

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

Graber et al.

Recommendation: major revisions

General Comments

Thank you to the authors for responding to my concerns. See remaining comments / response below.

Major comments

1. For readability and brevity, authors should consider focusing on TO prediction results for remainder of paper's main text after Figure 3. A statement near the end of section on L203-211 about why they focus on TO prediction for remainder of paper is justified by TD lower prediction skill and weaker connection to WRs.

Related- Figures 5 and S3: The TD probability anoms are weak/noisy for some panels in Fig S3, consistent with the generally weaker relation between TD and climate modes in Fig 5 a,c,e,g (except maybe NAO?). The authors should consider plotting TO anoms in Figure S3 to accompany discussions, e.g. L252-253.

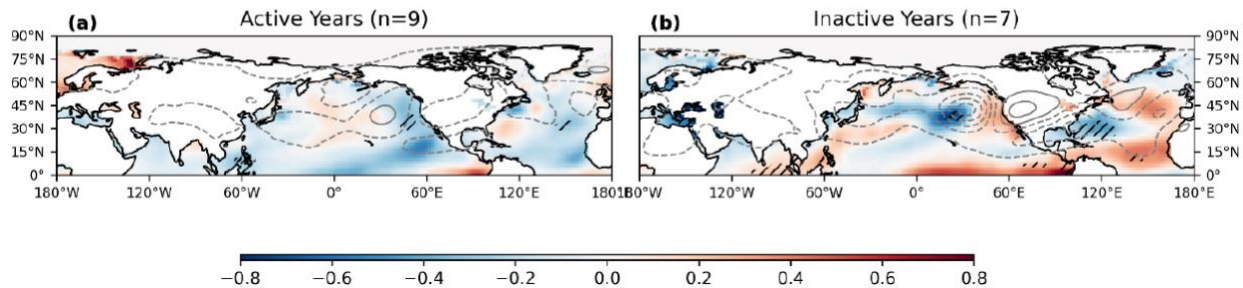
2. Figure 6 and L327-328: "However, WR-E's occurrence during the negative phase of ENSO is still less than any other WR." The authors are asking if a climate mode (like ENSO) makes a WR more or less likely, to explain if tornado outbreaks are more or less likely, so they should be comparing these fractions/occurrences to the *WR climatological frequency*. WR-E only occurs 14% of the time (Fig. 1), so it makes sense that it would almost always be the lowest percentage regardless of climate mode. In the case of ENSO as example, WR-E occurring ~20% during El Niño are robust increases in WR-E's frequency. This could partially explain predictive skill for negative years since WR-E on average relates to -50% reduction in (persistent?) outbreak days (cf. Figure 1).

In fact, for this figure and section, it might make more sense for the authors to compare the frequency/fraction of each WR given climate mode to the WR climatological frequency. E.g. Relative likelihood = $P(\text{WR} \mid \text{climate mode}) / P(\text{WR})$. This ties back to Figure 1's relative TD/TO frequencies more cleanly and the raw percentages can be misinterpreted easily.

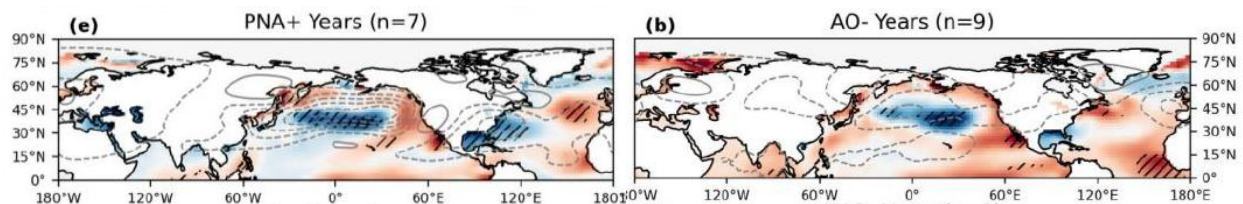
3. Figure 7 and Reviewer Response to Referee 2 #1d (and related Referee #1 comments): I still think there needs to be more refining of the potential role of SSTs as source of predictability. Its current presentation/discussion does not flow from previous sections, and most of the patterns discussed are not statistically significant so statements are inflated/not justified.

I liked Referee #1's suggestion of comparing SST/H500 anomalies of active and inactive TO years (new Figure S6, pasted below), which I think is the main signal the authors are trying to extract.

Composite SST Anomalies by Active and Inactive Tornado Outbreak Phases



Robust SST signals are present only for the inactive years, which resembles SST anomalies related to PNA+ / AO- (Figure 2 from Reviewer Response to Referee #2, pasted below). It does not particularly resemble any of the SST patterns from the active/inactive WR composites except maybe inactive WR-A.



Another suggestion is that the authors could attempt to perform pattern correlation between the SST anomaly composites for the active/inactive TO years and the SST anomaly composites for the climate modes/WRs, which would help quantify/summarize/explain similarities.

Minor comments

L62: Capitalize Northern Hemisphere

L89: Change “year-long” to “year-round”

L91: Sentence needs to be reworked for clarity as it is unclear what the authors mean by “results mainly driven by years with more Pacific Ridge WR days.”

149-151: State the tercile thresholds that dictate if tornado index in upper and lower tercile.

L179-180: WR-A resembles Pacific *Trough* from Lee et al. (2023).

L208: The authors should elaborate on what a “weak” pattern means... Non-persistent? Weak geopotential height anomalies? Patterns that do not cleanly project onto the calculated WRs?

L287: Threat Score should briefly be explained in Section 2.3. In addition, was the Threat Score only evaluated for NAO+ and PNA+ years? It is unclear why only one number for upper tercile and one number for lower tercile is reported though PC was reported for each climate mode phase.

Fig. 1: Panels g and i have the p-value in title displayed as one decimal point (0.0 instead of 0.0x). Can the authors check the string formatting, such that a p-value like 0.01, 0.02, ..., isn't displaying as 0.0? It looks correct in later figures.

The authors should consider including these missed references on climate modes-tornado/outbreaks over last few years:

Hoogewind, K. A., Galarneau Jr, T. J., & Gensini, V. A. (2025). Severe Convective Weather Outbreaks on 10 and 15 December 2021: Large-Scale Antecedent Conditions. *Monthly Weather Review*, 153(7), 1171-1194.

Hoogewind, K. A., Gensini, V. A., & Brooks, H. E. (2025). On the relationship between monthly mean surface temperature and tornado days in the United States. *npj Climate and Atmospheric Science*, 8(1), 119.

Kim, D., Lee, S. K., Lopez, H., Jeong, J. H., & Hong, J. S. (2024). An unusually prolonged Pacific-North American pattern promoted the 2021 winter quad-state tornado outbreaks. *npj Climate and Atmospheric Science*, 7(1), 133.

Malloy, K., & Tippet, M. K. (2024). A stochastic statistical model for US outbreak-level tornado occurrence based on the large-scale environment. *Monthly Weather Review*, 152(5), 1141-1161.

Malloy, K., & Tippet, M. K. (2025). Forecasting US Tornado Outbreak Activity and Associated Environments in the Global Ensemble Forecast System (GEFS). *Weather and Forecasting*, 40(4), 593-608.