

The Extended-Range Tornado Activity Forecast (ERTAF) Project

Vittorio A. Gensini, Bradford S. Barrett,
John T. Allen, David Gold, and Paul Sirvatka

ABSTRACT: Large-scale weather patterns favorable for tornado occurrence have been understood for many decades. Yet prediction of tornadoes, especially at extended lead periods of more than a few days, remains an arduous task, partly due to the space and time scales involved. Recent research has shown that tropical convection, sea surface temperatures, and the Earth-relative atmospheric angular momentum can induce jet stream configurations that may increase or decrease the probability of tornado frequency across the United States. Applying this recent theoretical work in practice, on 1 March 2015, the authors began the Extended-Range Tornado Activity Forecast (ERTAF) project, with the following goals: 1) to have a map room–style discussion of the anticipated atmospheric state in the 2–3-week lead window; 2) to predict categorical level of tornado activity in that lead window; and 3) to learn from the forecasts through experience by identifying strengths and weaknesses in the methods, as well as identifying any potential scientific knowledge gaps. Over the last five years, the authors have shown skill in predicting U.S. tornado activity two to three weeks in advance during boreal spring. Unsurprisingly, skill is shown to be greater for forecasts spanning week 2 versus week 3. This manuscript documents these forecasting efforts, provides verification statistics, and shares the challenges and lessons learned from predicting tornado activity on the subseasonal time scale.

<https://doi.org/10.1175/BAMS-D-19-0188.1>

Corresponding author: Vittorio A. Gensini, vgensini@niu.edu

In final form 24 January 2020

©2020 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: **Gensini**—Department of Geographic and Atmospheric Sciences, Northern Illinois University, DeKalb, Illinois; **Barrett**—Oceanography Department, United States Naval Academy, Annapolis, Maryland; **Allen**—Department of Earth and Atmospheric Science, Central Michigan University, Mount Pleasant, Michigan; **Gold**—Weather Analytics Center of Competency, Global Business Services, IBM, Armonk, New York; **Sirvatka**—Meteorology Program, College of DuPage, Glen Ellyn, Illinois

As forecast skill continues to improve (Bauer et al. 2015), especially in the extended range (Wanders and Wood 2016; Tian et al. 2017; Vigaud et al. 2017; Wang and Robertson 2019), interest from stakeholders is rapidly growing for forecasts that cover the subseasonal lead window (typically defined as the period spanning two weeks to two months into the future; Schubert et al. 2002; Robertson et al. 2015). This is particularly true for severe convective storms (SCSs; defined as thunderstorms that produce tornadoes, large hail, and/or damaging convective wind gusts) given the significant impacts on various economic and societal sectors (Simmons and Sutter 2013; Smith and Matthews 2015; NCEI 2019). Recent research has shown skill in forecasting U.S. tornado and hail frequency at various lead times in the subseasonal period (Allen et al. 2015; Lepore et al. 2017; Baggett et al. 2018; Lepore et al. 2018; Gensini and Tippett 2019), including a recently documented operational forecast of an anomalously high period of tornado frequency in late May 2019 issued more than three weeks in advance (Gensini et al. 2019).

These contemporary demonstrations of forecast skill for SCS events at subseasonal lead times is likely a testament to the increasing accuracy of global numerical weather prediction models (Bauer et al. 2015) and the increased understanding of weather and climate factors that explain significant amounts of the variance in subseasonal SCS frequency. For example, climatological studies have identified clustering behaviors of periods of high tornado frequency, which often occur over two or more days (Verbout et al. 2006; Trapp 2014). In addition, various weather and climate oscillations have strong links to U.S. tornado and hail frequency. The Madden–Julian oscillation (MJO) and the global wind oscillation (GWO) are two such modes of atmospheric variability with documented periodicity on the subseasonal time scale (Weickmann and Berry 2009; Zhang 2013). While not mutually exclusive (the GWO encompasses the MJO through incorporation of the tropical zonal wind fields), both processes have been shown to modulate tornado and hail frequency across the United States (Barrett and Gensini 2013; Thompson and Roundy 2013; Barrett and Henley 2015; Gensini and Marinaro 2016; Baggett et al. 2018; Gensini and Allen 2018; Tippett 2018; Moore 2018; Gensini et al. 2019; Moore and McGuire 2019). A known physical pathway by which the MJO and GWO influence CONUS severe weather is by generating positive mountain and frictional torques, particularly as convection associated with the MJO increases across the Maritime Continent and moves toward the international date line. This forces a stronger Hadley cell circulation, causing anomalous fluxes in the tropical meridional tropospheric wind component. As air is displaced poleward, it moves closer to Earth’s axis of rotation, thus causing increases in the zonal wind component due to the conservation of angular momentum. An increase in Northern Hemisphere westerly momentum results, which produces an extension of the polar jet stream over the Pacific Ocean. This gives rise to positive atmospheric angular momentum anomalies, characteristic of GWO phases 5 and 6, with potential for jet extension that eventually leads to wave breaking and a reduction of atmospheric angular momentum. It is this decreasing tendency of atmospheric angular momentum that favors synoptic weather patterns supportive of tornadoes east of the U.S. Rocky Mountains (Weickmann and Berry 2009; Gensini and Marinaro 2016). These synoptic weather patterns often include a midtropospheric trough in the western United States and a poleward flux in surface moisture across the Great Plains. More simply, GWO and MJO have been shown to alter the 300-hPa height fields, with

positive severe weather frequency anomalies generally located southeast of negative height anomalies, in regions favorable for ascent and for greater vertical wind shear (Barrett and Gensini 2013; Barrett and Henley 2015; Gensini and Marinaro 2016).

From an operational perspective, early modern forecasts for SCS events were confined to the same or next day (Grice et al. 1999), but it is now commonplace to see official forecast outlooks for SCSs—issued by the NOAA/NWS’s Storm Prediction Center (SPC)—with leads of 4–8 days. SPC forecasts have shown increasing skill with respect to time since the mid-1990s (Hitchens and Brooks 2012, 2014; Herman et al. 2018); however, consistent skillful forecasts of *daily* tornado and hail activity are not likely to be found beyond day 9 and 12, respectively, if solely using current atmospheric numerical weather prediction systems (Gensini and Tippett 2019). Despite the limitations of dynamical prediction, other subseasonal forecasting techniques can act as a “bridge” to extend prediction beyond the traditional limits of NWP (Zhang 2013).

