

**Examining the use of weather forecasts in decision scenarios: Results from a U.S. survey
with implications for uncertainty communication**

Rebecca E. Morss*

Jeffrey K. Lazo

Julie L. Demuth

National Center for Atmospheric Research, Boulder, CO

Submitted to *Meteorological Applications*

5 November 2009

*Corresponding author address:
Rebecca E. Morss
National Center for Atmospheric Research
P.O. Box 3000
Boulder, CO USA 80307-3000
Phone: 1-303-497-8172
morss@ucar.edu

Abstract

In order to communicate weather, climate, and hydrological forecast information effectively, it is important to understand how intended audiences are likely to interpret the information and use it in decision making. Such understanding is lacking, particularly for forecast uncertainty information. This article addresses these issues in the context of weather prediction, using empirical data from decision scenario questions asked in a nationwide U.S. survey. In the first set of decision questions, respondents were asked their probabilistic threshold for taking action to protect against potential rain or frost. In the second set, respondents were asked to make yes/no protective decisions in a potential reservoir flooding or fruit frost scenario given different types of forecast information, including deterministic (single-value) forecasts and forecasts that conveyed uncertainty. The analysis examines respondents' interpretations of different types of forecasts and their ability to use uncertainty forecast information in decision making. It also examines the extent to which respondents made decisions consistent with the cost-loss decision model. Further, the decision question responses are analyzed in conjunction with other survey responses to explore how respondents' information use relates to other measures of their perceptions and interpretations of forecast uncertainty. The results add to fundamental knowledge about people's interpretations and use of weather forecasts — especially forecasts containing explicit uncertainty information — and provide a starting point for future related work using survey and experimental approaches.

Keywords: probability of precipitation, quantitative precipitation forecasts, temperature forecasts, flooding, frost, forecast interpretation

1. Introduction

The goal of most hydrometeorological forecast communication is to provide information that people can interpret and use beneficially. Communicating forecasts effectively therefore requires understanding how people interpret and use forecast information presented in different ways. As capacity to estimate uncertainty in hydrometeorological forecasts has increased, the weather, climate, and hydrology communities have become particularly interested in effective communication of forecast uncertainty (e.g., Manning, 2003; Ryan, 2003; NRC, 2003, 2006; Demeritt et al., 2007; AMS, 2008; Budescu et al., 2009). Although hydrometeorological forecast interpretation and use has been studied in some real-world cases, there has been limited empirical study across broader populations. To help fill this knowledge gap, this article examines how members of the broad U.S. public interpret and use different types of weather forecasts, including those containing uncertainty information, based on their responses to decision scenario questions.

The article focuses on data from two sets of decision questions included in a nationwide survey of the U.S. public's perceptions, interpretations, uses, and values for weather forecast information (Morss et al., 2008; Lazo et al., 2009; Demuth et al., in preparation). The first set asked respondents their threshold, in terms of a percentage chance of rain or temperature below freezing, for taking protective action in a picnic or garden scenario. The second set asked respondents to use various quantitative precipitation or temperature forecasts to make binary (yes/no) protective decisions in a potential reservoir flooding or fruit frost scenario. The latter decision scenarios included a protective action component involving monetary costs and an outcome component involving monetary losses, similar to the cost-loss decision model employed in some studies of hydrometeorological forecast use and value. Although the decision scenarios tested here are simpler than most real world decision settings, examining use of information in a controlled context allows for replicability and more focused study.

The binary decision questions test forecasts that are either single-value (deterministic) or convey uncertainty in one of two ways: as a range of possible values or a percentage chance of

exceeding a damage threshold. The forecast information is presented using fairly simple, text-based formats, designed to allow examination of fundamental aspects of forecast interpretation and use. These findings can then be built on in research with more complex communication formats, including graphs and icons.

We examine respondents' stated probabilistic thresholds for taking protective action in the garden and picnic scenarios and their use of different forecast information in the reservoir and fruit scenarios. The results provide insight into respondents' interpretations of deterministic and uncertainty forecasts and their ability to use uncertainty information — including probabilistic forecasts — in decision making. They also indicate the extent to which respondents made decisions according to the cost-loss model. Each set of decision questions is tested in two scenarios to allow an initial examination of how results are similar and different across contexts. Further, by analyzing the decision question responses in conjunction with responses to other questions on the survey, we explore relationships between respondents' perceptions and interpretations of forecasts and their forecast use.

The decision questions are related to experimental economics and psychology approaches that empirically assess human behavior in controlled contexts. While such approaches have been widely used to study other topics, few studies have implemented them to examine use of weather forecast information. The present study complements this previous work (e.g., Roulston et al., 2006; Roulston and Kaplan, 2008; Joslyn et al., 2007, 2009ab; Nadav-Greenberg et al., 2008; Joslyn and Nichols, 2009) by asking related, but different, research questions about weather forecast interpretation and use, with the decision questions implemented in a survey rather than a laboratory setting. The survey implementation presents some potential disadvantages (see section 2.3) but also advantages (Fehr et al., 2003; Naef and Schupp, 2009). First, the survey provides a respondent population that is larger and more representative of the general public than most related experimental work. This enhances the power of the analysis and the applicability of the results to the broader public. The survey also provides a large body of attitudinal, sociodemographic, and other data that is often not available in experimental decision studies.

The findings add to existing knowledge about communication, interpretation, and use of hydrometeorological forecasts in several ways. First, despite decades of daily of forecast provision to the public, little is known empirically about how laypeople interpret different types of weather forecasts, particularly those conveying uncertainty information. Studies such as Murphy et al. (1980), Gigerenzer et al. (2005), Morss et al. (2008), and Joslyn et al. (2009a) have begun to address this by asking people their stated interpretations, but many aspects of interpretation are still poorly understood, and different aspects of interpretation may be revealed by studying information use. In addition, recent studies employing hypothetical models of how people make decisions (such as the cost-loss model) suggest that uncertainty information has potential value to hydrometeorological forecast users (e.g., Richardson, 2000; Zhu et al., 2001; Mylne, 2002; Palmer, 2002). However, for real decision makers, a variety of constraints limit effective communication and use of uncertainty information (e.g., Morss et al., 2005; NRC, 2006). Thus, empirical studies are needed to understand when communicating forecast uncertainty is desirable and how to best do so. Studies that empirically compare use of different types of weather forecasts, as we do here, are especially lacking. Further, weather forecasting is a domain in which risk information is communicated and used in decisions every day. Thus, studies such as the current one provide opportunities to apply concepts from risk communication and decision making under uncertainty to weather prediction and to inform these fields.

Section 2 reviews the methodology, including the survey implementation and the survey questions focused on in the article. Sections 3 and 4 analyze results from the threshold and binary decision questions, respectively. Section 5 summarizes and discusses the major findings.

2. Methodology

The data examined in this article was collected in 2006 as part of a nationwide U.S. survey. Section 2.1 briefly reviews the survey design and implementation. Sections 2.2 and 2.3 present the two sets of survey questions on which the analysis focuses, including descriptions of the decision scenarios and associated questions, brief discussion of related previous work and key

aspects of question design, and relevant hypotheses tested in sections 3 and 4.

The survey also included questions on respondents' sources, uses, and values of weather forecasts; perceptions and interpretations of weather forecasts, including uncertainty information; sociodemographic characteristics, and other topics (see Morss et al., 2008; Lazo et al., 2009; Demuth et al., in preparation). Several analyses were performed in conjunction with data from these other questions, in order to examine possible influences on use of forecast information. Analysis of the data was performed using Matlab and SAS.

2.1. Survey development, implementation, and respondent population

The survey instrument was developed and pretested by the authors using standard principles for developing survey questions and conducting survey research (Dillman, 2000; Schuman and Presser, 1996; Tourangeau et al., 2000). The survey was implemented in November 2006 on the Internet with a sample provided by a survey sampling company, designed to be representative of the U.S. population reachable online. The respondent population includes people from every U.S. state and the District of Columbia. Its sociodemographic characteristics are generally similar to that of the U.S. population, except that it is somewhat older and more educated and underrepresents people with very low and high incomes. While our respondent population is not a random sample of the general U.S. population, it is more diverse and representative than previous related work with convenience samples or students.

