How Forecasters Anticipate Nocturnal, Cool-Season Southeastern Tornado Events

David P. Nowicki[1,2], Kimberly E. Klockow[3,4], Alan E. Gerard[4], Ariel E. Cohen[5], and John A. Hart[5]

[1]National Weather Center Research Experiences for Undergraduates Program
Norman, Oklahoma
[2]University of Mississippi, Oxford, Mississippi
[3]Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma
[4]NOAA National Severe Storms Laboratory, Warning Research & Development Division,
Norman, Oklahoma
[5]Storm Prediction Center, Norman, Oklahoma

ABSTRACT

In this study, forecasters across the Southeast were surveyed to find out current practices/tools used when issuing forecasts and warnings during nocturnal (0300 UTC-1200 UTC), cool-season (November through May) tornado events. Additionally, forecasters were asked to rate personal beliefs regarding the possible forecasting utility added by novel statistical models as well as beliefs about potential personal use of such models both before and after viewing output from an existent statistical model. Readings from the well-calibrated, climatologically based Statistical Severe Convective Risk Assessment Model (SSCRAM) were shown to forecasters halfway through the survey to serve as the treatment for the sample. SSCRAM output was analyzed by Bunker (2017) to discover conditional probabilities of tornado occurrence – given certain environmental parameters in varying ranges – for this region during this specific kind of event. SSCRAM averages of conditional probabilities of tornado occurrence were found for six parameters in specific ranges; then, these averages were compared to forecasters' subjective estimates of conditional probabilities, given the same parameters in the same ranges. Results of the study show a gap between forecasters' knowledge and their calibration with the environment as well as a shift in personal beliefs regarding SSCRAM's potential utility and use after being shown an example of that model's output.

1. BACKGROUND

Among the many weather events forecasters must be prepared for, nocturnal tornado events are some of the most difficult to predict. The current system used for the issuance of tornado warnings (TORs) is warn-on-detection. According to current National Weather Service (NWS) directives, "[Weather forecast office]s should issue TORs when there is radar indication and/or reliable reports of a tornado or developing tornado" (NOAA 2017b). The warn-on-detection system is clearly limited in its utility to provide ample time for the general public to become storm-ready and safe before tornadoes occur. For example, the average lead time for TORs has increased from the 1980s to the 2010s from about 6 minutes to around 14 minutes; however, over this time period, the non-zero lead time average for TORs (i.e., the average lead time for all tornadoes warned for before touchdown) has vacillated around the 18-20 minute mark from 1986-2011 (Carbin et al. 2012). In

---

Corresponding author address:
David P. Nowicki,
University of Mississippi, University,
Mississippi 38677
email: dpnowick@go.olemiss.edu

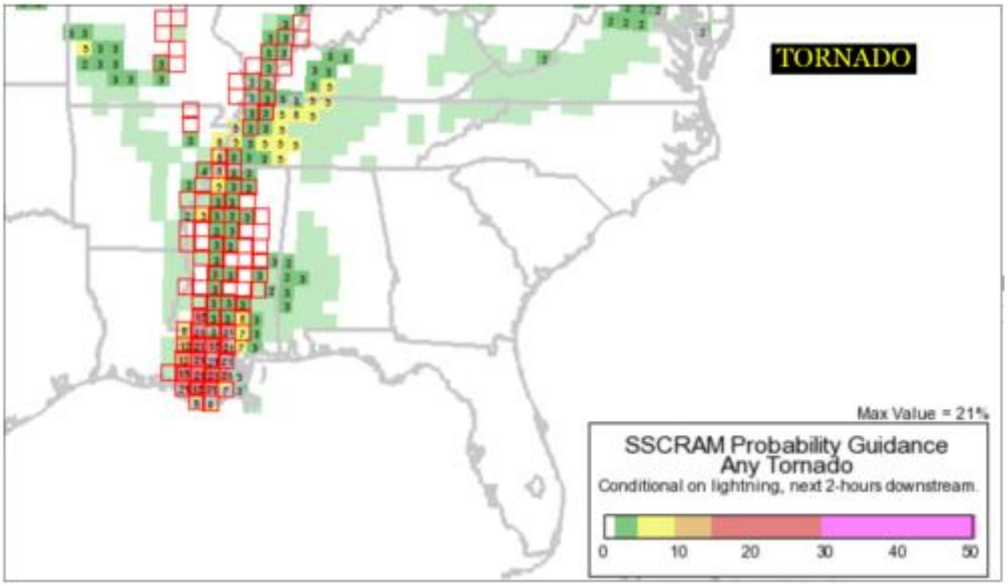

***Figure 1**: This figure shows real SSCRAM output data as it would appear to forecasters, if the model were used in operations currently. Red grid boxes indicate 40x40 km areas in which the condition of CG lightning has been met. Numbers within boxes indicate the conditional probabilities of 2-hour downstream tornado occurrence, using STP, 3 km AGL wind speed, and 100-mb MLCAPE as constraints (Hart & Cohen 2016).*

