River Forecast Application for Water Management: Oil and Water?

KEVIN WERNER

NOAA/Colorado Basin River Forecast Center, Salt Lake City, Utah

KRISTEN AVERYT

Cooperative Institute for Research in Environmental Sciences and Western Water Assessment, University of Colorado Boulder, Boulder, Colorado

GIGI OWEN

Climate Assessment for the Southwest, The University of Arizona, Tucson, Arizona

(Manuscript received 17 September 2012, in final form 13 March 2013)

ABSTRACT

Managing water resources generally and managing reservoir operations specifically have been touted as opportunities for applying forecasts to improve decision making. Previous studies have shown that the application of forecasts into water management is not pervasive. This study uses a scenario-based approach to explore whether and how people implement forecast information into reservoir operations decisions in a workshop setting. Although it was found that participants do utilize both forecast and observed information, they generally do not utilize probabilistic forecast information in a manner to appropriately minimize risks associated with the tail end of the forecast distribution. This study found strong tendencies for participants to wait for observed information, as opposed to forecast information, before making decisions. In addition, study participants tended to make decisions based on median forecast values instead of considering forecast probability. These findings support the development of quantitative decision support systems to optimally utilize probabilistic forecasts as well as for forecast agencies such as NOAA/NWS to continue investments in work to better understand contexts and environments where forecasts are used or have the potential for use in supporting water management decisions.

1. Introduction

Many factors influence water management decisions including political pressures, legal and policy constraints, infrastructure, and natural and managed water supply. The use of science and/or forecasts in informing water management decisions is often stymied by other considerations including these factors unrelated to science and/or forecasts as well as lack of knowledge or background in the science (Beller-Simms et al. 2008; Rayner et al. 2005). Previous research suggests that forecasts are not used frequently and that forecast usage, when it does occur, is not driven by improvements to forecast skill or enhancements to forecast services (O'Connor et al. 2005). Thus, there appears to be a significant and important gap between forecasts being produced and their actual or potential use by water management agencies. This study primarily aims to examine how water managers interpret and use probabilistic streamflow forecasts, which provide a probability distribution describing the likelihood of future runoff.

Forecast agencies such as the National Oceanic and Atmospheric Administration's National Weather Service (NOAA/NWS) continue to make significant investments in improving streamflow forecast capabilities and skill, often with the goal of improving forecast decision support for water management (Raff et al. 2012). While these types of investments have often resulted in improved forecast skill and enhanced forecast services (such as more forecast points or more frequent updates), it is much less clear how or if they are translated through to improvements to water management decision making.

Corresponding author address: Kevin Werner, NOAA/Colorado Basin River Forecast Center, 2242 West North Temple, Salt Lake City, UT 84116.
E-mail: kevin.werner@noaa.gov

DOI: 10.1175/WCAS-D-12-00044.1

In particular, NOAA/NWS has invested in developing ensemble streamflow forecasts including improved forecast methods (Day 1985; Schaake et al. 2007; Werner et al. 2005) and forecast verification (Brown et al. 2010; Demargne et al. 2010). These improvements have directly led to new probabilistic forecast services, more frequent forecast issuances, and possibly more skillful forecasts. However, both the literature reviewed here and NOAA/NWS experience has shown that only rarely have the improved forecasts translated into improved decision making or even greater forecast utilization by water management agencies.

Application of forecasts to water management decisions varies widely from agencies that seldom or never use forecasts to those that routinely utilize forecasts. Among agencies that explicitly utilize forecasts, applications range from subjective consideration of forecast information to objective utilization of forecast elements including the median forecast or a range of the forecast distribution. An increasingly common forecast application is to formally tie an operational decision to a median forecast value. An example of this is the coordinated operations guidelines for Lakes Powell and Mead on the Colorado River (USBR 2007). These guidelines set decisions related to interstate water allocation and reservoir operations to the median forecasted inflow volumes. Unfortunately this approach does not consider the range represented by the forecast distribution, which leads to large uncertainties when a forecast distribution straddles important threshold values. A less common but more rigorous approach is to consider the forecast distribution explicitly and objectively. Denver Water developed a capability to do this through incorporation of ensemble forecast time series into the spreadsheet model that individually considers each ensemble time series relative to their threshold values allowing the operator to minimize their risk of exceeding those thresholds (R. Steger 2012, personal communication). While this approach is desirable in that allows the operator to objectively manage their risks based on probabilistic forecasts, application of such systems remains rare.

This study examines how water managers and forecasters behave when presented with probabilistic forecasts and given the opportunity to apply those forecasts in a reservoir management decision-making environment. Through our scenario approach, we explore two questions: 1) To what degree do participants use streamflow forecasts to make reservoir management decisions? 2) How effectively do participants use the probabilistic information in the runoff forecasts to manage risk? Using results from a reservoir management scenario exercise, this paper will describe our study of how users apply probabilistic inflow forecasts to a simplified management scenario. We begin by presenting a literature review, followed by an explanation of our research design. Finally, we discuss results from and conclusions of our work.

2. Literature review

While the body of research on the use of forecasts in managing water resources is relatively young in its own right, it draws on diverse research from other disciplines including economics and public administration. One of the original uses proposed for seasonal climate forecasts was as a decision support tool for managing water resources (Glantz 1996). Through the use of forecasts water managers would be able to hedge their reservoir operations in order to, for example, hold more water back in years with dry forecasts to increase chances of filling reservoirs. Similarly, using climate forecasts to optimize water management has been suggested as an important practice for reducing risk in the face of climate change (Ludwig 2012).

