

RESEARCH ARTICLE

That's not what my app says: Perceptions of accuracy, consistency, and trust in weather apps

Cole Vaughn^{1,2}  | Kathleen Sherman-Morris² | Michael Brown² | Barrett Gutter²

¹National Oceanic and Atmospheric Administration, Global Systems Laboratory, Boulder, Colorado, USA

²Mississippi State University, Starkville, Mississippi, USA

Correspondence

Cole Vaughn, National Oceanic and Atmospheric Administration, Global Systems Laboratory, Boulder, CO, USA.
Email: cole.vaughn@noaa.gov

Funding information

Mississippi State University

Abstract

The usage of weather apps for forecast information has increased dramatically over the last 10–15 years. Ensuring that consumers value and trust weather apps is important to the integrity of weather forecasting. Public perception of weather app forecast accuracy and consistency undergirds the apps' value and trustworthiness. With app forecasts being interpreted solely by the app user, misunderstanding and consequent false expectations could jeopardize the public's perception of accuracy and consistency. Furthermore, weather apps often offer excessively—and potentially unrealistically—detailed forecasts on time and spatial scales, extending far into the future without sufficient disclaimers regarding the confidence level associated with such detailed forecasts. A survey of the public found perceived app accuracy and consistency to be positively correlated with the trust in an app. Participants indicated that they take at least modest consideration of uncertainty and spatial variability when assessing specific and longer range forecasts. On average, participants had low to moderate confidence in forecasts beyond 10 days, and a significant majority did not perceive a precipitation forecast as inaccurate, even when no rain occurred at their location, as long as it rained nearby. We tested for misinterpretation using a common expression of uncertainty in weather apps, namely probability of precipitation (PoP). A majority of participants made a correct interpretation of the two PoP values given, although, depending on the percentage, some misinterpreted the values as indicating precipitation intensity, totals, or duration. Overall, these findings offer encouragement for a society heavily reliant on weather apps while also encouraging more research on weather information interpretation.

1 | INTRODUCTION

Smartphone technology and the apps that run on these devices experienced significant uptake during the 2010s

(Pew Research Center, 2021). By around 2015, weather apps were becoming go-to sources for weather information alongside the traditional media like television (Hickey, 2015; Silver, 2015). In a study by Phan et al.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

(2018), college students ($N = 308$) attending one of three southeastern U.S. universities listed the weather app as their primary way of getting weather information, and four in five of those students claimed to use it every day. Although possessing a weather app does not indicate its use, roughly 90% of smartphone users have a weather app on their phone (Khamaj et al., 2019). The popularity of weather apps warrants research into not only their usage but also the public's perception of these apps and the factors that affect their perceptions.

2 | BACKGROUND

2.1 | Weather apps

Weather apps provide a variety of weather information including current and forecasted temperatures, precipitation, wind, humidity, sky conditions, and potentially more, depending on the app. Most weather apps have a radar, and some have forecast videos or articles for users to read (Figure 1). Users can set one or more locations for which to receive a forecast, and some apps have notifications that alert users with weather-related information that could potentially impact them.

Depending on the app, either computers or a combination of computers and humans create and input the forecast into the app (McGrath, 2019). The forecast display varies from app to app but most generally consist of a timeline that shows the forecast for the upcoming week or more (Figure 1). The level of detail that weather apps contain for each day of the forecast is again variable between apps, although some apps offer more precise details when a specific day or hour of the forecast is clicked on.

Most smartphone operating systems come with a basic weather app pre-downloaded on the phone; however, many people still choose to download a different weather app (Bryant et al., 2017). The Weather Channel and AccuWeather have been found to be two popular apps for users to download (Vaughn et al., 2023). While reasons for downloading a different weather app than the default one may vary, people expect higher accuracy and more information from an app that was not pre-downloaded on the phone (Bryant et al., 2017; Phan et al., 2018).

Convenience is a big reason people choose an app for forecast information, as apps provide immediate access to information (Nix-Crawford, 2017; Phan et al., 2018). This type of immediate access to the most up-to-date and



FIGURE 1 The image on the left is The Weather Channel's app (2023) interface showcasing the current conditions, today's forecast, and a weather news video. The middle image is the pre-downloaded app from an Apple iPhone (2023), which shows an extended forecast. The image on the right is the FOX Weather app (2023) showing the interactive radar display.

specific information is not generally available on television, at least not in the way that it is on an app, social media, or the Internet (Nix-Crawford, 2017). However, does this pursuit of convenience come with any costs regarding accuracy and public trust in and perception of weather apps?

2.2 | Relationship between forecast trust, accuracy, and consistency

The concept of trust has been studied in a wide range of contexts and scientific fields, and thus does not have a universal definition (Blomqvist, 1997). It, instead, involves different constructs depending on the field in which it is being studied (Blomqvist, 1997). Trust in relation to information has been defined as “reliance upon information received from another person about uncertain environmental states and their accompanying outcomes in a risky situation” (Schlenker et al., 1973, pp. 419). Applying this definition to a weather forecast, trust is revealed by a person’s willingness to rely on the forecast’s claims when making decisions about their life (whether in routine or severe weather contexts).

If a weather app forecast is to be considered valuable and useful, it has to be trustworthy. A forecast’s value and usefulness are based on the user having at least moderate confidence that the forecast information will be accurate and usable to improve decision making (Bryant et al., 2017; Demuth et al., 2011; Kay et al., 2015; Millner, 2008; Murphy, 1993; Voulgaris, 2019). The forecast must be worthy of the user’s reliance on it.

However, forecast users may be unaware of the true accuracy of the forecast data and instead base their judgment of accuracy on the “perceived accuracy” of the forecast—whether or not the consumer perceives the forecast to be accurate (Sherman-Morris, 2005). A forecast may have been “accurate” according to a forecaster using statistical analysis, but it may be interpreted as “inaccurate” by the consumer because they were measuring accuracy with two different standards (Murphy, 1993). For example, a forecast for rain may verify according to customary verification methods, but a forecast user may interpret the forecast as wrong if they did not observe rain at their location. Their version of accuracy originates in what they heard, what they then expected, and what they subsequently observed. Small deviations between observation and forecast are not expected to be noticed; however, large differences between the two, especially involving precipitation, are likely to be noted and perceived as inaccurate (Morrow, 2008; Murphy, 1993). Perceived accuracy is more subjective than accuracy because it is dependent on an individual’s own expectations and observations—

two things that are likely to vary from person to person. Thus, in the mind of the forecast user, perceived accuracy is accuracy.