addition, warn-on-detection uses a binary warn/no-warn method of warning issuance that lends itself to a large false alarm rate (currently over 75%) (Hoekstra et al. 2010). By contrast, the Warn-on-Forecast (WoF) initiative, described in more detail below, has the potential to give forecasters enough information to warn severe weather events (including tornado occurrence) an hour or more before the threat exists, as well as presenting those warnings with probabilistic fields showing a percentage chance of the severe weather event's occurrence as opposed to a simple will/won't happen measure. The WoF program, officially begun in 2010 by the National Oceanic and Atmospheric Administration (NOAA), aims to remedy the seeming stagnation of warn-on-detection improvability by "increas[ing] tornado, severe thunderstorm, and flash flood warning lead times…in order to reduce loss of life, injury, and damage to the economy" (NOAA 2017a). According to Stensrud et al. (2009), the WoF system is necessary to increase tornado lead times from the current average of around 14 minutes to at least 30 minutes, which would be the minimum time for many industries – including hospitals and sports arenas, among others – to adequately prepare for such an event. The mean ideal lead time according to a survey of laypeople conducted by Hoekstra et al. (2010) yielded a result of 34.3 minutes.

While the goal of WoF is simple, the methods/utilities being created that would aid such a system – notably dual-polarization radar and novel ensemble numerical weather prediction (NWP) models – have proven insufficient in their current iterations to yield consistently adequate convective-level predictions of tornado occurrence (Stensrud et al. 2012). Though some utility for WoF has been shown through testing in NOAA's Hazardous Weather Testbed (HWT), to date, no single NWP ensemble model or radar assimilation method has been found to adequately provide forecasters with enough accurate data and predictive capacity to both prove its utility in most warning situations and increase forecaster confidence in issuing those warnings. This shows that, while constantly improving through rigorous HWT testing and numerous innovative research projects aimed towards improving forecasters' ability to warn severe weather events up to an hour or more before their occurrence, the WoF system is far from complete. For much of the history of the WoF system, the focus has been placed, at least model-wise, on the creation of ensemble NWPs, which make numerical predictions for each event individually based on conditions of the environment and storm in question based on storm-current data (Stensrud et al. 2009). Statistical models, however, have been explored less in their utility to aid the WoF system. A novel statistical model, the Statistical Severe Convective Risk Assessment Model (SSCRAM, see Fig. 1) could be used to help predict the probability of downstream severe weather incidents (including tornadoes) produced by thunderstorms, based on the condition of cloudto-ground lightning (Hart & Cohen 2016). This model provides for predictability of these events up to two hours ahead of a thunderstorm as soon as the first lightning strike occurs. This is accomplished by analyzing specific environmental parameters in the immediate vicinity of the thunderstorm in question. All probabilities produced by the model are based upon real historic environments (including null events) from 2006-2014 (Hart & Cohen 2016). This mitigates the necessity for the "model [to] be started with a very accurate representation of ongoing convection to obtain the necessary one-to-one correspondence between model-predicted and observed thunderstorms," as the limitations of NWPs are described in Stensrud et al. (2009).

Of the several goals of this study, one major aim is to assess the potential utility and use of a statistical model whose output is conditional probabilities of severe weather occurrence. Therefore, it is necessary to review current use of conditional probabilities by forecasters. Model output given in conditional probabilities being utilized by forecasters today is exemplified by several variations of Model Output Statistics (MOS). These predictive models use various environmental parameters and data from other existent models to establish conditional probabilities of a thunderstorm's or severe thunderstorm's occurrence based on the environment and other model readings at the initiation time of the model's run (Hughes 2002). While similar in several ways to SSCRAM, MOS operates using equations derived from five years of collected data, as opposed to the nine years of data collected by SSCRAM for this study (Hart & Cohen 2016). Additionally, it only produces conditional probabilities of thunderstorm and severe thunderstorm occurrence (not more specific severe weather events), and it has various forms that use different models as the basis for prediction in various scenarios (e.g., medium-range GFS MOS output, Eta-based MOS output, etc.; Hughes 2002).

