BAMS Article

Innovations in Winter Storm Forecasting and Decision Support Services

David R. Novak, Sarah E. Perfater, Julie L. Demuth, Stephen W. Bieda III, Gregory Carbin, Jeffrey Craven, Michael J. Erickson, Matthew E. Jeglum, Joshua Kastman, James A. Nelson, David E. Rudack, Michael J. Staudenmaier, and Jeff S. Waldstreicher

ABSTRACT: Winter storms are disruptive to society and the economy, and they often cause significant injuries and deaths. Innovations in winter storm forecasting have occurred across the value chain over the past two decades, from physical understanding, to observations, to model forecasts, to postprocessing, to forecaster knowledge and interpretation, to products and services, and ultimately to decision support. These innovations enable more accurate and consistent forecasts, which are increasingly being translated into actionable information for decision-makers. This paper reviews the current state of winter storm forecasting in the context of the U.S. National Weather Service operations and describes a potential future state. Given predictability limitations, a key challenge of winter storm forecasting has been characterizing uncertainty and communicating the forecast in ways that are understandable and useful to decision-makers. To address this challenge, particular focus is placed on establishing a probabilistic framework, with probabilistic hazard information serving as a foundation for winter storm decision support services. The framework is guided by social science research to ensure effective communication of risk to meet users' needs. Solutions to gaps impeding progress in winter storm forecasting are highlighted, including better understanding of mesoscale phenomenon, the need for better ensemble calibration, a rigorous and consistent database of observed impacts, and linking multiparameter probabilities (e.g., probability of intense snowfall rates at rush hour) with users' information needs and decisions.

KEYWORDS: Ensembles; Winter/cool season; Operational forecasting; Probabilistic Quantitative Precipitation Forecasting (PQPF); Communications/decision making; Decision support

https://doi.org/10.1175/BAMS-D-22-0065.1
Corresponding author: David R. Novak, david.novak@noaa.gov
In final form 1 February 2023

AFFILIATIONS: Novak, Carbin, Kastman, and Nelson—NOAA/NWS/Weather Prediction Center, College Park, Maryland; **Perfater and Bieda**—NOAA/NWS Headquarters, Silver Spring, Maryland; **Demuth**—National Center for Atmospheric Research, Boulder, Colorado; **Craven and Rudack**—NOAA/NWS Meteorological Development Laboratory, Silver Spring, Maryland; **Erickson**—Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado, and NOAA/NWS/Weather Prediction Center, College Park, Maryland; **Jeglum and Staudenmaier**—NOAA/NWS Western Region Headquarters, Salt Lake City, Utah; **Waldstreicher**—NOAA/NWS Eastern Region Headquarters, Bohemia, New York

Winter storms are disruptive to society and the economy and are often deadly. Major storms can immobilize transportation, knock out power, and shutter stores and schools for days. In a study of winter storms affecting New York City, Hosterman et al. (2019) and Lazo et al. (2020) quantified the economic impact of just two storms at over $200 million for the aviation and energy sectors alone. The historic February 2021 Southern Plains snow and cold wave was the most destructive and costly winter storm to affect the United States in recorded history, with 210 deaths in Texas alone, and $24 billion in direct losses overall (NCEI 2021). Many of the deaths and damages in this event were caused by cascading failures in the power, water, and transportation infrastructure. Winter storms also have dramatic impacts on human health (falls, cardiac events, hypothermia) (Mills et al. 2020; Lin et al. 2021). In agricultural regions, severe winter storms lead to animal and crop losses (Zhang and Liang 2021).

Winter storms have a profound effect on transportation, leaving commuters cold, hungry, and stranded in vehicles, trains, and planes. Recent episodes include Chicago (February 2011); Atlanta (January 2014); Washington, D.C. (January 2011, 2016); New York (February 2013, November 2018); northern Virginia (January 2022); and Buffalo, New York (December 2022). Black and Mote (2015) documented an annual average of over 800 winter-related vehicle accident fatalities (direct and indirect), with a majority being caused due to road conditions or visibility reductions from blowing snow. In a study of 196 variable-length winter storm events, Mills et al. (2019) highlight that injury and noninjury vehicle collisions increased by 66% and 137%, during winter storms relative to dry weather conditions at comparable times of day and days of the week. Tobin et al. (2021) find that the largest fraction of winter-weather-related road fatalities is associated with deteriorating weather conditions, suggesting that adverse changes in weather may play a role in a large number of road fatalities. Call and Flynt (2021) studied the impact of snowfall on the New York State Thruway and found that for every 5.1 cm of snowfall, up to 2.6 additional crashes occurred. Thus, improved forecasts of the timing, location, and severity of winter storms can decrease traffic volume and improve public safety during these events (e.g., Knapp et al. 2000).

At the close of the twentieth century, numerous studies reported on the state of winter storm forecasting (Maglaras et al. 1995; Gurka et al. 1995; Keeter et al. 1995; Niziol et al. 1995; Kocin and Uccellini 2004a). Over the past two decades, new observations, improved models, and advanced forecasting tools have been developed. Emphasis on collaborative research partnerships has accelerated the transition of research to operations (R2O) (Waldstreicher 2005; Jacobs 2021). These efforts have translated into nearly a doubling of

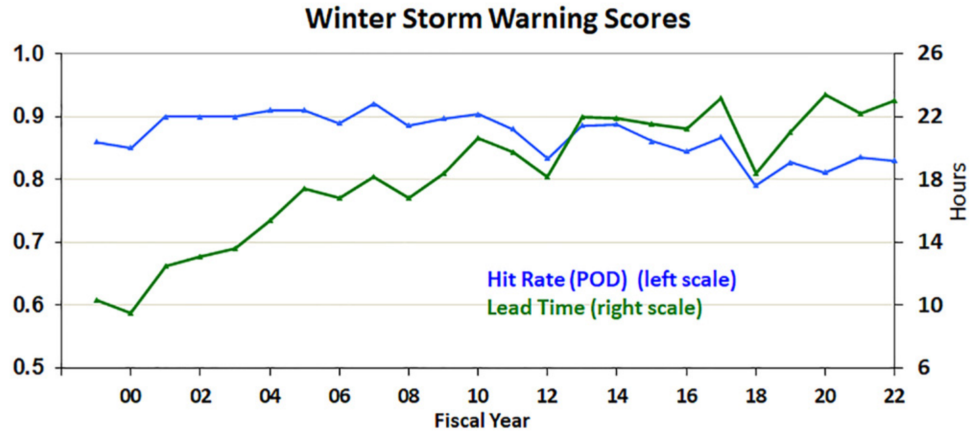

Fig. 1. Trend of the NWS Government Performance Results Act scores for Winter Storm Warning lead time (h; green) and POD (blue).

lead time for official NWS Winter Storm Warnings—approaching a full day—while maintaining a POD largely above 80% (Fig. 1).

Public safety officials are making more complex decisions and require more accurate, consistent, and specific forecasts with greater lead time. Uccellini and Ten Hoeve (2019) emphasize that accurate forecasts can improve preparedness and responsiveness to extreme weather events when the forecasts are understood and acted upon by public safety decision-makers. This connection between forecasts and key decision points in the emergency management, water resource management, and public safety communities has been termed impact-based decision support services (IDSS) and is the lynchpin that connects science, technology, forecasts, and warnings to societal outcomes (Uccellini and Ten Hoeve 2019). Specifically, IDSS is defined in NWS policy as "the provision of relevant information and interpretative services to enable core partners'[1] decisions when weather, water, or climate has a direct impact on the protection of lives and livelihoods" (NWS 2019). Such connection is built on a foundation of trusted relationships, including repeated in-person interactions well in advance of an event to understand partners' needs and decisions, "shoulder-to-shoulder" engagement during an event, and assessments after an event.

A key challenge of winter weather IDSS is characterizing and communicating uncertainty. As noted by Kocin and Uccellini (2004a, chapter 7), winter storms are the rare product of a combination of synoptic and mesoscale processes that come together at just the right time and place. Accordingly, the predictability of winter storms varies by time frame (typically more predictable closer to the event), by scale (an area of snowfall is more predictable than intense mesoscale snowbands embedded within the area of snow), by phenomenon (e.g., lake effect, orographic precipitation, mesoscale snowbands), and even among the same phenomenon from event to event [see predictability analysis of the December 2010 nor'easter (Zheng et al. 2013; Kocin et al. 2011) and the January 2015 and 2016 nor'easters (Greybush et al. 2017)]. Moreover, subtle changes in the meteorology can dramatically change the impacts. For example, a subtle shift of just 40 km in the rain–snow line along the Northeast urban corridor dramatically changes the impacts for ~40 million people. Effectively characterizing and communicating these meteorological and impact-based uncertainties in ways that are understandable is complex and challenging—but essential to do.

This paper highlights the current state of winter storm forecasting and the ongoing evolution toward a probabilistic framework to support user decisions. The probabilistic framework includes forecasts explicitly conveyed with probabilities as well as forecasts

[1] NWS Core Partners are defined as government and nongovernment entities that are directly involved in the preparation, dissemination, and discussions involving weather-, water-, or climate-related National Weather Service information that supports decision-making for routine or episodic high-impact events.

conveyed through ranges, scenarios, and categories based on the underlying probabilities. We further articulate a vision of providing decision-makers skillful and quantitative information on the likelihood of *impacts* from winter storms. This vision includes the role of social, behavioral, and economic science (SBES) research to identify users' perceptions, needs, and decisions so that effective probabilistic information is provided.

The concepts expressed herein are in the context of U.S. NWS operations. For the purposes of this paper, winter storms are defined as inclusive of the phenomenon warned for by NWS Winter Storm Warnings, including heavy snow, sleet, freezing rain, and blowing snow. Other key terms used in the work are summarized in Table 1.

Current state of winter storm forecasting

Numerous advances have occurred across the forecast value chain (WMO et al. 2015) over the last 20 years. For example, our physical understanding of mesoscale phenomenon embedded within winter storms (e.g., snowbands, lake effect snow, gravity waves, orographic precipitation processes) has improved. Field campaigns such as PLOWS (Profiling of Winter Storms; Rauber et al. 2014), OLYMPEX (Olympic Mountains Experiment; Houze et al. 2017), OWLeS (Ontario Winter Lake-effect Systems; Kristovich et al. 2017), and IMPACTS (Investigation of Microphysics and Precipitation for Atlantic Coast-Threatening Snowstorms; McMurdie et al. 2022) have provided new insight into thermodynamic and microphysical processes occurring within winter storms. This improved physical understanding is guiding the formulation of new model physics parameterizations, conceptual models, and forecaster tools.

Substantial improvements in operational observations have further advanced winter storm forecasting. Dual-polarization radar now provides the capability to detect the minute-by-minute evolution of the rain–snow line to the kilometer (e.g., Picca et al. 2014; Griffin et al. 2014). The fielding of new operational GOES satellite technology has enabled 1-min-interval rapid-scan imagery, lightning detection, improved vector winds,

Table 1. Definition of key terms used in this paper.

Term	Definition
Winter storm	An atmospheric disturbance causing hazards included in official NWS Winter Storm Warnings, including heavy snow, sleet, freezing rain, and blowing snow. For the purposes of this paper, other hazards such as coastal flooding and extreme cold are not included.
Risk	Possibility that an undesirable state (adverse effects or consequences) may occur as a result of a hazard event (Renn 2008; SRA 2022).
Hazard	A risk source where potential adverse consequences can cause harm (SRA 2022). For winter storms, the hazard risk source is a meteorological phenomenon (e.g., heavy snow, sleet, freezing rain, blowing snow).
Impact	The effects of adverse consequences on something of value (SRA 2022). Impacts of winter hazards may be, for instance, power outages, travel disruption, injuries, fatalities). Impacts of winter hazards can be affected by other societal factors (e.g., time of day, time of year).
Forecasting	General term inclusive of the prediction of meteorological variables, communication, and impact-based decision support.
Impact-based Decision Support Services (IDSS)	The provision of relevant information and interpretative services to enable core partners' decisions when weather, water, or climate has a direct impact on the protection of lives and livelihoods.
User	General term inclusive of all audiences of a service.
Decision-maker	Subset of users who have authority to make decisions that relate to public safety, protection of property, and economic well-being.
Core partner	Subset of decision-makers who are directly involved in the preparation, dissemination, and discussions involving weather-, water-, or climate-related NWS information that supports decision-making for routine or episodic high-impact events (NWS 2019).

and airmass products, among other advances (Goodman et al. 2012). The United States has also benefited from an increase in the number of mesonet observations (Mahmood et al. 2017). For example, the New York State mesonet, located in some of the snowiest regions of the United States, is helping reveal winter weather conditions on time and space scales not traditionally observed (Brotzge et al. 2020). Snowfall is notoriously difficult to measure (Hurwitz et al. 2020), but through a multiyear development effort, the NWS has established the first operational gridded snowfall analysis over the CONUS (NWS 2022a). The national gridded snowfall analysis is a combination of a model first-guess field and radar-derived quantitative precipitation estimates augmented by available in situ point observations. This analysis fosters an improved rigor of snowfall verification, which is essential for measuring forecast improvements and aiding model development.