While the maintenance of trust—and sometimes even the development of trust—is heavily reliant upon perceived accuracy, trust can also be developed from previous experience or relationships (Nix-Crawford, 2017; Wall et al., 2017) or the result of intensive interactions over time (Newton, 1997). Or, it can be more indirect and abstract, based on weaker yet regular associations with others (Newton, 1997). Examples of this in the context of weather and media include a person developing trust in a broadcast meteorologist because he or she regularly watches them (Sherman-Morris, 2005), or a person trusting a news organization because he or she recognizes or is familiar with its brand. A person may download a weather app from this news station because of that meteorologist or the brand recognition, meaning their initial trust is likely to be an extension of the existent trusting relationship prior to their assessment of the app’s accuracy.

There are three primary components often associated with trust—benevolence, integrity, and competence (Colquitt et al., 2007; McKnight & Chervany, 2001). Benevolence is caring for or acting in the trustor’s best interest, while integrity is acting in honesty and good faith (McKnight & Chervany, 2001). Integrity has been associated with consistency and reliability (Colquitt et al., 2007). Competency is having the ability to do something well (McKnight & Chervany, 2001). While benevolence may be more relevant when considering a human source, judgments about the consistency, reliability, and competence of a weather app may be reflected in perceptions about its accuracy and consistency.

When synthesizing this information to inform research into trust placed in weather apps, trust may initially be based on a previous relationship with a person or brand associated with an app, but the maintenance of that trust is likely to be based on regular perception of consistent, accurate forecasts over time. There is evidence of this in the literature where forecast accuracy has been suggested to impact trust in the forecast (Burgeno & Joslyn, 2020; Murphy, 1993). Interestingly however, failure to provide constant accuracy does not necessarily result in complete breach of trust (Keeling, 2011; Savelli & Joslyn, 2012). Many users still come back for another forecast even after an inaccuracy (Demuth et al., 2011). This may be because users expect there to be some error and uncertainty associated with forecasting (Savelli & Joslyn, 2012). Thus, while perceived accuracy is a main driver in keeping trust (Nix-Crawford, 2017), the accuracy does not have to be absolute. When studying perceived accuracy of weather apps in Australia, Bryant

et al. (2017) found that people generally perceived weather apps to be accurate. The current study analyzes this perception among weather app users in the United States.

Complementing accuracy, previous work has also shown that trust is impacted by forecast consistency (Losee & Joslyn, 2018; Murphy, 1993). Forecast consistency is defined in many differing ways. It has been referred to as the alignment between the forecast and what the forecaster actually thinks is going to happen (Murphy, 1993; Voulgaris, 2019). It has also been defined as the similarity of a message between two different sources or the uniformity of colors, symbols, and presentation between two different sources (Weyrich et al., 2019; Williams & Eosco, 2021). However, we define consistency as the similarity of the forecast from one forecast issuance to the next (Burgeno & Joslyn, 2020; Lashley et al., 2008).

Lashley et al. (2008) proposed that consistency is as significant as accuracy in keeping trust. And while the lack of consistency does result in lower trust, inaccuracy is found to be the more detrimental of the two when it comes to the forecast user's trust (Burgeno & Joslyn, 2020; Nix-Crawford, 2017).

2.3 | Weather app vulnerabilities to perceived inaccuracy and inconsistency

Having established that the perceived accuracy and consistency of a forecast are paramount to the trust and value assigned to it, we elaborate on potentially unique ways in which weather apps could detrimentally impact perceptions of accuracy and consistency.

Inaccuracy and inconsistency, or the perceptions thereof, can be found in all forecasts including those found on television. However, when the forecast is removed from television, the storytelling and context that accompany the forecast are often no longer present. Furthermore, it transfers the forecast interpretation from the broadcast meteorologist to the forecast user. A broadcast meteorologist—having at least some form of meteorological training—can interpret the forecast and then explain it to the viewer (Morrow, 2008). For weather app users, the responsibility to interpret the forecast falls on themselves, with little to no context on the confidence in and reliability of forecasts with high spatial and temporal resolutions as commonly found in weather apps. The forecast user may interpret the forecast differently than intended (Joslyn et al., 2009; Losee & Joslyn, 2018; Zabini et al., 2015), resulting in a higher likelihood of it being misunderstood (Zabini et al., 2015). A misunderstanding of the forecast can lead to false expectations, which can

lead to perceptions of inaccuracy when those false expectations do not verify.

Forecasts have been available away from the television format for decades. This is not a new development consequent to the invention of weather apps. However, the widespread applicability of this development has grown substantially in recent years. Television used to be the most common medium by which people acquired a weather forecast (Demuth et al., 2011). However, this has changed over the last 15 years, and the weather app is now considered to be the most popular source (Vaughn et al., 2023).

A common source of misunderstanding is found in the communication of uncertainty, which has a striking ability to create perceptions of inaccuracy in any forecast (Wall et al., 2017), no matter what medium it is taken from. However, because the app puts the interpretation of that uncertainty onto the forecast user, the forecaster does not have the ability to explain the intricacies of that uncertainty (Morrow, 2008).

Communication of uncertainty is common in weather forecasting. A simple example of uncertainty quantification used in most weather apps is the probability of precipitation (PoP), which expresses the uncertainty in the likelihood of precipitation (Zabini, 2016). Prior study has shown that forecast users prefer the use of PoPs (Morss et al., 2008) and that their use is associated with higher trust (Grounds, 2016). However, this does not mean that the percentage chance of rain given is being interpreted the way it was intended. In fact, research suggests that individuals tend to interpret the chance of rain in their own way (Morss et al., 2008). Though users may not grasp the concept of a 20% chance of rain, they can grasp the number 70 on a scale of 1 to 100. Percentages can serve as a sort of “code” or scale to define uncertainty (Zabini et al., 2015). They may understand that this is a “high” chance of rain, but they may also mistake it as meaning a long rain event or even one that will produce high rainfall totals (Joslyn et al., 2009; Zabini et al., 2015). The wide array of interpretations alone can lead to false expectations and consequent perception of inaccuracy.

