17A.1

UPGRADES TO THE STORM PREDICTION CENTER MESOANALYSIS: A LOW-LATENCY, 3-KM HRRR-BASED OBJECTIVE MESOSCALE ANALYSIS

Kelton T. Halbert, Nathan Dahl, Israel Jirak, Patrick Marsh

NOAA/NWS/NCEP Storm Prediction Center
The University of Oklahoma - Cooperative Institute for Severe and High Impact Weather Research and Operations

1. INTRODUCTION

The NOAA/NWS/NCEP Storm Prediction Center (SPC) produces a 40-km hourly surface objective mesoanalysis, first described by Bothwell et al. (2002). This hourly mesoanalysis uses Rapid Refresh (RAP; formerly RUC) 1-hour forecasts as a first guess, and then performs a 2-pass Barnes analysis (Barnes 1964) with available surface observations. This surface objective analysis is then merged with the RAP upper-air fields, and post processing of derived convective and sounding-based indices is performed by NSHARP (Hart and Korotky 1991). These hourly mesoanalysis fields are meant to be as low-latency as possible, typically available at 15-minutes past the top of the hour, and then made available on the SPC website: https://www.spc.noaa.gov/exper/mesoanalysis

The SPC Mesoanalysis (aka surface objective analysis or SFCOA) is the foundation of many operational tools at SPC, and has been a core dataset in many SPC publications. The output of SFCOA is used in real-time operations for Mesoscale Convective Discussion (MCD) issuance and public-facing graphics, as well as providing environmental information for issuing Severe Thunderstorm and Tornado Watches. The SPC Mesoanalysis has also provided near-storm environment data for deriving real-time tornado intensity estimates (Smith et al. 2020) and studies of convective mode (Thompson et al. 2012), among many other projects.

With the RAP and High Resolution Rapid Refresh (HRRR) forecast models slated for retirement, SPC has begun investigating how to upgrade and replace the legacy 40km analysis system, as well as expansion into territories outside of the continental US (OCONUS; Hawaii and Alaska). While both the RAP and HRRR will be retired, the HRRR will be replaced with another regional Convection-Allowing Model (CAM), but there is no mesoscale model equivalent replacement for the RAP. The 3D Realtime Mesoscale Analysis (3D-RTMA) produced experimentally by the NCEP Environmental Modelling Center (EMC) has been evaluated in the Hazardous Weather Testbed (HWT) for a number of years as a potential replacement. However, due to the computational cost of full 3D data assimilation, analyses are often not available until nearly the next analysis hour, negatively impacting forecast operations at SPC. Therefore, the HRRR serves as the basis for the current upgrade path for an updated SFCOA, which will be replaced eventually by the Rapid Refresh Forecast System Version 2 (RRFSV2) implementation.

An initial implementation of SFCOA-HRRR was built for the 2024 HWT Spring Forecasting Experiment (SFE), with hourly analyses being generated between 27 April 2024 and 07 June 2024. Implementation details and preliminary results are discussed in the following sections.

2. PRELIMINARY SFCOA-HRRR IMPLEMENTATION

The initial overall structure of the new SFCOA implementation is very similar to the original implementation from Bothwell et al. (2002):

1. The HRRR 1-hour forecast is used as a first-guess analysis for surface fields, and held static for upper-air fields.
2. A 2-pass Barnes analysis is performed on available surface observations. Many more surface observations are now included from state mesonets, RAWS, transportation networks, river authorities, etc., facilitated through the SynopticData API.
3. Rather than using 25-mb pressure level data for upper-air fields, SFCOA-HRRR is

** Corresponding Author Address:*
Kelton T. Halbert, NOAA/NWS/NCEP Storm Prediction Center, Norman, OK 73072;
email: kelton.halbert@noaa.gov

performed on the full, native vertical level data from the HRRR.

4. Post-processing of derived and sounding-based indices is done using SHARPlib, SPC's unified sounding post-processing library that is optimized for high-fidelity data. SHARPlib is available on GitHub: https://github.com/keltonhalbert/SHARPlib
5. When the HRRR 0-hour analysis becomes available, the mesoanalysis is recomputed using the analysis rather than the 1-hour forecast.

Through efficient use of parallel-friendly data structures and algorithms, the new SFCOA-HRRR is available as early as 5-10 minutes past the top of the hour, making it slightly faster than its predecessor despite being a higher-resolution analysis.

