The National Weather Service Storm Prediction Center's Convective Outlook: Conclusions from Past Research and Recommendations for Future Development

Makenzie Krocak[1,2,3], Sean Ernst[1,2,3], Joe Ripberger[1], Castle Williams[4], Joseph Trujillo-Falcón[2,3], Burkely Gallo[2,3], Patrick Marsh[3]

[1] Institute for Public Policy Research and Analysis, University of Oklahoma, Norman, OK
[2] Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, OK
[3] NOAA NWS Storm Prediction Center, Norman, OK
[4] Cherokee Nation Strategic Programs, Silver Spring, MD

November 23, 2021

1. Introduction and History

The convective outlook, a staple product of the National Weather Service (NWS) Storm Prediction Center (SPC), has been issued daily in one form or another since 1955. Since then, the outlook has undergone significant changes to its format and purpose. Originally the outlook highlighted areas with a Moderate and High risk for severe weather for the day of issuance, but a third category, Slight, was added in 1974 (Hitchens and Brooks 2012) and day 2, 3, and 4-8 forecasts were added in 1986, 2000, and 2005, respectively (Edwards and Ostby 2015). Originally these forecasts were issued based on subjective forecaster expectations about the coverage and intensity of severe weather during the forecast period, but the categories were married to probabilistic coverage forecasts in 2003 (P. Marsh 2020, personal communication, SPC 2020). This was achieved through the Practically Perfect Forecast (PPF), an estimate of what the "perfect" SPC outlook would appear like based on reports of convective wind, hail, and tornadoes received throughout the day (Brooks and Kay 1998).

Using the PPF as a framework in addition to input from veteran forecasters, the probability of severe weather reports occurring within 25 miles of a point was matched to the three categories that made up the outlook (see figs. 1 and 2, P. Marsh 2020, personal communication). An additional area of hatching could be contoured on these outlooks, as well, to denote areas where the chance of a significant severe weather report (hail of greater than 2 inches in diameter, winds of over 74 mph, or tornadoes of EF2 strength or greater) within 25 miles of a point was greater than 10% (Imy and Edwards 2013, Grams et al. 2014). For this initial version of the probabilistic outlook, the likelihoods of each of the three major convective hazards was only broken down for the day 1 period, with probabilities for days 2 and beyond forecasting the likelihood of any severe weather report within 25 miles of a point.

Day 1 Probability to Categorical Outlook Conversion

(SIGNIFICANT SEVERE area needed where denoted by hatching - otherwise default to next lower category)

Outlook Probability	TORN	WIND	HAIL
2%	SEE TEXT	NOT USED	NOT USED
5%	SLGT	SEE TEXT	SEE TEXT
10%	SLGT	NOT USED	NOT USED
15%	MDT	SLGT	SLGT
30%	HIGH	SLGT	SLGT
45%	HIGH	MDT	MDT
60%	HIGH	HIGH	MDT

Figure 1: The probabilistic breakdown of the original Day 1 Convective Outlook for tornado, wind, and hail threats.

Day 2 Probability to Categorical Outlook Conversion

(SIGNIFICANT SEVERE area needed where denoted by hatching - otherwise default to next lower category)

Outlook Probability	Combined TORN, WIND, and HAIL
5%	SEE TEXT
15%	SLGT
30%	SLGT
45%	MDT
60%	HIGH

Figure 2: The probabilistic breakdown of the original Day 2 Convective Outlook, with probabilities linked to the likelihood of any severe report occurring within 25 miles of a point.

These probabilistic outlooks have performed well since their deployment, as multiple verification efforts by Hitchens and Brooks have shown. By comparing PPFs derived from storm report data to previous outlooks, the pair showed that annual total outlook area has decreased significantly since 1973, while false alarm rates have decreased and probability of detection has remained unchanged (Hitchens and Brooks 2012). Further work identified that SPC forecasters' skill had increased over time since the mid-1990s, which explained the improvement in accuracy identified in the 2012 paper (Hitchens et al. 2013). This increase in skill was further confirmed by the identification of similar improvements in accuracy in the day 2 and 3 outlooks (Hitchens and Brooks 2014). However, these verification methods do have some biases, with missed events for significant severe weather, particularly significant wind, impacting skill estimates (Hitchens and

Brooks 2017). Other studies have shown that outlooks are also generally more skillful for more "textbook" severe weather events where instability and shear are high, versus more difficult to predict low-instability, high shear events such as those common to the Southeastern US (Herman et al. 2018). Overall, the modern convective outlook has been shown to possess skill for forecasting severe weather, though low-risk, high-impact events remain a challenge for severe weather forecasting.

