BAMS Article

National Weather Service (NWS) Forecasters' Perceptions of AI/ML and Its Use in Operational Forecasting

Christopher D. Wirz,[a,b] Julie L. Demuth,[a,b] Mariana G. Cains,[a,b]
Miranda White,[c,b] Jacob Radford,[b,d,e] and Ann Bostrom[b,f]

KEYWORDS:
Social Science;
Operational forecasting;
Communications/decision making;
Artificial intelligence;
Machine learning

ABSTRACT: Artificial intelligence and machine learning (AI/ML) have attracted a great deal of attention from the atmospheric science community. The explosion of attention on AI/ML development carries implications for the operational community, prompting questions about how novel AI/ML advancements will translate from research into operations. However, the field lacks empirical evidence on how National Weather Service (NWS) forecasters, as key intended users, perceive AI/ML and its use in operational forecasting. This study addresses this crucial gap through structured interviews conducted with 29 NWS forecasters from October 2021 through July 2023 in which we explored their perceptions of AI/ML in forecasting. We found that forecasters generally prefer the term "machine learning" over "artificial intelligence" and that labeling a product as being AI/ML did not hurt perceptions of the products and made some forecasters more excited about the product. Forecasters also had a wide range of familiarity with AI/ML, and overall, they were (tentatively) open to the use of AI/ML in forecasting. We also provide examples of specific areas related to AI/ML that forecasters are excited or hopeful about and that they are concerned or worried about. One concern that was raised in several ways was that AI/ML could replace forecasters or remove them from the forecasting process. However, forecasters expressed a widespread and deep commitment to the best possible forecasts and services to uphold the agency mission using whatever tools or products that are available to assist them. Last, we note how forecasters' perceptions evolved over the course of the study.

SIGNIFICANCE STATEMENT: Despite a range of familiarity with artificial intelligence and machine learning (AI/ML), forecasters are open to using AI/ML tools operationally. The extent of this openness ranged from being highly supportive to having some important concerns about how effective AI/ML can be and whether or not it would replace them. Although some forecasters see AI/ML products as the exciting cutting edge of science, others care little of the development approach and more about how well the product verifies and helps them do their job.

DOI: 10.1175/BAMS-D-24-0044.1

Corresponding author: Christopher D. Wirz, cdwirz@ucar.edu

Manuscript received 8 February 2024, in final form 21 August 2024, accepted 7 October 2024

AFFILIATIONS: [a] NSF National Center for Atmospheric Research, Boulder, Colorado; [b] NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES), Norman, Oklahoma; [c] Texas A&M-Corpus Christi, Corpus Christi, Texas; [d] Cooperative Institute for Research in the Atmosphere, For Collins, Colorado; [e] NOAA/Global Systems Laboratory, Boulder, Colorado; [f] University of Washington, Seattle, Washington

1. Introduction

Artificial intelligence and machine learning (AI/ML) have attracted a great deal of attention from the scientific community and general public over the last few years with advancements across many domains. Atmospheric science is no exception. The field has seen an explosion in AI/ML research (e.g., Chase et al. 2022), with applications ranging from severe convective weather (Flora et al. 2021) and atmospheric rivers (Chapman et al. 2022) to subseasonal-to-seasonal forecasts (Mayer and Barnes 2022). Traditionally, many of these AI/ML applications leverage existing physics-based models or output in some way (e.g., Gagne et al. 2017; Hill et al. 2023; Sobash et al. 2023); however, new "pure-AI" or "data-driven" AI/ML approaches are now showing skill at modeling certain parameters on a global level relying on only an initial atmospheric state, though often still dependent on physics-based data assimilation systems (Bi et al. 2023; Ben Bouallègue et al. 2024; Lam et al. 2022; Pathak et al. 2022). These approaches demonstrate how AI/ML are in some ways pushing the field into new, largely unknown places (Ebert-Uphoff and Hilburn 2023).

Despite the explosion of developmental attention on AI/ML, there has been disproportionately little research attention on the implications for the operational community. Early work has begun with questions about how the novel advancements will translate from the research community into practice. The National Weather Service (NWS) has begun addressing the growth of AI/ML through efforts such as educational lecture series for forecasters (Roebber 2022) and preliminary studies about the state of AI/ML readiness across the agency (Roebber and Smith 2023). Furthermore, the NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES) is working to develop trustworthy AI/ML applications that are user centered by conducting research with forecasters to better understand their needs and decision context (M. G. Cains et al. 2024) and by developing research agendas with additional user-centered needs in mind (Bostrom et al. 2023). Researchers are also evaluating how forecasters perceive specific AI/ML models through avenues like the hazardous weather testbed (HWT; e.g., Clark et al. 2023) and the hydrometeorology testbed (HMT; Schumacher et al. 2021).

Even with this important and growing area of research, we lack empirical evidence on how NWS forecasters, as key intended users, perceive AI/ML and its use in operational forecasting. This study aims to address this fundamental gap in our understanding through the analysis of interviews with 29 NWS forecasters from around the United States conducted from October 2021 through July 2023. In doing so, this work builds on past research that has demonstrated the importance and value of research on forecasters' needs and decision-making for improving the development of model guidance (e.g., Demuth et al. 2020; Henderson et al. 2022; Novak et al. 2008; Stuart et al. 2022) but that was not focused on AI/ML guidance. Our work complements M. G. Cains et al. (2024) by examining forecasters' general perceptions of AI/ML for operational forecasting, in contrast to the examination in M. G. Cains et al. (2024) of how forecasters assessed the trustworthiness of different attributes of two specific experimental AI/ML products (one predicting convective storm mode and one predicting severe hail).

Understanding forecaster's perceptions at this critical time in the AI/ML developmental life cycle is essential for advancing operational, developmental, and theoretical goals. *Operationally, forecasters are crucial decision-makers and communicators in the weather enterprise, so how they perceive and interact with new technologies, like AI/ML, influences the services provided by the NWS*. Furthermore, forecasters are also the sole authority for issuing mission-critical products like watches and warnings. Developmentally, the creation and refinement of new AI/ML tools and technology can leverage forecasters' perceptions to develop products in ways that meet forecasters' needs, address potential concerns, and enhance benefits, all of which increase the likelihood that products are both useful and used. Theoretically, a better understanding of how forecasters perceive AI/ML advances our understanding of expert decision-making using new technologies in high-stakes environments. To address these areas, we set out to answer the overarching research question of how do NWS forecasters perceive AI/ML and its use in forecasting, in general.

2. Methods

To investigate our research question, we conducted qualitative research, which is well suited for providing a deep, nuanced understanding of a research area such as forecasters' perception of AI/ML about which so little is known (Merriam and Tisdell 2016). These detailed insights are an essential first step for advancing social scientific knowledge in new areas because they provide a strong, empirical foundation for continued examinations. Without the nuanced understanding derived from the interviews, it can be difficult to assess what measures are important to study and how with quantitative research.

a. Sampling and overview of the final sample of NWS forecasters. We analyzed qualitative interview data from two related data collections with NWS forecasters. Both data collections are related efforts of the NSF AI2ES aimed to develop use-inspired AI/ML that users deem both trustworthy and useful for their decision-making (McGovern et al. 2022). The two sets of interviews used the same general design but were focused on different forecasting challenges and had related but distinct aims. The first set of structured interviews (n = 16) was conducted from October 2021 to May 2022 and focused on how NWS forecasters assessed the trustworthiness of and made use of decisions about new guidance or products in the context of forecasting for severe convective weather (hereafter referred to as the "severe interviews"). The second set of interviews with NWS forecasters (n = 13) focused on the same aspects of trustworthiness and use of decisions but did so using coastal fog as a hazard and was conducted from December 2022 to July 2023 (hereafter referred to as the "coastal interviews"). Both sets of interviews used the same approach to assess forecasters' perceptions of AI/ML, which allowed us to combine the datasets for this study.

We developed our sampling frame[1] of NWS Weather Forecasting Offices (WFOs) to consider interviewing based on the geographic areas that climatologically experience either severe convective weather (see M. G. Cains et al. 2024 for details) or coastal fog (Hardwick 1973). After we randomly selected offices from our sampling frame, we coordinated with the NWS regional offices and then randomly selected one of the three positions of interest from that office to interview. These positions were Science and Operations Officers (SOOs), lead forecasters, and general forecasters [General Schedule (GS) 5-11]. We then contacted the management of the selected WFO to see whether

[1] A sampling frame is a full list of members of the population of interest from which researchers select participants for a study.

there were interest and capacity for the interviews. If the office did not have the capacity or an interested forecaster, we randomly selected a new office and repeated the process. This sampling approach was designed to provide us with a range of perspectives and experiences across forecasting roles and regions.