2.1 Defining an Appropriate "Radius of Influence" for Observations

Eq. 1: A formulation of the interpolation weights used in Barnes 1964.

The Barnes analysis technique assumes that data and grid spacing are approximately equal. Of particular importance with this assumption is the "Radius of Influence" (RoI) used when computing the observation weights within the Barnes analysis. This RoI term (**K** in Eq. 1) is effectively a weight falloff parameter, with observations outside of that radius receiving increasingly smaller weights. Regrettably, surface observing networks in the US are not at a near-constant 3-km spacing to match the HRRR native resolution. A naive remedy to this is to use a gridpoint-dependent RoI based on the surface observing network density local to that gridpoint. For more dense observing networks, the RoI will be smaller and distant observations will have less impact, while in data sparse areas, observations at greater distances are included. The initial implementation of the variable RoI used for the HWT SFE was the median distance from the gridpoint to the nearest 10 observing sites, computed dynamically for each analysis hour and for each analysis variable. This allows for the Barnes weights to reflect network density changes dynamically from missing observations, observations thrown out by QC, or network outages. An example of the variable RoI is displayed in Figure 1.

3. PRELIMINARY SUBJECTIVE VERIFICATION

During the 2024 HWT SFE, three high-resolution mesoanalysis versions were subjectively evaluated by participants and forecasters: SFCOA-HRRR, 3D-RTMA HRRR, and 3D-RTMA RRFS. Additional comparisons were conducted by upscaling these analyses to a 40-km grid to match the original SFCOA-RAP at SPC. During daily evaluations, HWT participants were asked to compare the quality of SFCOA-HRRR analyses to 3D-RTMA HRRR, and for the upscaled analyses, participants were asked to rank analyses from best to worst. A ranking of 1 was considered the "best", while a ranking of 4 was considered the "worst". The evaluation results between SFCOA-HRRR and 3D-RTMA HRRR are displayed in Figure 2, while the results from the evaluations of upscaled analyses are summarized in Table 1.

3.1 Subjective Evaluations of SFCOA-HRRR and 3D-RTMA HRRR

The most common subjective rating given for SFCOA-HRRR when compared to 3D-RTMA HRRR was that they were "about the same", selected by 58% of SFE participants. Participants selected "slightly better" 16% of the time, and "slightly worse" 23% of the time, with only 2% of respondents saying SFCOA-HRRR was "much worse" than 3D-RTMA HRRR. These results suggest that, while there is room for improvement with the initial implementation of SFCOA-HRRR, the simpler, low-latency approach to the objective mesoanalysis is producing qualitatively comparable results to a full 3D data assimilation system when used in real-time forecasting environments.

3.2 Upscaled Mesoanalysis Subjective Evaluations

When the high-resolution mesoscale analyses were upscaled to a 40-km grid, SFCOA-HRRR and 3D-RTMA HRRR were effectively tied for first, with average rankings of 1.93 and 1.94 for the respective analysis systems. The current SPC Mesoanalysis, SFCOA-RAP, was ranked 3rd with an average ranking of 2.86, and in last place was 3D-RTMA RRFS with an average ranking of 3.27. This result is consistent with the evaluations of the high-resolution systems, but also shows that the new SFCOA-HRRR is adding additional value over the current SFCOA-RAP system.

4. SFCOA-HRRR PRELIMINARY OBJECTIVE VERIFICATION

Preliminary objective verification statistics were computed for the hourly 2-meter temperature and dewpoint analyses generated during the HWT SFE period, separating out analyses generated using 1-hour HRRR forecasts and 0-hour HRRR analyses. ASOS/AWOS observations across the CONUS that passed QC were used for the truth observations, and displayed/aggregated two ways: 1) temporally, with hourly distributions of error for all sites across the CONUS, and 2) spatially at each observing site, taking the median error across all hourly analyses during the SFE. The objective was to highlight any systematic geographic and diurnal errors in SFCOA-HRRR. Summary RMSE and bias statistics for all sites and all hourly analyses are displayed in Table 2, separated by field and which first-guess analysis is used.

4.1 SFCOA-HRRR Temporal Verification

After computing analysis errors at each ASOS/AWOS site across the CONUS, those errors were binned by UTC hour for analyses generated using the 1-hour and 0-hour HRRR surface fields as a first guess. The hourly-binned errors for temperature and dewpoint using both first guesses are displayed in Figures 3 and 4.

