UNDERSTANDING USER DECISION MAKING AND THE VALUE OF IMPROVED PRECIPITATION FORECASTS

LESSONS FROM A CASE STUDY

BY THOMAS R. STEWART, ROGER PIELKE JR., AND RADHIKA NATH

A study of snow removal decision making shows that improved, ongoing, systematic data collection is needed to support studies of the use and value of forecasts.

Scientific research on atmospheric processes and weather forecasting frequently leads to improvements in weather forecasts [such as quantitative precipitation forecasts (QPF)], as measured by scientific criteria. Furthermore, effective use of such improvements has the *potential* for avoiding injury and death, averting property and environmental damage, and other societal benefits. Although the potential societal value of improved weather forecasts is substantial, the realization of that potential is not automatic (Pielke and Carbone 2002). The scientific community has a responsibility to work toward that realization.

The success of that work depends on a number of factors. Obtaining the actual (not just the potential or theoretical) value of improved forecast technology requires (a) a forecasting process that translates improved science and technology into improved forecast products that are targeted to user needs, (b) effective communication of forecast information to users in a timely fashion and in a form useful for making weather-information-sensitive decisions, and (c) users who incorporate the forecast product into their decisions in order to make better choices (on the processes of forecasts, communication, and use, see Glantz and Tarleton 1991; Piekle and Carbone 2002; Sarewitz et al. 2002).

Meeting these requirements begins with the detailed understanding of user needs and decision processes that is the outcome of systematic study. Such study also provides a foundation for studies of the value of current forecasts and forecast improvement. There is a large literature on forecast value (see, e.g.,

AFFILIATIONS: STEWART AND NATH—Center for Policy Research, Rockefeller College of Public Affairs, The University at Albany, State University of New York, Albany, New York; PIELKE—CIRES Center for Science and Technology Policy Research, University of Colorado, Boulder, Colorado
CORRESPONDING AUTHOR: Thomas R. Stewart, Center for Policy Research, Rockefeller College of Public Affairs, The University at Albany, State University of New York, Albany, NY 12222
E-mail: t.stewart@albany.edu
DOI: 10.1175/BAMS-85-2-223

In final form 22 July 2003

Katz and Murphy 1997; much of this literature is available in an online bibliography on "The Use and Value of Weather and Climate Forecasts" at the University of Colorado Center for Science and Technology Policy Research at http://sciencepolicy.colorado.edu/biblio/index.html).

SNOW REMOVAL CASE STUDY. This project began as an exploratory case study of the use of weather information for decision making and the economic value of improved QPF for a surface transportation activity in the northeastern United States. The study was exploratory because it focused on the use and value of actual and potential forecast products. Thus, we consider that the lessons of general relevance learned from the methodology, data collection, and analysis for the conduct of future studies of weather forecast use and value to be more important than the particular details of the case study itself.

We chose to study snow removal on the New York Thruway. Snow removal is important for thousands of miles of highway in the United States. The snow removal budget for the New York Thruway alone is approximately $10 million per year. The benefit to travelers in safety and convenience and the environmental costs due to the use of salt on the roadways (e.g., McKeever et al. 1998; Vitaliano 1992) may be many times that amount.

We obtained an understanding of the decisions involved in snow removal through interviews, observations, and other sources. We sought data from a number of sources for modeling the effect of improved QPF on the decision process and for estimating the value of an improved forecast. The difficulties of reaching satisfactory estimates based on this modeling effort, however, helped demonstrate a major lesson about the critical importance of forecast verification data, data on actual decisions, and cost data for supporting studies of the value of weather information.

Interviews. Snow removal is the responsibility of the 23 thruway section maintenance facilities that are distributed along the 496-mile thruway. We interviewed supervisors at 5 of 23 facilities during November 1998.[1] The five facilities were chosen to represent the various weather and traffic conditions along the thruway. The facilities chosen (with mile markers, measured from New York City) were 1) Buffalo (mile 423.19, urban, substantial snow, much of it from lake-effect processes), 2) West Henrietta (mile 362.44, a major exit for Rochester), 3) Verona (mile 252.71, mostly rural, snowiest station on the thruway), 4) Albany (mile 141.92, urban, beyond most lake-effect snow), 5) Harriman (mile 45.20, north of New York City, heavy traffic at times, less snow than farther north).

In the interviews, we asked about steps taken to prepare for the snow season, responsibility and authority for the decision to send out trucks, choices available when snow is threatening, sources of weather information, required lead times or action, timeliness and accuracy of feedback about results, types of errors made, rewards and incentives, and desired weather information.

Observations. We observed snow-fighting operations at the Albany maintenance facility during the 14–15 March 1999 storm. We also obtained weather reports from a network of trained spotters maintained by John Quinlan [National Weather Service (NWS) Albany, New York], as well as reports from drivers in several sections of the thruway during this storm.

Other information. We also examined the "logs" (i.e., journals recording driver activity) that drivers keep during snow removal operations and reviewed maintenance manuals and other documents related to winter highway maintenance. The staff at the maintenance facilities and the thruway headquarters in Albany was cooperative and supportive of our study.

Data on forecast quality and costs were obtained from thruway data on costs and operations, national accident data, toll barrier weather observations, Rapid Update Cycle (RUC) model runs (November 1998–March 1999), NWS storm data for New York, and New York climatological snowfall data from the National Oceanic and Atmospheric Administration (NOAA).

