52 PERFORMANCE OF OPERATIONAL CONVECTION-ALLOWING MODELS BY REGION AND SEASON: NEAR-SURFACE STORM ENVIRONMENTS AND UPDRAFT HELICITY

Andrew R. Wade*
Cooperative Institute for Severe and High-Impact Weather Research and Operations (CIWRO), Norman, OK

Israel L. Jirak
Storm Prediction Center, Norman, OK

Anthony W. Lyza
CIWRO, Norman, OK

1. INTRODUCTION

The High Resolution Ensemble Forecast (HREF; Roberts et al. 2019) system, comprising ten solutions from five member models, is the primary convection-allowing model (CAM) guidance in operational use in the United States. HREF forecasts verify favorably compared to other CAM ensembles (Clark et al. 2021) and benefit from its model diversity (Roberts et al. 2020). Recently, the lead author's real-time subjective verification of near-surface HREF forecast fields against Real-Time Mesoscale Analysis (RTMA; De Pondeca et al. 2011) fields anecdotally suggested different behavior in southeastern U.S. cool-season severe events than in other regimes, particularly a cool bias in 2-meter temperature in severe weather environments. Quantifying any such bias may be particularly important for this region and season because of the prevalence of high-shear, low-CAPE severe weather (e.g., Sherburn et al. 2014) and accompanying convective forecast sensitivity to minor thermodynamic errors. We use two distinct methods of identifying potential inflow environments in observations/RTMA to establish biases in 2-meter temperature and dewpoint fields, and compare results in the Southeast cool season to those in the Great Plains warm season.

Corresponding author: Andrew R. Wade, andrew.wade@noaa.gov

2. DATA AND METHODS

2.1 *Tornado-centered errors vs. RTMA*

In the first method, tornado reports are spatiotemporally filtered so that no remaining reports fall within 250 km and 3 h of each other, increasing independence of the mesoscale environments sampled. From HREF member forecasts valid 0–1 h before the report time, and initialized 6–18 h prior to that time at either 00 or 12 UTC, 2-meter temperature and dewpoint fields are extracted in 400 x 400-km regions centered on each tornado report. Then, RTMA fields valid at the same time are regridded to the coarser HREF grid. Resulting difference fields for each member model are averaged across cases for the Southeast cool season (15 October–15 March, 2019–2022) and for the Plains warm season (15 March–15 October, 2019–2021), with geographic domains defined in Fig. 1. This method has the benefit of directly capturing the environments of most importance to these CAMs' operational utility.

2.2 *Warm sector errors vs. ASOS*

The former method is limited by the sample size available for the two new members in version 3 of the HREF, the High-Resolution Rapid Refresh (HRRR; Smith et al. 2008) version 4 and a convection-allowing version of the FV3 model. Furthermore, in the Southeast cool season, the tornado-centered method may be biased toward

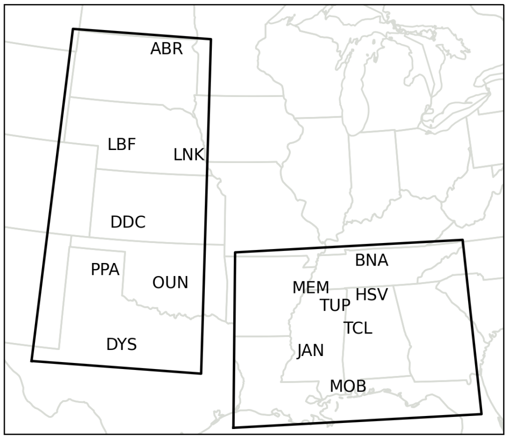

Fig. 1. Great Plains and Southeast domains (boxes) of tornado reports used for the method detailed in section 2.1, and the locations of ASOS sites (text) used for the method detailed in section 2.2.