Though the probabilistic forecasts issued by the SPC in the convective outlook have not changed since their introduction in 2003, in 2014 the categorical system wedded to those probabilities was expanded from three to five categories (Edwards and Ostby 2015). This change was driven in part by the concerns of emergency managers across the US, who wanted the multiple probability levels encompassed by the Slight risk category to be immediately seen in the categorical outlook as well (P. Marsh, personal communication). However, managers did not want to change the higher end categories in the outlook, as these were closely related to many decision triggers in advance of severe weather events. The SPC suggested adding the terms "Marginal" and "Enhanced Slight" to cover the "See Text" and higher end Slight risk probabilities and used a customer survey on their website to collect user feedback. Though this survey did not conclusively show that users were able to successfully interpret the new categorical outlook, the SPC moved forward with the change while renaming "Enhanced Slight" to "Enhanced" in the final product (P. Marsh, personal communication). This new design was introduced in 2014 and has been used in the outlook product since (see figs. 3 and 4).

Day 1 Outlook Probability	TORN	WIND	HAIL
2%	MRGL	Not Used	Not Used
5%	SLGT	MRGL	MRGL
10%	ENH	Not Used	Not Used
10% with Significant Severe	ENH	Not Used	Not Used
15%	ENH	SLGT	SLGT
15% with Significant Severe	MDT	SLGT	SLGT
30%	MDT	ENH	ENH
30% with Significant Severe	HIGH	ENH	ENH
45%	HIGH	ENH	ENH
45% with Significant Severe	HIGH	MDT	MDT
60%	HIGH	MDT	MDT
60% with Significant Severe	HIGH	HIGH	MDT

Figure 3: The probabilistic breakdown of the current Day 1 and 2 Convective Outlooks for tornado, wind, and hail threats. Note the inclusion of the ENH and MRGL, or Enhanced and Marginal, tiers.

Day 3 Outlook Probability	Combined TOR, WIND, HAIL
5%	MRGL
15%	SLGT
15% with Significant Severe	SLGT
30%	ENH
30% with Significant Severe	ENH
45%	ENH
45% with Significant Severe	MDT

Figure 4: The probabilistic breakdown of the current Day 3 Convective Outlook with probabilities linked to the likelihood of any severe report occurring within 25 miles of a point. Note SEE TEXT and the second SLGT segments have been replaced by MRGL and ENH.

Although the change to the outlook was made operational, SPC leadership expected there could be some controversy around the new design (P. Marsh 2020, personal communication). To help educate users, the SPC adopted a NWS graphic that describes the impacts to be expected from each level of the outlook (see fig. 5). In addition, to help bilingual partners in communicating risk categories, the NWS graphic was also translated to Spanish (see fig. 6). The color and word scale used by the SPC has further become the de-facto scale used by multiple government organizations, including the NWS Weather Prediction Center excessive rainfall outlook (WPC 2020). Aside from extending the probability by hazard breakdown beyond the day 1 forecast to day 2 in 2020, no additional changes have been made to the outlook since the new categories were added in 2014 (P. Marsh, personal communication, Grams et al. 2014).

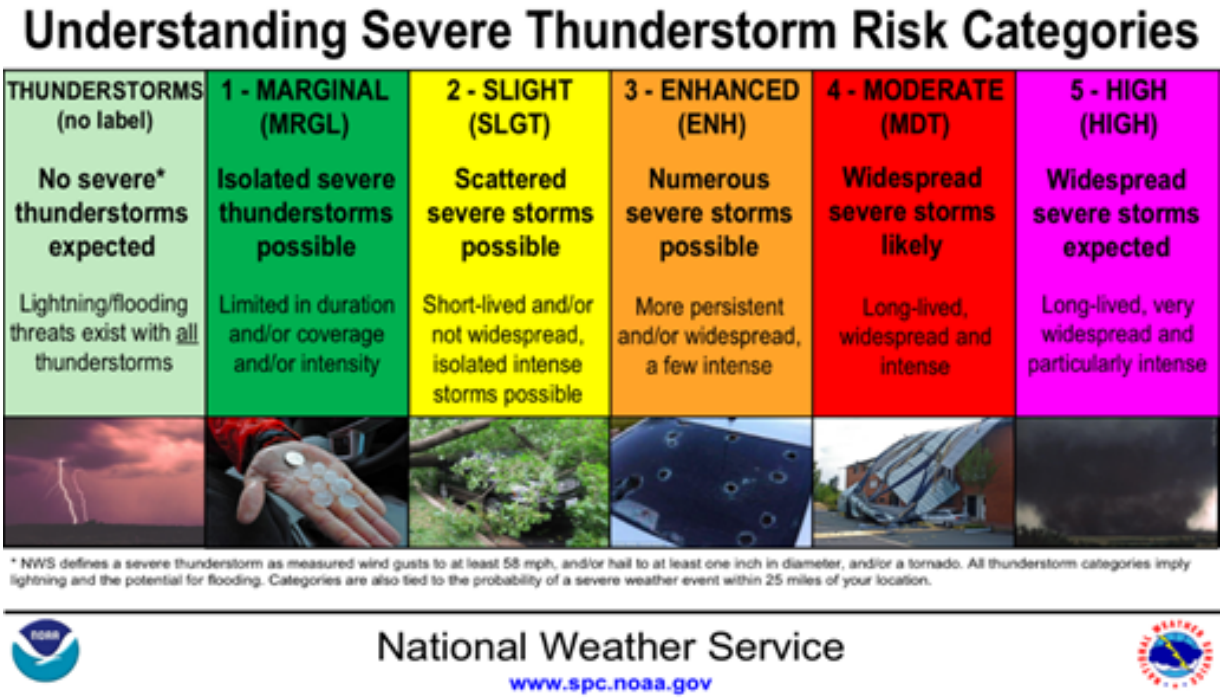

Figure 5: The SPC's online guide to the meaning of each tier of the convective outlook. Note that the image closely links the words and colors of each outlook tier, also adding numbers and coverage word descriptions to each level.

