BAMS Article

The Evolving Role of Humans in Weather Prediction and Communication

Neil A. Stuart, Gail Hartfield, David M. Schultz, Katie Wilson, Gregory West, Robert Hoffman, Gary Lackmann, Harold Brooks, Paul Roebber, Teresa Bals-Elsholz, Holly Obermeier, Falko Judt, Patrick Market, Daniel Nietfeld, Bruce Telfeyan, Dan DePodwin, Jeffrey Fries, Elliot Abrams, and Jerry Shields

ABSTRACT: A series of webinars and panel discussions were conducted on the topic of the evolving role of humans in weather prediction and communication, in recognition of the 100th anniversary of the founding of the AMS. One main theme that arose was the inevitability that new tools using artificial intelligence will improve data analysis, forecasting, and communication. We discussed what tools are being created, how they are being created, and how the tools will potentially affect various duties for operational meteorologists in multiple sectors of the profession. Even as artificial intelligence increases automation, humans will remain a vital part of the forecast process as that process changes over time. Additionally, both university training and professional development must be revised to accommodate the evolving forecasting process, including addressing the need for computing and data skills (including artificial intelligence and visualization), probabilistic and ensemble forecasting, decision support, and communication skills. These changing skill sets necessitate that both the U.S. Government's Meteorologist General Schedule 1340 requirements and the AMS standards for a bachelor's degree need to be revised. Seven recommendations are presented for student and forecaster preparation and career planning, highlighting the need for students and operational meteorologists to be flexible lifelong learners, acquire new skills, and be engaged in the changes to forecast technology in order to best serve the user community throughout their careers. The article closes with our vision for the ways that humans can maintain an essential role in weather prediction and communication, highlighting the interdependent relationship between computers and humans.

KEYWORDS: Operational forecasting; Neural networks; Numerical analysis/modeling; Communications/decision making; Societal impacts; Artificial intelligence

https://doi.org/10.1175/BAMS-D-20-0326.1

Corresponding author: Neil A. Stuart, neil.stuart@noaa.gov

In final form 7 May 2022

AFFILIATIONS: Stuart—NOAA/National Weather Service, Albany, New York; **Hartfield**—NOAA/National Weather Service, Raleigh, North Carolina; **Schultz**—Centre for Atmospheric Science, Department of Earth and Environmental Sciences, and Centre for Crisis Studies and Mitigation, University of Manchester, Manchester, United Kingdom; **Wilson**—Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, and NOAA/OAR/NSSL, Norman, Oklahoma; **West**—BC Hydro, Burnaby, British Columbia, Canada; **Hoffman**—Florida Institute for Human and Machine Cognition, Pensacola, Florida; **Lackmann**—North Carolina State University, Raleigh, North Carolina; **Brooks**—National Severe Storms Laboratory, and University of Oklahoma, Norman, Oklahoma; **Roebber**—University of Wisconsin–Milwaukee, Milwaukee, Wisconsin; **Bals-Elsholz**—Department of Geography and Meteorology, Valparaiso University, Valparaiso, Indiana; **Obermeier**—Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, Oklahoma; **Judt**—Mesoscale and Microscale Meteorology Laboratory, NCAR, Boulder, Colorado; **Market**—University of Missouri, Columbia, Missouri; **Nietfeld**—NOAA/OAR/ESRL/GSL, Boulder, Colorado; **Telfeyan***—557th Weather Wing, Offutt AFB, Nebraska; **DePodwin and Abrams***—AccuWeather, State College, Pennsylvania; **Fries**—Operations Standards and Tactics, 1st Weather Group (ACC), Offutt AFB, Nebraska; **Shields**—Ontario Ministry of Natural Resources and Forestry, Peterborough, Ontario, Canada
*Retired

In a rapidly changing world, scientific disciplines must evolve. In this regard, the weather, water, and climate enterprise has been a success. The tremendous improvements in human forecasting skill in recent decades have largely been the result of applying advances in atmospheric science and integrating these advances into the forecast process. These advances include the development and implementation of dense data networks, improved numerical weather prediction (NWP) models, and faster computers. More recently, artificial intelligence (AI), machine learning, and deep learning are improving both weather predictions and the communication of weather impacts. Even though these technologies can relieve humans of repetitive, boring, or onerous tasks (see the "Definitions" sidebar for types of AI), their accuracy may surpass human skill (e.g., Baars and Mass 2005; Novak et al. 2008; Erickson et al. 2021). Indeed, weather forecasting has evolved far beyond hand analysis of data, interpreting NWP model output, and generating text-based forecasts sent to users at specific times. Operational meteorologists use smart tools to selectively edit graphical and/or digital databases initialized by NWP guidance. But is editing automated forecasts really "forecasting"? How much continuing education, training, and resultant human-added skill is enough to justify humans remaining in the forecast process (e.g., Huntemann et al. 2015)?

Definitions

AI is an increasingly important approach to prediction, but describing every specific technique is beyond the scope of this study. Here, we explain the relationship between AI, machine learning, and deep learning.

AI refers to computational systems able to perform tasks that normally require human intelligence, but with increased efficiency, precision, and objectivity (NOAA 2020a).

A subset of AI called *machine learning* refers to mathematical models able to perform a specific task without using explicit instructions, instead relying on patterns and inference. The use of labeled training data can further improve the AI predictive capability through *supervised machine learning*.

Deep learning is a subset of machine learning that uses artificial neural networks capable of learning from unstructured or newly added data.

Self-reflection by the meteorological community on how the role of humans in weather prediction would evolve is not new (e.g., Snellman 1977; Brooks et al. 1996; McIntyre 1999; Bosart 2003; Baars and Mass 2005; Stuart et al. 2006; Sills 2009). Some have even stated that forecasting might be completely automated in 10–20 years (e.g., Rose et al. 2015), resulting in reductions in staffing of operational meteorologists. Proposed plans to adjust or reduce operational staffing within the National Weather Service (NWS 2017; NOAA 2020b) can also feed into operational meteorologists' concerns about their roles in a future forecasting environment. These concerns can lead to poor workforce morale, reduced performance, and difficulty retaining forecast staff (Carrington 2016; Frone and Blais, 2020). However, others argue against humans being replaced by automation (e.g., McIntyre 1999; Bosart 2003; Doswell 2004; Baars and Mass 2005; Erkkilä 2007).

There is an ever-evolving balance between the accuracy and speed of producing the forecast versus providing decision support for the user community (Murphy 1993), but this balance is especially crucial as emphasis shifts to more impact-based decision support services (Uccellini and Ten Hoeve 2019). History has shown that automation is not autonomous (see the "Misconceptions about the human–computer relationship" sidebar). The human–computer relationship is, and will remain, one of interdependence. The discussion should be about human *plus* machine (Karstens et al. 2018), not human *versus* machine (Hoffman et al. 2017).

Given the continued evolution of weather prediction and communication that has taken place since the Symposia on the Future Role of Humans in the Forecast Process in 2004–05 (Stuart et al. 2006, 2007), an update was due. The 100th anniversary of the American Meteorological Society in 2019 was appropriate for such self-reflection. This article provides that update through discussions of what tools are being developed, how they are being developed, and how the tools will potentially affect duties for operational meteorologists in the various sectors of the profession.

Misconceptions about the human–computer relationship

The meteorological community must maintain clarity concerning the language that is typically used to discuss computers and AI, as well as understand the nature of the human–computer relationship. A number of misconceptions in terminology have arisen (Bradshaw et al. 2013; Hoffman et al. 2014). One misconception surrounds the word "automation." It is often taken for granted that automation is autonomous, but it never is. A computational system is never entirely self-sufficient, and its ability to self-direct is always context bound and fragile.

"More automation means we need fewer people" is a misconception, as well. In fact, as the workplace becomes more computerized, more people will be needed who are experts at creating, using, and repairing the technology (Blackhurst et al. 2011). That means more technical training, more expertise, and greater labor costs, not lower labor costs.

Another noteworthy misconception is that all we need is more computer memory and more processing power, then a miracle will occur, and the human will no longer be needed. This misconception runs counter to established knowledge within the field of AI and leads directly to what may be the most important point to make here. Since the days of World War II, the premise of computer technology has been that computers have desirable characteristics that compensate for the weakness of humans. Machines are rational, humans are emotional; computers have large memory capacity, humans have limited memory capacity. Historically, this approach has manifestly ignored the flip side: the limitations of computers and the strengths of humans. What we have witnessed is that when the computer's model of the world becomes misaligned, it is the humans who jump in to fix things. In fact, humans and computers are interdependent. Computers have limitations *and* depend on humans to keep them aligned with the world. Humans have strengths *and* need computers to help them maintain situational awareness. As is true for all domains of human activity in which computers and AI are becoming a more engulfing aspect of the workplace, the meteorological community must escape the "human versus machine" mentality and embrace the "human plus machine" mentality.

A final important point has to do with the apparent promulgation of new AI systems, new data types, new visualization tools, etc. This article provides many examples of new systems and products, both operational and experimental. The abundance of such systems in an operational forecasting environment may lead to data overload. Such overload must be mitigated through operational contingencies and considerations, such that operational meteorologists can utilize available information to enhance their situational awareness and promote expertise.

In this article, we generalize *weather prediction* to forecasts for all hydrometeorological phenomena. We contend that the future forecast process will be dependent on computers and AI in the development and use of tools and forecast guidance that will shape the role of humans in the weather prediction and communication process. Weather prediction and communication will vary from sector to sector within the weather, water, and climate enterprise, and space limits the scope of this article. Our examples will draw mainly from the United States and the National Weather Service (NWS), although some private-sector and military examples are also provided. Even as AI increases automation, we acknowledge that humans will determine new ways to guide an ever-changing forecast process. We recognize that the future brings considerable uncertainty, and that new developments will inevitably bring changes that we cannot fully anticipate at present, in the United States and around the world. We envision a range of possible outcomes. One scenario may be that some decades into the future, the role of humans in weather forecasting will be largely confined to algorithm development, verification activities, and stakeholder relations. However, our consensus view reflects a more human-engaged scenario, particularly for communication and decision support.

Methods

A team of 18 meteorologists and 1 cognitive-systems engineer representing all sectors of the weather, water, and climate enterprise planned, created, and conducted four webinars that were delivered between September 2019 and January 2020 (Table 1). This team agreed on four overarching topics related to the evolving role of humans in weather prediction and communication into the future:

- development of automated forecasting tools,
- use of automated forecasting tools,
- training and proficiency for future forecasting, and
- envisioning the future of weather prediction and communication.

Each of the four webinars (presented in the order above) was attended by users via 60–80 independent connections (some connections included multiple viewers, so the actual audience size was larger). The audience represented a diverse cross section of the weather, water, and climate enterprise. After each webinar, questions, comments, and suggestions were provided by each audience (www.ametsoc.org/index.cfm/ams/webinar-directory/).

Panel discussions on the same four topics were conducted at the 100th AMS Annual Meeting in January 2020 to extend the webinar audience and gather additional perspectives (Table 2). There were ~40–80 attendees at each panel discussion. Email correspondence was also received from some webinar and panel-discussion attendees, with added input from face-to-face meetings with some people whose schedules could not accommodate attendance at the panel discussions. In the sections that follow, we summarize these webinars and panel discussions, combining the first two subjects (development and use of automated forecasting tools) into one section.

Table 1. Information on the four webinars that informed this article. Neil Stuart was the moderator for all webinars.

Date	Title	Presenters	URL
23 Oct 2019	Developing Tools for Forecasting and Communication: The Human Role in Their Design	Dan Nietfeld, Falko Judt, Greg West, Patrick Market, and Robert Hoffman	https://bit.ly/3k6xj21
22 Nov 2019	The Evolving Role of Humans in Weather Prediction and Communication: The Human/Automation Relationship—How can we best use AI tools?	Gail Hartfield, Jeffrey Fries, Bruce Telfeyan, and Robert Hoffman	https://bit.ly/33bzoCP
4 Dec 2019	The Evolving Role of Humans in Weather Prediction and Communication: Training and Proficiency for Future Forecasting	David Schultz, Harold Brooks, Gary Lackmann, and Paul Roebber	https://bit.ly/2R8vu89
8 Jan 2020	The Evolving Role of the Human in Weather Prediction and Communication: Envisioning the Future Forecast Process	Dan DePodwin, Holly Obermeier, Katie Wilson, and Elliot Abrams	https://bit.ly/3hf88IG

Table 2. Information on four panel discussions at the 100th AMS Annual Meeting on 15 Jan 2020 that informed this article. Neil Stuart was the moderator for all panel discussions. No recordings were made of the panel discussions. Any differences between the list of panelists in in the panel discussions and the official AMS Annual Meeting program (https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Program/1441) are due to listed presenters not in attendance and being replaced with others.

Time	Title	Panelists
0830–1000 LT	Development of Automated Forecasting Tools: Types and the Human Role in Their Design	Gail Hartfield, Falko Judt, Daniel Nietfeld, and Gregory West
1030–1200 LT	The Evolving Role of the Human in Weather Prediction and Communication: Use of Automated Forecasting Tools vs Humans	Jeffrey Fries, Gail Hartfield, Jerry Shields, and Bruce Telfeyan
1330–1430 LT	The Evolving Role of the Human in Weather Prediction and Communication: Training and Proficiency for Future Forecasting	Teresa Bals-Elsholz, Harold Brooks, Gary Lackmann, and Paul Roebber
1500–1600 LT	The Evolving Role of the Human in Weather Prediction and Communication: Envisioning the Future Forecast Process	Elliot Abrams, Daniel DePodwin, Gail Hartfield, Holly Obermeier, and Katie A. Wilson

Development and use of automated forecasting tools

Automated tools can be divided into three broad categories representing the operational forecast process: those that improve data quality, analysis, and assimilation; those that improve the accuracy of weather forecasts and warnings; and those that improve communication. In this section, we provide an overview of some of the operational and experimental tools being used within each of these categories.

Automated tools that improve data quality, analysis, and assimilation. These tools identify unrepresentative observations and determine an optimal interpolation between representative observations, while combining them with first-guess model fields. They also provide error statistics, dynamical relationships, and quality control (i.e., data assimilation). Once datasets have been quality controlled and considered the best representation of the current state of the atmosphere, NWP models can be initialized and integrated forward in time, and derived fields can be generated from their output. These steps are mostly automated, with such tools being used to quality control radar data (e.g., Lakshmanan et al. 2007), quality control hydrological data (e.g., Zhao et al. 2018), correct model bias in data assimilation (e.g., Berry and Harlim 2017), and incorporate remote sensing data into NWP (e.g., Boukabara et al. 2019).