RESULTS—THE DECISION-MAKING PROCESS. Describing the decision-making process re-

[1] Interviews lasted 1–2 h and were recorded and transcribed. Interviewers used a list of open-ended questions as guidelines, but were allowed the freedom to ask follow-up questions and explore issues that had not been anticipated when the interviews were planned. Interviewees were those who had the title of supervisor I (the head supervisor) and, in the case of Buffalo and Albany, supervisor IIs also participated. These supervisors are responsible for snow removal actions taken before, during, and after a storm. We also interviewed the Albany division engineer who oversees maintenance for all facilities in his division. (The thruway is divided into four divisions.)

quires, at a minimum, describing the key decision makers, their goals and the context in which they operate (i.e., the decision environment), the information they use to make decisions, the alternative actions available to them, and the important decision points. These requirements are illustrated below for our case study.

Key decision makers. The key decision makers are as follows:

- *Supervisors.* Each thruway maintenance office has three supervisors who cover the three shifts each day. Supervisors are responsible for a section of the thruway (usually about 30 miles). They call in extra staff if needed and decide when and where to send out trucks to plow and salt.
- *Drivers.* Drivers are responsible for a specific route. They make observations and can make some independent decisions about when to plow and apply salt. Obviously, they do not plow if they do not see snow on the roadway. More experienced drivers are given greater freedom to decide how to treat their route. Drivers are in radio contact with supervisors and report weather conditions regularly.
- *System division engineer.* Each of the four thruway division engineers is responsible for several maintenance offices. Generally, they do not get involved in decisions about specific storms. After the storm has ended, they review the outcomes of each storm and can influence the supervisors' actions in subsequent storms.
- *Thruway headquarters staff.* Staff at thruway headquarters in Albany makes policy, determines the maintenance budget, approves equipment purchases, and makes other decisions that affect snow-fighting capability in the long run.

We chose to focus exclusively on the thruway maintenance supervisors. They have the decision-making responsibility for each storm, and their judgments and decisions are directly affected by weather forecasts.

Goals and decision environment. Understanding a decision process requires that we first address the question, "What are the decision makers trying to accomplish, and what are the opportunities and constraints that they have to cope with in order to achieve this goal?" Based on our interviews and observations, we conclude that the following describes the key elements of snow fighting.

The primary goal of snow fighting is to serve public safety and convenience by keeping the maximum possible traction on the roadway. Supervisors want to be proactive. They try to have trucks loaded and on the road before precipitation starts. The purpose of snow fighting is to keep the roadway clear by plowing and keeping the freezing temperature of the pavement below the actual pavement temperature by applying salt. If there is rain or snow present, salt must be applied (if possible) before the pavement temperature drops below the freezing temperature. This means, in effect, that the supervisor must implicitly predict pavement temperature. Due to uncertainty about the weather, some errors are inevitable. The imperfect link between air temperature and pavement temperature adds to the uncertainty.

A highly simplified illustration (Table 1) represents two possible types of errors that maintenance supervisors must consider in making decisions. This classic 2 × 2 decision table is familiar to students in introductory decision theory courses. It has been applied in studies of value of weather information by Murphy (1977) and recently by Smith and Vick (1994), Richardson (2000), Thornes and Stephenson (2001), Wilks (2001), and Mylne (2002). See Katz and Murphy (1997) for a summary of methods for estimating forecast value.

The two possible types of errors represented in Table 1 are (a) sending out trucks too soon or unnecessarily or applying too much salt; and (b) not sending trucks, sending too few trucks, sending trucks too late, or not applying enough salt when it is necessary. Given their primary goal, decision makers considered the second error much more serious than the first. Unfor-

Table 1. Simplified decision table describing the errors that decision makers must consider.

		Decision	
		Do not send trucks	**Send trucks**
Event	**Snow**	**Error [type (b)]** Costs: Risk of accidents and inconvenience	**Good decision** Benefits: Public safety and convenience Costs: Payroll, fuel, maintenance, salt
	No snow	**Good decision** Roads are safe No costs	**Error [type (a)]** Costs: Payroll, fuel, maintenance, salt

tunately, there are no data on which kind of error is more frequent, but it seems likely that it is type (a).

Supervisors cope with the unavoidable uncertainty in their job by being risk averse. They want to avoid type (b) errors if at all possible. Their snow removal budget, and the incentives and culture of the thruway encourage them to use salt, trucks, and personnel as necessary in order to avoid this type of error. This does not mean that they are not concerned about saving money and salt. Given the inherent uncertainty they have to deal with, it is better to risk several type (a) errors than to experience one type (b) error. As a result, the best opportunity to benefit from an improved weather forecast would be in the reduction of errors of type (a). In other words, there will be an economic benefit to improved forecasts if they result in savings associated with staffing, fuel, maintenance, and salt without sacrificing public safety or convenience.

Information used for decision making. Studies of the use and value of weather forecasts must recognize that decisions are rarely based on weather information alone (Stewart et al. 1984; Stewart 1997; Katz and Murphy 1997). They must also recognize that weather forecasts are often not the decision makers' only source of weather information. Furthermore, weather information is typically embedded in a "matrix" of other relevant information.

Since thruway supervisors are faced with decisions that involve uncertainty about the weather and their decisions can have serious consequences, they constantly seek weather information. They are aware of local weather conditions and forecasts during the snow season. As snow approaches, they use National Weather Service zone forecasts, Data Transmission Network (DTN) Corporation radar displays and other information, contacts with other thruway facilities, and personal observation to make decisions. Based on our interviews, it is clear that supervisors obtain weather information from a variety of sources, but they still want better, locally specific forecasts of storm onset, storm intensity, and storm duration.

At the time of our study, the DTN displays had been installed in each maintenance office for less than a year. The supervisors expressed enthusiasm for this system. The display they used most (in the case of some supervisors, the only display they used) was the Northeast regional radar (covering the northeastern United States). Most supervisors felt that they could use this display to make their own forecasts that were more specific and timely than the forecasts that had been previously provided by a private weather-forecasting firm. Their preference for the radar display, even though they are not trained in the subtleties of interpreting radar, is evidence of their desire for more specific local information than is generally provided by forecasters.

