

AMS Conference Paper

<https://doi.org/10.25923/vcs3-gn92>

Developing an Updated Analog-Based Severe Weather Forecast System

Cameron J. Nixon^{1*}, Israel L. Jirak², Ryan E. Jewell², Andrew D. Moore²

¹Cooperative Institute for Severe and High-Impact Weather Research and Operations, Norman, OK

²NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK

32nd Conference on Severe Local Storms

August 2026

Link: <https://ams.confex.com/ams/2026Summit/meetingapp.cgi/Paper/482328>

P225 Developing an Updated Analog-Based Severe Weather Forecast System

Cameron J. Nixon^{1*} Israel L. Jirak², Ryan E. Jewell², Andrew D. Moore²

¹Cooperative Institute for Severe and High-Impact Weather Research and Operations, Norman, OK

²NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK

1. Background and Motivation

Predicting the conditional intensity of severe hazards (tornado rating, hail size, and wind speed) has become an integral part of convective forecasts, especially since the operationalization of the Storm Prediction Center's Conditional Intensity Group (CIG) system (Jirak et al. 2024). As such, there is an increasing need for conditional intensity guidance, which depends at least in part on the background environment. In addition, there is an increasing need for better predictions of the *behavior* of the hazard – for instance, could tornadoes be particularly long lived? Could hail swaths be particularly wide? Could any tornadoes acquire deviant motion? Currently, little if any guidance exists for such characteristics.

Though composite parameters such as STP, SHIP, and DCAPE can be related to potential hazard intensity (Jirak et al. 2024), they can still struggle to capture features in a sounding that may be important, and as such, forecasting using these alone is usually discouraged, especially in regions or regimes in which the parameter was not developed to represent (Taszarek et al., 2026).

The original Sounding Analog Retrieval System (SARS, Jewell, 2010) was developed with this in mind, and sought to use more specific characteristics of the environmental sounding profile to find similar analogs, and derive hazard intensities from those. These indices include MUCAPE, 500mb temperature, 700-500mb temperature, and 0-6 km bulk shear (chosen mostly because of their skill in hail prediction). It is frequently used in Storm Prediction Center operations, even still today.

That said, since its development, a number of studies have pointed out inconsistencies in the usefulness of these particular parameters (e.g., large hail does not always occur with large MUCAPE and/or steep low-level lapse rates), and have found other parameters that might be more important (Johnson & Sugden, 2014, Nixon & Allen, 2023, Taszarek et al., 2026). Other studies have even advocated for the use of full-profile analysis (Nowotarski & Jones, 2018, Warren et al., 2021, Nixon & Allen, 2023) to avoid simplifying soundings down to a few simple indices. In other words, while indices and parameters may be skillful for *most* severe events, they may struggle in environments that are different from most severe environments, or in locations that don't receive the most severe events. To illustrate this, **Fig. 1** depicts the wide variety of environments supportive of tornadoes across the U.S., and that there is no one environment that is most favorable for tornadoes.

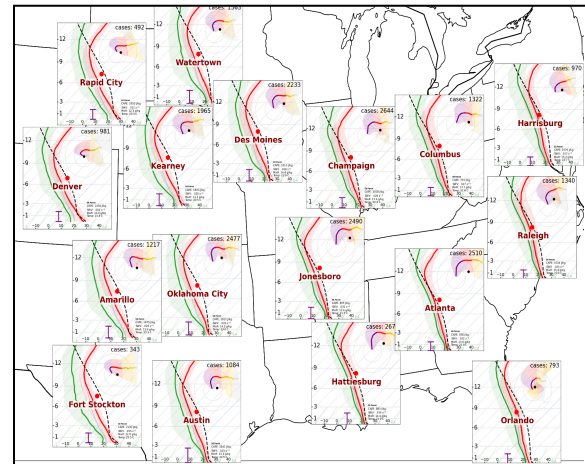


Fig. 1. Median soundings (skew-T and hodograph) for all confirmed tornadoes within 200 miles of each labeled point/city from 2003-2024, illustrating the wide variety of environments that produce tornadoes in the U.S.