Weather apps also have the tendency to forecast for a time period or at a level of specificity for which reasonable accuracy cannot be expected. Weather apps provide hyperlocal forecasting—a forecast that is given for a specific town or even sometimes a specific GPS location—while television tends to give a forecast for a metropolitan area or region (Zabini, 2016). The high specificity of a forecast for a specific town may not adequately capture or communicate the fact that weather can be highly variable spatially. We investigate whether participants in this study account for this spatial variability.

Aside from being too specific spatially, weather apps also tend to have too much specificity on a temporal scale. In some cases, hourly forecasts are available for 5–10 days ahead (Du et al., 2018; Zabini, 2016). The 2023 version of two of the most popularly downloaded apps as identified in Vaughn et al. (2023) offer a 15-day and 45-day forecast, respectively. While seasonal forecast decisions may be made during this timeframe, are daily forecasts at this time range useful—much less accurate? We add additional context to this research by studying people's confidence levels based on forecast length.

This research examines the American public's perceived accuracy and consistency of their preferred weather app and how those variables relate to forecast trust. This study also examines some ways in which weather apps are vulnerable to the promotion of perceived inaccuracy and inconsistency. The following research questions were the focus of this study:

RQ1: Are the perceived accuracy of and trust in a weather app correlated?

RQ2: Are the perceived consistency of and trust in a weather app correlated?

RQ3: How does the public interpret the quantification of uncertainty in a weather app?

RQ4: Does the public consider spatial variability when getting a weather app forecast?

3 | METHODS

This project used a common survey with Vaughn et al. (2023). The survey (supplemental material) used Likert-style questions with five answer choices to gauge the public perceptions of weather app accuracy, consistency, and trust. After converting the answer choices to 1–5 interval data (1 = lowest, 5 = highest), the relationship between perceived accuracy and trust as well as perceived consistency and trust was analyzed with Spearman correlation. A question also asked participants to rate the accuracy of weather apps in general. The mean ranks of the accuracy rating for weather apps in general and the specific app the participant uses were compared using the Wilcoxon signed-rank test. Respondents were also asked about their confidence levels (“very low” to “very high”) in the forecast at different time intervals—1 day out, 3 days out, 5 days out, 7 days out, and 10 days out. Each participant was asked only three of the possible five questions in a random order to help avoid any ordinal bias. The data was then recoded as 1–5 interval data (1 = very low, 5 = very high), and the Friedman test was used to compare the mean confidence rating between the different

days. This was done in order to understand the public's confidence in forecasts of varying length. Additional questions were included to evaluate consideration of spatial variability in the forecast and the interpretation of uncertainty information to help understand possible sources of misinterpretation of the forecast.

The survey sample ($N = 600$) was collected using Prolific, a company with registered survey takers who predominantly participate in research surveys for compensation. Five people reported not having a smartphone, and 32 people reported not using a weather app (Tables S1 and S2). The resultant sample size for most questions of the survey was 563 participants, 386 of which use a weather app at least daily (Table S2). A sample representative of the U.S. demographics was requested from Prolific, although the sample varied from this standard on age (younger), education level, and race and ethnicity (Table 1). Participants were recruited from all over the United States (Figure 2), with 41 of the states and the District of Columbia being represented. While the sample is inclusive of all age brackets, the higher concentration of responses among younger age brackets provides greater detail regarding the weather app perceptions of age ranges with the most users (Vaughn et al., 2023).

4 | RESULTS

4.1 | Perceived accuracy of weather apps and trust

Slightly more than half of respondents rated their weather app as having “high” accuracy; that went up to 70% of the sample when combined with those who answered “very high” (Table S3). The mean rating for perceived accuracy of the specific weather app the participant used (3.81, $N = 563$, Table S3) was greater than the mean rating for perceived accuracy of weather apps in general (3.70, $N = 561$, Table S4) although not by much. The Wilcoxon signed-rank test indicated a significant difference ($Z = -5.40$, $p < 0.001$), and thus respondents thought that the weather app that they specifically use was more accurate than weather apps in general.

Participants were also asked to rate their trust on a 5-point Likert scale (1 = no trust, 5 = highest level of trust, Table S5). When checking for the association between trust and perceived accuracy, a one-tailed Spearman correlation was significant, and the correlation was high [$r_s(557) = 0.766$, $p < 0.001$]. Thus, the greater the perceived accuracy of a weather app, the greater the trust a person puts in the app.

TABLE 1 Survey sample demographics versus 2020 U.S. Census demographics of ages 18+ for gender (U.S. Census Bureau, 2020), race and ethnicity (U.S. Census Bureau, n.d.a), education level (U.S. Census Bureau, 2021), urban/rural living environment (U.S. Census Bureau, n.d.b), and age (U.S. Census Bureau, 2020).

	Demographic characteristics	Survey participants	2020 Census data
Gender, $N = 600$	Male	289 (48.2%)	48.5%
	Female	292 (48.7%)	51.5%
	Transgender male	3 (0.5%)	-
	Transgender female	1 (0.2%)	-
	Gender variant/non-conforming	14 (2.3%)	-
	Prefer not to identify	1 (0.2%)	-
	Race and ethnicity ^a , $N = 600$	White	424 (70.7%)
Black or African American		74 (12.3%)	13.0%
Asian		39 (6.5%)	6.2%
Hispanic or Latino		39 (6.5%)	16.6%
Mixed race		19 (3.2%)	2.0%
Middle Eastern or North African		3 (0.5%)	-
American Indian or Alaska Native		1 (0.2%)	1.2%
Other		1 (0.2%)	-
Education level, $N = 600$	Some or no high school	7 (1.2%)	6.5%
	High school graduate	90 (15.0%)	27.8%
	Some college	189 (31.5%)	17.5%
	Associate's degree	53 (8.8%)	10.1%
	Bachelor's degree	176 (29.3%)	22.1%
	Advanced Degree	82 (13.7%)	12.7%
Urban/ rural living area, $N = 600$	Urban area	183 (30.5%)	Urban
	Suburban area	320 (53.3%)	80.0%
	Rural small town	65 (10.8%)	Rural
	Rural outside of town	29 (4.8%)	20.0%
	Not sure	3 (0.5%)	
Age, $N = 600$	18–29	334 (55.7%)	20.7%
	30–39	147 (24.5%)	17.4%
	40–49	58 (9.7%)	15.7%
	50–59	46 (7.7%)	16.3%
	60+	15 (2.5%)	30.0%

^aMore than one choice was possible for race and ethnicity in both survey and census, although no person in the survey picked more than one choice.

