14B.6 Exploring the Optimal Configuration of the High Resolution Ensemble Forecast System

Israel L. Jirak[1*], Adam J. Clark[2], Brett Roberts[1,2,3], Burkely Gallo[2,3], and Steven J. Weiss[1]
[1]NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK
[2]NOAA/OAR/National Severe Storms Laboratory, Norman, OK
[3]CIMMS, University of Oklahoma, Norman, OK

1. INTRODUCTION

The High Resolution Ensemble Forecast system version 2 (HREFv2) was implemented in the National Weather Service (NWS) on 1 November 2017 as an operational version of the Storm Prediction Center (SPC) Storm-Scale Ensemble of Opportunity (SSEO). Since 2011, the SSEO showed utility in SPC operations on a year-round basis and generally performed as good as or better than formally designed convection-allowing model (CAM) ensembles in providing severe weather guidance during past Spring Forecasting Experiments (SFEs) in the NOAA Hazardous Weather Testbed (HWT; Jirak et al. 2016). During the 2017 HWT SFE, the HREFv2 performed similarly to the SSEO, including higher subjective ratings and slightly better objective verification metrics than the other CAM ensembles examined (Clark et al. 2017). Thus, the operational HREFv2 serves as a meaningful baseline against which experimental and future CAM ensembles should be compared for consideration of operational implementation.

With the July 2018 operational implementation of the extended High Resolution Rapid Refresh (HRRR) runs at 0000, 0600, 1200, and 1800 UTC to 36 hours, there is an opportunity to include the HRRR model as an additional member of the HREF. Multiple experimental configurations of the HREF were tested and evaluated during the 2018 HWT SFE to help inform how to best configure the next operational version of HREF (i.e., v2.1). The candidate HREFv2.1 configurations included versions that add the extended HRRR runs to the HREFv2, as well as versions that remove some or all of the time-lagged members.

This paper will focus on the subjective evaluation of the different experimental HREF configurations during the 2018 HWT SFE. The current and candidate HREF configurations are described in the following section. Results from the comparison of the different HREF configurations during the 2018 HWT SFE are presented in the third section, followed by a summary and conclusions.

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

2. HREF CONFIGURATION

The HREFv2 consists of eight members with half of the members being time-lagged runs. The models are run at ~3-km grid spacing, using a multi-model (WRF-ARW & NMMB), multi-initial condition (NAM & RAP), and multi-physics approach to diversify forecast solutions (Table 1).

The implementation of the HRRRv3 into National Weather Service operations includes extended-length forecast runs out to 36 hours every six hours (i.e. at 0000, 0600, 1200, and 1800 UTC). This offers an opportunity to include the HRRRv3 as an additional member to the HREF (Table 2). To provide an evidence-based approach for making configuration decisions at the Environmental Modeling Center (EMC), several potential HREF configurations were examined and evaluated during the 2018 HWT SFE. The evaluation focused on HREF configurations that would maintain forecast diversity (i.e., multi-core, multi-IC). These different HREF configurations (Table 3) included the current HREFv2 configuration for comparison with five other candidate HREF configurations that added the HRRRv3, as well as four versions that removed selected time-lagged members.

Table 1. HREFv2 member configuration showing initial conditions (ICs)/lateral boundary conditions (LBCs), planetary boundary layer (PBL) schemes, and microphysics schemes. *SPC uses the 12-h time-lagged NAM Nest while NCO uses the 6-h time-lagged NAM Nest in HREFv2 products.

Member	ICs/LBCs	PBL	Micro
HRW NSSL	NAM/NAM -6h	MYJ	WSM6
HRW NSSL -12h	NAM/NAM -6h	MYJ	WSM6
HRW ARW	RAP/GFS -6h	YSU	WSM6
HRW ARW -12h	RAP/GFS -6h	YSU	WSM6
HRW NMMB	RAP/GFS -6h	MYJ	F-A
HRW NMMB -12h	RAP/GFS -6h	MYJ	F-A
NAM Nest	NAM/NAM	MYJ	F-A
NAM Nest -12h*	NAM/NAM	MYJ	F-A

Table 2. Same as Table 1, except for the HRRRv3 configuration, as potential addition(s) to the HREFv2.1

Member	ICs/LBCs	PBL	Micro
HRRRv3	RAP/RAP -3h	MYNN	Thompson
HRRRv3 -6h	RAP/RAP -3h	MYNN	Thompson

Table 3. Different HREF configurations explored during the 2018 HWT SFE. Left column includes the name of the configuration, including time-lagged (TL) members, a description of the configuration, and the total number of ensemble members.