Extended anomalous periods of tornado frequency, such as the events of May 2003 (Hamill et al. 2005), April 2011 (Doswell et al. 2012; Simmons and Sutter 2012; Knupp et al. 2014), and more recently in May 2019 (Gensini et al. 2019) strongly motivate both this project and future work focusing on subseasonal prediction of SCSs. While dynamical model skill begins to decay during the early portion of this forecast window, there are still forecast approaches (e.g., statistical analogs, model blends, linear inverse models, machine learning) that can achieve statistically significant skill relative to a persistence or climatological forecast. In addition, forecast lead times beyond one week would benefit from relaxing the time and space constraints of the traditional daily forecast target (e.g., larger verification area, forecast target of multiple days or a week), constraints that are related to increasing error growth and uncertainty with respect to time. This has been demonstrated with dynamical forecasts from NCEP’s Climate Forecast System, version 2, by using monthly forecast targets (Lepore et al. 2018).

With these thoughts in mind, the Extended-Range Tornado Activity Forecast (ERTAF) project was formed in March 2015 with the primary goal of assessing the feasibility of operational week 2 and week 3 forecasts of U.S. tornadic activity. Over the past five years, during the period 1 March–31 May, ERTAF participants have been using a variety of dynamical and statistical predictors to produce simple tercile forecasts of tornado activity for the week 2 and week 3 lead periods. In this article, we highlight both the successes and failures of our approaches, the tools that have been found to be useful, lessons we have learned throughout this process, and reflect on whether such forecasts could be introduced into a more formal operational setting.

Forecasting approaches

ERTAF forecasting methodologies leverage both dynamical and statistical guidance, with dynamical guidance playing a greater role in early portions of week 2 due to the waning skill in week 3 (Gensini and Tippett 2019). The general approach focuses on the planetary- to synoptic-scale physical mechanisms, rather than potential specific mesoscale details of any given event. Our experience suggests that a skillful pathway to forecast SCS events at subseasonal lead times is to rely on synoptic-scale Rossby wave configurations and the tropical/extratropical processes that influence them.

The forecast target is predicting the terciles of weekly CONUS tornado counts relative to climatology (1986–2015) during weeks 2 and 3. Possible forecasts include below average (“BA”; <75% of climatology), average (“A”; 75%–125% of climatology), and above average (“AA”; >125% of climatology). A typical forecast discussion for a given week begins with the discussion of verification from the previous week, and a score of –1, 0, or 1 is assigned to the previous week’s forecast. Scores of –1 are recorded when tercile forecasts are two categories off (e.g., forecast of “BA”, but verified as “AA”), 0 when tercile forecasts are one category off (e.g., forecast of

“AA”, but verified as “A”), and 1 when the correct tercile forecast was made (e.g., forecast and verification of BA conditions). The team revisits its reasoning from the forecast (from two and three weeks prior) and discusses the evolution of the global circulation related to what the team expected to occur.

The ERTAF team then shifts its focus on the creation of the week 2 forecast, and that process begins with analysis and discussion of major features in the observed global circulation. This includes examination of current MJO and GWO phases, spatial patterns of global SST anomalies, global and regional mountain/friction torque budgets, global satellite imagery, background seasonal oceanic/atmospheric base states (e.g., ENSO), and deliberation about how these features may impact the forecast periods. Dynamical forecasts are considered from the NCEP’s Global Ensemble Forecasting System, the European Centre for Medium-Range Weather Forecasts Ensemble, and the NCEP’s Climate Forecast System, version 2. In conjunction with the previous products, these models are investigated and aid in the subjective tercile forecast for week 2. Ensemble output is heavily used in the assessment of forecast confidence. As an example, we have found great utility in examining “chiclet”-style plots (Carbin et al. 2016) that display standardized anomalies of the daily CONUS coverage of the supercell composite parameter (SCP; Fig. 1a) and spatial plots of member SCP accumulation for the week (Fig. 1b). The SCP is a dimensionless combination of 0–6-km vertical wind shear, convective available potential energy, and 0–3-km storm-relative helicity (Thompson et al. 2003).