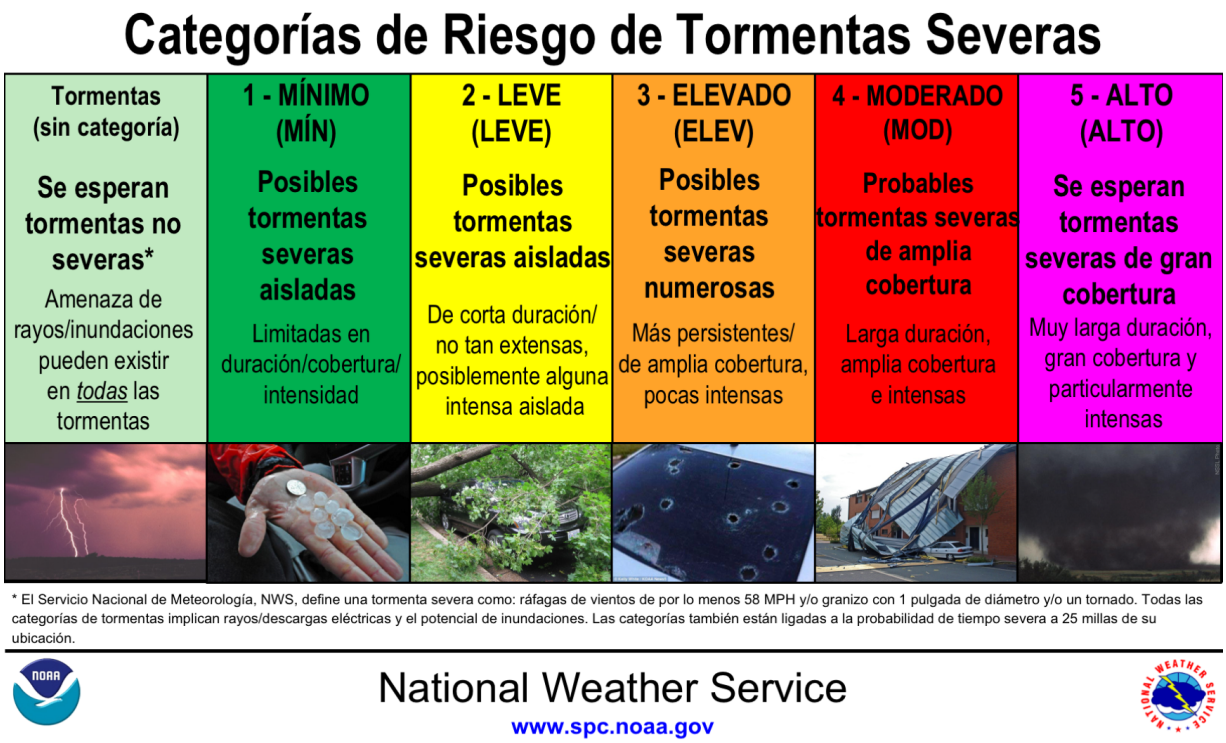

Figure 6: The SPC's online guide in defining risk categories for Spanish-speaking partners and audiences.

2. Research on challenges with the current outlook

1. Categorical labels and colors

One of the primary areas of study into the SPC's products has been how users interact with the words and colors that communicate risk levels in the modern convective outlook product. This investigation identified nine separate studies that investigated SPC outlook risk indicator interpretation, and there is a clear consensus among these studies that the words used in the outlook (Marginal, Slight, Enhanced, Moderate, and High in increasing order of likelihood of severe weather) are unintuitive to the public and dissatisfactory to core partners of the NWS (Abt Associates 2016; Ernst 2020; Ernst et al. 2021; Forbes 2020; Grundstein et al. 2019; Krocak et al. 2021; Lenhardt et al. 2020; Trujillo-Falcón et al. 2021; Williams 2021). These studies generally took the form of either interviews or surveys with broadcasters, emergency managers, and/or members of the general public.

Surveys where participants are asked to rank the SPC outlook words have shown that members of the public most commonly order the words from least to greatest risk as Slight, Marginal, Moderate, Enhanced, and High. This ordering was found in three independent national surveys and one Facebook survey (Ernst 2020; Ernst et al. 2021; Forbes 2020; Krocak et al. 2021; Williams 2021). Broadcast meteorologist interviews have found similar results, with participants reporting concern with how the public interprets Marginal, Slight, Enhanced, and Moderate, suggesting these words were inappropriate for the level of risk they were meant to communicate (Abt Associates 2016; Ernst 2020). Some broadcasters reported using numbers instead of words

(eg. Level 1, Level 2, etc) to present the outlook to their users, which in studies has been found to be more easily ranked and result in greater perceived risk by participants than the outlook words (Abt Associates 2016; Ernst 2020; Krocak et al. 2021).