These findings and recommendations have encouraged the development of a new sounding analog system that finds matching profiles based on the full shape of the sounding (both skew-T and hodograph), such that does not depend on indices or parameterizations. This system can also utilize an updated and much larger sounding database than the original SARS, and can be trained on tornadoes and wind in addition to hail. Perhaps most powerfully, however, this analog system can be enriched with track data, such as SPC's OneTor (Schaefer & Edwards, 1999) and hail tracks (Nixon et al. 2026). This would allow it to predict a variety of hazard characteristics such as track length, width, duration, and motion, thus potentially providing considerably more holistic value than traditional forecast parameters.

2. Data and Methodology

2.1 Data sources

The sounding data used for initial development were derived from the RAP/RUC SFCOA (Bothwell et al. 2002), a dataset of surface-adjusted SFCOA profiles from 2003-present. Event data included tornado tracks from SPC's OneTor dataset, as well as hail and wind tracks (Nixon et al. 2026) from 2003-2024. A proximity 00hr analysis SFCOA sounding was taken for each event (for over 480,000 events), at the latitude/longitude and time of each

event's start point, and at least 1 hour before the event's start time (in accordance with Jewell, 2010).

2.2 Sounding matching algorithm

The analog system considers input traces of θ (theta), w (mixing ratio), u/v (wind), and p (surface pressure). A rotated 0-6-km shear-relative approach (similar to Nixon & Allen, 2023) resulted in more skilful analog matching, per initial performance tests. At its core, the algorithm computes a simple Euclidean distance, at each sounding level, between the input sounding and all database soundings. It then computes a single similarity score for each database sounding by taking the sum of these distances over the relevant depth of the sounding (9 km was found to be an optimal balance of depth and computation time). In this way, the database soundings with the smallest total distances can be considered the “closest” analogs.

Of course, tuning is necessary to achieve optimal performance. First, not all levels should be weighted equally; for instance, it is well known that low-level features such as shear and buoyancy are more skilful in tornado prediction. Second, not all input variables are the same units; therefore, some weighting of the θ , w , u/v , and p input traces is necessary.

Post-processing of the closest analog matches was also tested to improve the skill of the system. For example, omitting analog soundings that were worse than a user-defined “bad” threshold (such that did not resemble the input sounding) led to more relevant sounding matches, and gave the system the ability to predict null events where close analogs did not exist. In addition, expanding the analog events by grabbing all nearby events within 60 km and 60 minutes of each analog sounding's associated event significantly increased sample size without the need to add more “bad” sounding matches.

2.3 Data Visualization

Since this analog system has the ability to predict a variety of hazard characteristics holistically, a few novel data visualizations are presented here that demonstrate some of its capabilities, many of which were inspired by initial testing in the 2026 SPC/NSSL Hazardous Weather Testbed Spring Forecasting Experiment.

First, the system can provide considerable value when examining modeled (or observed) atmospheric point soundings (an example is shown in Fig. 2). Potential output can include the frequency of each hazard occurring in the analogs (for instance, to assess whether the environment supports a hail and tornado threat, or if tornadoes may occur without hail), and the conditional intensity distributions for each hazard. Track data can also be used to predict track characteristics, most likely storm motion, and most likely track footprints (such that can depict deviant tornado motions).

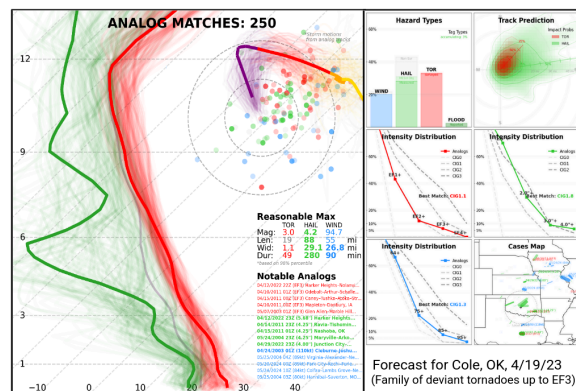


Fig. 2. Example of a variety of possible analog system outputs, including skew-Ts and hodographs, notable analog descriptions, reasonable max event characteristics, hazard types, track prediction (using 2D KDE over all analog tracks), conditional intensity distributions and CIG predictions, and cases map. Though this is not an operational sounding viewer, it is meant to show many of the potential ways analog data can be visualized. Output is for 0000 UTC on 4/20/23.

Secondly, the system can provide 2-dimensional spatial fields of these variables, such that mimic traditional mesoscale analysis fields. This can be used to show areas where conditional intensity is highest, or where each hazard is most favored given the environment (see Fig. 3).