AI tools can be used to improve analyses that serve as the analysis of record for model verification. These include advancements within two critical NWS analysis programs, the Real-Time Mesoscale Analysis and the Unrestricted Mesoscale Analysis, which are used as the training datasets and verifying observational datasets for the surface forecasts for the National Blend of Models (Tew et al. 2016; Craven et al. 2020). Both the Real-Time and Unrestricted Mesoscale Analyses are also used to bias correct inputs into the National Blend of Models. In addition, the Storm Prediction Center (SPC) produces automated objective analyses of upper-air and surface data (www.spc.noaa.gov/exper/mesoanalysis/help/begin.html) that are initialized using NWP (RAP model output). Operational meteorologists use analyses such as these from the SPC to assess the synoptic-scale and mesoscale environments, although these analyses are not without error (e.g., Privé and Errico 2013). AI will also be used increasingly in the future of data quality and processing, identifying, and correcting erroneous observations, and improving interpolation between observation points. Machine learning is also required to optimize initial conditions for convection-allowing models (i.e., models with convective parameterizations turned off so that they explicitly simulate convection, typically at horizontal grid spacing of 4 km or less and time scales of 0–48 h).

Operational meteorologists realize tools that operate behind the scenes, such as quality control of observations and interpolation between observation points, are critical to the performance of NWP models. Without such tools, humans would struggle to quality control the vast amounts of data going into NWP models. However, operational meteorologists must still look critically at real-time observations including examining satellite data, surface and upper-air analyses, radar data, and soundings. Operational meteorologists bring value to the forecast process through quality control of data and NWP model guidance (such as satellite data overlaid with model initialization), identifying areas where model initialization is poor or where models are not performing well. They also must be vigilant to recognize when an automated process has gone awry (see the “Misconceptions about the human–computer relationship” sidebar).

Automated tools that improve weather forecasts and warnings. The interfaces that humans are, and will be, using in the forecast process influence the development of tools for weather predictions and warnings. In the NWS and U.S. Air Force, an emerging generation of smart tools is becoming available to enable forecaster manipulation of deterministic gridded weather elements, based on preferred model or ensemble output. The AWIPS Graphical Forecast Editor allows a gridded database to be initialized by, or blended with, many choices of data and NWP deterministic and ensemble guidance. The U.S. Air Force’s Forecaster-in-the-Loop (FITL)-to-Grib (F2G) toolkit (Fries 2014) optimizes characterization of aviation hazards. Although NWS and Air Force operational meteorologists can still control the forecast by deciding on the deterministic model or model blend with which to produce an initial set of forecast fields and making alterations, the increasing accuracy of input sources will likely reduce this role in the future to only high-impact events.

In contrast, the private sector uses different interfaces and tools. One example is the Applications Programming Interface using Microsoft Windows in the “Human Over the Loop” paradigm at The Weather Channel (Rose et al. 2015). Operational meteorologists use filters and qualifiers to edit geonavigated digital polygons where real-time observations and NWP output flow through the database (Fig. 1), highlighting the interdependence between humans and machines. Another example is IBM’s Deep Thunder, which uses deep learning to create localized forecasts (Treinish et al. 2003; Ferrucci 2012). Additionally, Google uses radar data and convolutional neural networks for nowcasting (0–6-h forecasting) precipitation (Agrawal et al. 2019).

Different types of machine learning (e.g., neural networks and random forests) and adjusted multiple linear regression equations [e.g., model output statistics (MOS)] have been employed as tools to improve weather forecasts for decades (e.g., Glahn and Lowry 1972; Benjamin et al. 2019). In the past 20 years, however, research into AI has evolved from simple neural networks that are saturated in skill with relatively small amounts of training data to complex machine learning and deep learning applications. These machine learning applications are able to infer more complex relationships and continue gaining skill when given far greater amounts of training data (e.g., Ng 2018).

Other efforts to develop machine learning models have been increasing in the research sector (e.g., Roebber 2013; Herman and Schumacher 2018; McGovern et al. 2019). In the Global Synthetic Weather Radar (GSWR) project, the U.S. Air Force is employing AI and machine learning to fuse sensed data and output from the Air Force’s Global Air-Land Weather Exploitation Model (GALWEM; Stoffler 2017). Further advances include severe weather probabilities (ProbSevere; Cintineo et al. 2020) from the Multi-Radar/Multi-Sensor (MRMS) system (Smith et al. 2016). ProbSevere is a statistical model that uses the RAP model, GOES satellite data, and radar data to improve short-term probabilistic forecasting guidance in severe-weather situations (Fig. 2). Other efforts include the Warn-on-Forecast System (WoFS; Skinner et al. 2018), which consists of an ensemble of 36 individual members, and warning-scale probabilistic hazard information (PHI; Karstens et al. 2015). Experiments with WoFS, PHI, and other tools have been conducted at the Hazardous Weather Testbed (HWT; Clark et al. 2012; Calhoun et al. 2021), in which operational meteorologists evaluate

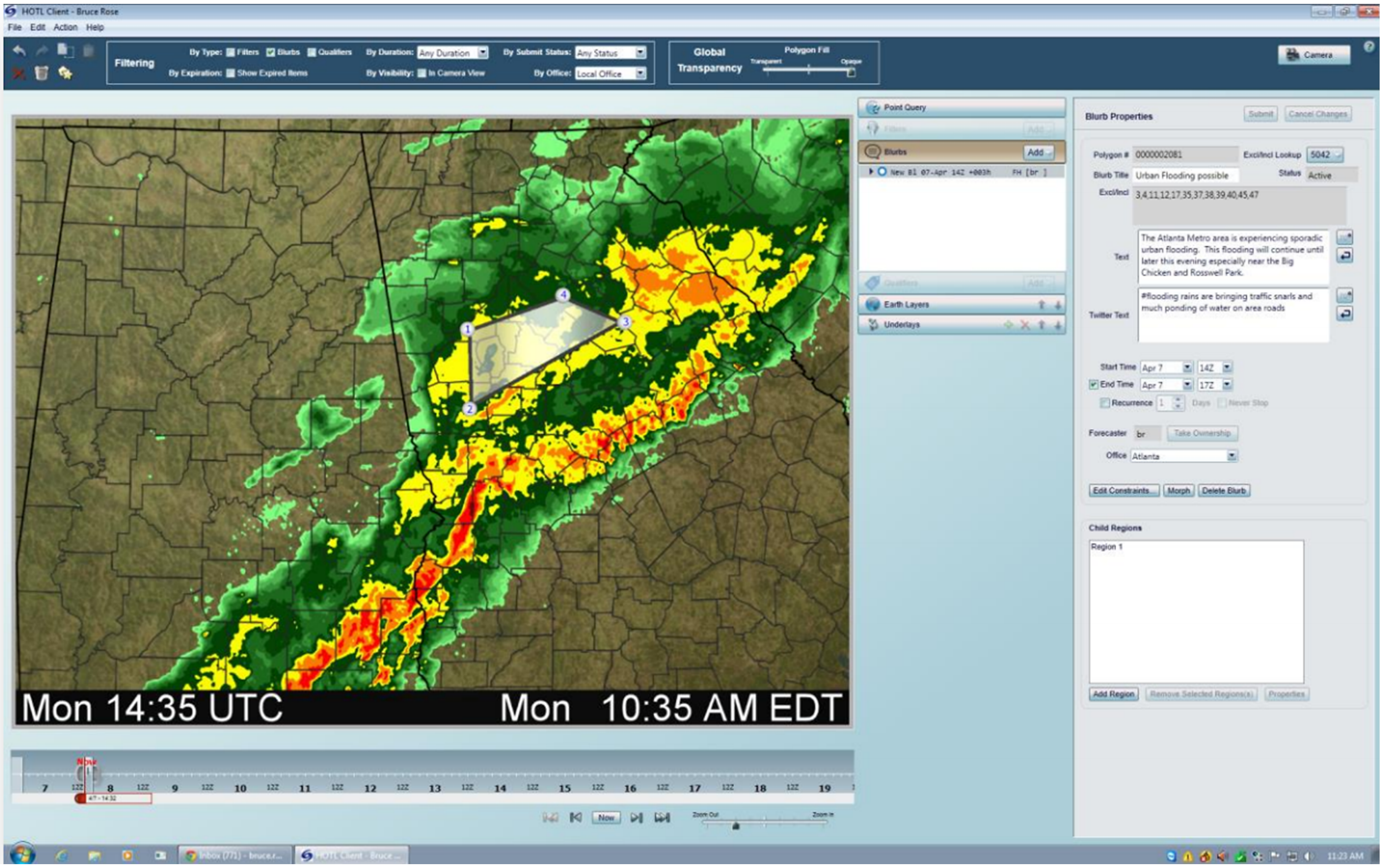

Fig. 1. **The "Human Over the Loop" graphical user interface runs on Windows machines and is composed of a graphics work area showing some portion of Earth, GIS layers, and weather "underlays." Forecasters draw polygons or clusters of polygons and then assign properties to those polygons that filter, qualify, or further amplify the underlying gridded forecasts. The forecaster instructions embodied in the polygonal objects are only applied when a forecast is requested from the API thus parallelizing the flow of forecast guidance and forecast intervention from the human (Fig. 2.1 from Rose et al. 2015).**

guidance and tools in development by using them during real-time and simulated severe weather events (Fig. 3).

Recent cross-sector collaboration efforts have provided new opportunities to accelerate NWP innovations. New models are being developed for operational meteorologists under the umbrella of the Unified Forecast System (UFS; Tallapragada 2020). The National Center for Atmospheric Research (NCAR) and IBM are collaborating on the Model for Prediction Across Scales (MPAS 2013), which uses a hexagonal mesh resembling a honeycomb that can be stretched wide in some regions and compressed for higher resolution in others. Additionally, initiatives are in development in which a shared cloud environment would facilitate experimentation and modification of NWP model code to improve the accuracy of the models (including the Earth Prediction Innovation Center; Jacobs 2021). Last, NOAA has unveiled an AI strategy (NOAA 2020a) designed to increase development and use of AI across its varied mission areas.

After decades of employing larger-scale models with parameterized convection, operational meteorologists are increasingly using output from convection-allowing models (e.g., Roebber et al. 2004; Clark et al. 2016; Sobash et al. 2020). Such models include the deterministic HRRR (Alexander et al. 2020) and NAM (3-km horizontal grid spacing). Simulated reflectivity products generated from convection-allowing models closely resemble real-time radar imagery in many cases. Although they can often accurately depict the mode, timing, and placement of precipitation features, the realistic representations can lull a forecaster into accepting it as fact, if the forecaster fails to maintain an appropriate level of critique. Additionally, the WoFS and the High Resolution Ensemble Forecast (HREF) systems produce ensemble products from several different

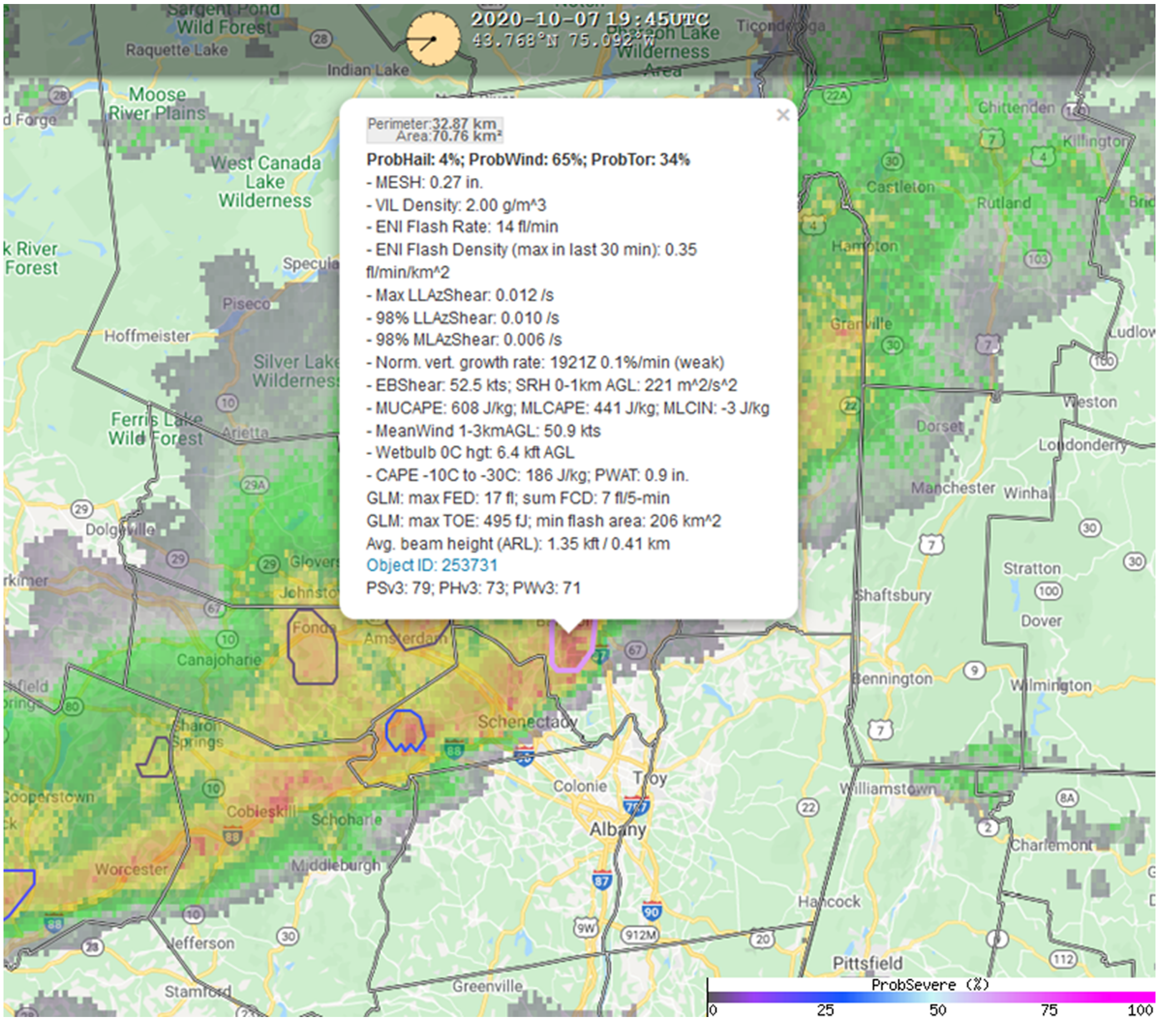

Fig. 2. Multi-Radar/Multi-Sensor (MRMS) system ProbSevere display with radar reflectivity overlay at 1945 UTC 7 Oct 2020 showing the probability of severe hail, wind, and tornadoes for a thunderstorm near the Capital Region of New York. Color contours represent the maximum probability for severe weather type in a radar-derived thunderstorm centroid. Note the 4% probability for hail, 65% probability for severe winds, and 34% probability for tornadoes. Other severe weather parameters are included such as maximum estimated hail size (MESH), vertically integrated liquid density, instability and shear parameters and lightning data. [Figure courtesy of the Cooperative Institute for Satellite Studies (CIMMS) at the Science and Space Engineering Center (SSEC) at the University of Wisconsin–Madison.]

models with around 3-km horizontal grid spacing (Roberts et al. 2019). Automated output based on these high-resolution models, such as predicted neighborhood probabilities of updraft helicity, can help operational meteorologists mitigate the concern of forecasters accepting a single deterministic solution and gauge the risk of particular weather hazards.