Although some studies have identified words that may communicate risk more effectively (Trujillo-Falcón et al. 2021), others have identified that words that imply probability may generally not have a widely agreed-upon risk level associated with them (Lenhardt et al. 2020). Further, there is little agreement in the literature that ranking the SPC words and colors together improves interpretation accuracy. Williams (2021) found that participants most often ranked the scale correctly when colors and words were presented together, while Grundstein et al. (2019) found that presenting both colors and words to participants did not consistently improve their ranking ability. The colors used in the SPC outlook have also been presented separately from the categorical labels in interpretation ability studies, but users have generally been more able to rank the colors in the same way as the SPC, with the exception of the color magenta (Ernst 2021; Williams 2021). In addition, these colors are generally only seen combined with the words in the visual outlook product, with the colors creating a "bullseye" around the area of greatest risk, providing context clues that help users interpret their ranking. The categorical labels, on the other hand, can be encountered by non-users without these contextual cues, meaning that errors in ranking the labels on their own are more likely to occur in day to day life.

Conclusion #1: There is strong evidence that the words used in the convective outlook are not easily understood by users without training, which may limit the value of the product for non-experts and increase the workload of expert users that must repackage the product for their non-expert partners.

Recommendation #1: Develop and empirically evaluate an alternative to the verbal risk categories in the convective outlook. In the development stage, clearly identify the metrics that will be used to assess the relative effectiveness of alternative options.

2. Product consistency

Another major theme across studies of the SPC outlook highlights concerns about consistency in the presentation of the outlook across the SPC itself, local NWS offices, and broadcast outlets both nationally and locally (Williams and Eosco 2021). Survey experiments and interviews have identified that inconsistency in the areas included in the outlook and the colors used in the outlook lead to the greatest perceived inconsistency amongst members of the public (Grundstein et al. 2019; Williams et al. 2020). When presented with inconsistent graphics based on the SPC outlook, participants tended to trust and use graphics that displayed higher risk for their area or used the SPC color palette (Williams et al. 2020). Emergency managers have also shared concerns about consistency and reported that local NWS office graphics can fail to line up with the national SPC product (Abt Associates 2016). Broadcast meteorologists have highlighted how High and Moderate SPC outlook areas have never been issued or are extremely rare in some parts of the country, and that it can be difficult to explain to their audiences that a "Slight" or

"Moderate" day is a very significant severe weather event for those regions (Ernst 2020). Finally, within the Spanish weather enterprise, bilingual meteorologists have provided different translations of the risk categories, causing inconsistency among television stations across the United States (Trujillo-Falcón et al., 2021).

Conclusion #2: There is strong evidence of inconsistency in the presentation of the convective outlook across the NWS and partners in risk communication in English and Spanish, but little evidence on the implications of this inconsistency and almost no research on how to encourage more consistent presentation formats across the weather enterprise.

Recommendation #2: Encourage research on the causes and implications of inconsistent information presentation. Explore avenues that may encourage more consistent presentation of core messages.

3. Utilization

Other research efforts have focused on NWS Core Partner use of the convective outlook, specifically with regards to outlook use by emergency managers and broadcast meteorologists. Both have been found to use the outlook as a decision trigger for protective actions, and also share concerns that non-expert users in the public struggle to interpret the convective outlook product (Abt Associates 2016; Ernst et al. 2018; Ernst 2020). In particular, both groups report that the outlook is a helpful aid for making staffing decisions in advance of severe weather (Stough et al. 2012; Ernst 2020). Broadcasters report that the text contained within the outlook product helps them organize their weather story more effectively (Stough et al. 2012) but have shared concerns about how to explain the outlook in parts of the country where Enhanced or Moderate risks are the highest outlook categories issued for them (Ernst 2020).

Emergency managers use the outlook slightly differently, explaining that to them the outlook is a method for quickly analyzing a severe weather threat and that they prefer simplified graphics like the Public Severe Weather Outlook (Stough et al. 2012). The outlook has also been found to help managers narrow down the details of a severe weather threat with time (Ernst et al. 2018), with a recent study suggesting they use the Day 2 outlook most frequently to trigger protective actions (Cross et al. 2021). This study also found that managers generally take protective action when an Enhanced outlook is issued for their area, with the majority of managers surveyed waiting until at least a Slight risk outlook before making high-value decisions (Cross et al. 2021). However, emergency managers also report concern about national forecast products not lining up with local products, which they generally prioritize in their information consumption (Abt Associates 2016).

Conclusion #3: There is strong evidence that NWS core partners (broadcast meteorologists and EMs) use the convective outlook but less evidence on how, why,

and when they use it to make decisions. As a result, it is difficult to identify opportunities to improve the utility of the outlook for core partners. Aside from ad hoc stories and experiences, there is little evidence on if and how members of the public use the convective outlook to make decisions.