For 2-m air temperature analyses using the HRRR 1-hour forecast as a first-guess, errors appear normally distributed and, for the most part, symmetric around the zero-error line. The interquartile range for the temperature errors are largely within +/- 1 ºF (+/- 0.56 ºC), and a very slight diurnal signal is present in the median errors. Median errors are minimized between 03-10 UTC, and again between 21-23 UTC. There appears to be a slight cold bias in the analyses from 11-20 UTC, and a slight warm bias between 00-02 UTC. The interquartile range is narrowest between 11-14 UTC, and largest from 18-04 UTC. When looking at mesoanalysis fields using the HRRR 0-hour analysis at the same valid time, the primary result is that the interquartile range and tails of the error distributions shrink, with some slight reductions in the bias of median errors. The warm bias from 00-02 UTC, as well as the cool bias from 11-16 UTC, are slightly reduced and closer to the zero-error line.

It is unknown whether these biases are from the HRRR itself, perhaps evidenced by the fact the analysis time step has a narrower interquartile range of errors, or with how observations themselves are being assimilated. Additional impacts, such as complex terrain and thunderstorm outflow could contribute to these errors, and further investigation is required. It should be noted, however, that these biases are on the order of tenths of a degree Fahrenheit, and of more importance is the interquartile range and tails of the error distributions. Even in the case of the 1-hour HRRR first guess, these errors being within +/- 1 ºF (+/- 0.56 ºC) is encouraging for a preliminary implementation of SFCOA-HRRR.

Looking at the 2-m dewpoint temperature errors binned by hour, the interquartile range of error is generally larger than the 2-m temperature analyses, and with a stronger bias present. The minimum in both bias and interquartile range errors lies between 9-12 UTC, and a general moist bias is observed at all other times. The errors within the interquartile range are mostly within +/- 1 ºF (+/- 0.56 ºC), though slightly exceeds + 1 ºF (+ 0.56 ºC) between 18-23 UTC. There is less variability between the 0-hour and 1-hour HRRR based analyses, suggesting that perhaps there is room for improvement on how moisture variables are being assimilated and objectively analyzed, rather than the errors being from the forecast-based data.

As a quality-control check, there was a restriction on the magnitude of the mixing ratio analysis increment, potentially contributing to the moist bias observed by not allowing for observations to fully modify the first guess. This hypothesis is further supported by the spatial distribution of these errors analyzed in the following section. Moisture analyses are generally known within the field to be challenging to create accurately, so while it is unsurprising that the errors are larger than the air temperature field, it is still encouraging that these moisture errors are relatively well contained within +/- 1 ºF (+/- 0.56 ºC) range, with potential for future improvement.

4.2 SFCOA-HRRR Spatial Verification

For preliminary spatial verification statistics, the temperature and dewpoint error for all hourly analyses were grouped by the station identifier, with the median error across all hourly analyses chosen as the evaluation metric. The spatial distribution of errors between the objective analyses using the 0-hour and 1-hour HRRR data as a first guess did not vary substantially, with the 1-hour HRRR first guess having slightly larger error magnitudes, reflecting the results of the temporal evaluation. For simplicity, only the errors for the 1-hour HRRR first-guess analyses are displayed in Figure 5.

The distribution of median 2-meter temperature errors across the CONUS appear largely correlated with areas of complex terrain, such as the Intermountain West and the Appalachian Mountains. Specifically, SFCOA-HRRR appears to be too warm in the presence of complex terrain, though there are some regions where the analyses are too cold. This is likely in part, if not entirely because, the Barnes technique does not traditionally account for varying terrain heights between observations. Observations within valleys, or on sloped terrain, are likely unduly influencing the gridpoints around them due to the isotropic nature of Barnes interpolation. Temperature data are assimilated by computing Dry Static Energy, which is analogous in its properties to Potential Temperature, but is height dependent instead of pressure dependent. The terrain height is used within the Dry Static Energy computation before performing the analysis, and then is reduced back to the terrain height, but this does not completely eliminate the issues posed by complex terrain.