When considering the Southeast U.S., the possibility of increased predictability of tornado occurrence potentially has greater societal implications than in other parts of the country. Research has shown that this particular region of the country is more vulnerable to tornado events year-round. Some issues facing the Southeast that are not faced to the same extent in other areas of the country include greater prevalence of nocturnal

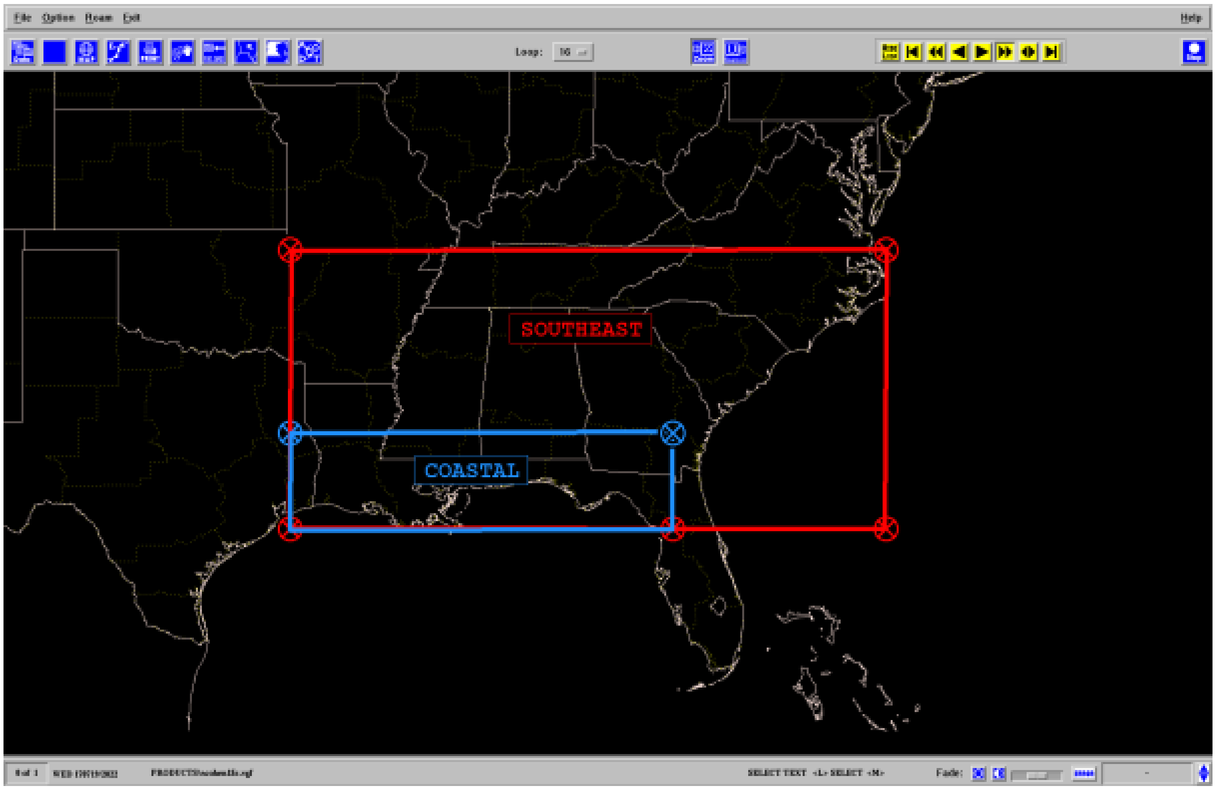

Figure 2: *The "Southeast" region as it is defined by Bunker (2017). Surveys were sent to the 22 WFOs with county warning areas lying within this region.*