Perhaps most dramatically, the ongoing revolution of numerical weather prediction (Bauer et al. 2015; Benjamin et al. 2019), improved data assimilation, and advent of ensembles (Bougeault et al. 2010; Palmer 2017) has fueled improvement of winter storm forecasts. For example, the average global model low track error of extratropical cyclones over the East Coast of the United States has decreased ~10% over the 2007–14 period (Korfe and Colle 2018, Fig. 9). Failures in prediction have inspired the establishment of operational ensemble systems. The "Surprise Snowstorm" of 2000 spurred advancement of the NCEP Short Range Ensemble Forecast (SREF) system (Tracton 2008). Winter storm forecasting is increasingly dependent on ensemble forecasts, including convection-allowing ensembles (e.g., Roberts et al. 2019; Schwartz et al. 2019). Displays of winter storm elements derived from model output, such as precipitation type, snowfall, and ice accumulation amount are now ubiquitous.

Despite these advances, forecasting winter storms remains challenging due to the sensitivity of snow and ice accumulations to storm track, precipitation type, precipitation rate, and snow-to-liquid ratios. Small model errors can have a disproportionate effect on the forecast and anticipated impacts. Figure 2 shows the soundings and associated news headlines and social media memes for two snowstorms. In these examples, just a 1°C cold error in the forecast temperature near 850 hPa resulted in an overforecast of snowfall in New York City in 2017, whereas just a 1°C warm error in the forecast temperature near 850 hPa resulted in an underforecast of snowfall in New York City in 2018. Further, the mesoscale nature of precipitation embedded within winter storms results in extreme gradients in winter precipitation type, intensity, and amounts that are difficult to predict even at short lead times (e.g., Kocin and Uccellini 2004a, 177–206; Market and Cissell 2002; Novak et al. 2006; Kenyon et al. 2020), and in some cases can elude model prediction altogether. In such cases, accurate nowcasts rely on forecaster experience and the monitoring of observational trends to anticipate and react to such features as the event is unfolding. The research community in general recognizes the need for better data assimilation, model physics, and ensemble system design (NOAA Science Advisory Board 2021; Magnusson et al. 2022), and specifically for winter storms, the need for better simulation of mesoscale extremes, better estimating mesoscale predictability, improving model microphysics, and improving prediction of precipitation type.

Another key challenge is that the predictability of winter storms varies from event to event and by scale. For example, the historic 22–24 January 2016 East Coast blizzard was anticipated days in advance, with forecasts of 1–2 ft (30–61 cm) of snow from Washington, D.C., to New York City verifying with stunning accuracy. In contrast, forecasts for the 17 November 2018 New York City snow storm predicted 1–2 in. (2.5–5 cm) as the snow began to fall, but ultimately over 6 in. (15.2 cm) was recorded (the heaviest November daily snowfall for New York City), immobilizing the city (Fig. 2). Such variation in predictability from event to event can influence public expectations and preparedness. A nor'easter successfully forecast days in advance can foster the expectation that *all* nor'easters can

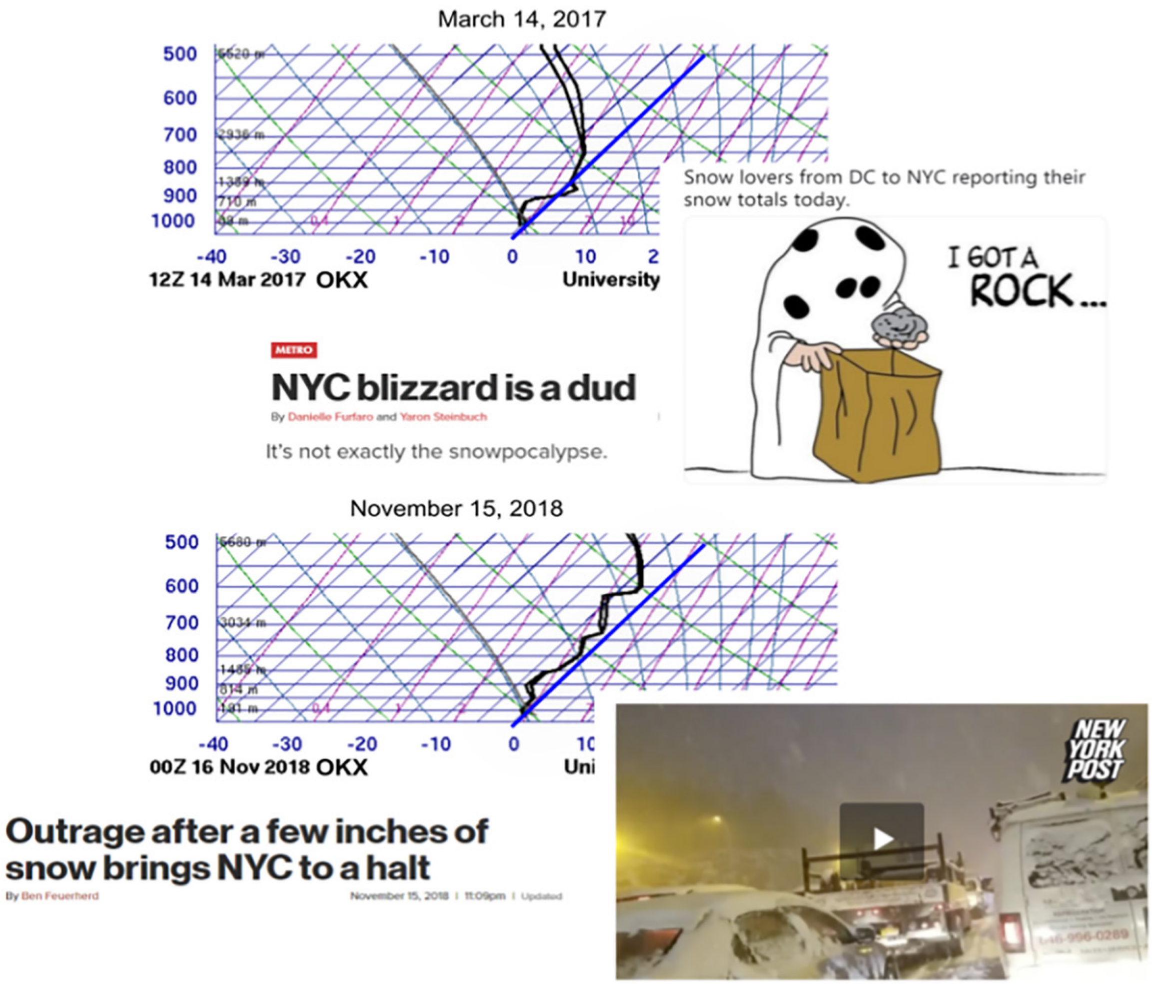

Fig. 2. Temperature soundings (below 500 hPa) at Brookhaven, NY (OKX), during (top) the overforecast 14 Mar 2017 snowstorm and (bottom) the underforecast 15 Nov 2018 snowstorm. Associated news media headlines (*New York Post*) and social media meme (courtesy Peter Mullinax) are shown for each event.

be forecast days in advance. Experimental research in the context of drought suggests that probabilistic forecast information attenuates the recency bias effect on participants' decision-making (Demnitz and Joslyn 2019), but it is unknown whether this extends to the real-world context for winter storms.

Finally, emerging research suggests that users weigh myriad factors—including different forecast elements—in assessing the risks posed by winter storms when making decisions (Barjenbruch et al. 2016; NWS 2018; Hosterman et al. 2019; Lazo et al. 2020; Morss et al. 2022). Figure 3 shows one representation of the various weather factors considered. Moreover, such factors can vary by users and by the risk scenario. Research by Morss et al. (2022) showed that, even when users have codified winter storm criteria for decision-making, they deviate from criteria as necessary based on other factors that influence situational risk. The intersection of users' varying complex risk management and decision-making contexts with inherent predictability limitations of winter weather is where probabilistic forecast and impact-based information has tremendous potential utility.

Evolving toward a probabilistic framework

Recommendations by Rothfusz et al. (2018), the National Research Council (NRC 2006), the American Meteorological Society (Hirschberg et al. 2011), the National Institute of Standards and Technology (NIST 2013), and the Priorities for Weather Research Report (NOAA Science Advisory Board 2021) encourage the expanded evaluation and use of probabilistic information to convey weather forecast uncertainty. Forecasting a Continuum of Environmental Threats (FACETs) is a proposed framework to use probabilistic hazard information (PHI) for high-impact weather and water events (Rothfusz et al. 2018). The NWS recently commissioned researchers to review over 300 papers pertaining to forecast communication. A key finding is that, assuming appropriate presentation, probability information generally improves decision quality (Ripberger et al. 2022). Demnitz and Joslyn (2019) found that probabilistic predictions inspired greater trust and allowed participants to make better economic decisions (measured as expected value) overall than did deterministic predictions. In a survey of nearly 500 businesses regarding tornado warnings, Howard et al. (2021) found that a probabilistic information system would produce an annual cost avoidance of \$2.3–\$7.6 billion compared to the current deterministic warning paradigm—largely due to reductions of false alarms. There has been less research about application of a probabilistic framework specifically for winter weather as compared to other hazardous weather (e.g., tornadoes, hurricanes), and accordingly, the Priorities for Weather Research report calls for the prioritization of research on development, communication, and use of uncertainty information for many hazards, including winter storms.

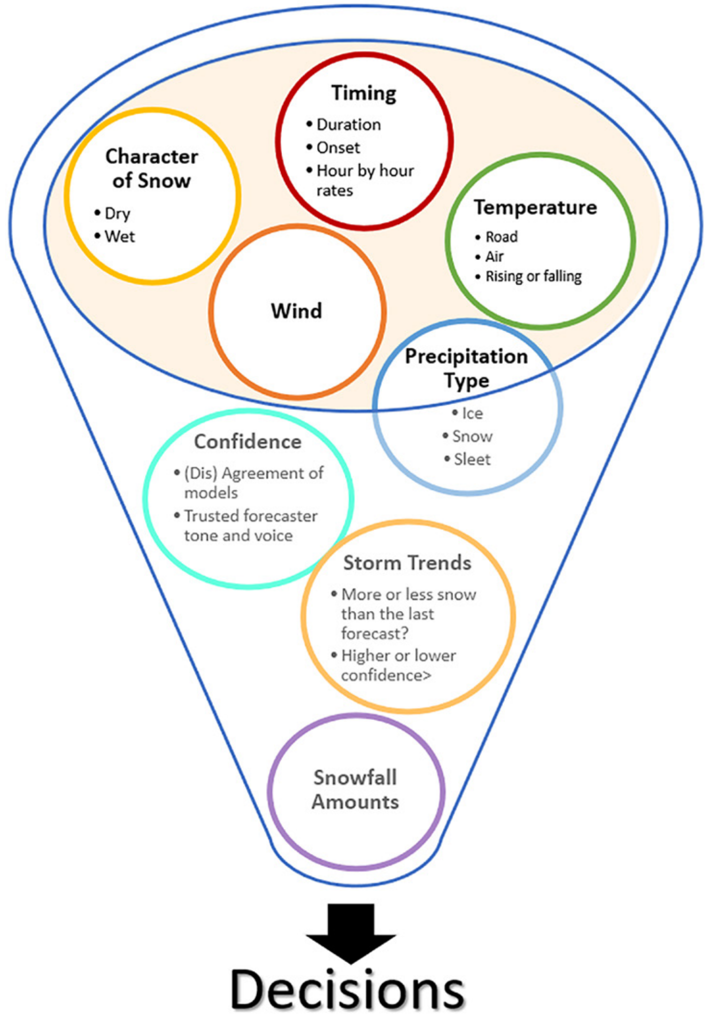

Fig. 3. **A representation of the various weather factors used in winter storm decision-making (adapted from NWS 2018).**

Probabilistic framework.