The development and use of ensemble forecast systems have grown considerably over the last 25 years (e.g., Buizza 2018). However, most ensemble forecasts must be calibrated for their probabilities to be reliable. While run-to-run continuity and ensemble member agreement may inform forecaster confidence, a probabilistically reliable forecast system is key. Thankfully, the need for calibration is well understood and research into optimizing calibration techniques is becoming more commonplace (e.g., Duan et al. 2017; Voudouri et al. 2017; Buizza 2018). But operational meteorologists still need to remain aware of the continued limitations of ensemble forecasts.

Regardless of the application, a basic understanding of these tools is required to gain operational meteorologists' confidence. They must also be able to recognize unrealistic output,

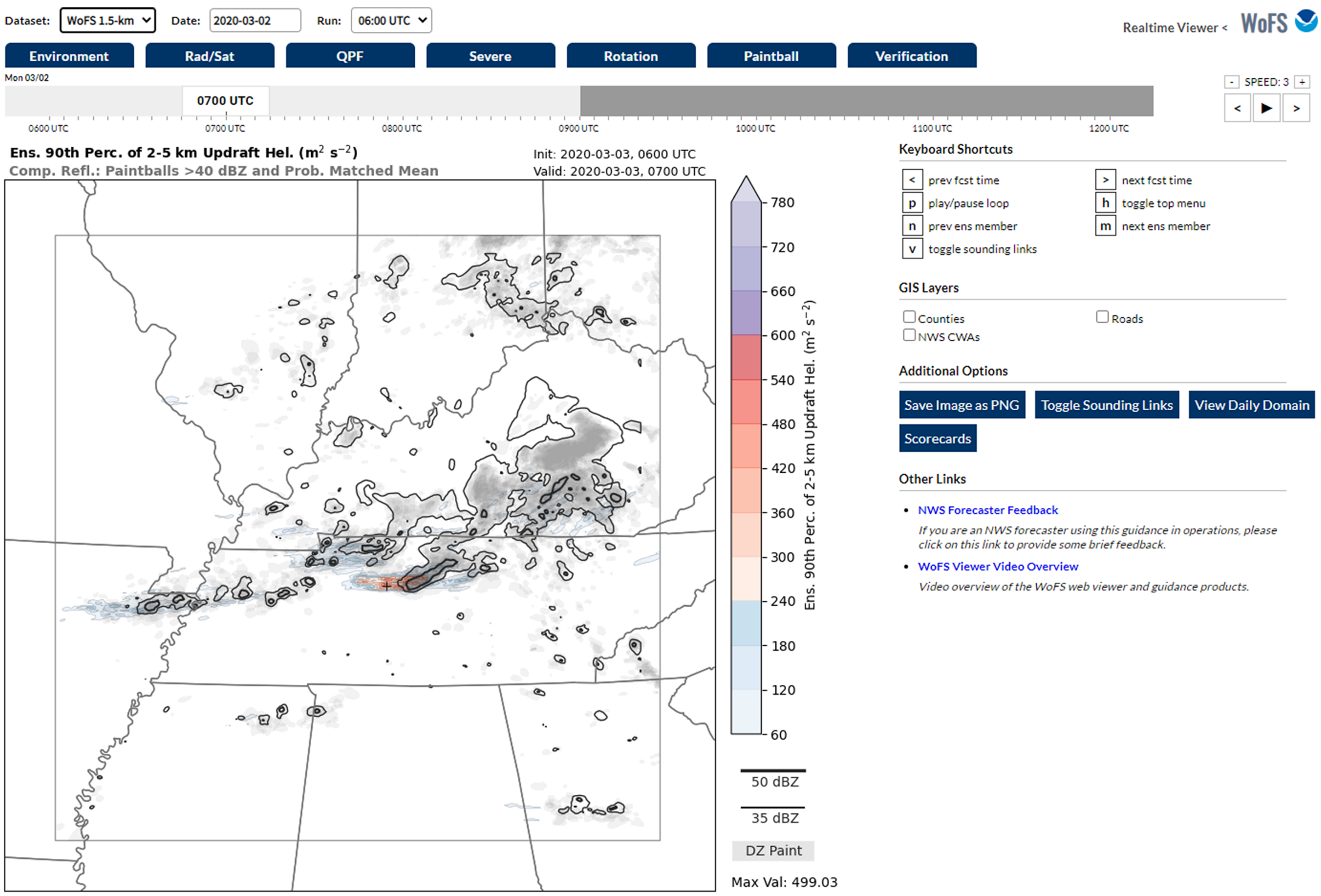

Fig. 3. Warn-on-Forecast System (WoFS) display of the ensemble 90th percentile of 2–5-km updraft helicity, with composite reflectivity paintballs > 40 dB*Z* and probability matched mean, initialized 0600 UTC 3 Mar 2020, valid 0700 UTC 3 Mar 2020. An EF3 tornado occurred in the Nashville, Tennessee, area. (Figure courtesy of the National Severe Storms Laboratory.)

which may result from, for example, an extreme event that has little to no representation in the machine learning model's training dataset. This recognition requires humans to maintain expert skills in mesoanalysis to conduct this quality control and identify "targets of opportunity" to improve upon model forecasts, most likely near- and short-term forecasts.

Automated tools that improve communication. Murphy (1993) stated that "Forecasts possess no intrinsic value. They acquire value through their ability to influence the decisions made by users of the forecasts." In other words, the forecast itself cannot stand alone; it must be followed by the creation and presentation of understandable and actionable weather information to customers, partners, and the public. Broadcast meteorologists, along with the weather companies that provide graphics support for broadcasters, have been leaders in developing technology and tools to facilitate and optimize weather communication. Methods of delivering forecasts and warnings to the end users are becoming increasingly automated, with applications for smart devices (e.g., phones, watches, Google Assistant, Amazon Alexa), social media, and streaming services growing in number. In one recent advance, The Weather Channel uses a suite of tools called Unreal Engine to create realistic visualizations, termed immersive mixed reality, which places weather anchors virtually within simulated weather hazards to help viewers better understand weather threats (Fig. 4) and their risks (Barrett 2018).

Other research is being conducted on improving the interpretation of severe-weather text and graphics on social media. Sutton and Fischer (2021) conducted a study in which they tracked the study group members' eye movements on National Weather Service Twitter and

Fig. 4. Immersed mixed reality used by The Weather Channel. Meteorologist Chris Warren is showing how different depths of water can affect safety in vehicles and personal safety. (Figure courtesy of The Weather Company/Channel, available at www.youtube.com/watch?v=12nlTn4G5Xg.)

Facebook severe weather posts to determine level of interest and comprehension of severe weather information. This type of research is becoming increasingly important to determine how users view and interpret weather graphics and text, with the goal of helping operational forecasters create more effective messaging to improve users' decision-making in the face of dangerous weather.

Convection-allowing models are also assisting operational meteorologists as a weather information communication tool. NWS forecast offices include simulated reflectivity and other products from convection-allowing models in briefings to partners and the public, as a way of showing the expected timing and pattern of precipitation. Broadcast meteorologists use model output similarly, including merging real-time radar information with this high-resolution model output to provide details on expected storm movement and behavior. Tools like these are used with geographic information system (GIS) maps to provide spatial detail and to help orient users and improve interpretation of weather information.

An AI-rooted communication tool, Probabilistic Winter Precipitation Forecast (PWPF), is helping the NWS convey winter-weather threats by combining deterministic and ensemble guidance to produce probabilities for snow and ice. After consulting additional sources of guidance such as empirical orthogonal functions (EOF) and ensemble cluster analyses (Zheng et al. 2019), operational meteorologists complete their deterministic forecasts of snow and ice accumulation. They use a set of tools that alter the probabilities of snow or ice, along with exceedance probabilities (e.g., Novak et al. 2014a; Waldstreicher et al. 2018). Maps and tables generated from this output include the probability of exceedance of various thresholds, as well as the 10th and 90th percentile accumulations (Fig. 5), which are given the titles of "expect at least this much" and "worst case scenario," respectively. Providing these ranges helps communicate uncertainty and confidence to NWS partners, helping them with their decision-making and preparations. In the future, these forecasts are expected to include alternative scenarios and probabilities for different storm tracks.

Finally, to foster improved public safety and decision-making, it is crucial to provide weather information in multiple languages. A recent project at the National Hurricane Center is using machine learning algorithms to provide on-the-fly English-to-Spanish translations of NWS free-text products (M. Bozeman 2021, personal communication). One major challenge that the group has faced includes training the algorithm to interpret slight variations in language,

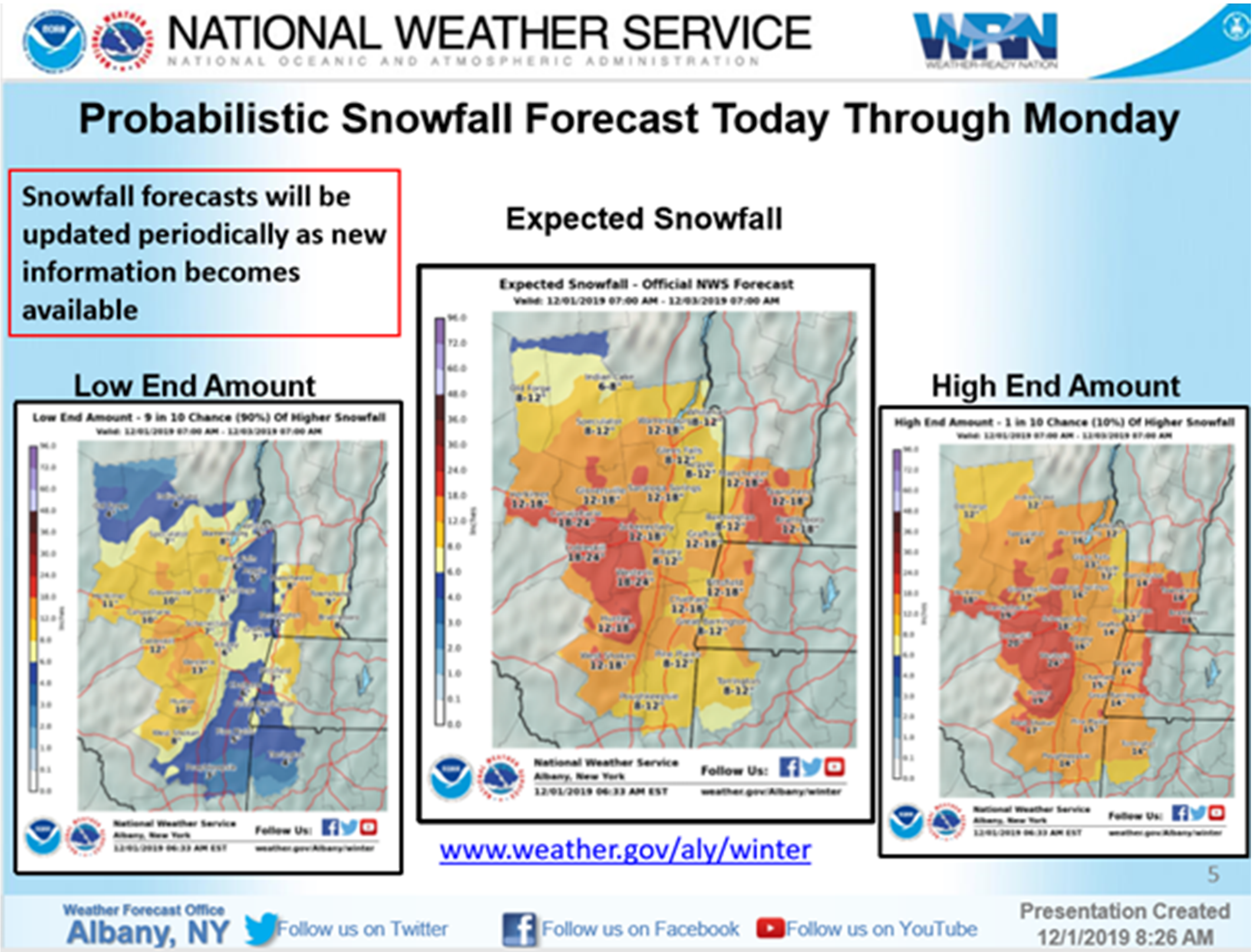

Fig. 5. Part of an impact decision support services briefing issued by the National Weather Service in Albany, New York, prior to the 1–3 Dec 2019 snowstorm. The 10th percentile, 90th percentile, and most likely scenarios are displayed with the disclaimer that snowfall forecasts will be updated as new information is available. (Figure courtesy of the National Weather Service in Albany, New York.)

despite numerous regional dialects of Spanish. Although clearly a time saver in the future, there is a tremendous amount of work to develop a satisfactory, robust, and accurate product. Thus, this project shows the importance of the human–machine mix in developing new tools for communicating forecasts in multiple languages due to the increasingly diverse populations in many communities around the world.

Training and proficiency for future forecasting

Training and talent development must change to meet the demands of an evolving forecast process. Current faculty are likely to be teaching content for which they have never received formal training. Students that entered universities in 2020 will complete their careers around 2060—Is anyone capable of knowing what suite of skills will be needed this far in the future? Consequently, universities must deliver a balance of sufficiently specialized training in meteorology (or atmospheric, Earth, or geosciences, more generally) and a broader university education. Consider that the U.S. Government's Meteorologist General Schedule 1340 (GS-1340) requirements are 30 years old. These standards describe a homogeneous workforce (www.weather.gov/jetstream/careers) in the face of the need for diverse skills (Environment and

Climate Change Canada has similar outdated requirements). At some institutions, the GS-1340 position description can approximate a kind of accreditation, but it can also be viewed as an overly rigid baseline of standards. The AMS standards for a bachelor's degree in atmospheric science are similar (www.ametsoc.org/ams/index.cfm/about-ams/ams-statements/statements-of-the-ams-in-force/bachelor-s-degree-in-atmospheric-science/), though they are more flexible and have been updated more recently. In contrast, other sectors of the government including the Air Force have recently revised their commissioned officer accession guidelines, adding physicists, mathematicians, and other geoscience specialties to the list of qualified applicants. Hence, NOAA/NWS should revise the standards offered by the NWS to promote greater flexibility and expansion of skills into the future workplace.