In addition to weather information, supervisors must consider the time of day and week that the storm occurs. Snow during rush hour may require special attention due to heavy traffic that both exposes more motorists to risk and can inhibit the movement of the trucks. Snow on weekends may require more lead time to call in extra drivers because weekend maintenance crews are generally smaller. In general, the supervisors must take into account the number of trucks and drivers that are available, as well as other local road conditions, such as bridges and overpasses that may need early and repeated treatment.

Figure 1 places the supervisors' decision process in the context of weather information and situational variables. We assume that improved QPF capabilities would affect the decision process through improved NWS forecasts or by being incorporated into a weather information system such as the DTN system. It is also possible that improved QPF might have a direct effect on the decision process. Clearly, any evaluation of improved QPF must be made in the context both of other information that is available to the decision maker and of the situational factors (e.g., time and location of storm) that affect their decisions.

Overview of decision process. Snow plowing involves a dynamic decision process. The decision to devote resources to snow fighting, and the level of those resources is made continuously preceding and during the storm. The outcomes of decisions made at one time will depend on weather events and will affect decisions made at future times.

Understanding the supervisors' decision process involves understanding what alternatives they have available, the multiple decision points in the process, and the information that they use in making decisions at each point in the process. These are described in the following sections.

Decision points. Dealing with snow on the thruway begins with preseason preparation, for example, setting up trucks with spreaders, plows, and wings, and training drivers. These activities are important, but are not sensitive to weather forecasts (although they may be sensitive to seasonal forecasts).

Action for a specific weather event is triggered by a forecast of both cold weather and precipitation (risk of snow or ice on roadway). This usually occurs a few days in advance of the storm and the information may

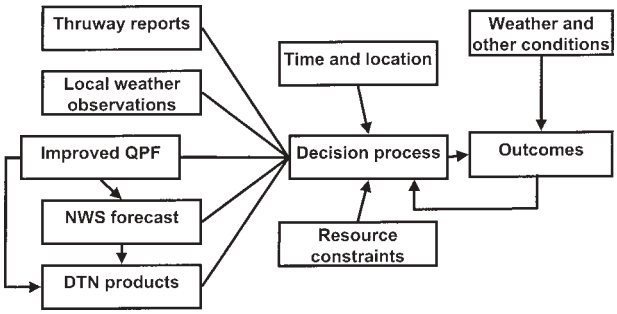

Fig. 1. Overview of weather information and situational variables that affect snow-fighting decision making and the role of improved QPF.

come from a commercial radio or TV broadcast, the DTN terminal, or other means of communication. The supervisors constantly monitor weather information. The supervisors' actions at this lead time include increased vigilance, planning, and preparation (e.g., checking equipment and supplies).

The alternative actions that the supervisor must consider (Table 2) vary as the storm progresses. Benefits of improved forecasts are likely to be obtained by influencing the decisions most sensitive to weather information, that is, those involving adding crew, sending out trucks, plowing and salting, and stopping operations. The information used, actions, lead time, and costs and benefits associated with each step in Table 2 are described in Table 3.

Based on our interviews and observations, several conclusions about the nature of the dynamic decision processes involved in snow fighting can be drawn. First, thruway supervisors tend to operate intuitively rather than analytically (Hammond 1996). As a result, they cannot always describe in precise detail their rules for initiating action when snow threatens. Modeling their decision processes requires observations and experiments. Interviews alone will not suffice.

Second, although supervisors are not sophisticated about the intricacies of weather forecasting, they have considerable expertise; they are experienced and vigilant. Their goal is to have trucks out on the road well before a storm starts, and they are willing to bear the costs of "false alarms" in order to avoid being caught by a storm and not responding quickly enough.

Third, supervisors must, implicitly or explicitly, make their own forecasts of a critical variable—pavement temperature. At the time of our study, no one provided this information for them. Currently the technology exists, several vendors sell such services, and the Road Weather Maintenance Decision Support System Project at the National Center for Atmospheric Research (NCAR; www.rap.ucar.edu/projects/rdwx_mdss/index.html) includes road surface temperature forecasts. It is obvious that pavement temperature is not the same as air temperature, and the treatment of the roadway, particularly the application of salt, depends on pavement temperature.

The skill of the predictions of pavement temperature and other weather factors made by the supervisors is critical to valuing forecasts, and is unknown.

In an effort to improve decision making and to reduce costs and environmental damage, technology is changing. For example, Road Weather Information Stations (RWIS) are being tested and may be installed. Alternatives to salt that are less corrosive, but more expensive, are also being considered. "Zero-velocity spreaders," which keep salt from scattering off the road are being tested. At least one decision support system for snow fighting is under development.

Since NWS forecasts are embedded in a web of other, related, information (Fig. 1), some of which is more timely and site specific than can be provided by NWS, extracting the impact of an improved forecast is complicated, even in this relatively simple case. The likelihood that improved forecasts will be implemented in a future when the decision process itself fundamentally changes, for example, with new technology that presents different opportunities and costs to the supervisor, further complicates the problem.

Table 2. Summary of alternative actions available to decision makers.

- When snow is anticipated, the supervisor may need to call in extra drivers. He must decide when to do so and how many to call.
- Before plowing has started, the supervisor must consider a number of decision alternatives:
  - ✓ Do nothing
  - ✓ Send a reduced number of trucks
  - ✓ Send out all trucks
- After plowing has started, alternatives are as follows:
  - ✓ Keep trucks out
  - ✓ Recall some trucks (or keep them in when they come in to refill with salt)
  - ✓ Recall all trucks
- During plowing and salting, it is possible to modify the route and the amount of salt applied.