storms (Ashley et al. 2008), greater frequency of violent tornadoes per 10,000 km^2 (Gagan et al. 2010), greater forward speed of storms (particularly during the cool season), and greater forest coverage, which reduces visibility (Ashley 2007, Gagan et al. 2010). In addition, this region faces socioeconomic factors that increase the potential for tornados to result in fatalities, including high percentages of the general population that are elderly, live in mobile homes or other unsubstantial structures, or are impoverished (Ashley 2007). The nocturnal nature of these storms enhances their potential danger to the public. As noted in Ashley et al. (2008), “most watches and warnings occurred well after prime-time television and late local news” in some of the deadliest tornado outbreaks over the last twenty years. This “break” in communication between those disseminating warnings and the general public greatly increases the risk of injury or fatality during nocturnal events, as compared to those that occur during daylight hours. This is particularly true in the “Dixie Alley” – contained fully within the “Southeast,” as defined in Bunker (2017, see Fig. 2) – in which nearly 30% of violent tornadoes occur during the hours of 9 PM and 7 AM CST, according to Gagan et al. (2010). This paper goes on to explain why the cool season is of great importance to the Southeast. In this region, unlike in the Great Plains region, tornado occurrence in much higher during this time of year, with up to 40% of strong tornadoes occurring during the October-February months. In general, the environment of much of the Southeast compounds the difficulty of detecting and predicting tornado events, as stated in Ashley (2007), because “the region’s close proximity to the Gulf of Mexico enhances low-level moisture availability and may make spotting tornadoes more difficult because of haze, lower lifting condensation levels, and a tendency for high-precipitation supercells that often contain ‘rain wrapped’ and difficult-to-identify tornadoes.” For these reasons, the Southeast is ideal for a study in tornado predictability, given the relative difficulty of detection/warning of tornadoes and heightened societal vulnerability. This project carried a set of hypotheses:

H1: Forecasters are generally able to discern environments that are predictive from those that are not;

H2: Forecasters lack well-calibrated subjective estimates of predictability for those environments;

H3: Factors such as years of forecasting experience overall, forecasting experience in the Southeast, and tendency to rely on environmental data for warning decision-making will improve subjective estimates of predictability.

H4: Forecasters will find value in a tool that helps them more faithfully estimate the predictability of these environments.

To test these hypotheses, forecasters across the Southeast were surveyed for their beliefs about predictive environments.

2. METHOD

To obtain the data necessary for assessing SSCRAM’s utility, a survey was distributed which collected various data concerning current forecaster methods utilized in nocturnal, cool-season severe weather events; subjective forecaster confidence in decisions regarding tornado probability of occurrence in these environments, given environmental parameters; and several demographic factors regarding professional and personal influencing factors on forecasters surveyed. Using data from Bunker (2017) – which attempts to show which environmental parameter(s) and ranges of those parameters show the highest conditional probability of downstream tornadic activity – it was possible to construct questions which addressed forecasters’ practices and confidences with regards to which environmental parameters are currently analyzed in nocturnal, cool-season thunderstorm events. Following a brief description of the novel SSCRAM model, which was presented mid-survey along with an image from the SSCRAM model, questions were asked to gauge how/if forecasters’ confidence in SSCRAM to improve their accuracy in warning for tornadoes changed. In addition, potential use of SSCRAM by forecasters in a hypothetical future situation in which that model would be available was gauged. This first section was comprised of nine questions of varying formats. Demographic

Parameter (Range)	Effective Bulk Shear (65-70 kts)	0-6 km AGL shear (65-70 kts)	Significant Tornado Parameter (4)	100-mb MLCAPE (2500-2750 J/kg)	Effective Helicity (550-600 (m/s)^2)	100-mb ML LCL height (700-800 m)
Forecasters' Average Subjective Estimate of Conditional Probability	42.8%	40.8%	57.7%	46.8%	56.5%	41.3%
SSCRAM-derived Average Conditional Probability	13.5%	3.4%	19.0%	3.1%	17.1%	2.2%
Subjective Inflation Factor	**217.3%**	**1111.3%**	**204.2%**	**1429.8%**	**230.7%**	**1817.8%**

Table 1: *This table shows results from the second question of the survey, which was asked before forecasters were shown SSCRAM output. Shown are forecasters' average subjective estimate of the conditional probability of a tornado's occurrence, conditional probability of a tornado's occurrence SSCRAM data averaged over the specified parameter range, and the resulting subjective inflation factor for each parameter.*

and personal forecasting practice beliefs were recorded (in a suite of seven questions) in an attempt to find any correlation between forecaster understanding, confidence (or change in confidence), or belief in SSCRAM's utility (or lack of this belief), and any of these demographic or personal practice identifiers. Questions regarding time forecasters have spent in the NWS, as well as those regarding existence of or adherence to a personal or "best" forecasting schema were devised using ideas from Dr. Daphne LaDue's Ph.D. dissertation (2011).