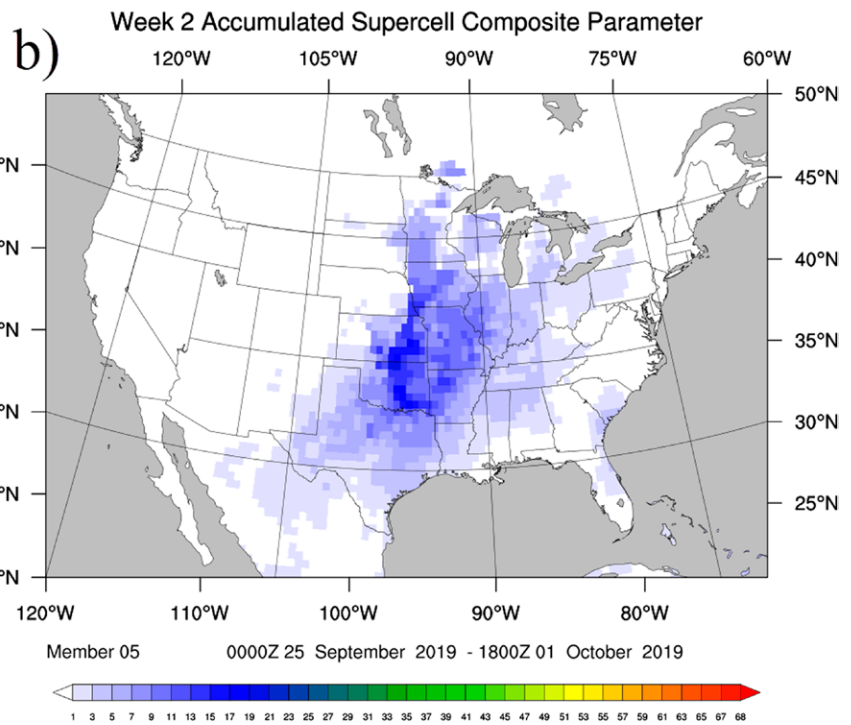
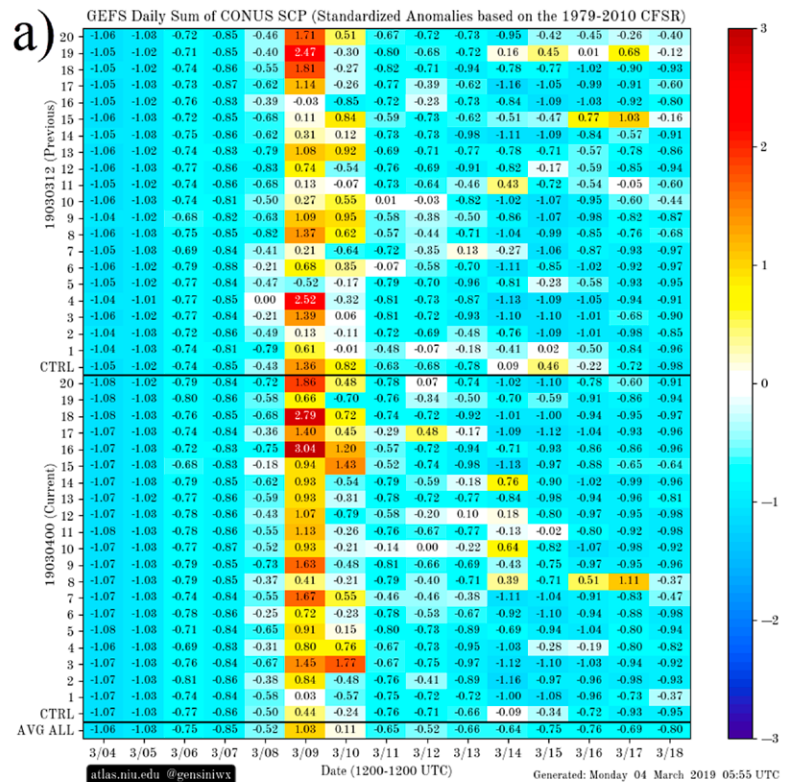


Fig. 1. Examples of (a) Global Ensemble Forecast System (GEFS) chiclet chart. Standardized anomalies of the daily sum of the supercell composite are plotted as a function of daily lead time (x axis) and ensemble member (y axis) and (b) a spatial map of GEFS weekly supercell composite accumulation. Both products (and their derivatives) have helped the ERTAF project make skillful forecasts for week 2 tornado activity.

For week 3, the forecast focus is typically on anticipated changes to the underlying physical “base state” of the atmosphere, where we tended to give greater weight to the leading modes of subseasonal variability (e.g., MJO, GWO) through use of dynamical forecasts and statistical analogs. Forecasts of Real-time Multivariate MJO (RMM) phase space from the Climate Prediction Center are heavily utilized, in addition to some experimental products developed by the ERTAF team (e.g., Climate Forecast System, version 2, predictions of GWO phase space; Fig. 2). The team tended to default to “A” conditions for week 3 if the leading modes of subseasonal variability were not coherent (e.g., a neutral MJO or large spread in the ensemble members), as this uncertainty reduced the team’s forecast confidence. In addition to the tercile forecasts for both weeks, textual discussions were also provided as reasoning for the forecasts. For example, on 28 April 2019, the forecast discussion issued read as follows:

Week 3 (and beyond) continues to focus on the propagation of the current convective activity over the Indian Ocean. This signal is forecast to shift into the western hemisphere by week 2 / end of week 3. We are unsure of the exact timing that could accompany an associated jet extension / collapse and associated potential for AA conditions at this time. GEFS and ECMWF ensembles both suggest that this process will be starting around the middle part of week 3. This would suggest that any subseasonal signal for an AA forecast would hold off until beyond 18 May. We will watch this process closely and update next week accordingly.

—Forecasters: Gensini, Allen, Gold, Barrett

This discussion was the second in a series of forecast discussions leading up to the anomalously high tornado activity of late May 2019 (Gensini et al. 2019).

In the spring of 2018, the ERTAF team added areal outlooks for week 2 tornado frequency if the week 2 tercile was forecast to be “AA.” This was introduced in an effort to explore the feasibility of forecasting the spatial locations of tornado occurrence during the week 2 period. While no objective definition exists for the week 2 areal outlook, the ERTAF team decided to simply outline a polygonal area across the CONUS where forecasters felt the greatest density of tornadoes would occur. While this is the newest aspect of the project, areal outlooks highlighting the potential for tornado events in week 2 do appear