For this research, we analyzed both sets of interviews, resulting in a final sample size of 29 interviews conducted from October 2021 to July 2023. Over this time, the field of AI/ML continued its rapid growth both in the atmospheric science domain (Chase et al. 2022; Roebber and Smith 2023) and in general with advancements like the release of ChatGPT in November 2022 (Heaven 2023). Our data provide a unique opportunity to see how NWS forecasters perceived the increasing development and advancements of AI/ML over this key period of AI/ML.

To further understand how AI/ML were perceived during this pivotal time, we used two interview versions to assess the potential effects of labeling a new product as "AI/ML." To study this, we had a version of the interview in which we said AI/ML throughout the interview (referred to as the "AI/ML version") and one where we did not mention AI/ML until the very end of the interview, instead referring to the AI/ML guidance as "new guidance" throughout the interviews without changing any other details about the AI/ML products (referred to as the "No AI/ML version"). Forecasters were randomly assigned to either the AI/ML version or the No AI/ML version of the interview. Figure 1 illustrates this distinction for one piece of information from the coastal interviews. Note that even though we do not explicitly say AI/ML in the No AI/ML version, we still discuss the techniques in the same way. As we discuss in the results, some forecasters in the No AI/ML interviews did recognize "convolution neural network (CNN)" as AI/ML.

A: Excerpt from AI/ML interview Version

AI/ML Technique Used

- The machine learning approach used to develop the FogNet guidance is called a convolutional neural network (CNN).
- A CNN is a set of spatial filters that can learn to detect patterns and correlations at various scales. These filters are similar to those used for smoothing or neighborhood averaging, except that weights in the filters are learned.
- In this case, the 3D-CNN algorithm stacks horizontal maps, each with different atmospheric and oceanic variables, allowing the filters to detect patterns across combinations of NAM outputs and satellite-derived sea surface temperatures over different time steps to predict the occurrence of fog.

B: Excerpt from No AI/ML interview Version

New Technique Used

- The new approach used to develop the FogNet guidance is called a convolutional neural network (CNN).
- A CNN is a set of spatial filters that can learn to detect patterns and correlations at various scales. These filters are similar to those used for smoothing or neighborhood averaging, except that weights in the filters are learned.
- In this case, the 3D-CNN algorithm stacks horizontal maps, each with different atmospheric and oceanic variables, allowing the filters to detect patterns across combinations of NAM outputs and satellite-derived sea surface temperatures over different time steps to predict the occurrence of fog.

Fig. 1. Excerpt from the materials reviewed in the coastal interviews from the (A) AI/ML and (B) No AI/ML versions of the interviews. The added highlighting shows the only differences between the versions; the highlight was not present during the interviews.

All interviews were conducted virtually and recorded via Google Meet. The severe interviews were led by authors MGC and CDW, and the coastal interviews were led by authors CDW, MW, JR, and MGC. The lead author (CDW) either led or observed all of the 29 interviews. The breakdown of the final sample used for this study can be found in Table 1.

Table 1. Summary of the sample of forecasters by the region in which they work, their role within the forecast office, and the version of the interview they received (the version that mentioned AI/ML or the one that did not).

	Severe	Coastal	Total
Region			
Alaska	0	2	2
Central	7	0	7
Eastern	2	3	5
Southern	7	4	11
Western	0	4	4
Role			
SOO	5	2	7
Lead	4	3	7
General	7	8	15
Version			
No AI/ML	6	6	12
AI/ML	10	7	17
Total	16	13	29

b. Relevant data: Interview materials and questions of interest. The two data collections for this study were largely focused on assessing early stage products derived using AI/ML for severe convective weather (Burke et al. 2020; Sobash et al. 2023) and coastal fog (Kamangir et al. 2021, 2022) forecasting to better understand which aspects of new products influenced how forecasters assessed new guidance. The main emphasis of the interviews was on forecasting processes for these hazards and an in-depth review of specific AI/ML products. An example of the content reviewed by the forecasters is presented in Fig. 1. We also asked forecasters a series of general questions that were not specific to the AI/ML products they were reviewing. For this study, we focus on just a few questions from the interviews focused on forecasters' perceptions of AI/ML. These questions were as follows:

- **Q1: All interviews:** When you hear the terms AI and ML, what comes to mind? Is one of the terms more meaningful to you with regard to your work?
- **Q2: All interviews:** How do you feel about the use of AI/ML in forecasting?

We asked all forecasters these questions in this order. For the AI/ML version of the interviews, we asked these questions early in the interview, before we introduced the forecasters to the example guidance. For the No AI/ML interviews, we asked these questions at the end of the interview because we did not want to prime how forecasters reviewed the guidance. The No AI/ML version provided us with an interesting opportunity to ask different follow-up questions, so we asked these forecasters the following questions after the two listed above:

- **Q3: No AI/ML version only**: Some developers and meteorologists consider the new guidance you just reviewed to be AI/ML. Does knowing this change how you think about the guidance?
- **Q4: No AI/ML version only**: Relatedly, there are developers and meteorologists who also consider some commonly used guidance AI/ML. Would finding out that guidance you regularly use was considered AI/ML change how you think about or use it?

Our analyses for this paper focus on forecasters' responses to these four questions. All interview materials, including the full interview protocols and information boards for the AI/ML and No AI/ML versions of the interviews, can be found in Cains et al. (2024a,b) for the severe weather interviews and in Wirz et al. (2024a,b) for the coastal fog interviews. Note that the data analysis of the four questions reported here is original, as these questions and data are not analyzed in the related papers (Cains et al. 2024a,b; Wirz et al. 2024a,b).

All data collections and materials were approved by the NSF National Center for Atmospheric Research Human Subjects Committee.

c. Qualitative data analysis. In addition to our qualitative research design, we analyzed our data using a qualitative approach to synthesize the rich, detailed data our interviews generated. Qualitative data analysis, broadly defined, is how researchers classify and interpret data in order to make claims about what is represented, both explicitly and implicitly, within that data (Flick 2014). This process is highly iterative, involving many reviews of the data to generate and refine sets of themes and categories across the data (Merriam and Tisdell 2016). This general approach has a long history in social science (Glaser and Strauss 1967) and has been refined into methods like reflexive thematic analysis (Braun and Clarke 2006; Clarke and Braun 2013, 2017).

For this study, all analyses were led by the first author (CDW) and later jointly interpreted iteratively with coauthors. This process began with taking field notes, which are unstructured written notes meant to capture in situ observations and impressions of key ideas, themes, observations, and impressions (Bogdan and Biklen 1998). These field notes were regularly reviewed and used to contextualize the forecasters' responses to the specific questions of interest for this study listed in the previous section. To examine these questions, the first author reviewed the video recordings of each of the 29 interviews, isolated the section of the video where the four questions were asked and answered, made a copy of just that segment for each interview, and then compiled all the segments into one video file of all the severe interview clips and one of all the coastal interviews. This allowed the first author to readily analyze the data of interest, including hearing the tone and observing nonverbal cues of the forecasters. The first author began by reviewing his field notes for each interview, analyzing the video recording for the key questions, and noting observations and impressions in an analytic memo. He repeated this process several times, iteratively developing and refining the analytic memos into themes.

Themes were developed, refined, and prioritized based on their significance to the overarching research question and study context (Richards 2020). Some themes were designed to be comprehensive and mutually exclusive, meaning they were applicable to all forecasters and consisted of categories that each forecaster would fit into only one Marshall and Rossman (2014). For example, as discussed below, one resultant theme represents the patterns that emerged in the ways that forecasters talked about their experiences with and/or understanding of AI/ML. We developed four groups—*Do not know much*, *Some but fuzzy*, *Experienced dabbler*, and *Formal training* (Table 2)—and categorized each forecaster into one of them based on their experience with AI/ML. Other resultant themes are based on the data from multiple forecasters but not necessarily from all forecasters.