However, a collection of mostly interview-based studied has shown that water management agencies are largely unwilling or unable to adopt forecasts into their operations (Lemos 2008; Pulwarty and Redmond 1997; Rayner et al. 2005). For example, Rayner et al. (2005) identified institutional constraints and priorities as the major factors limiting the adoption of climate forecasts in water management agencies in the Pacific Northwest, Southern California, and the Washington, DC, metropolitan areas of the United States. Using a semistructured, snowballing interview approach, they found that water management agencies of various sizes and in all three regions valued delivery reliability and water quality above all other considerations. Respondents indicated that failure to deliver water was simply not tolerated by the agencies, their elected officials, or the public that they serve. Similarly, water delivered must always be of sufficient quality to meet the end use. Although cost was frequently an important consideration, it played a much less significant role than reliability and quality. With the emphasis on reliability and quality, the water management agencies studied were resistant to adapt new technologies to lower costs if there was any potential increase in risk of failure to deliver reliability and quality.

In 2000, researchers working in South Carolina and Pennsylvania's Susquehanna River basin conducted a large-scale survey designed to assess respondents' 1) size of water management responsibilities, 2) perception of their own risk, and 3) perception of forecast skill for both weather and climate forecasts (Dow et al. 2007; O'Connor et al. 2005). The survey populations were substantial in both study areas ($n = 405$ in Pennsylvania and 269 in South Carolina), allowing for significant statistical analysis of the results (O'Connor et al. 2005). The results showed that only small percentages of respondents reported use of climate and/or weather forecasts for most purposes in both states. Only 10%–25% of respondents reported usage of forecasts for planning future water storage needs, expanding distribution systems, adjusting reservoir levels, adjusting inventory supply needs, and similar uses. By contrast, however, managers were more likely to use forecast information for two other purposes: scheduling personnel for maintenance and construction and starting public information campaigns to conserve water. Analysis of survey data also produced another interesting result: perception of risk significantly correlated with respondents' willingness to use forecasts in decision making. In other words, if a particular agency were at some risk of not meeting its delivery or quality requirements, it would be willing to seek new information that it otherwise would not be willing to use. On the other hand, the analysis found that agency size and perception of forecast skill did not correlate well with actual use of or willingness to use climate forecasts. Therefore, forecast usage is more determined by perception of risk and recent experiences with observed weather and climate than by any improvements in forecast skill or the potential value of the forecasts to the end user.

Previous research has also examined the difficulties in communicating probabilistic forecasts both to forecasters (Demeritt et al. 2010) and to forecast consumers (Ramos et al. 2010). These studies document the difficulty in communicating probabilistic forecasts. Demeritt et al. (2010), for example, interviewed operational river forecasters on their use and perceptions of probabilistic forecasts in Europe. They found many forecasters were skeptical of the ability of forecast users to understand the uncertainty expressed in the probabilistic forecasts. Many forecasters also expressed skepticism themselves with the meaning of the uncertainty estimates produced by the forecast models. Economists such as Daniel Kahneman have similarly noted that as humans "we easily think associatively, we think metaphorically, we think causally, but statistics requires thinking about many things at once" (Kahneman 2011, p. 13). Better understanding how forecast consumers understand and potentially use probabilistic forecasts is clearly important in this study.

Summarizing the state of practice for application of seasonal to interannual forecasts in water resources management, Beller-Simms et al. (2008) noted the traditional "loading dock" model in which forecasts are produced and placed on a loading dock for consumers to use has not effectively promoted forecast usage. Instead of a loading dock model, Beller-Simms et al. (2008) and others note the importance of improving our collective understanding of how forecasts in particular and science more generally can be effectively applied to decision making in a collaborative manner that includes the forecast/science producer. Other studies (e.g., Feldman and Ingram 2009) have similarly examined the application of science to water resources management decision making and have concluded that more emphasis should be placed on employing social science methods to understanding and improving that process.

One common element from previous studies is that rational choice theory does not appear to apply. Rational choice theory, sometimes called rational-comprehensive, requires that all options be systematically studied before decisions are made at every time step relevant to the decision-making unit (Lindblom 1959). Rayner et al. (2005) explicitly described and addressed rational choice theory as it applies to using forecasts. The early optimism from the climate research and forecasting communities was essentially based on rational choice theory whereby decision makers would be strongly motivated by an interest to optimize performance through the application of all available and relevant information to the required decision (Glantz 1996). Under rational choice theory, water management agencies would continually search out and apply new information and new forecasts. New information sources would be analyzed and weighted according to their value such that forecasts—even those with considerable uncertainty—could be leveraged into improving the decision-making process incrementally. However, as virtually all the previous studies have found, water management agencies do not do this. These results are consistent with examples from other areas of public administration showing that rational choice theory is rarely adhered to in real-world decision making (Lindblom 1959; Simon 1946; Wildavsky 1969). Instead, as Simon (1946) and others have described, decision makers more typically rely on a combination of previous experience, political considerations, and consider incremental changes to the status quo based on their organizational limitations, time availability, and other practical considerations.

Another common thread to previous research was that forecast usage, when it did occur, was largely motivated by the perception of risk to climate or weather on meeting water delivery requirements or water quality standards. O'Connor et al. (2005) documented this explicitly through their survey results. Rayner et al. (2005) speculated that water shortages or long-term drought could cause agencies to seek out new information sources such as climate forecasts. Water management agencies that do not perceive a risk from future climate or weather are very unlikely to seek out forecasts to integrate into their operations.