Idealized description. In recognition that winter storms are inherently uncertain, the NWS Winter Weather Program has embraced the expansion of PHI to aid user decisions. Fundamentally, the Winter Weather Program is working toward a consistent probabilistic framework. Consider Fig. 4, which is adapted from Rothfusz et al. (2018) for winter weather hazards. The figure depicts an idealized scenario where certainty increases as the event comes closer in time. As the certainty increases, the urgency and specificity of the forecast increases. Importantly, the foundational PHI is frequently updated. In this idealized framework, probabilistic thresholds can be used to consistently trigger and inform products and messaging. For example, NWS Policy cites a 50% probability of an event as the trigger for a Winter Storm Watch and 80% probability of an event as the trigger for a Winter Storm Warning (NWS 2022b). For storms with greater predictability, these thresholds may be met with long lead time (days), while in storms with less predictability these thresholds

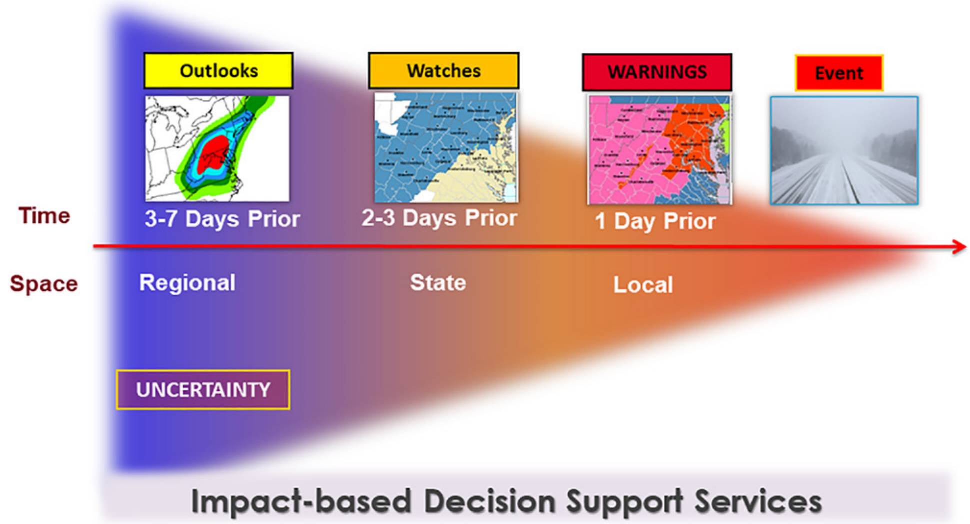

Fig. 4. **Idealized depiction of the winter storm probabilistic framework based on the 23–26 Jan 2016 snowstorm. Uncertainty is represented by the colored envelope, which narrows as the event approaches. Frequently updated probabilistic hazard information informs the issuance of Winter Storm Outlooks, Watches, and Warnings as uncertainty decreases (or alternatively, as certainty increases). Impact-based decision support services occur in advance and during the event. [Adapted from H. Lazrus, NCAR, and from Fig. SB1 in Rothfusz et al. (2018).]**

may be met just hours prior to the onset of the storm. Consistently anchoring decisions at a probabilistic threshold (e.g., warnings at 80% probability) helps guard against over- and underwarning. As long as the probabilities are calibrated, this threshold approach can build partner trust and confidence as the partner knows what the expected hit and miss ratio will be for a given product (for a Winter Storm Warning, ideally 80% POD and 20% FAR). This is also an example of how yes–no deterministic products (a warning) can be supported by the underlying frequently updating PHI.

IDSS underpins this framework, with meteorologists assisting decision-makers in the interpretation of the forecast, including frequent updates as the PHI and associated situation evolve. The probabilistic framework offers multiple ways of supporting IDSS for core partners. A partner could have a probabilistic threshold of occurrence as the primary factor that influences their decision-making. Or, because probability and lead time tend to be inversely correlated, a partner might have a lead time when they need to make a decision, in which case the probability at that lead time is provided to support decisions. The probabilistic framework also includes forecasts conveyed through ranges, scenarios, and categories based on underlying probabilities (e.g., Novak et al. 2014; Demuth et al. 2020), such as the lowest and highest amounts of snowfall expected or the earliest and latest times of snowfall onset. Finally, the probabilistic framework could improve partners' awareness of and preparatory responses for low-probability, high-impact threats.

Practical realities of the framework. In practice, strict adherence to probabilistic thresholds for Winter Storm Watches, Warnings, and messaging remains challenging. An essential assumption of the framework is that the PHI is calibrated (i.e., when an 80% chance of an event is predicted, it occurs 80% of the time). Modern-day ensemble systems are typically underdispersive, with observed events occurring outside the envelope of solutions more than should be expected (e.g., Buizza et al. 2005; Buizza 2018; Romine et al. 2014; Zheng et al. 2019). Underdispersive ensemble forecasts can lead users to underestimate the probability of extreme events. In essence these are "surprise" storms, when heavy snow occurs and there was no ensemble prediction of the event. Underdispersive ensembles also erode trust in ensembles among forecasters (Novak et al. 2008). However, improvements in model physics, ensemble design, and resolution have improved the calibration of operational ensemble output over the past decade (Swinbank et al. 2016), and forecasters are gradually becoming more comfortable using ensemble information (Demuth et al. 2020; Tripp et al. 2023). Importantly, statistical postprocessing can further improve the calibration of ensemble precipitation forecasts (e.g., Voisin et al. 2010; Hamill et al. 2017; Buizza 2018; Scheuerer and Hamill 2019) and is an area ripe for machine learning (ML) approaches (McGovern et al. 2019, 2022; Handler et al. 2020). The NWS has embraced statistical postprocessing through the National Blend of Models (Craven et al. 2020; Hamill et al. 2017), including numerous statistically postprocessed probabilistic winter weather variables.

Another assumption and associated challenge is that uncertainty monotonically decreases (certainty increases) as the event draws near. However, uncertainty can actually *increase* for some winter storm events at shorter lead times. For example, nonlinear shifts in the possibility of a strong nor'easter can suddenly increase the threat of mesoscale snowbands (that were less likely when a strong nor'easter was unlikely). In this situation, the ensemble spread of snowfall amounts might increase due to the greater possibility of mesoscale snowfall bands either hitting or missing a location. Nevertheless, probabilistic guidance is useful for characterizing this possibility and associated messaging. Research into the causes of increased uncertainty at short lead times and how decision-makers interpret and respond to increased uncertainty at short lead times is encouraged.

Another key assumption is that there even *is* a probabilistic threshold that aligns with user decisions and that users solely adhere to thresholds (Morss et al. 2022). Users have a wide variety of needs and risk tolerances, and there likely is no single threshold sufficient to address them all (Morss et al. 2010; Senkbeil et al. 2013). For example, a warning may be appropriate for a high probability of a minor snowfall event near rush hour, even if probabilities of warning criteria do not meet an established threshold (Demuth et al. 2020). Being able to skillfully derive and provide information to meet this array of needs and decision contexts is an important part of the subjective aspect of IDSS, and is a uniquely human role (Stuart et al. 2022). Nevertheless, a probabilistic framework can be used to root decisions in skillful, objective information. Ultimately, more sophisticated multiparameter PHI aligned with impact thresholds to derive the explicit probability of impacts is desired.

One of the most challenging cases is the occurrence of a low-probability, high-impact event with short lead time. Consider the occurrence of an ice storm. The framework assumes that as the event draws close in time, the objective probabilities will eventually increase to a critical user threshold. For example, 36 h in advance there may be a 10% chance of damaging ice, but 12 h in advance the probabilities will increase to 50%. The framework assumes users will tolerate shorter lead time for sufficient certainty in these cases. Regardless of if a specific trigger is met, the framework supports frequent updates of probabilistic information that may be especially critical in these rapidly changing situations for forecasters and users alike. The use of probabilistic tools to alert users to low-probability, high-impact events may be necessary in these cases, and work to improve short-range predictions is encouraged.

In recognition of the need to root decisions in objective probabilistic information while having flexibility to address ensemble shortcomings and complex user decisions, the Winter Weather Program is framing probabilistic thresholds and criteria as a "first guess," with flexibility to use other tools to address subjective factors and anticipated impacts.

Real-world example. An example of the framework applied to the 10–11 December 2021 snowstorm is shown in Fig. 5. The experimental Winter Storm Outlook is used to illustrate one example of winter PHI. The WSO uses a multimodel ensemble to quantify the probability

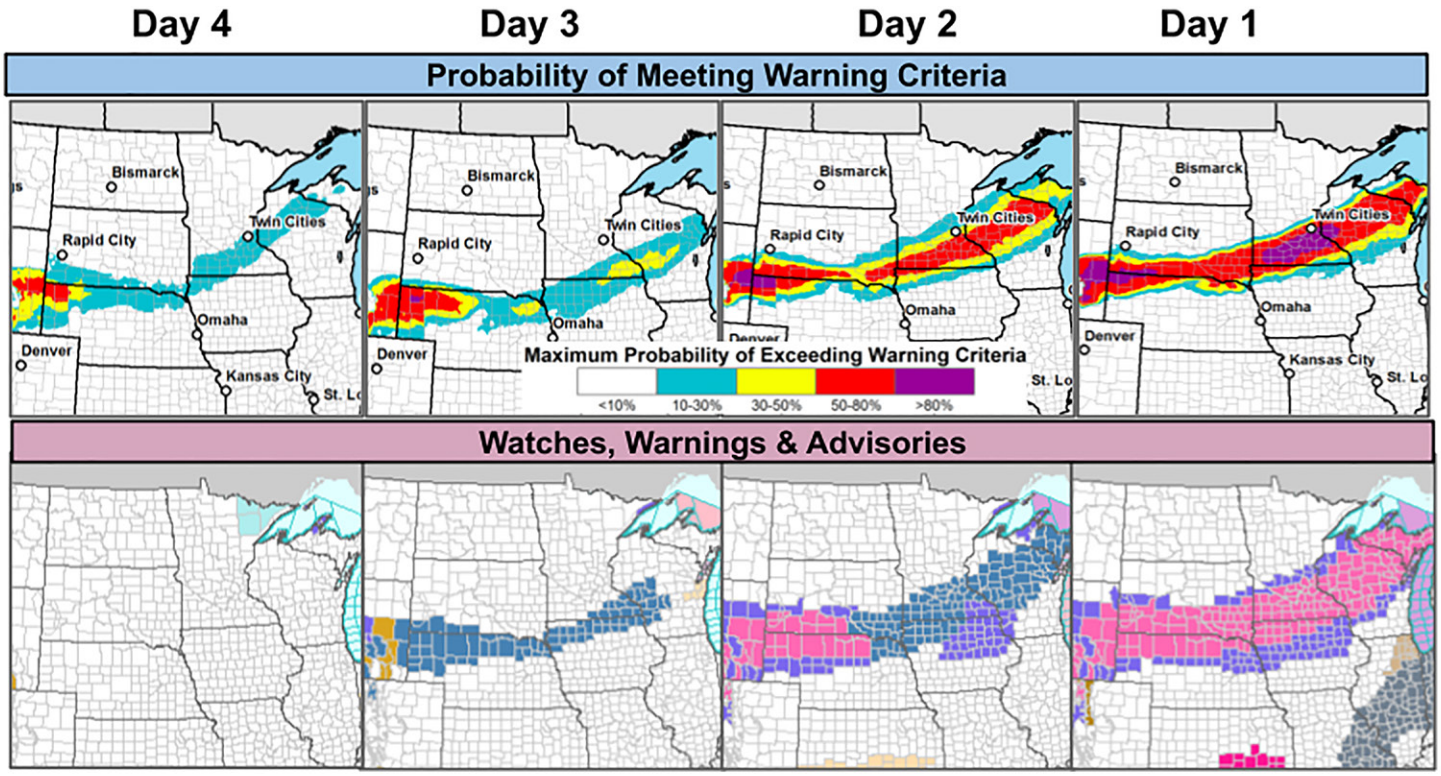

Fig. 5. Example application of a probabilistic framework to the 10–11 Dec 2021 snowstorm. (top) Probability of exceeding snowfall accumulation warning criteria from 4 days in advance (1200 UTC 7 Dec) to 1 day in advance (1200 UTC 10 Dec) of the event. (bottom) NWS Winter Storm Watch (blue), Warning (pink), and Advisory (purple) at corresponding times.

of exceeding local Winter Storm Warning criteria (NWS 2022c). Snowfall warning criteria across the north-central United States is 6 in. $(12\ h)^{-1}$ [15.2 cm $(12\ h)^{-1}$]. Thus, in this case the WSO is showing the probability of exceeding 6 in. $(12\ h)^{-1}$ [15.2 cm $(12\ h)^{-1}$].