The themes listed in the results below were refined and reevaluated for validity and comprehensiveness through iterative reviewing of the data, analytic memos, and field notes. The first author then identified exemplar quotations for each theme and its different dimensions. These exemplar quotations are key units for representing and sharing qualitative research findings. The goal for these quotations is that they are pieces of data that convey something broader about the study and are the most concise pieces of data that convey this broader meaning without the need of additional information (Lincoln and Guba 1985). The initial sets of exemplar quotations and themes were shared with and examined by the second author (JLD), which led to further refinement of the key themes. Next, the themes were shared with members of NWS management across all the regions represented in the study in a presentation that served as a way to both share the results of the work and assess the face validity of the selected themes. The themes were then reviewed by other members of the study team and the forecasters who were interviewed. Together, these approaches helped finalize and

Table 2. Overview of the thematic codes for NWS forecasters' familiarity with AI/ML. Each thematic code (*Do not know much, Some but fuzzy, Experienced dabbler, and Formal training*) is accompanied by a brief description and two exemplar quotations from forecasters to demonstrate how the theme manifested in the interviews.

Familiarity	Description and exemplar quotations
Do not know much	Description: Forecasters who reported they were not familiar with AI/ML or did not know much about it
	SF10: I don't know a lot about [AI/ML]. I know it's way beyond like Isaac Asimov and all that stuff and now. But it's not something I learned about in school, so I don't know a lot about it either CF01: I guess I'm just not exposed to it as much. I don't know exactly what has been machine learning. I really don't—I don't know
Some but fuzzy	Description: Forecasters who knew some details about AI/ML or had some experience with a model but were unsure about specifics or details
	SF14: And this isn't completely true—it's still sort of statistics. I don't understand the interior guts of it, but it's some statistical method of some type that's doing it CF06: What I would think is that the AI is using a lot of weather models and then comparing it to surface observations to try to develop which biases the models might hold and how to forecast it better. I imagine the AI would be looking at basically every variable—satellite models and surface observations—to try to then forecast out in time
Experienced dabbler	Description: Forecasters who reported having more experience than the "some but fuzzy" group with an AI/ML model and/or had investigated AI/ML through reading papers or listening to talks
	SF05: I did participate in the Hazardous Weather Testbed a couple years ago. So I got versed in a few [AI/ML] things there. I've also just read a lot of articles about [AI/ML] on my own time CF03: I am one of the participants in the severe forecast challenge in the Weather Service and last year especially started using [the CSU severe weather forecasts] quite a bit to start to narrow my focus into a certain area for where I want to target my forecast. So that's the first thing that pops out into my head because they use machine learning to generate those forecasts
Formal training	Description: Forecasters who reported having detailed knowledge about AI/ML through courses and experiences in graduate school
	CF11: I've had familiarity with its applications, things like neural networks and forecasting for quite some time now. Just since being exposed to some of the concepts while I was getting my masters and looking at things like that in grad school CF09: I happen to do machine learning for my master's degree research. So, I'm relatively familiar with it and so when I hear machine learning, I think it's a really useful tool to create certain things and to analyze data

refine the themes, as well as assess their significance and meaningfulness (Patton 2014). This interactive and collaborative process yielded the themes and exemplar quotes provided in the results section. Quotations from the severe interviews are denoted by "SF" and those from the coastal interviews are noted by "CF." The number following each code reflects a single forecaster.

3. The power of the labels: General preference for the term machine learning over artificial intelligence and the "AI/ML effect"

When asked what comes to mind when they hear the terms artificial intelligence or machine learning, forecasters provided a range of associations. These associations spanned science fiction, descriptions of how AI/ML work, fears of being replaced by technology, and specific AI/ML forecasting tools they have used or heard about. Several forecasters also explicitly mentioned they had different associations for AI and ML, which underlies our first key finding.

a. Forecasters generally prefer the term machine learning over artificial intelligence, but the reasons for this vary. When asked which term was more meaningful to their work, forecasters overwhelmingly selected machine learning over artificial intelligence, with some having no preference for one term over the other and only 1 of the 29 preferring artificial intelligence. This pattern strengthened slightly over the course of the interviews, with fewer forecasters having no preference and more preferring the term machine learning. However, there were differences in *why* forecasters stated they preferred machine learning.

Some forecasters reported ML was what they had seen or heard in professional contexts, such as at conferences, in academic papers, and around the office. The professional familiarity led forecasters to prefer the use of the term machine learning over artificial intelligence when discussing applications related to forecasting. For example, when we asked forecaster SF10 which term they preferred, they said:

> SF10: Machine learning—I've heard that in some conferences, like when I've been to the AMS[2] annual meetings. I've seen papers and presentations talking about [using machine learning] so I'm a little bit more familiar [with it] between the two. I don't know if they're two sides of the same coin or anything like that.

[2] American Meteorological Society.

This forecaster's response demonstrates how influential the terminology used to describe AI/ML in contexts like conferences and publications is for their preference, despite not being familiar with any differences between the two terms. This suggests the preference among forecasters is at least partially the result of how the research community has set the agenda, intentionally or not, pertaining to AI/ML. Further evidence of this subtle but powerful agenda setting was that several forecasters easily reported a preference for machine learning over artificial intelligence but were unsure as to exactly why they held that preference.

Other forecasters reported that the term machine learning sounded more consistent with their conceptual understanding of how other techniques used to understand or predict the atmosphere work. In other words, some forecasters thought the term machine learning better represented their understanding of how model guidance and other forecasting tools are generally developed than the term artificial intelligence. For example, when we asked forecaster CF08 which term they preferred, they said:

> CF08: Machine learning seems like a better phrase for me personally. I guess I haven't really dabbled much at all into it. But to me it seems like it's just a more advanced technique of what we've been doing for a long time. We're just trying to find correlations between certain things to come up with a certain outcome, and for the human to do it, it takes a lot of effort, work, and number crunching. But I can imagine having a computer doing that aspect on its own multiple times a day or second or whatever. It's just gonna be more efficient, eventually.

Notice how this forecaster first makes the connection to the ways that the term ML connects to their understanding of how guidance is typically derived. However, they then expand on that idea and describe how the term evokes a positive image of computers efficiently aiding the forecasting process. This quote represents not only how some forecasters found the term machine learning to be an intuitive extension of tools developed to assist with forecasting but also how the term was, as forecaster CF01 put it, "more enticing" than artificial intelligence.

In addition to a preference for the term machine learning, several forecasters reported an aversion to the term artificial intelligence. The negative association stemmed from their belief that AI was used more in the public domain and seemed to have a more negative connotation than ML. Forecasters often reported this came from how AI/ML were represented in mainstream media. As forecasters CF12 and SF14 describe:

> CF12: You see AI used a lot now in the common mainstream media, and I think [sometimes] AI is cast positively and sometimes it's cast through the lens of "the robots are going to take over the world."

SF14: AI I think has publicly and socially gotten to be a negative—or has a negative connotation. You never see in the movies AI going right.

These quotes demonstrate how forecasters are sensitive to the complex and varied coverage AI receives in the media (Korneeva et al. 2023; Ouchchy et al. 2020). Conversely, forecasters also noted that they did not see the term machine learning come up in the media, or at least in the same ways, which further reinforced their preference for ML over AI.

Overall, forecasters' perceptions of AI and preference for the term machine learning have been influenced by how the news and entertainment industries have represented the technology. These findings suggest that the language used by the research and operational communities has subtly led to these preferences, which have been reinforced by a perceived negative portrayal and response to AI in the public domain.

b. The AI/ML effect: The implications of labeling something as AI/ML. We were also interested in understanding whether explicitly labeling a product as AI/ML affected how forecasters perceived it. As a reminder, the main portion of these interviews involved forecasters reviewing information about new prototype products and evaluating how trustworthy they thought the products were. Thus, we analyzed how the 12 forecasters who were not initially told the product they reviewed was AI/ML (Table 1 and Fig. 1) responded to two questions about whether or not finding out the product they had just looked at (Q3 in section 2b) or a product they had used before (Q4 in section 2b) was considered AI/ML would affect the way they viewed the product. The results in this section synthesize how these 12 forecasters responded to these two questions.

Several forecasters were able to figure out the product they were reviewing for the earlier portion of the interview was AI/ML, largely based on the technique name and description that were provided (Fig. 1), despite not being explicitly told it was an AI/ML product. Even though we purposefully did not reveal the product was AI/ML nor mention ML in any part of the interview, there is enough growing familiarity with AI/ML that several forecasters ascertained it was an AI/ML product. Some did not, however. There was a trend throughout the study period with initially no forecasters realizing the products they were reviewing were AI/ML to many more of the later interviews identifying the product as AI/ML. This further reinforces forecasters' growing familiarity with AI/ML from 2021 to 2023. When asked whether they were to find out the product they had just reviewed or one they regularly used was AI/ML would change how they thought about it or used it, none of the forecasters reported that it would negatively impact their perceptions of the product. However, there were notable differences in *how* forecasters responded to these questions.