Another reason for flexibility in education is the expansion of the private sector in meteorology. Because some academic meteorology programs have historically been oriented toward educating operational meteorologists, particularly for government service, universities should question whether their programs are fit for this new era. Some of the obstacles toward reinvention of curricula include the amount of material and experiences that students need to possess versus the amount of availability in a typical undergraduate curriculum. Consequently, the idea of curricular tracks or concentrations has been implemented at some institutions, allowing students to focus on particular topics or add courses focusing on related skills such as statistics or GIS. Extracurricular experiences that further broaden students, such as paid and unpaid internships, independent studies, work study programs (e.g., Roebber et al. 2010), and research experiences for undergraduates (e.g., Gonzalez-Espada and Zaras 2006; Burnett and LaDue 2011) can be valuable and should be made available to interested students.

It is incumbent on the public and private sectors to get more directly involved with universities, and vice versa, to make such opportunities available to students, who will be their future employees. But university education is more than just preparing a student for job placement; ideally, university education teaches students how to learn and think critically. The employer will also be required to deliver a certain amount of on-the-job training at various stages of their employment, perhaps as continuing professional development—another reason for employers and universities to be discussing how best to partition education and training. Below, we identify particular skills needed for students and operational meteorologists to succeed in the future atmospheric-science workforce.

1) *Computing and data skills*. Many curricula require at least one course in computing/programming, but this may not provide students with sufficient skills to understand NWP, AI, and other techniques. Although the nature of programming has evolved from Fortran and C+ to higher-level languages such as Python and R, coding and other skills (e.g., web-page design, GIS) are needed by meteorologists to analyze, visualize, forecast, and communicate. For example, GIS courses and certificates are increasingly being sought by students in lieu of minors in more traditional disciplines such as mathematics, physics, or computer science. Specifically, GIS and other visualization skills can be valuable in graduate school, the NWS, television operations, self-employment, and private industry, as well as outside of meteorology.
2) *Probability, statistics, and ensemble forecasting*. Research reveals the advantages of a probabilistic approach to forecasting, with the use of ensemble forecasts being a key application (e.g., Murphy 1993; Fritsch et al. 1998). Students must have a foundational understanding of ensembles, including their construction, interpretation of output, the importance of probabilistic reliability, data-visualization techniques, and communication to an end user. Institutions are increasingly incorporating ensemble techniques and products within their synoptic meteorology curriculum and weather briefings. The foundation for ensemble techniques is rooted in probability and statistics, which can also carry over to

careers in other sectors of the weather, water, and climate enterprise and weather-adjacent fields (e.g., risk assessment, insurance/reinsurance, and finance).

3) *Artificial intelligence*. AI methods, as well as their subsets machine learning and deep learning (see the "Definitions" sidebar), can be viewed as leverage for a forecaster. One does not need an NWP model to make weather forecasts, but through its use, one can advance understanding and forecast skill. Similarly, postprocessing of NWP with AI methods can enhance those products (e.g., by revealing and correcting model biases). These methods can also be used as tools to quickly explore the sensitivity of a forecast to atmospheric parameter adjustments and their effects on a range of future atmospheric conditions. Even if students do not necessarily implement AI methods themselves, students need to know how they are built, what their strengths and weaknesses are, and ethical considerations when implementing and using AI methods (McGovern et al. 2022).
4) *Decision support and communication skills*. One of the most important value-added skills is the ability to communicate effectively with users of weather information. This skill set involves two interrelated concepts. The first is the ability to communicate in a variety of ways (e.g., oral, written, graphical) and within a range of personality types, audiences, and professional settings. The second is the understanding of how people, including operational meteorologists, make decisions. What kind of information is needed, and how do human thought processes work? Decision support may be the most important role that operational meteorologists of the future will undertake: taking the forecast product and crafting the best message to be delivered at the right time to prompt the desired response. Related topics of study include societal response to forecasts, psychology of decision-making and risk assessment, and the application of statistics/probability.
5) *Research skills and interdisciplinarity*. As models improve, operational meteorologists will need to understand when automated tools work well and when their output needs to be modified. To do this, some research skills will be beneficial (e.g., Steeneveld and Vilà-Guerau de Arellano 2019). Specifically, the research should start with the problem identification. Moreover, research should improve the *value* of forecasts, not just their *quality*. Such research will increasingly be interdisciplinary and multidisciplinary, and it may require functional knowledge of non-physical-science fields, including social science.
6) *Effective group work*. Collaboration will be required to tackle increasingly challenging and multidisciplinary problems. Diverse teams composed of individuals of varying perspectives and heuristic techniques often outperform teams of high-ability problem solvers (e.g., Hong and Page 2004). Schools can conduct group exercises that require students to work together to produce an accurate and consistent forecast and/or message through consensus and working together through the scientific process. Students with expertise in physical science, social science, communications, GIS, and graphical design can all contribute to the forecast and communication process, reinforcing the importance of diverse teams in solving multidisciplinary problems.
7) *Management of work–life balance with rotating or nonstandard hours*. Although not a "skill" per se, operational meteorology can be a demanding job, sometimes much more than we acknowledge. Weather happens all the time, and providing 24-h support requires nonstandard work hours, which can affect physical/mental health (Costa 2010) and decision-making abilities. Add in working from home (i.e., teleworking) as a more frequent and attractive option for some, and such constraints require a holistic perspective of both the employer to build a skilled and effective team and the employee to maintain work–life balance. Students considering such careers would benefit from job shadowing or internships to evaluate the demands of changing sleep patterns, work, physical/mental health, and social/family life.

The above set of skills should be considered in any revision to AMS guidelines and/or the GS-1340 position description. Atmospheric science students will graduate into a broad range of potential career trajectories, so universities need to ensure student competitiveness in the short-term job market. In addition, operational meteorologists need continuing professional development and self-directed learning (e.g., LaDue and Cohen 2018) to keep abreast of new science and applications. Educators are encouraged to provide a solid base of fundamentals, along with enough breadth, to allow versatility. Thus, we make the following recommendations.

- The public sector, private sector, and professional organizations, such as the AMS and National Weather Association, work with universities to revisit curriculum requirements perhaps on a decadal or subdecadal interval.
- Students/operational meteorologists take the initiative to gain skills (and sharpen existing skills). Job shadowing, internships, mentoring, career fairs, professional conferences, and networking all help to prepare for diverse careers. Mentoring should be pursued. These activities must be facilitated by university programs in ways that maintain diversity.
- Students/operational meteorologists seek opportunities to distinguish themselves by pursuing their primary interests and skills at which they excel. Examples include double majors, advanced degrees, and developing leadership skills through volunteering, university clubs, or other opportunities. Recognizing that this places additional financial burdens on students, resources must be provided.
- Students training to be operational meteorologists spend more time at university doing actual forecasting tasks, as well as training in positions outside of traditional forecasting routes, with some specializations. Future positions will likely be centered on high-impact weather, communication, and forecast tool/automated forecast development.
- Operational meteorologists engaging in ongoing training to improve themselves require verification and feedback. Individual feedback should be provided in a private, constructive, nonjudgmental manner without professional repercussions. Aggregated, anonymous verification is also important for a forecast office. Verification can help determine where humans are adding value, where they are not, and how to add value, especially for high-impact events.
- Trainers, developers, and research-to-operations-to-research (R2O2R) personnel sit with operational meteorologists at the forecast desk, both for the development side to see what they need and how they can use new tools, but also to help them best use these new tools.

Students/operational meteorologists will have to take initiative and be creative with their education, self-directed learning, and continuing professional development. Correspondingly, universities and employers should help by developing and publicizing opportunities for all students—regardless of background, financial situation, or other circumstances—to tailor and expand their skill sets to be as prepared as possible for the continually evolving weather enterprise. However, universities and employers also hold some responsibility in making sure all students or operational meteorologists have these opportunities and should not assume that students who are not taking advantage of these opportunities lack initiative.

Envisioning the future of weather prediction and communication

In this section, we lay out our vision for the future of weather prediction and communication, focusing on the role of the operational meteorologists in the process. This vision is based on the webinars, panel discussions, and other in-person communications, as well as numerous advancements and developments in weather forecasting and information technology already discussed in the previous sections. We also extrapolate existing trends into the future, where possible. Our vision is organized into periods before the forecast, while making the forecast, communicating the forecast, and after the forecast.

Before the forecast. Testbeds and proving grounds focused on operational forecasting environments will remain a vital part of our vision for the future. Testbeds allow for input into products by end users before the transition to operations (e.g., Sørensen and Torfing 2011). NOAA operates several testbeds (NOAA 2020b) such as the Weather Prediction Center Hydrometeorology Testbed and the NWS Operations Proving Ground (OPG), which evaluate experimental guidance products for winter weather (www.wpc.ncep.noaa.gov/hmt/11th_Annual_HMT_WWE_2021_Final_Report.pdf), severe weather (www.nssl.noaa.gov/news/factsheets/hwt.pdf), and extreme precipitation (https://vlab.noaa.gov/documents/214451/5286029/CFP+%28QPF%29+Virtual+Experiment+Report.pdf/0a42a6de-4254-c3df-ccf3-8cb9f7bd3583?t=1605800868969). Testbeds also allow operational meteorologists and researchers to evaluate products extracted from experimental sources of data and guidance to determine which of these derived fields might be helpful in operations. Examples of products that have come from these and other testbeds include WoFS, PHI, rapidly updating phased-array radar (Heinselman et al. 2015; Wilson et al. 2017a), convection-allowing models and ensembles (e.g., Gallo et al. 2017; Demuth et al. 2020), and flash-flood prediction products (e.g., Martinaitis et al. 2017; Yussouf et al. 2020). Experimental products are also evaluated in a field office environment, acting as a local or regional testbed with local and regional user groups and media. For example, probabilities for intense hourly snowfall (Fig. 6) could help predict impacts of snowfall rates on transportation interests, school superintendents, and other winter weather-sensitive user groups, products that may become operational in the future.

Another benefit of testbeds is that they bring various partners together to work on common problems, which is the essence of R2O2R. When partners are allowed to take part in

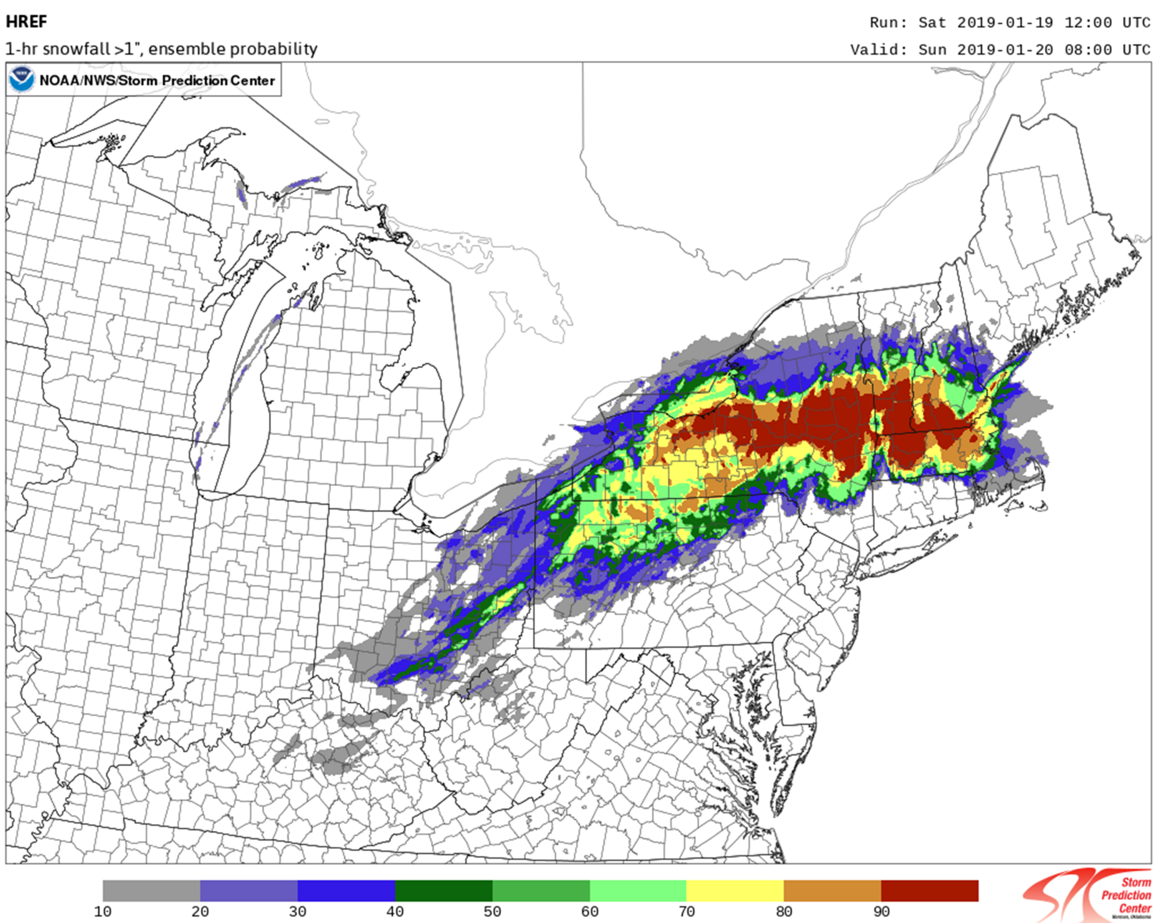

Fig. 6. High Resolution Ensemble Forecast (HREF) 20-h forecast probability for 1-h snowfall > 1 in. initialized 1200 UTC 19 Jan 2019 and valid 0800 UTC 20 Jan 2019. (Figure courtesy of the Storm Prediction Center.)

the creation process, the likelihood of future adoption and use of a product both increase (e.g., Nygren et al. 2018). As a result of coproduction, the designs of many products continue to mature (e.g., Kuster et al. 2017; Calhoun et al. 2018; Meyer et al. 2019; Klockow-McClain et al. 2020; Obermeier et al. 2020). Similarly, managers and policy-makers should periodically shadow operational meteorologists in a field office or forecast floor, similar to operational meteorologists evaluating experimental guidance at testbeds. The policy-makers can then have direct experience and make more informed funding and policy decisions to improve the forecast and communication process.

More broadly, these adaptations and benefits will depend on a multisector interdependence in a R2O2R framework, including social-science fields such as human factors and communication. One illustration of that framework is the concept known as Forecasting a Continuum of Environmental Threats (FACETS; Rothfusz et al. 2018). FACETS emphasizes the need for social, behavioral, and economic sciences research in probabilistic hazardous-weather forecasting and is evident in recent research with NWS operational meteorologists, core partners, and the public during HWT experiments.