Four days in advance of the event the WSO exhibits a stripe of probabilities exceeding 10% from eastern Wyoming toward the Twin Cities of Minneapolis and Saint Paul, Minnesota, and into northern Wisconsin. Three days in advance of the event, probabilities increased, with over a 50% chance of exceeding local snowfall criteria in southeast Wyoming and the Nebraska panhandle, and embedded areas exceeding 30% in northeast Nebraska and southeast Minnesota. Consistent with this increase, Winter Storm Watches were issued from eastern Wyoming into Wisconsin. Two days in advance of the event, the probabilities increased to over 50% in many areas and shifted north. Warnings were issued in parts of southeast Wyoming, southwest South Dakota, and the Nebraska panhandle where probabilities approached 80%. Farther east, Winter Storm Watches were expanded northward, consistent with the northward shift in probabilities. The Twin Cities now had probabilities exceeding 30%. One day in advance, probabilities further increased to over 80% in southern Minnesota, including the southern Twin Cities metro area. Winter Storm Warnings now stretched continuously from eastern Wyoming into central Wisconsin. Warning criteria snowfall was observed in a majority of the warning area. For example, the Minneapolis–Saint Paul International Airport (MSP) observed 11.7 in. (29.7 cm). In response to the warnings, more than 190 flights were proactively canceled at MSP, and there were numerous school closures and early dismissals across the region.

Probabilistic tools. In recognition that users have a wide variety of needs and risk tolerances, tools to sample and extract information from the full probability distribution are necessary, including ranges, scenarios, and feature-based displays. An experimental tool to succinctly communicate the range of possibilities calculates and displays the 10th- and 90th-percentile values of the snowfall distribution (Novak et al. 2014; Waldstreicher and Radell 2018; Demuth et al. 2020). This helps provide quantitative information on low-probability, high-impact events. Guided by research interviews with core partners, the graphical depiction provides the official "expected" snowfall forecast, the 10th percentile as the "low end" amount and the 90th percentile as the "high end" amount. In practice, the 90th-percentile value has become a proxy for communicating a reasonable worst-case scenario to decision-makers (NWS 2018). By definition, the 90th-percentile value has a 10% chance of being exceeded. This approach has similarities with other hazards, such as storm surge, where the 10% chance of exceeding ~1 m of inundation is used in Storm Surge Watch and Warning decisions.

An example of this information for the 10–11 December 2021 snowstorm across the Twin Cities, Minnesota, metro area is shown in Fig. 6. In this example, the expected forecast for MSP was 10.4 in. (26.4 cm). However, given uncertainty in the location of an expected snowband, the forecast low-end amount was 6 in. (15.2 cm) and the high-end amount was 13.7 in. (34.8 cm). The observed snowfall was 11.7 in. (29.7 cm). A mesoscale snowband indeed formed in this event, creating a sharp snowfall gradient of 3–20 in. (7.6–50.8 cm) across the metro area. The 10th- and 90th-percentile forecasts largely captured this range (Fig. 6).

Active work continues on refining the derivation, communication, and verification of percentile information. For example, verification of the snowfall percentiles across the CONUS during the 2021/22 season showed that the forecast 90th percentile was within 5% to the observed 90th percentile; however, the forecast 10th percentile was not well calibrated to the observed 10th percentile at higher snowfall values. Regular verification helps drive further development and adjustments to improve statistical reliability.

Another emerging tool to sample and extract information from ensembles is cluster analysis. Cluster analysis quickly and objectively groups ensemble members with similar forecasts together, thereby reducing a large set of ensemble forecasts down into the most prevalent forecast scenarios (Brill et al. 2015; Zheng et al. 2017, 2019). Forecasters can then quickly view these ensemble clusters to better understand and communicate forecast uncertainty and the range of possible forecast outcomes (Lamberson et al. 2023).

Tools extracting information from ensembles on mesoscale features are also being developed. Building on the recommendations from Novak et al. (2012), the Weather Prediction Center hosts a tool that displays the evolution of snowbands from the High Resolution Ensemble Forecast (HREF; Roberts et al. 2020). This approach uses the Method for Object-Based Diagnostic Evaluation Time-Domain (MODE-TD) which is part of the Model

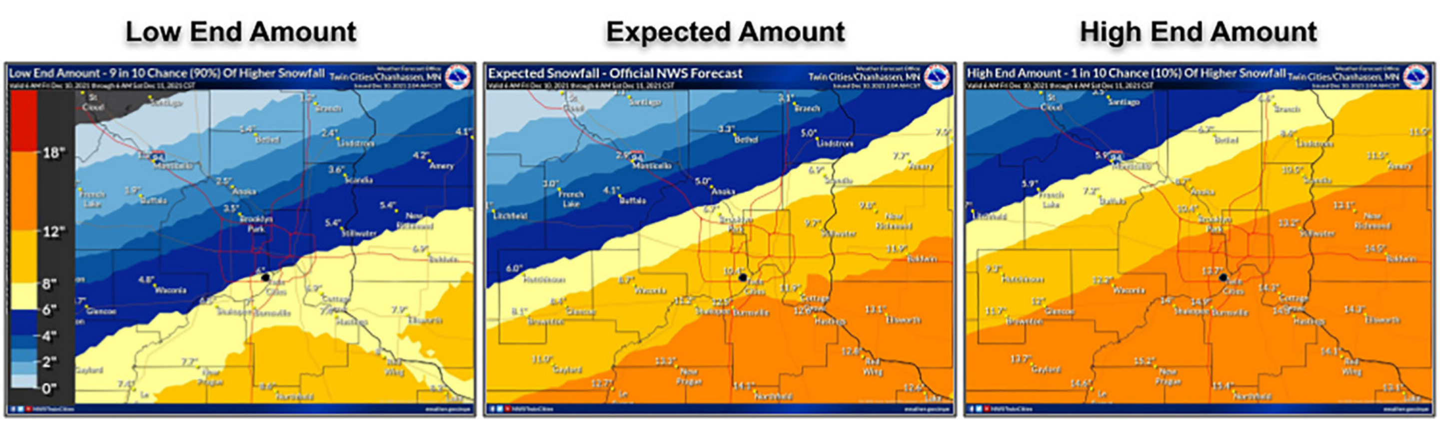

Fig. 6. Forecasted snowfall range (in.) over the Twin Cities, Minnesota, metro area issued 0700 UTC 10 Dec 2021 showing the forecast (a) low-end amount, (b) expected amount, and (c) high-end amount. The MSP station location is shown by black dot.

Evaluation Tools (MET; Brown et al. 2021), to identify and track ensemble member QPF objects. Only objects that have 1-h accumulated precipitation exceeding 0.10 in. (2.5 mm) and that are snow (identified by the ensemble member categorical snow precipitation type grid) are displayed. The outline (i.e., object area) and intensity of the object are displayed. The intensity is calculated as the 90th percentile of hourly QPF within each object (colored outline). Tools such as these allow forecasters to quickly synthesize the forecast probability of occurrence, timing, and location of mesoscale snowbands. Application of this tool during the historic 17 December 2020 snowstorm shows that the occurrence and timing of an intense mesoscale snowband was predicted by the HREF 20 h in advance [most members with at least 0.15 in. h^{-1} (3.8 mm h^{-1}) liquid equivalent; Fig. 7a]; however, the observed location was on the northwest edge of the ensemble envelope (Figs. 7a,b).

Envisioning the future: Probabilistic impact information

Foundational PHI and associated tools are necessary, but not sufficient, for the most effective IDSS. Rather, the next step is aligning PHI with user impact thresholds, ultimately providing decision-makers quantitative information on the probability of impacts relevant to their decisions.

An example of this concept comes from the aviation sector, which has identified specific decision thresholds that affect safety and business efficiency. The NWS/NCEP's Aviation Weather Center (AWC) developed the Aviation Winter Weather Dashboard (AWWD) (Steiner et al. 2015). The AWWD was designed to provide automated guidance on winter weather impacts at major U.S. airports. Specific winter weather hazard thresholds were developed for each of the 29 major airports in collaboration with the Federal Aviation Administration (FAA) and industry partners. The dashboard displays the potential impact on each airport based on the chance of snowfall, freezing rain, or visibility thresholds being exceeded, as predicted by the NCEP SREF system. The dashboard uses a multicolored matrix to display the likelihood of the predicted impact, including nominal (green), slight (yellow), moderate (orange), and high (red) impact. This translation from a probabilistic forecast to a categorical impact takes into account relevant operational thresholds and other considerations for terminal operations. In general, threat levels are triggered at the 40% probability of exceedance of the critical threshold for that threat level. Figure 8 shows the AWWD outlook issued Friday, 20 February 2015, for the Denver area with moderate operational

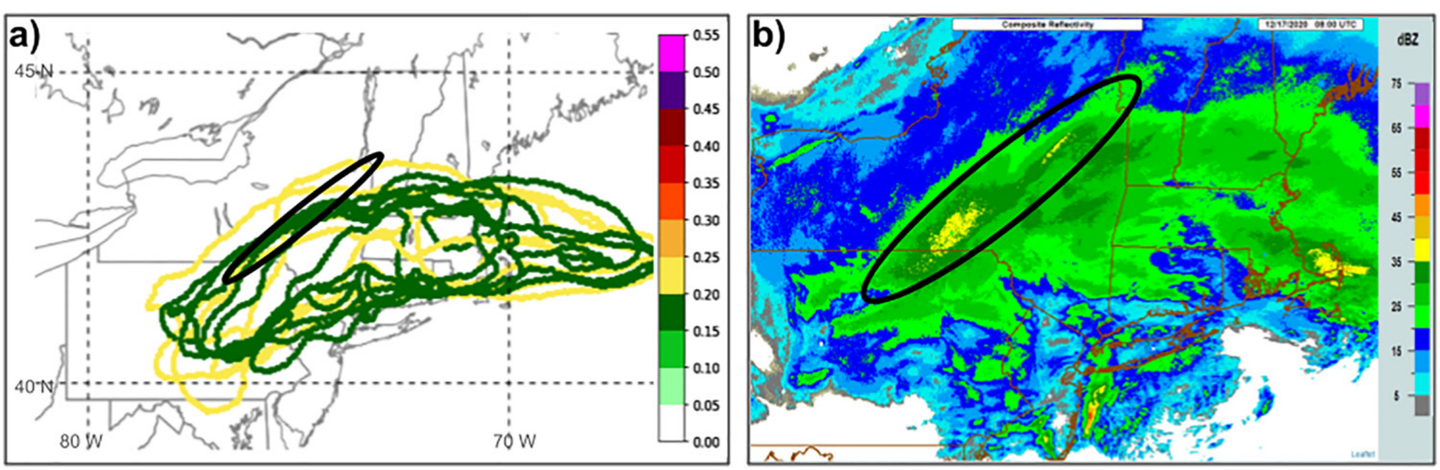

Fig. 7. **(a) 20-h forecast of snowband objects from the NCEP HREF valid 0800 UTC 17 Dec 2020. Forecast objects are outlined in color corresponding to the 90th percentile of intensity according to color bar (liquid equivalent, in. h^{-1}). An ellipse representing the approximate observed band position is outlined in black. (b) Observed radar composite at 0800 UTC 17 Dec 2020, showing observed snowband (black ellipse) verifying along the northwest side of the envelope of solutions.**