errors of a particular sign. Warm sectors that are cooler than forecast may not produce tornadoes at all in a CAPE-limited winter regime, escaping the sample. Both of these limitations may be avoided by verifying forecasts against Automated Surface Observing Systems (ASOS) observations across warm sectors, regardless of whether severe storms are nearby. Seven sites spread across each region are used for verification: KBNA, KHSV, KJAN, KMEM, KMOB, KTCL, and KTUP for the Southeast, and KABR, KDDC, KDYS, KLBF, KLNK, KOUN, and KPPA for the Plains (Fig. 1). Observations at 00, 06, 12, 18, and 21 UTC over the seasons defined in the previous section are retrieved and ad hoc minimal "warm sector" criteria are applied. For the Southeast cool season, temperature must be at least 16° C, dewpoint at least 13° C, and the southerly component of 10-m wind at least 3 m s^{-1}. For the Plains warm season, temperature must be at least 21° C, dewpoint at least 16° C, and the southerly component of 10-m wind at least 3 m s^{-1}. Observations meeting these criteria are then compared to 6–18-hour HREF forecasts valid at the nearest gridpoint at the same time.

3. RESULTS

3.1 *Tornado-centered errors vs. RTMA*

In the tornado-centered mean error fields, the region of interest is not the location of the tornado itself but the broad inflow sector to the east and south. Here, all HREF member models have a cold bias (Fig. 2) as hypothesized. This bias is on the order of 1 K across members. Two members, the North American Mesoscale Model (NAM) convection-allowing nest and the National Severe Storms Laboratory's (NSSL) version of the Weather Research and Forecasting (WRF; Skamarock et al. 2008) model, are colder than the others, likely because of undermixing in the Mellor-Yamada-Janjić (MYJ; Janjić 1994) planetary boundary layer (PBL) parameterization used by those members. However, mixing is not the only source of these errors, and perhaps not even the primary one; the cold bias appears in members using other PBL schemes (including a nonlocal scheme in the WRF-ARW) and the near-surface moisture field is essentially unbiased in the inflow sector for most members (Fig. 3).

In contrast to the nearly identical biases present in all members in the Southeast cool season, the Plains warm season produces a variety of distinct model behaviors in both temperature (Fig. 4) and dewpoint (Fig. 5) fields. While some of the local biases, particularly in moisture forecasts, are larger, they vary considerably across members and do not all have the same sign. This wider range of behaviors is more suitable for a mixed-physics ensemble. Furthermore, inflow-sector errors in most members appear spatially heterogeneous and likely are influenced in places by large individual errors, such as those resulting from a modeled or observed convective cold pool or a misplaced mesoscale boundary, unlike the broadly homogeneous cold errors in the Southeast cool-season inflow sectors.

3.2 *Warm sector errors vs. ASOS*

Though targeting a much broader set of environments and verifying with point observations instead of an objective analysis, the ASOS-based method produces a result very