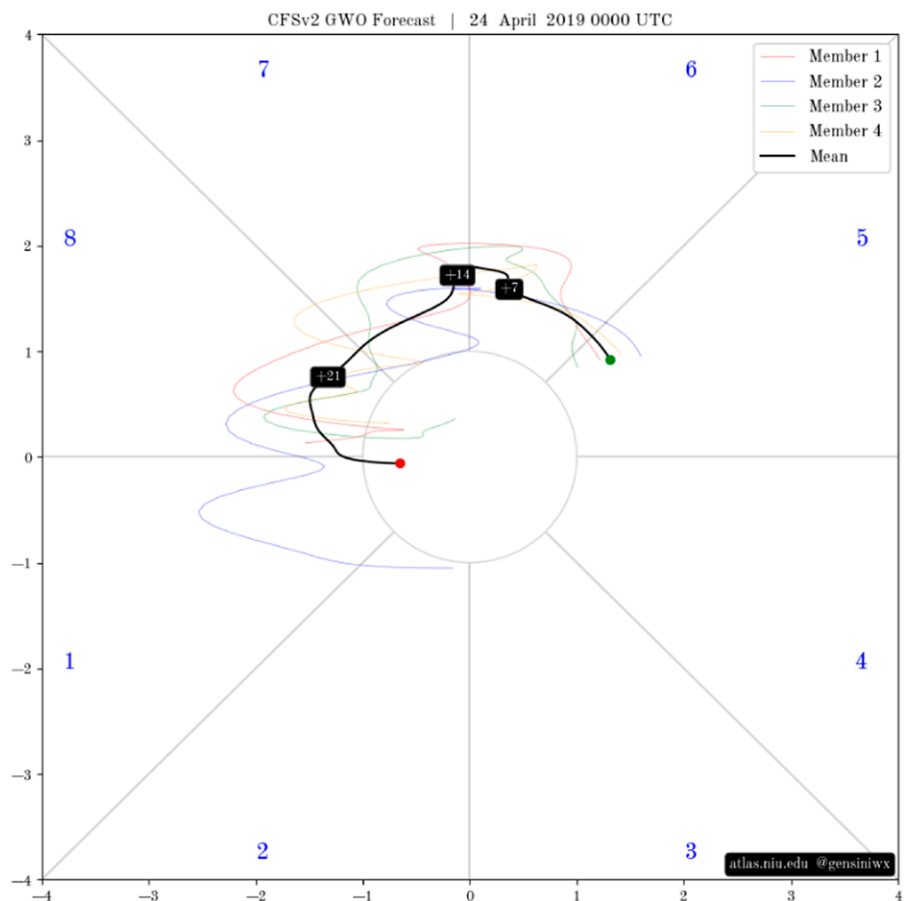


Fig. 2. The 4-week CFSv2 prediction (four-member ensemble) of the global wind oscillation (GWO) initialized 0000 UTC 24 Apr 2019. The green (red) dot indicates the start (end) of the 28-day forecast period. The blue number corresponds to the GWO phase.

to be feasible (Fig. 3), especially during forecasts of opportunity (Gensini et al. 2019). In the future, we plan to calculate weekly climatological probabilities of tornadoes within a specified radius of a point, and then create areal outlooks based on the probability of climatological exceedance to make this a more objective and meaningful process.

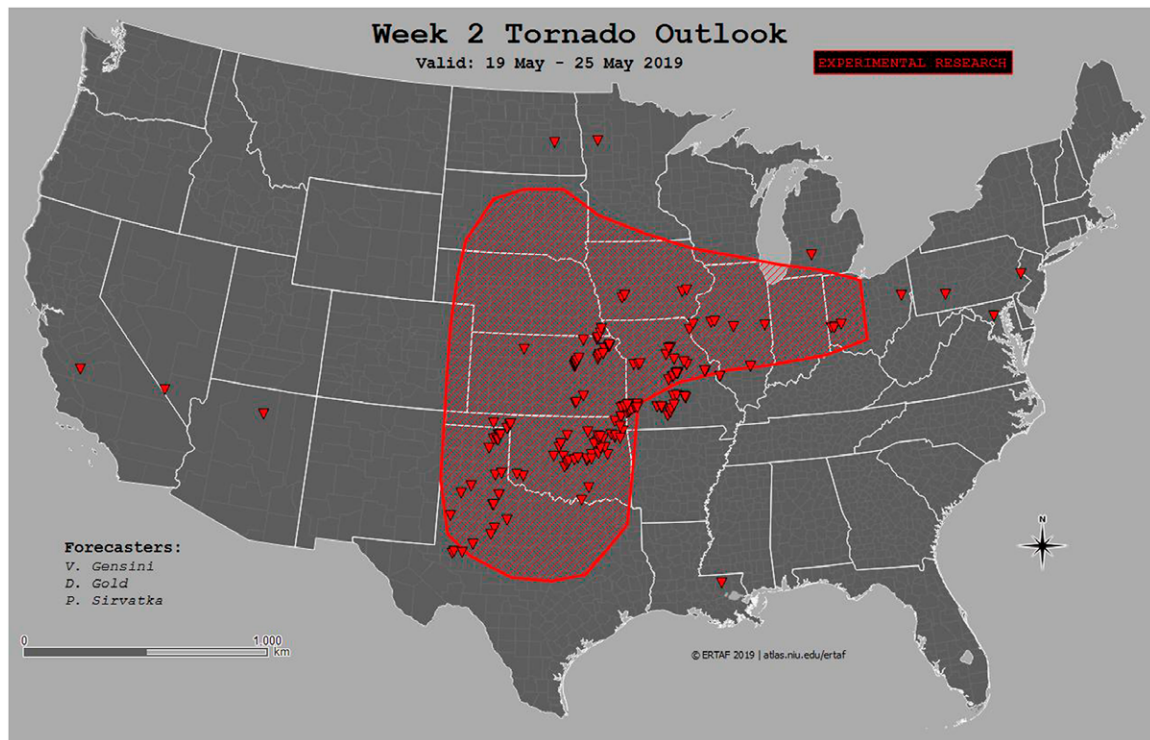


Fig. 3. ERTAF areal tornado outlook (red polygon) issued on 12 May 2019, valid for 19–25 May 2019. Red triangles indicate 19–25 May 2019 preliminary tornado reports.