Ultimately, however, testbeds are about the forecast process and the operational meteorologist. Thus, testbeds should continue to consider the human side of operational meteorologists—the impact of the mental and physical workload of the operational environment on the human. Workload has already been examined in NOAA testbeds (Wilson et al. 2017a) and is projected to increase as new weather data and forecasting tools become available. Testbeds can consider approaches to mitigating potential workload impacts of new developments prior to their operational deployment, to help protect the future performance and well-being of operational meteorologists. Furthermore, experiences such as the Mesoanalyst Workshop at the NWS OPG (Runk et al. 2020) help to designate forecaster responsibilities as new analysis, prediction, and communication techniques force the forecast process to evolve. As a result, the impact of new forecast approaches on the current forecast process, particularly optimizing the balance between automation and human intervention in the prediction and communication process, can be assessed.

Such concerns go beyond time allocation, but also mental workload, which is the amount of attention resources required to meet a performance criterion. Mental workload is influenced by operational meteorologists' task demands and their past experience. For example, in an experiment assessing operational meteorologists' use of rapidly updating phased-array radar data, Wilson et al. (2017a) applied the Instantaneous Self-Assessment Tool to measure cognitive workload. In follow-up focus groups (Wilson et al. 2017b), these operational meteorologists were then able to suggest strategies to help minimize instances of cognitive overload (e.g., using algorithms to track trends in radar signatures, using more efficient visualization techniques, redistributing forecaster responsibilities within offices).

We also envision that testbeds could examine the composition of the forecast teams. Alternative staffing models could be proposed where high-impact weather events require a team of expert operational meteorologists with various strengths to sort through the multitude of forecast and warning information and present specific information that meets customers' needs. For example, one forecaster might excel at mesoanalysis and verbal communication, whereas another might be better at radar and NWP model analysis and graphic production. These members could fill multiple critical roles within an effective warning, forecast, and communication team. Event-driven team building would seek to optimize the forecast, warning, and communications process for each event.

Finally, while end-user inclusion at the development and testing stages of future forecast tools is imperative, consideration of the needs of the public is equally important. The future forecast process will result in more publicly available probabilistic forecasts. Therefore, the extent of understandability and usability of these forecasts must be evaluated. Past studies indicate that forecasts that include uncertainty may improve the decision-making of laypeople (e.g., Joslyn and LeClerc 2012), and multiple entities have recommended that future forecast products contain uncertainty (e.g., Rothfusz et al. 2018). In fact, the public prefers forecasts that include an expression of uncertainty over deterministic forecasts (e.g., Morss et al. 2008; Peachey et al. 2013; Demnitz and Joslyn 2020). Although numerical probabilities are a good starting point (Gulacsik 2019), new methods will likely be needed because even operational meteorologists have stated challenges interpreting probabilities and probabilistic guidance (Demuth et al. 2020). Simple changes in color-coding scheme, map display, layering, or type of information (e.g., qualitative versus quantitative) can have an impact on how the public may perceive, interpret, and use a forecast (e.g., Sherman-Morris et al. 2015; Miran et al. 2017; Klockow-McClain et al. 2019; Meyer et al. 2019).

Making the forecast. In the future, we still anticipate that a fundamental understanding of the atmosphere based on operational experience will continue to be needed by operational meteorologists, including the diagnostic expertise at applying past weather events to evaluate forecast guidance (e.g., Novak et al. 2008; Daipha 2012, 2015; Hoffman et al. 2017). Operational meteorologists' abilities to recognize synoptic patterns will be needed to identify errors at all stages in the automation process (e.g., Bosart 2003), correct them, and adjust them based upon a variety of factors (e.g., input from calibrated probabilistic forecast systems, recognizing similar atmospheric conditions in past events). Such abilities are a significant benefit that humans can provide in the forecast process (e.g., Doswell 2004; Stuart et al. 2007; Hoffman and Fiore 2007; Novak et al. 2014b; Hoffman et al. 2017). For example, someone that uses conceptual models based on learning from past events could determine which solution, cluster of solutions, or probability distribution within an ensemble represents the most likely outcome or range of outcomes.

However, operational meteorologists will need to shift paradigms from a binary decision of warn–not-warn to a continuous flow of information (Rothfusz et al. 2018). Downstream severe weather threats are likely to be depicted as plumes of probabilities on spatial scales of less than 10 km and time scales of minutes (e.g., Fig. 8 of Karstens et al. 2018; Stumpf and Gerard 2021), derived from convection-allowing models or ensembles. Operational meteorologists would still need to perform quality control on these predictions, because even convective allowing models improved by AI would still have errors due to subgrid convective processes in the atmosphere. Alerting and messaging would be determined by the changing probabilities at locations within the polygons and plumes, based on the motion and evolution of the thunderstorms. For example, people in a location in the path of a possible tornado could be made aware of thunderstorms upstream hours prior to impact, to encourage people to be prepared to take action should the thunderstorms become tornadic. Information on the location and severity of the threat would be updated frequently until the tornado is within 15–45 min of impact, when the probability threshold for a warning is met. The warning would consist of unique visual and/or audio alerts with specific graphical, audio or text information available on all media platforms. The continuous-flow-of-information paradigm is not confined to severe weather: A winter weather version is currently being tested in some regions of the National Weather Service (Graham et al. 2021).

Furthermore, we anticipate that advances in big-data analytics provide an opportunity for AI to provide meaningful feedback to the operational meteorologist on their decisions relative to the degree of uncertainty in a forecast problem. This "virtual assistant" would analyze tendencies relative to results and verification. AI could also provide quantitative feedback to an individual, researchers in human cognition, and forecast-process managers to guide improvement. AI systems will always contain errors, including some that a human would be unlikely to make (see the "Misconceptions about the human–computer relationship"

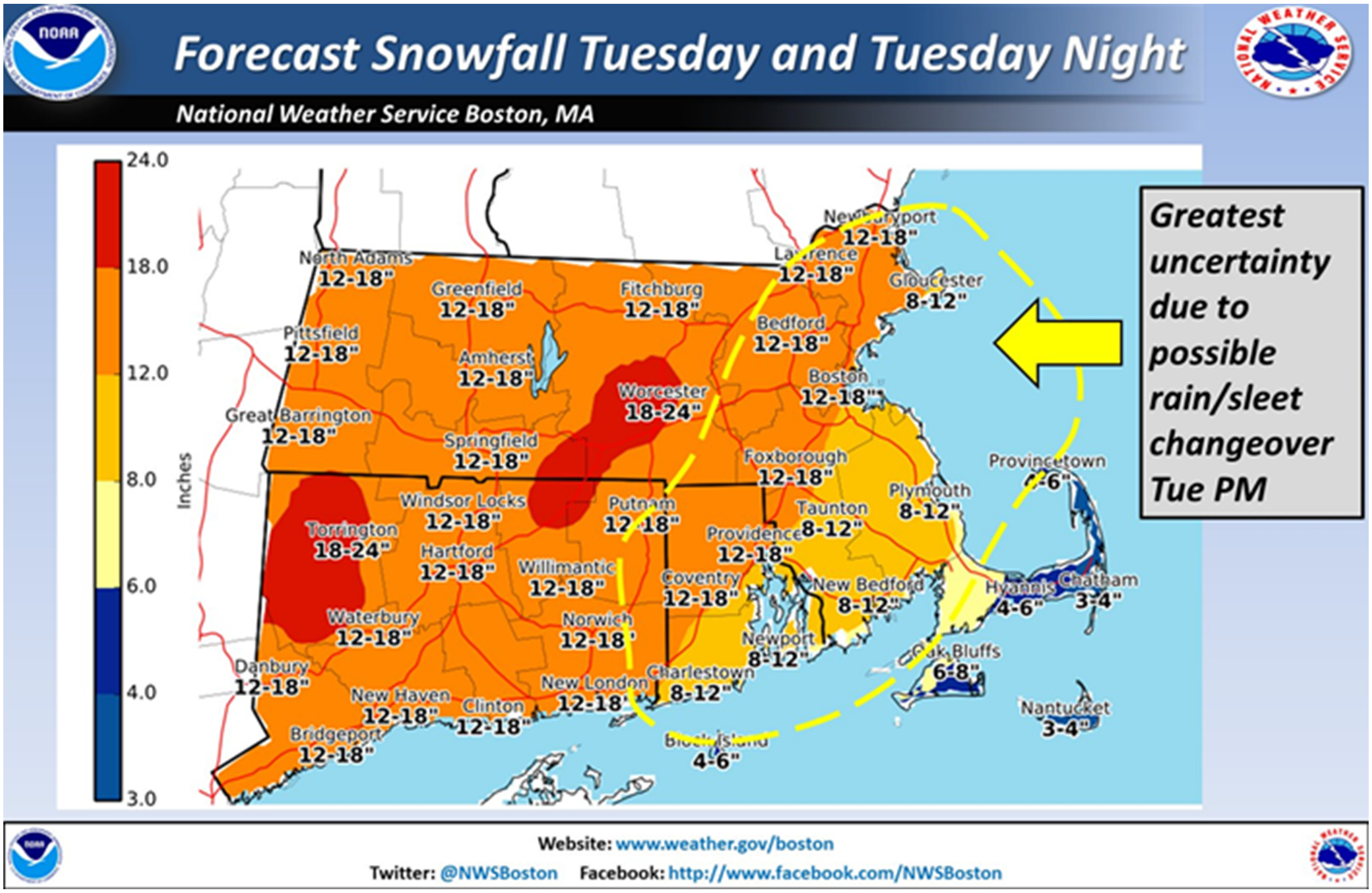

Fig. 7. Part of an impact decision support services briefing issued by the NWS in Norton/Boston, Massachusetts, explaining sources of uncertainty in a region where rain and sleet could significantly affect snow amounts prior to a major snowstorm. (Figure courtesy of the National Weather Service in Boston, Massachusetts.)

sidebar). In situations where an anticipated forecast event lies outside of the AI training dataset, human intervention may be required. Because of these issues of appropriate trust and reliance, forecaster–computer relations should improve as operational meteorologists learn the strengths and weaknesses of ever-improving AI systems. New data sources such as drones and unmanned autonomous vehicles will need to be quality controlled. Recognizing erroneous or anomalous data will continue to be an important step prior to and within the forecast process.

For example, variables contributing to a machine learning application can be listed along with the weights of each variable in different forecast scenarios. This is being done in some R2O settings already (K. Corbosiero and B. Filipiak 2021, personal communication). Operational meteorologists can learn what variables and weights contribute to the optimized forecasts, plus identify when observed atmospheric variables are outside of the machine learning database, to know when to deviate from the machine learning–based forecasts.

We envision that the private sector will continue to grow in number of employees and in importance to the global economy. They will continue to innovate, producing specialized forecast and warning information, including for the many industries that can be impacted by adverse weather. Although some private-sector groups are changing the role of humans in the forecast process (e.g., Rose et al. 2015), other private-sector groups are seeing demand for other applications of improved meteorological data outside the weather enterprise. Humans have a unique role in conveying and explaining weather threats and the sometimes-nuanced role that weather can play within industry operations. Weather and climate predictions for commodities/derivatives, risk management (e.g., Brockett et al. 2005), and reinsurance (e.g., Murnane 2004) have expanded in recent years, and the human role in the application of weather information into weather-adjacent industries will continue to expand.

Communicating the forecast. In the future, we anticipate that communication of the forecasts will also change. On the forecast floor, operational meteorologists can add value to automated probabilistic output by explaining to users that this presents a range of reasonably possible solutions, along with the most likely forecast. Precipitation transition zones in the winter provide a particular challenge in communicating uncertainty. Figure 7 is from an impact-based decision support services briefing slide issued by the NWS Forecast Office in Norton/Boston, Massachusetts. This image shows the greatest uncertainty in snow totals by describing the potential for reduced snowfall amounts with rain and sleet due to the uncertain evolution of a precipitation transition zone.

Communication of the forecast must be done in a manner that can be easily understood and which spurs appropriate actions. Figure 8 shows a social media post describing a warning for a deadly heat wave in Southern California. Included is a map showing color-coded threat levels across the warning area, along with a description of what each color means and protective actions that should be taken. Social-science researchers have documented the varied ways in which people interpret presentations of weather information (e.g., Morss et al. 2008; Peachey et al. 2013; Sherman-Morris et al. 2015; Miran et al. 2017), and this necessitates that operational forecasters understand the varied needs of their customers and communicate in a clear and effective way. The NWS Evolve initiative focuses on the importance of connecting weather information to decision-makers through impact-based decision support services

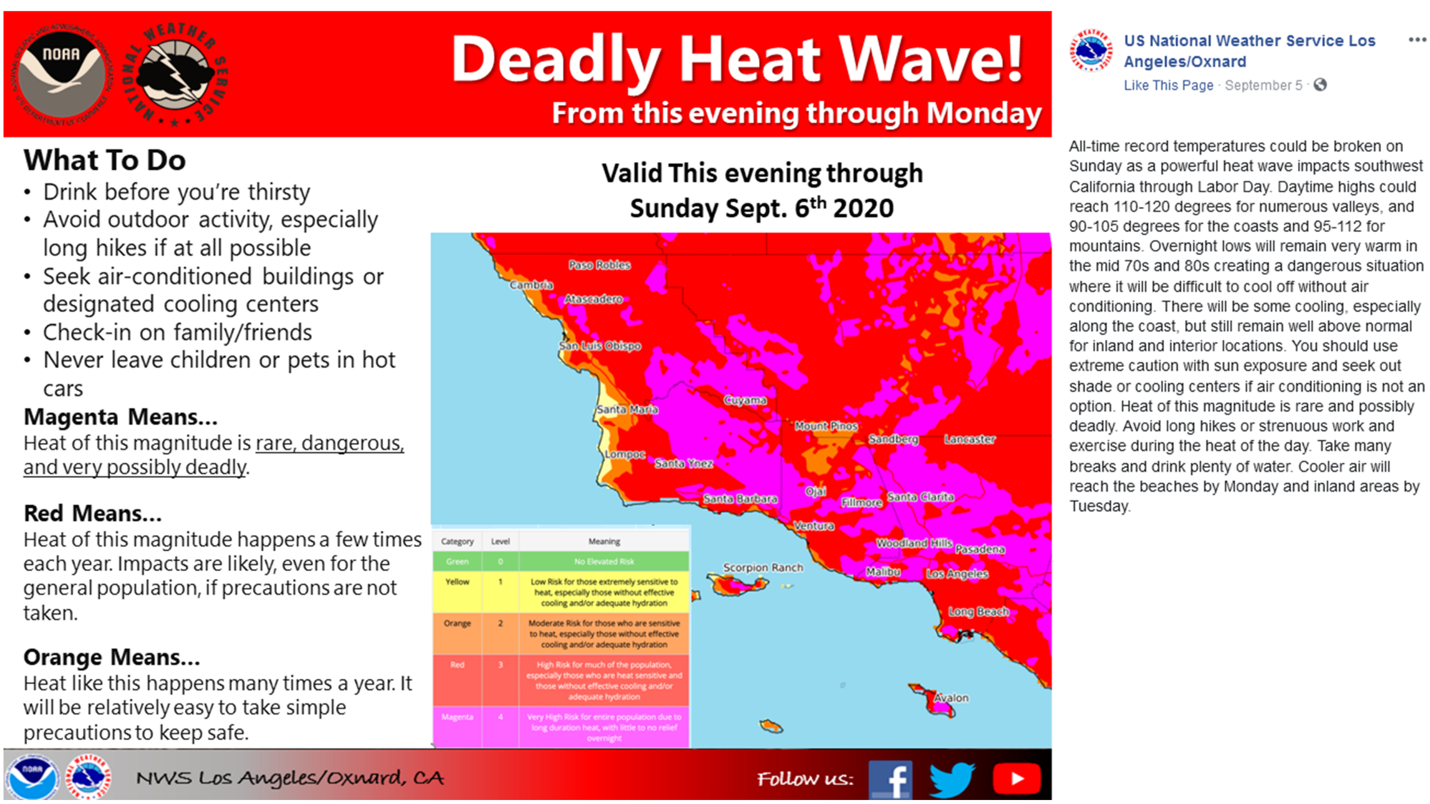

Fig. 8. **An impact decision support services post to social media followers, issued by the National Weather Service in Los Angeles/Oxnard, California, on 5 Sep 2020 describing a deadly heat wave in Southern California with a graphical depiction of color-coded threat levels and recommended actions. (Figure courtesy of the National Weather Service in Los Angeles/Oxnard, California.)**

(Uccellini and Ten Hoeve 2019), underscoring the need to tailor weather information and messaging to the needs of the users. Such forecast messaging must also be adapted to smart devices, social media, streaming services, virtual/immersed reality, and the ever-changing ways people receive the weather information, as use of these avenues for forecast information continues to increase.