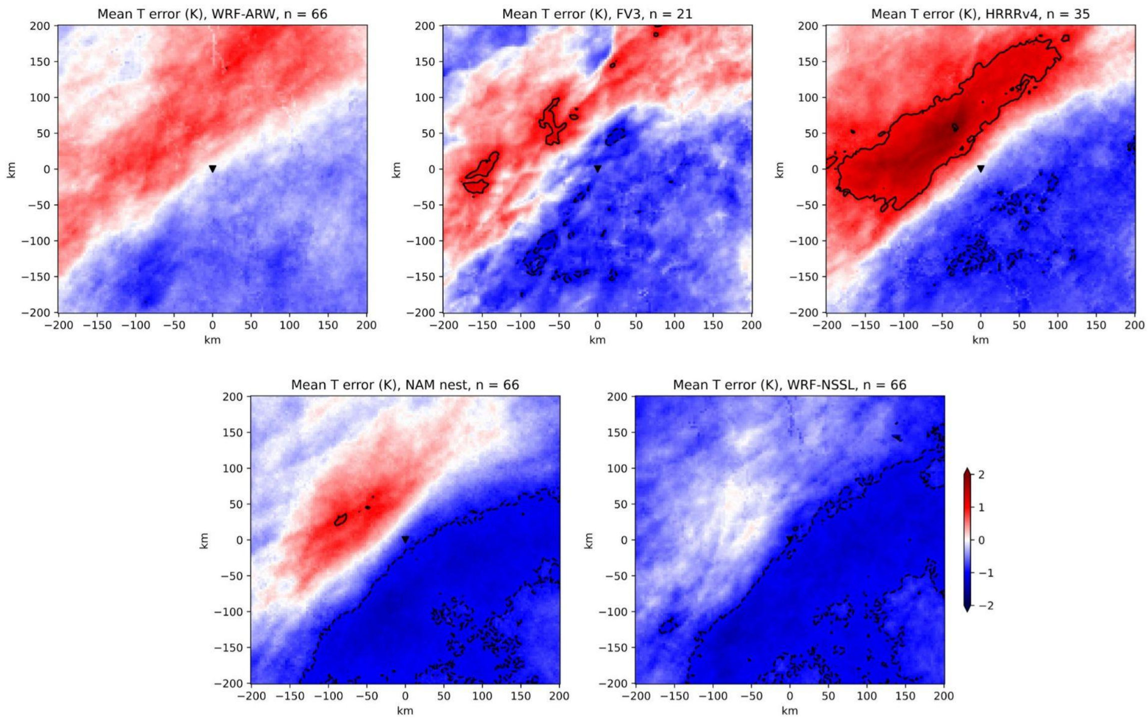

Fig. 2. Mean 2-m temperature errors (K) centered on Southeast cool-season tornado reports (black triangle at center) for each HREF member, across rows from top left: WRF-ARW, FV3, HRRRv4, NAM nest, and WRF-NSSL. Black contours are intervals of 1 K (dashed if negative) excluding 0 K.

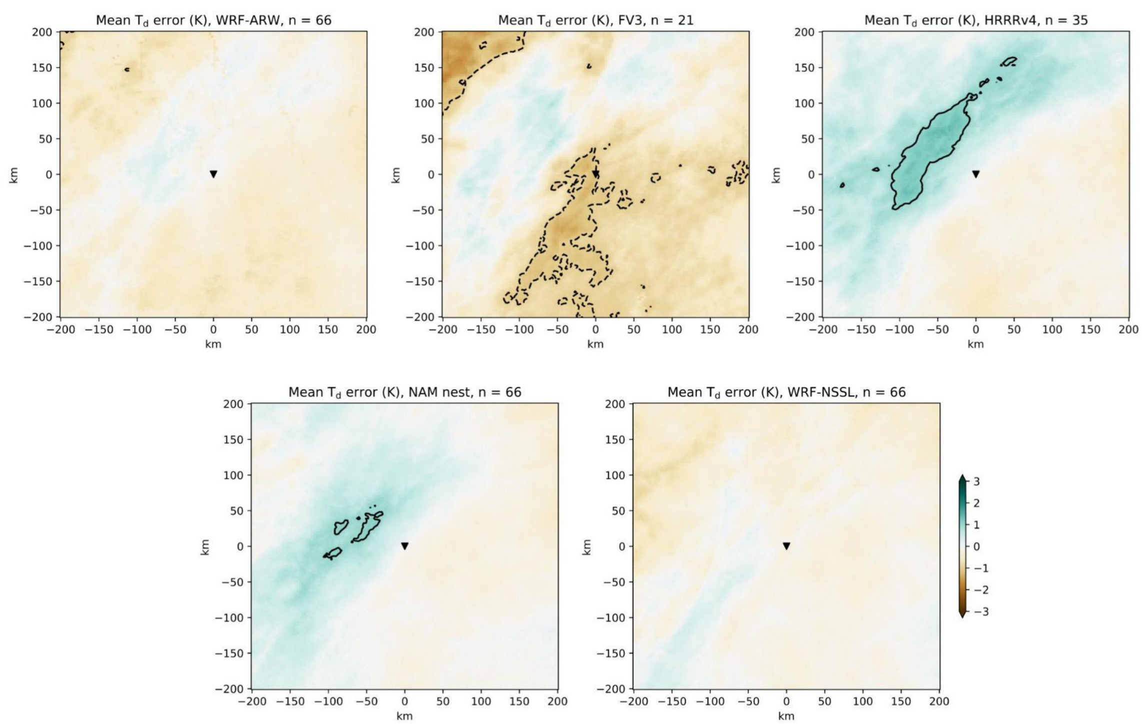

Fig. 3. As in Fig. 2, but for 2-m dewpoint errors (K).