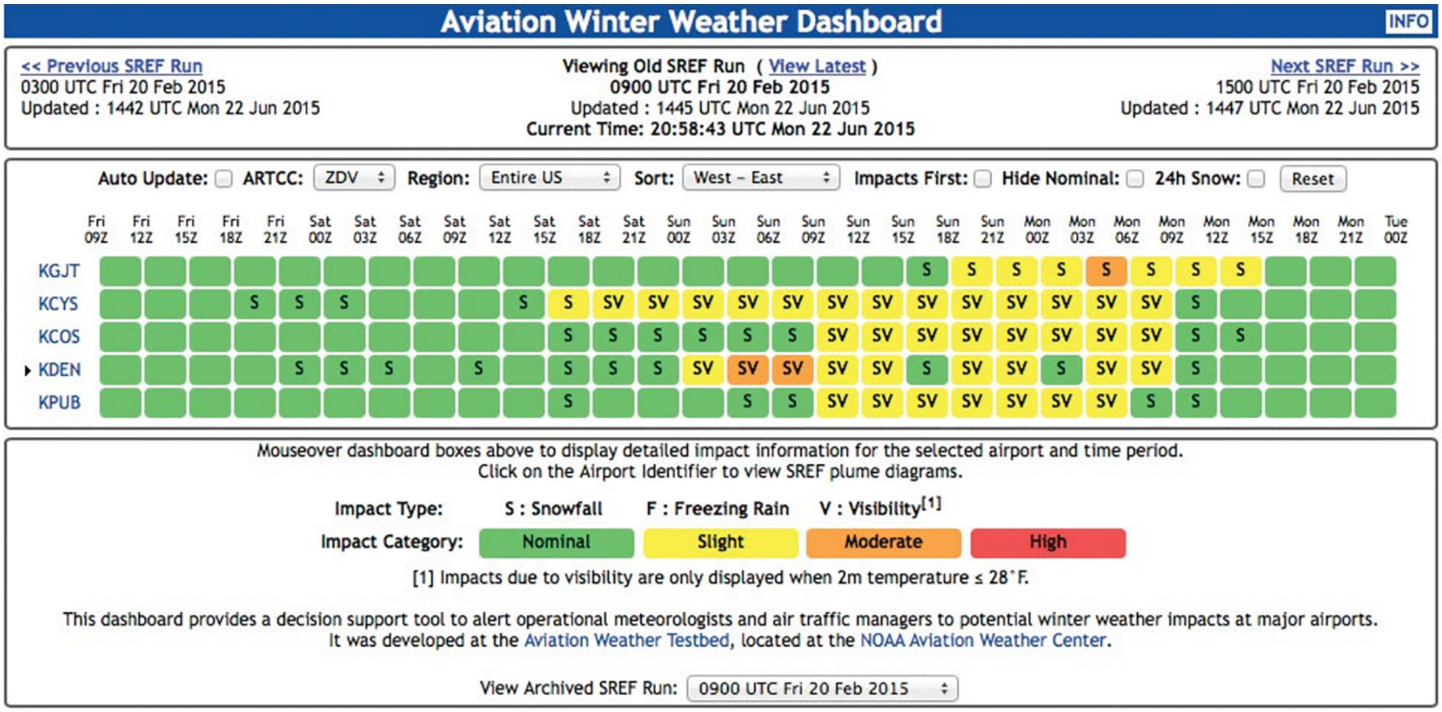

Fig. 8. Example of Aviation Winter Weather Dashboard for forecasts for Colorado airports valid 20 Feb 2015 (from Steiner et al. 2015).

impacts from a greater than 40% chance of snowfall rates exceeding 0.5 in. h^{-1} (1.3 cm h^{-1}) ("S" in the table) and a greater than 40% chance of visibility < 1 n mi (<1.85 km) ("V" in the table) expected on Sunday. This dashboard provides a decision support tool to alert operational meteorologists and air traffic managers to potential winter weather impacts at major airports. Beyond the aviation sector, one can imagine similar impact assessments and dashboards for other sectors of the economy.

Although sophisticated and quantitative winter hazard impact models are likely to be developed through time, it remains a nascent field (Merz et al. 2020). To communicate winter impacts more generally across a range of decision contexts, various scales have been developed to qualitatively estimate the severity of winter storms. For example, the Northeast Snowfall Impact Scale (NESIS) (Kocin and Uccellini 2004b) is presented as a poststorm measure of the impact of heavy snowfall in the Northeast urban corridor and is based upon the geographical area of snowfall—weighted by population. The Regional Snowfall Index (RSI; Squires et al. 2014) is an evolution of the NESIS which applies the technique to other portions of the country. The NESIS and RSI do not take into account impacts from ice, wind, and blowing snow, and importantly, these indices are calculated *after* a storm. The Local Winter Storm Scale (LWSS; Cerruti and Decker 2011) is created using observed conditions of winds, snow, ice, and visibility and applying a score and weighting function for each element depending upon the specific value. The result is a numerical score from 0 to 6 depicting the "severity" of the storm over the user defined time frame. The LWSS can further be correlated to a Rooney Disruption Index (RDI) (Rooney 1967) relating societal impacts by using historical weather observations and storm impact accounts from newspapers.

Building on the foundation of the above indices, the NWS has recently developed the Winter Storm Severity Index (WSSI; NWS 2021). The WSSI is a predictive tool designed to enhance communication of the expected winter storm severity (potential societal impacts) to users. The WSSI is comprised of subcomponent algorithms that use official NWS National Digital Forecast Database forecast values, climatology, and nonmeteorological data (land cover, tree type, population, etc.) to determine the level of potential societal impact based upon specific characteristics of winter storms. The subcomponents are

- Snow Amount Index: potential of impacts due to the total amount of snow or accumulation rate.
- Snow Load Index: potential of impacts due to the weight of the snow.
- Ice Accumulation: potential of impacts due to combined effects of ice and wind.
- Blowing Snow Index: potential of impacts due to blowing snow.
- Flash Freeze Index: potential of impacts due to quick-forming ice from rapid temperature drops during or after precipitation.

Each of the components produces a value that equates to the potential impact. The final WSSI value is the maximum value from all the subcomponents. The five levels are labeled Winter Weather Area, Minor, Moderate, Major, and Extreme, which are mapped to provide a graphical depiction of anticipated overall impacts to society due to winter weather. A retrospective example of the summary forecast output from the WSSI is shown in Fig. 9a for the historic High Plains Blizzard of 12–14 April 2022. This storm immobilized the region, with drifts to 15 ft (4.6 m)—stopping snowplows in their tracks Fig. 9b (insert).

The next frontier of such impact scales is to highlight the *probability* of categories (and associated specific impacts)—ultimately contributing to the vision of providing community decision-makers quantitative information on the likelihood of impacts from winter storms. For example, one can calculate a WSSI value from each member of a meteorological ensemble to derive the probability of an impact category. Figure 9b shows the result of this process for the 12–14 April 2022 blizzard. This probabilistic WSSI is formally experimental during the 2022/23 winter season. One can envision incorporating more sophisticated factors, such as the time of day and the day of the week.

Importantly, a rigorous and consistent database of observed impacts is needed to verify impact forecasts. Current impact datasets, such as the NOAA Storm Events Database, are a general starting point, but consistent, official, and timely information on power outages, road closures, traffic accidents, airport delays, hospital admissions, and insurance claims, among other information is needed within a common platform. This more detailed and complete impact data are essential to calibrating impact predictions and building decision-makers' trust in the predictions.

Integrating social science for improved winter storm information

The evolution toward a probabilistic framework for winter storm forecasting is envisioned to help users better understand winter weather risks and improve their decision-making. Although some empirical research has been conducted pertaining to different users' access, interpretation, and use of winter weather forecast information (Rice and Spence 2016; NWS 2018; Burgeno and Joslyn 2020; Call and Flynt 2021; Demuth et al. 2020; Su et al. 2021; Morss et al. 2022; Tripp et al. 2023), additional research is needed in several critical areas to guide development and operationalization of probabilistic winter storm forecast information that is user-oriented and actionable. Several topics are outlined below, all of which are supported by recent reports that provide guidance to NOAA and the broader Weather Enterprise (NASEM 2018; NOAA Science Advisory Board 2021).

There are a variety of users of probabilistic winter forecast information. NWS forecasters themselves are critical, first-order users of probabilistic guidance. Recent research with forecasters, including interviews (Demuth et al. 2020) and surveys (Tripp et al. 2023), have found that forecasters want new tools, including probabilistic information, and training that help them understand challenging and high-impact forecast situations, but that these tools

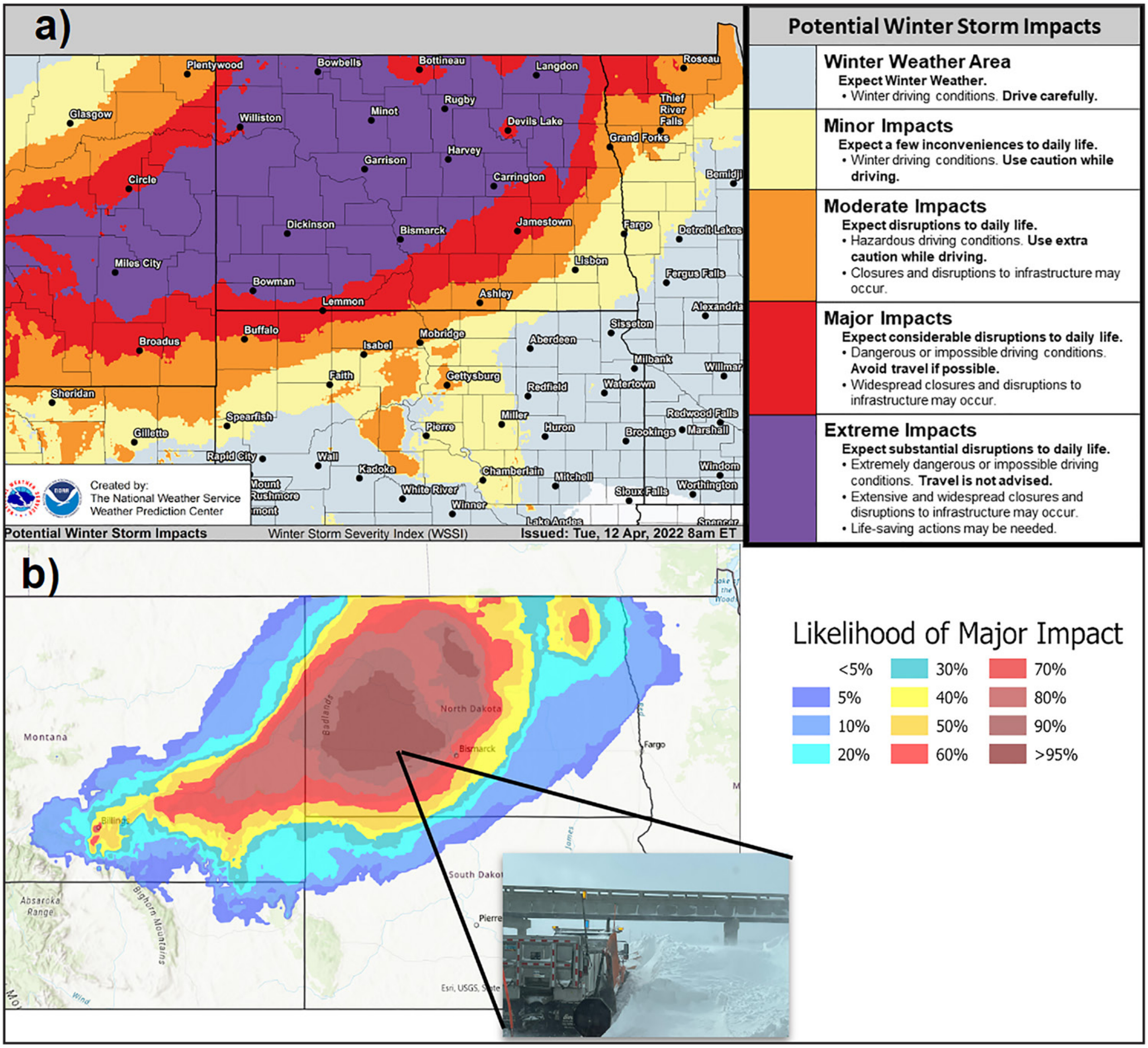

Fig. 9. **Example of the summary forecast output from the Winter Storm Severity Index for the historic High Plains Cyclone of 12–14 Apr 2022 showing (a) the most likely impact category and (b) the probability of exceeding major impacts. (insert) Picture of a stranded plow on Interstate 94.**

need to be guided by their forecast and IDSS needs. More research is needed to understand how forecasters interrogate and interpret probabilistic guidance, integrate it into their assessment of risks of different winter hazards, and translate it into forecast and IDSS for partners. As discussed above, there are predictability challenges (e.g., of heavy snowfall rates, uncertainty *increasing* as an event nears, flatter versus sharper probability distributions). How forecasters assess and use probabilistic guidance in these different situations must also be considered. Among the most important users, NWS's core partners make mitigating and protective decisions when a winter storm threatens. It is critical to understand how far in advance these users must make decisions, and how different winter storm forecast parameters (e.g., timing, amount, intensity) influence decisions. It is further critical to explore these users' interpretations of, and preferences for, probabilistic information conveyed in different ways (e.g., probability of exceedance, ranges, scenarios) and how these different forms of information intersect with their decision-making. Last, but not least, members of the public are a broad, diverse user group that the NWS serves, and thus it is also essential to build on the research referenced above to further develop robust knowledge about how they interpret and use probabilistic information conveyed in different ways and for different winter storm risk scenarios.