Before distribution, the survey was put through a focus group of three NWS forecasters from WFOs in various regions of the country. The forecasters only knew before the focus group that they would be assisting in the vetting of a survey created to aid in improving the WoF system. The focus group was instructed to read through the survey one question at a time, attempting to answer the question, as written, to the best of their abilities. They were then asked their answers; what each question was aiming to ask, in their opinions; and how well the scale for each answer suited their personal preference for answering the question (as they understood it) most accurately. After this process was completed for each question, the research team revealed to the group its intent for the question considered and asked how to better phrase the question, frame the situation to be considered, or alter the scale to yield the least confusing questions and most useful data for our study. Following the focus group, heavy edits were made to the survey (per focus group discussion/suggestions) before being distributed to forecasters. This survey was given to the Southern and Eastern Region Headquarters of the NWS to distribute to select Weather Forecasting Offices (WFOs) within the Southeastern region of the United States, as defined by the geographic range established as "Southeast" by Bunker (2017).

3. ANALYSIS & RESULTS

To analyze the first survey question – which asked forecasters to rate how often different types of tools/parameters were consulted during the type of event in question – ratios between radar (the most-used tool) and parameter consultation were calculated. These were computed to see how much more/less often each parameter was consulted compared to how often radar was consulted. These calculations were completed using the average ratings of the entire sample for each measure. The results for question one were quite telling with regards to current use of tools and consultation of environmental parameters. Radar was consulted considerably more often than each parameter listed [effective bulk shear (3.72x), effective helicity (2.94x), 0-6 km shear (5.14x), 100-mb MLCAPE (3.78x), 100-mb ML LCL height (4.36x), significant tornado parameter (3.06x)]. These results showed not only high relative use of radar compared to consultation of environmental parameters but also that environmental parameters were consulted very seldom, on average. This data was useful in testing correlations pertaining to H3 ("tendency to rely on environmental data for warning decision-making"), though that hypothesis was not supported, as no meaningful correlation between tendencies shown in this question and improved subjective estimates of conditional probability.

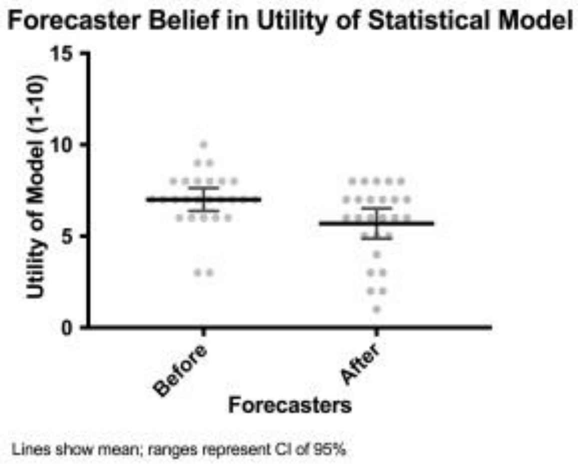

Figure 3: *This graph shows the results of a Wilcoxon signed-rank test performed to assess whether forecasters' beliefs regarding SSCRAM's potential utility (i.e., "value") changed after they were shown an example of SSCRAM output (see Fig. 1) and a brief description of how the model operates. Graph shows a statistically significant decrease (W = -136, p = 0.0016) in forecasters' belief in the model's utility.*

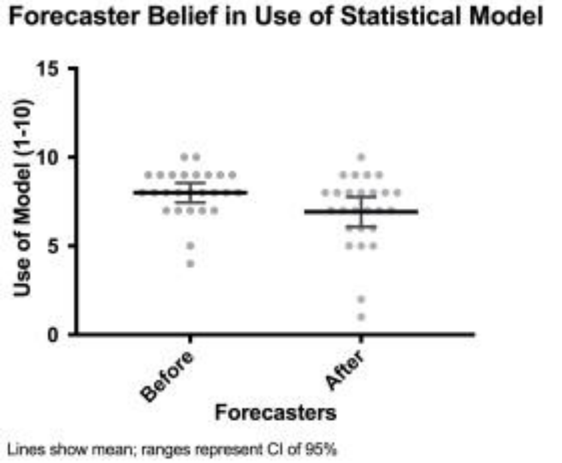

Figure 4: *This graph shows the results of a Wilcoxon signed-rank test performed to assess whether forecasters' beliefs regarding their propensity to use SSCRAM after they were shown an example of SSCRAM output (see Fig. 1) and a brief description of how the model operates. Graph shows a statistically significant decrease (W = -99, p = 0.0029) in forecasters' belief that they would potentially use the model.*