Table 3. Information, actions, lead time, and costs and/or benefits for each decision point.

Decision point	Information	Actions	Lead time	Costs and/or benefits
Vigilance	Forecast (NWS zone forecast), observations, DTN radar displays.	Check equipment, including lights, oil, fuel, brakes, plows, cutting edges, etc.	Up to 8 h	No additional direct costs, but people may be diverted from other jobs, possibly resulting in work delays
Add crew if necessary	Forecast (NWS zone forecast), DTN radar display, reports from other thruway facilities (which are polled hourly), observation	Take crews off other tasks, call in extra crew, occasionally trucks can be called in from neighboring sections.*	30–120 min, depending on where people have to travel from	Payroll, overtime
Load and send trucks out	Observation by supervisor, DTN radar display, forecast	Load trucks with salt, dispatch to preassigned route, stand by if necessary	To load: 15 min to reach route: 0–30 min	Fuel and maintenance
Plow and apply salt (or other material)	Observation by drivers and supervisors on patrol (some supervisors have pavement temperature sensors), DTN radar display	Set rate of salt application, change rate of application during storm, prewet salt, use Magic Salt (a product that is more environmentally friendly, but also more expensive than regular salt)	0	Salt and fuel and maintenance, vehicle corrosion ($113 ton^{-1}), highway structure corrosion ($615 ton^{-1}) for bridge repairs), aesthetics ($75 ton^{-1} for tree damage in Adirondack park region), health damage from sodium in drinking water (speculative; estimates from Vitaliano 1992)
Stop applying salt	Observation by drivers/ supervisor, DTN radar display	Stop spreading	0	Salt savings
Recall trucks	Observation by drivers/ supervisor, DTN radar display	Recall trucks to maintenance office	0	Fuel and maintenance
Send extra people home	Observation by drivers/ supervisor	Send people home, assign them to other jobs	0	Potential savings in labor costs
Clean up	Observation of snow on thruway	Empty spreaders, plow shoulders, median, etc.	0	Payroll, fuel, maintenance

*Actions are constrained by union requirement of a minimum 4-h shift for crew that is called in. A minimum 2-week notice is required to change someone's shift without having to pay them extra, virtually eliminating the possibility that shifts can be rearranged to accommodate a storm.

ATTEMPT TO CREDIBLY ESTIMATE THE VALUE OF AN IMPROVED FORECAST. From the case study, we learned about the decision processes of the thruway supervisors, and the context in which their decisions are made. This is but one component required for an estimate of forecast value.

Estimation of the value of weather forecasts is a special case of the problem of valuing information (see

Katz and Murphy 1997 for a detailed discussion of the economic value of weather information). Decision theory provides an elegant method for making estimates of the value of information, but the validity of the resulting estimates rests on assumptions that are not always valid. The shakiest assumption of decision analytic models is that users of the information act optimally based on the information they are given. *Descriptive methods* are needed to incorporate realistic assumptions about decision-making behavior into models for estimating forecast value. Our goal was to go beyond traditional economic models, exemplified by the snow removal models of Howe and Cochrane (1976), Smith and Vick (1994), and Thornes and Stephenson (2001), and to base our estimate on a detailed case study of current practice. That case study has been described above. Studying a specific decision problem, limited to a particular geographical region, provides the greatest likelihood to accomplish this goal.

Estimating the value of an improved forecast requires an estimate of the value of the current forecast as a baseline. Once that baseline is established, deriving an estimate of the value of an improved forecast requires a prediction of how the improved forecast will be used in decision making, and a prediction of any future changes in technology that might affect decision making, among other factors. Such predictions are inherently clouded by uncertainty. Any model for estimating the value of an improved forecast therefore requires four components:

1) a decision-making model capable of mapping both current and improved forecasts, along with other important information, into decisions and, ultimately, actions—this requires data on forecasts (i.e., verification) that can in fact be mapped into decision processes;
2) a detailed model of current and future forecast skill that specifies both the distribution of forecasts and the conditional probability of specific weather events, given specific forecasts—the closer the match of specific weather events with decision criteria, the better the potential linkage of weather information with its decision value;
3) a model of outcomes resulting from combinations of actions and subsequent weather events, under a variety of scenarios (e.g., storms, traffic levels, section of roadway)—this model would include the effect of alternative decisions on traffic flow and accident risk, and the fate and environmental and health effects of salt used on the roadway;
4) a detailed model for valuing outcomes—this would include not only estimates of cost reduction, but also estimates of the value of benefits such as reduced travel time and reduced accident rates.

We will summarize the data gathered for each of these model components here. There were serious complications at each step due both to the complexity of the problem and the lack of datatsets both suitable for analysis and large enough to provide reliable estimates.

Decision-making model. Our qualitative description of the decision-making process (outlined above) was a necessary foundation for developing the quantitative model needed for calculating forecast value. Such models can be derived from the record of decisions made if that record also contains sufficient detail about the conditions and other information at the time the decision was made. Appropriate data were not available (they rarely are). A second approach is based on simulated decisions based on hypothetical situations (see Cooksey 1996). That would require more time and resources than were available for this study.

Model of current and future forecast skill. A basic requirement of any model of the value of an improved forecast is a quantitative estimate of that improvement. Furthermore, the detail of the estimate of skill must match the detail of the decision model. Such an estimate did not exist at the time we began the project, and it proved exceedingly difficult to obtain.