Several forecasters were largely unphased when we told them the product they had just looked at was AI/ML and said they would not care if they found out they were already using an AI/ML product in their current work. Often these forecasters reported being more concerned with how well the product verified and/or helped them with key forecasting challenges. If the product consistently performed well and was helpful, the specific details on development were not that important to them. Forecaster SF13 summarizes this perspective well when we asked whether finding out a product was AI/ML changed their view or use of it:

SF13: Probably not. I mean to me guidance is just that, guidance. It's something that I look at to help come up with a most likely solution for our customers and partners. Whether or not it's AI, for me personally, it really wouldn't make a difference.

As this forecaster describes, a key part of how they evaluate products is how it supports their forecasting and decision support services. These results that illustrate forecasters' primary concern about the skill and utility of the product rather than having full details about how the guidance was developed are not specific to AI/ML; they echo forecaster perspectives about numerical weather prediction (NWP) as discussed in Demuth et al. (2020). Similarly, forecaster CF03 said they saw the AI/ML product "*as just another tool in my toolbox*." However, other forecasters reported that finding out a product was AI/ML would shift their considerations. They particularly noted the importance of knowing more about the data used to train the AI/ML product, specifically that the training data were sufficient, as mentioned by forecaster CF09:

> CF09: I don't think [learning it's AI/ML] really changes it outside of knowing that it's extra important to understand what it was trained on and where, perhaps even more so than a normal weather model, though it's important to know that too. Beyond that, I don't think it really changes my view of it. If anything, I like it better because I know that it's actually meaningfully weighting these variables rather than the HREF or something which might have some very crude weighting between its model members to produce something like fog.

This forecaster demonstrates how they would evaluate AI/ML products differently than traditional NWP products. They note that training data are important for the performance of AI/ML products, but they also suggest there could be strengths over traditional approaches.

There were also some forecasters who said they would be more excited about the AI/ML products and might even assume it was better in some way. This was largely consistent over the course of the study, but it was slightly more common in the second half of the study period. The more positive perceptions tended to be associated with excitement and interest in getting experience with AI/ML products. As forecaster SF16 discussed, there is a certain hype around AI/ML in the research and development community that influenced how some forecasters perceive AI/ML tools:

> SF16: If I learned that something was machine learning, this probably wouldn't be the right way to think about it, but I would probably think it's better than one without [ML]. Just because I kind of associate machine learning with the "state of the art" trend in science. I think [of ML as] newer, better, shinier—that sort of thing.

Of note, one forecaster who described themself as being somewhat skeptical of using AI/ML in forecasting reported that learning the product they had just reviewed was AI/ML did not change their view of the product but rather made them feel more positive about AI/ML. As this forecaster describes, they did not really understand AI/ML, but seeing how similar the AI/ML product was to traditional guidance made them view AI/ML more favorably:

> CF05: This is similar to some guidance that we use and knowing that what went into it was machine learning stuff [...] makes you just think, "Oh, well, that's okay. That's cool. So, maybe machine learning can be more useful." I think before hearing that, my thoughts on machine learning were neutral to slightly negative, whereas hearing that some of the guidance that I've been using was machine learning would not make me feel badly. It would make me feel like, "Oh, well what else can machine learning do?" It would make it more positive.

As this forecaster suggests, for those who were more hesitant toward or skeptical of AI/ML, showing products first and then communicating that they are AI/ML may be more helpful and reassuring. This idea was further reinforced by another forecaster who thought using existing or familiar products to communicate how AI/ML was just another tool in the toolbox that they had already used could help ease the concerns of more skeptical forecasters:

> SF10: I could see that being a great way to introduce a training slide on this, like "Did you know that this model that you've already used that's been out for years has machine learning?" I think that would be a great way to start. I think that would be good for [those hesitant or skeptical of AI/ML] to see.

Overall, labeling a product as AI/ML does not appear to deter forecasters from reviewing it, but it raises different considerations and responses from forecasters.

4. Forecasters' familiarity with and openness to AI/ML

Forecasters' discussions of AI/ML revealed a range of familiarity and openness regarding AI/ML; however, there were notable patterns across the interviews that we synthesize in this section. We categorized all forecasters based on our interpretation of how familiar they were with AI/ML, as well as how open to AI/ML they appeared to be based on their responses. The categorizations in this section are mutually exclusive, meaning that each of the 29 forecasters fits into only one of the categories for each grouping.

a. Forecasters have a wide range of familiarity with AI/ML. We observed a wide range in how familiar forecasters reported being with AI/ML. There were generally four different types of forecasters with respect to their familiarity, which we have termed *Do not know much*, *Some but fuzzy*, *Experienced dabbler*, and *Formal training* (a short description and example quotations are provided in Table 2). As expected, there were some forecasters who reported being largely unfamiliar with AI/ML (*Do not know much*) and others who had a few ideas about how AI/ML worked or had heard just a bit about it but lacked a deep understanding or experience (*Some but fuzzy*). Conversely, two of the younger forecasters reported getting experience with ML in graduate school (*Formal training*), which may reflect a potential change in the NWS workforce if more forecasters begin their careers with AI/ML experience.

There was also a sizable group of forecasters who reported taking the initiative to learn more about AI/ML through reading papers, attending conference sessions, or searching online (*Experienced dabbler*). Another source of familiarity with AI/ML was from hands-on experience with or hearing about an AI/ML product. Models were often referred to generally, but those specifically mentioned by name were Pangu-Weather (Bi et al. 2023), GraphCast (Lam et al. 2022), Nadocast (Hempel 2024), and the CSU severe model (Hill et al. 2023). The mentions of these specific models by name happened primarily in the more recent interviews. Of note, positive past experience—whether direct personal experience or indirectly hearing about others' experience—appears to have strong, positive effects on how forecasters view the use of AI/ML in forecasting. Forecasters not only mentioned indirect or direct experience with the CSU severe model and Nadocast models by name but also mentioned AI/ML experience more generally. For example,

> SF01: I was actually part of the Hazardous Weather Testbed this past spring where they were using some Machine Learning for some algorithms. And I left that experiment with a positive impression on AI because I was seeing an improvement versus the version that we use now.

Interestingly, there was a shift from the earlier interviewees mentioning experience with AI/ML products in experimental-only contexts, such as the HWT, to some experience reviewing and using AI/ML products in real-world, operational settings. Although not the focus of this paper, this change over time illustrates that some AI/ML products are already making their way into operational settings, through different mechanisms. Further efforts are needed to understand and improve the research-to-operations transition, for AI/ML and otherwise.

b. Forecasters are open to the use of AI/ML in forecasting, but to different degrees. Overall, forecasters report being open to the use of AI/ML, and no one reported being opposed to the idea. The lack of opposition could, however, be the result of the interview setting in which forecasters may not have been comfortable expressing this opposition for fear of how it would be perceived by the interviewing researchers. Furthermore, there may have been a selection bias in our sampling process with those who are opposed to new or different forms of guidance opting out of participating in our interviews. Nonetheless, forecasters expressed different levels of openness to the use of AI/ML for forecasting. As with familiarity, we identified three different types of forecasters with respect to their openness, which we have termed *Highly supportive*, *Generally open*, and *Open, but*... Short descriptions and example quotations are provided in Table 3.

Some forecasters were highly supportive about the use of AI/ML in operational forecasting (*Highly supportive*). Within this group, there were a small number of forecasters who were extremely interested in AI/ML and could not wait for more of it to be used in operational contexts. These forecasters were not shy about sharing the ways they expected AI/ML to change forecasting for the better. For example, forecaster SF15 reported:

> SF15: Well, I will tell you I absolutely love [AI/ML], and I can't wait for more of it and more of it. I think it's one of the neatest and most opportunistic ways to improve forecasting going forward.

However, most forecasters fell into the remaining categories, with some reporting a general openness to and interest in the use of AI/ML but not reporting any more details or feelings

Table 3. Overview of the thematic codes for NWS forecasters' openness to the use of AI/ML in forecasting. Each thematic code (Open, but…, Generally open, and Highly supportive) is accompanied by a brief description and two exemplar quotations from forecasters to demonstrate how the thematic manifested in the interviews.