The spatial distribution of dewpoint errors largely displays the same signal, where the bulk of the errors are concentrated in the presence of complex terrain, and particularly in the Western US. These dewpoint errors in the West reflect that the analyses generated by SFCOA-HRRR are too moist, which is possibly due to a restriction on the magnitude of the mixing ratio analysis increment allowed, as discussed in the temporal verification section. This was initially implemented as a quality-control check, but it may be preventing the surface observations from appropriately correcting the first-guess analysis. Additionally, the 1-hour first-guess is potentially too close to the model initialization time to have accurately mixed and representative boundary layers, though further work is needed to confirm this hypothesis. Portions of the Central Plains exhibit a fairly notable dry bias, which is consistent with a known afternoon dry bias in HRRR forecasts during the warm season for this region. The way mesonet observations are being assimilated could also be a contributing factor, and warrants further investigation.

5. DISCUSSION AND FUTURE WORK

The initial implementation of a 3-km scale surface objective mesoanalysis using the HRRR as a foundation showed promise as a viable replacement for the current RAP-based SPC Mesoanalysis. In subjective evaluations, SFCOA-HRRR showed improvement over the current SFCOA-RAP system, and was considered comparable to 3D-RTMA HRRR by HWT SFE participants. This is a particularly notable result since SFCOA-HRRR is a simpler and lower-latency system than the full 3D data assimilation used by 3D-RTMA. Objective verification of 2-meter temperature and dewpoint analyses at ASOS/AWOS sites highlight some areas for improvement, particularly in the presence of complex terrain and with the moisture analyses. Additional work to further develop, tune, and improve SFCOA-HRRR are expected to continue in future HWT experiments, eventually replacing SFCOA-RAP before the retirement of the RAP model.

In addition to improving the performance of SFCOA-HRRR, additional work is needed to better understand how changing from a 40km-based system using 25-mb pressure level data in the vertical, to a 3km-based system using native model level data, impacts distributions of sensitive integrated quantities such as Storm Relative Helicity (SRH) or computed storm-motion vectors. Using native vertical level data substantially improves the number of samples in the lowest levels of profiles, which in turn results in more integrated area on the hodograph and Skew-T. In order to address these questions, historical mesoanalyses for both SFCOA-HRRR and SFCOA-RAP will need to be compared to determine if certain parameters, such as the Significant Tornado Parameter, need recalibration when moving to higher spatial resolutions. In future work, archived native-level HRRR data and archived surface observations will be used to generate historical mesoanalysis data for comparison with SFCOA-RAP.

Re-processing historical mesoanalysis data also provides an opportunity to quality-control and share both historical SFCOA-RAP and SFCOA-HRRR hourly mesonanalysis data with the wider meteorological community. Ongoing efforts are actively seeking to eventually make this data publicly available through public dataset cloud providers. These efforts would include producing new variables and fields based on recent and current research in order to facilitate transitioning new products into operations, as well as providing a long-term dataset for other research topics.

6. REFERENCES

Barnes, S. L. (1964). A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology (1962-1982)*, 396-409.

Bothwell, P. D., Hart, J. A., & Thompson, R. L. (2002). An integrated three-dimensional objective analysis scheme in use at the Storm Prediction Center. In *CONFERENCE ON WEATHER ANALYSIS AND FORECASTING* (Vol. 19, pp. J117-J120). AMS.

Hart, J. A., & Korotky, W. (1991). The SHARP workstation v1. 50 users guide. *NOAA/National Weather Service*, *30*.

Smith, B. T., Thompson, R. L., Speheger, D. A., Dean, A. R., Karstens, C. D., & Anderson-Frey, A. K. (2020). WSR-88D tornado intensity estimates. Part I: Real-time probabilities of peak tornado wind speeds. *Weather and Forecasting*, *35*(6), 2479-2492.

Thompson, R. L., Smith, B. T., Grams, J. S., Dean, A. R., & Broyles, C. (2012). Convective modes for significant severe thunderstorms in the contiguous United States. Part II: Supercell and QLCS tornado environments. *Weather and forecasting*, *27*(5), 1136-1154.

7. TABLES AND FIGURES

Upscaled Analysis Average Rankings

Analysis System	Ranking (1-4)
SFCOA-HRRR	1.93
3D-RTMA HRRR	1.94
SFCOA-RAP	2.86
3D-RTMA RRFS	3.27

Table 1: The average rankings of the mesoanalysis systems evaluated in the 2024 HWT SFE when upscaled to 40km. A ranking of 1 was considered the best, and a ranking of 4 was considered the worst.