Recommendation #3: Encourage more research on how core partners and members of the public use the convective outlook to make decisions. This research is necessary to identify opportunities to improve the value of the product and identify unknown product dependencies (e.g., EM decision triggers and thresholds) that may be affected by product alterations.

4. Serving vulnerable populations

Further work has tested ways of incorporating community-level data into the convective outlook. Schneider et al. (2009) suggest that the probabilities of severe reports provided in the convective outlook could be linked with population estimates to create a social impact scale. The researchers proposed that this scale could be used to estimate the likelihood of a catastrophic tornado event based on the SPC outlook forecast, which could help organizations like FEMA as well as local emergency managers prepare resources before storms form. Such a product or change to the outlook would not be without downsides, as Schneider et al. (2009) address that this scale would be biased towards higher risks for larger population centers and lower risks in rural areas, where factors beyond population density can increase individual risk of harm from severe weather.

Other social factors beyond population can lead to vulnerabilities to severe convective storms. Though few studies identified here have addressed SPC outlook interpretation across the US public, those that have found inequality in comprehension across demographic groups. Ernst et al. (2021) found that numeracy, or the ability to understand probabilities and math, was a strong predictor of individuals' ability to correctly rank the words and colors used in the SPC outlook. Additionally, Black and Non-White participants did not perform as well on the SPC color ranking task as White participants, suggesting that marginalized groups in the US may be underserved by the current design of the outlook (Ernst et al. 2021). Along these lines, Trujillo-Falcón et al. (2021) showed that language barriers exist for users that do not speak or read English as their primary language. Further, there can exist dialects within languages that lead to words having different meanings across different regions, meaning that simply directly translating English words in the outlook may not be an effective solution. Therefore, future studies should incorporate these marginalized groups and guarantee that they are not analyzed under umbrella terms (e.g., instead of simply asking whether someone is "Hispanic or Latino," also inquire where their Hispanic/Latino heritage comes from).

Conclusion #4.1: There is emerging evidence that the convective outlook can be coupled with community-level data (e.g., Census data) to better forecast the societal impacts of severe storms.

Conclusion #4.2: There is a growing body of evidence suggesting that comprehension of the convective outlook is not equitable across different groups in society. Comprehension is lowest in vulnerable and marginalized communities, especially if these communities have relatively low levels of numeracy. This may compound the impact of disasters in these areas.

Recommendation #4.1: Encourage more research on how to integrate the convective outlook with community-level data. This research is likely to have the most impact if efforts include but look beyond population density to identify and include other social vulnerabilities that may impact community response.

Recommendation #4.2: Insist that researchers include vulnerable and marginalized communities when empirically evaluating alternatives to the risk categories in the convective outlook. For example, empirical evaluation must include assessments of Spanish-language alternatives to the current risk categories.

3. Research on changes to the current outlook

1. Categorical Labels

As the sections above demonstrate, previous research has generally focused on documenting problems (challenges) with the current convective outlook. There is considerably less research on possible solutions to these problems. One exception is a study by Krocak et al. (2021), which examined the potential impact of modest changes to the SPC categories on public risk perceptions and protective action intentions. The proposed alternatives included: (1) replace the risk words with numeric levels (e.g., replace "slight risk" with "level 2 of 4 risk"); (2) replace the risk words with percentages (e.g., replace "slight risk" with "a 5% chance"); (3) supplement the risk words with levels (e.g., replace "slight risk" with "slight risk (level 2 of 5)"); or (4) supplement the risk words with percentages (e.g., replace "slight risk" with "slight risk (5% chance)"). As found in Ernst et al. (2020), participants' perceived risk was greater for the word Enhanced than Moderate, with Enhanced consistently scoring higher on likelihood of response than Moderate, even when these labels were combined with level numbers or percentages. In other words, supplementing the current words with numeric information (levels/percentages) does not fix the underlying confusion caused by the current words. Replacing the words with numeric information alone seems to correct the confusion, but the levels and percentages were not equivalent in their inducement of risk perceptions and intended responses. Participants estimated their concern and likelihood of response consistently lower for the percentages than the levels, providing preliminary evidence that levels alone may be a preferred format. This finding was consistent across demographic groups, suggesting that numeric risk levels do not cause the inequalities in comprehension that appear to accompany the current words.

Conclusion #5: There is not enough evidence on alternatives to the convective outlook to clearly identify the most effective solution to the problems that previous researchers have documented. However, research is moving in this direction and early evidence suggests that replacing the risk words with numeric risk levels may assist in overcoming the confusion caused by the current words.

Recommendation #5: Risk levels represent a promising alternative to the risk categories in the convective outlook. Before adopting this alternative, it is imperative to encourage more research on the generalizability of these findings to other populations and settings, and on different metrics of effectiveness. It is also imperative to identify the downstream implications of such a change on core partners and others who rely on convective outlook.