Openness	Description and exemplar quotations
Highly supportive	Description: Forecasters who said they were very supportive of using AI/ML
	SF11: [AI/ML is] where we need to go. I mean ultimately we're trying to get the best guidance out there because it gets us closer to the more likely solution of what reality is going to be
	CF02: I think it's great. I'm glad. It's obviously the next step in the evolution of how we process data. I think it, like everything, it's going to have positive negatives, and we're going to learn that with time. But I think it's great
Generally open	Description: Forecasters who said they were open to the idea without adding major caveats, but also did not communicate much more beyond that
	SF10: I mean anything to make—especially the models that we use better, I think that's a good thing. I think the model data itself is useful to us
	CF05: I think it'd be interesting if I could see how it would be utilized. I think it could be really interesting. I'm open to it. I mean, I can't say I've had a lot of experience with it though
Open, but…	Description: Forecasters who reported being open to the use of AI/ML, but had reservations or concerns associated with that openness
	SF12: From a standpoint of job security and those types of things [AI/ML replacing forecasters is] kind of a scary thought. But also it's an exciting idea to think about how that type of technology may be able to help us
	CF08: There's the human aspect where it probably is also going to take a lot of what my job is. I've seen some of the aspects of [AI/ML] and applying that into the forecasting world is, to me, pretty exciting. I just do not know if I've seen any of the fruits, or any early fruits, of it yet

beyond that (*Generally open*) and others adding some concerns or caveats to their openness (*Open, but…*). This last group of forecasters reported concerns, some of which we detail in section 5.

During the interviews conducted earlier in the study period, a common theme was that forecasters were excited about the prospect of AI/ML and were looking forward to seeing what it could do. Conversely, the interviews conducted later in the study period started showing that AI/ML had more fully arrived, which could be seen by forecasters providing more details about the role of AI in forecasting and more of their hands-on experience. The interviewees from later in the study period had more caveats to add to their openness to AI/ML, as represented by more forecasters falling into the *Open, but…* group. However, there were roughly even numbers of *Highly supportive* forecasters throughout the study.

5. What forecasters view as the positives and negatives regarding AI/ML and forecasting

Some forecasters were excited and intrigued by the potential advances that AI/ML may bring to operational meteorology. Many see the field of AI/ML as still developing but that it is "the future" of forecasting. This openness to AI/ML is paired with some important and relevant concerns, such as fears of losing their jobs or being removed from the forecasting process and overhyping what the models are actually able to do. In this section, we synthesize several key themes regarding what forecasters viewed as the positives and negatives of AI/ML (Table 4), which we then contextualize with specific quotations. Of note, these are perceptions and ideas that were shared in the interviews by forecasters with a range of experiences and openness to AI/ML, as we described in the previous section. Thus, some points in this section may not be directly connected to current AI/ML capabilities or may even represent potential disconnects between forecasters' perceptions and how AI/ML work.

a. What forecasters view as positives of AI/ML. Overall, most forecasters were excited and hopeful that AI/ML techniques would lead to guidance that performs better in some way than the products and tools they currently have access to. Some said this generally, talking

Table 4. Overview of the themes for what NWS forecasters view as the positives and negatives of AI/ML and forecasting.

Specific applications and implications of AI/ML that forecasters discussed
What forecasters view as positives of AI/ML
• Better-performing guidance • Enhanced pattern recognition across large amounts of data • Increased computational efficiency and reduced latency of model guidance • Increased spatial and temporal resolution of guidance and downscaling • Bias correction • Limiting forecasters' biases • Guidance that continually improves over time as it "learns" from more cases • Increased confidence in forecasts and improvements in their ability to message that would come with better and more efficient guidance
What forecasters view as negatives of AI/ML
• Not being able to catch extreme or rare events given the lack of cases models are trained on • Overreliance on AI/ML products beyond their application areas or training data • Have not seen the AI/ML products in action and would need hands-on experience before really evaluating how they feel about them • Might be too black boxed for some to feel confident using • Guidance might not be smooth over time, but rather jump around run to run • Replacing or removing forecasters from the forecasting process

mostly about being open to any guidance that helped produce a more accurate forecast and their confidence in these forecasts, while others provided more specific areas they thought AI/ML guidance could lead to improvements for. In some cases, like those we detail below, forecasters offered more details on why they thought AI/ML could have a positive or negative impact on forecasting, but others offered points like "downscaling" without elaborating more on what they meant.

The first positive was enhancing pattern recognition across large amounts of data. Many forecasters reported that AI/ML would be much better at sifting through the massive amounts of data that are now available to see patterns more effectively and efficiently than human forecasters could. For example, when asked why they said they might trust an AI/ML product, forecaster CF10 responded:

> CF10: Just because it's doing computations—it's recognizing patterns and learning from that. Which I try to do and forecasters try to do, but we're not that as good as maybe we think we are at it.

Relatedly, several forecasters were excited about the prospect of models that could run faster and more efficiently than their current guidance to decrease latency in getting the most up-to-date runs and data. Forecaster SF16 specifically discussed the potential for more up-to-date and timely data being an exciting aspect of AI/ML guidance:

> SF16: The one thing that I thought was really interesting [about AI/ML models] was it seems like you can get quicker run times—I don't know if that's the right word—from the machine learning models than running like an explicit HRRR. We're always looking for the latest model run. If we could get model output more often, that's something we're going to look at. That's something that we're going to be interested in.

Several forecasters also mentioned they thought AI/ML could be useful for increasing the resolution of guidance, as forecaster CF11 reported:

> CF11: As AI technology continues to improve, we can be more confident in the products that are being produced and at a higher resolution.

Another broad area of interest pertained to addressing different types of bias in forecasting. Some forecasters believed AI/ML could be a useful tool for bias correction. As forecaster CF06 describes this process:

> CF06: The AI is using a lot of weather models and then comparing it to surface observations to try to kind of develop which biases the models might hold and how to forecast it better.

Others saw it as a way to limit potentially negative biases that forecasters may have. Forecaster SF15 described how, in their opinion, AI/ML guidance lacks "bias" that human forecasters have:

> SF15: [AI/ML can produce] very scientifically sound solutions absent of human bias. That's what I like. There is so much human bias in forecasting.

This forecaster appears to view the AI/ML models as being more objective than human decision-makers. We caution on this point because it is largely acknowledged that AI/ML models carry biases, and the weather and climate domain is no exception (McGovern et al. 2022). For example, some of the forecasters who reviewed the storm mode probability product developed using a convolutional neural network noted that the human hand labeling of the storm model training images was a source of uncertainty given the known difficulty and subjectivity of categorizing storm mode (M. G. Cains et al. 2024). This notion is one that should be addressed in future discussion and training of AI/ML applications in the meteorological domain.

Interestingly, several forecasters also mentioned being intrigued by the idea of guidance that was always "learning" and improving over time. They saw AI/ML guidance as being able to continually learn from its past predictions and verification to improve its performance. As forecaster CF01 described:

> CF01: I mean machine learning is just like, "Oh, okay. It's learning about what it's doing and how to fix itself."

This relatively common idea that AI/ML products would continually learn and improve over time represents a potential tension point between what forecasters, as users, are expecting of AI/ML products and their current capabilities. Although this continual learning may be true of some AI/ML models, the current practice is generally to train a model initially with a subset of data, after which the model is used to make predictions based on the insights gained from that training data. However, it may be episodically retrained to maintain or improve stability of model performance, rather than being continually training the AI/ML model.

Last, many forecasters mentioned they were hopeful for any guidance, like AI/ML-derived ones, that could improve their forecasts and ability to message to their core partners and the public. They saw the efficiency and increased performance of AI/ML as a way to provide more time and confidence for this messaging, which we will discuss in more detail later on.

b. What forecasters view as negatives of AI/ML. In addition to the excitement and hopes, forecasters also had concerns and worries related to AI/ML. Several forecasters discussed concerns about AI/ML not being able to catch extreme or rare events given the lack of cases to train the models on. Relatedly, some forecasters were also concerned AI/ML models could be overrelied on, especially in cases beyond the models' application areas or training data. Forecaster CF09 discusses and contextualizes these concerns beautifully:

> CF09: I think it's important that [AI/ML] doesn't get overused or that it's important to understand what it's doing. Because machine learning is, for lack of a better word, can be really narrow in scope. It only understands the training set you've given in, and it's only valid for certain parameters, it's only valid within that certain training set, and I think it's really easy for those tools to become trusted and then taken out of the scope of where they are. And then there's no laws of physics within those tools restraining them to a scenario. If you take it out of something that was in the training set, you're trying to use it in different circumstances, it could totally fall apart. More so than a physics-based model could. So, that's my biggest fear, that kind of thing. And that's why I'm personally really skeptical. And we start talking about replacing physically based models totally with machine learning. Our biggest thing in weather forecasting is predicting the extremes, right? And I, personally, am going to trust a physics-based model when predicting the extremes versus the machine learning model that has one example in its whole test case for that county.