Summary Statistics for 2-m Temperature and Dewpoint Analyses

	F00 2-m Temperature	F01 2-m Temperature	F00 2-m Dewpoint	F01 2-m Dewpoint
RMSE	1.10 ºF (0.61 ºC)	1.25 ºF (0.69 ºC)	1.44 ºF (0.8 ºC)	1.52 ºF (0.84 ºC)
Bias	-0.03 ºF (-0.017 ºC)	-0.02 ºF (-0.011 ºC)	0.19 ºF (0.11 ºC)	0.18 ºF (0.1 ºC)

Table 2: Summary RMSE and bias statistics using all ASOS/AWOS sites for all generated hourly analyses, broken up by whether the analysis was created using the HRRR analysis or 1-hour forecast as a first guess.

Gridded Variable Radius of Influence for Barnes Interpolation Weights

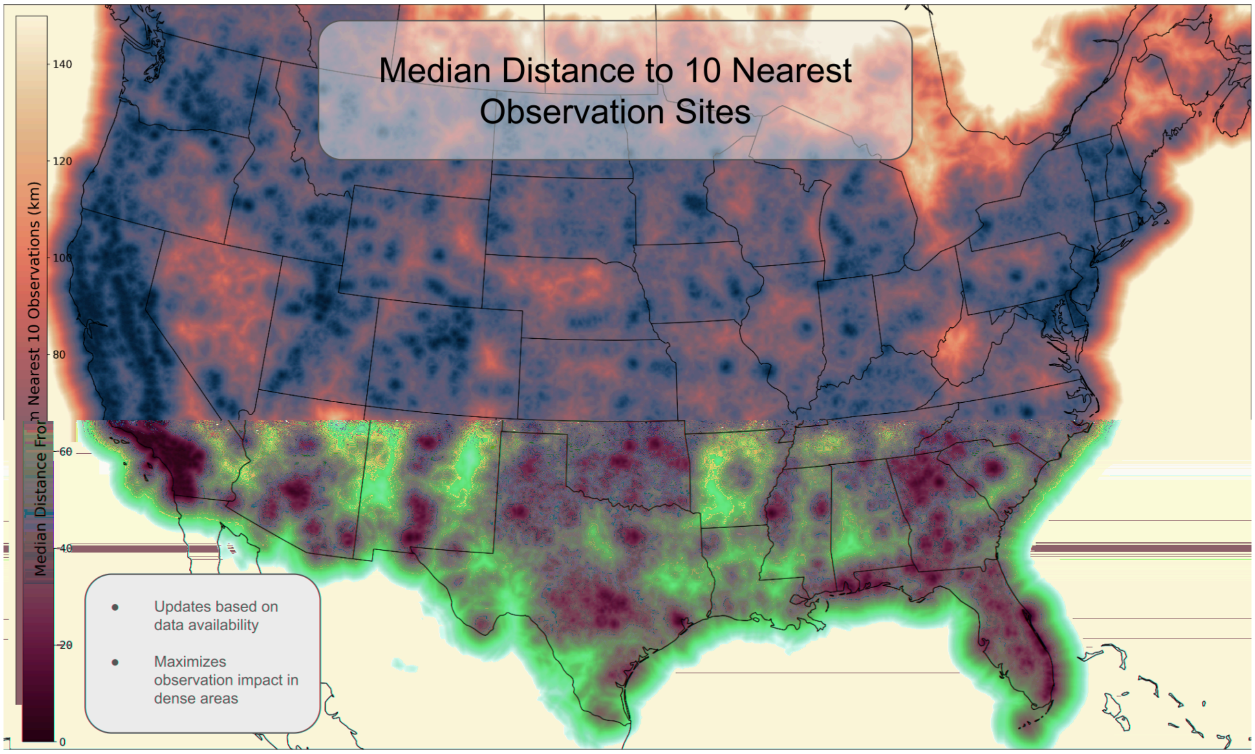

Figure 1. An example of a variable Radius of Influence used for 2-meter air temperature Barnes interpolation weights, using the median distance from each analysis gridpoint to the nearest 10 observation sites.