2. Use of timing information

Studies into how individuals use the SPC convective outlook have also identified a need for storm timing information and have highlighted both successes and challenges in using the outlook product to communicate timing of severe weather threats. Importantly, improved storm timing information has been directly requested by NWS partners in previous studies. Core NWS partners, such as broadcast meteorologists and emergency managers, report using the convective outlook to set staffing levels and plan a response to potential impacts (Ernst 2020; Stough et al. 2012). However, these users also wanted timing information to be conveyed in local instead of UTC time and reported better response outcomes after requesting more timing information from local offices (Abt Associates 2016; Ernst, Ladue, and Gerard 2018, Krocak 2020). Further, Grundstein et al. (2019) found that members of the public could misinterpret the outlook as the threat at the current moment, rather than a 24-hour summary of risk, asking questions like "I wonder if this purple area three hours from now is going to be over us? Will it go away?" The researchers suggest that explicit timing information may reduce confusion about changes in risk throughout the day for non-expert users.

The most recent attempt to package storm timing information in a product easily accessed and understood by core users is described by Krocak (2020). In a series of experiments from 2016 to 2019, the Potential Severe Timing (PST) forecast was included in simulations involving NWS meteorologists and emergency managers at the HWT (Clark et al. 2018, 2019). Initially this product took the form of lines suggesting times that storms would arrive along their length, called isochrones, but in practice it was soon found that forecasters and users preferred polygon areas that captured four-hour times when storms could be expected. Emergency managers lauded the product for helping them make decisions for time-sensitive actions like school closures, but had some concerns about when forecasters drew overlapping PST polygons, and could be confused by the colors used to differentiate the polygons (Clark et al. 2018, 2019). Forecasters felt the product was helpful, but that the forecast process was difficult, and that there could be challenges working with local offices on defining polygon bounds if the product were to be centered at the SPC

(Krocak 2020). Similarly, Jirak et al. (2012) developed a method for creating more specific timing information in the form of 4-hour periods resulting from disaggregation of the full outlook. Regardless of the packaging, there is a desire for more specific timing information on the convective outlook scale.

Conclusion #6: There is strong evidence that core partners and members of the public desire and may benefit from the inclusion of sub-daily timing information in the convective outlook. There is some evidence that forecasters can skillfully provide such information, but it is difficult to produce these forecasts, and there may be challenges when coordinating the product with local forecast offices.

Recommendation #6: Encourage more research on the inclusion of sub-daily timing information in the convective outlook. Develop and empirically evaluate alternative communication formats to include this information in public facing products (e.g., outlook graphics).

3. Probabilistic information

Beyond Krocak et al. (2021), there have not yet been studies into how users interpret the absolute probabilities that underlie the SPC outlook, and this study was limited to the information that is currently present in the convective outlook. There are numerous other ways to present the likelihood of severe weather occurrence, such as relative probabilities and odds ratios. Literature reviews on how to communicate probability information in different domains (Bonner et al. 2021; Ripberger et al. 2020) suggest that use of numeric probabilities and clear reference classes are best for communicating risk to the public in an unbiased manner. Pitfalls can emerge from the use of "1-in-x" frequency formats and the use of words and phrases like "there is a chance of rain" to communicate probability information.

Another potential area where the SPC's probabilistic communication could change is in moving from the current system of discrete probabilities (e.g., areas denoting 2%, 5%, 10%, etc. levels of tornado likelihood with clearly drawn lines between levels) to continuous probabilities (e.g., areas that increase in probability more gradually across area, such as in the National Hurricane Center's hurricane wind likelihood product). Although studies such as Grunstein et al. (2019) have looked at how being on the threshold between discrete outlook levels impacts perceived risk, there have yet to be studies of what happens to user comprehension and risk perception when probabilities or outlook levels are smoothed to a continuous scale.

Conclusion #7: There is considerable evidence that including probabilistic information in forecasts improves decision quality. There is also strong evidence that this relationship depends on how probabilistic information is presented. There are a few general rules (e.g., use numeric information when possible), but systematic case

studies are necessary to empirically determine the most effective presentation formats for different products.

Recommendation #7: Encourage more research on the most effective way to include probabilistic information in the convective outlook.[1]

4. Conditional intensity

In its current form, the SPC convective outlook relies on probability information (probabilistic coverage) to define the risk of a given event. Little or no attention is given to the possible severity (intensity) of the event when determining the risk category. This can create risk communication challenges when forecasters believe that a storm is unlikely to form (low probability) but that it might produce a significant tornado if it were to form (high intensity). To overcome this challenge, SPC forecasters are investigating the possibility of separating probabilistic coverage from intensity to include "conditional intensity" information in future iterations of the convective outlook. Some studies, such as those discussed by R. Clark (2020), suggest that forecasters have some skill in creating the product, but many questions remain about verification of such forecasts. Likewise, many questions remain about the value of this information to local forecast offices, core partners, and members of the public and the most effective to provide it.

Conclusion #8: There is a growing body of evidence that SPC forecasters could skillfully provide conditional intensity information in future versions of the convective outlook. More evidence is necessary on the value of this information to end users and the most effective way to provide it.

Recommendation #8: Encourage research on the value of conditional intensity information and the most effective way to include it in the convective outlook.[2]

We recognize that the current body of research is ongoing and actively changing. There are projects underway and others that will be started in the near future that should be taken into consideration before any changes are adopted. Furthermore, questions remain about how alterations to the current system (like numeric category labels) will impact interpretation by different users. Therefore, our final recommendation is to create a space where new questions, ideas, and conclusions can be discussed as additional relevant work is completed.