We received 1520 completed surveys. Fifty-five respondents said that they did not ever use weather forecasts and were not asked most of the remaining questions, including those examined here. Thus, the analysis presented here begins with data from 1465 respondents. For further detail on the survey development, implementation, and respondent population, see Morss et al. (2008).

2.2. Threshold decision questions

The threshold decision questions asked respondents their probabilistic forecast threshold for taking protective action in two scenarios, referred to as the picnic and garden scenarios. The picnic scenario involves protection against precipitation, and the garden scenario involves

protection against low temperature. These questions were motivated by a related question asked in Gigerenzer et al. (2005). Each respondent was randomly assigned to receive one of the two scenarios and the associated threshold question.

In the picnic scenario, respondents were told to suppose they have an outdoor picnic planned for tomorrow. They were then asked: “At what forecast chance of rain for tomorrow would you decide today to move your picnic indoors?” There were 11 response options: forecast chance of rain from 10% to 100%, in intervals of 10%, or not moving the picnic indoors (i.e., take no action).

In the garden scenario, respondents were told to suppose they have a garden with plants that will die if the temperature drops below freezing (32°F). They were then asked: “At what forecast chance that the temperature will be below freezing (32°F) tonight would you decide today to cover your plants?” There were 11 response options: forecast chance of temperature below freezing (32°F) from 10% to 100%, in intervals of 10%, or not covering the plants (i.e., take no action).

We hypothesized that respondents would select a range of thresholds, but otherwise we had no specific hypotheses about the responses.

2.3. Binary decision questions

The binary decision questions asked respondents whether or not they would take protective action given different forecast information in two scenarios, referred to as the reservoir and fruit scenarios. The two scenarios are summarized in Table 1 and presented in further detail below. The scenarios were patterned after the cost-loss decision situation (Thompson 1952; Thompson and Brier 1955), in the sense that they include two decision alternatives (protective action at a cost, or no protective action) and two possible outcomes involving monetary losses (\$100,000 damage, or no damage). For each scenario, two cost conditions (\$10,000 or \$20,000) for protective action were tested. The questions were developed to compare respondents’ use of different types of forecast information and to examine their decisions using weather forecasts from a cost-loss perspective.

The reservoir scenario involves protection against precipitation and the fruit scenario involves protection against low temperature, similar to the picnic and garden scenarios, respectively, discussed for the threshold question in section 2.2. Each respondent was (randomly) assigned to receive either precipitation or temperature scenarios for both sets of questions. In other words, respondents who received the picnic (garden) scenario for the threshold decision question received the reservoir (fruit) scenario for the binary decision questions. Each respondent was also randomly assigned to one of the two protective-action cost conditions. Thus, for the binary decision questions, the respondent population was divided into four groups (Table 1).

In the reservoir scenario, respondents were told: “Suppose you are a manager of a local water reservoir. If there are 4 inches or more of rain tomorrow, your reservoir will overflow and flood the town, causing \$100,000 in damages (but no injuries or deaths) that your company must pay for. You can prevent a potential flood by releasing water from your reservoir today, but releasing water will cost your company {\$10,000 or \$20,000}.” The four possible combinations of decisions (action or no action) and outcomes (damage or no damage) were explained. Respondents were then presented nine forecast conditions, one at a time, in random order. For each forecast condition, they were asked “Would you spend the {\$10,000 or \$20,000} to release water from your reservoir?”

In the fruit scenario, respondents were told: “Suppose you are a fruit grower and your crop is nearly ripe. If the temperature drops below freezing (32°F) tonight and your crop is unprotected, it will be damaged and you will lose \$100,000. You can prevent potential freeze damage by protecting your crop today, but protecting your crop will cost you {\$10,000 or \$20,000}.” As in the reservoir scenario, the four combinations of decisions and outcomes were explained. Respondents were then asked “Would you spend the {\$10,000 or \$20,000} to protect your crop?” for each of nine forecast conditions.

The response options were “yes” or “no” for each question. Respondents were not allowed to return to previous questions, nor were they given information about which outcome occurred.

The nine forecast conditions tested in the reservoir and fruit scenarios are shown in Table 2.

The forecast conditions were designed to manipulate a few relatively simple dimensions of information presentation, using text-based formats. The “single-value” forecast conditions provide deterministic forecast information, similar to that currently available in most forecasts (NRC, 2006); the forecasted values are different distances from the damage threshold. Two forms of uncertainty communication were tested: “range” forecasts and “percentage-chance” forecasts. The range forecast conditions were designed to represent a fairly simple form of uncertainty communication, one that does not involve probabilities. The range forecasts are symmetrical about the third single-value forecast, with the first range meeting the damage threshold and the second range exceeding the threshold. The percentage-chance forecast conditions were designed to represent a somewhat more complex form of uncertainty communication; they present different probabilities of reaching or exceeding the damage threshold.

Note that each respondent received only one of the two scenarios and one of the two cost conditions. For the scenario and cost condition they were given, however, each respondent received all nine forecast conditions (in random order). Thus, the study employs a combination of between- and within-subject design.

While the two scenarios and associated forecast conditions are similar in many ways, they are not directly parallel. Differences include the scenario content, the phrasing of the protective decision, the positive vs. negative perspective of the damage thresholds and uncertainty forecast conditions, and the different numerical values in the thresholds and forecasts. A large body of psychology and related research suggests that such differences can affect people’s responses (e.g., Tversky and Kahneman, 1981; Kuhberger, 1998; Levin et al., 1998; Windschitl and Weber, 1999; Joslyn et al., 2009b). Given this, the primary goal of testing two scenarios is to examine how results are consistent or different across the two, as a first exploration of the extent to which our findings might apply across contexts.

The binary decision questions were motivated by related experimental work in behavioral economics and psychology that examines individuals’ decisions (e.g., Kagel and Roth, 1995;

Loewenstein, 1999; Hertwig and Ortmann, 2001; Croson, 2005). The questions have elements of both approaches; for example, the presentation of monetary costs and benefits in the scenarios is more similar to economics experiments, while the context provided in the scenarios is more similar to psychology experiments (Croson, 2005; Ariely and Norton, 2007). Our survey implementation also differs from the typical implementation of such experiments in a laboratory setting. From an experimental economics perspective, a major limitation of our study related to the survey implementation is that subjects did not receive real monetary payoffs related to their decisions. Such monetary incentives are often not used in psychology studies, and previous work suggests that while monetary incentives can influence some types of findings, in many situations not using them does not substantially alter subjects' average behavior. Monetary incentives do tend to reduce variability in subjects' responses (Smith and Walker, 1993; Camerer and Hogarth, 1999; Hertwig and Ortmann, 2001). On the other hand, one advantage of implementing the decision questions in a survey setting is obtaining a larger, more diverse respondent population than typical experiments, which usually involve smaller samples of students (Fehr et al., 2003; Naef and Schupp, 2009). With this in mind, our findings are informative as a first study of some of the issues examined here, and they can help design future related survey and laboratory-based work.

Regarding the single-value forecast conditions, we hypothesized that some respondents would choose to protect when the forecast had not reached the damage threshold; we also hypothesized that more respondents would protect as the forecast became closer to the threshold. Regarding the range forecast conditions, we hypothesized that as one of the extreme values in the range reached and then exceeded the damage threshold, more respondents would protect. Regarding the percentage-chance forecasts, we hypothesized that as the likelihood of exceeding the threshold increased, more respondents would protect, and that the responses would provide evidence of decision making according to the cost-loss model. We also hypothesized that respondents would be less likely to choose protective action in the higher cost condition.

3. Results: Probabilistic thresholds for protection

Figure 1 presents results from the decision threshold questions, in which respondents were asked to indicate their probabilistic threshold for protection in either the picnic or garden scenario (described in section 2.2). Comparing responses within each graph in Fig. 1 indicates that given a similar situation, different individuals can have very different probability thresholds for decision making, supporting our hypothesis. These differences among individuals likely result from a combination of people's different tolerances for weather-related risk, their different perceptions and interpretations of weather forecasts, and other factors (discussed further below).