HREF Config	Description	#
HREFv2	Current Config (Table 1)	8
HREFv2+HRRR	Add HRRR & HRRR TL	10
HREFv2+HRRR (No TL)	Remove **all** TL members	5
HREFv2+HRRR (No HRRR TL)	Remove HRRR TL member	9
HREFv2+HRRR (No NMMB TL)	Remove NMMB TL members	8
HREFv2+HRRR (No ARW TL)	Remove ARW TL members	7

3. RESULTS

Forecasts from the different HREF configurations were available for next-day evaluation during the 2018 HWT SFE, providing an opportunity for subjective comparisons among the configurations with regard to providing severe weather guidance. The HWT SFE participants examined the forecasts from the different HREF configurations using a multi-panel plot with observational overlays. Then, the participants provided a subjective rating of the forecasts based on their assessment of the utility of this guidance for a severe weather forecaster. Overall, the forecasts from the different HREF configurations appeared qualitatively similar on most days. While there were some differences among the forecasts, a careful examination was typically required to see the details.

On the first day of the 2018 HWT SFE, there were notable differences in the forecasts from the HREF configurations (Fig. 1). In the 26-hour forecast valid at 0200 UTC on 1 May 2018, the forecast from the HREFv2+HRRR (No ARW TL) (Fig. 1, bottom right panel) was rated higher by most participants than the HREFv2+HRRR (No NMMB TL) (Fig. 1, bottom middle panel) forecast. The HREFv2+HRRR (No ARW TL) better captures the axis of severe hail across central Nebraska within higher updraft helicity (UH; Kain et al. 2008) probabilities and also has an extension of low UH probabilities into southwest Kansas, where isolated severe hail was reported.

Other representative examples from the 2018 HWT SFE are shown in Figs. 2 and 3. On 2 May 2018, the forecasts from the HREFv2+HRRR (No TL) (Fig. 2; top-right panel) were subjectively rated higher (primarily for higher UH probabilities in southwest Oklahoma) by most participants than forecasts from the other HREF configurations. More typical, however, were the forecasts for 23 May 2018, where all HREF configurations generated very similar forecasts of severe wind potential (Fig. 3).