4.2 | Perceived consistency of weather apps and trust

Participants were asked how often their weather app tends to make big jumps in the forecast—how often they are noticing large changes in the forecast from issuance to issuance (Table S8). An ordinal scale was used ranging from “never” to “almost always”, and 82.5% fell in the “sometimes” or “seldom” categories. However, 15.5% said that their app “often” or “almost always” made big jumps in the forecast. This question measured inconsistency as opposed to

consistency, as it seemed easier to measure and explain to participants.

The results of the one-tailed Spearman correlation for perceived inconsistency and trust revealed a weak negative association between the two [$r_s(557) = -0.215$, $p < 0.001$]. While the questions inquiring about perceptions of accuracy and consistency are not necessarily equally comparable in their measurement of each construct, these correlation results offer support to the findings of previous research that perception of accuracy may be the stronger predictor of trust than perception of consistency (Burgeno & Joslyn, 2020; Nix-Crawford, 2017).

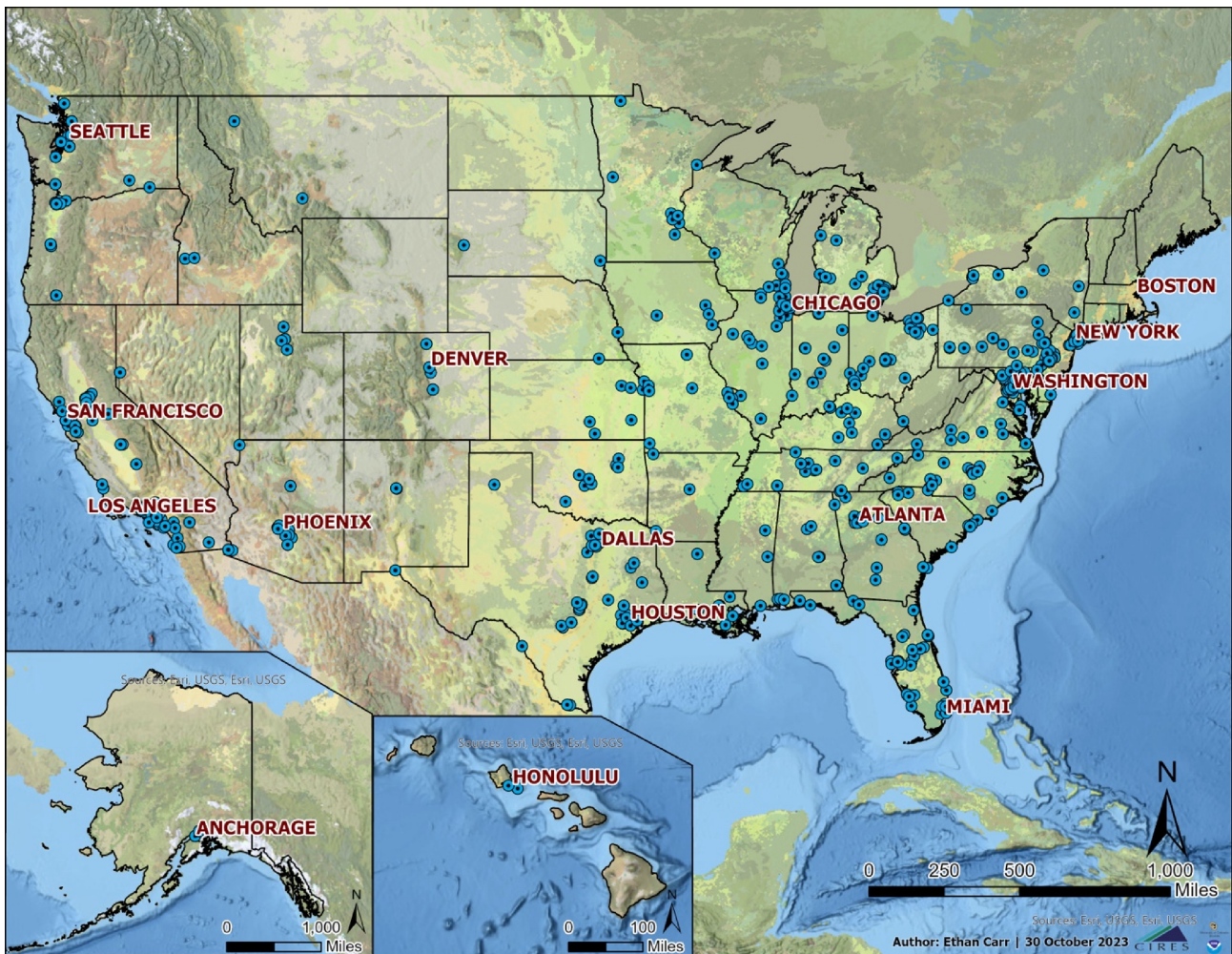


FIGURE 2 Each dot on the map indicates a survey participant's location.

4.3 | Public interpretation of uncertainty in weather apps

For the questions regarding participants' confidence levels for forecasts at different time intervals, the mean rating for each question was calculated and compared using a Freidman test (Tables S9-S13). Figure 3 shows the mean confidence rating decreased with time (Day 1 = 4.07, $N = 374$; Day 3 = 3.54, $N = 358$; Day 5 = 3.09, $N = 350$; Day 7 = 2.83, $N = 359$; Day 10 = 2.54, $N = 359$). The Freidman test result was statistically significant, indicating that there was a significant difference between at least some of the means ($\chi^2 = 497.39$, $p < 0.001$). A Wilcoxon signed-rank test was then run post hoc on each of the consecutive relationships (i.e., Day 1 vs. Day 3, Day 3 vs. Day 5, etc.). A Bonferroni adjustment was used to account for the repeated comparisons made using the Wilcoxon signed-rank test. Since four comparisons were being made, the p -value required for significance fell to 0.0125. The mean confidence rating of day 1 was significantly higher than that of day