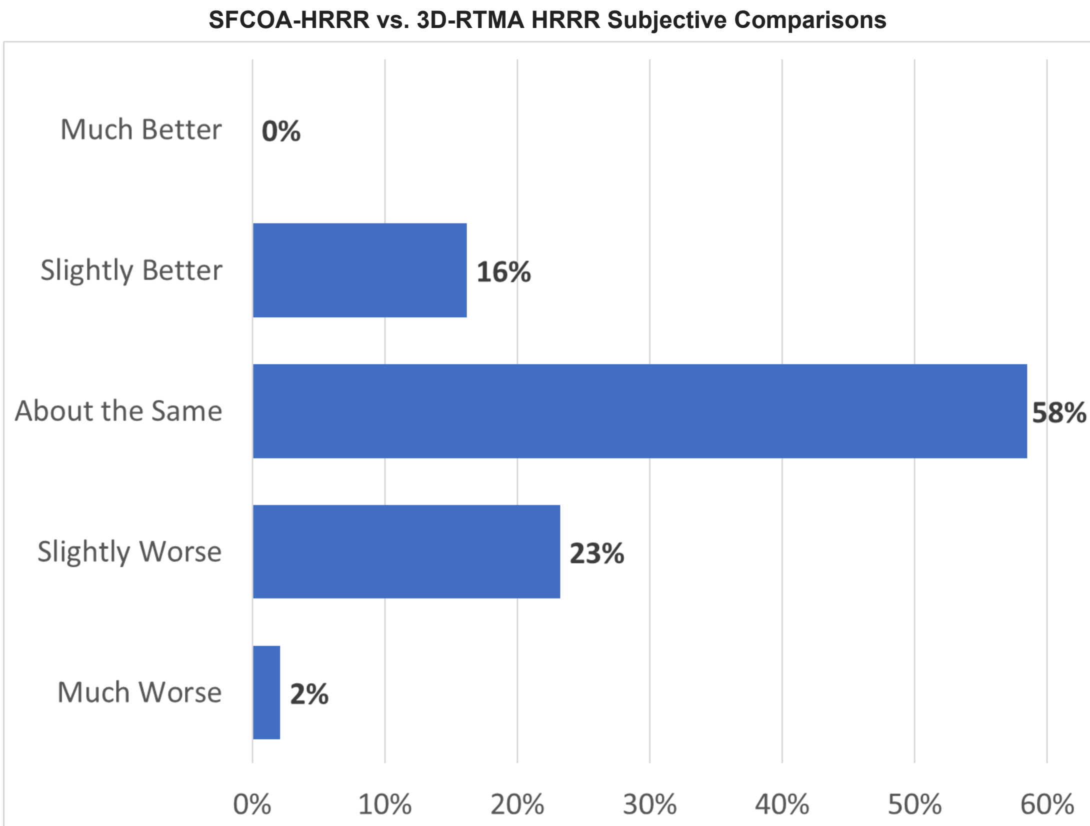

Figure 2. 2024 HWT SFE Participant Evaluations of SFCOA-HRRR when compared to 3D-RTMA HRRR.