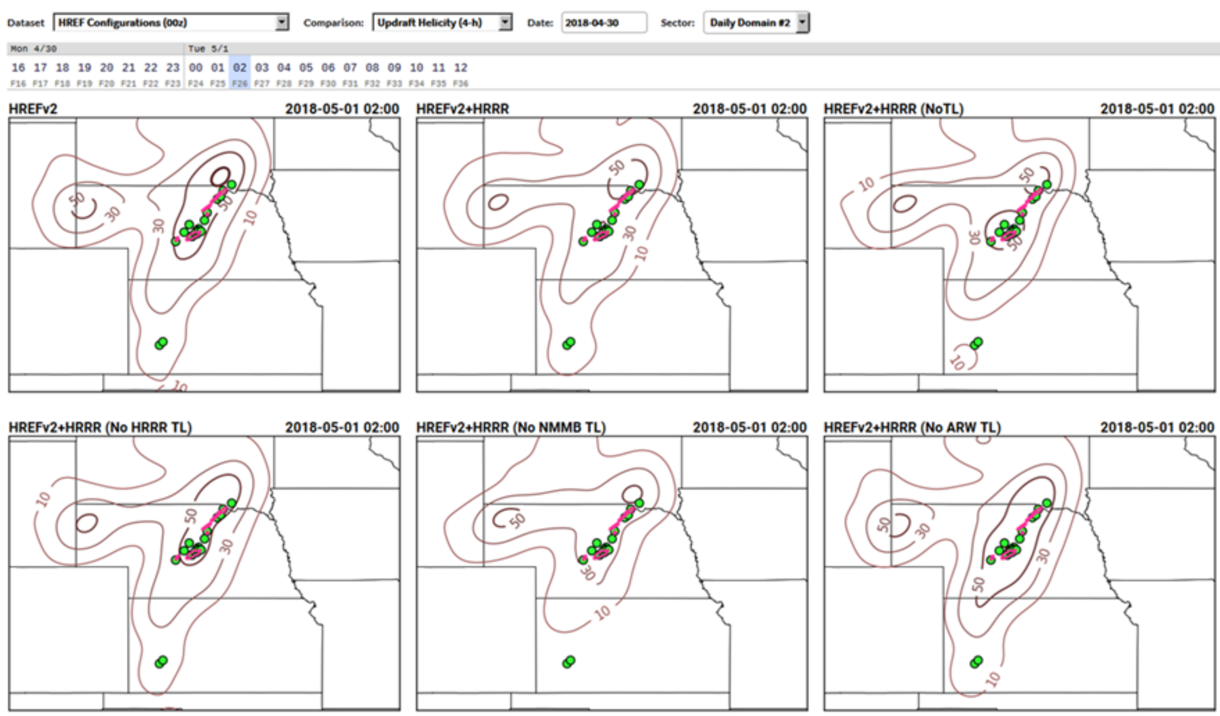

Figure 1. Six-panel comparison plot used to conduct the evaluation of the 0000 UTC HREF configurations during the 2018 HWT SFE. The 4-h neighborhood UH probability forecasts exceeding 75 m^2/s^2 valid for 0200 UTC on 1 May 2018 are shown for the current HREFv2 configuration (top-left panel), HREFv2+HRRR (top-middle panel), HREFv2+HRRR (No TL) (top-right panel), HREFv2+HRRR (No HRRR TL) (bottom-left panel), HREFv2+HRRR (No NMMB TL) (bottom-middle panel), and HREFv2+HRRR (No ARW TL) (bottom-right panel). The observed severe hail reports (green circles) and observed radar-derived maximum estimated size of hail (MESH; pink swaths) are overlaid as a reference for subjective verification.