The second question – in which forecasters were asked to give their subjective estimates for the conditional probability of a tornado's occurrence, given specific parameter ranges – was analyzed by finding the average conditional probability for each parameter and comparing that average to the data-derived correct probability given by SSCRAM output. Comparison for each parameter was given as a subjective inflation factor (SIF) of the average forecaster's answer compared to the data-derived answer. When asked to give their subjective estimations, forecasters surveyed showed great lack of calibration with the environment, which supports H2. This is shown in Table 1, which shows the average of forecasters' subjective conditional probabilities, the average probabilities deemed "correct" from SSCRAM data, and the SIF comparison for each parameter, given a specific range. Though forecasters surveyed showed lack of calibration with regards to their subjective estimates, they did rank the parameters given in nearly the correct order of most- to least-predictive. The three most-predictive parameters according to SSCRAM data were among the top four most-predictive according to forecasters' subjective estimates, with the only incorrect ranking being MLCAPE (usually a strong predictor of severe weather occurrence, in the range given). Additionally, the SIFs for those three most-predictive parameters were the lowest three, meaning forecasters were more well-calibrated with those parameters that were most-predictive. Thus, though H2 was supported by the data collected, H1 was also supported (albeit with high SIFs).

To analyze the results regarding forecasters' beliefs of the potential utility of statistical models and their potential use of these models before and after introduction of the SSCRAM model (questions 3 and 4 vs. questions 8 and 9), a Wilcoxon signed-rank test was used to gauge the difference between the population's

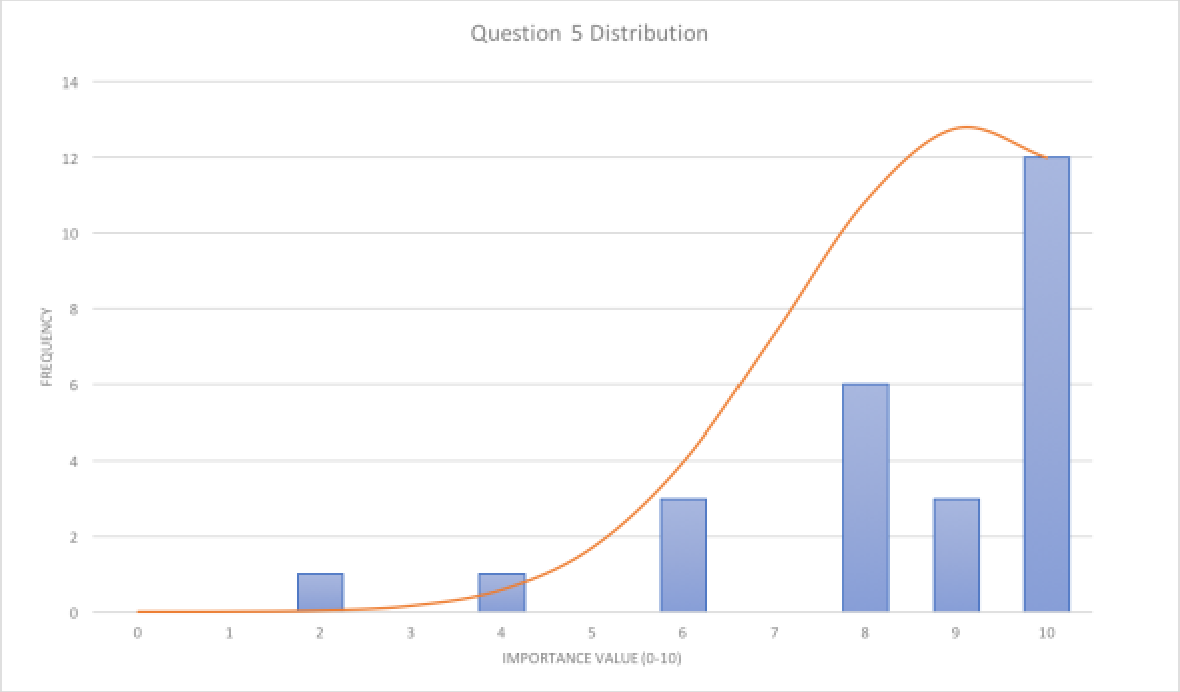

Figure 5: *This figure shows a histogram and accompanying distribution curve for forecasters' answers to question 5. The value forecasters recorded (0-10) regarding how important they considered environmental parameters to be during events with nocturnal convection is shown on the x-axis, while the number of forecasters that answered a certain value is shown on the left y-axis.*

average confidences in both of these factors. This comparison showed a statistically significant decrease in forecasters' average response for both perceived utility and potential use of SSCRAM – a result that runs completely counter to H4 (see Figs. 3 & 4). Three outliers – those with the largest differences in before/after response pairs – were found for the pair of questions regarding potential utility of a statistical model, and two were found for the pair of questions regarding potential use. The downward trend held and was still statistically significant in both instances when these outliers were removed from the sample (W = -85, p = 0.0129; W = -71, p = 0.0115; respectively).