In a major effort to assess current skill in the vicinity of the thruway, we merged a set of RUC 3-h forecasts provided by B. Schwartz of NOAA's Forecast Systems Laboratory (FSL) with corresponding hourly observations made by toll barrier operators that were provided to us by the thruway authority. This resulted in a dataset of 4532 forecast–observation pairs between 20 November 1998 and 31 March 1999. Since the forecasts and observations were not strictly comparable, assumptions had to be made in order to use the toll barrier observations to verify the RUC model forecasts. The forecasts were 3-h accumulations of total precipitation (ptot) ending at the time of the observations. The observers used verbal categories such as clear, flurries, light snow, rain, etc. After some experimentation, we categorized the RUC model output into two categories: "no precipitation" (ptot = 0) and "any precipitation" (ptot > 0). We also combined the verbal categories used by the toll barrier observers into two categories: "clear" versus "rain

or snow." (Note that, for toll barrier observers, clear describes the conditions of the road, not the sky.) Based on these categories, we were able to calculate an equitable threat score of 0.25 and a Brier skill score of 0.403.

Although this analysis consumed many months, the results were unsatisfying. The data on current forecasts lack the detail we need to develop a model. For example, we have data for precipitation and snow amounts for only one lead time (3 h), but the actions of supervisors require lead times of 0–8 h, depending on the action, as well as forecasts of storm onset and duration. In addition, we need a finer-grained analysis (more categories of forecasts corresponding to more categories of observations) in order to represent the forecasts that are used by the supervisors. We would also need to examine the data for more than one season. Furthermore, projections of future improvements in skill, resolution, and dissemination techniques are purely speculative. This leads to an important conclusion: before we can be confident about projections of future skill, a number of questions about *present* skill in forecasting onset, severity, and duration of snow and ice events relevant to decision making remain to be addressed.

In order to model both forecast skill and snow-fighting activity, we needed a climatology of storms that affect the thruway. We found no appropriate climatological studies, so we developed one from toll barrier weather observations. It was surprising to us that data on the climatology of the thruway and data on the skill of current forecasts were not only unavailable, but proved very difficult to generate (see appendix for methods and results of this effort). We had assumed when we started the project that obtaining weather data necessary to assess forecast value would be relatively easy, and that we would concentrate on our strength, which is studying the way information is used. It turned out that much of our time and resources were consumed in obtaining and analyzing weather data (with disappointing results) in order to make it suitable for our analysis. We strongly recommend a review of weather data archiving procedures and verification studies for supporting future studies of the use and value of forecasts. Such data and verification must be at the scale of the decision process in order to effectively contribute to forecast value studies grounded in reality, rather than assumptions.

Model of outcomes resulting from combinations of actions and subsequent weather events. We adopted a scenario approach to modeling outcomes. A scenario is constructed by specifying (a) a section of roadway, time of day, and week; (b) storm intensity and duration over time; (c) a forecast of storm intensity and duration, which is updated regularly; (d) a model of the supervisor's response to the forecast and the current weather; (e) assumptions about technology used; and (f) a model of the outcome of various combinations of storm intensity and snow fighting activity.

Because of the variety of types of storms, road sections, and situations that affect the value of forecasts, many different scenarios would be required. Not only would storms of various durations and intensities have to be simulated, the conditions under which those storms occur have a significant effect on the supervisors' actions. For example, weekend storms are treated differently from those occurring during the week. Storms in urban areas during rush hour are treated more aggressively than the same storm in a rural area at night.

If a representative set of scenarios could be constructed, and each scenario evaluated (assuming appropriate data were available with the requisite level of detail), then the results for each scenario could be appropriately weighted (weights depending on the relative frequency of the scenarios) and combined into an overall estimate of value. We realized that we would not be able to construct and evaluate all possible scenarios, so we focused on constructing a few important ones. During the research, it became apparent that the data needed to evaluate even one scenario were lacking.

Evaluating outcomes. We were able to obtain some information that would be useful in valuing outcomes from thruway information, Federal Highway Administration data, and other sources. A paper by Vitaliano (1992) on the social costs of highway salting was particularly helpful. Information on a number of critical factors was not found, including the role of weather in accidents on the thruway or similar highways, exposure to risk of accidents (i.e., per vehicle miles traveled) under different weather conditions, weather effects on travel time, and the improvements in accident rates and travel time resulting from snow removal. In addition, the historical weather record is based on a sparse network and does not include pavement conditions. The effect of changing technology on the value of outcomes is very difficult (perhaps impossible) to accurately predict. As is true for many activities, it is easier to quantify the costs of snow fighting than the benefits.

Summary of modeling effort. The limitations of the data we were able to collect included (a) the only data avail-

able were secondary data that were not collected for modeling; (b) the sample sizes of the datasets we generated were small (one winter is a small sample); (c) datasets are on different geographic scales (thruway section, county, NWS region, NWS station, toll barrier), so it is necessary to make interpolating assumptions in order to combine sets; (d) similarly, the datasets are on different time scales (days, hours, storm), and (e), the data included significant gaps, missing data, and inconsistencies.

Due to the lack of relevant and useful weather information, the complexity of the decision process, and uncertainty about potential outcomes and their value, there is a substantial gap between the available data and the data needed to establish reliably the value of an improvement in QPF. An estimate of value would therefore have to be based in large part on untested assumptions and judgments, and would have an extremely wide confidence interval if, indeed, a confidence interval could be obtained. Such assumptions could easily be modulated within the bounds of credulity to result in an extremely wide range of outcomes. These factors led us eventually to abandon the goal of estimating in any meaningful way the dollar value of improved QPF.

CONCLUSIONS. The case study results provide insights, albeit qualitative ones, about the use and value of forecasts. We can make the following two statements about the potential value of improved QPF forecasts. First, improved forecasts could help supervisors achieve their goal of being proactive. They want to avoid deicing (treating a road that is already slick) in favor of anti-icing (preventing the road from getting slick in the first place). Second, improved forecasts could result in improved allocation of resources. Trucks could be sent to where they are needed and salt applied where it is most needed and not applied where it is not needed, with improved efficiency of operations and reduced use of chemicals, labor, and equipment.