Innovations in virtual reality and immersed mixed reality (as demonstrated in Fig. 4) will increasingly aid in communicating weather risk and impacts (e.g., Crout 2019). These innovations could become important public education tools, including virtual-reality video games to teach people how to prepare for various weather-related hazards by placing them in and around the virtual hazard (Bernhardt et al. 2019, 2020; Zhu and Li 2021).

The future forecast and communication process will also inform a greater range of decision makers. Currently, broadcast meteorologists communicate NWS forecasts and warnings, and they continue to be the leading source of tornado warning information for the public (Silva et al. 2018). Involving the end users in the process of developing new systems and procedures has been a cardinal rule in cognitive systems engineering for decades (e.g., Woods and Hollnagel 2006). Similarly, research activities within the HWT and NWS OPG need to involve NWS core partners, emergency managers, and broadcast meteorologists (Nemunaitis-Berry et al. 2020) in the development and testing cycles of future forecast tools and products (e.g., Fig. 9).

After the forecast. Even after the forecast is made, we anticipate changes in the future. Forecaster performance (e.g., decision accuracy, lead time) will be routinely evaluated using a sophisticated verification system. Performance could be measured in a controlled situation, such as a testbed that uses 360 feedback. For example, Karstens et al. (2018) showed how forecaster feedback during probabilistic hazard information experiments can be incorporated to improve the presentation and reliability of severe weather automated probabilistic guidance,

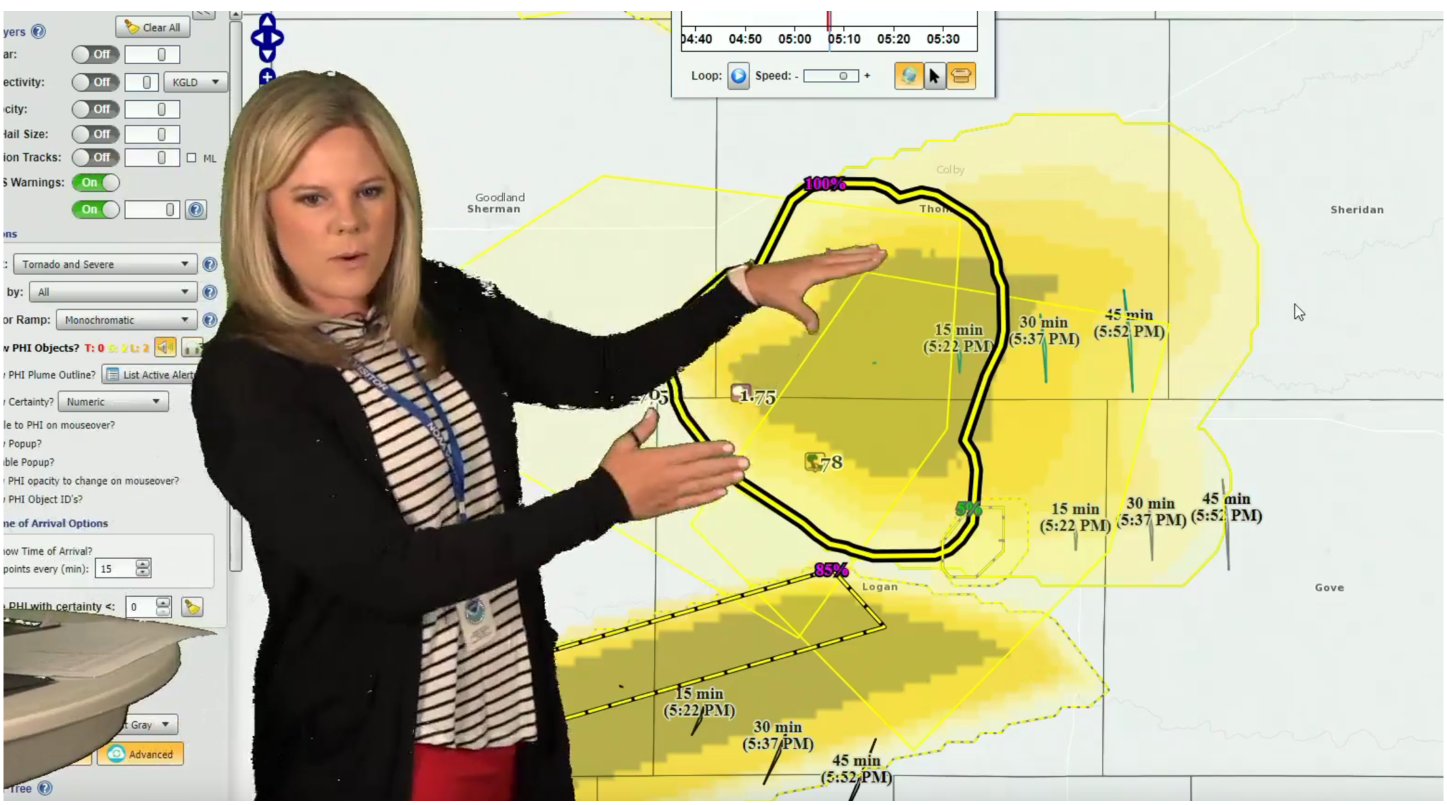

Fig. 9. Broadcast meteorologist Katy Morgan at the Hazardous Weather Testbed evaluating experimental methods of broadcasting severe weather probabilities in a continuous flow of information. [Figure courtesy of the National Severe Storms Laboratory Probabilistic Hazards Information (PHI) Project.]

thus resulting in increased acceptance and use of the new tool. Providing information about how the experimental automated guidance has been developed can help ensure the success of automation use within the future forecast process.

Furthermore, these new forecast and warning paradigms will require new and creative methods of verification. Cost/benefit analyses and verification can illuminate areas where humans are adding value. More analysis of the value humans add to deterministic and probabilistic forecast products is needed and can inform budgets in all sectors. Verification will also determine where operational meteorologists' efforts need to be directed to improve the forecasting process. The Model Evaluation Tools (MET) package developed by NCAR is an object-oriented verification program to evaluate forecast model and grid-to-grid verification (e.g., Clark et al. 2014; Brown et al. 2021). Usage of MET is expanding and could become the standard verification method in the NWS, as well as in other sectors of the enterprise. Some local NWS offices are performing internal verification of predictions of specific weather phenomena in a GIS environment. For example, Fig. 10 shows a GIS graphical display of the difference in snow predictions and observed snowfall over the NWS Albany, New York, forecast area during 1–3 December 2019. This GIS analysis calculates the difference between gridded snowfall forecasts at a 2.5-km grid spacing and observed snowfall. Observed snowfall determined by the National Operational Hydrologic Remote Sensing System (NOHRSC) is supplemented by snowfall values received from observers, and inverse distance-weighting interpolation is used. These types of analyses provide verification information that can be used as input into research to improve future predictions and to determine whether the forecasts addressed the most important impacts from an event.

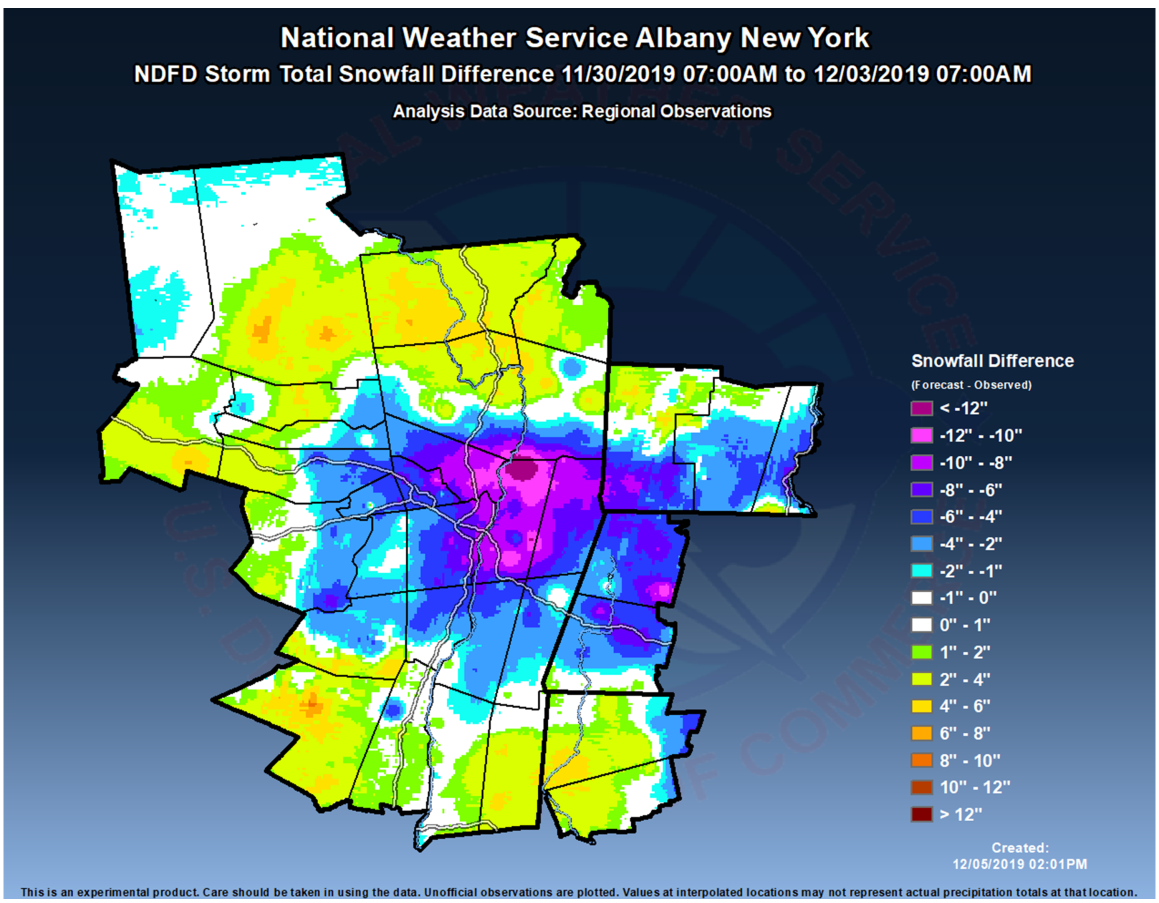

Fig. 10. GIS analysis of the difference between forecasted snowfall in the National Weather Service National Digital Forecast Database and observed in the NWS Albany, New York, forecast area initialized at 1200 UTC 30 Nov 2019 and valid 1200 UTC 3 Dec 2019. (Figure courtesy of The National Weather Service in Albany, New York.)

To summarize, we have laid out our vision of the future forecasting enterprise, showing that humans will maintain an essential role in weather prediction and communication. Although this paper presents a snapshot of how the human role in forecasts and warnings is changing and where it is likely to go, how this vision evolves will require regular reevaluation, requiring input and collaboration from across the weather, water, and climate enterprise and beyond. R2O2R must be embraced by all sectors of the enterprise, and "stove-piping" (i.e., where research is conducted in relative isolation with limited consideration or coordination with parallel efforts) must be eliminated (e.g., McNeeley et al. 2012). Our enterprise must seek out experts in the physical, social, and cognitive sciences for every research initiative, coordinating and collaborating in a shared environment.

In fact, the greatest future advancements assisted by AI technology may be through collaborations with weather-adjacent or weather-dependent industries. Creative and futuristic visionaries could develop algorithms that will automatically order more winter sleds for a hardware store based on a cold and wet 2-week outlook. Perhaps a smart device will assess the forecast and automatically select your clothes for the day. Streaming services may be fully equipped to communicate weather alerts for the viewer's area. Roadside salt spreaders might automatically activate at the first forecast of wintry weather. Ethics, of course, must be considered as new technologies are employed to produce new products and services.

Conclusions

Weather forecasts are not perfect, and the weather, water, and climate enterprise has been tremendously successful at demonstrating consistent improvement in forecasts for many years. From empirical approaches in the first half of the twentieth century to numerical weather prediction models, satellites, radar, workstations, nowcasting, ensembles, data assimilation, and rapid updating models, advances in meteorology have been intricately entwined with technology harnessed for overwhelmingly positive benefit. Yet experts keep stating the misconception that technology will result in increased automation, displacing human workers, overdependence on technology, unintended consequences of the technology, and the threat of losing control of that technology (see the "Misconceptions about the human–computer relationship" sidebar). Furthermore, there is the misuse, misunderstanding, and mistrust of AI technology and automation (Hoffman 2017).