Importantly, winter storms and the risks posed by them evolve in space and time, and associated forecast and impact probabilities evolve as well. The December 2021 real-world case described above provides an excellent example of the evolving risk as the probabilities changed through time (Fig. 5). Yet, most SBES research with users conducted during an event is collected at only one point in time, which cannot capture whether, when, and how people's interpretations and responses are changing with changing probabilistic forecast information. It therefore is important to investigate risk communication and decision-making as dynamic and to collect such social science observational data accordingly. Furthermore, it is important to collect data from users across events, to understand how different winter storm scenarios and associated differences in predictability influence users' interpretations and decision-making. For instance, how do users interpret and use the aforementioned low-end and high-end amounts based on the 10th and 90th percentiles from an ensemble when that distribution is sharper (narrower range) versus flatter (wider range)? As the December 2021 case shows, there are inherent spatial variations in probabilistic forecast and impact-based information (e.g., Fig. 6). Thus, it is important to study these spatial representations of probability and risk from users' perspectives.

To better motivate action, consistent messaging has been a recent focus (Weyrich et al. 2019; Burgeno and Joslyn 2020; Williams and Eosco 2021). Inspired by the success of the National Hurricane Center's use of key messages for tropical cyclones, the Winter Weather Program implemented key messages in 2022 to galvanize core partners and media around a synthesized and consistent message. The key messages often highlight probabilistic information (Fig. 10). Measuring the influence of these consistent, key messages on user decisions is necessary.

Summary and next steps

Innovations in winter storm forecasting have occurred across the value chain over the past two decades, from physical understanding, to observations, to model forecasts, to postprocessing, to forecaster knowledge and interpretation, to products and services, and ultimately to decision support. These innovations enable more accurate and consistent forecasts, which are increasingly being translated into actionable information for decision-makers. The NWS Winter Weather Program has embraced a probabilistic framework, with PHI serving as a foundation for decision support, and has a vision of providing community decision-makers skillful and quantitative information on the likelihood of impacts from winter storms.

However, to achieve this vision, there are several gaps that must be addressed. We recommend the following priority actions:

- Fundamentally improve model precipitation accuracy (occurrence, type, and amount) (NOAA Science Advisory Board 2021). This includes improved prediction of winter storm synoptic-scale patterns, mesoscale features, thermal profiles, and precipitation type.
- Improve ensemble underdispersion for winter storms. Minimize the number of events verifying outside the envelope of ensemble predictions to further build trust among forecasters and decision-makers. Beyond improved physical understanding and improved model systems, this action likely includes the use of ML and other postprocessing techniques to calibrate ensemble output.

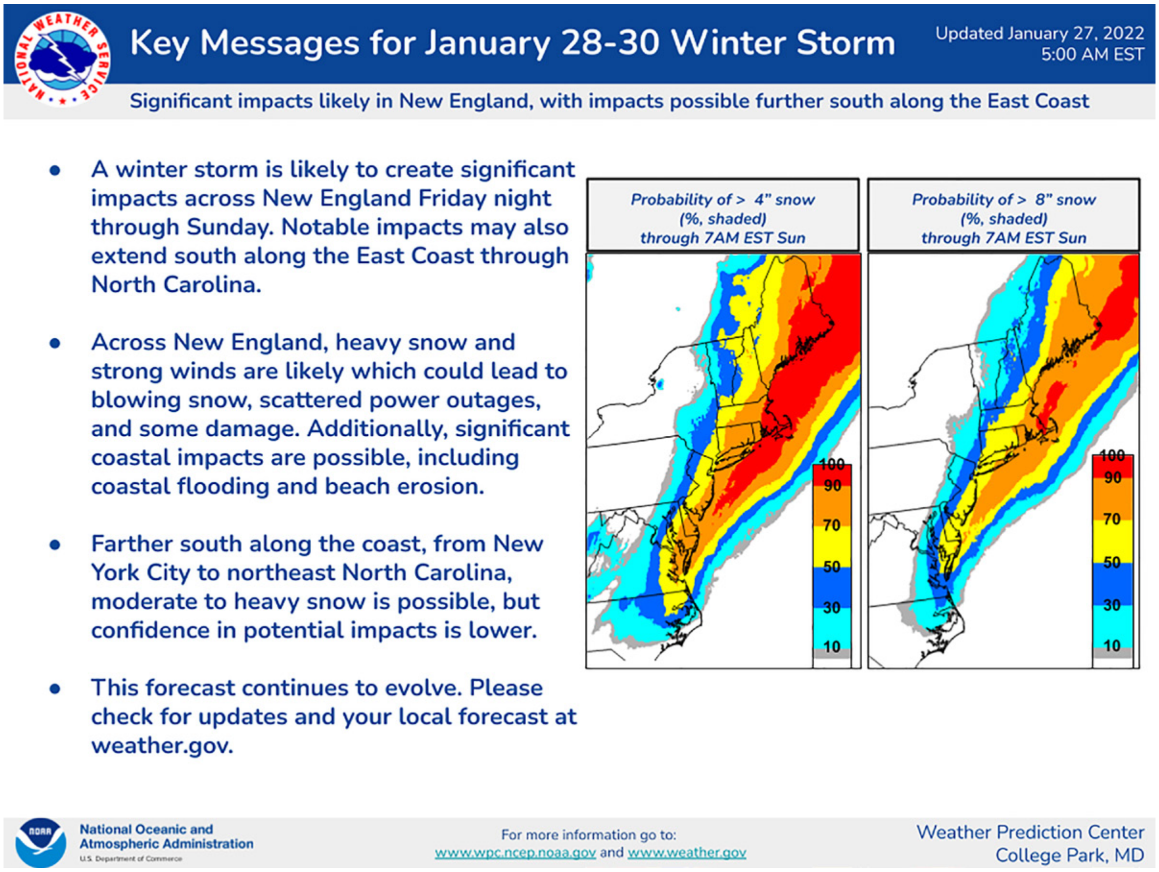

Fig. 10. Example key messages 2 days prior to the 28–30 Jan 2022 winter storm. The key messages highlight the increasing potential for a storm, with the greatest hazards along the coast. The probability of more than 4 in. (10.2 cm) and 8 in. (20.3 cm) of snow is shown.

- Focus on improving short-lead-time forecast failures, diagnosing root causes, and ultimately reducing the occurrence of low-probability, high-impact events with short lead time.
- Redouble efforts on characterizing the mesoscale aspects of winter storm threats including probabilistic information on timing, snow and ice accumulation rates, and precipitation type transitions—especially at short lead times.
- Extend probabilistic services further out in time to enhance situational awareness and advance lead time for winter storms.
- Establish a rigorous and consistent database of observed impacts to support impact verification and the development of impact-based tools. Establish a database of critical impact thresholds for transportation, energy, and other sectors.
- Develop probabilistic *impact* information that links multiparameter PHI (e.g., probability of intense snowfall rates at rush hour) with decision-maker thresholds.
- Expand support for SBES research with a range of users (forecasters, core partners, members of the public) to guide the evolution of probabilistic products and decision support, including improving understanding of users' risk assessment and decision-making contexts when winter storms threaten.

As the needs of decision-makers for accurate, consistent, timely, and actionable winter weather information increase, it is critical for the public, private, and academic sectors to work in partnership to make advances across the value chain. Burgeoning examples of this collaboration in the winter weather realm include the Pathfinder initiative, which facilitates collaboration between the NWS, the Federal Highway Administration (FHWA), state transportation departments, and private-sector companies to assess the impact of weather on transportation, and to develop consistent, concise messaging for motorists, including the use of variable-message signs along major roadways (FHWA 2018). The establishment of Earth Prediction Innovation Center (Jacobs 2021; Uccellini et al. 2022), the Unified Forecast System initiative (Jacobs 2021), and the Precipitation Prediction Grand Challenge (NOAA 2020) are promising collaborative R2O initiatives to improve winter storm accuracy and increase lead time for decision-makers. The Hydrometeorological Testbed (HMT) is a key forum fostering partnerships as it convenes forecasters, model developers, and academics to rigorously evaluate experimental model applications and new services through the lens of winter storm hazards and IDSS on a sustained basis (e.g., Ralph et al. 2005). Further, increasing integration of SBES into products and services will ensure understandable and actionable information. Recent workshops involving meteorologists and representatives in emergency management, media, and transportation have focused on effective forecast communication of probabilistic information for high-impact weather (Colle et al. 2021), and such engagements are encouraged. Through these and other collaborative efforts, the vision of providing community decision-makers quantitative information on the likelihood of impacts from winter storms can become reality.

Acknowledgments. The authors are grateful to Dr. Louis Uccellini for continued encouragement of both programmatic advances and the writing of this paper. This includes the creation of the Winter Weather Service Program as the 11th Service Program of the NWS in April of 2014. We also acknowledge numerous other individuals who embraced this important program, annually advancing winter weather observations, forecasts, warnings, assessments, research, and IDSS throughout the NWS and the weather enterprise. The authors are grateful to the anonymous reviewers for their insightful and constructive feedback. The views expressed in this article are those of the authors and do not necessarily represent those of the NWS.

Data availability statement. Much of the data analyzed in this paper was a review of existing data, which are openly available at locations cited in the reference section. Other data are available from the following public domain resources: University of Wyoming Sounding Archive (http://weather.uwyo.edu/upperair/sounding.html), Iowa State archived data and plots (https://mesonet.agron.iastate.edu/archive/), and NWS Performance Management (https://verification.nws.noaa.gov/services/public/).

References

Barjenbruch, K., and Coauthors, 2016: Drivers' awareness of and response to two significant winter storms impacting a metropolitan area in the intermountain west: Implications for improving traffic flow in inclement weather. *Wea. Climate Soc.*, **8**, 475–491, https://doi.org/10.1175/WCAS-D-16-0017.1.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, https://doi.org/10.1038/nature14956.

Benjamin, S. G., J. M. Brown, G. Brunet, P. Lynch, K. Saito, and T. W. Schlatter, 2019: 100 years of progress in forecasting and NWP applications. *A Century of Progress in Atmospheric and Related Sciences: Celebrating the American Meteorological Society Centennial*, *Meteor. Monogr.*, No. 59, Amer. Meteor. Soc., https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0020.1.

Black, A. W., and T. L. Mote, 2015: Characteristics of winter-precipitation-related transportation fatalities in the United States. *Wea. Climate Soc.*, **7**, 133–145, https://doi.org/10.1175/WCAS-D-14-00011.1.

Bougeault, P., and Coauthors, 2010: The THORPEX Interactive Grand Global Ensemble. *Bull. Amer. Meteor. Soc.*, **91**, 1059–1072, https://doi.org/10.1175/2010BAMS2853.1.

Brill, K. F., A. R. Fracasso, and C. M. Bailey, 2015: Applying a divisive clustering algorithm to a large ensemble for medium-range forecasting at the Weather Prediction Center. *Wea. Forecasting*, **30**, 873–891, https://doi.org/10.1175/WAF-D-14-00137.1.

Brotzge, J. A., and Coauthors, 2020: A technical overview of the New York State Mesonet standard network. *J. Atmos. Oceanic Technol.*, **37**, 1827–1845, https://doi.org/10.1175/JTECH-D-19-0220.1.

Brown, B., and Coauthors, 2021: The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bull. Amer. Meteor. Soc.*, **102**, 782–807, https://doi.org/10.1175/BAMS-D-19-0093.1.

Buizza, R., 2018: Statistical postprocessing of ensemble forecasts. *Ensemble Forecasting and the Need for Calibration*, S. Vannitsem, D. S. Wilks, and J. W. Messner, Eds., Elsevier, 15–48, https://doi.org/10.1016/B978-0-12-812372-0.00002-9.

——, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**, 1076–1097, https://doi.org/10.1175/MWR2905.1.

Burgeno, J. N., and S. L. Joslyn, 2020: The impact of weather forecast inconsistency on user trust. *Wea. Climate Soc.*, **12**, 679–694, https://doi.org/10.1175/WCAS-D-19-0074.1.

Call, D. A., and G. A. Flynt, 2021: The impact of snowfall on crashes, traffic volume, and revenue on the New York State Thruway. *Wea. Climate Soc.*, **14**, 131–141, https://doi.org/10.1175/WCAS-D-21-0074.1.

Cerruti, B. J., and S. G. Decker, 2011: The local winter storm scale: A measure of the intrinsic ability of winter storms to disrupt society. *Bull. Amer. Meteor. Soc.*, **92**, 721–737, https://doi.org/10.1175/2010BAMS3191.1.