SFCOA-HRRR Temperature Errors Binned by UTC HOUR

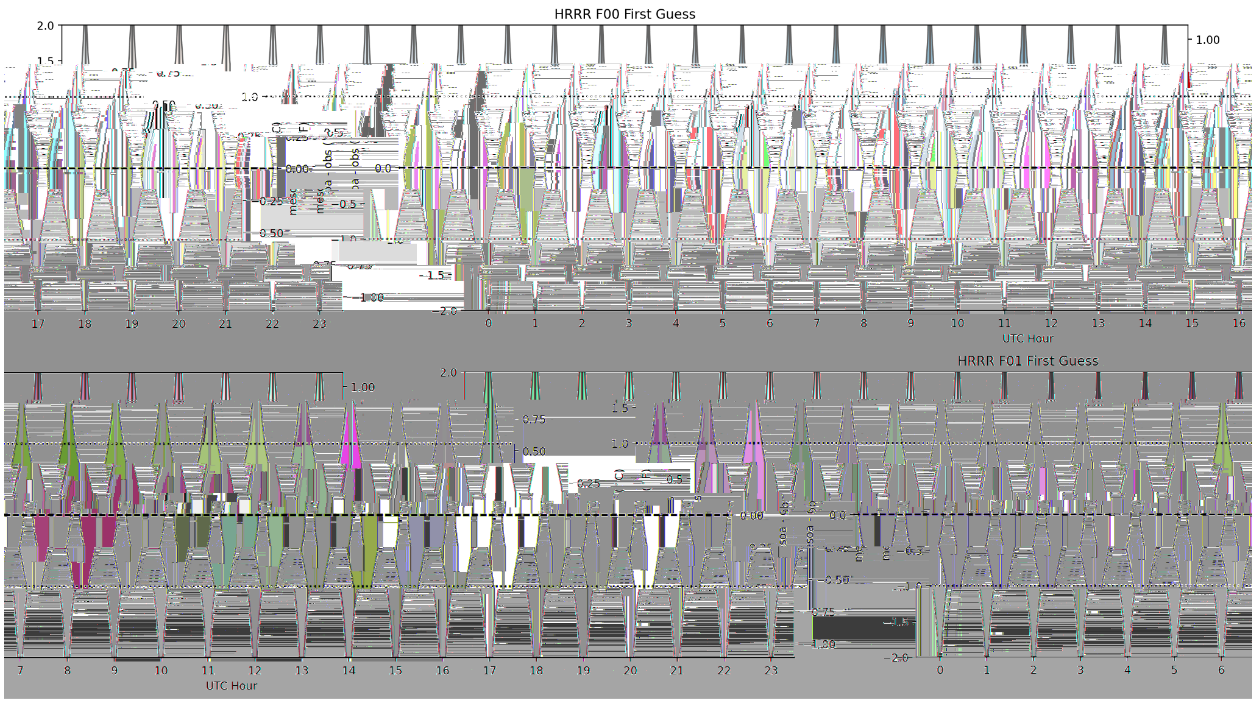

Figure 3. SFCOA-HRRR temperature errors (left axis: ºF, right axis: ºC) across the CONUS binned by UTC hour during the HWT SFE, using analysis generated from a HRRR 0-hour analysis time first-guess (**a; upper**) and a 1-hour forecast first guess (**b; lower**). Violin hues correspond to the value of the median error for the distribution.

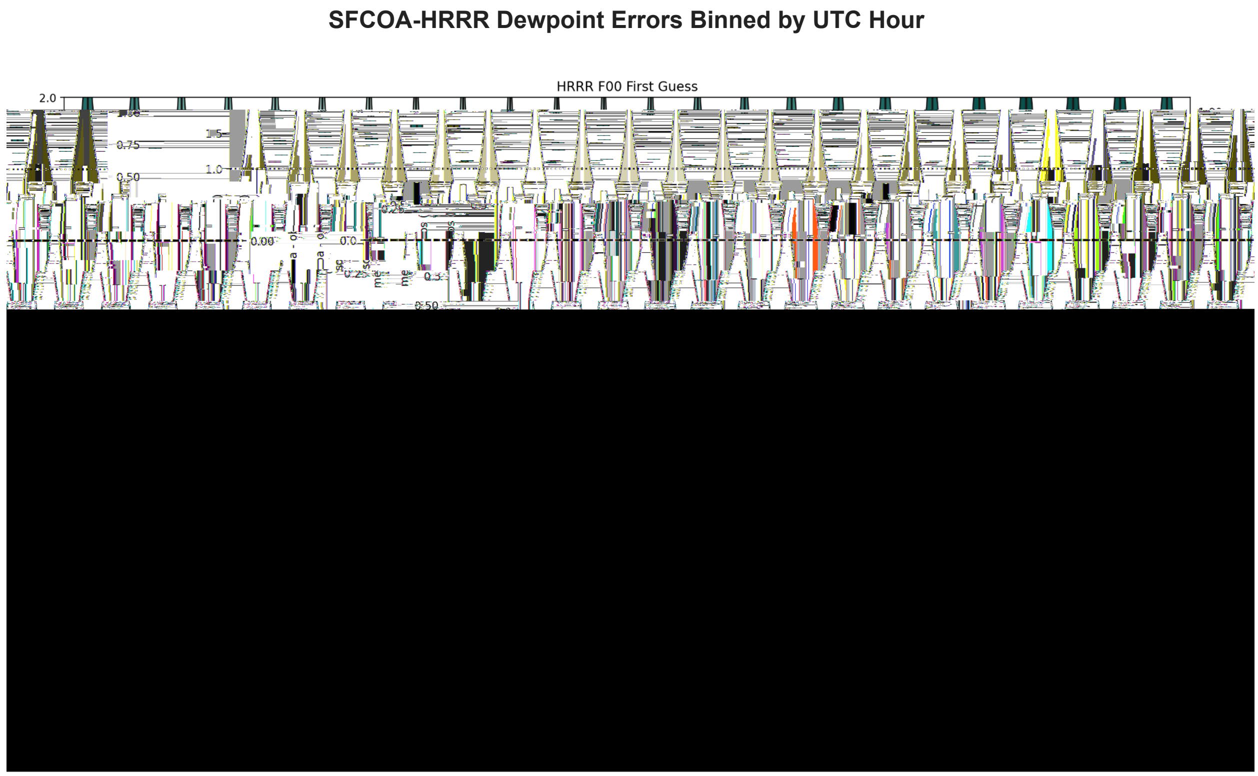

Figure 4. Same as Fig. 2, except for 2-m dewpoint errors.