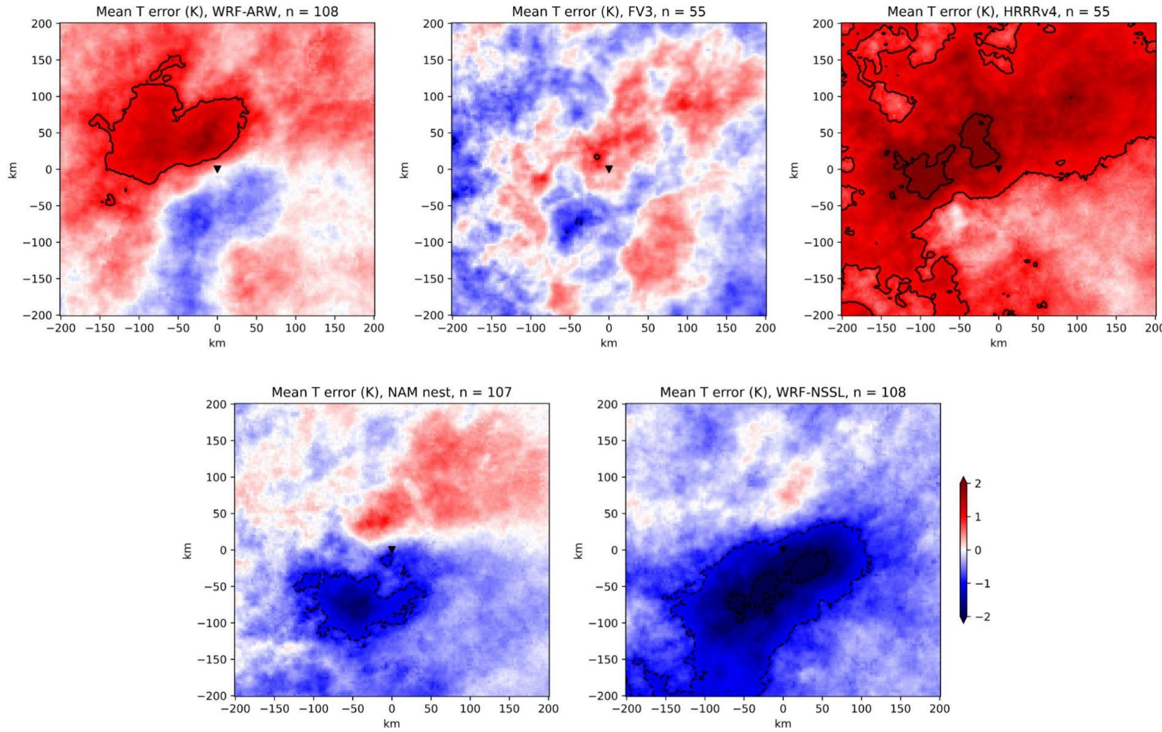

Fig. 4. As in Fig. 2, but for the Great Plains warm season.