[1] Research done in the FY2021 JTTI project "Enhancing the Storm Prediction Center's Convective Outlook with Continuous Probabilities and Conditional Intensity Forecasts" will contribute to this recommendation.

[2] Research done in the FY2021 JTTI project "Enhancing the Storm Prediction Center's Convective Outlook with Continuous Probabilities and Conditional Intensity Forecasts" will contribute to this recommendation.

References

Abt Associates Inc., 2016: Interpretation and use of the SPC day 1 outlook and recommendations for increasing temporal and spatial resolution. #29103, 42pp, https://repository.library.noaa.gov/view/noaa/29103

Bonner, C., L. J. Trevena, W. Gaissmaier, P. K. J. Han, Y. Okan, E. Ozanne, E. Peters, D. Timmermans, and B. J. Zikmun-Fisher, 2021: Current best practice for presenting probabilities in patient decision aids: Fundamental principles. Med. Decis. Making, 1-13, https://doi.dox.org/10.1177/0272989X21996328

Brooks, H. E., M. Kay, and J. A. Hart, 1998: Objective limits on forecasting skill of rare events. Preprints, 19th Conf. on Severe Local Storms, Minneapolis, MN, Amer. Meteor. Soc., 552–555. [Accessed online at: https://www.nssl.noaa.gov/users/brooks/public_html/papers/prague2k1.pdf]

Clark, R., I. L. Jirak, P. T. Marsh, R. Schneider, and B. T. Gallo, 2020: Revising hazard intensity information in Storm Prediction Center outlooks: A hazardous weather testbed experiment. 30th Conf. on Weather Analysis and Forecasting/26th Conf. on Numerical Weather Prediction, Boston, MA, Amer. Meteor. Soc., 8B.1, https://ams.confex.com/ams/2020Annual/meetingapp.cgi/Paper/365388

Clark, A., I. Jirak, B. Gallo, B. Roberts, K. Knopfmeier, R. Hepper, A. Dean, P. Heinselman, M. Krocak, J. Choate, K. Wilson, P. Skinner, Y. Wang, G. Creager, L. Wicker, S. Dembek, and J. Hales, 2018: Spring Forecasting Experiment 2018 preliminary findings and results. Hazardous Weather Testbed, 69pp, https://hwt.nssl.noaa.gov/sfe/2018/docs/HWT_SFE_2018_Prelim_Findings_v1.pdf

Clark, A., I. Jirak, B. Gallo, B. Roberts, K. Knopfmeier, J. Vancil, A. Dean, P. Heinselman, M. Krocak, J. Choate, K. Wilson, P. Skinner, Y. Wang, G. Creager, L. Reames, L. Wicker, S. Dembek, and S. Weiss, 2019: Spring Forecasting Experiment 2019 preliminary findings and results. Hazardous Weather Testbed, 77pp, https://hwt.nssl.noaa.gov/sfe/2019/docs/HWT_SFE_2019_Prelim_Findings_FINAL.pdf

Cross, H. A., D. Cavanaugh, C. C. Buonanno, and A. Hyman, 2021: The Impact of the Storm Prediction Center's Convective Outlooks and Watches on Emergency Management Operational Planning. J. Operational Meteor., 9(3), 36-46, https://doi.org/10.15191/nwajom.2021.0903

Edwards, R. and F. Ostby, 2015: Time line of SELS and SPC. Storm Prediction Center, accessed Feb 28 2020, https://www.spc.noaa.gov/history/timeline.html

Ernst, S., D. LaDue, and A. Gerard, 2018: Understanding emergency manager forecast use in severe weather events. J. Operational Meteor., 6(9), 95-105, https://doi.org/10.15191/nwajom.2018.0609

Ernst, S., 2020: Colorful language: Investigating the interpretation of the Storm Prediction Center's convective outlook by broadcast meteorologists and members of the public. Thesis, College of Atmospheric and Geographic Sciences, University of Oklahoma, 124pp, https://shareok.org/handle/11244/325336

Ernst, S., J. T. Ripberger, M. J. Krocak, H. Jenkins-Smith, and C. Silva, 2021: Colorful Language: Investigating Public Interpretation of the Storm Prediction Center Convective Outlook, Wea. Forecasting, DOI: https://doi.org/10.1175/WAF-D-21-0001.1

Forbes, A., 2020: Analyzing Severe Weather Terminology Interpretation in the Southeast United States. Senior Capstone, Mississippi State University.

Grams, J., B. Bunting, and S. Weiss, 2014: SPC Convective Outlooks. Storm Prediction Center, April 9th, 2019, DOI: https://www.spc.noaa.gov/misc/SPC_probotlk_info.html

Grundstein, A., J. So, and C. Williams, 2019: A qualitative look at SPC's day 1 convective outlook among members of the public. Phase 1 final report, 68pp.