Verification results

ERTAF tercile forecasts were first assessed for skill using the Heidke skill score (HSS). HSS measures the fraction of correct forecasts after eliminating those forecasts that would be correct due purely to random chance (Wilks 2011). HSS values for week 2 forecasts ranged from 0.11 in 2016 to 0.49 in 2019, with a value of 0.37 for all forecasts over the 5-yr period (Fig. 4a). HSS values for week 2 were positive for all years, suggesting skill in all years over a random forecast. HSS values for week 3 forecasts ranged from -0.20 in 2018 to 0.44 in 2017, with a value of 0.23 for all forecasts over the 5-yr period (Fig. 4b). To provide further context, HSS values were also

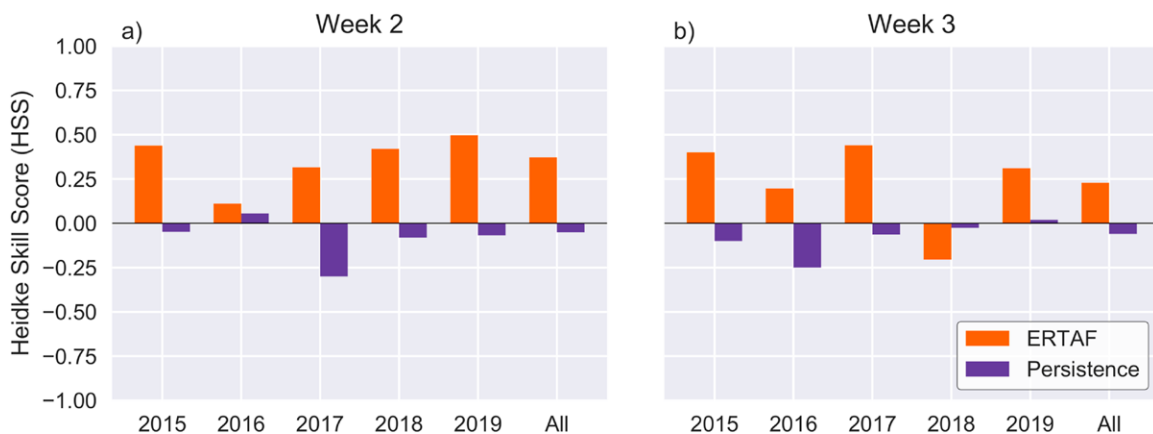


Fig. 4. ERTAF and persistence Heidke skill scores for (a) week 2 and (b) week 3.

calculated for a persistence forecast. A persistence forecast as a reference makes more sense for this application, as a climatology forecast would always be “A” and have a HSS of zero. The persistence forecast used herein was defined by the tornado activity tercile (“BA,” “A,” or “AA”) valid for the 7-day period leading up to the date the ERTAF team met to issue the week 2 and week 3 forecasts. ERTAF week 2 forecasts were more skillful than persistence in each year, with HSS values ranging from 0.06 to 0.62 greater than persistence. HSS values were found to be 0.42 greater than persistence for all week 2 forecasts during the 2015–19 period. ERTAF week 3 forecasts were found to have a greater HSS value than persistence for all years except 2018, with all forecasts over the 5-yr period (0.23) being 0.29 greater than persistence (–0.06). The driver of the poor performance for 2018 was a tendency for ERTAF forecasts to default to a forecast of “A” due to the absence of strong signals from the leading modes of variability. This resulted in a lower success rate in the face of many “BA” verifications.

In addition to HSS, the critical success index (CSI) was calculated as an objective measure of categorical forecast performance. CSI is calculated as the ratio of correct forecasts to the sum of the correct forecasts, misses, and false alarms. Thus, a collection of perfect forecasts (with no misses or false alarms) would have a CSI of 1 (Wilks 2011). CSI values for week 2 “BA” (Fig. 5a) and “AA” (Fig. 5c) forecasts were higher than for “A” forecasts (Fig. 5b). When examining all years and forecast terciles, the greatest fractional improvement over persistence was found for week 2 “AA” forecasts. Persistence correctly forecast these events 8% of the time, whereas ERTAF recorded correct week 2 “AA” forecasts 38% of the time (Fig. 5c). Week 2 CSI values ranged from 0.83 (week 2 “BA” forecasts in 2015) to 0 (week 2 “A” forecasts in 2016). For week 3, the greatest improvement in ERTAF versus persistence CSI values was found for “AA” forecasts (27% improvement; Fig. 5f). When all years are aggregated, the ERTAF team outperformed (a higher CSI) a persistence forecast for all three terciles and for both weeks.

As a final evaluation metric—and to help uncover a source of forecast error—bias was calculated for each forecast tercile (“BA,” “A,” and “AA”) at each lead period (week 2 and 3) in each year (2015–19), and then compared to bias in the persistence forecast. Bias measures the fraction of events forecast compared to the total number of events that occur. An unbiased collection of forecasts has a value of 1.0, while overforecasting would have a bias >1.0 and underforecasting would have a bias <1.0 (Wilks 2011). While interannual variability is

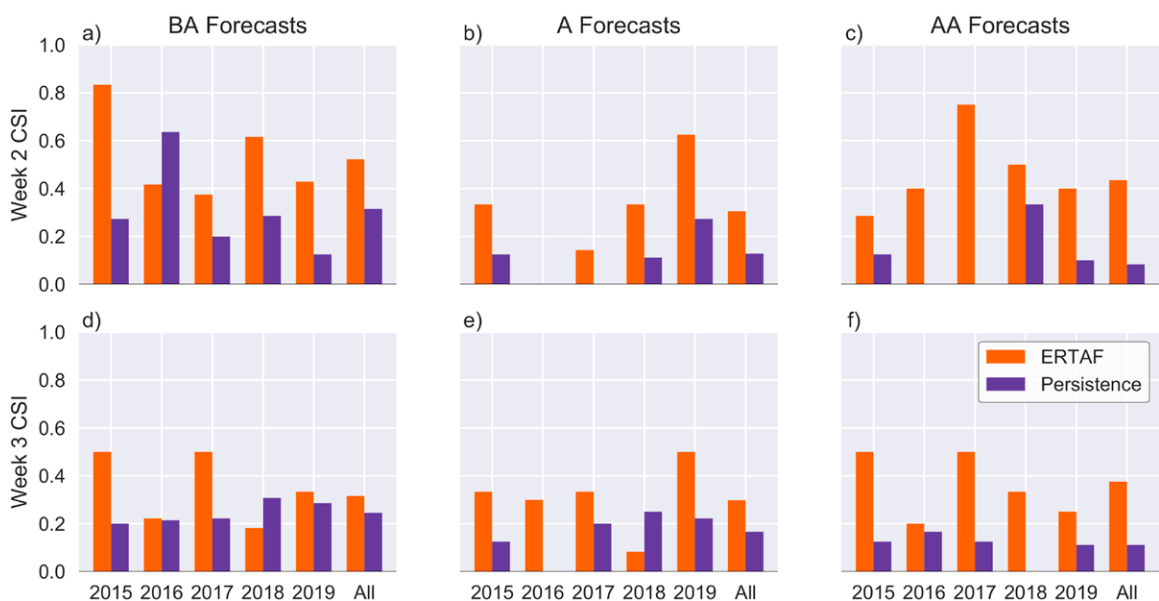


Fig. 5. ERTAF and persistence critical success index (CSI) scores for (a) “BA” week 2, (b) “A” week 2, (c) “AA” week 2, (d) “BA” week 3, (e) “A” week 3, and (f) “AA” week 3 forecasts.

