Preprints, 16th AMS Conference on
Weather Analysis and Forecasting
Phoenix, Arizona, 11-16 January 1998, J24-J27

J2.4 VALUE OF WEATHER FORECASTS FOR ELECTRIC UTILITY LOAD FORECASTING

Harold E. Brooks
NOAA/ERL/National Severe Storms Laboratory
Norman, OK 73069

Andrew P. Douglas
School of Electrical and Computer Engineering
University of Oklahoma
Norman, OK 73019

1. INTRODUCTION

Weather forecasts can be evaluated in a variety of ways. Murphy (1993) defined three kinds of forecast "goodness": consistency, quality, and value. Consistency refers to the relationship between the forecasts and the "true beliefs" of the forecaster, quality refers to the relationship between the forecasts and weather events, and value refers to the relationship between the forecasts and the benefits or losses accrued by users. Most forecast evaluation studies have focused on the quality question, in large part because it is, in some sense, the most tractable and requires only the collection of information about forecast and observed weather. The potential for great economic value of weather information has been highlighted recently (Katz and Murphy 1997) The complexity of the relationship between quality and value of forecasts has been illustrated by Roebber and Bosart (1996). Two primary results can be drawn from their studies. First, high (low) quality does not necessarily correspond to high (low) value for different forecast users. Secondly, benefits are often concentrated in a small number of forecast situations. Thus, even though weather information may not be significant on a large number of days, it may be important enough on a small number to have large economic impact.

Brooks et al. (1997) have evaluated the quality of forecasts presented by a variety of media sources and the National Weather Service (NWS) for Oklahoma City for a period of one year, using a distributions-oriented approach (Murphy and Winkler 1987). The forecasts include maximum and minimum temperatures out through five days lead time and precipitation probability forecasts from two of the stations out to seven days. In summary, the different forecast sources all have strengths and weaknesses and selecting an overall "best" forecast based upon forecast quality is extremely problematic.

On the other hand, it is possible to determine the relative value of the forecasts for individual users, assuming that the sensitivity of those users to particular combination of forecasts and weather events is known. Douglas et al. (1998) have adapted a model of the scheduling of plants for the local electrical utility to account for the use of weather information in the scheduling process via Bayesian estimation. To test the model, Douglas et al. used the forecasts and observations of Brooks et al. (1997) and actual customer demand for the local utility. Electric utilities have to make decisions about what generation plants to turn on or off based upon anticipated usage by customers. Depending on the nature of the generation plants, it may take two or more days to bring the plant on-line or take it off-line. Generating too little power locally to meet customer demands may require the utility to buy power from other utilities on a short-term basis, usually at unfavorable rates. Generating too much power would be wasteful. Either too little or too much generation thus becomes a loss for the utility. As a result, weather forecasts have the potential to have economic impacts on utilities. Temperature, in particular, is an important weather variable. Electrical demand varies significantly with temperature because of the need for cooling in hot weather and heating in cold weather.

Here, we present the results of Brooks et al. (1997) and Douglas et al. (1998) for the various forecast systems and lead times. Of particular interest to us is the relationship of forecast quality and value for this user and the effect of forecast lead time on the value of the forecasts. These points give insight into possible important areas for concentration for specialized weather forecasts tailored for the utility.

2. ACCURACY OF FORECASTS

In order to ease comparisons with the output of the load forecasting model, we have computed the mean absolute error of the temperature forecasts for

Corresponding author address: Harold E. Brooks, National Severe Storms Laboratory, 1313 Halley Circle, Norman, OK 73069

each forecast day, combining the maximum and minimum temperature forecasts. Note that maximum forecasts in this data set are less accurate than the corresponding minimum forecasts. (For a more complete treatment of the verification of the forecasts, see Brooks et al. (1997).) A total of 321 days of observations with corresponding forecasts from all 5 sources with all lead times out to 5 days are available. The forecasts, except for the newspaper source, are all available in the early evening for forecast periods beginning with the overnight minimum temperatures. The newspaper forecasts are not available until the overnight hours, when the earliest editions are printed.