This quotation further shows this forecaster's complex and nuanced knowledge and perceptions of AI/ML. For instance, their mention of "no laws of physics" is a nod to the emergence of pure-AI or data-driven AI/ML models and the associated concerns about their potential limitations for operational forecasting.

Another prominent theme was forecasters adding the caveat that they would need to see AI/ML products in action and would need hands-on experience before really evaluating how they felt about them, a finding that is consistent with related research (M. G. Cains et al. 2024). Some forecasters, like forecaster CF10, mentioned that AI/ML tools might be too black boxed for some forecasters to feel confident using operationally. Furthermore, as forecaster CF10 also describes, verification is helpful, but they also want to know the inputs of model guidance and what it is or is not picking up on:

> CF10: I think there's the big risk of [AI/ML] being a black box, which forecasters don't trust. And I think [this interview] showed some ways to get more trust in [an AI/ML product] by looking at the ingredients and […] what the AI is picking up on. And you kind of got to do that. Otherwise it's really tough for the forecaster to trust it even if they do see the verification, which is useful. But yeah, you need a bit of a peek under the hood. That's kind of been the case with the NBM. Once people can peek under the hood a little more and see the ingredients that are going into it, there's more trust in it.

Beyond the more technical dimensions, one overarching concern is that forecasters will be replaced or removed from the forecasting process. Many acknowledged this was a personal concern they had felt, as forecaster CF08 reported:

> CF08: And then there's the human aspect where it's like, well [AI/ML] probably is also going to take a lot of what my job is.

Others reported the concern of being replaced was one they had heard but did not share, as forecaster SF10 described:

> SF10: I know some people—especially in the Weather Service—are like, "Oh, they're gonna take our jobs and everything's gonna be automated!" Well, I'm not that doom and gloom when it comes to [AI/ML].

As these quotations demonstrate, the fear of being replaced or removed from the forecasting process is very real and prominent in the minds of forecasters.

6. Forecasters' mission-driven orientation shapes their openness to AI/ML

Overall, there is a widespread and deep commitment to the best possible forecasts and services to uphold the organizational mission among forecasters, specifically to provide information and services for "the protection of life and property and enhancement of the national economy" (NWS 2024). Although forecasters mentioned the fear of replacement that AI/ML poses, they also clearly articulated that this fear did not outweigh their deep commitment to providing the best possible products and services to their partners and the public in order to uphold the organizational mission. The forecasters conveyed this point in different ways.

Even forecasters who had reservations about the use of AI/ML in forecasting were open to using the technology and seeing what improvements it could provide. As forecaster CF06 describes, the potential risks to job security are definitely real and present in their mind, but they are also motivated to see past these risks if the technology can help improve the services and communication the agency can provide:

> CF06: Throughout our field, there's always the question about job security, and in relation to models becoming better than forecasters. But at the end of the day, it just is about the shift in what our job really is, as the National Weather Service, being able to communicate, whether to

people, to emergency managers, to the state, all of that type of stuff. So, If it can help us do our jobs better, then in the long run it'll be good for meteorology.

Forecaster SF04 is similarly open to AI/ML, despite some skepticism, and discusses finding ways for the human expert and AI/ML to work together to produce the best possible outcomes:

SF04: I'm a little skeptical on what can be gained from [AI/ML]. But again we won't know until we try. So I am supportive of it. I think the human has some expertise. Humans always are going to play a part, but I do think there are some areas where AI will definitely have an advantage, and then there's areas where the humans still have the advantage. And trying to meld those two—it's interesting and challenging.

This notion is often referred to as "human–AI teaming" and is supported by a large body of research [National Academies of Sciences, Engineering, and Medicine (NASEM) 2022] and serves as a useful lens for understanding how to deliberately design the integration of guidance in ways that enhance human and technological strengths.

Forecasters also commonly discussed how improvements in guidance through approaches like AI/ML would likely change and shift what their jobs look like in the future. Many discussed how they expected AI/ML improvements to shift their roles to more communication with partners and the public. Several discussed how this shift was an exciting opportunity and one that would help them improve their services and products. For example, forecaster CF11 details how they see this moment as a great opportunity for change:

CF11: I think [AI/ML] offers enormous potential for our field, and it's going to change how our job works. I think within the next decade or two, my job as a forecaster is going to be completely different. As AI technology continues to improve, we can be more confident in the products that are being produced and at a higher resolution and at higher confidence. Again we're almost going to be like translators of science, [...] so we'll be able to spend more time communicating with our partners—people like the Coast Guard or other people involved in marine operations down in the bay—than we're gonna have to be poring over guidance.

No matter if forecasters were hesitant or excited about AI/ML, they agreed they were open to using whatever tools available to assist them in fulfilling the mission that motivates them. As forecaster SF12 describes:

SF12: The mission of the National Weather Service is to save lives and protect property, and I think all of us who take that mission to heart, we really do want to message our forecast with confidence, especially leading up to these higher-impact events so that people will take action. So if there's any advantage from new technology, with AI and others, then certainly I am all in favor of that because we want the public to be more confident in our forecast, not less confident. And so with that I say I'm all in favor of ways that we can improve, and AI seems to be a very good possibility to get there.

7. Conclusions

As AI/ML development continues to grow and data-driven AI/ML models ignite a type of arms race across the public and private sectors (e.g., Ebert-Uphoff and Hilburn 2023), we have provided an essential, in-depth analysis of how the NWS forecasters we interviewed view these tools and their use in operational forecasting. This work builds on a growing body of research on how forecasters make decisions and navigate a highly complex and challenging decision space (e.g., Demuth et al. 2020; Henderson et al. 2022; Novak et al. 2008;

Stuart et al. 2022). Through the research reported here, we have expanded this domain of research to include AI/ML (see also M. G. Cains et al. 2024). We further build on the work of M. G. Cains et al. (2024) by looking beyond specific AI/ML products to the broader context of how forecasters are thinking about AI/ML and operational forecasting more generally in their everyday decision-making. The intersection of decision-making and AI/ML is an important area for research (Bostrom et al. 2023), yet in the atmospheric science community, the attention has almost exclusively been on the development of AI/ML tools. In this paper, we have expanded previous work that focuses on evaluations of specific products to provide an overview of how forecasters, as a key frontline user group, perceive the technology and its role in operational forecasting. Several key results emerged from our study.

We demonstrated how the forecasters interviewed overwhelmingly prefer the term machine learning over artificial intelligence and how the reasons for this preference range from their familiarity with the use of ML in professional settings to their perceptions of how AI has been portrayed in the media. Furthermore, labeling products as AI/ML does not appear to deter most forecasters from reviewing the product, but it does prompt varied responses. Although some of the forecasters see AI/ML products as the exciting cutting edge of science, others care little about the development approach and more about how well the product verifies and helps them do their job.

We also found that the forecasters are quite varied in their familiarity with AI/ML. Forecasters range from knowing little to nothing about it to having had graduate courses on AI/ML. Many forecasters also mentioned being interested in learning more about AI/ML and expressed interest in having training modules or resources on the topic. Despite the differences in familiarity, the forecasters are open to using AI/ML tools operationally; however, the extent of this openness ranged from being highly supportive to having some important concerns about how effective AI/ML can be. Relatedly, the forecasters expressed their ideas about many positive and negative impacts of AI/ML on forecasting, including key areas where they saw the technology improving current capabilities (increased model efficiency, improved spatial and temporal resolution, etc.) and areas about which they expressed caution and concerns (inability for AI/ML to catch extremes, limited generalizability of AI/ML models, etc.). One of the major concerns raised by the forecasters was the fear of being replaced or removed from the forecasting process by AI/ML models. Nonetheless, the forecasters are open to whatever methods and tools can improve forecast communication and decision support services, particularly if it allows them more time and ability to message risks, uncertainty, and confidence. Figure 2 represents a high-level overview of these results.