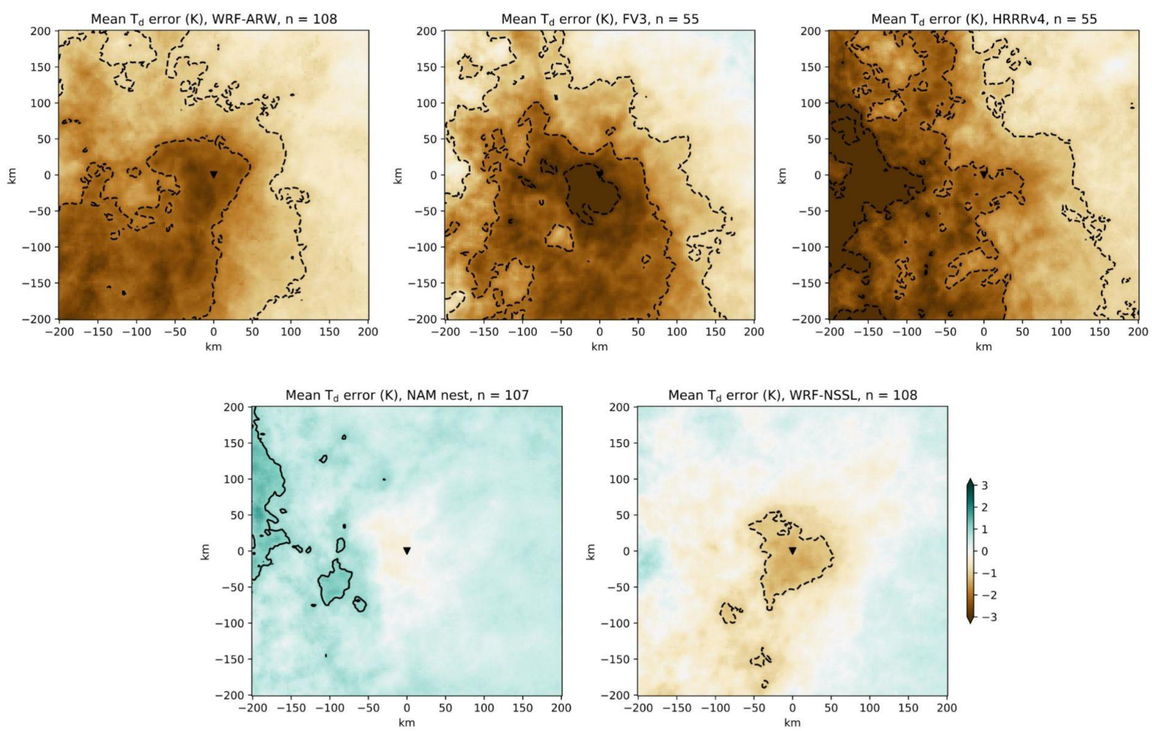

Fig. 5. As in Fig. 3, but for the Great Plains warm season.

similar to the tornado-centered method. For Southeast cool-season warm sectors, a cold bias on the order of 1 K exists across members (Fig. 6), with the MYJ members somewhat colder. Dewpoint errors (Fig. 7) deviate slightly more from the tornado-centered results with moist biases evident in the MYJ members. The larger sample allows stratification by time of day. 12 UTC errors (Fig. 8) are nearly uniform across members, but 21 UTC errors (Fig. 9) reveal a sizable gap between MYJ members, which have a ~2 K cold bias near peak heating, and the other members. However, the non-MYJ members retain a cold bias even at 21 UTC. Errors may also be stratified by ASOS sites' reported low-level sky cover categories, with no meaningful differences (not shown here); the cold bias persists from clear skies or few clouds all the way to overcast conditions. Meanwhile, results from the Plains warm season indicate minimal temperature biases (Fig. 10), especially across members. Dewpoint errors (Fig. 11) are consistent with the tornado-centered framework and show a neutral-to-dry bias for all members. The Plains warm-season results are probably less specific to convective environments than the Southeast cool-season results, since the warm sector criteria used here are near climatology for much of the Plains during the warm season.

4. CONCLUSIONS AND ONGOING WORK

4.1 *Cold bias in Southeast cool-season warm sectors*

Two distinct and independent approaches agree that all HREF members suffer from a near-surface cold bias in Southeast cool-season severe weather environments. While many forecasters responsible for the region are already aware that this bias exists anecdotally, confirmation of mean/median errors of roughly –1 K (or as large as –2 K if using the MYJ scheme during diurnal heating) provides a quantitative baseline for evaluating model guidance in these environments. For further context, an idealized low-CAPE thermodynamic profile emulating that of Sherburn and Parker (2019) and containing 470 J kg^{-1} surface-based CAPE yields only 326 J kg^{-1} if the parcel temperature is reduced by 1 K, so that a cold error of 1 K in such a scenario would reduce the forecast CAPE by 30 percent of the amount observed.

The reason for this bias is still unknown. The HREF membership comprises multiple model cores and initial conditions, four PBL schemes, and four microphysics schemes, so none of these are adequate explanations. The HREF archive does not contain enough vertical levels to assess characteristics of modeled cloud layers compared to upper-air observations; however, any systematic error in cloud thickness should lead to a diurnal cycle of 2-meter temperature bias, not a persistent cold bias at both 12 and 21 UTC.