Comparing responses between the two graphs in Fig. 1 indicates that our survey population responded differently to the two scenarios, in two ways. First, the distribution of respondents' thresholds for taking protective action is significantly¹ different in the two scenarios (Wilcoxon rank-sum test², respondents who chose not to take protective action removed; N=1389, U=26.389, p<.0001). On average, respondents selected a higher threshold in the picnic scenario (58.3%) than in the garden scenario (53.3%). Given the multiple differences between the scenarios, understanding why respondents' thresholds differed would require further research. Overall, however, this finding suggests that context can influence people's use of weather forecasts.

Second, the distribution of responses in the garden scenario exhibits secondary peaks at the lowest (10%) and highest (100%) thresholds, suggesting that people tended to respond less consistently than in the picnic scenario. The garden scenario asked about probability of temperature exceedance forecasts, which are generally not currently publicly available, while the picnic scenario asked about probability of precipitation (PoP) forecasts, which have been

¹ Throughout the manuscript, the term "significant" is used to denote statistical significance.

² The Wilcoxon rank-sum test (also called the Mann–Whitney–Wilcoxon test) is a non-parametric test of whether two independent samples of observations come from identical populations (have the same underlying distribution).

regularly provided to the U.S. public for decades. One possibility is that people responded less consistently in the garden scenario because of their relative lack of familiarity with the forecast format. Further research would be required to test this explanation and, more generally, to examine how experience with a type of forecast information affects its use.

As noted above, we expect that people's use of forecast information is related to their perceptions and interpretations of forecasts, including forecast uncertainty. To begin exploring these relationships, we performed two analyses examining people's responses to the threshold decision questions in conjunction with their responses to other survey questions. The first analysis examined whether respondents who reported different levels of confidence in 1-day PoP or temperature forecasts (see Morss et al., 2008) had different distributions of protective-action thresholds in the picnic or garden scenarios, respectively. The analysis was performed separately for the two scenarios, and no significant differences were found (Kruskal-Wallis test³, respondents who chose not to take protective action removed; picnic: $N=698$, $\chi^2=6.370$, $df=4$, $p=0.1732$; garden: $N=691$, $\chi^2= 2.314$, $df= 4$, $p=0.6783$). This could be because people's confidence has a complex relationship with their thresholds for protective action. For example, as people's confidence decreases, they may have lower confidence that a damaging event will be detected, leading to a lower threshold for protection; alternatively, their lower confidence may cause them to be less likely to take action based on a forecast, leading to a higher threshold or a greater likelihood to not take protective action.

The second analysis examines relationships between people's interpretations of PoP and their use of PoP forecasts in decision making, following Gigerenzer et al. (2005) and Joslyn et al. (2009a). These two studies as well as Morss et al. (2008) found that many laypeople still do not know the technically correct definition of PoP forecasts, corroborating earlier work by Murphy et al. (1980). Gigerenzer et al. found that people who interpreted PoP as meaning percent region or

³ The Kruskal-Wallis test is a non-parametric test of whether three or more independent samples come from identical populations.

time of rain had a slightly lower PoP threshold for taking an umbrella than those with a “correct” interpretation. Joslyn et al. found that “misinterpreting the forecast as more than half the area or time affects decision making by increasing the tendency to take precautionary action” (p. 188). Both discuss this result in terms of people’s tendency to misinterpret PoP as a deterministic forecast of rain, with the only uncertainty being about where or when; this leads them to take protective action at lower PoPs, or, equivalently, to be more likely to take protective action for a given PoP.

Morss et al. (2008) examined people’s PoP interpretations using data from the survey discussed here. While they found that many respondents selected a “region” or “time” interpretation of PoP when asked in a closed-ended format, other aspects of their results suggest that many people do not have preconceived interpretations of PoP of the types emphasized in previous studies. Further, despite difficulty clearly defining PoP forecasts, most people find PoP forecasts important (Morss et al., 2008; Lazo et al., 2009). Based on their results, Morss et al. suggest that a correct technical understanding of PoP may not be important to many people in using PoP. Given this different view of people’s interpretations and use of PoP and our larger, more representative respondent population, we decided to examine the relationship between PoP interpretation and use in our data. Our decision question also differs from that asked by Joslyn et al., so that we are examining this relationship from a complementary decision-making angle.

To analyze this relationship, we used data from the 629 respondents who received both the picnic scenario and a closed-ended question asking their interpretation of the PoP forecast “60% chance of rain tomorrow” (see Morss et al., 2008). We examined respondents’ PoP thresholds for taking protective action, grouped according to their selected interpretation of PoP. For comparison with Joslyn et al., we also examine results with the “region” and “time” interpretation groups combined. Results are shown in Table 3, including the mean PoP threshold for each group and a statistical test of whether each group’s distribution of PoP thresholds is significantly different from that of the group who selected the “correct” (“days”) interpretation. The mean thresholds do differ across the groups with different interpretations. However, only

one group — the “time” interpretation group — has a threshold distribution that is significantly different (at the 5% level) than the “correct” group. The “time” group does have a lower mean PoP threshold, which would support Joslyn et al.’s conclusion about “deterministic” misinterpretations of PoP influencing decision making. However, the “region” group does not have significantly different thresholds from the “correct” group, nor do the “region” and “time” groups combined. This, along with other results in Table 3, suggests that the relationship between PoP interpretation and decision making is more complex.

Overall, our result that people’s decision thresholds depend on the individual and the context suggest the importance of giving people forecast information that they can use to make decisions based on their own criteria given the situation, rather than recommending decisions based on assumed thresholds for protection. Our findings also indicate the importance of understanding and considering contextual and individual influences on forecast interpretation and use for developing effective forecast communication formats. Further work is needed to understand the potentially complex relationships between people’s perceptions and interpretations of forecasts (including PoP) and their use.

4. Results: Binary protective decisions with different forecast information

This section examines results from the binary decision questions, in which respondents were asked to make yes-no protective decisions in either the reservoir or fruit scenario (described in section 2.3 and Table 1). Respondents were given one of the two scenarios and one of the two cost conditions and were asked whether they would take protective action under each of the nine forecast conditions in Table 2.

Of the 1465 respondents, 183 answered “no” (take no protective action) to all nine forecast conditions, and 49 answered “yes” (take protective action) to all nine. These respondents’ repeated yes or no decisions suggest that they may have had difficulty understanding or answering the questions or lacked motivation to expend the cognitive energy required to do so (e.g., Jackson, 1967; Krosnick, 1991). These responses also provide little information for the

analyses presented here. Consequently, we removed these 232 respondents (15.8%) from the analyses in this section⁴. The number of respondents for each of the scenarios and cost conditions, before and after removal of these respondents, is provided in Table 1.

We start by comparing results across the scenarios and cost conditions, in section 4.1. Sections 4.2, 4.3, and 4.4 then compare respondents' decisions among the nine forecast conditions, categorized as shown in Table 2. Recall that respondents received the forecast conditions in random order and were not permitted to return to previous questions, which reduces the potential for order effects on the major findings. Most of the results presented are for comparisons across the subject panel; a few within-subject comparisons are also discussed.

4.1. Response to decision context and cost of protection

Two general patterns appear across the binary decision question results. We analyze these patterns here by pooling results across all nine forecast conditions, but they are also evident in responses to individual forecast conditions (see Fig. 2 and subsequent figures).

First, protective action was chosen significantly more often in the fruit scenario than in the reservoir scenario (Pearson chi-squared test⁵; $N=11,097$, $\chi^2=137.315$, $df=1$, $p<.0001$). As discussed in section 2.3, the scenarios and associated decisions and forecasts differ in multiple aspects of content and presentation that could have contributed to this difference in tendency to protect. Consequently, further research would be needed to understand this result. More generally, this result suggests that, as discussed in section 3, context can affect the interpretation and use of weather forecasts.