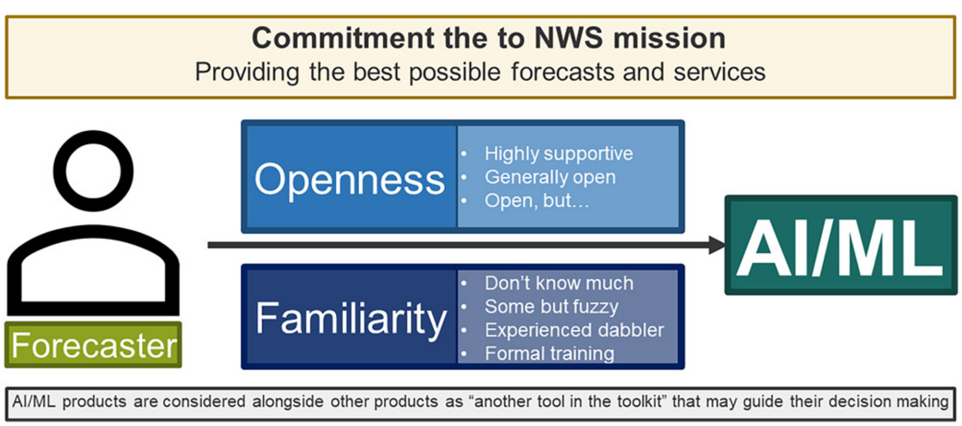

Fig. 2. A conceptual diagram representing how forecasters' perceptions of AI/ML are shaped by their commitment to the NWS mission and how the AI/ML tools are considered alongside other products they have access to. Forecasters' overall familiarity with an openness to AI/ML varies.

Forecasters' interests, concerns, and commitments related to AI/ML and forecasting represent an opportunity for NWS and NOAA to support the continued training of forecasters in ways that integrate the strengths of new technologies with the strengths of forecasters. Providing basic training on AI/ML for forecasters and clearly communicating how new technologies will or will not affect forecasters' roles in the weather enterprise would be important first steps for addressing concerns and moving toward an effective integration of AI/ML.

From an AI/ML development perspective, our results demonstrate some fruitful areas for future work. As we outlined in section 5, there are many areas in which the forecasters see AI/ML contributing to operational forecasting in positive ways. Specific opportunities for developers to better meet forecasters' needs exist around producing guidance that improves upon the performance of current guidance and their confidence in guidance output. Beyond performance, forecasters were also interested in the ways AI/ML could enhance their current guidance options by reducing the latency of model guidance, correcting model biases, and increasing model resolution. These represent areas where future work would be particularly well received by forecasters. A key point for developers to consider is that some forecasters also believed AI/ML models would be continually updating and learning with each forecast and that they were free from bias. Developers should be aware of and prepared to respond to these expectations among some forecasters.

Conversely, the negative dimensions represent areas where more work or more effective communication would be particularly valuable. Several forecasters commented on the importance of the training data for AI/ML guidance being adequate and clearly communicated. Specific concerns were the ability of AI/ML to capture extreme or rare events and issues with AI/ML being applied outside of the domains of which it was trained. The forecasters also communicated concerns about AI/ML potentially being too "black boxed" for them to feel confident using operationally. Furthermore, many of the forecasters reported they would need hands-on experience with AI/ML models to be able to meaningfully evaluate them. Last, there was a concern among forecasters that AI/ML might replace or remove them from the forecasting process, both personally felt and that they had heard from others. Addressing and speaking to both the positives and negatives we outlined can result in AI/ML products that can more effectively traverse the notorious valley of death from research to operations (NASEM 2000).

From an operational perspective, our results suggest that the forecasters interviewed are generally open to and ready for AI/ML. However, improved communication is needed about the vision of how AI/ML will be integrated into operational forecasting and how this will or will not affect forecasters. Similarly, the weather community, and NWS in particular, would strongly benefit from developing structured and systematic mechanisms to understand forecasters' perspectives, incorporate their needs, and expose and train them on these tools. These mechanisms should be codeveloped with forecasters to meet their needs without overburdening them.

In other words, we suggest that the narrative around forecasters and AI/ML ought not to be framed as questions about how to get forecasters ready for or onboard with AI/ML. Rather, the key questions should include how to effectively meet the needs, concerns, and hopes forecasters have around AI/ML given the context of their work environment and job roles and how to create a system that successfully integrates forecasters' domain knowledge and expertise with new AI/ML developments. This user-centered narrative is essential for approaching AI/ML development in a way that can improve the forecasts and services provided by forecasters and the broader weather enterprise.

It is further important to acknowledge that this space is rapidly evolving. Even during the course of our study, we noted differences in how forecasters were discussing AI/ML. For example, the forecasters in the severe interviews, which were conducted first, focused more on the abstract level, talking about being open to whatever led to the best forecast and most effective communication in any way, whereas the forecasters in the coastal interviews, which were conducted later, expressed more concrete ideas and visions for how AI/ML will affect their workflow—namely by shifting their responsibilities to more communication, especially when challenging events threaten their county warning areas. Despite these differences, there was a clear and unanimous commitment to providing the best possible forecasts and products to those they served.

Although our interviews focused specifically on AI/ML tools, the forecasters' deep commitment to the NWS mission is not unique to a specific tool but rather is an ethic that should be considered with any efforts to develop new or improved forecast guidance. Further, forecasters take a holistic approach to the broad range of information and guidance they have access to, of which AI/ML is just one. Moving forward, development and operational practices that speak to how new technologies will facilitate meeting the NWS mission in ways that emphasize and support the important role forecasters play can lead to more effective outcomes for the weather enterprise.

Acknowledgments. We are grateful for all the forecasters who took the time to do these interviews with us and for providing essential feedback on this work. We also thank the NWS regional management for their support of this research. We would like to acknowledge the support and insights from all of our colleagues from the NSF AI Institute for Research on Trustworthy AI in Weather, Climate, and Coastal Oceanography (AI2ES), especially those involved with developing the interview protocols. We would also like to acknowledge the Machine Integration and Learning for Earth Systems (MILES) group at the NSF National Center for Atmospheric Research for their valuable insights on early iterations of this work. This material is based on work supported by the National Science Foundation under Grant RISE-2019758. This material is based on work supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under the Cooperative Agreement 1852977.

Data availability statement. The interview data are not available in accordance with our Human Subjects Committee approval.

References

Ben Bouallègue, Z., and Coauthors, 2024: The rise of data-driven weather forecasting: A first statistical assessment of machine learning-based weather forecasts in an operational-like context. *Bull. Amer. Meteor. Soc.*, **105**, E864–E883, https://doi.org/10.1175/BAMS-D-23-0162.1.

Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, **619**, 533–538, https://doi.org/10.1038/s41586-023-06185-3.

Bogdan, R. C., and S. K. Biklen, 1998: *Qualitative Research in Education. An Introduction to Theory and Methods.* 3rd ed. Allyn and Bacon, 276 pp.

Bostrom, A., and Coauthors, 2023: Trust and trustworthy artificial intelligence: A research agenda for AI in the environmental sciences. *Risk Anal.*, **44**, 1498–1513, https://doi.org/10.1111/risa.14245.

Braun, V., and V. Clarke, 2006: Using thematic analysis in psychology. *Qual. Res. Psychol.*, **3**, 77–101, https://doi.org/10.1191/1478088706qp063oa.

Burke, A., N. Snook, D. J. Gagne II, S. McCorkle, and A. McGovern, 2020: Calibration of machine learning–based probabilistic hail predictions for operational forecasting. *Wea. Forecasting*, **35**, 149–168, https://doi.org/10.1175/WAF-D-19-0105.1.

Cains, M., and Coauthors, 2024a: "Interviews with NWS forecasters related to severe weather and new artificial intelligence/machine learning (AI/ML) guidance predicting severe hail and storm mode: Interview materials for "AI/ML" version", in Interviews with National Weather Service (NWS) forecasters related to severe weather and new artificial intelligence/machine learning (AI/ML) guidance predicting severe hail and storm mode. DesignSafe-CI, accessed 16 July 2024, https://doi.org/10.17603/ds2-8mgd-2j44.

——, and Coauthors, 2024b: "Interviews with NWS forecasters related to severe weather and new artificial intelligence/machine learning (AI/ML) guidance predicting severe hail and storm mode: Interview materials for "No AI/ML" version", in Interviews with National Weather Service (NWS) forecasters related to severe weather and new artificial intelligence/machine learning (AI/ML) guidance predicting severe hail and storm mode. DesignSafe-CI, accessed 16 July 2024, https://doi.org/10.17603/ds2-nfgd-3g25.

Cains, M. G., C. D. Wirz, J. L. Demuth, A. Bostrom, A. McGovern, D. J. Gagne, R. Sobash, and D. Madlambayan, 2024: Exploring NWS forecasters' assessment of AI guidance trustworthiness. *Wea. Forecasting*, **39**, 1219–1241, https://doi.org/10.1175/WAF-D-23-0180.1.

Chapman, W. E., L. D. Monache, S. Alessandrini, A. C. Subramanian, F. Martin Ralph, S.-P. Xie, S. Lerch, and N. Hayatbini, 2022: Probabilistic predictions from deterministic atmospheric river forecasts with deep learning. *Mon. Wea. Rev*., **150**, 215–234, https://doi.org/10.1175/MWR-D-21-0106.1.