present, week 2 forecasts for all years showed little to no bias (Fig. 6a). For week 3, a common theme of overforecasting “A” conditions is noted, at roughly the equal expense of “BA” and “AA” forecasts (Fig. 6b). This is due to an ERTAF team forecasting philosophy of issuing an “A” forecast (i.e., climatology) for the week 3 period in the absence of strong signals from the leading modes of subseasonal variability.

In summary, the first five years of the ERTAF project have demonstrated skill in the week 2 and 3 forecasting of CONUS tercile tornado frequency categories as compared to a persistence forecast. Skill scores are generally higher for week 2 versus week 3, with a 14% reduction in the HSS from week 2 to week 3. Encouragingly, the ERTAF project has shown some of the highest skill scores during periods of “AA” conditions, which could be argued as the most important category for potential impacts to lives and property. Finally, while ERTAF forecasts could be considered unbiased for week 2, our results indicate that “A” forecasts were overused in the week 3 period. This could be targeted as a source of forecast error for the future improvement of week 3 forecasts.

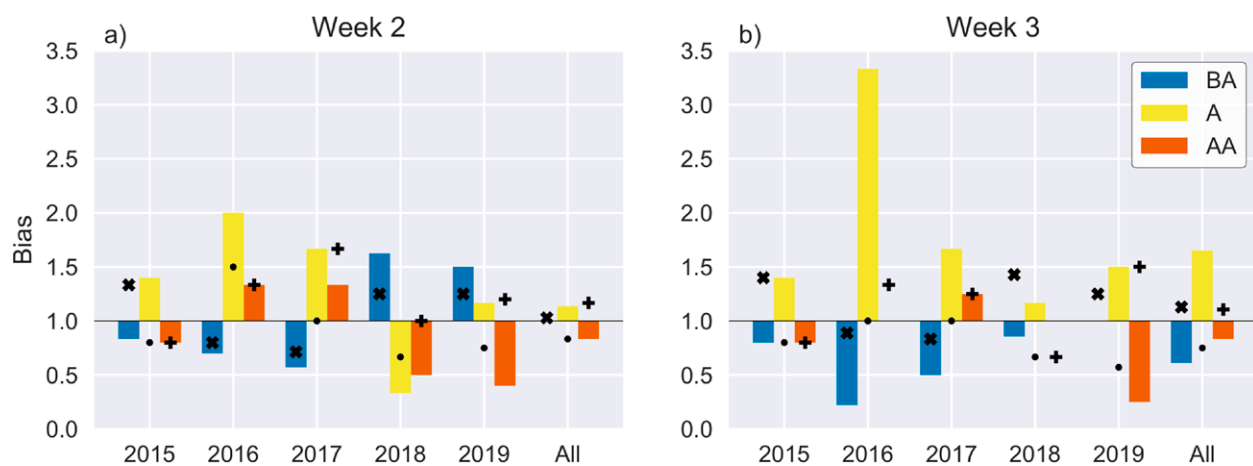


Fig. 6. ERTAF tercile forecast biases by year for (a) week 2 and (b) week 3 leads. Persistence forecast biases are shown as black x (“BA”), • (“A”), and + (“AA”) symbols for comparison.

Concluding remarks

As Brooks (2007) stated, “forecasters can be put in the difficult position of having to issue forecasts when the state of knowledge is less than perfect.” At the time scales involved in the ERTAF project, the authors find that an ingredients-based forecast approach must be employed in conjunction with the understanding of synoptic- and planetary-scale Rossby wave configurations that augment or diminish tornado event probabilities. In addition to the simple severe thunderstorm ingredients (i.e., source of lift to the level of free convection, adequate surface moisture, convective instability, and sufficient vertical wind shear), tornado occurrence also requires the understanding of capping inversion sources, thunderstorm morphology, mesoscale forcing for ascent, variability in lifting condensation level heights, low-level storm relative helicity, and more. ERTAF members had to synthesize multiple, sometimes offsetting, scenarios where tornado occurrence might be minimized or maximized. Discussions often included assessing such favorable environments heuristically, including the subjective assessments of 1) quality of moisture return, 2) an overly meridional jet configuration, 3) potential timing of significant shortwaves, and 4) extreme capping associated with elevated mixed layers that may present in the CONUS Rossby wave configuration.

Much of the discussions for the week 3 period centered on trying to find a forecast of opportunity (Gensini et al. 2019), a moment where such ingredients may be favorably juxtaposed to support tornado activity, given pattern recognition or statistical links to known sources

of subseasonal tornado frequency variability. Extrapolation of long-range deterministic and ensemble NWP guidance allowed for subjective estimates of the favorability of large-scale environments to local tornado events (e.g., confidence of a longwave thermal trough entering the western CONUS). In addition, while we have no way to objectively measure this at the current time, we believe that a contributing factor to the skill found in week 2 and week 3 forecasts is related to forecaster experience (Hoffman et al. 2017). Each forecaster's individual past experience and heuristic insight contributed akin to an ensemble member, resulting in a final blended forecast analogous to crowd-sourcing methods (Muller et al. 2015).

