} L_i n_j - L_i \sum_{j=i+1}^{N} n_j,$$

(3)

where $i$ identifies the chosen protective action.

If the frequency distribution of weather states, $n_i$, is given by climatological expectations, one may determine in advance the expected expenses of each possible protective action if continually applied over a long period of time. One may then choose the action with the lowest expected expense for uniform application to every situation; this is the climatology strategy. This protective action depends upon $n_i$ which varies from place to place and $E_{ij}$ which varies from user to user.

The constant choice of a particular strategy, either arbitrarily or with reference to climatology, is made without the benefit of forecast information. With either categorical or probability forecasts, the protective action is expected to change from time to time. For categorical forecasts, the user is assumed always to follow the forecast so that protective action $i$ is selected whenever weather state $i$ is predicted. If $n_{ij}$ is the joint frequency of a forecast of state $i$ and the occurrence of state $j$, then the expense associated with following a categorical forecast is

$$X_{cat} = \sum_{i=1}^{N} \sum_{j=1}^{N} E_{ij} n_{ij}.$$  

(4)

Considering probability forecasts, the estimated probability of each weather state $p_j$ is substituted for $n_j$ in Eq. (3). Then the action with the lowest expected expense is chosen.

The assumption is now made that all weather states have a common cost-loss ratio $a$. One may then show from Eq. (3) that

$$(L_i - L_{i+k})(a - \sum_{j=i+1}^{N} p_j) \leq \sum_{j=i+1}^{N} (L_j - L_{i+k}),$$

$$k = 1, 2, \ldots, N - i$$

(5)

$$(L_i - L_{i-k})(a - \sum_{j=i-k+1}^{N} p_j) \leq \sum_{j=i-k+1}^{N} (L_i - L_j)p_j,$$

$$k = 1, 2, \ldots, i - 1$$

where $i$ identifies the action with the lowest expected
expense, and may then show by induction that these expressions are equivalent to
\[ \sum_{j=i+1}^{N} p_j \leq a \leq \sum_{j=1}^{N} p_j. \] (6)

This is the same expression given by Epstein (1969) for the special case
\[ L_i = \frac{i - 1}{N - 1} L. \] (7)

This result extends the applicability of Eq. (6) to all monotonically increasing \( L \), and not just those values that satisfy Eq. (7). For example, the \( 5 \times 5 \) cost–loss matrix specified by Rapp (Thompson, 1972) for a construction operation satisfies Eq. (2) for all \( a_i = 0.333 \). Therefore, despite the fact that the unprotected loss values \( L_i \) do not satisfy Eq. (7), the optimum strategy for each series \( p_j \) can still be determined with Eq. (6).

3. Economic value calculations

The value of public forecasts is estimated by comparing the expenses of a user who follows a forecast strategy (either categorical or probability information), with the expenses of a user whose strategy is independent of such information (either "never protect" or climatology). Likewise, the value of educated users is estimated through a comparison of strategies which require probability information (either climatology or probability forecast strategies) with the expenses of less complex strategies (never protect or follow the categorical forecast). There is a protective strategy associated with each combination of forecasts/no forecasts and an education program/no program as is illustrated in Table 1.

If both probability and categorical forecasts are available, it is assumed that without an education program only the categorical information will be received. Murphy et al. (1980) have questioned the assumption that only probability forecasts are misinterpreted, but we will assume categorical forecasts are followed perfectly. In so doing, we will tend to overestimate the value of the present forecast format and underestimate the value of an education program. If forecasts were removed in favor of an education program, the user could apply climatology information, via Eq. (3), to determine the best protective action. With neither forecasts nor education, the user is assumed to ignore the weather, never take protective action, and to cover his losses as they occur.

The expense of never taking protective action is given by Eq. (3) with \( i = 1 \); Eq. (4) defines the expenses associated with categorical forecasts. In this paper, categorical forecasts are derived from probability forecasts in that the weather state with the highest probability is chosen for the categorical pre-

| Table 1. Each of the four methods for choosing a protection strategy is related to the presence or absence of forecasts or an education program as shown. |
|----------------|-----------------|-----------------|
| No forecast | Forecast |
| No education program | Never protect | Categorical forecast |
| Education program | Climatology strategy | Probability forecast |

...dition. The expenses of the climatology strategy are given by Eq. (3); probability forecast expenses are given by Eq. (3), with \( p_j \) substituted for \( n_j \). This approach to the problem of calculating \( \text{ex post} \) economic value is identical to that taken by Thompson (1962), Murphy (1977) and others.

It remains to specify \( L_i \) and \( a \). We will adopt Epstein’s (1969) formulation for \( L_i \) [Eq. (7)] and will permit the cost–loss ratio to take on the values 0.1–0.9 in turn. This assumes that only these cost–loss values are possible and that for the entire population of forecast consumers taken together all cost–loss ratio values occur with equal frequency. Both the two-state precipitation/no precipitation case and the four-state POPA scheme will be evaluated using the four-month data set of 24-h forecasts and the three-month data set of 48-h forecasts. Since expenses are calculated after the fact, this is an "ex-post" analysis.

It is convenient to relate the performance of these procedures with easily recognized bench marks. The "never-protect" strategy was chosen as the zero level of economic utility since it requires neither a forecast nor an educated user. Perfect categorical forecasts were chosen to represent maximum utility. If a protection strategy has an expense \( E \) then the value of this strategy is \( V \) where
\[ V = (E_{np} - E)/(E_{np} - E_p), \] (8)

where \( E_{np} \) and \( E_p \) are the expenses associated with the never-protect strategy and with perfect forecasts respectively. In this formulation, the never-protect strategy has zero value while perfect forecasts have a value of unity. It is noted that for the uniform distribution of cost–loss value \( a \), used here and indeed for any distribution symmetrical about \( a = 0.5 \), it may be shown that
\[ E_{np} = 2E_p. \]

4. Results and conclusions

The value, as defined by Eq. (8), of each protective strategy is given in Table 2. Also indicated in this table is the percentage of correct categorical forecasts in the data sets used. The climatological expectation probabilities of (0.68, 0.20, 0.10, 0.02) were approximated by actual frequencies of (0.70, 0.17, 0.09, 0.04) for the 24-h forecast data set and (0.69, 0.17, 0.10, 0.04) for 48-h forecast data set for
TABLE 2. Value of each strategy for each leadtime and number of precipitation categories. Also indicated is the percentage of correct categorical forecasts in each case.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Lead time</th>
<th>24 h</th>
<th>48 h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>POPA</td>
<td>POPA</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Never protect</td>
<td>0.12</td>
<td>0.22</td>
<td>0.13</td>
</tr>
<tr>
<td>Follow climatology</td>
<td>0.48</td>
<td>0.52</td>
<td>0.21</td>
</tr>
<tr>
<td>Follow categorical forecast</td>
<td>0.59</td>
<td>0.62</td>
<td>0.41</td>
</tr>
<tr>
<td>Percentage correct</td>
<td>78</td>
<td>86</td>
<td>72</td>
</tr>
</tbody>
</table>

The economic value of precipitation forecasts can be greatly enhanced provided that probability forecasts are no longer denied to the people of Canada and that the public understands these forecasts and uses them in making decisions in weather sensitive activities.

Acknowledgments. The author wishes to recognize the diligence of an anonymous referee who was extremely helpful in the development of the manuscript. He is also indebted to Tom Agnew who provided the data base, to INTERA Environmental Consultants Ltd. for support in the final stages of this work and to Pat Duncan for typing the manuscript.

REFERENCES