Chase, R. J., D. R. Harrison, A. Burke, G. M. Lackmann, and A. McGovern, 2022: A machine learning tutorial for operational meteorology, Part I: Traditional machine learning. arXiv, 2204.07492v2, https://doi.org/10.48550/arXiv.2204.07492.

Clark, A. J., and Coauthors, 2023: The first hybrid NOAA Hazardous Weather Testbed Spring Forecasting Experiment for advancing severe weather prediction. *Bull. Amer. Meteor. Soc.*, **104**, E2305–E2307, https://doi.org/10.1175/BAMS-D-23-0275.1.

Clarke, V., and V. Braun, 2013: Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *Psychologist*, **26**, 120–123.

——, and ——, 2017: Thematic analysis. *J. Positive Psychol.*, **12**, 297–298, https://doi.org/10.1080/17439760.2016.1262613.

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.

Ebert-Uphoff, I., and K. Hilburn, 2023: The outlook for AI weather prediction. *Nature*, **619**, 473–474, https://doi.org/10.1038/d41586-023-02084-9.

Flick, U., 2014: Mapping the field. *The SAGE Handbook of Qualitative Data Analysis*, SAGE, 3–18.

Flora, M. L., C. K. Potvin, P. S. Skinner, S. Handler, and A. McGovern, 2021: Using machine learning to generate storm-scale probabilistic guidance of severe weather hazards in the warn-on-forecast system. *Mon. Wea. Rev.*, **149**, 1535–1557, https://doi.org/10.1175/MWR-D-20-0194.1.

Gagne, D. J., II, A. McGovern, S. E. Haupt, R. A. Sobash, J. K. Williams, and M. Xue, 2017: Storm-based probabilistic hail forecasting with machine learning applied to convection-allowing ensembles. *Wea. Forecasting*, **32**, 1819–1840, https://doi.org/10.1175/WAF-D-17-0010.1.

Glaser, B., and A. Strauss, 1967: *The Discovery Grounded Theory: Strategies for Qualitative Inquiry.* Aldine Transaction, 271 pp.

Hardwick, W. C., 1973: Monthly fog frequency in the continental United States. *Mon. Wea. Rev.*, **101**, 763–766, https://doi.org/10.1175/1520-0493(1973)101<0763:MFFITC>2.3.CO;2.

Heaven, W. D., 2023: The inside story of how ChatGPT was built from the people who made it. *MIT Technology Review*, 3 March, https://www.technologyreview.com/2023/03/03/1069311/inside-story-oral-history-how-chatgpt-built-openai/.

Hempel, B., 2024: Nadocast: Tornado probabilities via post-processing weather model outputs with machine learning. Github, accessed 1 March 2024, https://github.com/brianhempel/nadocast.

Henderson, J., J. Spinney, and J. L. Demuth, 2022: Conceptualizing confidence: A multisited qualitative analysis in a severe weather context. *Bull. Amer. Meteor. Soc.*, **104**, E459–E479, https://doi.org/10.1175/BAMS-D-22-0137.1.

Hill, A. J., R. S. Schumacher, and I. L. Jirak, 2023: A new paradigm for medium-range severe weather forecasts: Probabilistic random forest–based predictions. *Wea. Forecasting*, **38**, 251–272, https://doi.org/10.1175/WAF-D-22-0143.1.

Kamangir, H., W. Collins, P. Tissot, S. A. King, H. T. H. Dinh, N. Durham, and J. Rizzo, 2021: FogNet: A multiscale 3D CNN with double-branch dense block and attention mechanism for fog prediction. *Mach. Learn. Appl.*, **5**, 100038, https://doi.org/10.1016/j.mlwa.2021.100038.

——, E. Krell, W. Collins, S. A. King, and P. Tissot, 2022: Importance of 3D convolution and physics on a deep learning coastal fog model. *Environ. Modell. Software*, **154**, 105424, https://doi.org/10.1016/j.envsoft.2022.105424.

Korneeva, E., T. O. Salge, T. Teubner, and D. Antons, 2023: Tracing the legitimacy of Artificial Intelligence: A longitudinal analysis of media discourse. *Technol. Forecasting Soc. Change*, **192**, 122467, https://doi.org/10.1016/j.techfore.2023.122467.

Lam, R., and Coauthors, 2022: GraphCast: Learning skillful medium-range global weather forecasting. arXiv, 2212.12794v2, https://doi.org/10.48550/arXiv.2212.12794.

Lincoln, Y. S., and E. G. Guba, 1985: *Naturalistic Inquiry.* SAGE, 416 pp., https://play.google.com/store/books/details?id=2oA9aWlNeooC.

Marshall, C., and G. B. Rossman, 2014: *Designing Qualitative Research.* SAGE, 352 pp., https://play.google.com/store/books/details?id=-zncBQAAQBAJ.

Mayer, K. J., and E. A. Barnes, 2022: Quantifying the effect of climate change on midlatitude subseasonal prediction skill provided by the tropics. *Geophys. Res. Lett.*, **49**, e2022GL098663, https://doi.org/10.1029/2022gl098663.

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

Merriam, S. B., and E. J. Tisdell, 2016: Designing your study and selecting a sample. *Qualitative Research: A Guide to Design and Implementation*, Vol. 67, Wiley, 73–104.

NASEM, 2000: *From Research to Operations in Weather Satellites and Numerical Weather Prediction: Crossing the Valley of Death.* National Academies Press, 96 pp., https://doi.org/10.17226/9948.

——, 2022: *Human-AI Teaming: State-of-the-Art and Research Needs.* The National Academies Press, 140 pp., https://doi.org/10.17226/26355.

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.

NWS, 2024: About the NWS. National Weather Service, https://www.weather.gov/about/.

Ouchchy, L., A. Coin, and V. Dubljević, 2020: AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI Soc.*, **35**, 927–936, https://doi.org/10.1007/s00146-020-00965-5.

Pathak, J., and Coauthors, 2022: FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. arXiv, 2202.11214v1, https://doi.org/10.48550/arXiv.2202.11214.

Patton, M. Q., 2014: *Qualitative Research & Evaluation Methods: Integrating Theory and Practice.* SAGE, 832 pp., https://play.google.com/store/books/details?id=ovAkBQAAQBAJ.

Richards, L., 2020: *Handling Qualitative Data: A Practical Guide.* SAGE, 336 pp., https://www.torrossa.com/gs/resourceProxy?an=5018096&publisher=FZ7200.

Roebber, P., 2022: The Roebber Lectures. National Weather Service Virtual Lab Forum, https://vlab.noaa.gov/web/vlab-forum/the-roebber-lectures.

Roebber, P. J., and S. Smith, 2023: Prospects for machine learning activity within the United States National Weather Service. *Bull. Amer. Meteor. Soc.*, **104**, E1333–E1344, https://doi.org/10.1175/BAMS-D-22-0181.1.

Schumacher, R. S., A. J. Hill, M. Klein, J. A. Nelson, M. J. Erickson, S. M. Trojniak, and G. R. Herman, 2021: From random forests to flood forecasts: A research to operations success story. *Bull. Amer. Meteor. Soc.*, **102**, E1742–E1755, https://doi.org/10.1175/BAMS-D-20-0186.1.

Sobash, R. A., D. J. Gagne, C. L. Becker, D. Ahijevych, G. N. Gantos, and C. S. Schwartz, 2023: Diagnosing storm mode with deep learning in convection-allowing models. *Mon. Wea. Rev.*, **151**, 2009–2027, https://doi.org/10.1175/MWR-D-22-0342.1.

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

Wirz, C., and Coauthors, 2024a: "Interviews with National Weather Service (NWS) Forecasters related to coastal fog and new artificial intelligence/machine learning (AI/ML) guidance: Interview materials for "No AI/ML" version", in Interviews with National Weather Service (NWS) Forecasters related to coastal fog and new artificial intelligence/machine learning (AI/ML) guidance. DesignSafe-CI, accessed 16 July 2024, https://doi.org/10.17603/ds2-6wz9-zd32.

——, and Coauthors, 2024b: "Interviews with National Weather Service (NWS) Forecasters related to coastal fog and new artificial intelligence/machine learning (AI/ML) guidance: Interview materials for the "AI/ML" version", in Interviews with National Weather Service (NWS) Forecasters related to coastal fog and new artificial intelligence/machine learning (AI/ML) guidance. DesignSafe-CI, accessed 16 July 2024, https://doi.org/10.17603/ds2-3qzm-0464.