Because of the strong negative skew to the answer distribution on question five regarding how important forecasters believe the environmental conditions to be in a nocturnal severe convective event, simple population mean, median, and mode were gleaned from the data for that question. Though questions one and two together show both that forecasters consult environmental parameters relatively infrequently compared to how often they consult radar or satellite data and that they have little calibration with those same parameters, question five showed that forecasters do believe that environmental conditions during nocturnal severe convective events are important (mean: 8.42, median: 9, mode: 10; see Fig. 5). The answers to this question were tested for correlation with various demographic factors to test H3, though the hypothesis was not supported, as no correlations were found.

Gauging forecasters' understanding of conditional probabilities as a concept was accomplished using a multiple-choice question to avoid any sort of language in a question that would seem patronizing to forecasters' intelligence or experience. Though question two showed poor calibration with the environment when forecasters were required to report their thinking in terms of conditional probabilities, ~88% of those surveyed showed understanding of conditional probabilities as a general concept by answering the multiple-choice question correctly, which supports H1.

During analysis, the data gathered for question six was deemed too convoluted to analyze with any certainty as to the validity of the analysis. Though the wording of the question was deemed clear by both the research team and those forecasters consulted during the survey validation focus group, the answers given by forecasters surveyed varied to such a large degree that it was not clear to the research team whether or not the forecasters all read the question the same way. Specifically, the question asked forecasters to give subjective estimates, as in question 2. However, this question asked for answers in odds ratios – “rate how much more/less likely you would be to issue a tornado warning (e.g., half (0.5x) as likely, twice (2x) as likely, etc.)” – with the same parameters given as in the previous question. Each parameter was given in two different ranges for comparison (e.g., 100-mb MLCAPE of 2500-2750 J/kg vs. 1000-1250 J/kg). While the intent of the question was to compare the second choice to the first and give odds ratios based off that comparison, it became clear after the survey’s distribution that, looking back at the question, even members of the research team interpreted the correct order of interpretation (i.e., first range over second range vs. second range over first range) differently. For this reason – as well as the fact that the range of answers within each parameter choice was staggering – the answers to question six were deemed too inconsistent to analyze as validated answers.

Additionally, the only correlation with demographics that proved useful were correlations that showed that answers intended to be mutually exclusive were reported as such by those surveyed. No significant correlations were found between any answers in the first part of the survey and any of the demographic measures tested, including level of education, location of WFO, time in the NWS, time at current WFO, frequency of outside learning opportunities taken, and personal beliefs regarding existence of personal or “best” forecasting schemas. These results run completely counter to H3, as no correlations were discovered. It is important to note that, though forecasters’ opinions regarding potential utility and use declined after seeing and reading about SSCRAM, both “after” measures were still positive (i.e., above a 5, “neutral,” average score).

4. CONCLUSIONS & FUTURE RESEARCH

This project began with an expectation – both informed by past literature and hoped for in the creation of SSCRAM – that probabilistic guidance would aid forecasters’ warning and forecast decisions, especially in particularly dangerous and difficult-to-read situations such as nocturnal, cool-season severe weather that has the potential to produce tornadoes. In light of this study, such an assumption is not necessarily substantiated by data, at least at this point in time. This is supported by the fact that the null hypothesis that forecaster beliefs in the utility/use of statistical models would not change after being shown real statistical model data was rejected because of a lowering in these beliefs. Currently, forecasters show lack of sufficient calibration with environment to adequately understand, trust, and effectively use the SSCRAM model with the result of more effectively warning and protecting the extremely vulnerable population in the Southeast. From analysis of survey questions, it is clear that forecasters do have good knowledge of conditional probability as a general concept and have reasonably well-calibrated intuition regarding the most-predictive environmental parameters in these situations. Forecasters surveyed were also in strong agreement that the environmental conditions during these situations are of great importance. Additionally, forecasters do still have a positive overall belief (on average) that SSCRAM would potentially add to forecasters’ personal utility and would be used. Furthermore, forecasters agreed in their belief that novel statistical models, generally, would add to their possible utility as a forecaster.

While forecasters gave positive results regarding their base knowledge, intuition, and goals, they become more reluctant to accept a novel model after being shown readings that disagree greatly with their subjective estimates. It is our conjecture that the reluctance of forecasters to alter their original beliefs likely stems from the fact that their beliefs are largely inflated with respect to the objective values shown to them during the course of the survey. This implies that revealing a description and image of SSCRAM gave those surveyed more doubt about either their ability to make adequate use of such a model or the validity of the model itself, neither of which supports this project’s hypothesis.