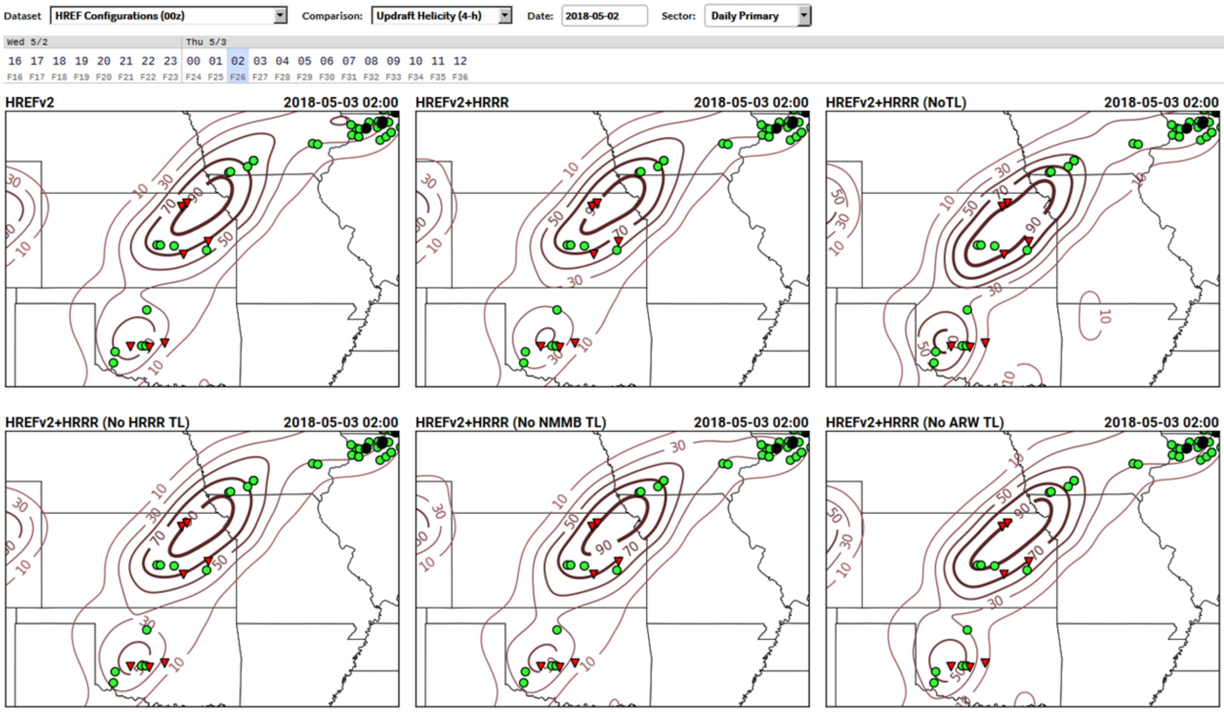

Figure 2. Same as Fig. 1, except for forecasts valid 0200 UTC on 3 May 2018. The upside-down red triangles represent tornado reports, and the black circles represent significant hail (i.e., ≥2" diameter) reports.

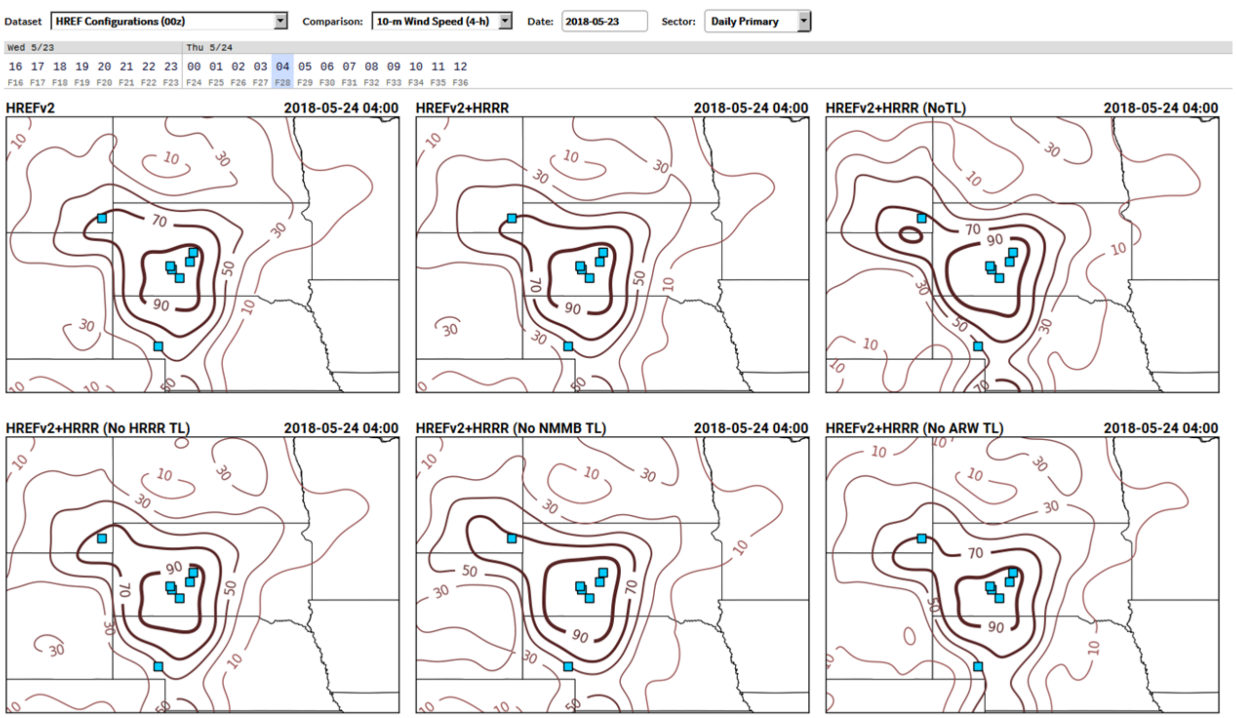

Figure 3. Same as Fig. 1, except for 4-h neighborhood probabilities of 10-m wind speeds exceeding 30 kts valid at 0400 UTC on 24 May 2018. The blue squares represent severe/damaging wind reports.