ERTAF members found that one of the greatest challenges in producing long-lead tornado activity-level forecasts was the definition of what to use as a forecast target. ERTAF members have discussed the use of tornado days, significant tornado days, tornado counts, tornado environments, standardized count anomalies, and even targeting temporal rolling windows instead of fixed calendar weeks. Specifically, the exact timing of an event has proven to be an elusive quantity to target at weeks 2 and 3, which is partly driven by the increasing forecast uncertainty as a function of lead time. Anticipated week 2 conditions to support "AA" often "split" across weeks 2 and 3, leading to a verification of "A" conditions in both periods. The authors continue to discuss best practices and future options for forecast targets, both in space and time.

Finally, the authors believe skillful predictions of U.S. tornado frequency at lead times of 2 and 3 weeks are operationally feasible with current tools and the current state of scientific knowledge. While the historical forecast approach has been admittedly simple (i.e., tercile forecasts of CONUS activity levels for an entire week), it has proven to be skillful versus both persistence and random reference forecasts. Nevertheless, a significant amount of work will be needed in the future to refine forecast targets, understand the best predictors, create new forecast methods, and identify relationships with other modes of climate variability—with the overall goal of improving SCS prediction in the subseasonal forecast horizon.

Acknowledgments. ERTAF has been an unfunded effort, which has meant that members of the group have changed through time and even from week to week based on availability, as committing to a meeting every Sunday in the spring for an hour or more has proven to be a significant responsibility. We acknowledge the contributions made by both Mr. Al Marinaro and Dr. Mike Ventrice, who have both taken part in ERTAF efforts over the years. The authors wish to acknowledge the pioneering work of Dr. Klaus Weickmann and Mr. Ed Berry, whose research and contributions more than a decade ago provided the basis for many of the ideas implemented to make the ERTAF project possible. We also acknowledge the honest feedback, perspectives, and comments that have helped encourage and refine our approach to extended-range tornado forecasting, including conversations with numerous operational forecasters and interactions with several individuals interested in subseasonal forecasting. Finally, three anonymous reviewers provided constructive feedback on an initial draft of this manuscript.