3. Method

This study utilizes a scenario-based research design to understand how reservoir managers and forecast providers interact with probabilistic forecasts of reservoir inflow to make operational decisions managing a reservoir. We were interested in exploring two questions: 1) To what degree would participants use streamflow forecasts to make reservoir management decisions? 2) How effectively participants would use the probabilistic information in the runoff forecasts to manage risk? We developed two reservoir operations scenarios that prescribed operations risks through reservoir capacity and monthly release constraints with the goal of keeping the reservoir as full as possible. We simulated time by alternately providing probabilistic monthly inflow forecasts and actual inflow volumes iteratively through a runoff period. By comparing participant release schedules with historical median inflows and forecasted inflows, we are able to qualitatively assess the extent to which participants relied on forecasts in their decision making and, in particular, using probabilistic forecasts to inform reservoir management. While our data do not represent laboratory-quality data, our quasi-experimental approach provides a rare opportunity to explore improving the use of probabilistic forecasts in reservoir management.

Our data come from three workshops conducted between January 2011 and November 2011. These workshops were intended primarily as educational opportunities for forecast users and forecast providers. In all three instances of the workshop, the scenario exercise was conducted as part of a larger effort to inform forecast users of the breadth of forecast information available. Other workshop topics included lectures explaining the science behind ensemble streamflow forecasts and the weather and climate forecasts that support the streamflow forecasts, forecast verification results to provide context on forecast skill, and usability exercises on NOAA/NWS websites where forecasts and other information may be obtained. Participants were largely self-selecting based on their interest in learning more about forecast methods and applications. Given both the context of the workshops and the self-selection nature of participation, we expect that the rate of forecast usage in our scenario exercise should be greater than in the real world.

In total there were 60 participants in three workshops. Of those, 51 participants completed the scenario exercise whose results are reported here. The workshops themselves were conducted in different environments and attracted different types of participants. The first workshop in January 2011 was a short course held in conjunction with the American Meteorological Society annual meeting. The second workshop in August 2011 was targeted toward water managers in the state of Utah and followed an anomalously large runoff year. The third workshop in November 2011 was part of a NOAA/NWS training course and attracted mostly forecast providers. Demographic information including professional affiliation, geographic area, and experience with forecasting and/or water management were collected from each participant for the first two workshops but not the third where the format of the workshop precluded the collection of this information. This information was used to assess correlations with successful forecast usage and relevant work experience. Examining both proposed and actual reservoir releases relative to the historical averages and capacity of the reservoir assessed forecast usage.

The exercise scenarios were developed to simulate decision making in reservoir operations. Participants were presented with reservoir management scenarios and a series of forecasts describing probability functions for monthly runoff volumes for a mountain basin whose runoff is dominated by snowmelt. At the beginning of the exercise, participants were told that their "job" was to fill the reservoir as much as possible by the end of the runoff season without exceeding monthly minimum and maximum release constraints. An inexpensive prize was offered as an incentive to the participant who had the highest reservoir at the end of the runoff season without exceeding the maximum or minimum release constraints in the exercise or overtopping the reservoir. Participants were only allowed to set monthly releases within a prescribed range meant to reflect minimum instream flow requirements on the low end and flooding mitigation on the high end. The range of releases allowed in the exercise was 15 KAF to 60 KAF month^{-1} (1 KAF = 1000 acre-feet). In addition, participants that exceeded the capacity of the reservoir (500 KAF) at the end of any month were excluded from consideration from the winning incentive.

At the beginning of the exercise, participants were provided with the 1 March reservoir state, a probabilistic forecast of monthly volumes extending into the summer months, and the historical average inflows for each month and the operation constraints for the reservoir (e.g., minimum and maximum releases). Reservoir volumes are provided and tracked in units of KAF and presented using a series of box-and-whisker plots similar to those shown in Fig. 1. These conventions were chosen specifically to match the forecast paradigm adopted by NOAA's experimental water resources outlook as described by Beller-Simms et al. (2008) and outlined on the