Overall, the different HREF configurations were rated similarly in terms of providing severe weather guidance during the five-week 2018 HWT SFE with mean subjective ratings ranging between 6.1 to 6.4 (Fig. 4). In fact, all of the HREF configurations had a median rating of 7 (out of 10), except for the HREFv2+HRRR (No NMMB TL) configuration, which had a lower median rating of 6. Subjectively, there was day-to-day variability in the performance of the various HREF configurations with forecasts on most days appearing similar enough to not provide a practical difference to a forecaster (i.e., differences not large enough to change an outlook). This result was not necessarily expected for the 10-member HREFv2+HRRR configuration compared to the 5-member HREFv2+HRRR (No TL) configuration, but it does highlight the resiliency of an ensemble to membership changes.

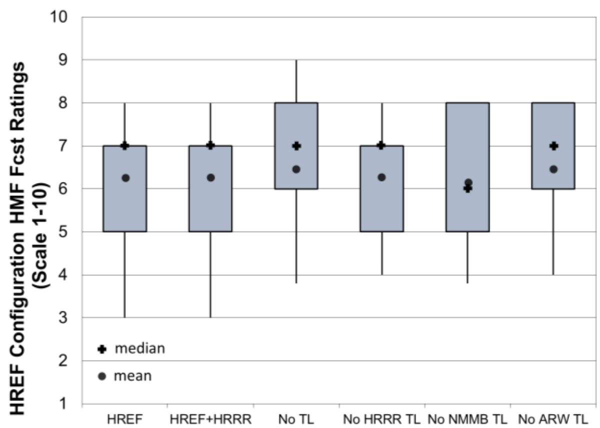

Figure 4. Box-and-whiskers plot of subjective ratings (1-10) for ensemble neighborhood probabilistic forecasts of hourly maximum fields (Kain et al. 2010; e.g., UH) from the HREF configuration experiment during the five-week 2018 HWT SFE. The boxes represent the interquartile range, and the whiskers represent the 10th and 90th percentiles. The crosses represent the median ratings, and the circles represent the mean ratings.

To investigate another perspective of the subjective ratings, the number of times that each HREF configuration was given the single-highest rating for a particular forecast was recorded. This indicated when a HWT SFE participant felt that one HREF configuration stood out as the top performer for a particular forecast. For the majority of forecasts, no HREF configuration stood out as the top performer (Fig. 5). The HREF configuration without any time-lagged members [HREFv2+HRRR (No TL)] was rated as the top performer more often than the other configurations, but it was only for a small percentage (~15%) of the ratings. On occasions when older runs performed poorly, removing them from the HREF improved the probabilistic forecast. However, for most cases, the time-lagged members did not degrade and actually improved the probabilistic ensemble forecast.

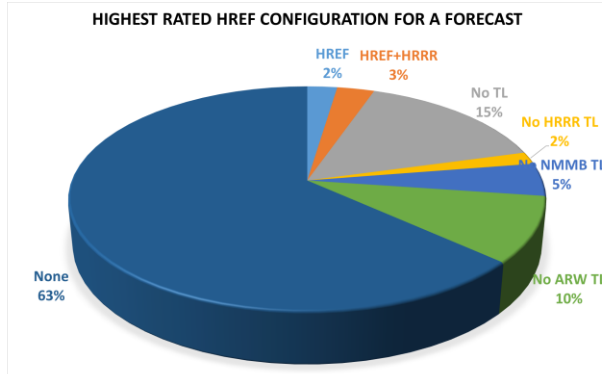

Figure 5. Pie chart showing the percentage of forecasts for which an HREF configuration received the *highest* subjective rating during the 2018 HWT SFE.

Similarly, the number of times that each HREF configuration was given the single-*lowest* rating for a particular forecast was documented. It was even more common for none of the HREF configurations to stand out as the worst performer, as more than three-fourths of the ratings did not highlight a single poorest-performing configuration (Fig. 6). The HREFv2+HRRR (No NMMB TL) configuration was rated as the worst performer more often than any other configuration (i.e., ~9% of the ratings), which is somewhat surprising given the perception that NMMB members do not perform as well as the ARW members for convective weather forecasting. The additional spread provided by time-lagged NMMB members occasionally contributed to improving the probabilistic severe weather guidance in HREF forecasts (e.g., southward convective initiation along a dryline).

