BAMS
Meeting Summary

Advancing Hazardous Weather Prediction in the 2024 NOAA Hazardous Weather Testbed Spring Forecasting Experiment

Adam J. Clark,[b,d] Israel L. Jirak,[a] Timothy A. Supinie,[a] Kent H. Knopfmeier,[b,c] David Harrison,[a,c] Jake Vancil,[a,c] David Jahn,[a,c] Christopher D. Karstens,[a,d] Eric D. Loken,[b,c,d] Miranda Silcott,[b,c] Ryan Martz,[b,c,d] Nathan A. Dahl,[a,c] David Imy,[b] Andrew R. Wade,[a,c] Jeffrey M. Milne,[a,c,d] Kimberly A. Hoogewind,[b,c] Sean Ernst,[a,e] Joseph C. Picca,[a,c] Matthew Flournoy,[a,d] Michael Baldwin,[a,c] Pamela L. Heinselman,[b,d] Montgomery Flora,[b,c] Joshua Martin,[b,c] Brian C. Matilla,[b,c] Kristin Calhoun,[b] Thomas J. Galarneau Jr.,[b] Derek Stratman,[b,c] Corey Potvin,[b,d] Patrick S. Skinner,[b,c,d] and Patrick Burke[b]

KEYWORDS:
Atmosphere; Ensembles; Forecasting techniques; Numerical weather prediction/forecasting; Operational forecasting

The 2024 NOAA Hazardous Weather Testbed Spring Forecasting Experiment

What: Over 160 forecasters and researchers convened in-person and virtually to engage in real-time severe weather forecasting and evaluation activities aimed at accelerating research-to-operations and informing NOAA's Unified Forecast System. Major emphases of SFE 2024 included 1) deterministic and ensemble components of the Rapid Refresh Forecast System, 2) the Model for Prediction Across Scales, 3) global artificial intelligence (AI)–based NWP emulators, 4) the Warn-on-Forecast System, and 5) innovative AI-based postprocessing strategies.

When: 29 April–31 May 2024

Where: Norman, OK, and Online

DOI: 10.1175/BAMS-D-24-0249.1

Corresponding author: Adam J. Clark, adam.clark@noaa.gov

Manuscript received 6 September 2024

AFFILIATIONS: [a] NOAA/NWS/NCEP Storm Prediction Center, Norman, Oklahoma; [b] NOAA/OAR National Severe Storms Laboratory, Norman, Oklahoma; [c] Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman, Oklahoma; [d] School of Meteorology, University of Oklahoma, Norman, Oklahoma; [e] Institute for Public Policy Research and Analysis, University of Oklahoma, Norman, Oklahoma

1. Introduction

The 2024 NOAA Hazardous Weather Testbed Spring Forecasting Experiment (2024 SFE) marked the second consecutive hybrid SFE, with 68 of 160 participants contributing remotely. SFEs are coled by the NWS/Storm Prediction Center (SPC) and OAR/National Severe Storms Laboratory (NSSL) and since the early 2000s have played a key role for the U.S. severe weather enterprise by testing and optimizing new prediction tools through end-user-focused research. The wealth of severe weather forecasting and research expertise at the National Weather Center in Norman, OK, combined with state-of-the-art visualization tools, well-designed experiments, and valuable collaborations have made the annual SFEs one of the most productive and well-respected weather forecasting experiments in the world. SFE 2024 was particularly important because the configuration of NOAA's first formally designed convection-allowing model (CAM) ensemble known as REFS (Rapid Refresh Ensemble Forecast System) was finalized in fall 2023, so the SFE 2024 was vital to conducting a thorough subjective evaluation during the peak severe weather season before a proposed operational implementation in 2025.

SFE 2024 forecasting activities included 1) severe weather outlooks for days 1–4, 2) 0–1 and 1–2 h severe weather outlooks made with and without a new machine learning tool that combines the Warn-on-Forecast System (WoFS; Heinselman et al. 2024) and ProbSevere output (Cintineo et al. 2020), and 3) short-term mesoscale discussions using WoFS. Model evaluations used the 33-member Community Leveraged Unified Ensemble (CLUE; Clark et al. 2018), which included REFS and RRFS,[1] medium-range CAM ensembles with 7-day forecasts, the Model for Prediction Across Scales (MPAS; Skamarock et al. 2018), and a 1-km grid-spacing Weather Research and Forecasting model configuration. Other evaluations explored artificial intelligence (AI)–based NWP emulators, mesoscale analyses, and other machine learning (ML) approaches for generating severe weather guidance from CAMs and global models.

[1] RRFS is the control member for REFS.