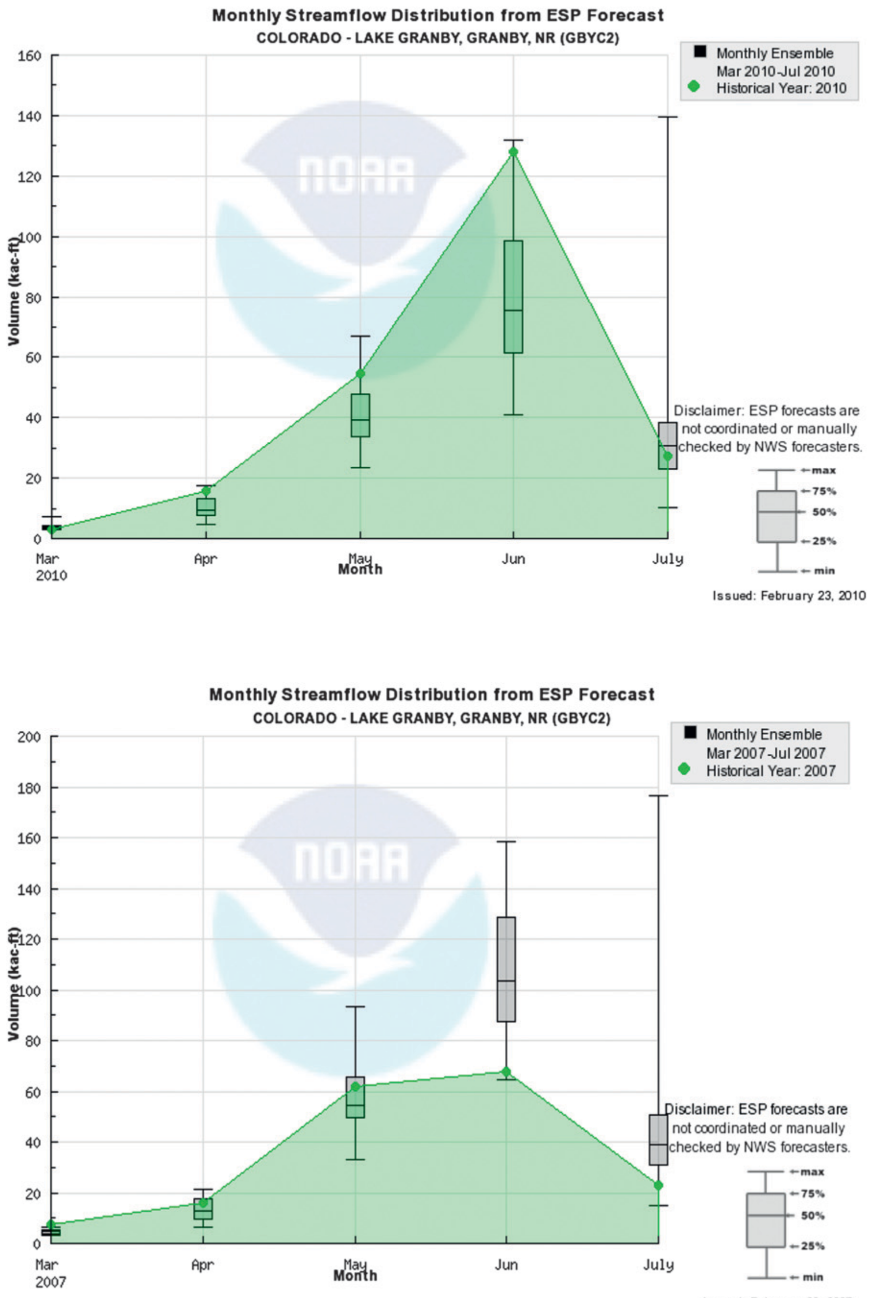

FIG. 1. Monthly inflow forecasts for scenarios (a) A and (b) B illustrating forecast information as presented to participants. The box-and-whisker plots depict the forecast probability for monthly inflows. The green line is the observed inflow. Figures reprinted with permission from NOAA.

NOAA water resources outlook (e.g., wateroutlook.nwrfc.noaa.gov). Both a description of and a hands on opportunity to use the water resources outlook website was included as part of all three workshops as well. The 1 March reservoir state provided was 40 KAF. Participants were asked to provide their monthly release schedule. Each participant would then be given the actual inflow for March. From this, participants would calculate the 1 April reservoir contents using a simple water balance equation:

$$S_j = S_i + Q_{\mathrm{in}} - Q_{\mathrm{out}},$$

where, for any month i, S is the reservoir storage, Q_{in} is the monthly inflow volume, and Q_{out} is the monthly release volume. After calculating their new reservoir content for each simulated month, participants were given an updated forecast from that month forward for the remainder of the runoff season and asked to update their release schedule accordingly. The exercise continued in this manner through the end of July at which point the ''winner'' was determined by comparing the participants' final reservoir levels among those that did not exceed the release constraints.

The scenarios themselves were chosen to demonstrate cases where the observed inflows were near the tail of the forecast probability distribution in order to test our second goal (e.g., how effectively do participants manage risk in a probabilistic forecast environment). The scenarios were based on the actual forecasts and inflows for Lake Granby, Colorado, in 2010 (scenario A) and 2007 (scenario B). Minor modifications such as rounding were made to the actual forecasts and data for those years to simplify the math required in the actual exercise. Figure 1 shows the 1 March forecasts and corresponding observed inflow volumes for both scenarios. A cool, wet May in 2010 contributed to observed inflows at the high end of the forecast distribution whereas a dry spring in 2007 contributed to observed inflows at the low end of the forecast distribution in that year. In 2010, not only were the inflows at the high end of the forecast distribution, but they were also somewhat late in materializing. June 2010 experienced the largest inflows and these were at the very highest end of the forecast distribution. Thus, the reservoir operation challenge for scenario A was to recognize the possibility of the high end of the forecast distribution actually occurring in the reservoir release schedule. Scenario A presented a significantly more challenging reservoir management situation in which we expected to see more variance in the results developed by participants. Therefore, we assigned roughly twice as many participants to scenario A than to scenario B. Participants were assigned to scenarios without regard to any other characteristic; 35 participants completed scenario A while 16 completed scenario B.

The format of the scenario exercise was largely consistent between workshops with two exceptions. First, based on input from the first workshop where participants requested historical contextual information about the inflows, participants in the second two workshops were provided with the historical median inflows. Second, third workshop participants worked together in teams of two whereas they worked individually in the first two workshops. This change was intended to simulate real-world decision making where reservoir operation decision makers more commonly work in group settings than individually. In all three workshops, participants largely worked in isolation from the other participants. The reward structure was also very similar between the three workshops.