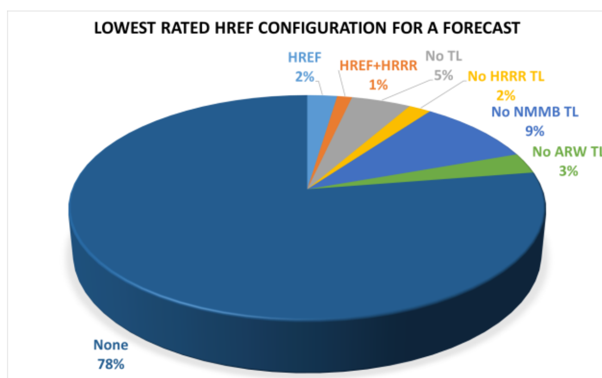

Figure 6. Same as Fig. 5, except for the percentage of forecasts in which a particular HREF configuration received the *lowest* subjective rating.

4. SUMMARY AND CONCLUSIONS

During the 2018 HWT SFE, a variety of HREF configurations were examined and evaluated to provide information and evidence to EMC on configuring the next operational version of HREF (i.e., v2.1). A multi-panel webpage allowed for the subjective evaluation of the different HREF configurations for severe weather forecasting. The HWT SFE participants provided subjective ratings of the HREF configurations in a next-day subjective evaluation with regard to the guidance provided to a severe weather forecaster.

The main takeaway from the HREF configuration comparison is that the various ***HREF configurations looked very similar overall*** on most days for severe weather guidance (i.e., the practical difference to a forecaster was small). On some days during the SFE, the time-lagged members did not perform as well as more recent convection-allowing model runs, so removing them improved the probabilistic ensemble forecast. Unexpectedly, the time-lagged NMMB members appeared to add more value (through increased ensemble diversity/spread) than the time-lagged ARW members during the 2018 HWT SFE, as the HREFv2+HRRR (No NMMB TL) configuration was overall rated the lowest. More work is needed in objective verification [i.e., surrogate severe (Sobash et al. 2016), reflectivity] of these HREF configurations to confirm the results from the subjective evaluations during the 2018 HWT SFE.

Acknowledgements. The evaluation of the HREF configurations would not have been possible during the 2018 HWT SFE without dedicated participants and the support and assistance of numerous individuals at SPC, NSSL, and the Cooperative Institute for Mesoscale Meteorological Studies (CIMMS) at the University of Oklahoma.

REFERENCES

Clark, A. J., I. L. Jirak, and Co-authors, 2017: 2017 HWT Spring Forecasting Experiment: Preliminary findings and results. Available online: https://hwt.nssl.noaa.gov/Spring_2017/HWT_SFE_2017_Prelim_Findings_Final.pdf

Jirak, I. L., C. J. Melick, and S. J. Weiss, 2016: Comparison of the SPC Storm-Scale Ensemble of Opportunity to other convection-allowing ensembles for severe weather forecasting. Preprints, 28th Conf. Severe Local Storms, Portland, OR. Amer. Meteor. Soc., P102.

Kain, J. S., S. J. Weiss, D. R. Bright, M. E. Baldwin, J. J. Levit, G. W. Carbin, C. S. Schwartz, M. L. Weisman, K. K. Droegemeier, D. B. Weber, K. W. Thomas, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931-952.

Kain, J. S., S. R. Dembek, S. J. Weiss, J. L. Case, J. J. Levit, and R. A. Sobash, 2010: Extracting unique information from high resolution forecast models: Monitoring selected fields and phenomena every time step. *Wea. Forecasting*, **25**, 1536–1542.

Sobash, R. A., C. S. Schwartz, G. S. Romine, K. Fossell, and M. Weisman, 2016: Severe weather prediction using storm surrogates from an ensemble forecasting system. *Wea. Forecasting*, **31**, 255–271.