In order to highlight differences in the forecasts, we have chosen to use the NWS forecasts as a baseline for consideration. This is a reasonable approach since the NWS forecasts are available for use as an input into the media forecasts. It also mimics a potentially real scenario for a decision maker using weather information--a comparison of a private sector forecast for which they might have to pay with the "free" forecast available from the NWS. In the section that follows, then, it would be logical for that decision maker to determine how much a particular forecast system is worth to them.

As expected, the NWS forecasts are less accurate as the lead time increases (Fig. 1). The mean absolute error (MAE) increases from 3.3 •F to 6.6 •F as the lead time goes from 1 to 5 days, with the biggest increase coming between day 2 and day 3. This latter point may reflect the change in the nature of the guidance products available with the end of the valid period of MOS guidance from the NGM and the end of the valid period of the Eta model forecasts.

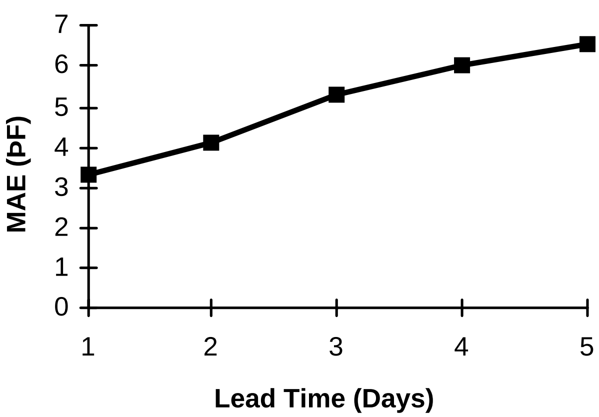

Fig. 1: Accuracy of NWS temperature forecasts by lead time expressed as Mean Absolute Error.

The media forecasts show that the MAE of those sources differs from the NWS by less then 0.4 •F, typically (Fig. 2). Note that two of the sources (forecast source [FS 2] and FS 3) are as accurate or more accurate than the NWS at all lead times, whereas the other two are more accurate only at long lead times. In fact, FS 2 is the most accurate system at all lead times,

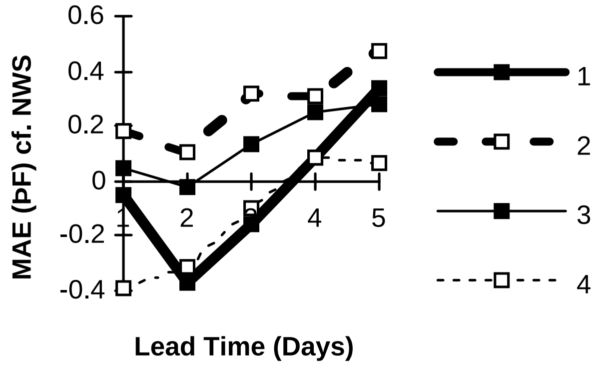

Fig. 2: Comparative accuracy of forecasts from four media sources. MAE from NWS forecasts subtracted out, so that positive values indicate that forecast source is more accurate than NWS at that lead time.

when the maximum and minimum temperature forecasts are combined. We wish to highlight the day 2 forecasts in particular. Tests of the significance of the difference between FS 1 and FS 2, using a bootstrapping technique, indicate that FS 2 is significantly more accurate (as measured by MAE) than FS 1 at a 95% confidence level. Also, note that the difference between FS 1 and FS 4 is only 0.06 •F at day 2.