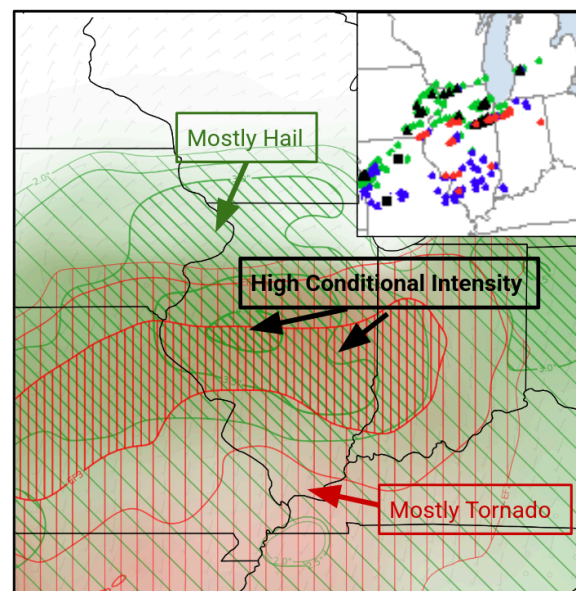


Fig. 3. Example of spatial output depicting predominant analog hazard type (color shading) and 98th percentile reasonable max intensity (hatching) for both tornadoes (red) and hail (green). In this example, the highest conditional intensities for tornadoes and hail overlap, but the highest tornado potential is along and south of a warm front in central IL, where the highest hail potential is along and north of this front. Output is for RAP/RUC SFCOA analysis for 2300 UTC on 3/10/25.

3. Initial Testing

Calibrating this sounding analog system requires finding the “optimal” weights for each input vertical profile (θ , w , u/v , and p) with height. Of course, defining “optimal” is not trivial; unlike most machine-learning algorithms that are trained off a single target variable (such as EF rating), the authors seek to predict a wide variety of storm characteristics (such as length, width, duration, storm motion, etc.). Initial testing considered a random sample of 5,000 tornadoes and ran the matching algorithm 250 times with slightly perturbed weights, to find the configuration that resulted in the lowest (best) ranked probability score (Epstein, 1969) for 4 tornado intensity bins: EF0-1, EF2, EF3, and EF4+. Perhaps unsurprisingly, this matching configuration placed the highest importance on wind profiles in the lowest 500-1000m (such that affect low-level shear), where weights then fell off quickly with height. Temperature was also weighted highest at 1 km, but remained at least moderately important through the depth of the profile (such that was likely related to CAPE). Moisture was weighted highest at the surface level (perhaps reflecting low-level CAPE and/or LCL height), and decreased rapidly with height.

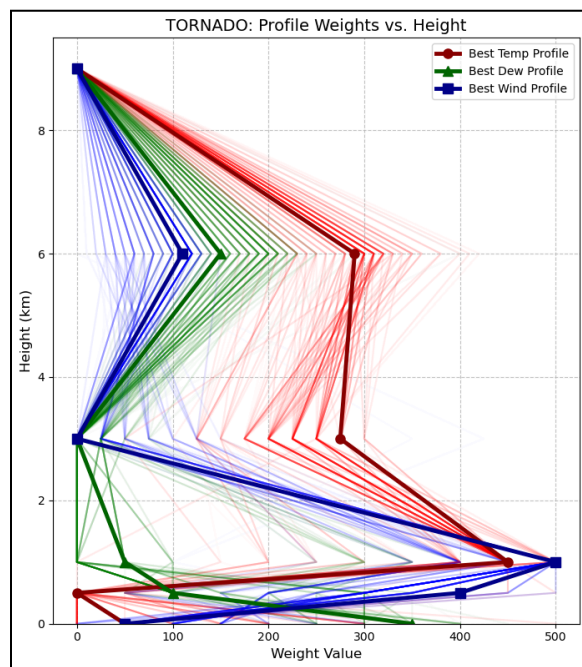


Fig. 4. Optimal weights profile for each input vertical profile such that resulted in the lowest (best) ranked probability score, after 250 tests on a random sample of 5,000 tornadoes.