Second, there were no significant differences in how often protective action was chosen between the two cost conditions ($N=11,097$, $\chi^2=0.073$, $df=1$, $p=.7865$; see also the logistic regression analysis in section 4.4). This result rejects our hypothesis that respondents would be

⁴ Removing these respondents does not change the overall findings.

⁵ The Pearson chi-squared test tests whether the frequency distributions of two or more samples come from identical populations.

more likely to protect when the cost of protection was lower. We suspect that we did not detect a response to cost because, given the hypothetical scenarios and the lack of real monetary payoffs in the study design, the monetary values in the questions were not substantially meaningful to many respondents. This finding may also have been influenced by our between-subject design for the cost condition. Given this result, we combine the \$10,000 and \$20,000 cost conditions in most of the remaining analysis and discussion (although the cost conditions are still depicted separately in the figures).

4.2. Interpretation and use of single-value forecasts

This section compares respondents' decisions in the single-value (deterministic) forecast conditions tested (Fig. 2). The results indicate how people responded when the deterministic forecast was different distances from the damage threshold, providing information about people's interpretations of deterministic forecasts as revealed by their stated decisions.

Figure 2 shows that, as hypothesized, some respondents chose to protect when the forecasted value had not reached or exceeded the damage threshold. This indicates that many people infer uncertainty into deterministic low temperature and quantitative precipitation forecasts, corroborating and extending Morss et al.'s (2008) similar finding for high temperature forecasts based on people's stated interpretations.

In addition, as hypothesized, significantly more respondents chose to protect as the single-value forecast became closer to the damage threshold, in both scenarios (Pearson chi-squared test, cost conditions combined; reservoir: $N=1821$, $\chi^2=427.042$, $df=2$, $p<.0001$; fruit: $N=1878$, $\chi^2=335.443$, $df=2$, $p<.0001$; see also the logistic regression model results discussed later in this section). Within-subject comparisons indicate that only 8% of individuals did not follow this pattern, in other words, changed their response from "yes" to "no" as the forecast grew closer to the threshold. Together, these results suggest that many if not most of our respondents were able to interpret and use the deterministic forecasts in the general manner we expected when posing the questions.

This data also provides some information about respondents' interpretations of uncertainty in

single-value low-temperature and quantitative precipitation forecasts. For example, since the damage threshold in the fruit scenario is 32°F, the 72% of respondents who chose to take protective action given a forecast of 33°F indicated a belief that there was at least 1°F of uncertainty in the deterministic forecast. More specifically, if one imagines that respondents infer some uncertainty distribution about a deterministic forecast, this 72% believed there was sufficient probability the temperature would drop below 32°F to merit protective action. Some of the remaining 28% may also have believed temperatures might fall below 32°F, but the probability was not sufficiently large to merit protection. The other data in Fig. 2 can be used to make similar inferences about respondents' interpretations of uncertainty in the other deterministic forecasts tested; for example, 59% of respondents indicated an interpretation of at least 1 in. of uncertainty in a deterministic forecast of rain amount. Because the probability of damaging weather required for a person to take protective action depends on several factors about which we do not have complete information, further research would be needed to use data on people's decisions to analyze their interpretations of uncertainty in greater detail.

Following on the discussion in section 3, we began exploring relationships between people's decisions using deterministic forecasts and their perceptions and interpretations of forecast uncertainty. To do so, we analyzed the data from Fig. 2 in conjunction with data from the survey on respondents' stated confidence in 1-day forecasts of different parameters and their stated expectations of high temperature given a 1-day forecast (Morss et al. 2008). Relationships were examined in several different ways, and no significant relationships were found. The most comprehensive analysis was construction of logistic regression models (see section 4.4). Data for the reservoir and fruit scenarios were modeled separately, with cost conditions combined. The dependent variable was yes/no response; the independent variables were: 1) distance of the deterministic forecast from the damage threshold; 2) confidence in 1-day forecasts of precipitation amount (reservoir scenario) or temperature (fruit scenario); and 3) for the fruit scenario, interpretations of 1-day high temperature forecasts. Interactions between variables were included in the models, to test, e.g., whether people with lower confidence were more likely than

others to protect when the deterministic forecast was closer to the threshold. As expected given Fig. 2, the distance of the forecast from the threshold was a highly significant predictor in both scenarios. The other variables tested were not significant predictors. Thus, as discussed in section 3, further work is needed to explore relationships between people's perceptions and interpretations of weather forecasts and their forecast use.

4.3. Interpretation and use of range forecasts

Next we examine respondents' decisions in the range forecast conditions tested, compared with the single-value forecast conditions with equivalent midpoints (Fig. 3). The comparison provides information about how people interpreted forecasts that conveyed uncertainty using ranges, as revealed by their stated decisions.

A significantly different number of respondents chose to protect in the three forecast conditions shown in Fig. 3, in both scenarios (Pearson chi-squared test, cost conditions combined; reservoir: $N=1821$, $\chi^2=19.022$, $df=2$, $p<.0001$; fruit: $N=1878$, $\chi^2=11.318$, $df=2$, $p=.0035$). However, the differences in response to the three forecast conditions in Fig. 3 are small, compared, e.g., to the differences in Fig. 2. We suspect that this is because the ranges tested are fairly narrow (and have the same midpoint). Recall from section 4.2 that many respondents appear to infer at least 1°F of uncertainty into temperature forecasts and at least 1 in. of uncertainty into quantitative precipitation forecasts. Given these interpretations, adding a range of +/- 1-2°F or 1-2 in. of rain to a forecast appears not to have altered many respondents' decisions. This is corroborated by a within-subject comparison, which indicates that 59% of respondents (70% in the fruit scenario and 47% in the reservoir scenario) selected the same decision in all three conditions shown in Fig. 3.

Despite the small differences in Fig. 3, the fruit scenario results support our hypothesis that as one of the extreme values in the range reached and then exceeded the damage threshold of 32°F, more respondents would choose to protect. However, in the reservoir scenario, fewer respondents chose to protect given a forecast of 1-5 in. than a forecast of 2-4 in. This may be because the placement of "4 in." at the end of the 2-4 in. forecast emphasizes the 4 in. damage

threshold, whereas in the 1-5 in. forecast condition, respondents have to conduct a simple subtraction to compare the forecast with the threshold. As discussed earlier, previous psychology research has shown that such details in presentation can be important. Overall, therefore, these results suggest that further research is needed to understand how people interpret and use forecasts expressing uncertainty as ranges similar to those tested. This is particularly important given that, besides PoP, such ranges (or implied ranges, such as “low temperature in the low 30s”) are one of the more common forms used to communicate weather forecast uncertainty to the public today.

4.4. Interpretation and use of percentage-chance forecasts

Finally, we examine respondents’ decisions in the percentage-chance forecast conditions tested: a 5%, 10%, 20%, or 40% chance that the damage threshold will be reached or exceeded (Fig. 4). Studies indicating laypeople’s lack of numeracy in general, along with anecdotal experiences from weather forecast providers, has led to discussion in the weather prediction community about whether probabilities are a suitable form for presenting uncertainty to the general public (e.g., Murphy et al., 1980; Sink, 1995; NRC, 2006; Roulston and Kaplan, 2008). Here we contribute to this discussion by empirically examining respondents’ ability to understand and use one form of probabilistic weather forecasts. The data also allows comparison of people’s decisions in the scenarios with decisions hypothesized in the cost-loss model, which has been used as a prototype decision model for studying use and value of probabilistic hydrometeorological forecasts (Katz and Murphy 1997).

Figure 4 indicates that in both scenarios, more respondents chose to take protective action as the forecasted chance of exceeding the damage threshold increased. To test this relationship, we combined data from the two scenarios and two cost conditions and constructed a logistic regression⁶ model (O’Connell, 2005), with yes/no response as the dependent variable (Table 4). Because this analysis combines individuals’ responses to four questions (one for each

⁶ Logistic regression is used because the response variable is a discrete qualitative variable.

percentage-chance forecast), we accounted for potential intra-subject correlation using the method of generalized estimating equations (Allison, 1999, Ballinger, 2004; Lazo et al., in press). Percentage-chance forecast was a highly significant predictor of respondents' decisions: the higher the percentage chance of damage, the more likely they were to take protective action. Respondents were also significantly more likely to protect in the fruit scenario than the reservoir scenario, and cost of protection was not a significant predictor, corroborating results in section 4.1. We tested incorporating various sociodemographic characteristics and measures of forecast perception and interpretation as additional model predictors, using data from other questions on the survey, but found no stable, significant relationships.