With regards to future research on this subject, the most obvious alteration to the study design that would likely aid the validity of the results gathered for this study would be to give the potential respondents more time to answer the survey, thereby increasing the sample size (n = 26, the key limitation to the generalizability of this study). Additionally, conducting post-survey cognitive interviews would allow the research team to better assess the validity of answers – especially answers to questions like number six, whose answers appeared too convoluted to possibly be valid. Further research beyond the original scope of this study could be done in the HWT to test how if forecasters show significant change in their accuracy or confidence of forecasts/warnings in the types of situation considered in this paper when given conditional probabilities. Certainly, more exposure to conditional probability outputs could potentially change the negative trend in forecaster beliefs of statistical model utility/use shown in this study. This research also opens the door to explore related issues that WoF aims to address, such as the currently high FAR. Given the high subjective inflation factor (SIF) shown by forecasters in this study, it could be useful to research a potential link between high SIF with regards to environmental parameters and high FAR.

5. ACKNOWLEDGMENTS

The author would like to thank Dr. Daphne LaDue (OU CAPS) for guidance through the many trials encountered during a first research experience as well as exceptional leadership and accessibility as the director of the National Weather Center Research Experience for Undergraduates program. Additionally, thanks are extended to Dr. James Correia, Jr., (CIMMS) for his help in organizing the author’s thoughts and narrowing the focus of this study to a manageable level and to Ryan Bunker (NWC REU intern) for his data contributions without which this project would have been impossible, as well as for his continued capacity as a productive and determined research partner. This work was prepared by the authors with funding provided by National Science Foundation Grant No. AGS-1560419, and NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement #NA11OAR4320072, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, NOAA, or the U.S. Department of Commerce.

6. REFERENCES

Ashley, W. S., 2007: Spatial and temporal analysis of tornado fatalities in the United States: 1880-2005. *Wea. Forecasting,* **22,** 1214-1228, doi: 10.1175/2007WAF2007004.1.

Ashley, W. S., A. J. Krmenec, and R. Schwantes, 2008: Vulnerability due to nocturnal tornadoes. *Wea. Forecasting,* **23,** 795-807, doi: 10.1175/2008WAF2222132.1.

Bunker, R. C., cited 2017: Illustrating predictability for nocturnal tornado events in the southeastern United States. [Available online at http://www.caps.ou.edu/reu/reu17/finalpapers/Bunker-Paper.pdf.]

Carbin, G., P. Heinselman, and D. Stensrud, 2012: White paper 2: current challenges in tornado forecast and warning. *Proc. Weather-Ready Nation: Science Imperatives for Severe Thunderstorm Research,* Birmingham, AL, National Oceanic & Atmospheric Administration, A-9-A-16.

Gagan, J. P., A. Gerard, and J. Gordon, 2010: A historical and statistical comparison of "tornado alley" to "dixie alley". *National Wea. Digest,* **34,** 145-155.

Hart, J. A., and A. E. Cohen, 2016: The Statistical Severe Convective Risk Assessment Model. *Wea. Forecasting,* **31,** 1697-1714, doi: 10.1175/WAF-D-16-0004.1.

Hoekstra, S., and Coauthors, 2010: A preliminary look at the social perspective of warn-on-forecast: preferred tornado warning lead time and the general public’s perceptions of weather risks. *Wea. Climate Soc.*, **3,** 128-140, doi: 10.1175/2011WCAS1076.1.

Hughes, K. K., 2002: Automated gridded forecast guidance for thunderstorms and severe local storms based on the eta model. Preprints, *21st SOS Conf.,* San Antonio, TX, Amer. Meteor. Soc., J3.4.

LaDue, D. S., 2011: How meteorologists learn to forecast the weather: social dimensions of complex learning. Ph.D. dissertation, University of Oklahoma, 240 pp.

National Oceanic & Atmospheric Administration, cited 2017: Warn on forecast. [Available online at http://www.nssl.noaa.gov/projects/wof/.]

National Oceanic & Atmospheric Administration, cited 2017: WFO severe weather products specification. [Available online at http://www.nws.noaa.gov/directives/sym/pd01005011curr.pdf.]

Stensrud, D. J., and Coauthors, 2009: Convective-scale warn-on-forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.,* **90,** 1487-1499, doi: 10.1175/2009BAMS2795.1.