Mean Confidence Rating in Forecast

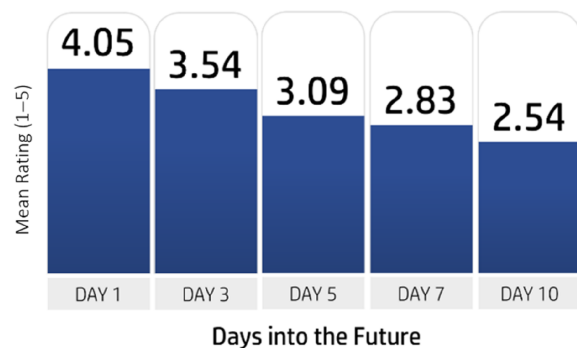
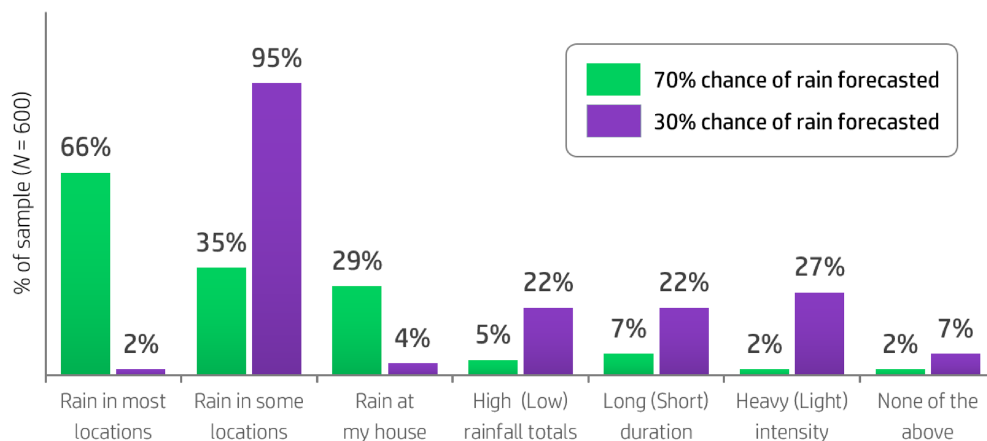


FIGURE 3 Mean rating of confidence in the forecast for Days 1-10.

3 ($Z = -8.138$, $p < 0.001$). The same trend was observed for the other comparisons made (Day 3 vs. Day 5: $Z = -6.663$, $p < 0.001$; Day 5 vs. Day 7: $Z = -4.155$, $p < 0.001$; Day 7 vs. Day 10: $Z = -4.119$, $p < 0.001$). Thus, the public's confidence in a forecast decreases as



What does the public expect for a 70% chance of rain vs. 30% chance of rain?
(Multiple selections possible)

FIGURE 4 Green bars represent the percentage of the sample that answered in each category when asked about a 70% chance of rain. The purple bars represent the percent of the sample that answered in each category when asked about a 30% chance of rain.

time goes by, which seems to indicate that the public understands that there is more uncertainty in the forecast with time.

Additional understanding of uncertainty interpretation was pursued. Two questions were posed to participants asking about two different scenarios—one where a weather app forecasted a 70% chance of rain and another where an app forecasted a 30% chance of rain. Each percentage was equidistant from 50%, with one representing a higher chance of rain and the other representing a lower chance of rain.

Participants were asked to check all responses from the answer choices that they expected to occur in each situation. The possible responses represented the areal coverage of rain, rain at a specific location, the rainfall totals, the duration of rain, and the intensity of the rain fall. For the 70% chance of rain question (Figure 3, Table S6), 66.4% of people answered correctly that most locations in the area would get rain, and 29.1% of people expected rain at their house. The percentage of people who chose responses related to rain totals, duration, or intensity was less than 7% for each.

The question for the 30% chance of rain (Figure 4, Table S7) yielded different results. Less than 2% of participants thought that most locations in the area would get rain, and 94.8% correctly thought that “some locations” would get rain. Only 3.5% of the sample thought it would rain at their house. Interestingly, the frequency of responses relating to rain totals, duration, and intensity increased rather dramatically, from 22% to 26.9% of the sample.

4.4 | Public consideration of spatial variability in weather apps

Participants were told to recollect the last time that their weather app forecasted rain but it did not rain at

their location. They were asked whether it rained nearby. After answering yes, no, or unsure, they were asked if they thought the forecast was accurate, inaccurate, or neither for that day. These questions worked together to understand participants' consideration of spatial variability of weather when evaluating whether the forecast was accurate.

Three-hundred and thirty-four people said that it had rained nearby (Table S14). Of those people, 71.3% said the forecast was accurate for that day, and another 19.2% said that it was neither accurate nor inaccurate (Table S15). Thus, even though it did not rain at their house, 90.5% of the people would not say that the forecast was inaccurate.

Now, for those who did not get rain and there was no nearby rain, 26.8% still said the forecast was accurate while 68.3% said it was inaccurate. Of those who were unsure if it rained nearby, 48.4% said the forecast was neither accurate nor inaccurate, and 29.3% said it was accurate. Overall, even when a person did not get rain at their house—regardless of whether it rained nearby or not—53.9% still said the weather forecast was accurate. Based on these results, the public does seem to be considering at least some spatial variability and uncertainty when considering a forecast's accuracy and verification.

Two questions were then used to understand whether the consideration of spatial variability has changed as consumers have trended away from television forecasts and toward app forecasts. Participants were asked what geographic area they thought a forecast covered when it came from a weather app or from television. The choices consisted of a range that grew in spatial coverage including “your specific location”, “your town”, “your county”, and “your county and the neighboring counties”. Only 22.3% of respondents thought that an app forecast was for their specific location, and a majority (51.5%) thought the forecast was for their town (Table S16).

In contrast, when asking about the geospatial extent of a television weather forecast, 44.7% said it was for their county and the neighboring counties (Table S17). However, 24% still said it was for their town. This seems to indicate that some of the public understands that a forecast on television is for a broader locale, even though the extended forecast near the end is typically for the main city where the news station is located. They also seem to understand that weather app forecasts tend to be more location-specific than television forecasts. Thus, with the weather app becoming the dominant medium for getting a weather forecast, the spatial variability of weather that is being considered when evaluating forecast verification may have decreased from a time when the television was the main source for a weather forecast.

5 | DISCUSSION

With the growth in popularity of weather apps, they now serve as a mediator between forecasters and the public for much of society. Similar to the findings of Bryant et al. (2017), most participants in this study considered their weather app to be highly accurate, which is notable given that the value a forecast holds is largely based on its accuracy (Bryant et al., 2017; Demuth et al., 2011; Kay et al., 2015). This study also found that the trust placed in a weather app and perceived accuracy of an app are highly correlated. However, it should be noted that the present study did not attempt to measure the actual accuracy of any of the weather apps used.

