Forecasting Tornado Activity Weeks in Advance

Adapted from "The Extended-Range Tornado Activity Forecast (ERTAF) Project," by **Vittorio A. Gensini** (Northern Illinois University), **Bradford S. Barrett**, **John T. Allen**, **David Gold**, and **Paul Sirvatka**. Published online in *BAMS*, June 2020. For the full, citable article, see [DOI:10.1175/BAMS-D-19-0188.1](https://doi.org/10.1175/BAMS-D-19-0188.1).

A spatial map of GEFS weekly supercell composite accumulation. This product (and its derivatives) helped the ERTAF project make skillful forecasts for week-2 tornado activity.

Skill in forecasting severe convective storms (SCSs) and associated tornado and hail frequency is increasing at various subseasonal lead times, from two weeks to two months. Interest from stakeholders is rapidly growing for these forecasts.

Subseasonal forecast skill is likely a testament to improved global numerical weather prediction and better understanding of the variance in subseasonal SCS frequency—for example, the clustering of periods of high tornado frequency over two or more days. In addition, the Madden–Julian oscillation (MJO) and the global wind oscillation (GWO) can lead to synoptic patterns supportive of tornadoes east of the U.S. Rocky Mountains, such as a midtropospheric trough in the western United States and consequent polar flux in surface moisture across the Great Plains.

Operational 4–8-day SPC outlooks for SCSs now have measurable skill. Extended anomalous periods of tornado frequency, as occurred in May 2003, April 2011, and May 2019, strongly motivate our work in developing skillful subseasonal outlooks. While dynamical model skill begins to decay during the early portion

of this forecast window, there are still alternative approaches (e.g., statistical analogs, model blends, linear inverse models, machine learning) that can produce extended forecasts superior to persistence or climatology. In addition, forecast lead times beyond one week would benefit from using broader forecast targets.

With these thoughts in mind, we formed the Extended-Range Tornado Activity Forecast (ERTAF) project in March 2015 with the primary goal of assessing the feasibility of operational week 2 and week 3 forecasts of U.S. tornadic activity. In this article, we highlight both the successes and failures of our approaches to this challenge, the useful tools, and the lessons learned, and we reflect on whether such forecasts could be introduced into a more formal operational setting.

Forecasting approaches

ERTAF leverages both dynamical and statistical guidance, with the former emphasized in week 2 forecasts. The general approach focuses on the planetary- to synoptic-scale physical mechanisms rather than mesoscale details. Our experience suggests that relying on synoptic-scale Rossby wave configurations and

▲ * 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.

the tropical/extratropical processes that influence them yield more skillful results.

We predict weekly continental United States (CONUS) tornado counts relative to climatology (1986–2015) during weeks 2 and 3, using terciles of below average (BA; <75% of climatology), average (A; 75%–125% of climatology), and above average (AA; >125% of climatology). Verification scores are –1 for when forecasts are two terciles off (e.g., forecast of BA, but verified as AA), 0 for one tercile off, and 1 for the correct tercile.

The ERTAF team then creates a subjective tercile forecast for week 2, beginning with analysis and discussion of the global circulation. This includes current MJO and GWO phases, spatial patterns of global SST anomalies, global and regional mountain/friction torque budgets, global satellite imagery, and background seasonal oceanic/atmospheric base states (e.g., ENSO). Dynamical forecasts are considered from the National Centers for Environmental Prediction's (NCEP's) Global Ensemble Forecast System (GEFS) and Climate Forecast System, version 2, and the European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble. Ensemble output is used to assess forecast confidence. As an example,

▲ * ERTAF and persistence Heidke skill scores for (a) week 2 and (b) week 3.

we have found great utility in examining standardized anomalies of the daily CONUS coverage of the supercell composite parameter, which combines 0–6-km vertical wind shear, convective available potential energy, and 0–3-km storm-relative helicity.

For week 3, the forecast typically focuses on anticipated changes to the underlying physical “base state” of the atmosphere, where we tend to give greater weight to the leading modes of subseasonal variability (e.g., MJO, GWO) through use of dynamical forecasts and statistical analogs. The team tends to default to “A” conditions for week 3 if the leading modes of subseasonal variability are not coherent (e.g., a neutral MJO or large spread in the ensemble members). In addition to the tercile forecasts for both weeks, we provide textual reasoning with the forecasts. For example, on 28 April 2019, the forecast discussion read:

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.