3. VALUE OF FORECASTS

A similar assessment of the value of the forecasts from the different systems can be carried out using the utility load forecasting model. The "free" NWS forecasts provide between approximately $50,000 and $200,000 of value per year to the modelled utility, compared to using climatological information, as the lead time goes from 5 to 1 days (Fig. 3). The media

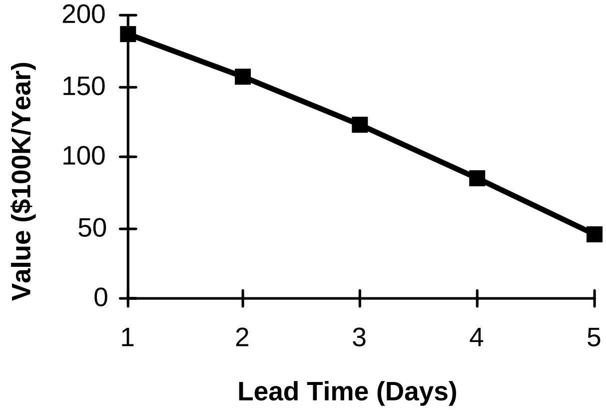

Fig. 3: Annual value of NWS forecasts compared to long-range climatological information for utility load application. Value in $100,000 per year.

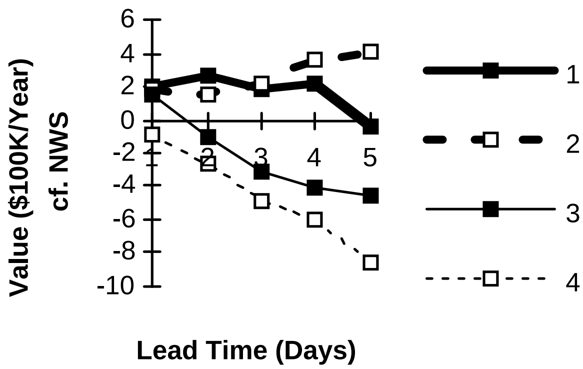

Fig. 4: Comparative annual value of forecasts from four media sources. Value of NWS forecasts subtracted out, so that positive values indicate that forecast source is more valuable than NWS at that lead time.

forecasts offer a considerable range of value (Fig. 4). It is particularly interesting to compare Figs. 2 and 4. Although FS 2 is more valuable than the NWS at all lead times, many other aspects are very different between the two figures. For example, FS 3 is less valuable than the NWS for days 3-5, despite being more accurate during that time period.

The most striking difference between the accuracy and value figures is at day 2 lead time. The *most* valuable forecast (to this forecast user) is that from FS 1, which was the *least* accurate forecast. In fact, mean expected losses from using FS 1 are on the order of $100,000 less per year than from using FS 2, even though FS 2 was statistically significantly more accurate, according to mean absolute error. In fact, the very small difference in accuracy between FS 1 and 4, with FS 4 being 0.06 •F more accurate is associated with more than $500,000 difference in value over a year, with FS 1 being the more valuable.

4. ACCURACY/VALUE RELATIONSHIP

The complex relationship between accuracy and value has been brought out before (Murphy 1993; Roebber and Bosart 1996). This case provides another example of this situation. By breaking down the forecasts into subsets, depending on the temperature, we can begin to get an idea of why this occurs. In particular, we will look at forecast performance during three temperature regimes, considering sensitivities that are relevant to the utility company. The regimes are separated by the mean daily temperature with cold days being defined as less than 65 •F, non-sensitive days between 65 and 75 •F, and hot days above 75 •F. Using this division for this data set, 63% of the days are cold, 24% are non-sensitive, and 13% are hot. The distinctions are based primarily on the demands for heating and cooling by customers.