Perceived consistency of an app forecast was also related to consumer trust in the app. Lack of consistency was sometimes noticed by survey participants and it was found to negatively impact their trust. It should be noted that our measure and discussion of perceived consistency does not refer to consistency from one forecast issuance to the next. This would not fully address perceived consistency by the public, as a person may not check the weather forecast after each issuance. Thus, consistency to them—and the definition we used—is the similarity of the forecast between the forecast issuances that they observe. It is also important to mention that forecast consistency and accuracy can sometimes be mutually exclusive. In the event that a big change in the forecast is necessary to increase accuracy with the advent of new information, consistency may not be achievable. Sacrificing a more accurate forecast to lessen the negative impact on consistency could be detrimental. However, pursuance of accuracy and consistency in weather app forecasts should still both be prioritized to the extent possible.

This study found that the public's confidence in a forecast wanes the further out the forecast extends.

A forecast for 10 days out received a mean confidence rating between low and moderate. This implied questionable confidence in the whole forecast for that day, much less any spatially or temporally high-resolution details that the forecast may contain. Yet, Zabini (2016) found that over 50% of the weather apps they analyzed had forecasts that extended between 10 and 15 days out. While some weather apps may broaden their level of detail on specific forecast information for longer range forecasts, this was not the case for the two most popularly downloaded weather apps mentioned by Vaughn et al. (2023). Presentation of forecast information far into the future may provide the appearance that forecasters are confident enough to give specific details for very far into the future without acknowledging any actual confidence levels. Survey respondents knew to be hesitant of this, expressing at best moderate confidence in the forecast that far out. This lack of high confidence may contribute to Myers's (2019) finding that people do not make decisions based on forecasts at that range. If decisions are not being made for that time period, and a forecast's value is rooted in its ability to enhance decision making (Millner, 2008; Voulgaris, 2019), the need for formal 10-to-15-day forecasts is drawn into question. The weather forecasting community must re-evaluate whether these forecasts are necessary and wise and whether their motivation is rooted in science or in commercialism—offering more than the competition—as suggested by Morrow (2008) and Demuth et al. (2009).

In an age of hyperlocal and highly personalized content, where a smartphone's location-based services are incorporated into every app and every search, the weather app industry seemed to have no other option but to join the trend. Weather apps can now provide a forecast for every user based on their location. While convenient and accessible, weather is spatially variable and may differ between locations that are even short distances apart. Providing point-specific forecasts can risk communicating that a certain weather condition will be present at a specific location, not that it will just be present in the area. When evaluating whether a forecast is verified, does a person consider what went on around them as opposed to just what happened at their location? According to this study's results, it appears at least some people do.

While both media can provide regional forecasts and point-specific forecasts, the survey consensus was that a television forecast is for a region and an app's forecast is for a town. Matched with the understanding that weather apps have become the dominant forecast medium over television, this does lend credence to the idea that the public's consideration of weather's spatial variability may be decreasing in spatial extent. However, these results also indicate that a majority of people assume their app's

forecast is for their town as a whole instead of a specific location. This finding further supports the idea that at least some people factor in the spatial variability of weather as they determine if a forecast verified. Further research will be necessary to determine the intricacies of this consideration and how far it extends spatially.

Found in most weather apps (Zabini, 2016), PoP is an obvious example of uncertainty quantification. The results of this study showed that interpretation of a percentage varied between two examples—30% and 70%. Simply changing the percentage changed the expectations for what it meant. When respondents were asked about their interpretation of a 70% chance of rain, most thought it had something to do with what area and locations would get rain (e.g., most locations would get rain, some locations would get rain, or it would rain at their house). While this finding still held true when respondents were asked about a 30% chance of rain, significantly more people also made assumptions about the expected rainfall duration, totals, and intensity for 30%. This excellently illustrates the findings of Morss et al. (2008) that forecast users interpret PoP in their own way. It also lends credence to the idea of Zabini et al. (2015) and Joslyn et al. (2009) that rainfall totals and duration may be perceived simply based on the PoP value.

This finding concludes that an objective measure like PoP can be subjectively interpreted (or misinterpreted), indicating that the vast efforts exploring the communication and interpretation of probabilistic information should continue (Ripberger et al., 2022), potentially even focusing on best practices for communicating probabilistic information on weather apps. Understanding interpretation by the public is vital to having appropriate messaging that avoids communicating inaccurate expectations. This should be a priority considering the ubiquity of PoPs, especially in weather apps.

6 | CONCLUSION

With weather apps having become an increasingly normal and popular way to get a weather forecast, this study paid deserved attention to their perceived accuracy and consistency and the relation between those variables to trust in the apps. Given the correlations that accuracy and consistency had with trust, the high perceived accuracy ratings from this study's participants are encouraging, while the mediocre forecast consistency ratings caused some concern. The correlations observed in this research indicate that creating weather apps with both high accuracy and high consistency is important to the future of weather forecasting. This study found generally positive results regarding the interpretation of PoP and

consideration of spatial variability while also providing recommendations for future research in these areas, especially in relation to weather apps. If the public's view of weather forecasting now rests heavily on the shoulders of a computer interface, it is vital that research continues to ensure that these apps are helping to advance forecasting and maintain scientific integrity.

7 | LIMITATIONS

With the rapid development and advancement of technology, the findings of this research may not be generalizable to weather apps in the future or even those in the past. Continual testing and research of this topic will be necessary to keep up with the technological progression. Furthermore, in this research, all weather apps are treated as generally equal in terms of format, information provided, and so on. A previous study has shown that a majority of people use one out of a small group of weather apps (Vaughn et al., 2023), and these apps are all relatively similar in their format, function, and services offered. However, treating all weather apps as one is not fully realistic, as there is variation between the multitude of weather apps on the market. This research only asked participants about the accuracy of apps in general. Yet they were not asked about the accuracy of different components of the forecast (e.g., temperature, precipitation, etc.) individually. There are also many different types of uncertainty information. This research only asked participants about forecast confidence at different time intervals and about PoPs. These are very tangible expressions of uncertainty that were easily measured, but they do not represent all forms of uncertainty. Future work should explore how uncertainty in the forecast can be better communicated through weather apps. Finally, when participants were asked to recall a time when their weather app forecasted rain for the spatial variability section of the survey, recall bias may have impacted these results.