Beyond the potential to improve the efficiency of thruway operations, improved forecasts have benefits for the end users of the thruway—travelers on the road. These benefits include improved traffic flow during bad weather, reduced fuel consumption, reduced accident rates, reduced environmental impact, and reduced health impact.

The data (including meteorological data) necessary to support credible and defensible quantitative estimates of the value of improved forecasts are lacking, returning us to where we had begun with forecast value conclusions dependent almost entirely upon basic assumptions and not the empirical data of the case. Responsible estimates would have confidence bounds so wide as to be useless.

An anonymous reviewer observed that "many of the problems encountered would not be such roadblocks if adequate resources (both time and funding) were allocated to this line of research." We agree completely. However, the level of investment required to overcome the problems encountered in this study would all but certainly exceed the value of having credible estimates of the value of improved forecasts for a single case study of the New York Thruway system. The experiences in this project point to a systemic problem in developing credible estimates of forecast value as the obstacles encountered in this project are typical of those found in studies of forecast value (compare Katz and Murphy 1997).

Overwhelmingly, because resources are limited, researchers can only avoid such obstacles by relying on assumptions rather than empirical data. Reliance on such assumptions detracts from the real-world applicability and meaningfulness of forecast value studies. But this circumstance is a consequence not only of the limited resources devoted to forecast value, but also to the considerable effort that would be required by both atmospheric scientists as well as forecast value experts to change procedures for the collection, analyses, and distribution of weather data and forecasts to be commensurate with the scale of the decision processes in question. This effort would represent a marked change in approach to verification from the present approach of considering weather data and forecasts in terms of phenomena and forecasts but with little or no consideration of how that data might be used in the context of understanding particular decision processes.

To realize the potential of improved forecasts would require careful attention to the forecast products themselves (and their production and dissemination through the public and private sectors) *as well as* the role of forecast products in the specific context of actual decision-making processes. The major implication of this study is that a more holistic approach to understanding and realizing forecast value is needed, that is, one in which information (both of forecast skill and usage) centered on the decision process is collected in a much more intensive manner than is presently the case. Just as meteorologists need detailed knowledge of atmospheric conditions to project atmospheric conditions, so too do students of forecast value need detailed observational knowledge in order to develop realistic expectations for changes in forecast value under alternative courses of action.

Absent such detail, studies of forecast value will simply be extensions of the assumptions brought to bear upon the study, which then runs the risk of the tailoring of such studies to suit particular desired outcomes.

One international effort has great potential to provide such a holistic approach. The World Meteorological Organization (WMO) World Weather Research Programme (WWRP) recognizes in its pro forma for the development of proposals for research and development projects and forecast demonstration projects under its purview that advances in science and technology "will lead to societal benefits if and only if that research is successfully turned into products that are used by decision makers. Thus, the societal aspects of weather are an essential area of complementary research." The WWRP has recommended four areas of investigation necessary to gain WMO approval of projects under the WWRP.

- *Obtaining an improved understanding of the nature of the problem and the opportunity.* These include the costs of weather-related events and who incurs those costs. Results obtained from this research, in conjunction with knowledge of predictability, etc., can help scientists to more effectively prioritize research objectives. More broadly, such research can provide information to help policy makers focus national priorities.
- *Use of forecasts by decision makers.* Even the most accurate forecast is of little value if it is not well used. To this end, it is important to understand what information decision makers could effectively use and also effective ways to communicate that information. Research in this area can help to identify those conditions necessary and sufficient for forecasts to contribute to the needs of decision makers.
- *The process of transitioning research to the operational community.* This process focuses on the needs of forecasters seeking to provide information of use to decision makers. Both the structure of the process and the content of the information being transferred should be evaluated from the standpoint of the penultimate goal of producing useful products. Thus, in addition to research on the use of forecasts, appropriate research might include the institutional structures through which the transfer process takes place.
- *Evaluation of forecasts.* There are many measures of forecast "goodness". Such evaluations are an important component of the program's ability to assess progress with respect to its goals.

The results of such research have significant potential to inform decision makers who seek to prioritize public expenditures among myriad competing objectives, of which weather is but one, as well as those decision makers within the weather community who face difficult choices among many meritorious competing priorities. Absent such information, weather-related priorities will likely be set in an ad hoc fashion, perhaps based on the self-interests of those involved in the weather policy process, potentially limiting the ultimate societal benefits associated with investments in research and technology (cf. Pielke and Carbone 2002).

Thus, the major conclusion of this study is guidance to the meteorological community as to what sorts of data and methods would be necessary to improve the knowledge base of forecast use and value. If the potential societal value of meteorological research is to be approached, much less realized, and if knowledge of forecast value is to play anything more than a minor role in weather policy decisions, then an ongoing effort to support studies of forecast use and value will be required.