However, technology and automation has guided humans to construct the forecast process, reconstructed it anew, and continue to be integral in the forecast process. Throughout each advance and with each new step of increasing levels of automation and sophistication, humans still remain better than the machines in many tasks. Increased automation need not be the final stage of "meteorological cancer" of which Snellman (1977) warned. Maybe in the 45 years since Snellman's (1977) longing for the ideal human–machine mix, have we finally arrived at the era of "meteorological transformation" we have sought? Operational meteorologists can give up tedious tasks that can be automated and take on more challenging tasks that computers cannot do, such as communication and interpretation of automated products. In turn, operational meteorologists learn new skills and new hires bring new skills, as both adapt to the evolving forecast system and evolve the forecast system. Those whose careers will span the next few years and decades should not bemoan the situation, but embrace automation as a partner, develop lifelong learning to maximize one's marketability, and drive forward one's career path.

Acknowledgments. We thank all the attendees of the webinars, and the workshops at the AMS Annual Meeting in January 2020 for their comments and feedback, which have informed this article. The following AMS personnel assisted in conducting the webinars: Kelly Savoie, Brian Mardirosian, Kristen Knight, Brandon Crose, Sangjun Lee, Jason Emmanuel, Alex Garcia, and the Board of Continuing

Professional Development. Becky Adams-Selin and the AMS Committee on Weather and Forecasting provided support for the panel discussions at the AMS Annual Meeting. Thanks also to Dr. Peter Neilley and Joe Koval of The Weather Channel/Company; Steve Weiss, retired Chief of the Science and Support Branch, NWS Storm Prediction Center; Kennard “Chip” Kasper, Meteorologist in Charge, NWS Forecast Office in Key West, Florida; Kyle Larkin, Customer Success Meteorologist at Climacell Company in Boston; John Ten Hoeve, Deputy Director, NWS Office of Organizational Excellence, Silver Spring, Maryland; and Dr. David Topping, University of Manchester, for more specific input into this study. We thank Editor Jeff Waldstreicher, Martin Baxter, Philip Schumacher, and an anonymous reviewer for their comments that improved the manuscript. Finally, we thank Brian Motta at the Forecast Decision Training Division of the Cooperative Program for Meteorological Education and Training in Boulder, Colorado, for prompting the lead author to update this research effort. Partial funding for Schultz was provided to the University of Manchester by the Natural Environment Research Council Grant NE/N003918/1. Reference to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its recommendation, or favoring by the U.S. Government or NOAA/National Weather Service. Use of information from this publication shall not be used for advertising or product endorsement purposes. The views expressed in this article are those of the author(s) and do not necessarily represent those of the National Weather Service.

Data availability statement. No datasets were generated or analyzed during the current study.

References

Agrawal, S., L. Barrington, C. Bromberg, J. Burge, C. Gazen, and J. Hickey, 2019: Machine learning for precipitation nowcasting from radar images. arXiv, 1912.12132, https://doi.org/10.48550/arXiv.1912.12132.

Alexander, C., and Coauthors, 2020: Rapid Refresh (RAP) and High Resolution Rapid Refresh (HRRR) model development. *26th Conf. on Numerical Weather Prediction*, Boston, MA, Amer. Meteor. Soc., 8A.1, https://ams.confex.com/ams/2020Annual/webprogram/Paper370205.html.

Baars, J., and C. Mass, 2005: Performance of National Weather Service forecasts compared to operational, consensus, and weighted model output statistics. *Wea. Forecasting*, **20**, 1034–1047, https://doi.org/10.1175/WAF896.1.

Barrett, B., 2018: How the weather channel made that insane storm surge animation. Wired, accessed 10 November 2020, www.wired.com/story/weather-channel-hurricane-florence-storm-surge-graphic/.

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.

Bernhardt, J., J. Snellings, A. Smiros, I. Bermejo, A. Rienzo, and C. Swan, 2019: Communicating hurricane risk with virtual reality: A pilot project. *Bull. Amer. Meteor. Soc.*, **100**, 1897–1902, https://doi.org/10.1175/BAMS-D-17-0326.1.

——, G. Dusek, A. Hesse, W. Santos, T. Jennings, A. Smiros, and A. Montes, 2020: Developing a virtual reality video game to simulate rip currents. *J. Vis. Exp.*, **161**, e61296, https://doi.org/10.3791/61296.

Berry, T., and J. Harlim, 2017: Correcting biased observation model error in data assimilation. *Mon. Wea. Rev.*, **145**, 2833–2853, https://doi.org/10.1175/MWR-D-16-0428.1.

Blackhurst, J. L., J. S. Gresham, and M. O. Stone, 2011: The autonomy paradox: Why 'unmanned systems' don't shrink manpower needs. Armed Forces Journal, http://armedforcesjournal.com/the-autonomy-paradox/.

Bosart, L. F., 2003: Whither the weather analysis and forecasting process? *Wea. Forecasting*, **18**, 520–529, https://doi.org/10.1175/1520-0434(2003)18<520:WTWAAF>2.0.CO;2.

Boukabara, S., V. Krasnopolsky, J. Q. Stewart, E. S. Maddy, N. Shahroudi, and R. N. Hoffman, 2019: Leveraging modern artificial intelligence for remote sensing and NWP: Benefits and challenges. *Bull. Amer. Meteor. Soc.*, **100**, ES473–ES491, https://doi.org/10.1175/BAMS-D-18-0324.1.

Bradshaw, J. M., R. R. Hoffman, M. Johnson, and D. D. Woods, 2013: The seven deadly myths of "autonomous systems." *IEEE Intell. Syst.*, **28**, 54–61, https://doi.org/10.1109/MIS.2013.70.

Brockett, P. L., M. Wang, and C. Yang, 2005: Weather derivatives and weather risk management. *Risk Manage. Insur. Rev.*, **8**, 127–140, https://doi.org/10.1111/j.1540-6296.2005.00052.x.

Brooks, H. E., J. M. Fritsch, and C. A. Doswell, III, 1996: The future of weather forecasting: The eras of revolution and reconstruction. *15th Conf. on Weather Analysis and Forecasting*, Norfolk, VA, Amer. Meteor. Soc., 523–526.

Brown, B., and Coauthors, 2021: The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bull. Amer. Meteor. Soc.*, **102**, E782–E807, 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.

Burnett, M., and D. S. LaDue, 2011: How the NWC REU and other extra-curricular experiences impact undergraduates' early career paths. *20th Symp. on Education*, Seattle, WA, Amer. Meteor. Soc., 2.5, http://ams.confex.com/ams/91Annual/webprogram/Paper184229.html.

Calhoun, K. M., and Coauthors, 2018: Cloud-to-ground lightning probabilities and warnings within an integrated warning team. *Special Symp. on Impact-Based Decision Support Services*, Austin, TX, Amer. Meteor. Soc., 4.4, https://ams.confex.com/ams/98Annual/meetingapp.cgi/Paper/329888.

——, K. L. Berry, D. M. Kingfield, T. Meyer, M. J. Krocak, T. M. Smith, G. Stumpf, and A. Gerard, 2021: The Experimental Warning Program of NOAA's Hazardous Weather Testbed. *Bull. Amer. Meteor. Soc.*, **12**, E2229–E2246, https://doi.org/10.1175/BAMS-D-21-0017.1.

Carrington, L., 2016: A qualitative phenomenological study of employee perceptions of the impact of layoffs. Ph.D. dissertation, Walden University, 234 pp., https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=3903&context=dissertations.

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, L. Cronce, and J. Brunner, 2020: NOAA ProbSevere v2.0—ProbHail, ProbWind, and ProbTor. *Wea. Forecasting*, **35**, 1523–1543, https://doi.org/10.1175/WAF-D-19-0242.1.

Clark, A. J., and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, https://doi.org/10.1175/BAMS-D-11-00040.1.

——, R. Bullock, T. Jensen, M. Xue, and F. Kong, 2014: Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Wea. Forecasting*, **29**, 517–542, https://doi.org/10.1175/WAF-D-13-00098.1.

Clark, P., N. Roberts, H. Lean, S. P. Ballard, and C. Charlton-Perez, 2016: Convection-permitting models: A step-change in rainfall forecasting. *Meteor. Appl.*, **23**, 165–181, https://doi.org/10.1002/met.1538.

Costa, G., 2010: Shift work and health: Current problems and preventive actions. *Saf. Health Work*, **1**, 112–123, https://doi.org/10.5491/SHAW.2010.1.2.112.

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**, 1–14, https://doi.org/10.15191/nwajom.2020.0801.

Crout, S., 2019: The importance of including interactive immersive technology in instructional design to enhance learner and user experience. M.S. thesis, Dept. of English, Minnesota State University, 33 pp., https://cornerstone.lib.mnsu.edu/etds/944/.

Daipha, P., 2012: Weathering risk: Uncertainty, weather forecasting, and expertise. *Soc. Compass*, **6**, 15–25, https://doi.org/10.1111/j.1751-9020.2011.00437.x.

——, 2015: From bricolage to collage: The making of decisions at a Weather Forecast Office. *Sociol. Forum*, **30**, 787–808, https://doi.org/10.1111/socf.12192.

Demnitz, R., and S. Joslyn, 2020: The effects of recency and numerical uncertainty estimates on overcautiousness. *Wea. Climate Soc.*, **12**, 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.

Doswell, C. A., 2004: Weather forecasting by humans—Heuristics and decision making. *Wea. Forecasting*, **19**, 1115–1126, https://doi.org/10.1175/WAF-821.1.

Duan, Q., and Coauthors, 2017: Automatic model calibration: A new way to improve numerical weather forecasting. *Bull. Amer. Meteor. Soc.*, **98**, 959–970, https://doi.org/10.1175/BAMS-D-15-00104.1.

Erickson, M. J., B. Albright, and J. A. Nelson, 2021: Verifying and redefining the Weather Prediction Center's excessive rainfall outlook forecast product. *Wea. Forecasting*, **36**, 325–340, https://doi.org/10.1175/WAF-D-20-0020.1.

Erkkilä, T., 2007: About the nature of the forecaster profession and the human contribution to very short range forecasts. *European Forecaster*, No. 14, 6–11, http://euroforecaster.org/gpeasy/gpEasy_CMS/Newsletter_14_(2009).

Ferrucci, D. A., 2012: Introduction to "This is Watson." *IBM J. Res. Dev.*, **56**, https://doi.org/10.1147/JRD.2012.2184356.

Fries, J., 2014: USAF use of an editable weather system for decision support services. *26th Conf. on Weather Analysis and Forecasting/22nd Conf. on Numerical Weather Prediction*, Atlanta, GA, Amer. Meteor. Soc., P167, https://ams.confex.com/ams/94Annual/webprogram/Paper241743.html.

Fritsch, J., and Coauthors, 1998: Quantitative precipitation forecasting: Report of the Eighth Prospectus Development Team, U.S. Weather Research Program. *Bull.*

Amer. Meteor. Soc.*, **79**, 285–323, https://doi.org/10.1175/1520-0477(1998)079<0285:QPFROT>2.0.CO;2.

Frone, M. R., and A.R. Blais, 2020: Organizational downsizing, work conditions, and employee outcomes: Identifying targets for workplace intervention among survivors. *Int. J. Environ. Res. Public Health*, **17**, 719, https://doi.org/10.3390/ijerph17030719.

Gallo, B. T., and Coauthors, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/Hazardous Weather Testbed spring forecasting experiment. *Wea. Forecasting*, **32**, 1541–1568, https://doi.org/10.1175/WAF-D-16-0178.1.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor. Climatol.*, **11**, 1203–1211, https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Gonzalez-Espada, W., and D. S. Zaras, 2006: Evaluation of the impact of the NWC REU program compared with other undergraduate research experiences. *J. Geosci. Educ.*, **54**, 541–549, https://doi.org/10.5408/1089-9995-54.5.541.

Graham, R. A., J. Wiedenfeld, P. Schumacher, M. Fowle, J. Gagan, C. Greif, P. Nutter, and B. Smith, 2021: Evolving the NWS to more probabilistic and ensemble-based forecasts and services. *16th Symp. on Societal Applications*, Online, Amer. Meteor. Soc., 8.3, https://ams.confex.com/ams/101ANNUAL/meetingapp.cgi/Paper/380477.

Gulacsik, G., 2019: The impact of color-coded uncertainty on understanding and decision-making. M.S. thesis, University of Washington, 63 pp., https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/45262/Gulacsik_washington_0250O_20902.pdf?sequence=1&isAllowed=y.

Heinselman, P., D. LaDue, D. M. Kingfield, and R. Hoffman, 2015: Tornado warning decisions using phased array radar data. *Wea. Forecasting*, **30**, 57–78, https://doi.org/10.1175/WAF-D-14-00042.1.

Herman, G. R., and R. S. Schumacher, 2018: "Dendrology" in numerical weather prediction: What random forests and logistic regression tell us about forecasting extreme precipitation. *Mon. Wea. Rev.*, **146**, 1785–1812, https://doi.org/10.1175/MWR-D-17-0307.1.

Hoffman, R. R., 2017: A taxonomy of emergent trusting in the human-machine relationship. *Cognitive Systems Engineering: The Future for a Changing World*, 1st ed. P. Smith and R. R. Hoffman, Eds., CRC Press, 137–164.

——, and S. M. Fiore, 2007: Perceptual (re)learning: A leverage point for human-centered computing. *IEEE Intell. Syst.*, **22**, 79–83, https://doi.org/10.1109/MIS.2007.59.

——, J. K. Hawley, and J. M. Bradshaw, 2014: Myths of automation, part 2: Some very human consequences. *IEEE Intell. Syst.*, **29**, 82–85, https://doi.org/10.1109/MIS.2014.25.

——, D. S. LaDue, H. M. Mogil, P. J. Roebber, and J. G. Trafton, 2017: *Minding the Weather: How Expert Forecasters Think.* MIT Press, 488 pp., https://doi.org/10.2307/j.ctt1t88w2v.

Hong, L., and S. E. Page, 2004: Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc. Natl. Acad. Sci. USA*, **101**, 16 385–16 389, https://doi.org/10.1073/pnas.0403723101.

Huntemann, T. L., D. E. Rudack, and D. P. Ruth, 2015: Forty years of NWS forecasts: Past performance and future advances. *Harry R. Glahn Symp.*, Phoenix, AZ, Amer. Meteor. Soc., 449, https://ams.confex.com/ams/95Annual/webprogram/Manuscript/Paper258606/Huntemann_449_NDFD-GFSMOS-ECMWFMOS_preprint.pdf.

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.

Joslyn, S. L., and J. E. LeClerc, 2012: Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *J. Exp. Psychol. Appl.*, **18**, 126–140, https://doi.org/10.1037/a0025185.

Karstens, C., and Coauthors, 2015: Evaluation of a probabilistic forecasting methodology for severe convective weather in the 2014 Hazardous Weather Testbed. *Wea. Forecasting*, **30**, 1551–1570, https://doi.org/10.1175/WAF-D-14-00163.1.

——, and Coauthors, 2018: Development of a human–machine mix for forecasting severe convective events. *Wea. Forecasting*, **33**, 715–737, https://doi.org/10.1175/WAF-D-17-0188.1.

