egusphere-2026-536

Seasonal prediction of springtime tornado activity in the United States using a hybrid model

Graber et al.

Recommendation: major revisions

General Comments

The purpose of this study was to evaluate spring prediction skill of U.S. tornado activity by using April-initialized forecasts of weather regime (WR) frequency to model U.S. tornado and tornado outbreak activity in April-May. Prediction skill is explained and interpreted by stratifying forecast skill by modes of climate variability, linking modes of climate variability to favorability of tornado-related environments, and linking WR frequency to the modes of favorability. Finally, the authors examine SST patterns for WR phases. In general, seasonal WR information was most skillful for tornado outbreak predictions, and, when stratifying by climate mode, was more skillful during AO+, NAO+, PNA+, and El Niño years. The WRs associated with greater tornado outbreak activity was associated with respective climate mode that was linked to favorable environments, and vice versa. This is valuable research for the community. However, the authors need to address several concerns before publication. I outline them below.

Major comments

1. The delineation between what information comes from WRs and what information comes from the AO/NAO/PNA climate modes is unclear and arbitrary. I would like the authors to give more thought into how to present the "pathway" or route from more low-frequency modes (ENSO, SSTs) to daily synoptic setup for a tornado (outbreak) day.
    a. The authors aggregate AO/NAO/PNA over April-May to define a low-frequency climate mode. However, AO/NAO/PNA have daily indices too that, similar to WR, explain more synoptic-scale atmospheric variability. Therefore, are the WRs even necessary, i.e., do they give redundant information that forecasts of daily index of AO/NAO/PNA would provide?
    b. If the WRs are providing the synoptic-scale information, the authors should composite the CAPE and shear environments for the different WRs (rather than composite over seasonal AO/NAO/PNA phase) to compare to Figure S2 anomalous tornado (outbreak) days. Results like L230-232 (low CAPE + high VWS = active TO) is not clear/confusing and I think it might be from the way these are composited.
    c. ENSO modulates PNA too. How to detangle? This should at least be discussed in Summary section.

d. Figure 7 shows the (potential) modulation of seasonal WR via SSTs. Would it be more appropriate/accurate that the pathway is that the SSTs are modulating the seasonal AO/NAO/PNA which are influencing the daily WR frequency? What would Figure 7 look like for AO/NAO/PNA phases?

2. The most skillful forecasts correspond to the phases that are unfavorable to tornado outbreaks (El Niño and PNA+), which suggests that the hybrid model performs better during inactive outbreak years, contrary to studies like Lepore et al. (2017). Could this be an artifact of using PC = correct predictions/total predictions? Or could it be that the variance is lower during inactive years, so greater signal-to-noise ratios? The authors should justify their skill metric or discuss this result more.

Lepore, C., M. K.Tippett, and J. T.Allen (2017), ENSO-based probabilistic forecasts of March–May U.S. tornado and hail activity, Geophys. Res. Lett., 44, 9093–9101, doi:10.1002/2017GL074781.

3. Is all the skill coming from April WRs? I would be curious if skill is very low/negative for May WRs (initialized in April).

4. ENSO can still experience a transition in phase during April. Does it make sense to use a threshold for EN3.4 that is not fixed, e.g. 25th and 75th percentiles based on month?

Minor comments

The overuse of acronyms hurts its readability. TO vs. TD vs. TI specifically gets confusing.

L85-95: Are the anomalies detrended for k-means as in Graber et al. (2025) and Tippett et al. (2024) WR studies? I assume so based on L90 but should be stated anyway.

L221-222: Enhanced outbreak and AO- link different from other studies (Tippett et al. 2022), and the discussion of Fig 5b results does not match what is shown in Figure S2 in terms of environments.

L287: “WR-E occurs more frequently in the positive phase of ENSO…” I also note that WR-E is more frequent during La Niña (stat. sig. too) and should be mentioned.

L397: Add “statistically” in front of “insignificant” and consider changing instances of “patchy” to “incoherent”/”inconsistent”

Figure 1: Fix panels f and h legend.

Figure 7: Perhaps show sample sizes on panels.