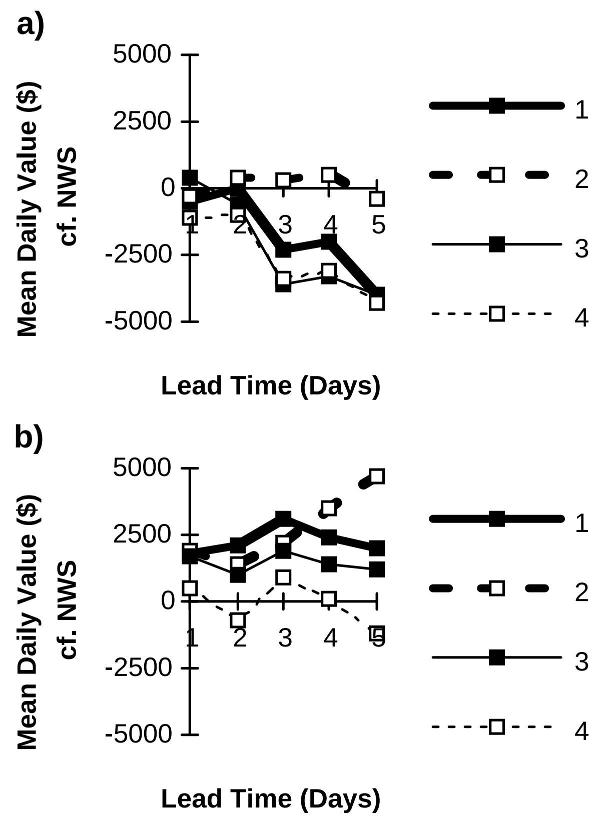

Fig. 5: Comparative average daily value of forecasts from four media sources. Value of NWS forecasts subtracted out, so that positive values indicate that forecast source is more valuable than NWS at that lead time. a) Cold days. b) Non-sensitive and hot days.

The relatively large value of FS 1 for the utility can be seen to be concentrated in the non-sensitive and warm days (Fig. 5). FS 1's warm forecasts are worth approximately $700/day more than FS 2 and $2800/day more than FS 4. The differences in value are much smaller at that lead time for the cold day forecasts and FS 2's cold day forecasts are more valuable than FS 1's, even though the differences in accuracy are larger (FS 2 has an MAE 0.6 •F lower than FS 1 for cold days, while the MAE for hot days is virtually equal). As a result, when considering the value of the forecasts for the entire year, the cold day forecasts have less of an impact for this application. The value of the forecasts are concentrated on the relatively small number (37%) of the cases when the temperatures are high. Detailed analysis of the difference in the forecasts will be carried out at a later date.

5. DISCUSSION

While the result that value and accuracy of forecasts are not the same thing is not a new one, this example is particularly dramatic in that there is a case in which the least accurate of a set of forecast systems is the most valuable for a real user. (We note in passing that we do not have any way of knowing what weather information, if any, the actual utility used during the course of this work, so that we cannot compare the results to actual use.) As with some of the examples of Roebber and Bosart (1996), the value of the forecast systems is concentrated in a relatively small number of forecasts. This also allows us to identify situations in which forecasters could potentially provide great value for the user. For example, it seems clear that improvements in the 3-5 day range could make a huge difference in value. Annual differences between the 1-2 day and 3-5 day forecasts are on the order of $5-10 million (see Fig. 3). Even the extreme difference in the value of different media sources at day 5 is on the order of $15 million. It seems obvious that a niche exists where both the forecaster and user could make significant financial gains in this arena. Similar analysis and consideration of the sensitivity of various users to weather information should allow forecasters to find situations in which they can maximize the effect their forecasts have on their users. It may provide an opportunity to focus marketing efforts on identifying customer need.

6. ACKNOWLEDGMENTS

We thank Arthur Witt, Mike Eilts, and J. T. Johnson for collecting the media forecasts. Oklahoma Gas and Electric and the Electric Power Research Institute provided support for the development and use of the utility load forecast model. Allan Murphy made suggestions about the interpretation of the results during a visit to the National Severe Storms Laboratory.

7. REFERENCES

Brooks, H. E., A. Witt, and M. D. Eilts, 1997: Verification of public weather forecasts available via the media. *Bull. Amer. Meteor. Soc.*, **78**, in press.

Douglas, A. P., A. M. Breipohl, F. N. Lee, R. Adapa, 1998: The impacts of temperature forecast uncertainty on Bayesian load forecasting. Submitted to 1998 IEEE Winter Meeting.

Katz, R. W., and A. H. Murphy, 1997: *Economic value of weather and climate forecasts.* Cambridge Univ. Press, 237 pp. (ISBN 0-521-43420-3)

Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281-293.

_____, and R. L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.*, **115**, 1330-1338.

Roebber, P. J., and L. F. Bosart, 1996: The complex relationship between forecast skill and forecast value: A real-world comparison. *Wea. Forecasting*, **11**, 544-559.