The study design used here does not allow a strong test of whether people can use uncertainty-explicit forecasts to make better decisions than they would using deterministic forecasts. However, Fig. 4 and Table 4 do provide evidence that much of our respondent population was able to interpret probabilistic forecast information of the type provided well enough to use it in the binary decision questions. This is further supported by a within-subject analysis, summarized in Table 5. Looking at the right-most column (all respondents combined), 55% of individuals ordered their responses to the four percentage-chance conditions in a manner that suggests understanding of the forecasts; in other words, their decision changed from no (don't protect) to yes (protect) as the chance of exceeding the damage threshold increased (rows b-d). An additional 30% made the same decision under all four percentage-chance conditions (rows a and e). Only 15% chose to protect at a lower chance and not to protect at a higher chance (row f); we refer to these as "unordered" responses⁷.

The 15% of respondents who provided unordered responses may not have understood the

⁷ Approximately one-quarter of these respondents (3.5% of all respondents) gave a pattern of responses that changed from take protective action to do not take protective action as the percentage chance increased (YNNN, YYNN, or YYYN using the format in Table 5) and thus are not strictly "unordered." See also footnote 10.

scenario or the percentage-chance forecasts, or their decisions may have been influenced by other factors. Such variability or inconsistency in responses is typical in laboratory experiments of this type (e.g., Camerer, 1989; Starmer and Sugden, 1989; Hey, 2001; Blavatsky, 2007). Recall that these four forecast conditions were randomly ordered among the nine presented, and that the survey implementation provided no monetary incentives for or feedback on decisions. Further, previous research suggests limitations in numeracy among much of the public (e.g., Lipkus et al., 2001; Peters et al., 2008), and the types of percentage-chance weather forecasts tested here are unfamiliar to most people. Given these considerations, a notable percentage of individuals provided responses suggesting that they understood probabilistic weather forecasts of the type tested well enough to use them to make simple protective decisions. One possible implication is that it may not be entirely accurate to characterize members of the public as “less sophisticated” users of uncertainty-explicit weather forecasts, compared to “more sophisticated” users such as emergency managers or members of the private sector (as discussed in, e.g., NRC, 2003; Novak et al., 2008; Santos et al., 2008).

The within-subject analysis in Table 5 also provides information about respondents’ thresholds for taking protective action. Looking again at the rightmost column with all respondents combined, 6% of respondents changed their response from no at 5% chance of damaging weather to yes at 10% chance, indicating a decision threshold between 5% and 10% (row b). An additional 13% and 36% indicate thresholds between 10-20% and 20-40%, respectively (rows c and d). Assuming they understood the forecasts and questions, respondents who said yes or no to all four forecast conditions indicate thresholds below 5% or above 40%, respectively (rows a and e). As expected from Fig. 4 and respondents’ tendency to protect more often in the fruit scenario (section 4.1), a within-subject analysis of the two scenarios separately (not shown) suggests that respondents had higher thresholds for protection in the reservoir scenario than the fruit scenario. These results further support those discussed in section 3, that different individuals have different percentage-chance thresholds for decision making based on their risk tolerance, the context, and other factors.

The expected value of taking protective action and not taking action are equal when there is a 10% (20%) chance of exceeding the damage threshold in the \$10,000 (\$20,000) cost condition. Thus, according to the simplest formulation of the cost-loss decision model, respondents are expected to protect (not protect) when the forecasted chance of damage is above (below) this level.⁸ One piece of evidence that many respondents were not making cost-loss decisions is our finding that cost of protection did not significantly affect decisions. Further, as Fig. 4 and the middle two columns of Table 5 show, many respondents chose to protect (not protect) when the expected cost was greater (less) than the expected loss. A portion of respondents indicated risk aversion, but many indicated risk-seeking behavior. Some of this behavior could be due to the variability in subject responses typical in experimental studies, exacerbated by the lack of real monetary incentives (Smith and Walker, 1993; Camerer and Hogarth, 1999). However, core theories in economics and psychology (such as expected utility and prospect theory) suggest that people in general do not make strict expected-value decisions, and that people tend to be risk-seeking in the domain of moderate-probability losses (e.g., von Neumann and Morgenstern, 1944; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992).⁹ Thus, our data provides evidence, supported by numerous studies in other settings, that many people do not make decisions as hypothesized in the simplest form of the cost-loss decision making model, even in idealized scenarios modeled after static cost-loss decisions.

⁸ The decision criterion in the cost-loss model is: protect when $p > C/L$ and do not protect when $p < C/L$, where p is the probability of damaging weather, C is the cost of protection, and L is the loss suffered when damaging weather occurs. When $p = C/L$, decision makers can choose to either protect or not protect.

⁹ Prospect theory also suggests that people tend to shift from risk-averse to risk-seeking as the probability of loss increases from low to moderate. Thus, our characterization of “unordered” respondents and of risk-averse/risk-seeking respondents in Table 5 may be oversimplified.

5. Summary and discussion

This article builds knowledge about communication, interpretation, and use of hydrometeorological forecasts — including forecast uncertainty information — by analyzing data from decision scenario questions incorporated into a nationwide survey of the U.S. public. The decision questions are related to experimental economics and psychology approaches to empirically studying people’s use of information in decisions. Testing forecast use in a controlled context permits simplification of the decision setting and the information available, facilitating focused study. Use of forecasts of low temperature, probability of precipitation, and precipitation amount is explored.

The results indicate that, given the same decision scenario, different members of the public have a variety of probabilistic thresholds for taking protective action. Given the same scenario and the same forecast information, different people also make different protective decisions. Both findings are consistent with individuals’ different orientations towards weather-related risk, varying interpretations of forecasts and situations, and other differences. A potential implication is that, at least for forecasts with indefinite or inhomogeneous audiences, it may be important to provide forecast information that people can use to make decisions based on their own criteria, rather than recommending decisions based on pre-defined thresholds.

Respondents’ thresholds for and likelihood of taking protective action also varied with the decision scenario. This is consistent with a large body of research in non-weather settings indicating that context, presentation, and framing can affect information interpretation and use. This suggests that designing effective communication formats will require testing forecast interpretation and use in different contexts and considering presentation and framing effects.

When given single-value (deterministic) forecasts, the majority of respondents chose to take protective action even when the forecasted value did not reach the damage threshold. This suggests that most respondents interpreted the deterministic forecasts as uncertain, corroborating and extending the results on stated interpretations reported in Morss et al. (2008). As in Morss et al. (2008), different people interpreted uncertainty in deterministic forecasts differently. The

results presented here provide some information about people's interpretations of uncertainty in deterministic forecasts of rain amount and low temperature, but more in-depth study is needed to understand these interpretations in greater detail.

When given forecasts that conveyed uncertainty as ranges, more respondents chose to take protective action when the forecasted range reached or exceeded the damage threshold than when given a single-value forecast with the midpoint of the range. However, considering the uncertainty that most respondents are already interpreting in single-value forecasts, ranges with the spreads we tested did not provide many respondents with information that altered their decision. Moreover, our findings suggest possible presentation issues in interpretation of range forecasts. Consequently, further investigation of how people interpret and use range forecasts is needed, especially since these are one of the more common forms currently used to communicate weather forecast uncertainty.