Given the study design, results are not generalizable to all water managers. We do not claim representativeness or attempt to calculate statistical significance. Importantly, participants in this exercise were specifically asked to use forecasts and were intentionally not provided with the plethora of other contextual information that surrounds water management decision making in the real world. This factor together with the format and design of the workshops, undoubtedly, led to a higher rate of forecast usage than would be observed in the real world. Participants were not randomly selected from the larger population of forecast consumers and stakeholders. Rather, they voluntarily chose to participate in one of the three workshops. Because of the geographic nature of the workshops, participants from the western United States were overrepresented. Also, the snowmelt nature of the exercise lends itself to people with experience in the western United States more so than other regions. Nonetheless, the results from this study have currency and relevance in improving understanding of how people use probabilistic forecasts.

4. Results and discussion

This section analyzes both the initial release schedules proposed by participants and the release schedule actually implemented by participants to assess their use of forecasts. The analysis of the releases proposed and implemented by the participants allows us to gain an understanding of how participants are using the information provided to them in the scenario (including forecasts) to make reservoir operations decisions. Through comparisons over time, between groups of participants, with climatological inflows, and with the forecasted inflows, we are able to qualitatively assess which information sources participants are most commonly using as well as how effective those information sources are for the decision space within the scenario.

Actual reservoir contents at the end of each month were calculated based on the participants' actual releases as part of the exercise. Figure 2 shows the reservoir contents separately for all scenario A (Fig. 2a) and B (Fig. 2b) participants.

For scenario A, only 5 out of 35 participants (14%) were able to avoid overtopping the reservoir. This result demonstrates that most participants did not plan for the worst-case scenario or even a high extreme scenario within the forecast distribution. Instead, most participants

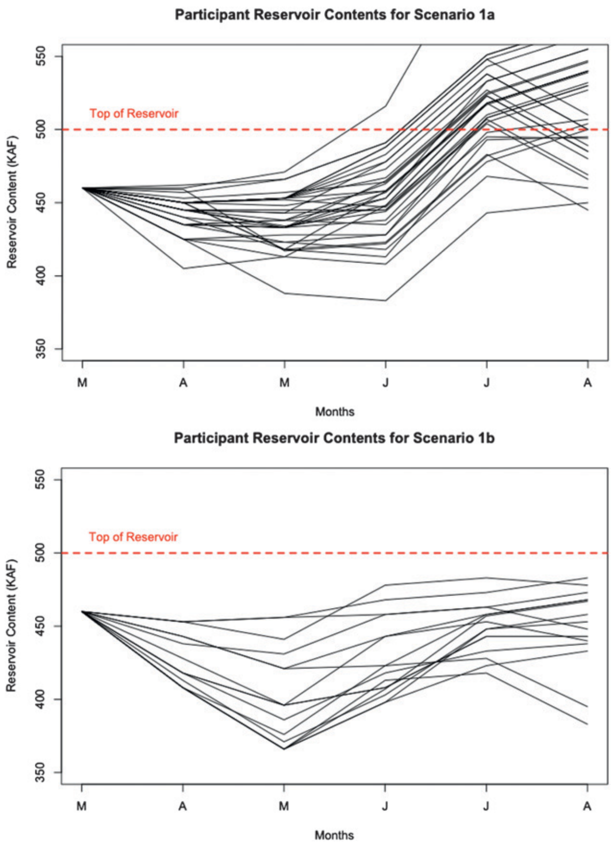

FIG. 2. Monthly reservoir storage amounts for scenario (a) A and (b) B participants. The red line in (a) indicates the top of the reservoir.

in the exercise took a ''wait and see'' approach toward responding to extreme forecasts. Rather than a robust consideration of available data, most participants instead relied on small subset. Once high inflows became apparent in May, most participants were not able to evacuate sufficient space from their reservoir to contain the very high inflows in June. This observed tendency in the exercise is also apparent in many water management practices that prescribe management actions only for water on the ground or in the channel rather than based on forecasted flows (Raff et al. 2012).

In scenario B (Fig. 2b), none of the participants overtopped their reservoir. Interestingly, by May in this scenario, the participants had formed distinct groups. One group drew the reservoir down early in the exercise (e.g., March and April releases were large) while the other group held their reservoirs at about the same level as the start of the exercise. Our demographic data show that the former group was composed of participants with reservoir management experience while the second group was not, suggesting the importance of reservoir operation experience. Unfortunately, we were limited by our

TABLE 1. Participants (left) proposing and (right) implementing sufficient total releases to avoid dam overtopping. Note that dam may still overtop with sufficient total releases if release timing is not optimal.

	Number and percentage of participants proposing sufficient releases	Number and percentage of participants implementing sufficient releases
Scenario A (n = 35)	9 (26%)	14 (40%)
Scenario B (n = 16)	14 (88%)	16 (100%)

demographic data that only coarsely described our participants and the overall sample size in further pursuing how participant background influenced their choices in the scenario exercise. Clearly this is a fertile ground for further investigation.