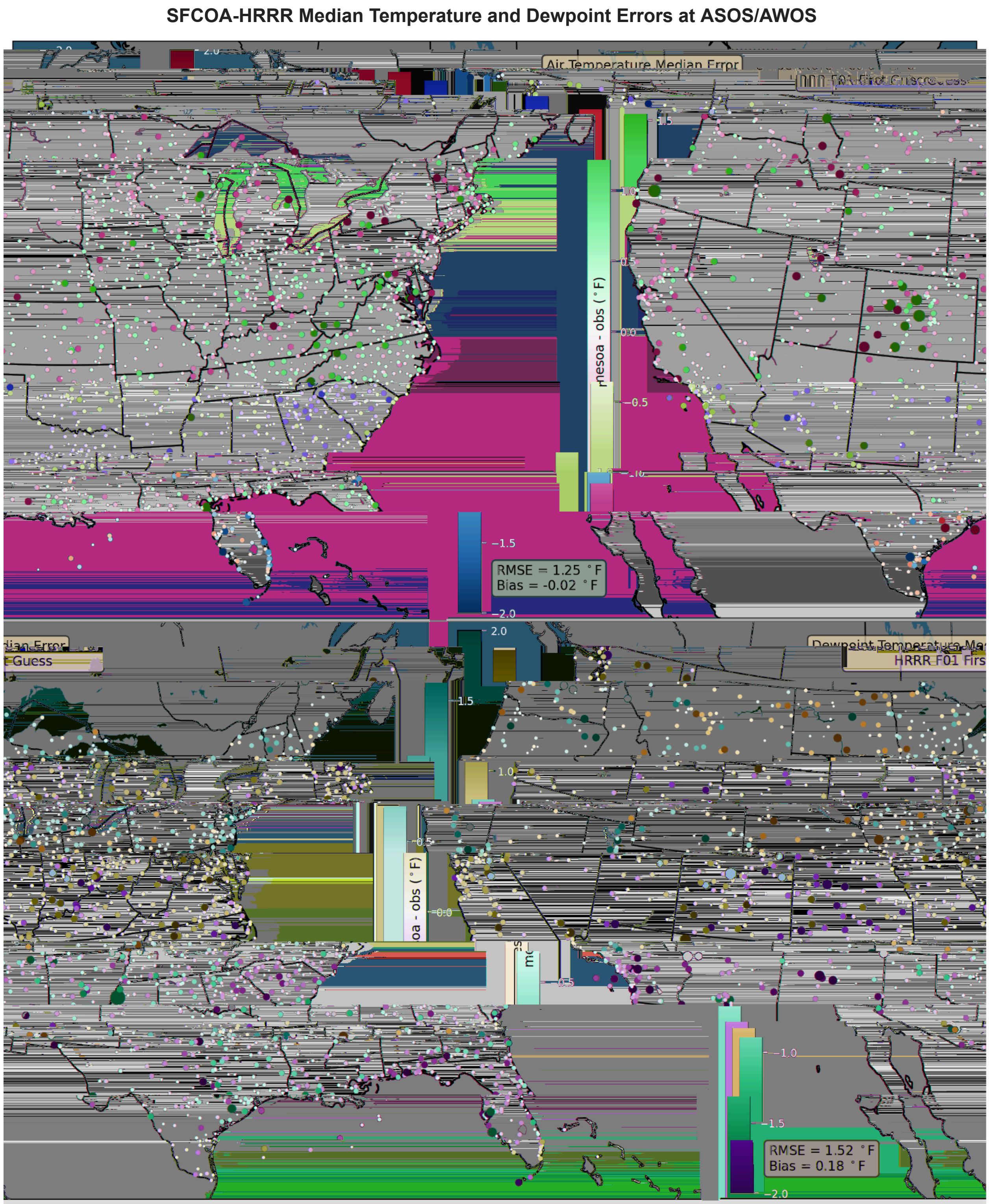

Figure 5. Median 2-m temperature (**a; upper**) and 2-m dewpoint (**b; lower)** errors at ASOS/AWOS sites during the HWT SFE.