When given forecasts of the percentage chance of exceeding the damage threshold, more respondents took protective action as the percentage chance increased. This result and a within-subject analysis suggest that many respondents understood probabilistic forecasts of the type tested well enough to use them in the decisions presented. This was true for probabilistic forecasts of precipitation amount and low temperature, neither of which is commonly available to members of the U.S. public (the population sampled here). This provides some evidence that, unlike in some other areas of risk communication, weather and weather forecasts may be sufficiently common experiences that probabilities may be suitable for conveying uncertainty in weather forecasts to many members of the public, at least for everyday weather (e.g., NRC, 2006; Handmer and Proudly, 2007). This is supported by previous findings that the majority of laypeople understand the probabilistic component of PoP and like PoP forecasts conveyed using probabilities (Murphy et al., 1980, Morss et al., 2008). Yet related research suggests potential issues with laypeople's specific interpretations of probabilistic weather forecasts, as well as the importance of information presentation (e.g., Gigerenzer et al. 2005; NRC 2006; Morss et al. 2008; Joslyn et al. 2009a). Thus, further work is needed to understand when probabilistic and

other forms of uncertainty communication are valuable across weather prediction contexts, as well as how to communicate weather forecast uncertainty in ways that promote effective interpretation and use.

Our analysis also indicates that many respondents did not make decisions according to the simplest form of the cost-loss decision model used in some meteorological studies of forecast use and value. This occurred even though we presented decision makers with a one-time, binary decision given only one piece of forecast information and perfect knowledge of costs and losses; most real decisions are more complex. While the lack of real monetary payoffs in our study design may have increased variability in respondents' decisions, this finding is consistent with a large body of work from economics, psychology, and other fields on the importance of risk perception and other factors in people's decision making. Thus, while the cost-loss model can be useful for hypothetical study, understanding real-world forecast interpretation, use, and value requires empirical studies of people's decisions using weather forecasts. Such studies should build on knowledge of how people make decisions under risk and uncertainty developed in other disciplines and other settings.

Finally, we explored relationships between subjects' responses to the decision questions and their perceptions and interpretations of forecasts, using data from other questions posed in the survey (Morss et al., 2008; Lazo et al., 2009). One analysis examined whether respondents who selected different interpretations of PoP had different PoP thresholds for taking protective action. This analysis uses data from a larger, more representative sample than previous related work by Gigerenzer et al. (2005) and Joslyn et al. (2009). The results, along with those in Morss et al. (2008), suggest that the relationship between PoP interpretation and use is more complex than discussed in these two previous studies. Our other cross-analyses found no notable relationships between individuals' forecast perceptions and interpretations (as measured in the survey) and their forecast use. We suspect that such relationships exist but are more complex than we can examine using the current data set. Thus, further study of such relationships is needed.

Future work, implemented in surveys or laboratory settings, can build on the questions asked

here and extend empirical comparison of forecast interpretation and use to more complex presentation formats. By design, work in controlled settings simplifies the decision context and information available. Thus, such work is complementary to “real world” empirical studies employing case analyses, interviews, other types of survey questions, naturalistic decision making, and other methodologies to examine forecast communication, interpretation, and use. Together, the knowledge gained can help the meteorology community learn whether, when, and how to provide different types of forecast information — particularly uncertainty information — to various audiences, in order to enhance forecast interpretation and beneficial use.

Acknowledgments The authors thank Jamie Kruse for helpful discussion on aspects of the analysis, Susan Joslyn for helpful comments on an earlier version of the manuscript, and Kelsey Mulder for assistance with literature review. This research was partially sponsored by NCAR’s Collaborative Program on the Societal Impacts and Economic Benefits of Weather Information, which is funded by the National Science Foundation and the National Oceanic and Atmospheric Administration through the U.S. Weather Research Program. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

6. References

- Allison PD. 1999. *Logistic Regression Using the SAS System: Theory and Application*. SAS Institute, Inc.: Cary, NC; 308pp.
- Ariely D, Norton MI. 2007. Psychology and experimental economics: A gap in abstraction. *Current Directions in Psychological Science* **16**: 336–339.
- AMS (American Meteorological Society). 2008. Enhancing weather information with probability forecasts. *Bulletin of the American Meteorological Society* **89**: 1049–1053.
- Ballinger GA. 2004. Using Generalized Estimating Equations for longitudinal data analysis. *Organizational Research Methods* **7**: 127-150.
- Blavatskyy P. 2007. Stochastic expected utility theory. *Journal of Risk and Uncertainty* **34**: 259-286.
- Budescu DV, Broomell S, Por HH. 2009. Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. *Psychological Science* **20**: 299-308.
- Camerer CF. 1989. An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty* **2**: 61-104.
- Camerer CF, Hogarth R. 1999. The effects of financial incentives in economics experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* **19**: 7-42.
- Croson R. 2005. The method of experimental economics. *International Negotiation* **10**: 131–148.
- Demeritt D, Cloke H, Pappenberger F, Thielen J, Bartholmes J, Ramos M-H. 2007. Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards* **7**: 115–127.
- Dillman, DA. 2000. *Mail and Internet Surveys: The Tailored Design Method*. 2d ed. John Wiley & Sons, Inc.: New York, NY; 464 pp.
- Fehr E, Fischbacher U, von Rosenbladt B, Schupp J, Wagner GG. A nation-wide laboratory: examining trust and trustworthiness by integrating behavioral experiments into representative survey. CESifo Working Paper Series No. 866.

- Gigerenzer G, Hertwig R, van den Broek E, Fasolo B, Katsikopoulos KV. 2005. A 30% chance of rain tomorrow: How does the public understand probabilistic weather forecasts? *Risk Analysis* **25**: 623–629.
- Handmer J, Proudly B. 2007. Communicating uncertainty via probabilities: The case of weather forecasts. *Environmental Hazards* **7**: 79-87.
- Hertwig R, Ortmann A. 2001. Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences* **24**: 383-403.
- Hey J. 2001. Does repetition improve consistency? *Experimental Economics* **4**: 5-54.
- Jackson DA. 1967. Acquiescence response styles: Problems of identification and control. In *Response Set in Personality Assessment*, Berg IA (ed.). Aldine Publishing Company: Chicago, IL; 71-114.
- Joslyn SL, Nichols RM. 2009. Probability or frequency? Expressing forecast uncertainty in public weather forecasts. *Meteorological Applications* **16**: 309-314.
- Joslyn S, Nadav-Greenberg L, Nichols RM. 2009a. Probability of precipitation: Assessment and enhancement of end-user understanding. *Bulletin of the American Meteorological Society* **90**: 185-193.
- Joslyn S, Nadav-Greenberg L, Taing MU, and Nichols RM. 2009b. The effects of wording on the understanding and use of uncertainty information in a threshold forecasting decision. *Applied Cognitive Psychology* **23**: 55-72.
- Joslyn S, Pak K, Jones D, Pyles J, Hunt E. 2007. The effect of probabilistic information on threshold forecasts. *Weather and Forecasting* **22**: 804–812.
- Kagel JH, Roth AE (eds.). 1995. *Handbook of Experimental Economics*. Princeton University Press: Princeton, NJ; 721 pp.
- Kahneman D, Tversky A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**: 263-291.
- Katz RW, Murphy AH. Forecast value: Prototype decision-making models. In *Economic Value of Weather and Climate Forecasts*, Katz RW, Murphy AH (ed.). Cambridge University Press:

Cambridge, UK, 183-217.

- Krosnick JA. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* **5**: 213-236.
- Kühberger A. 1998. The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes* **75**: 23–55.
- Lazo JK, Morss RE, Demuth JL. 2009. 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bulletin of the American Meteorological Society* **90**: 785-798.
- Lazo, JK, Waldman DM, Morrow BH, Thacher JA. In press. Household evacuation decision making and the benefits of improved hurricane forecasting: Developing a framework for assessment. *Weather and Forecasting*.
- Levin IP, Schneider SL, Gaeth GJ. 1998. All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes* **76**: 149-188.
- Lipkus IM, Samsa G, Rimer BK. 2001. General performance on a numeracy scale among highly educated samples. *Medical Decision Making* **21**:37-44.
- Loewenstein G. 1999. Experimental economics from the vantage-point of behavioural economics. *Economic Journal* **109**: F23-34,
- Manning MR. 2003. The difficulty of communicating uncertainty. *Climatic Change* **61**: 9-16.
- Morss RE, Demuth J, Lazo JK. 2008. Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Weather and Forecasting* **23**: 974–991.
- Morss RE., Wilhelmi OV, Downton MW, Grunfest E. 2005. Flood risk, uncertainty, and scientific information for decision-making: Lessons from an interdisciplinary project. *Bulletin of the American Meteorological Society* **86**: 1593–1601.
- Murphy AH, Lichtenstein S, Fischhoff B, Winkler RL. 1980. Misinterpretations of precipitation probability forecasts. *Bulletin of the American Meteorological Society* **61**: 695-701.
- Mylne K. 2002. Decision-making from probability forecasts based on forecast value. *Meteorological Applications* **9**: 307–315.