4.2 *Updraft helicity layers and thresholds*

Ongoing work focuses on the skill of the available updraft helicity (UH) integration layers, 0–3 and 2–5 km, and the variability of percentile thresholds of UH by region/season. Traditional UH-based guidance may not perform as well in high-shear, low-CAPE environments as in higher-CAPE environments (Sobash 2018) and lower thresholds may be required (Graham and Lackmann 2020). Preliminarily, over 27 severe weather days in the Southeast cool season, the 0–3-km integration layer has a higher median fractions skill score than the 2–5-km layer for forecasting all severe hazards. In these events, skill appears to be optimized at thresholds of 24-hour maximum UH around 30–40 m^2 s^{-2} for the 0–3-km layer and 50–60 m^2 s^{-2} for the 2–5-km layer (except in the FV3 core, which systematically produces much higher values of UH in both layers). This is around the 99.95th percentile for each layer across the Southeast cool-season events, but the climatological distributions vary considerably by season. Future work will treat tornadoes separately once an adequate sample of days is available, and may attempt to relate UH values more directly to environmental CAPE rather than inferring CAPE from the region and season.

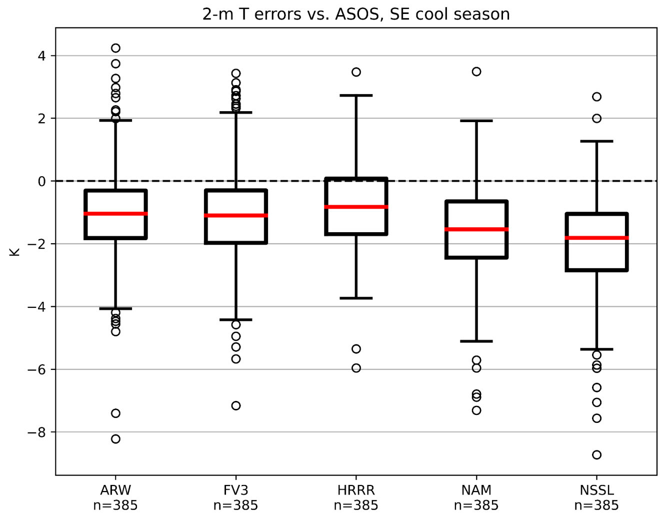

Fig. 6. Boxplots of 2-meter temperature errors (K) for each HREF member, compared to ASOS observations meeting warm sector criteria in the cool season at 7 selected sites in the Southeast.

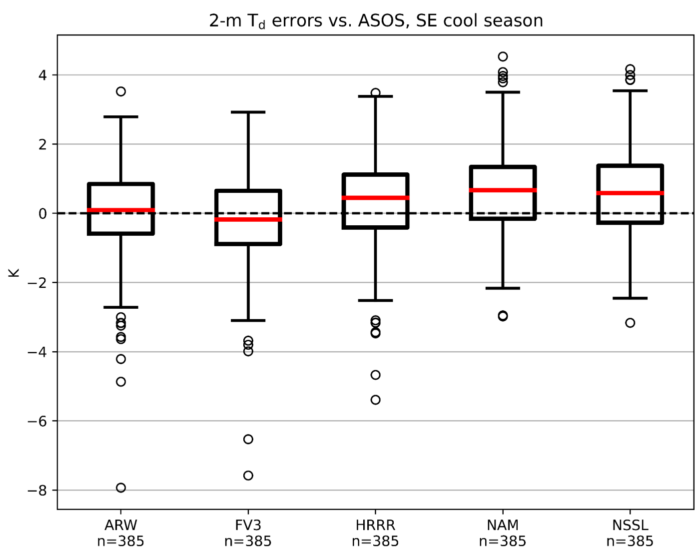

Fig. 7. As in Fig. 6, but for 2-meter dewpoint errors (K).