ACKNOWLEDGMENTS. The research reported here was supported by a grant from Forecast Systems Laboratory (FSL), NOAA/Oceanic and Atmospheric Research. We thank Sandy McDonald for his support and encouragement. The opinions expressed here are those of the authors and do not necessarily reflect those of NOAA or FSL. We want to express our thanks to the thruway maintenance supervisors and the Albany division engineer who consented to be interviewed for this study. Bob Puzier was extremely helpful in obtaining access to the supervisors and in providing us with data maintained at Thruway Authority Headquarters in Albany. We are also grateful to George Tanner and Arthur O'Donnell of the Thruway Authority for their council, support, and encouragement. Dick Westergard, of the Albany NWS Weather Forecasting Office, provided invaluable weather advice and access to data. He also accompanied us on one of the interviews. John Quinlan kindly activated his spotter network for us during the 14–15 March 1999 storm. We thank Tom Schlatter of FSL for his support in obtaining RUC data, and Barry Schwartz for designing and supervising the model runs. We are grateful to Mary Downton of the Environmental and Societal Impacts Group at NCAR for extracting New York climatological records. Laurie Johnson provided advice on potential sources of economic value of forecasts. Tony Sambuca of the New York State Department of Transportation generously shared his knowledge of winter maintenance and provided useful state documents. Ann McCartt of the Institute for Traffic Safety and Management at the

University at Albany also provided valuable advice and helped us avoid extensive searches for data that do not exist. Joann Orologio patiently transcribed the interviews and entered large amounts of data from toll barrier observations. An earlier version of this paper was presented by the first author at the WMO WWRP Quantitative Precipitation Forecasting Verification Workshop, held in Prague on 14–16 May 2001. We are also grateful to the anonymous reviewers who provided many helpful comments.

APPENDIX: ANALYSIS OF WINTER WEATHER NEAR THE NEW YORK THRUWAY. The New York State Thruway Authority made available to us the toll barrier weather observations made by the four thruway divisions for part of the winter 1997–98 season and the entire winter 1998–99 season.

For the purposes of this analysis we focused on the toll barrier reports for the 1998–99 winter season. The toll barrier reports provide hourly information about air temperature, pavement temperature, and other significant weather factors. These reports were filled out by hand by the toll barrier staff. Notations for information other than temperature were not always entered and were sometimes impossible to decipher. Cut-offs for rain versus snow versus light snow versus sleet are based on the judgment of the observer, thus adding a dimension of subjectivity to the reports.

A preliminary visual examination of the toll barrier reports for the winter 1998–99 season provided us with a list of dates when significant weather events may have taken place. The selection of these events was based on an extended notation of rain, snow, or similar weather events over many hours and sometimes days. Events that seemed to last only a few hours were excluded.

As a next step storm data publications were gathered from the NWS. These storm reports list all extreme events for the United States. We focused on the data dealing with New York State for the months ranging from November 1998 to March 1999. These data are organized by the regions covered by each Weather Forecasting Office (WFO). New York State is divided into five regional WFOs.

Forty-seven storms were identified on both our preliminary search of toll barrier data and the NWS storm data. There were about eight storms listed on the NWS storm data publication reports that did not show up in our examination of the toll barrier data. Conversely there were about 30 possible snow/rain events on the toll barrier data that did not show up in the NWS storm data publications. This discrepancy is not surprising because of the subjectivity in storm identification and the different geographical coverage of the data sources.

The events that were picked on the basis of their presence in both lists were then examined. According to the NWS' categorization, the events ranged from winter/wintry mix, freezing rain, ice storm, snow, heavy snow, to blizzard. Snowfall amounts were not always listed. Where listed they did not follow the same format; for instance, in some cases averages were given, whereas in others specific snow amounts for specific locations were provided. In some cases no amounts were provided.

Figures A1 and A2 describe selected results from this analysis. Figure A1 shows the frequency distribu-

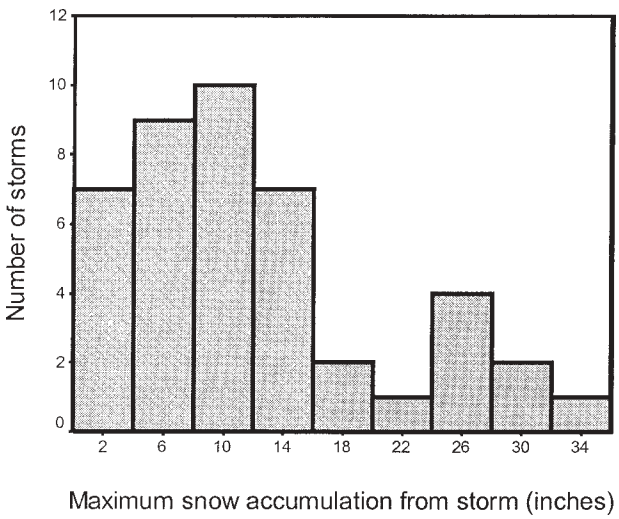

Fig. A1. Summary of maximum snow accumulation from 43 storms occurring near the New York Thruway in the winter of 1998–99.

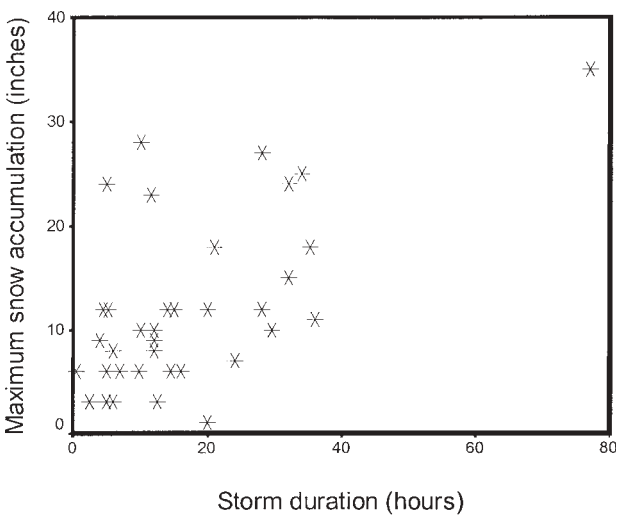

Fig. A2. Scatterplot of the relation between storm duration and maximum snow accumulation for 37 storms that had data on both variables.

tion for maximum snow accumulation for the 43 storms with available data. The maximum accumulation during this period was 35 in. (from a storm in western New York between 22 and 25 December 1998), and the median was 10 in. Figure A2 is a scatterplot showing the relation between duration and accumulation for the 37 storms having data on both variables. The correlation between duration and accumulation is 0.601. However, when the 77-h outlier storm is dropped, the correlation drops to 0.434. The slope of the regression line for all 37 storms is 0.341 in. h^{-1}. Dropping the 77-h storm reduces the slope to 0.303 in. h^{-1}.