Next we compare participants' proposed reservoir operation at the start of the exercise to their implemented operation at the end of the exercise. Table 1 reports number of participants proposing and implementing sufficient releases to avoid overtopping the reservoir for each scenario. In both scenarios, the number of participants implementing sufficient releases is greater than the number of participants initially proposing sufficient releases. This indicates that at least some participants adapt their plans as new information becomes available through the course of the exercise. Given that the vast majority of participants in scenario A were not able to modify their plans sufficiently to avoid overtopping and that the inflow forecasts did not change dramatically, this result suggests that participants were more likely modifying their plans based on actual inflows reported rather than on changing forecasts.

Release schedules were compared with historical median inflows and with forecasted inflows to assess the extent to which participants hedged risk or followed historical or forecasted inflows. Figure 3 shows the total March through July seasonal reservoir releases initially proposed by each participant. In the exercise (as would be the case in the real world), participants were free to change their future schedules as simulated time advanced and they were given new information (e.g., revised monthly forecasts and actual inflows). Proposed seasonal releases are plotted together with the excess forecasted inflow, actual excess inflow, and historical median excess inflow. Excess inflow is calculated by subtracting the beginning available reservoir storage capacity (40 KAF) from the inflow value.

For scenario A, the actual excess inflow was 62 KAF more than the median forecast but 125 KAF less than the 10% forecast. Reservoir release plans with volumes

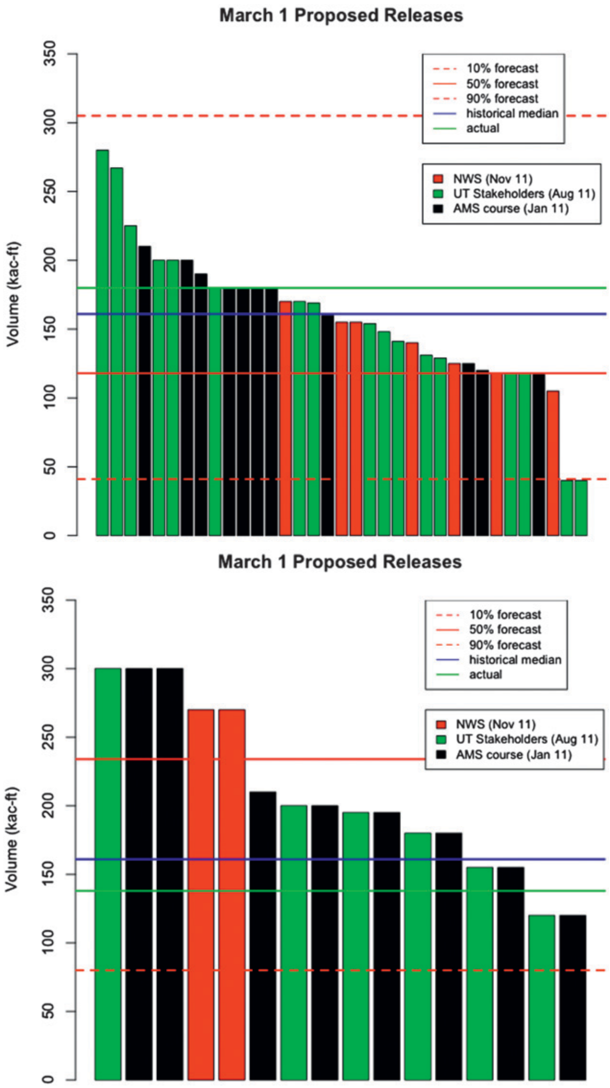

FIG. 3. Total releases in 1 March release schedules developed by participants ranked from largest to smallest. Participants are color coded by workshop, for scenarios (a) A and (b) B. Also included are the seasonal inflow volume forecast (in red horizontal lines), actual seasonal inflow (green), and historical median inflow (blue). These reference lines account for the initial 40 KAF of space by subtracting that value.

less than the actual excess inflow would overtop the reservoir. A majority of participants (22 of 35) in scenario A proposed releases less than the actual excess inflow in their 1 March plan. An additional five participants proposed releases that equaled the actual excess inflow. The timing of the inflows relative to the proposed releases would determine whether the participants that proposed releases equal to or greater than the actual excess inflow would overtop the dams. For example, a participant with a total proposed release of 190 KAF would release more water over the season than the excess actual inflow. However, if the timing of the releases was such that insufficient space was evacuated from the reservoir to contain the 120 KAF of inflow in June, their reservoir would still overtop.

While the majority of scenario A participants proposed to release less than the actual excess inflow, all but three participants (8.5%) proposed to release at least the median forecasted excess inflow. Several participants proposed scheduled releases nearly equal to the median forecast. Fewer participants proposed releases nearly equal to the historical median runoff. This indicates that most participants were aware of and incorporated the forecasted inflow and fewer participants accounted for historical streamflow.

In scenario B the actual excess inflow (178 KAF) was much less than the median forecast (274 KAF). In this situation, all participants were able to avoid overtopping their reservoirs. Interestingly, most participants initially proposed releasing less water than the excess median forecast value. Had the median forecast verified, these participants would have overtopped their reservoirs. One possible rationale for proposing releases less than the median excess inflow forecast could be that participants were instead keying off the median historical excess inflow (160 KAF). Many of the participants proposed initial releases near this historical median value. However, only participants in the second two workshops had access to the historical data.