AUTHOR CONTRIBUTIONS

Cole Vaughn: Conceptualization (equal); data curation (equal); formal analysis (equal); funding acquisition (equal); investigation (equal); methodology (equal); project administration (equal); resources (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (equal); writing – review and editing (equal). **Kathleen Sherman-Morris:** Project administration (equal); supervision (equal); writing – review and editing (equal). **Michael Brown:** Supervision (equal); writing – review and editing (equal). **Barrett Gutter:** Supervision (equal); writing – review and editing (equal).

ACKNOWLEDGEMENTS

We thank Dr. Philip Poe for his guidance and advice on this research and to the Mississippi State University Department of Geosciences for funding the survey. We also thank Dr. Jamie Vickery from NOAA GSL for reviewing this article prior to submission and Ethan Carr from the Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA GSL for providing the GIS map.

DATA AVAILABILITY STATEMENT

The survey data used to conduct this research is included in the supplemental material. For further questions regarding the data, you may contact the lead author at cole.vaughn@noaa.gov.

FUNDING INFORMATION

The Mississippi State Department of Geosciences funded the survey for this research.

ORCID

Cole Vaughn  <https://orcid.org/0000-0003-4606-8265>

REFERENCES

- Blomqvist, K. (1997) The many faces of trust. *Scandinavian Journal of Management*, 13(3), 271–286. Available from: [https://doi.org/10.1016/S0956-5221\(97\)84644-1](https://doi.org/10.1016/S0956-5221(97)84644-1)
- Bryant, M.J., Wilde, S.J. & Smart, W.J. (2017) Taking the weather with you: user acceptance, trust and value of weather apps on smartphones. *International Journal of Social and Humanistic Computing*, 2(3–4), 247–260. Available from: <https://doi.org/10.1504/IJSHC.2017.084758>
- Burgeno, J.N. & Joslyn, S.L. (2020) The impact of weather forecast inconsistency on user trust. *Weather, Climate, and Society*, 12(4), 679–694. Available from: <https://doi.org/10.1175/WCAS-D-19-0074.1>
- Colquitt, J.A., Scott, B.A. & LePine, J.A. (2007) Trust, trustworthiness, and trust propensity: a meta-analytic test of their unique relationships with risk taking and job performance. *The Journal of Applied Psychology*, 92(4), 909–927. Available from: <https://doi.org/10.1037/0021-9010.92.4.909>
- Demuth, J.L., Lazo, J.K. & Morss, R.E. (2011) Exploring variations in people's sources, uses, and perceptions of weather forecasts. *Weather, Climate, and Society*, 3(3), 177–192. Available from: <https://doi.org/10.1175/2011WCAS1061.1>
- Demuth, J.L., Morrow, B.H. & Lazo, J.K. (2009) Weather forecast uncertainty information: an exploratory study with broadcast meteorologists. *Bulletin of the American Meteorological Society*, 90(11), 1614–1618. Available from: <https://doi.org/10.1175/2009BAMS2787.1>
- Du, H., Jones, P., Segarra, E.L. & Bandera, C.F. (2018) Development of a REST API for obtaining site-specific historical and near-future weather data in EPW format. In: *Building simulation and optimization 2018*. Cambridge: International Building Performance Simulation Association England. Available from: <https://orca.cardiff.ac.uk/id/eprint/111475/>.
- Grounds, M. (2016) Communicating weather uncertainty: an individual differences approach. Dissertation. University of Washington, p. 141, Available from: <https://digital.lib.washington.edu/researchworks/handle/1773/38200>.
- Hickey, W. (2015) Where people go to check the weather. FiveThirtyEight. Available from: <http://fivethirtyeight.com/datalab/weather-forecast-news-app-habits/>.
- Joslyn, S., Nadav-Greenberg, L. & Nichols, R.M. (2009) Probability of precipitation: assessment and enhancement of end-user understanding. *Bulletin of the American Meteorological Society*, 90(2), 185–194. Available from: <https://doi.org/10.1175/2008BAMS2509.1>
- Kay, M., Patel, S.N. & Kientz, J.A. (2015) How good is 85%? A survey tool to connect classifier evaluation to acceptability of accuracy. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 347–356. Available from: <https://doi.org/10.1145/2702123.2702603>
- Keeling, S. (2011) Weather forecasts—a matter of trust. *Geography*, 96(16), 16–21. Available from: <https://doi.org/10.1080/00167487.2011.12094304>
- Khamaj, A., Kang, Z. & Argyle, E. (2019) Users' perceptions of smartphone weather Applications' usability. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 2216–2220. Available from: <https://doi.org/10.1177/1071181319631098>
- Lashley, S.L., Fisher, L., Simpson, B.J., Taylor, J., Weisser, S., Logsdon, J.A. et al. (2008) Observing Verification Trends and Applying a Methodology To Probabilistic Precipitation Forecasts At A National Weather Service Forecast Office. Available from: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8b65f165a67c3b24e62e1871bd250514cb766d5b>.
- Losee, J.E. & Joslyn, S. (2018) The need to trust: how features of the forecasted weather influence forecast trust. *International Journal of Disaster Risk Reduction*, 30, 95–104. Available from: <https://doi.org/10.1016/j.ijdrr.2018.02.032>
- McGrath, K. (2019) Your weather app may not be as reliable as you thought. Here's why. Digital Trends. [Accessed on 3 July, 2023 at Available from: <https://www.digitaltrends.com/mobile/can-you-trust-weather-app/>].
- McKnight, D.H. & Chervany, N.L. (2001) Trust and distrust definitions: one bite at a time. In: Falcone, R., Singh, M. & Tan, Y. (Eds.) *Trust in cyber-societies: integrating the human and artificial perspectives*. Berlin Heidelberg: Springer, pp. 27–54. Available from: https://www.academia.edu/download/47013991/Trust_and_distrust_definitions_One_bite_a20160704-6727-gjwhjg.pdf
- Millner, A. (2008) Getting the most out of ensemble forecasts: a valuation model based on user–forecast interactions. *Journal of Applied Meteorology and Climatology*, 47(10), 2561–2571. Available from: <https://doi.org/10.1175/2008JAMC1899.1>
- Morrow, B.H. (2008) Communicating Weather Forecast Uncertainty: An Exploratory Study with Broadcast Meteorologists. pp. 28, Available from: https://www.researchgate.net/profile/Jeffrey-Lazo/publication/237288288_Communicating_Weather_Forecast_Uncertainty_An_Exploratory_Study_with_Broadcast_Meteorologists/links/570ebb3c08aee76b9dadfd47/Communicating-Weather-Forecast-Uncertainty-An-Exploratory-Study-with-Broadcast-Meteorologists.pdf.
- Morss, R.E., Demuth, J.L. & Lazo, J.K. (2008) Communicating uncertainty in weather forecasts: a survey of the U.S. public. *Weather and Forecasting*, 23, 5, 974–991. Available from: <https://doi.org/10.1175/2008WAF2007088.1>