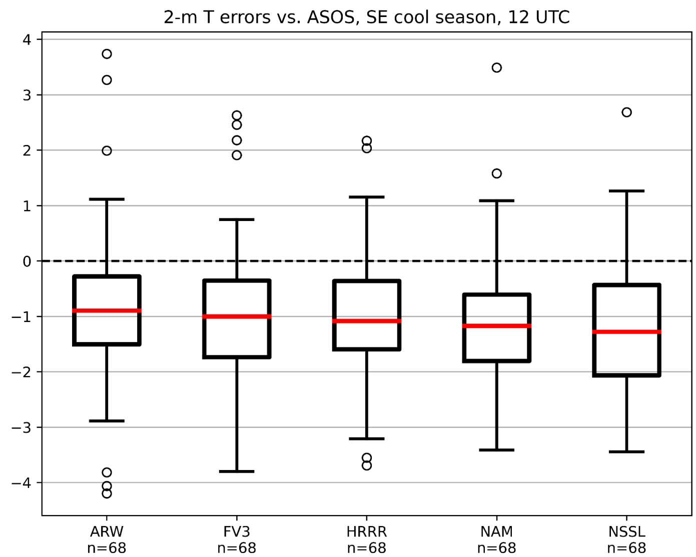

Fig. 8. As in Fig. 6, but for 12 UTC observations only.

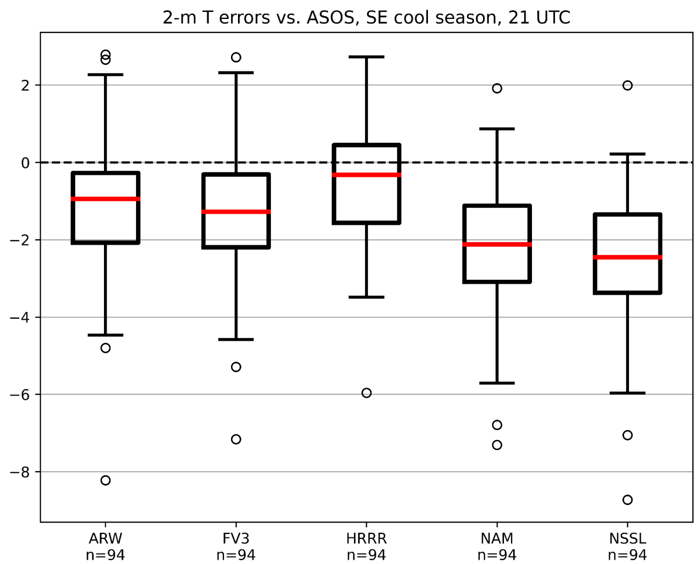

Fig. 9. As in Fig. 6, but for 21 UTC observations only.

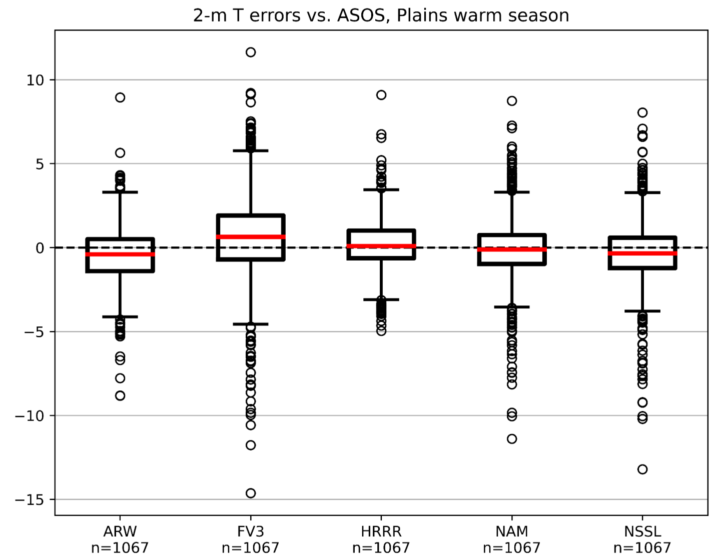

Fig. 10. As in Fig. 6, but for the Great Plains warm season.