Colle, B. A., R. Auld, K. Johnson, C. O'Connell, T. G. Taylor, and J. Rice, 2021: Improving communication of uncertainty and risk of high-impact weather through innovative forecaster workshops. *Bull. Amer. Meteor. Soc.*, **93**, E1424–E1430, https://doi.org/10.1175/BAMS-D-20-0108.1.

Craven, J. P., D. E. Rudack, and P. E. Shafer, 2020: National Blend of Models: A statistically post-processed multi-model ensemble. *J. Oper. Meteor.*, **8** (1), 1–14, https://doi.org/10.15191/nwajom.2020.0801.

Demnitz, R., and S. Joslyn, 2019: The effects of recency and numerical uncertainty estimates on overcautiousness. *Wea. Climate Soc.*, **11**, 309–322, https://doi.org/10.1175/WCAS-D-19-0115.1.

Demuth, J. L., and Coauthors, 2020: Recommendations for developing useful and usable convection-allowing model ensemble information for NWS forecasters. *Wea. Forecasting*, **35**, 1381–1406, https://doi.org/10.1175/WAF-D-19-0108.1.

FHWA, 2018: Pathfinder: Implementation recommendations. FHWA Doc. FHWA-HOP-18-059, 2 pp., https://ops.fhwa.dot.gov/publications/fhwahop18059/fhwahop18059.pdf.

Goodman, S. J., and Coauthors, 2012: The GOES-R proving ground: Accelerating user readiness for the next-generation geostationary environmental satellite system. *Bull. Amer. Meteor. Soc.*, **93**, 1029–1040, https://doi.org/10.1175/BAMS-D-11-00175.1.

Greybush, S. J., S. Saslo, and R. Grumm, 2017: Assessing the ensemble predictability of precipitation forecasts for the January 2015 and 2016 East Coast winter storms. *Wea. Forecasting*, **32**, 1057–1078, https://doi.org/10.1175/WAF-D-16-0153.1.

Griffin, E. M., T. J. Schuur, A. V. Ryzhkov, H. D. Reeves, and J. C. Picca, 2014: A polarimetric and microphysical investigation of the Northeast Blizzard of 8–9 February 2013. *Wea. Forecasting*, **29**, 1271–1294, https://doi.org/10.1175/WAF-D-14-00056.1.

Gurka, J. J., E. P. Auciello, A. F. Gigi, J. S. Waldstreicher, K. K. Keeter, S. Businger, and L. G. Lee, 1995: Winter weather forecasting throughout the eastern United States. Part II: An operational perspective of cyclogenesis. *Wea. Forecasting*, **10**, 21–41, https://doi.org/10.1175/1520-0434(1995)010<0021:WWFTTE>2.0.CO;2.

Hamill, T., E. Engle, D. Myrick, M. Peroutka, C. Finan, and M. Scheuerer, 2017: The U.S. National Blend of Models for statistical postprocessing of probability of precipitation and deterministic precipitation amount. *Mon. Wea. Rev.*, **145**, 3441–3463, https://doi.org/10.1175/MWR-D-16-0331.1.

Handler, S. L., H. D. Reeves, and A. McGovern, 2020: Development of a probabilistic subfreezing road temperature nowcast and forecast using machine learning. *Wea. Forecasting*, **35**, 1845–1863, https://doi.org/10.1175/WAF-D-19-0159.1.

Hirschberg, P., and Coauthors, 2011: A weather and climate enterprise strategic implementation plan for generating and communicating forecast uncertainty information. *Bull. Amer. Meteor. Soc.*, **92**, 1651–1666, https://doi.org/10.1175/BAMS-D-11-00073.1.

Hosterman, H. R., J. K. Lazo, J. M. Sprague-Hilderbrand, and J. E. Adkins, 2019: Using the National Weather Service's impact-based decision support services to prepare for extreme winter storms. *J. Emerg. Manage.*, **6**, 455–467, https://doi.org/10.5055/jem.2019.0439.

Houze, R. A., Jr., and Coauthors, 2017: The Olympic Mountains Experiment (OLYMPEX). *Bull. Amer. Meteor. Soc.*, **98**, 2167–2188, https://doi.org/10.1175/BAMS-D-16-0182.1.

Howard, S. P., K. E. Klockow-McClain, A. P. Boehmer, and K. M. Simmons, 2021: Firm behavior in the face of severe weather: Economic analysis between probabilistic and deterministic warnings. *Wea. Forecasting*, **33**, 757–767, https://doi.org/10.1175/WAF-D-20-0107.1.

Hurwitz, M. M., and Coauthors, 2020: Six priorities for investment in snow research and product development. *Bull. Amer. Meteor. Soc.*, **101**, E2025–E2029, https://doi.org/10.1175/BAMS-D-20-0218.1.

Jacobs, N. A., 2021: Open innovation and the case for community model development. *Bull. Amer. Meteor. Soc.*, **102**, E2002–E2011, https://doi.org/10.1175/BAMS-D-21-0030.1.

Keeter, K. K., S. Businger, L. G. Lee, and J. S. Waldstreicher, 1995: Winter weather forecasting throughout the eastern United States. Part III: The effects of topography and the variability of winter weather in the Carolinas and Virginia. *Wea. Forecasting*, **10**, 42–60, https://doi.org/10.1175/1520-0434(1995)010<0042:WWFTTE>2.0.CO;2.

Kenyon, J. S., D. Keyser, L. F. Bosart, and M. S. Evans, 2020: The motion of mesoscale snowbands in northeast U.S. winter storms. *Wea. Forecasting*, **35**, 83–105, https://doi.org/10.1175/WAF-D-19-0038.1.

Knapp, K. K., L. D. Smithson, and A. J. Khattak, 2000: The mobility and safety impacts of winter storm events in a freeway environment. *Proc. Mid-Continent Transportation Symp.*, Ames, IA, National Transportation Board, 71 pp., https://trid.trb.org/view.aspx?id=654630.

Kocin, P. J., and L. W. Uccellini, 2004a: *Northeast Snowstorms: Overview*. Vol. 1, *Meteor. Monogr.*, No. 54, Amer. Meteor. Soc., 296 pp.

——, and ——, 2004b: A snowfall impact scale derived from northeast storm snowfall distributions. *Bull. Amer. Meteor. Soc.*, **85**, 177–194, https://doi.org/10.1175/BAMS-85-2-Kocin.

——, ——, J. Alpert, B. Ballish, D. Bright, R. Grumm, and G. Manikin, 2011: The blizzard of 25–27 December 2010: Forecast assessment. Weather Prediction Center, 57 pp., www.wpc.ncep.noaa.gov/winter_storm_summaries/event_reviews/2010/December25_27_2010_Blizzard.pdf.

Korfe, N. G., and B. A. Colle, 2018: Evaluation of cool-season extratropical cyclones in a multimodel ensemble for eastern North America and the western Atlantic Ocean. *Wea. Forecasting*, **33**, 109–127, https://doi.org/10.1175/WAF-D-17-0036.1.

Kristovich, D. A. R., and Coauthors, 2017: The Ontario Winter Lake-Effect Systems field campaign: Scientific and educational adventures to further our knowledge and prediction of lake-effect storms. *Bull. Amer. Meteor. Soc.*, **98**, 315–332, https://doi.org/10.1175/BAMS-D-15-00034.1.

Lamberson, W. S., M. J. Bodner, J. A. Nelson, and S. A. Sienkiewicz, 2023: The use of ensemble clustering on a multimodel ensemble for medium-range forecasting at the Weather Prediction Center. *Wea. Forecasting*, https://doi.org/10.1175/WAF-D-22-0154.1, in press.

Lazo, J. K., H. R. Hosterman, J. M. Sprague-Hilderbrand, and J. E. Adkins, 2020: Impact-based decision support services and the socioeconomic impacts of winter storms. *Bull. Amer. Meteor. Soc.*, **101**, E626–E639, https://doi.org/10.1175/BAMS-D-18-0153.1.

Lin, S., and Coauthors, 2021: The immediate effects of winter storms and power outages on multiple health outcomes and the time windows of vulnerability. *Environ. Res.*, **196**, 110924, https://doi.org/10.1016/j.envres.2021.110924.

Maglaras, G. J., J. S. Waldstreicher, P. J. Kocin, A. F. Gigi, and R. A. Marine, 1995: Winter weather forecasting throughout the eastern United States. Part I: An overview. *Wea. Forecasting*, **10**, 5–20, https://doi.org/10.1175/1520-0434(1995)010<0005:WWFTTE>2.0.CO;2.

Magnusson, L., and Coauthors, 2022: Skill of medium-range forecast models using the same initial conditions. *Bull. Amer. Meteor. Soc.*, **103**, E2050–E2068, https://doi.org/10.1175/BAMS-D-21-0234.1.

Mahmood, R., and Coauthors, 2017: Mesonets: Mesoscale weather and climate observations for the United States. *Bull. Amer. Meteor. Soc.*, **98**, 1349–1361, https://doi.org/10.1175/BAMS-D-15-00258.1.

Market, P. S., and D. Cissell, 2002: Formation of a sharp snow gradient in a midwestern heavy snow event. *Wea. Forecasting*, **17**, 723–738, https://doi.org/10.1175/1520-0434(2002)017<0723:FOASSG>2.0.CO;2.

McGovern, A., D. J. Gagne II, R. Lagerquist, K. Elmore, and G. E. Jergensen, 2019: Making the black box more transparent: Understanding the physical implications of machine learning. *Bull. Amer. Meteor. Soc.*, **100**, 2175–2199, https://doi.org/10.1175/BAMS-D-18-0195.1.

——, and Coauthors, 2022: NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES). *Bull. Amer. Meteor. Soc.*, **103**, E1658–E1668, https://doi.org/10.1175/BAMS-D-21-0020.1.

McMurdie, L. A., and Coauthors, 2022: Chasing snowstorms: The Investigation of Microphysics and Precipitation for Atlantic Coast-Threatening Snowstorms (IMPACTS) campaign. *Bull. Amer. Meteor. Soc.*, **103**, E1243–E1269, https://doi.org/10.1175/BAMS-D-20-0246.1.

Merz, B., and Coauthors, 2020: Impact forecasting to support emergency management of natural hazards. *Rev. Geophys.*, **58**, e2020RG000704, https://doi.org/10.1029/2020RG000704.

Mills, B., J. Andrey, B. Doberstein, S. Doherty, and J. Yessis, 2019: Changing patterns of motor vehicle collision risk during winter storms: A new look at a pervasive problem. *Accid. Anal. Prev.*, **127**, 186–197, https://doi.org/10.1016/j.aap.2019.02.027.

——, ——, S. Doherty, B. Doberstein, and J. Yessis, 2020: Winter storms and fall-related injuries: Is it safer to walk than to drive? *Wea. Climate Soc.*, **12**, 421–434, https://doi.org/10.1175/WCAS-D-19-0099.1.

Morss, R. E., J. K. Lazo, and J. L. Demuth, 2010: Examining the use of weather forecasts in decision scenarios: Results from a US survey with implications for uncertainty communication. *Meteor. Appl.*, **17**, 149–162, https://doi.org/10.1002/met.196.

——, H. Lazrus, J. L. Demuth, and J. Henderson, 2022: Improving probabilistic weather forecasts for decision making: A multi-method study of the use of forecast information in snow and ice management at a major U.S. airport. NCAR Tech. Rep., NCAR/TN-573+STR, 43 pp., https://doi.org/10.5065/wyv1-wq11.

NASEM, 2018: *Integraphic Social and Behavioral Sciences within the Weather Enterprise.* National Academies Press, 198 pp.

NCEI, 2021: Billion-dollar weather and climate disasters. NOAA/NCEI, www.ncdc.noaa.gov/billions/events/US/2021.

NIST, 2013: Technical investigation of the May 22, 2011 tornado in Joplin, Missouri. National Institute of Standards and Technology Final Rep., 428 pp., www.nist.gov/publications/final-report-national-institute-standards-and-technology-nist-technical-investigation.

Niziol, T. A., W. R. Snyder, and J. S. Waldstreicher, 1995: Winter weather forecasting throughout the eastern United States. Part IV: Lake effect snow. *Wea. Forecasting*, **10**, 61–77, https://doi.org/10.1175/1520-0434(1995)010<0061:WWFTTE>2.0.CO;2.

NOAA, 2020: Precipitation prediction grand challenge strategy. NOAA Weather, Water, and Climate Board, 44 pp., www.noaa.gov/sites/default/files/2022-01/PPGC-Strategy_FINAL_2020-1030.pdf.