- Nadav-Greenberg L, Joslyn S, Taing MU. 2008. The effect of weather forecast uncertainty visualization on decision-making. *Journal of Cognitive Engineering and Decision Making* **2**: 24-47.
- Naef M, Schupp J. Measuring trust: Experiments and surveys in contrast and combination. SOEP paper No. 167.
- Novak DR, Bright DR, Brennan MJ. 2008. Operational forecaster uncertainty needs and future roles. *Weather and Forecasting* **23**: 1069-1084.
- NRC (National Research Council). 2003. *Communicating Uncertainties in Weather and Climate Information: A Workshop Summary*. National Academies Press: Washington, DC; 68 pp.
- NRC. 2006. *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts*. National Academies Press: Washington, DC; 124 pp.
- O'Connell, AA. 2005. *Logistic Regression Models for Ordinal Response Variables*. Sage Publications: Thousand Oaks, CA; 120 pp
- Palmer TN. 2002. The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quarterly Journal of the Royal Meteorological Society* **128**: 747-774.
- Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S. 2006. Numeracy and decision making. *Psychological Science* **17**: 407-413.
- Richardson DS. 2000. Skill and economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society* **126**: 649-668.
- Roulston MS, Bolton GE, Kleit AN, Sears-Collins AL. 2006. A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting* **21**: 116-122.
- Roulston MS, Kaplan TR. 2008. A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. *Meteorological Applications* **16**: 237-244.
- Ryan RT. 2003. Digital forecasts: Communication, public understanding, and decision making. *Bulletin of the American Meteorological Society* **84**: 1001-1003.

- Santos P, Sharp DW, Rader G, Volkmer M. 2008. Employing hurricane wind probabilities to convey forecast uncertainty and potential impact through NWS field office forecast products. In *Proceedings of 19th Conference on Probability and Statistics, New Orleans, LA, 21-24 Jan 2008*. American Meteorological Society: Boston.
- Schuman H, Presser S. 1996. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. Sage Publications: Thousand Oaks, CA; 372 pp.
- Sink SA. 1995: Determining the public's understanding of precipitation forecasts: Results of a survey. *National Weather Digest* **9**: 9-15.
- Smith VL, Walker J. 1993. Monetary rewards and decision cost in experimental economics. *Economic Inquiry* **31**: 245–261.
- Starmer C, Sugden R. 1989. Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty* **2**: 159–178.
- Thompson JC. 1952. On the operational deficiencies in categorical weather forecasts. *Bulletin of the American Meteorological Society* **33**: 223-226.
- Thompson JC, Brier GW. 1955. The economic utility of weather forecasts. *Monthly Weather Review* **83**: 249-254.
- Tourangeau R, Rips LJ, Rasinski K. 2000. *The Psychology of Survey Response*. Cambridge University Press: Cambridge, UK; 416 pp.
- Tversky A, Kahneman D. 1981. The framing of decisions and the psychology of choice. *Science* **211**: 453-458.
- Tversky A, Kahneman D. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5**: 297-323.
- von Neumann J, Morgenstern O. 1944. *Theory of Games and Economic Behavior*. Princeton University Press: Princeton, NJ.
- Windschitl PD, Weber EU. The interpretation of “likely” depends on the context, but “70%” is 70% — right?: The influence of associative processes on perceived certainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **25**: 1514-1533.

Zhu Y, Toth Z, Wobus R, Richardson D, Mylne K. 2001. On the economic value of ensemble-based weather forecasts. *Bulletin of the American Meteorological Society* **83**: 73–83.

Table 1. Overview of the reservoir and fruit scenarios used in the binary decision questions described in section 2.3 (each with two cost conditions).

	Reservoir scenario		Fruit scenario	
<i>Potential damage (monetary loss)</i>	\$100,000		\$100,000	
<i>Damage threshold</i>	4 in. or more of rain		low temperature below 32°F	
<i>Cost of taking protective action</i>	\$10,000	\$20,000	\$10,000	\$20,000
<i>Number of respondents after / before removing all-yes and all-no responses (total 1233 / 1465)</i>	307 / 362	300 / 367	315 / 363	311 / 373

Table 2. Forecast conditions presented in the binary decision questions.

	Reservoir scenario	Fruit scenario
<i>Single-value forecast conditions</i>	1 in. of rain	low temperature of 37°F
	2 in. of rain	low temperature of 35°F
	3 in. of rain	low temperature of 33°F
<i>Range forecast conditions</i>	2 to 4 in. of rain	low temperature of 32°F to 34°F
	1 to 5 in. of rain	low temperature of 31°F to 35°F
<i>Percentage- chance forecast conditions</i>	5% chance of 4 in. or more of rain	5% chance of 32°F or lower
	10% chance of 4 in. or more of rain	10% chance of 32°F or lower
	20% chance of 4 in. or more of rain	20% chance of 32°F or lower
	40% chance of 4 in. or more of rain	40% chance of 32°F or lower

Table 3. Results for the probabilistic threshold question in the picnic scenario, with subjects grouped by their responses to the survey’s closed-ended question on interpretation of 60% PoP (discussed in Morss et al., 2008, their section 3c and Table 2).

PoP interpretation	N	Mean PoP threshold for protection	Standard deviation of PoP threshold for protection	Probability that the group’s PoP-threshold distribution differs from that of “days” interpretation group (Wilcoxon two-sided rank-sum test, df=1)
<i>Percent of region</i>	116	57.8	19.5	0.44
<i>Percent of time</i>	70	53.4	19.7	0.04
<i>Percent of region and percent of time combined</i>	186	56.1	19.6	0.12
<i>Percent of days like tomorrow*</i>	124	59.0	17.8	
<i>Percent of forecasters</i>	127	57.3	19.8	0.43
<i>Don’t know</i>	54	56.1	18.4	0.30
<i>Other</i>	138	61.9	15.5	0.25

* Technically correct interpretation, according to how PoP forecasts are verified, as interpreted by Gigerenzer et al. (2005).

Table 4. Results from logistic regression model examining respondents' decision whether to take protective action (yes=1, no=2) as a function of percentage-chance forecast (5%, 10%, 20%, or 40%), decision scenario (reservoir=1 or fruit=2), and cost of protection (\$10,000 or \$20,000). The regression models the probability of the lower value response ("yes"); see section 4.4 for description of model (N=4932).

Parameter	Estimate	Standard Error	P
<i>Intercept</i>	-3.0038	0.2195	<.0001
<i>Percentage-chance forecast</i>	0.0756	0.0027	<.0001
<i>Scenario</i>	0.5274	0.0911	<.0001
<i>Cost of protection</i>	-0.0021	0.0090	.8196

Table 5. Within-subject analysis of responses to the four percentage-chance conditions, with data from the reservoir and fruit scenarios combined. Row a (e) represents subjects who said yes (no) to the four percentage-change forecast conditions. Rows b-d represent subjects who provided “ordered” responses, and row f represents subjects who provided “unordered” responses, as discussed in section 4.4. The percent of respondents who provided each pattern of responses is shown for the \$10,000 protective-action cost condition (N = 622), \$20,000 cost condition (N = 611), and the two cost conditions combined (N = 1233). For the \$10,000 and \$20,000 cost-condition results, {RA} denotes respondents who chose to protect when the expected cost of protection was greater than the expected loss (indicating risk-averse behavior), and {RS} denotes respondents who chose to protect when the expected cost of protection was less than the expected loss (indicating risk-seeking behavior).