References

- Allen, J. T., M. K. Tippett, and A. H. Sobel, 2015: Influence of the El Niño–Southern Oscillation on tornado and hail frequency in the United States. *Nat. Geosci.*, **8**, 278–283, <https://doi.org/10.1038/ngeo2385>.
- Baggett, C. F., K. M. Nardi, S. J. Childs, S. N. Zito, E. A. Barnes, and E. D. Maloney, 2018: Skillful subseasonal forecasts of weekly tornado and hail activity using the Madden-Julian oscillation. *J. Geophys. Res. Atmos.*, **123**, 12 661–12 675, <https://doi.org/10.1029/2018JD029059>.
- Barrett, B. S., and V. A. Gensini, 2013: Variability of central United States April–May tornado day likelihood by phase of the Madden-Julian oscillation. *Geophys. Res. Lett.*, **40**, 2790–2795, <https://doi.org/10.1002/grl.50522>.
- , and B. N. Henley, 2015: Intraseasonal variability of hail in the contiguous United States: Relationship to the Madden–Julian oscillation. *Mon. Wea. Rev.*, **143**, 1086–1103, <https://doi.org/10.1175/MWR-D-14-00257.1>.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Brooks, H. E., 2007: Ingredients-based forecasting. *Atmospheric Convection: Research and Operational Forecasting Aspects*, Springer, 133–140.
- Carbin, G. W., M. K. Tippett, S. P. Lillo, and H. E. Brooks, 2016: Visualizing long-range severe thunderstorm environment guidance from CFSv2. *Bull. Amer. Meteor. Soc.*, **97**, 1021–1031, <https://doi.org/10.1175/BAMS-D-14-00136.1>.
- Doswell, C. A., III, G. W. Carbin, and H. E. Brooks, 2012: The tornadoes of spring 2011 in the USA: An historical perspective. *Weather*, **67**, 88–94, <https://doi.org/10.1002/wea.1902>.
- Gensini, V. A., and A. Marinaro, 2016: Tornado frequency in the United States related to global relative angular momentum. *Mon. Wea. Rev.*, **144**, 801–810, <https://doi.org/10.1175/MWR-D-15-0289.1>.
- , and J. T. Allen, 2018: US hail frequency and the global wind oscillation. *Geophys. Res. Lett.*, **45**, 1611–1620, <https://doi.org/10.1002/2017GL076822>.
- , and M. K. Tippett, 2019: Global Ensemble Forecast System (GEFS) predictions of days 1–15 US tornado and hail frequencies. *Geophys. Res. Lett.*, **46**, 2922–2930, <https://doi.org/10.1029/2018GL081724>.
- , D. Gold, J. T. Allen, and B. S. Barrett, 2019: Extended U.S. tornado outbreak during late May 2019: A forecast of opportunity. *Geophys. Res. Lett.*, **46**, 10 150–10 158, <https://doi.org/10.1029/2019GL084470>.
- Grice, G., and Coauthors, 1999: The golden anniversary celebration of the first tornado forecast. *Bull. Amer. Meteor. Soc.*, **80**, 1341–1348, [https://doi.org/10.1175/1520-0477\(1999\)080<1341:TGACOT>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<1341:TGACOT>2.0.CO;2).
- Hamill, T. M., R. S. Schneider, H. E. Brooks, G. S. Forbes, H. B. Bluestein, M. Steinberg, D. Meléndez, and R. M. Dole, 2005: The May 2003 extended tornado outbreak. *Bull. Amer. Meteor. Soc.*, **86**, 531–542, <https://doi.org/10.1175/BAMS-86-4-531>.
- Herman, G. R., E. R. Nielsen, and R. S. Schumacher, 2018: Probabilistic verification of storm prediction center convective outlooks. *Wea. Forecasting*, **33**, 161–184, <https://doi.org/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, <https://doi.org/10.1175/WAF-D-12-00061.1>.
- , and ———, 2014: Evaluation of the Storm Prediction Center's convective outlooks from day 3 through day 1. *Wea. Forecasting*, **29**, 1134–1142, <https://doi.org/10.1175/WAF-D-13-00132.1>.
- Hoffman, R. R., D. S. LaDue, H. M. Mogil, P. J. Roebber, and J. G. Trafton, 2017: *Minding the Weather: How Expert Forecasters Think*. MIT Press, 488 pp.
- Knupp, K. R., and Coauthors, 2014: Meteorological overview of the devastating 27 April 2011 tornado outbreak. *Bull. Amer. Meteor. Soc.*, **95**, 1041–1062, <https://doi.org/10.1175/BAMS-D-11-00229.1>.
- Lepore, C., M. K. Tippett, and J. T. Allen, 2017: ENSO-based probabilistic forecasts of March–May US tornado and hail activity. *Geophys. Res. Lett.*, **44**, 9093–9101, <https://doi.org/10.1002/2017GL074781>.
- , ———, and ———, 2018: CFSv2 monthly forecasts of tornado and hail activity. *Wea. Forecasting*, **33**, 1283–1297, <https://doi.org/10.1175/WAF-D-18-0054.1>.
- Moore, T. W., 2018: Annual and seasonal tornado activity in the United States and the global wind oscillation. *Climate Dyn.*, **50**, 4323–4334, <https://doi.org/10.1007/s00382-017-3877-5>.
- , and M. P. McGuire, 2019: Tornado-days in the United States by phase of the Madden–Julian oscillation and global wind oscillation. *Climate Dyn.*, **54**, 17–36, <https://doi.org/10.1007/S00382-019-04983-Y>.
- Muller, C., L. Chapman, S. Johnston, C. Kidd, S. Illingworth, G. Foody, A. Vereem, and R. Leigh, 2015: Crowdsourcing for climate and atmospheric sciences: Current status and future potential. *Int. J. Climatol.*, **35**, 3185–3203, <https://doi.org/10.1002/joc.4210>.
- NCEI, 2019: U.S. billion-dollar weather and climate disasters. NOAA, www.nccdc.noaa.gov/billions/.
- Robertson, A. W., A. Kumar, M. Peña, and F. Vitart, 2015: Improving and promoting subseasonal to seasonal prediction. *Bull. Amer. Meteor. Soc.*, **96**, ES49–ES53, <https://doi.org/10.1175/BAMS-D-14-00139.1>.
- Schubert, S., R. Dole, H. vandenDool, M. Suarez, and D. Waliser, 2002: Prospects for improved forecasts of weather and short-term climate variability on sub-seasonal (2-week to 2-month) times scales. NASA Tech. Memo. NASA/TM-2002-104606, Vol. 23, 171 pp.
- Simmons, K. M., and D. Sutter, 2012: The 2011 tornadoes and the future of tornado research. *Bull. Amer. Meteor. Soc.*, **93**, 959–961, <https://doi.org/10.1175/BAMS-D-11-00126.1>.
- , and ———, 2013: *Economic and Societal Impacts of Tornadoes*. Springer Science and Business Media, 296 pp.
- Smith, A. B., and J. L. Matthews, 2015: Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Nat. Hazards*, **77**, 1829–1851, <https://doi.org/10.1007/s11069-015-1678-x>.
- Thompson, D. B., and P. E. Roundy, 2013: The relationship between the Madden–Julian oscillation and U.S. violent tornado outbreaks in the spring. *Mon. Wea. Rev.*, **141**, 2087–2095, <https://doi.org/10.1175/MWR-D-12-00173.1>.
- Thompson, R. L., R. Edwards, J. A. Hart, K. L. Elmore, and P. Markowski, 2003: Close proximity soundings within supercell environments obtained from the Rapid Update Cycle. *Wea. Forecasting*, **18**, 1243–1261, [https://doi.org/10.1175/1520-0434\(2003\)018<1243:CPSWSE>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<1243:CPSWSE>2.0.CO;2).
- Tian, D., E. Wood, and X. Yuan, 2017: CFSv2-based sub-seasonal precipitation and temperature forecast skill over the contiguous United States. *Hydrol. Earth Syst. Sci.*, **21**, 1477–1490, <https://doi.org/10.5194/hess-21-1477-2017>.
- Tippett, M. K., 2018: Robustness of relations between the MJO and U.S. tornado occurrence. *Mon. Wea. Rev.*, **146**, 3873–3884, <https://doi.org/10.1175/MWR-D-18-0207.1>.
- Trapp, R. J., 2014: On the significance of multiple consecutive days of tornado activity. *Mon. Wea. Rev.*, **142**, 1452–1459, <https://doi.org/10.1175/MWR-D-13-00347.1>.
- Verbout, S. M., H. E. Brooks, L. M. Leslie, and D. M. Schultz, 2006: Evolution of the U.S. tornado database: 1954–2003. *Wea. Forecasting*, **21**, 86–93, <https://doi.org/10.1175/WAF910.1>.
- Vigaud, N., A. W. Robertson, and M. K. Tippett, 2017: Multimodel ensembling of subseasonal precipitation forecasts over North America. *Mon. Wea. Rev.*, **145**, 3913–3928, <https://doi.org/10.1175/MWR-D-17-0092.1>.
- Wanders, N., and E. F. Wood, 2016: Improved sub-seasonal meteorological forecast skill using weighted multi-model ensemble simulations. *Environ. Res. Lett.*, **11**, 094007, <https://doi.org/10.1088/1748-9326/11/9/094007>.
- Wang, L., and A. W. Robertson, 2019: Week 3–4 predictability over the United States assessed from two operational ensemble prediction systems. *Climate Dyn.*, **52**, 5861–5875, <https://doi.org/10.1007/s00382-018-4484-9>.
- Weickmann, K., and E. Berry, 2009: The tropical Madden–Julian oscillation and the global wind oscillation. *Mon. Wea. Rev.*, **137**, 1601–1614, <https://doi.org/10.1175/2008MWR2686.1>.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Academic Press, 704 pp.
- Zhang, C., 2013: Madden–Julian oscillation: Bridging weather and climate. *Bull. Amer. Meteor. Soc.*, **94**, 1849–1870, <https://doi.org/10.1175/BAMS-D-12-00026.1>.