The results in Fig. 3 show a general tendency for people to focus on the median forecast when presented with a series of probabilistic forecasts, and the reliance on historical data in decision making. The cluster of proposed releases near the median forecasted excess inflow demonstrates this tendency. However, at least in the scenario A results, there was also some tendency for people to hedge somewhat—albeit not enough in most cases—either toward the higher end of the forecast distribution given the incentive structure to not overtop the dam or toward utilizing the historical data. Instead of following a 50% forecast they planned for the 40% forecast even though that meant at least a 40% chance of failure given the uncertainty associated with timing. These results suggest that most participants are aware of the need to hedge and will act on that awareness. This observation is consistent with previous literature finding that rather than a robust consideration of all available data and forecasts (e.g., rational choice theory), participants instead commonly consider recent past experience and simplified data (e.g., Lindblom 1959; Simon 1946).

Next, we analyze the distribution of participants' implemented (actual) releases. Figure 4 shows these distributions following the same convention as Fig. 3. These figures show that the range of implemented releases was

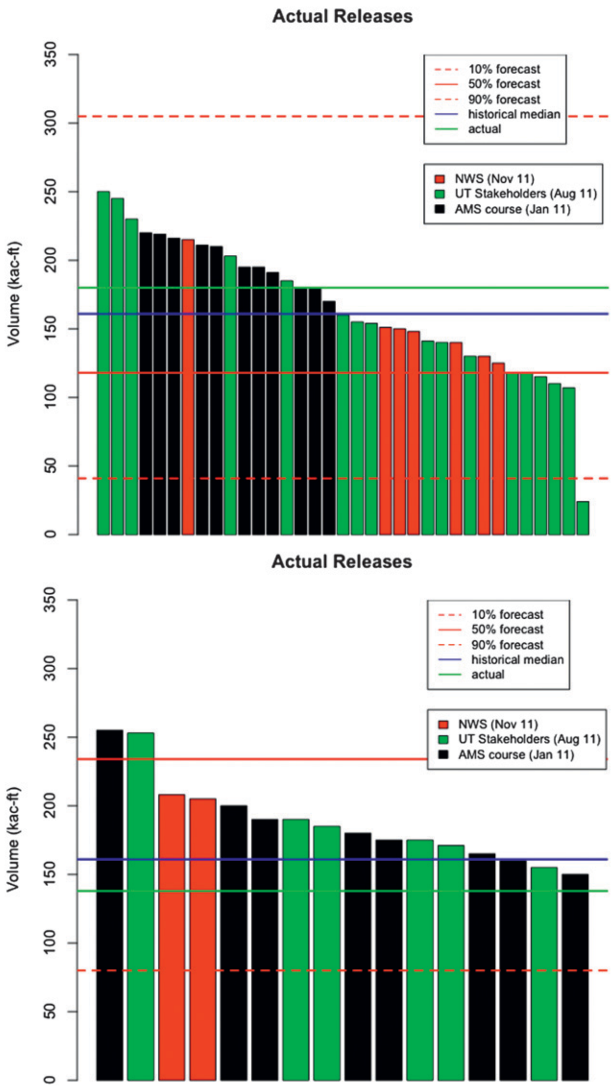

FIG. 4. Total implemented releases by participants ranked from largest to smallest. Plotting conventions are as in Fig. 3.

narrower than the proposed releases for both scenarios but especially so for scenario B. Consistent with the results in Table 1, these results show a tendency for participants to adapt their plans to new information as the exercise progressed. It also shows that the participants tended to coalesce around similar solutions especially for scenario B. This later result is not terribly surprising given the incentive to fill the reservoir and lower than forecasted volumes in scenario B.

Participants at the NWS training proposed and implemented releases with less variance between participants than those from the other two workshops. This may have been the result of the exercise structure wherein participants worked together in teams of two people instead of independently as in the other two workshops. These teams may have served to temper some of the outlying release schedules that were submitted by participants in the other two workshops.

Participants in the AMS course adapted their release schedules in scenario A more significantly than in the other two workshops. In their initial releases (Fig. 3a), AMS course participants were represented throughout the distribution of all participants. However, in their actual releases (Fig. 4a), AMS course participants had shifted within the total participant population such that all AMS course participants were in the higher half of the distribution. We speculate that these participants benefited from material presented elsewhere in their workshop that was not presented in the other two workshops.

Our results suggest a potential application for objective decision support tools for utilizing complex and uncertain information such as that employed by Denver Water. The incentive structure in our study encouraged participants to avoid overtopping their dams as a first priority while filling to the maximum height possible as a second priority. Incentive structures in the real world are often much more complex. In both cases, there is a clear role for both a human decision maker that can intelligently respond to a decision support system that effectively seeks to optimize possible action against predetermined criteria as well as an objective tool that utilizes probabilistic forecasts to minimize risk functions while maximizing opportunity functions. The scenario A results demonstrate the importance planning for instances where the tail of the forecast distribution is realized. Future work could employ a decision support tool to demonstrate effective management techniques that plan for the contingency realized in scenario A. Those results could be used together with this exercise to more rigorously test the use of decision support tools but also to demonstrate their value to future exercise participants.

Our study shows, consistent with previous work (e.g., O'Connor et al. 2005), that forecast agencies cannot take for granted that forecasts are understood or applied in the manner that forecast agencies intend. Second, people generally do not consider the tails of a forecast probability distribution in their decision making even in cases such as our scenario where penalties for system failure are extreme. In the real world, extreme events that are often responsible for failing systems are often represented in the tail ends of the forecast distribution. One effective strategy for managing extreme events is an appropriate consideration of the full forecast probability function as in the Denver Water capability (Steger 2011). Finally, in order for forecast agencies to maximize the value of their forecast products, they must invest in partnerships with forecast users to collaboratively understand how forecasts are being used, address misunderstandings, and look for opportunities for improvement. In the scenario presented in this case study, for example, the forecast agency could have played an important role in working with the reservoir operator to understand the potential impact of realizing the high end of the forecast distribution. Understanding and supporting decision making in this context is an important role for a forecast agency to play.