Row label	Pattern of responses for different percent chances of exceeding the damage threshold (Y = take protective action; N = do not take protective action)				Percent of respondents (\$10,000 cost condition)	Percent of respondents (\$20,000 cost condition)	Percent of respondents (cost conditions combined)
	5%	10%	20%	40%			
<i>a</i>	Y	Y	Y	Y	8% {RA}	6% {RA}	7%
<i>b</i>	N	Y	Y	Y	5%	6% {RA}	6%
<i>c</i>	N	N	Y	Y	11%	16%	13%
<i>d</i>	N	N	N	Y	39% {RS}	34%	36%
<i>e</i>	N	N	N	N	23% {RS}	23% {RS}	23%
<i>f</i>			other		13%	16%	15%

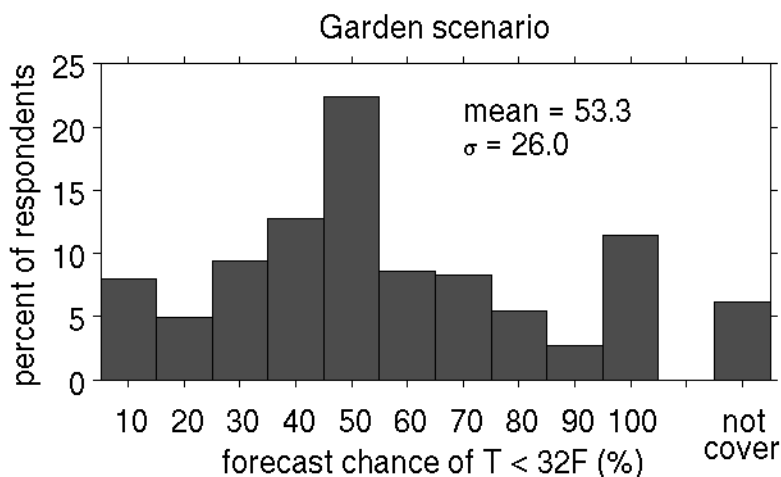
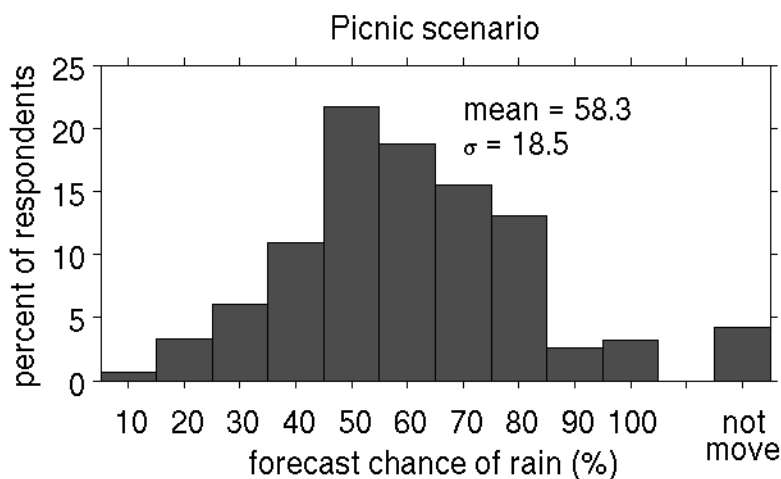


Figure 1. Percent of respondents who selected different probabilistic thresholds for taking protective action in the picnic scenario (upper, N=729) or garden scenario (lower, N=736), along with the percent of respondents who selected not taking protective action (not moving the picnic indoors or covering the garden plants). Means and standard deviations are shown for the respondents who selected 1 of the 10 percent-chance options.

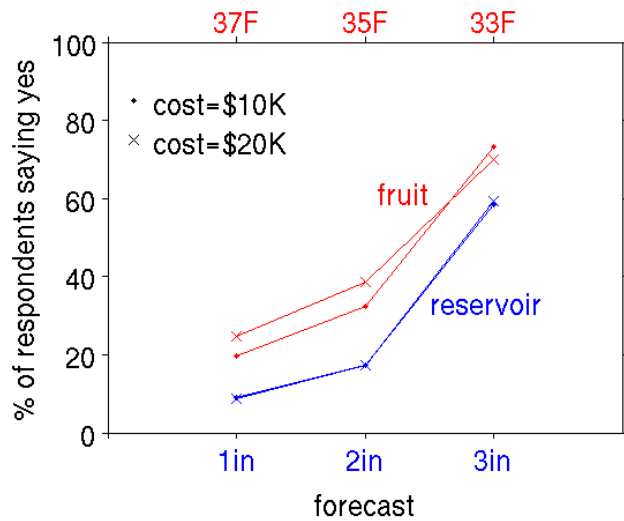


Figure 2. Percent of respondents who said they would take action to protect against possible frost in the fruit scenario (red) or flooding in the reservoir scenario (blue), given the single-value forecast conditions in Table 2. The forecast conditions for the reservoir and fruit scenarios are presented along the bottom and top of the figure, respectively. Results are shown separately for the \$10,000 (•) and \$20,000 (×) cost conditions, for each of the scenarios. Recall that the damage threshold is 4 inches of rain in the reservoir scenario and 32°F in the fruit scenario. (N=1233 for scenarios and cost conditions combined; number of respondents for each scenario / cost condition combination is shown in Table 1.)

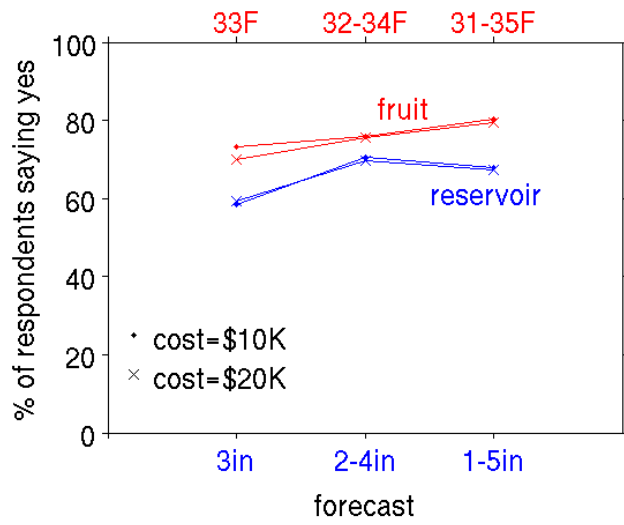


Figure 3. As in Fig. 2, for a single-value forecast and range forecasts symmetric about that single-value forecast (Table 2).

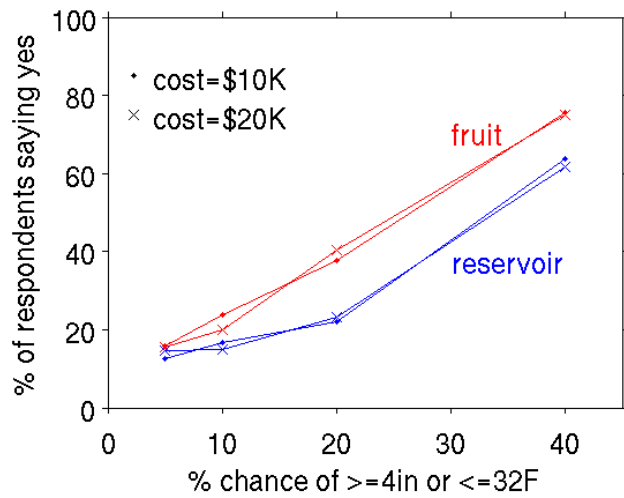


Figure 4. As in Figs. 2-3, for the percentage-chance forecast conditions in Table 2. The x-axis represents the percent chance that there will be four or more inches of rain (in the reservoir scenario) or that the low temperature will be 32°F or lower (in the fruit scenario). For the \$10,000 (\$20,000) cost-condition, the cost-loss decision model says that respondents will choose not to protect at 5% (5% or 10%) chance and to protect at 20% or 40% (40%) chance.