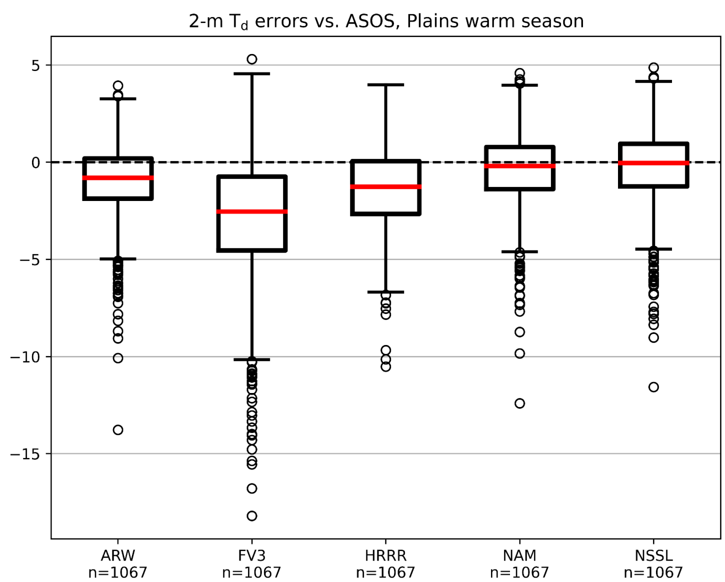

Fig. 11. As in Fig. 7, but for the Great Plains warm season.

REFERENCES

Clark, A. J., and Coauthors, 2021: Spring Forecasting Experiment 2021: Preliminary findings and results . NOAA Hazardous Weather Testbed, NOAA, 86 pp., https://hwt.nssl.noaa.gov/sfe/2021/docs/HWT_SFE_2021_Prelim_Findings_FINAL.pdf.

De Pondeca, M. S. F. V., and Coauthors, 2011: The real-time mesoscale analysis at NOAA's National Centers for Environmental Prediction: Current status and development. *Wea. Forecasting*, **26**, 593–612, https://doi.org/10.1175/WAF-D-10-05037.1.

Graham, C. S., and G. M. Lackmann, 2020: Verification of convection-allowing NWP in high-shear, low-CAPE environments. *30th Conf. on Weather Analysis and Forecasting*, Boston, MA, Amer. Meteor. Soc., 1B.4.

Janjić, Z. I., 1994: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon. Wea. Rev., 122, 927–945.

Roberts, B., I. L. Jirak, A. J. Clark, S. J. Weiss, and J. S. Kain, 2019: Postprocessing and visualization techniques for convection-allowing ensembles. *Bull. Amer. Meteor. Soc.*, **100**, 1245–1258, https://doi.org/10.1175/BAMS-D-18-0041.1.

Roberts, B., B. T. Gallo, I. L. Jirak, A. J. Clark, D. C. Dowell, X. Wang, and Y. Wang, 2020: What does a convection-allowing ensemble of opportunity buy us in forecasting thunderstorms? *Wea. Forecasting*, **35**, 2293–2316, https://doi.org/10.1175/WAF-D-20-0069.1.

Sherburn, K. D., and M. D. Parker, 2014: Climatology and ingredients of significant severe convection in high shear, low-CAPE environments. *Wea. Forecasting*, **29**, 854–877, https://doi.org/10.1175/WAF-D-13-00041.1.

Sherburn, K. D., and M. D. Parker, 2019: The development of severe vortices within simulated high-shear, low-CAPE convection. *Mon. Wea. Rev.*, **147**, 2189–2216, https://doi.org/10.1175/MWR-D-18-0246.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3 . NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Smith, T. L., S. G. Benjamin, J. M. Brown, S. Weygandt, T. Smirnova, and B. Schwartz, 2008: Convection forecasts from the hourly updated, 3-km High Resolution Rapid Refresh (HRRR) model. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 11.1., https://ams.confex.com/ams/24SLS/techprogram/paper_142055.htm.

Sobash, R. A., 2018: Severe weather forecast skill across different environments in an experimental convection-allowing model. *29th Conf. Severe Local Storms*, Stowe, VT, Amer. Meteor. Soc., 125.