More work is needed to demonstrate the value of decision support systems in decision-making contexts similar to that used in our exercise is needed. While water management agencies are increasingly relying on ensemble streamflow forecasts and calling for improvements in forecast skill and reliability, more work is needed to fully understand the value that a decision support system can add. Ultimately human operators and decision makers will continue to play important roles in water management. Understanding how to most effectively leverage science and forecasts into that context is important component of improving resiliency and reliability of our water management infrastructure.

Acknowledgments. The authors thank Eric Gordon and the anonymous reviewers for their invaluable insights on this manuscript.

REFERENCES

Beller-Simms, N. K., H. Ingram, D. Feldman, N. Mantua, K. L. Jacobs, and A. M. Waple Eds., 2008: Decision-support experiments and evaluations using seasonal-to-interannual forecasts and observational data: A focus on water resources. U.S. Climate Change Science Program Synthesis and Assessment Product 5.3 Rep., 192 pp.

Brown, J. D., J. Demargne, D. J. Seo, and Y. Liu, 2010: The Ensemble Verification System (EVS): A software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations. *Environ. Modell. Software,* **25,** 854–872.

Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plann. Manage.,* **111,** 157–170.

Demargne, J., J. Brown, Y. Liu, D. J. Seo, L. Wu, Z. Toth, and Y. Zhu, 2010: Diagnostic verification of hydrometeorological and hydrologic ensembles. *Atmos. Sci. Lett.,* **11,** 114–122.

Demeritt, D., S. Nobert, H. Cloke, and F. Pappenberger, 2010: Challenges in communicating and using ensembles in operational flood forecasting. *Meteor. Appl.,* **17,** 209–222.

Dow, K., R. E. O'Connor, B. Yarnal, G. J. Carbone, and C. L. Jocoy, 2007: Why worry? Community water system managers' perceptions of climate vulnerability. *Global Environ. Change,* **17,** 228–237.

Feldman, D. L., and H. M. Ingram, 2009: Making science useful to decision makers: Climate forecasts, water management, and knowledge networks. *Wea. Climate Soc.,* **1,** 9–21.

Glantz, M., 1996: *Currents of Change: El Niño's Impact on Climate and Society.* Cambridge University Press, 194 pp.

Kahneman, D., 2011: *Thinking, Fast and Slow.* Farrar, Straus and Giroux, 512 pp.

Lemos, M. C., 2008: What influences innovation adoption by water managers? Climate information use in Brazil and the United States. *J. Amer. Water Resour. Assoc.,* **44,** 1388–1396.

Lindblom, C. E., 1959: The science of ''muddling through.'' *Public Adm. Rev.,* **19,** 79–88.

Ludwig, F., 2012: Using seasonal climate forecasts for water management. *Climate Change and Adaptation in the Water Sector,* F. Ludwig et al., Eds, Earthscan, 79–86.

O'Connor, R. E., B. Yarnal, K. Dow, C. L. Jocoy, and G. J. Carbone, 2005: Feeling at risk matters: Water managers and the decision to use forecasts. *Risk Anal.,* **25,** 1265–1275.

Pulwarty, R. S., and K. T. Redmond, 1997: Climate and salmon restoration in the Columbia River basin: The role and usability of seasonal forecasts. *Bull. Amer. Meteor. Soc.,* **78,** 381–397.

Raff, D., L. Brekke, K. Werner, A. Wood, and K. White, 2012: Short-term water management decisions: User needs for improved climate, weather and hydrologic information. Climate Change and Water Working Group Rep., 233 pp. [Available online at http://www.ccawwg.us/index.php/activities/short-term-water-management-decisions-user-needs-for-improved-climate-weather-and-hydrologic-information.]

Ramos, M. H., T. Mathevet, J. Thielen, and F. Pappenberger, 2010: Communicating uncertainty in hydro-meteorological forecasts: Mission impossible? *Meteor. Appl.,* **17,** 223–235.

Rayner, S., D. Lach, and H. Ingram, 2005: Weather forecasts are for wimps: Why water resource managers do not use climate forecasts. *Climatic Change,* **69,** 197–227.

Schaake, J., and Coauthors, 2007: Precipitation and temperature ensemble forecasts from single-value forecasts. *Hydrol. Earth Syst. Sci. Discuss.,* **4,** 655–717.

Simon, H. A., 1946: The proverbs of administration. *Public Adm. Rev.,* **6,** 53–67.

USBR, 2007: Record of decision: Colorado River interim guidelines for lower basin shortages and the coordinated operations for Lake Powell and Lake Mead. U.S. Bureau of Reclamation, 59 pp. [Available online at http://www.usbr.gov/lc/region/programs/strategies/RecordofDecision.pdf.]

Werner, K., D. Brandon, M. Clark, and S. Gangopadhyay, 2005: Incorporating medium-range numerical weather model output into the ensemble streamflow prediction system of the National Weather Service. *J. Hydrometeor.,* **6,** 101–114.

Wildavsky, A., 1969: Rescuing policy analysis from PPBS. *Public Adm. Rev.,* **29,** 189–202