Data for one winter obviously do not represent climatology. We also selected data from a CD put out by NOAA/National Climatic Data Center (NCDC), "Cooperative Summary of the Day," TD3200–Period of record through 1993 (Vol. 17, New York and Pennsylvania). The data were extracted and analyzed by Mary Downton at the National Center for Atmospheric Research. These data are organized by daily snowfall rather than by storm, so they are of limited use without considerable additional work (such as cross matching with surface analyses) for creating a storm climatology. We used them to determine whether the winter of 1998–99 was atypical. Figure A3 is an example of a comparison between data for Albany accumulated over the winters of 1939–94 and the winter of 1998–99 (daily snowfall amounts obtained from NWS records). The curves follow each other closely, but indicate that the lower snowfall amounts (2–4 inches) were more prevalent in 1998–99 than was typical during the earlier period. Thus, any simulations would have to be corrected to reflect this bias.

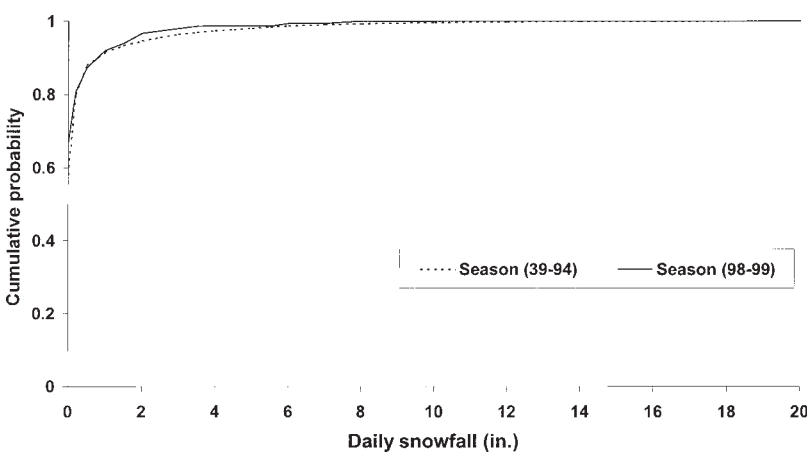

Fig. A3. Cumulative probability of daily snowfall amounts in Albany. A comparison of 1939–94 and 1998–99.

REFERENCES

Brier, G. W., and R. A. Allen, 1951: Verification of weather forecasts. *Compendium of Meteorology,* T. F. Malone, Ed., Amer. Meteor. Soc., 841–848.

Cooksey, R., 1996: *Judgment Analysis: Theory, Methods, and Applications.* Academic Press, 407 pp.

Glantz, M., and L. Tarleton, 1991: *Mesoscale Research Initiative: Societal Aspects.* National Center for Atmospheric Research, 35 pp.

Hammond, K. R., 1996: *Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error, Un-avoidable Injustice.* Oxford University Press, 436 pp.

Howe, C. W., and H. C. Cochrane, 1976: A decision model for adjusting to natural hazard events with application to urban snow storms. *Rev. Econ. Stat.,* **58,** 50–58.

Katz, R. W., and A. H. Murphy, Eds., 1997: *Economic Value of Weather and Climate Forecasts.* Cambridge University Press, 222 pp.

McKeever, B., C. Haas, J. Weissman, and R. Greer, 1998: Life cycle cost-benefit model for Road Weather Information Systems. *Transp. Res. Rec.,* **1627,** 41–48.

Murphy, A. H., 1977: The value of climatological, categorical and probabilistical forecasts in the cost-loss situation. *Mon. Wea. Rev.,* **105,** 803–816.

Mylne, K. R., 2002: Decision-making from probability forecasts based on forecast value. *Meteor. Appl.,* **9,** 307–315.

Pielke, R. A., Jr., and R. Carbone, 2002: Weather forecasts, impacts and policy: An integrated perspective. *Bull. Amer. Meteor. Soc.,* **83,** 393–403.

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **126,** 649–667.

Sarewitz, D., R. A. Pielke Jr., and. R. Byerly, Eds., 2000: *Prediction: Science,Decision-Making and the Future of Nature.* Island Press, 405 pp.

Schaefer, J. T.,1990: The critical success index as an indicator of warning skill. *Wea. Forecasting.,* **5,** 570–575.

Smith, K., and S. D. Vick, 1994: Valuing weather radar benefits for winter road maintenance: A practical case example. *Meteor. Appl.,* **1,** 173–177.

Stewart, T. R., 1997: Forecast value: Descriptive decision studies. In R. W. Katz & A. H. Murphy

(Eds.), *Economic Value of Weather and Climate Forecasts.* Cambridge University Press, 147–181.

Stewart, T. R., R. W. Katz, and A. H. Murphy, 1984: Value of weather information: A descriptive study of the fruit frost problem. *Bull. Amer. Meteor. Soc.,* **65,** 126–137.

Thornes, J. E., and D. B. Stephenson, 2001: How to judge the quality and value of weather forecast products. *Meteor. Appl.,* **8,** 307–314.

Vitaliano, D. F., 1992: An economic assessment of the social costs of highway salting and the efficiency of substituting a new deicing material. *J. Policy Anal. Manage.,* **11,** 397–418.

Wilks, D. S., 2001: A skill score based on economic value for probability forecasts. *Meteor. Appl.,* **8,** 209–219.