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

In the spring of 2018, the ERTAF team added areal outlooks for week 2 tornado frequency when the forecast is “AA.” This delineates locations under greatest tornado threats within polygons. 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 process more objective and meaningful.

Verification results

ERTAF tercile forecasts were first assessed with 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—calculated for a persistence forecast as opposed to climatology. Persistence was defined by the 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 during the 2015–2019 period, with HSS values ranging from 0.06 to 0.62 greater than persistence. ERTAF week 3 forecasts had 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). In 2018, ERTAF tended to default to a forecast of “A” due to the absence of strong signals from the leading modes of variability, when in fact “BA” verified more often.

In addition to HSS, we calculated the critical success index (CSI), the ratio of correct forecasts to the sum of the correct forecasts, misses, and false alarms. When the CSI for all years are aggregated, the ERTAF team outperformed a persistence forecast for all three terciles and for both weeks. CSI is generally higher for week 2 than week 3. Encouragingly, the ERTAF project has shown some of the highest skill for “AA” conditions, arguably the most impactful category.

Week 3 showed common overforecasting of “A” conditions at roughly the equal expense of “BA” and “AA” forecasts, because the ERTAF team issues an “A” forecast (i.e., climatology) for week 3 in the absence of strong signals from the leading modes of subseasonal variability. While ERTAF forecasts were unbiased for week 2, this overuse of “A” forecasts could be targeted for future improvements.

Concluding remarks

ERTAF project members agree that an ingredients-based forecast approach combined with an understanding of synoptic- and planetary-scale Rossby wave configurations is critical in assessing tornado event probability. ERTAF members often discussed favorable severe thunderstorm 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 Rossby wave configuration.

Many discussions for week 3 centered on trying to find a moment when such ingredients may favor tornado activity. Extrapolation of long-range deterministic and ensemble NWP guidance allowed for subjective estimates of the favorability of large-scale environments to local tornado events. We believe that collective forecaster experience and heuristic insights contribute to the skill of week 2 and week 3 forecasts, akin to an ensemble member.

Finally, the definition of what to use as a forecast target is crucial to producing long-lead tornado activity forecasts. 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. The authors continue to discuss best practices and future options for forecast targets, both in space and time.

While admittedly simple, the ERTAF approach has proven to be skillful and operationally feasible. 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. ●●

METADATA

BAMS: What would you like readers to learn from this article?

Vittorio Gensini (Northern Illinois University): *Our ability to predict severe thunderstorms at time scales ranging from hours to weeks has significantly advanced in the last 10 years. While we cannot forecast a specific time and location of tornadoes at a 2- to 3-week timescale, we have shown skill in predicting the general weather patterns that favor or disfavor severe hail and tornadoes.*

BAMS: How did you become interested in forecasting?

VG: *I've always been interested in forecasting, but I've become quite passionate about the emerging area of S2S (subseasonal-to-seasonal) prediction. Specifically, after co-authoring a paper in Monthly Weather Review in 2016, I became quite interested how Earth-relative atmospheric angular momentum may explain significant amounts of S2S climate variability.*

BAMS: What surprised you the most about the tornado predictions at this range?

VG: *I was surprised we were able to show skill over a persistence forecast for both weeks 2 and 3. Our skill scores were clearly better for week 2, and I think there is still more low-hanging fruit in the prediction of week 2 hazards.*

BAMS: What was the biggest challenge you encountered at extended range?

VG: *We tended not to use any dynamical prediction for week 3 due to the lack of skill, and instead defaulted toward the statistical correlations provided by subseasonal and seasonal drivers of variability. In addition, there were times when we knew a big outbreak of tornadoes was likely toward the end of the week 1 or week 2 periods. The clear temporal decision boundary of calendar weeks is something we hope to change in the future.*

BAMS: What's next?

VG: *Obtain funding for basic research. Unfortunately, we have been unsuccessful in obtaining any grants to answer some basic scientific questions related to S2S prediction of extreme weather events. We are hopeful this will change in the offing!*