Herman, G. R., E. R. Nielsen, and R. S. Schumacher, 2018: Probabilistic verification of Storm Prediction Center convective outlooks. Wea. Forecasting, 33, 161-184, DOI: 10.1175/WAF-D-17-0104.1

Hitchens, N. M. and H. E. Brooks, 2012: Evaluation of the Storm Prediction Center's day 1 convective outlooks. Wea. Forecasting, 27, 1580-1585, DOI: 10.1175/WAF-D-12-00061.1

Hitchens, N. M., H. E. Brooks, and M. P. Kay, 2013: Objective Limits on Forecasting Skill of Rare Events. Wea. Forecasting, 28, 525-534, DOI: 10.1175/WAF-D-12-00113.190

Hitchens, N. M. and H. E. Brooks, 2014: Evaluation of the Storm Prediction Center's convective outlooks from day 3 through day 1. Wea. Forecasting, 29, 1580-1585, DOI: 10.1175/WAF-D-13-00132.1

Hitchens, N. M. and H. E. Brooks, 2017: Determining criteria for missed events to evaluate significant severe convective outlooks. Wea. Forecasting, 32, 1321-1328, DOI: 10.1175/WAF-D-16-0170.1

Imy, D. and R. Edwards, 2013: SPC severe weather outlooks. Storm Prediction Center, accessed Feb 28 2020, https://www.spc.noaa.gov/misc/SPC_3-tier_probotlk_info.html

Jirak, I. L., C. J. Melick, A. R. Dean, S. J. Weiss, and J. Correia Jr., 2012: Investigation of an automated temporal disaggregation technique for convective outlooks during the 2012 Hazardous Weather Testbed Spring Forecasting Experiment. 26th Conf. on Severe Local Storms. Nashville, TN, Amer. Meteor. Soc., 10.2, https://ams.confex.com/ams/26SLS/webprogram/Paper211733.html

Krocak, M., 2020: If we forecast it, they may (or may not) use it: Sub-daily severe weather timing information and its utility for forecasters, stakeholders, and end users. Dissertation, University of Oklahoma, 144pp, https://shareok.org/handle/11244/324932

Krocak, M. J., S. Ernst, J. T. Ripberger, C. Silva, and H. Jenkins-Smith, 2021: Exploring the differences in SPC convective outlook interpretation using categorical and numeric information, Wea. Forecasting, Under Review.

Lenhardt, E. D., R. N. Cross, M. J. Krocak, J. T. Ripberger, S. R. Ernst, C. L. Silva, and H. C. Jenkins-Smith, 2020: How likely is that chance of thunderstorms? A study of how National Weather Service Forecast offices use words of estimative probability and what they mean

to the public. J. Operational Meteor., 8(5), 64-78, https://doi.org/10.15191/nwajom.2020.0805

Ripberger, J. T., A. Bell, A. Fox, A. Forney, W. Livingston, C. Gaddie, C. Silva, and H. Jenkins-Smith, 2020: Communicating probability information in weather forecasts: Findings and recommendations from a living systematic review of the research literature. Systematic Review Report v2, 57pp, https://repository.library.noaa.gov/view/noaa/29110/noaa_29110_DS1.pdf

Schneider, R. S., A. R. Dean, and H. Brooks, 2009: Estimating potential severe weather societal impacts using probabilistic forecasts issued by the NWS Storm Prediciton Center. 23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction, Omaha, NE, Amer. Meteor. Soc., 5B.5, https://ams.confex.com/ams/23WAF19NWP/techprogram/paper_154306.htm

Stough, S., E. Leitman, J. Peters, J. Correia Jr., 2012: The role of Storm Prediction Center products in decision making leading up to severe weather events. 18pp, https://www.spc.noaa.gov/publications/leitman/stough.pdf

Trujillo-Falcón, J. E., O. Bermúdez, K. Negrón-Hernández, J. Lipski, E. Leitman, and K. Berry, 2021: Hazardous weather communication en español: Challenges, current resources, and future practices. *Bull. Amer. Meteor. Soc.*, **102**, 765-773, https://doi.org/10.1175/BAMS-D-20-0249.1.

Weather Prediction Center (WPC), 2020: Excessive rainfall forecasts. Accessed Feb 28 2020, https://www.wpc.ncep.noaa.gov/qpf/excess_rain.shtml

Williams, C., and G. M. Eosco, 2021: Is a Consistent Message Achievable?: Defining "Message Consistency" for Weather Enterprise Researchers and Practitioners. Bulletin of the American Meteorological Society, 102, 2, DOI: https://doi.org/10.1175/BAMS-D-18-0250.1

Williams, C., A. J. Grundstein, and J. So, 2020: Should severe weather graphics wear uniforms? Understanding the effects of inconsistent convective outlook graphics on members of the public. 15 Symposium on Societal Applications: Policy, Research, and Practice, Boston, MA, Amer. Meteor. Soc., 11A.1

Williams, C., 2021: What is Being Enhanced? An Examination of the Storm Prediction Center's Risk Category System Among Members of the Public. 16th Symposium on Societal Applications: Policy, Research, and Practice/2nd Symposium on Diversity, Equity, and Inclusion. Amer. Meteor. Soc., J4.1