Klockow-McClain, K., R. A. McPherson, and R. P. Thomas, 2019: Cartographic design for improved decision-making: Trade-offs in uncertainty visualization for tornado threats. *Ann. Assoc. Amer. Geogr.*, **109**, 314–333, https://doi.org/10.1080/24694452.2019.1602467.

——, and Coauthors, 2020: Putting multiple probabilistic products before end users: The 2019 HWT emergency manager experiments. *15th Symp. on Societal Applications*, Boston, MA, Amer. Meteor. Soc., 12A.6, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/369752.

Kuster, C. M., P. L. Heinselman, J. C. Snyder, K. A. Wilson, D. A. Speheger, and J. E. Hocker, 2017: An evaluation of radar-based tornado track estimation products by Oklahoma public safety officials. *Wea. Forecasting*, **32**, 1711–1726, https://doi.org/10.1175/WAF-D-17-0031.1.

LaDue, D. S., and A. E. Cohen, 2018: Facilitating the self-directed learning efforts of professional meteorologists. *Bull. Amer. Meteor. Soc.*, **99**, 2519–2527, https://doi.org/10.1175/BAMS-D-17-0324.1.

Lakshmanan, V., A. Fritz, T. Smith, K. Hondl, and G. Stumpf, 2007: An automated technique to quality control radar reflectivity data. *J. Appl. Meteor. Climatol.*, **46**, 288–305, https://doi.org/10.1175/JAM2460.1.

Martinaitis, S. M., and Coauthors, 2017: The HMT Multi-Radar Multi-Sensor Hydro Experiment. *Bull. Amer. Meteor. Soc.*, **98**, 347–359, https://doi.org/10.1175/BAMS-D-15-00283.1.

McGovern, A., C. D. Karstens, T. Smith, and R. Lagerquist, 2019: Quasi-operational testing of real-time storm-longevity prediction via machine learning. *Wea. Forecasting*, **34**, 1437–1451, https://doi.org/10.1175/WAF-D-18-0141.1.

——, I. Ebert-Uphoff, D. J. Gagne II, and A. Bostrom, 2022: The need for ethical, responsible, and trustworthy artificial intelligence for environmental sciences. arXiv, 2112.08453, https://doi.org/10.48550/arXiv.2112.08453.

McIntyre, M., 1999: Numerical weather prediction: A vision of the future, updated still further. *The Life Cycles of Extratropical Cyclones*, M. A. Shapiro and S. Grønås, Eds., Amer. Meteor. Soc., 337–355, https://link.springer.com/chapter/10.1007/978-1-935704-09-6_22.

McNeeley, S. M., and Coauthors, 2012: Catalyzing frontiers in water–climate–society research: A view from early career scientists and junior faculty. *Bull. Amer. Meteor. Soc.*, **93**, 477–484, https://doi.org/10.1175/BAMS-D-11-00221.1.

Meyer, T. C., K. Berry, K. M. Calhoun, K. E. Klockow, H. Obermeier, A. Campbell, and J. Wolfe, 2019: Visualization and communication of probabilistic lightning hazard information for a broad spectrum of end users. *Ninth Conf. on the Meteorological Application of Lightning Data*, Phoenix, AZ, Amer. Meteor. Soc., 12.5, https://ams.confex.com/ams/29SLS/webprogram/Paper348668.html.

Miran, S. M., C. Ling, J. James, A. E. Gerard, and L. P. Rothfusz, 2017: User perception and interpretation of tornado probabilistic hazard information: Comparison of four graphical designs. *Appl. Ergon.*, **65**, 277–285, https://doi.org/10.1016/j.apergo.2017.06.016.

Morss, R. E., J. L. Demuth, and J. K. Lazo, 2008: Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Wea. Forecasting*, **23**, 974–991, https://doi.org/10.1175/2008WAF2007088.1.

MPAS, 2013: MPAS overview. Accessed 10 November 2020, https://mpas-dev.github.io/.

Murnane, R. J., 2004: Climate research and reinsurance. *Bull. Amer. Meteor. Soc.*, **85**, 697–707, https://doi.org/10.1175/BAMS-85-5-697.

Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281–293, https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2.

Nemunaitis-Berry, K., H. Obermeier, K. M. Calhoun, T. C. Meyer, K. E. Klockow-McClain, D. LaDue, Z. Stanford, and A. Gerard, 2020: Incorporating end users in Hazardous Weather Testbed experiments. *10th Conf. on Transition of Research to Operations*, Boston, MA, Amer. Meteor. Soc., 2.3, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/363946.

Ng, A., 2018: *Machine Learning Yearning: Technical Strategy for AI Engineers in the Era of Deep Learning.* Stanford University Press, 118 pp.

NOAA, 2020a: NOAA artificial intelligence strategy analytics for next-generation Earth science. NOAA Doc., 8 pp., https://sciencecouncil.noaa.gov/Portals/0/2020%20AI%20Strategy.pdf.

——, 2020b: Budget estimates: Fiscal year 2020. NOAA Doc., 946 pp., https://www.noaa.gov/sites/default/files/legacy/document/2019/Nov/NOAA-FY20-Congressional-Justification.pdf.

Novak, D. R., 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, 2014a: Using percentiles to communicate snowfall uncertainty. *Wea. Forecasting*, **29**, 1259–1265, https://doi.org/10.1175/WAF-D-14-00019.1.

——, C. Bailey, K. F. Brill, P. Burke, W. A. Hogsett, R. Rausch, and M. Schichtel, 2014b: Precipitation and temperature forecast performance at the Weather Prediction Center. *Wea. Forecasting*, **29**, 489–504, https://doi.org/10.1175/WAF-D-13-00066.1.

NWS, 2017: Operations and workforce analysis and catalog. NWS Doc., 132 pp., www.weather.gov/media/nws/OWA_Catalog_09072017.pdf.

Nygren, J. M., E. Zukauskaite, and N. Westberg, 2018: User participation in coproduction for health innovation. *JMIR Res. Protoc.*, **7**, e126, www.ncbi.nlm.nih.gov/pmc/articles/PMC5966657/.

Obermeier, H., K. Berry, K. E. Klockow-McClain, T. C. Meyer, P. A. Campbell, A. E. Gerard, J. E. Trujillo, and C. Carithers, 2020: Communicating probabilistic hazard information: Broadcast meteorologists in the 2018-2019 Hazardous Weather Testbed. *15th Symp. on Societal Applications*, Boston, MA, Amer. Meteor. Soc., 17.4, https://ams.confex.com/ams/2020Annual/webprogram/Paper368267.html.

Peachey, J. A., D. M. Schultz, R. Morss, P. J. Roebber, and R. Wood, 2013: How forecasts expressing uncertainty are perceived by UK students. *Weather*, **68**, 176–181, https://doi.org/10.1002/wea.2094.

Privé, N. C., and R. M. Errico, 2013: The role of model and initial condition error in numerical weather forecasting investigated with an observing system simulation experiment. *Tellus*, **65A**, 21740, https://doi.org/10.3402/tellusa.v65i0.21740.

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.

Roebber, P. J., 2013: Using evolutionary programming to generate skillful extreme value probabilistic forecasts. *Mon. Wea. Rev.*, **141**, 3170–3185, https://doi.org/10.1175/MWR-D-12-00285.1.

——, D. M. Schultz, B. A. Colle, and D. J. Stensrud, 2004: Toward improved prediction: High-resolution and ensemble modeling systems in operations. *Wea. Forecasting*, **19**, 936–949, https://doi.org/10.1175/1520-0434(2004)019<0936:TIPHAE>2.0.CO;2.

——, M. Westendorf, and G. R. Meadows, 2010: Innovative weather: A new strategy for student, university, and community relationships. *Bull. Amer. Meteor. Soc.*, **91**, 877–888, https://doi.org/10.1175/2010BAMS2854.1.

Rose, B., and Coauthors, 2015: A Human Over the Loop forecast paradigm at The Weather Company (TWC). *27th Conf. on Wea. Forecasting,* Chicago, IL, Amer. Meteor. Soc., 7B.5, https://ams.confex.com/ams/27WAF23NWP/webprogram/Paper273492.html.

Rothfusz, L. P., and Coauthors, 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.

Runk, K. J., and Coauthors, 2020: Evaluating the role of the mesoanalyst in severe weather impacts based decision support services. Part III—Messaging focus. *Eighth Symp. on Building a Weather-Ready Nation*, Boston, MA, Amer. Meteor. Soc., 6.3, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/366884.

Sherman-Morris, K., K. B. Antonelli, and C. C. Williams, 2015: Measuring the effectiveness of the graphical communication of hurricane storm surge threat. *Wea. Climate Soc.*, **7**, 69–82, https://doi.org/10.1175/WCAS-D-13-00073.1.

Sills, D. M., 2009: On the MSC forecasters forums and the future role of the human forecaster. *Bull. Amer. Meteor. Soc.*, **90**, 619–628, https://doi.org/10.1175/2008BAMS2657.1.

Silva, C. L., and Coauthors, 2018: Refining the baseline: Public reception, understanding, and responses to severe weather forecasts and warnings in the contiguous United States. University of Oklahoma Center for Risk and Crisis Management Rep., 29 pp., http://risk.ou.edu/downloads/news/WX18-Reference-Report.pdf.

Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype Warn-on-Forecast System. *Wea. Forecasting*, **33**, 1225–1250, https://doi.org/10.1175/WAF-D-18-0020.1.

Smith, T. M., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 1617–1630, https://doi.org/10.1175/BAMS-D-14-00173.1.

Snellman, L. W., 1977: Operational forecasting using automated guidance. *Bull. Amer. Meteor. Soc.*, **58**, 1036–1044, https://doi.org/10.1175/1520-0477(1977)058<1036:OFUAG>2.0.CO;2.

Sobash, R. A., G. S. Romine, and C. S. Schwartz, 2020: A comparison of neural-network and surrogate-severe probabilistic convective hazard guidance derived from a convection-allowing model. *Wea. Forecasting*, **35**, 1981–2000, https://doi.org/10.1175/WAF-D-20-0036.1.

Sørensen, E., and J. Torfing, 2011: Enhancing collaborative innovation in the public sector. *Adm. Soc.*, **43**, 842–868, https://doi.org/10.1177/0095399711418768.

Steeneveld, G., and J. Vilà-Guerau de Arellano, 2019: Teaching atmospheric modeling at the graduate level: 15 years of using mesoscale models as educational tools in an active learning environment. *Bull. Amer. Meteor. Soc.*, **100**, 2157–2174, https://doi.org/10.1175/BAMS-D-17-0166.1.

Stoffler, R., 2017: The USAF GALWEM: Improving military decision making, advancing national capability. *28th Conf. on Weather Analysis and Forecasting*, Seattle, WA, Amer. Meteor. Soc., 2B.2, https://ams.confex.com/ams/97Annual/webprogram/Paper313034.html.

Stuart, N. A., and Coauthors, 2006: The future of humans in an increasingly automated forecast process. *Bull. Amer. Meteor. Soc.*, **87**, 1497–1502, https://doi.org/10.1175/BAMS-87-11-1497.

——, D. M. Schultz, and G. Klein, 2007: Maintaining the role of humans in the forecast process: Analyzing the psyche of expert forecasters. *Bull. Amer. Meteor. Soc.*, **88**, 1893–1898, https://doi.org/10.1175/BAMS-88-12-1893.

Stumpf, G. J., and A. E. Gerard, 2021: National Weather Service severe weather warnings as Threats-in-Motion. *Wea. Forecasting*, **36**, 627–643, https://doi.org/10.1175/WAF-D-20-0159.1.

Sutton, J., and L. M. Fischer, 2021: Understanding visual risk communication messages: An analysis of visual attention allocation and think-aloud responses to tornado graphics. *Wea. Climate Soc.*, **13**, 173–188, https://doi.org/10.1175/WCAS-D-20-0042.1.

Tallapragada, V., 2020: Unified forecast system development and operational implementation plans at NCEP/EMC. *10th Conf. on Transition of Research to Operations*, Boston, MA, Amer. Meteor. Soc., 5A.2, https://ams.confex.com/ams/2020Annual/webprogram/Paper362787.html.

Tew, M. A., A. Stern, A. Horvitz, K. Gilbert, and D. T. Myrick, 2016: National Blend of Models: Transformational forecast change by the National Weather Service. *Fourth Symp. on Building a Weather-Ready Nation*, New Orleans, LA, Amer. Meteor. Soc., 6.2, https://ams.confex.com/ams/96Annual/webprogram/Paper286228.html.

Treinish, L. A., A. P. Praino, and Z. D. Christidis, 2003: Implementation of mesoscale numerical weather prediction for weather-sensitive business operations. *19th Conf. on IIPS*, Long Beach, CA, Amer. Meteor. Soc., 3.9, https://ams.confex.com/ams/annual2003/techprogram/paper_55570.htm.

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.

Voudouri, A., P. Khain, I. Carmona, O. Bellprat, F. Grazzini, E. Avgoustoglou, J. M. Bettems, and P. Kaufmann, 2017: Objective calibration of numerical weather prediction models. *Atmos. Res.*, **190**, 128–140, https://doi.org/10.1016/j.atmosres.2017.02.007.

Waldstreicher, J. S., D. R. Radell, and P. Schumacher, 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.

Wilson, K. A., P. L. Heinselman, C. M. Kuster, and D. M. Kingfield, 2017a: Forecaster performance and workload: Does radar update time matter? *Wea. Forecasting*, **32**, 253–274, https://doi.org/10.1175/WAF-D-16-0157.1.

——, ——, and ——, 2017b: Considerations for phased-array radar data use within the National Weather Service. *Wea. Forecasting*, **32**, 1959–1965, https://doi.org/10.1175/WAF-D-17-0084.1.

Woods, D. D., and E. Hollnagel, 2006: *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering.* CRC Press, 232 pp.

Yussouf, N., K. A. Wilson, S. M. Martinaitis, H. Vergara, P. L. Heinselman, and J. J. Gourley, 2020: The coupling of NSSL Warn-on-Forecast and FLASH systems for probabilistic flash flood prediction. *J. Hydrometeor.*, **21**, 123–141, https://doi.org/10.1175/JHM-D-19-0131.1.

Zhao, Q., Y. Zhu, D. Wan, Y. Yu, and X. Cheng, 2018: Research on the data-driven quality control method of hydrological time series data. *Water*, **10**, 1712, https://doi.org/10.3390/w10121712.

Zheng, M., E. K. M. Chang, and B. A. Colle, 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.

Zhu, Y., and N. Li, 2021: Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review. *J. Saf. Sci. Resilience*, **2**, 1–10, https://doi.org/10.1016/j.jnlssr.2020.11.004.