4. Summary and Future Applications

An updated sounding analog system that uses the full sounding profile to find close analogs and predict a variety of potential hazard characteristics shows promise in offering considerable value to forecasters. Much of this value comes not just from its ability to predict the conditional intensity of hazards, but its ability to provide more holistic predictions of hazard behavior (including track length, duration, and motion) than is possible using more traditional environmental parameters.

Some potential applications of this system include predicting which hazards are actually possible on a given day (is there a tornado and hail risk? A tornado and flash flood risk? A tornado risk only?), and predicting the conditional intensity distribution of hazards (such that can be fitted to and predict SPC's conditional intensity group first-guess distributions). Track-based data can be used to predict the most likely spatial impacts from a hypothetical event (e.g., could a hailstorm in this environment impact a long, wide area? Could tornadoes in this environment move northward, despite most right-moving supercells moving eastward?), and their potential longevity. The performance of the analog matching algorithm, and its optimal weighting configuration, will be continued to be tested in future iterations.

REFERENCES

- Bothwell, P. D., J. A. Hart, and R. L. Thompson, 2002: An integrated three-dimensional objective analysis scheme in use at the Storm Prediction Center. Preprints, *21st Conf. on Severe Local Storms/19th Conf. on Weather Analysis and Forecasting/15th Conf. on Numerical Weather Prediction*, San Antonio, TX, Amer. Meteor. Soc., JP3.1, <https://ams.confex.com/ams/pdfpapers/47482.pdf>
- Epstein, E. S. (1969). A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology* (1962-1982), 8(6), 985-987. [https://doi.org/10.1175/1520-0450\(1969\)008<0985:ASSFPF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1969)008<0985:ASSFPF>2.0.CO;2).
- Jewell, R. E. (2010, October). A sounding analog system to produce probabilistic forecasts of maximum hail size. In *25th Conf. on Severe Local Storms*, https://ams.confex.com/ams/25SLSTechprogram/paper_176260.htm.
- Jirak I. L., J. Vancil, J. Picca, R. S. Schneider, P. T. Marsh, 2024: Conditional Intensity Estimation of Severe Hail and Wind Using Environmental Parameters. 31st Conf. on Severe Local Storms, Virginia Beach, VA, Amer. Meteor. Soc., 17A.2.
- Johnson, A. W., & Sugden, K. E. (2014). Evaluation of sounding-derived thermodynamic and wind-related parameters associated with large hail events. *E-Journal of Severe Storms Meteorology*, 9(5), 1-42. <https://doi.org/10.55599/ejssm.v9i5.57>.

Nixon, C. J., Allen, J. T., & Taszarek, M. (2023). Hodographs and skew T s of hail-producing storms. *Weather and Forecasting*, 38(11), 2217-2236. <https://journals.ametsoc.org/view/journals/wefo/38/11/WAF-D-23-0031.1>.

Nixon, C. J., Jirak, I. L., and Wendt, N. A. (2026): A Track-Based Hail Dataset. 32st Conf. on Severe Local Storms, Madison, WI, Amer. Meteor. Soc., P224.

Nowotarski, C. J., & Jones, E. A. (2018). Multivariate self-organizing map approach to classifying supercell tornado environments using near-storm, low-level wind and thermodynamic profiles. *Weather and Forecasting*, 33(3), 661-670. <https://doi.org/10.1175/WAF-D-17-0189.1>.

Schaefer, J. T., and R. Edwards, 1999: The SPC tornado/severe thunderstorm database. Preprints, 11th Conf. on Applied Climatology, Dallas, TX, Amer. Meteor. Soc., 6.11.

Taszarek, M., Púčik, T., Allen, J. T., Nixon, C., Dowdy, A., Bechis, H., ... & Brooks, H. E. (2026). What do large hail, tornado, and severe thunderstorm wind environments have in common across continents?. *Bulletin of the American Meteorological Society*, BAMS-D. <https://doi.org/10.1175/BAMS-D-25-0268.1>.

Warren, R. A., Richter, H., & Thompson, R. L. (2021). Spectrum of near-storm environments for significant severe right-moving supercells in the contiguous United States. *Monthly Weather Review*, 149(10), 3299-3323. <https://doi.org/10.1175/MWR-D-21-0006.1>

* Corresponding author address: Dr. Cameron J. Nixon, NOAA/NWS/NCEP/Storm Prediction Center, 120 David L. Boren Blvd., Norman, OK 73072; e-mail: Cameron.Nixon@noaa.gov