- Murphy, A.H. (1993) What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and Forecasting*, 8(2), 281–293. Available from: [https://doi.org/10.1175/1520-0434\(1993\)008%3C0281:WIAGFA%3E2.0.CO;2](https://doi.org/10.1175/1520-0434(1993)008%3C0281:WIAGFA%3E2.0.CO;2)
- Myers, L. (2019) Understanding the role of weather culture on messaging. In: *17th southeast severe storms Symp.* Starkville: East Mississippi Chapter of the National Weather Association and the American Meteorological Society.
- Newton, K. (1997) Social capital and democracy. *The American Behavioral Scientist*, 40(5), 575–586. Available from: <https://doi.org/10.1177/0002764297040005004>
- Nix-Crawford, B. (2017) The trust factor between meteorologists and viewers and the effects on local television audience retention. Thesis, Television Management, Drexel University, 82, Available from: <https://www.proquest.com/docview/2014476891/fulltextPDF/CF1763B8B74148ACPQ/1?accountid=34815>.
- Pew Research Center. (2021) Mobile Fact Sheet. Accessed on 7 July 2023, [Available from: <https://www.pewresearch.org/internet/fact-sheet/mobile/>]
- Phan, M.D., Montz, B.E., Curtis, S. & Rickenbach, T.M. (2018) Weather on the go: an assessment of smartphone mobile weather application use among college students. *Bulletin of the American Meteorological Society*, 99(11), 2245–2257. Available from: <https://doi.org/10.1175/BAMS-D-18-0020.1>
- Ripberger, J., Bell, A., Fox, A., Forney, A., Livingston, W., Gaddie, C. et al. (2022) Communicating probability information in weather forecasts: findings and recommendations from a living systematic review of the research literature. *Weather, Climate, and Society*, 14(2), 481–498. Available from: <https://doi.org/10.1175/WCAS-D-21-0034.1>
- Savelli, S. & Joslyn, S. (2012) Boater safety: communicating weather forecast information to high-stakes end users. *Weather, Climate, and Society*, 4(1), 7–19. Available from: <https://doi.org/10.1175/WCAS-D-11-00025.1>
- Schlenker, B.R., Helm, B. & Tedeschi, J.T. (1973) The effects of personality and situational variables on behavioral trust. *Journal of Personality and Social Psychology*, 25(3), 419–427. Available from: <https://doi.org/10.1037/h0034088>
- Sherman-Morris, K. (2005) Tornadoes, television and trust—a closer look at the influence of the local weathercaster during severe weather. *Global Environmental Change Part B: Environmental Hazards*, 6, 4, 201–210. Available from: <https://doi.org/10.1016/j.hazards.2006.10.002>
- Silver, A. (2015) Watch or warning? Perceptions, preferences, and usage of forecast information by members of the Canadian public. *Meteorological Applications*, 22(2), 248–255. Available from: <https://doi.org/10.1002/met.1452>
- U.S. Census Bureau. (2020) Age and Sex Composition in the United States: 2020. Accessed on 10 March 2023, [Available from: <https://www.census.gov/data/tables/2020/demo/age-and-sex/2020-age-sex-composition.html>].
- U.S. Census Bureau. (2021) Educational Attainment in the United States: 2020. Accessed on 10 March 2023, [Available from: <https://www.census.gov/data/tables/2020/demo/educational-attainment/cps-detailed-tables.html>].
- U.S. Census Bureau. (n.d.a) National Population by Characteristics: 2020–2022. Accessed on 10 March 2023, [Available from: <https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-detail.html>].
- U.S. Census Bureau. (n.d.b) 2020 Census Urban Areas Facts. Available from: <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2020-ua-facts.html>.
- Vaughn, C., Sherman-Morris, K., Gutter, B. & Brown, M. (2023) A change in the weather: understanding public usage of weather apps. *Journal of Operational Meteorology*, 11, 140–160.
- Voulgaris, C.T. (2019) Crystal balls and black boxes: what makes a good forecast? *Journal of Planning Literature*, 34(3), 286–299. Available from: <https://doi.org/10.1177/0885412219838495>
- Wall, T.U., Brown, T.J. & Nauslar, N.J. (2017) Spot weather forecasts: improving utilization, communication, and perceptions of accuracy in sophisticated user groups. *Weather, Climate, and Society*, 9(2), 215–226. Available from: <https://doi.org/10.1175/WCAS-D-15-0055.1>
- Weyrich, P., Scolobig, A. & Patt, A. (2019) Dealing with inconsistent weather warnings: effects on warning quality and intended actions. *Meteorological Applications*, 26(4), 569–583. Available from: <https://doi.org/10.1002/met.1785>
- Williams, C.A. & Eosco, G.M. (2021) Is a consistent message achievable?: defining “message consistency” for weather Enterprise researchers and practitioners. *Bulletin of the American Meteorological Society*, 102(2), E279–E295. Available from: <https://doi.org/10.1175/BAMS-D-18-0250.1>
- Zabini, F. (2016) Mobile weather apps or the illusion of certainty. *Meteorological Applications*, 23(4), 663–670. Available from: <https://doi.org/10.1002/met.1589>
- Zabini, F., Grasso, V., Magno, R., Meneguzzo, F. & Gozzini, B. (2015) Communication and interpretation of regional weather forecasts: a survey of the Italian public. *Meteorological Applications*, 22(3), 495–504. Available from: <https://doi.org/10.1002/met.1480>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Vaughn, C., Sherman-Morris, K., Brown, M., & Gutter, B. (2024). That's not what my app says: Perceptions of accuracy, consistency, and trust in weather apps. *Meteorological Applications*, 31(3), e2205. <https://doi.org/10.1002/met.2205>