NOAA Science Advisory Board, 2021: Priorities for weather research. NOAA Science Advisory Board, 118 pp., https://sab.noaa.gov/wp-content/uploads/2021/11/PWR-Report_24Nov2021_Draft.pdf.

Novak, D. R., and B. A. Colle, 2012: Diagnosing snowband predictability using a multimodel ensemble system. *Wea. Forecasting*, **27**, 565–585, https://doi.org/10.1175/WAF-D-11-00047.1.

——, J. S. Waldstreicher, D. Keyser, and L. F. Bosart, 2006: A forecast strategy for anticipating cold season mesoscale band formation within eastern U.S. cyclones. *Wea. Forecasting*, **21**, 3–23, https://doi.org/10.1175/WAF907.1.

——, D. R. Bright, and M. J. Brennan, 2008: Operational forecaster uncertainty needs and future roles. *Wea. Forecasting*, **23**, 1069–1084, https://doi.org/10.1175/2008WAF2222142.1.

——, K. F. Brill, and W. A. Hogsett, 2014: Using percentiles to communicate snowfall uncertainty. *Wea. Forecasting*, **29**, 1259–1265, https://doi.org/10.1175/WAF-D-14-00019.1.

NRC, 2006: *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts.* National Academies Press, 124 pp., www.nap.edu/catalog/11699/completing-the-forecast-characterizing-and-communicating-uncertainty-for-better-decisions.

NWS, 2018: Communicating probabilistic information for decision-makers: A case study using experimental snow forecast products. NOAA NWS Final Rep., 48 pp., https://vlab.noaa.gov/web/nws-social-science/past-projects.

——, 2019: National Weather Service Policy directive 10-24: Impact-based decision support services. NOAA NWS, 19 pp., www.nws.noaa.gov/directives/sym/pd01024curr.pdf.

——, 2021: Winter storm severity index product description document. NWS, 7 pp., https://nws.weather.gov/products/PDD/PDD_OplWSSI_2020-2021.pdf.

——, 2022a: National gridded snowfall analysis product description document. NWS, 3 pp., https://nws.weather.gov/products/PDD/PDD_ExpNationalGriddedSnowfallAnalysis_2022.pdf.

——, 2022b: National weather service instruction 10-513: WFO winter weather products specification. NOAA NWS, 49 pp., www.nws.noaa.gov/directives/sym/pd01005013curr.pdf.

——, 2022c: Winter storm outlook product description document. NWS, 5 pp., https://nws.weather.gov/products/PDD/PDD_ExpWinter%20Storm%20Outlook_2020.pdf.

Palmer, T., 2017: The primacy of doubt: Evolution of numerical weather prediction from determinism to probability. *J. Adv. Model. Earth Syst.*, **9**, 730–734, https://doi.org/10.1002/2017MS000999.

Picca, J. C., D. M. Schultz, B. A. Colle, S. Ganetis, D. R. Novak, and M. J. Sienkiewicz, 2014: The value of dual-polarization radar in diagnosing the complex microphysical evolution of an intense snowband. *Bull. Amer. Meteor. Soc.*, **95**, 1825–1834, https://doi.org/10.1175/BAMS-D-13-00258.1.

Ralph, F. M., and Coauthors, 2005: Improving short-term (0–48 h) cool-season quantitative precipitation forecasting: Recommendations from a USWRP workshop. *Bull. Amer. Meteor. Soc.*, **86**, 1619–1632, https://doi.org/10.1175/BAMS-86-11-1619.

Rauber, R. M., and Coauthors, 2014: Stability and charging characteristics of the comma head region of continental winter cyclones. *J. Atmos. Sci.*, **71**, 1559–1582, https://doi.org/10.1175/JAS-D-13-0253.1.

Renn, O., 2008: Review of psychological, social and cultural factors of risk perception. *Risk Governance: Coping with Uncertainty in a Complex World*, O. Renn, Ed., Routledge, 98–148.

Rice, R. G., and P. R. Spence, 2016: Thor visits Lexington: Exploration of the knowledge-sharing gap and risk management learning in social media during multiple winter storms. *Comput. Human Behav.*, **65**, 612–618, https://doi.org/10.1016/j.chb.2016.05.088.

Ripberger, J., A. Bell, A. Fox, A. Forney, W. Livingston, C. Gaddie, C. Silva, and H. Jenkins-Smith, 2022: Communicating probability information in weather forecasts: Findings and recommendations from a living systematic review of the research literature. *Wea. Climate Soc.*, **14**, 481–498, https://doi.org/10.1175/WCAS-D-21-0034.1.

Roberts, B., I. L. Jirak, A. J. Clark, S. J. Weiss, and J. S. Kain, 2019: Postprocessing and visualization techniques for convection-allowing ensembles. *Bull. Amer. Meteor. Soc.*, **100**, 1245–1258, https://doi.org/10.1175/BAMS-D-18-0041.1.

——, B. T. Gallo, I. L. Jirak, A. J. Clark, D. C. Dowell, X. Wang, and Y. Wang, 2020: What does a convection-allowing ensemble of opportunity buy us in forecasting thunderstorms? *Wea. Forecasting*, **35**, 2293–2316, https://doi.org/10.1175/WAF-D-20-0069.1.

Romine, G. S., C. S. Schwartz, J. Berner, K. R. Fossell, C. Snyder, J. L. Anderson, and M. L. Weisman, 2014: Representing forecast error in a convection-permitting ensemble system. *Mon. Wea. Rev.*, **142**, 4519–4541, https://doi.org/10.1175/MWR-D-14-00100.1.

Rooney, J. F., Jr., 1967: The urban snow hazard in the United States: An appraisal of disruption. *Geogr. Rev.*, **57**, 538–559, https://doi.org/10.2307/212932.

Rothfusz, L. P., R. Schneider, D. Novak, K. Klockow-McClain, A. E. Gerard, C. Karstens, G. J. Stumpf, and, T. M. Smith, 2018: FACETs: A proposed next-generation paradigm for high-impact weather forecasting. *Bull. Amer. Meteor. Soc.*, **99**, 2025–2043, https://doi.org/10.1175/BAMS-D-16-0100.1.

Scheuerer, M., and T. M. Hamill, 2019: Probabilistic forecasting of snowfall amounts using a hybrid between a parametric and an analog approach. *Mon. Wea. Rev.*, **147**, 1047–1064, https://doi.org/10.1175/MWR-D-18-0273.1.

Schwartz, C. S., G. S. Romine, R. A. Sobash, K. R. Fossell, and M. L. Weisman, 2019: NCAR's real-time convection-allowing ensemble project. *Bull. Amer. Meteor. Soc.*, **100**, 321–343, https://doi.org/10.1175/BAMS-D-17-0297.1.

Senkbeil, J. C., D. A. Scott, P. Guinazu-Walker, and S. M. Rockman, 2013: Ethnic and racial differences in tornado hazard perception, preparedness, and shelter lead time in Tuscaloosa. *Prof. Geogr.*, **66**, 610–620, https://doi.org/10.1080/00330124.2013.826562.

Squires, M. F., J. H. Lawrimore, R. R. Jr. Heim, D. A. Robinson, M. R. Gerbush, and T. W. Estilow, 2014: The regional snowfall index. *Bull. Amer. Meteor. Soc.*, **95**, 1835–1848, https://doi.org/10.1175/BAMS-D-13-00101.1.

SRA, 2022: Society for risk analysis glossary. SRA, 9 pp., www.sra.org/wp-content/uploads/2020/04/SRA-Glossary-FINAL.pdf.

Steiner, M., A. R. S. Anderson, S. Landolt, S. Linden, and B. R. J. Schwedler, 2015: Coping with adverse winter weather: Emerging capabilities in support of airport and airline operations. *J. Air Traffic Control*, 36–45.

Stuart, N. A., and Coauthors, 2022: The evolving role of humans in weather prediction and communication. *Bull. Amer. Meteor. Soc.*, **103**, E1720–E1746, https://doi.org/10.1175/BAMS-D-20-0326.1.

Su, C., J. N. Burgeno, and S. Joslyn, 2021: The effects of consistency among simultaneous forecasts on weather-related decisions. *Wea. Climate Soc.*, **13**, 3–10, https://doi.org/10.1175/WCAS-D-19-0089.1.

Swinbank, R., and Coauthors, 2016: The TIGGE project and its achievements. *Bull. Amer. Meteor. Soc.*, **97**, 49–67, https://doi.org/10.1175/BAMS-D-13-00191.1.

Tobin, D. M., H. D. Reeves, M. N. Gibson, and A. A. Rosenow, 2021: Weather conditions and messaging associated with fatal winter-weather-related motor-vehicle crashes. *Wea. Climate Soc.*, **14**, 835–848, https://doi.org/10.1175/WCAS-D-21-0112.1.

Tracton, M. S., 2008: Must surprise snowstorms be a surprise. *Synoptic-Dynamic Meteorology and Weather Analysis and Forecasting - A Tribute to Fred Sanders*, L. F. Bosart and H. B Bluestein, Eds., Amer. Meteor. Soc., 251–268.

Tripp, D. D., J. E. Trujillo-Falcon, K. E. Kockow-McClain, H. D. Reeves, K. L. Berry, J. S. Waldstreicher, and J. A. Nelson, 2023: Foundational needs of forecasters for probabilistic winter forecasting. *Wea. Forecasting*, **38**, 3–15, https://doi.org/10.1175/WAF-D-22-0116.1.

Uccellini, L. W., and J. E. Ten Hoeve, 2019: Evolving the National Weather Service to build a weather-ready nation: Connecting observations, forecasts, and warnings to decision-makers through impact-based decision support services. *Bull. Amer. Meteor. Soc.*, **100**, 1923–1942, https://doi.org/10.1175/BAMS-D-18-0159.1.

——, R. W. Spinrad, D. M. Koch, C. N. McLean, and W. M. Lapenta, 2022: EPIC as a catalyst for NOAA's Future Earth Prediction Systems. *Bull. Amer. Meteor. Soc.*, **103**, E2246–E2264, https://doi.org/10.1175/BAMS-D-21-0061.1.

Voisin, N., J. C. Schaake, and D. P. Lettenmaier, 2010: Calibration and downscaling methods for quantitative ensemble precipitation forecasts. *Wea. Forecasting*, **25**, 1603–1627, https://doi.org/10.1175/2010WAF2222367.1.

Waldstreicher, J. S., 2005: Assessing the impact of collaborative research projects on NWS warning performance. *Bull. Amer. Meteor. Soc.*, **86**, 193–204, https://doi.org/10.1175/BAMS-86-2-193.

——, and D. Radell, 2018: The NWS probabilistic snowfall forecast experiment: 2016–17 review and future plans. *13th Symp. on Societal Applications*, Austin, TX, Amer. Meteor. Soc., 8.3, https://ams.confex.com/ams/98Annual/webprogram/Paper333150.html.

Weyrich, P., A. Scolobig, and A. Patt, 2019: Dealing with inconsistent weather warnings: Effects on warning quality and intended actions. *Meteor. Appl.*, **26**, 569–583, https://doi.org/10.1002/met.1785.

Williams, C. A., and G. M. Eosco, 2021: Is a consistent message achievable? Defining "message consistency" for weather enterprise researchers and practitioners. *Bull. Amer. Meteor. Soc.*, **102**, E279–E295, https://doi.org/10.1175/BAMS-D-18-0250.1.

WMO, WBG, GFDRR, and USAID, 2015: Valuing weather and climate: Economic assessment of meteorological and hydrological services. WMO-1153, 286 pp., https://library.wmo.int/doc_num.php?explnum_id=3314.

Zhang, Y., and B. Liang, 2021: Evaluating the vulnerability of farming communities to winter storms in Iowa, US. *Environ. Sustainability Indic.*, **11**, 100126, https://doi.org/10.1016/j.indic.2021.100126.

Zheng, M., E. K. M. Chang, and B. A. Colle, 2013: Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Wea. Forecasting*, **28**, 1133–1156, https://doi.org/10.1175/WAF-D-12-00132.1.

——, E. K. Chang, B. A. Colle, Y. Luo, and Y. Zhu, 2017: Applying fuzzy clustering to a multimodel ensemble for U.S. East Coast winter storms: Scenario identification and forecast verification. *Wea. Forecasting*, **32**, 881–903, https://doi.org/10.1175/WAF-D-16-0112.1.

——, ——, and ——, 2019: Evaluating U.S. East Coast winter storms in a multimodel ensemble using EOF and clustering approaches. *Mon. Wea. Rev.*, **147**, 1967–1987, https://doi.org/10.1175/MWR-D-18-0052.1.