2. Some highlights of the 2024 SFE

In detailed blind (i.e., model labels were hidden) subjective and objective assessments, significant issues were identified in RRFS that degraded forecast quality. While the current rapidly updating operational CAM, the High-Resolution Rapid Refresh (HRRR) model, improved relative to the previous year because of a weather regime associated with more predictable severe weather (i.e., strong vertical wind shear with large instability), the RRFS forecasts were actually worse in 2024, which led to a greater performance gap relative to 2023. It was found that efforts to reduce the high precipitation bias in RRFS by implementing the Grell–Freitas (GF) convective parameterization caused many high-impact severe weather events to be missed by the RRFS. In these cases, the composite reflectivity fields were often characterized by spurious, low-intensity precipitation that seemed to preclude more

realistic storm development. Interestingly, the REFS ensemble results did not reveal forecast degradation and subjective ratings showed performance on par with the High-Resolution Ensemble Forecast System version 3 (HREFv3), the current operational CAM ensemble. Upon closer examination of individual REFS members, it was found that performance was closely tied to the specific convective parameterization scheme, with members using the scale-aware Simplified Arakawa–Schubert (saSAS) convective scheme performing much better than those using GF, so that the saSAS members positively impacted ensemble performance. These results have led NOAA's Environmental Modeling Center to reconsider RRFS/REFS configurations, which will likely delay operational implementation while feasible options to improve performance are explored.

SFE 2024 marked the second consecutive year of tests with MPAS, which were expanded relative to 2023. Similar to SFE 2023, MPAS results for day 1 and 2 lead times were very encouraging. In almost every deterministic comparison, MPAS performance was runner-up to the HRRR and superior to every Finite-Volume Cubed-Sphere (FV3)–based system evaluated. MPAS was also tested within an alternative REFS configuration. While a postprocessing issue prevented meaningful subjective results, objective measures showed that this MPAS REFS ensemble configuration performed similarly to HREFv3 and two other alternative REFS configurations that were considered.

Real-time forecasts from AI-based NWP emulators were subjectively evaluated for the first time in SFE 2024. Facilitated by the availability of large reanalysis datasets like ERA5, companies like NVIDIA and Google have made rapid progress in developing purely AI-based algorithms for weather prediction that run faster and have better objective performance using standard statistical metrics than traditional NWP models. An activity was designed in which SFE participants assessed the skill of three AI global NWP emulators at 7-day lead times. These forecasts were compared to the operational Global Forecast System (GFS), and GFS analyses were used as the verification standard. The AI models included Google's GraphCast, Huawei Cloud's Pangu-Weather, and NVIDIA's FourCastNet v2. Evaluations focused primarily on 500-mb (1 mb = 1 hpa) geopotential heights and winds in the regional SFE domains. In this first-of-its-kind Hazardous Weather Testbed (HWT) evaluation, the AI models often demonstrated skill in predicting the synoptic-scale pattern, but participants still favored the GFS overall. Concern was expressed that AI forecasts were too smooth and often failed to resolve more subtle shortwave troughs that influenced the severe weather on a given day. Subsequent SFEs will continue to monitor and evaluate progress in this important area.

3. More information on the 2024 SFE

In the 2024 SFE, progress was made in key areas to accelerate R2O and assist evidence-based decision-making for models and tools that improve operational severe weather forecasts. The full 2024 SFE Summary Report is available at https://hwt.nssl.noaa.gov/sfe/2024/docs/SFE2024_tech_memo.pdf.

Acknowledgments. This work comprised regular duties at federally funded NOAA/NSSL for Clark, Imy, Heinselman, Burke, Calhoun, Galarneau, and Potvin and NOAA/SPC for Jirak, Supinie, Karstens, and Flournoy. Knopfmeier, Harrison, Vancil, Jahn, Loken, Silcott, Martz, Dahl, Wade, Milne, Hoogewind, Ernst, Picca, Baldwin, Flora, Martin, Matilla, Stratman, and Skinner were funded by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma (OU) Cooperative Agreement NA21OAR4320204, U.S. Department of Commerce. OU's Institute for Public Policy Research and Analysis also funded Ernst.

References

Cintineo, J. L., M. J. Pavolonis, J. M. Sieglaff, L. Cronce, and J. Brunner, 2020: NOAA ProbSevere v2.0—ProbHail, ProbWind, and ProbTor. *Wea. Forecasting*, **35**, 1523–1543, https://doi.org/10.1175/WAF-D-19-0242.1.

Clark, A. J., and Coauthors, 2018: The Community Leveraged Unified Ensemble (CLUE) in the 2016 NOAA/Hazardous Weather Testbed Spring Forecasting Experiment. *Bull. Amer. Meteor. Soc.*, **99**, 1433–1448, https://doi.org/10.1175/BAMS-D-16-0309.1.

Heinselman, P. L., and Coauthors, 2024: Warn-on-Forecast System: From vision to reality. *Wea. Forecasting*, **39**, 75–95, https://doi.org/10.1175/WAF-D-23-0147.1.

Skamarock, W. C., M. G. Duda, and S.-H. Ha, 2018: Limited-area atmospheric modeling using an unstructured mesh. *Mon. Wea. Rev.*, **146**, 3445–3460, https://doi.org/10.1175/MWR-D-18-0155.1.