13A.3 Operational Model Performance for Fire Weather Forecasting

Israel L. Jirak* and Matthew Elliott
NOAA/NWS/NCEP/Storm Prediction Center, Norman, OK

1. INTRODUCTION

The NOAA/NWS Storm Prediction Center (SPC) issues Fire Weather Outlooks for the contiguous United States (CONUS) for Days 1-8 (defined as 12-12 UTC). These Fire Weather Outlooks assess where pre-existing fuel conditions combined with forecast weather conditions will result in a significant threat for the spread of wildfires (i.e., strong winds concurrent with low relative humidity at the surface). Operational numerical weather prediction models play a vital role in informing these Fire Weather Outlooks, yet the performance of the models is rarely assessed for fire weather forecasting capabilities. This study aims to quantify and document the performance characteristics of operational models for fire weather forecasting.

The data and methodology used in this study can be found in the following section. Results of fire weather forecasting performance in 2021 from coarse and high-resolution operational models are presented in the third section, followed by a summary and conclusions.

2. DATA AND METHODOLOGY

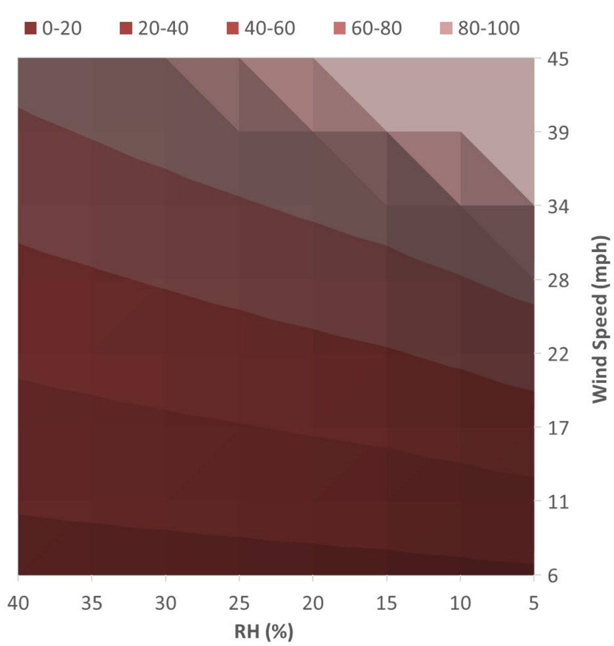
To assess the performance of operational models for fire weather forecasting, a relatively simple, straightforward approach was employed. The 21-hour forecast of the Fosberg Fire Weather Index (FFWI; Fosberg 1978) from 0000-UTC model runs (i.e., valid at 2100 UTC) during 2021 were verified across the CONUS. The focus on an afternoon valid time (i.e., 2100 UTC) was to highlight biases in surface relative humidity (RH) and wind speed that primarily arise from diurnal planetary boundary layer (PBL) mixing. The coarse models examined in this study include the North American Mesoscale (NAM) model, the Global Forecast System (GFS) model, and the European Center for Medium-range Weather Forecasting (ECMWF) model while the high-resolution models examined were the High-Resolution Rapid Refresh (HRRR), the HiRes Window (HRW) Advanced Research WRF (ARW), the HRW NSSL, the HRW Finite-Volume Cubed (FV3; available after 11 May), and the NAM CONUS Nest. The SPC RAP-based mesoanalysis (Bothwell et al. 2002) was utilized as the observational dataset to calculate the grid-based verification statistics on a common 40-km grid. The statistics were accumulated over the CONUS for each day of 2021, including root-mean-squared error (RMSE), mean error (ME), and 2x2 contingency table statistics [probability of detection (POD), false-alarm ratio (FAR), critical success index (CSI), and bias; Wilks 2006].

* *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

The FFWI, which is a non-linear combination of meteorological data that results in a linear relationship between meteorological conditions and wildfire behavior, was used a proxy in this study for fire weather forecasting. While the FFWI is not a widely used parameter, it does have the desired characteristic of combining wind speed and RH into a single variable for verification and is commonly examined by SPC forecasters for identifying critical fire weather areas. The FFWI is defined below:

$$\mathrm{FFWI} = [\eta\sqrt{(1+U^2)}]/3.002$$

Where U is the surface wind speed and η is the moisture damping coefficient that is a function of surface temperature and RH. The FFWI is scaled so that a value of 100 occurs with RH=0% and U=30 mph. See the graph of FFWI as a function of wind speed and RH in Fig. 1, and note the more rapid increase of FFWI values with increasing wind speed as opposed to decreasing RH. Please note that the FFWI only accounts for meteorological conditions and does not consider fuel conditions.

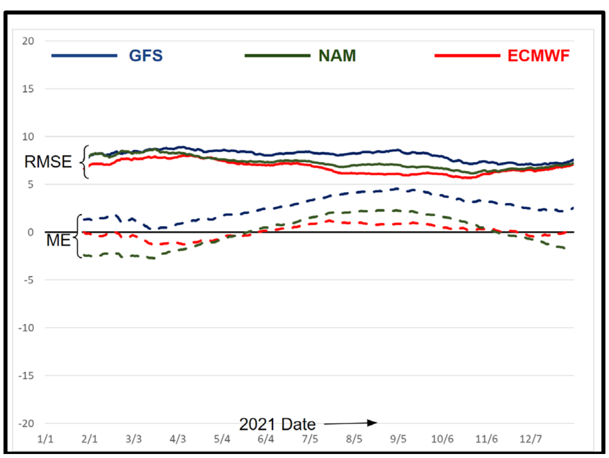


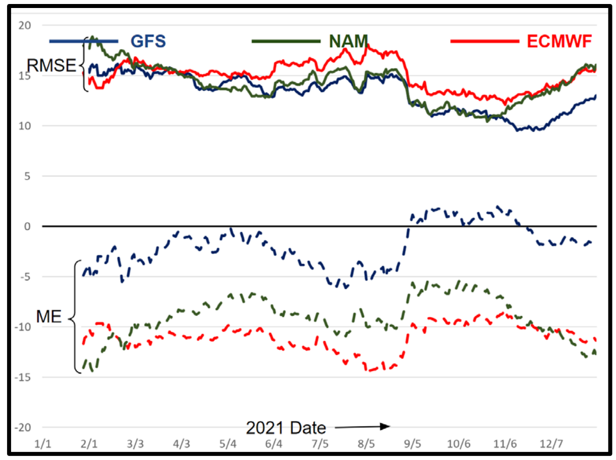
3. RESULTS

Forecasts of the FFWI from various operational 0000 UTC-initialized models were verified at 2100 UTC (i.e., 21-h forecast) for each day of 2021 to assess the performance for fire weather forecasting. There were three primary components to this objective verification of 1) coarse and 2) high-resolution models: a) RMSE and ME across the CONUS, where the 2-m temperature (T) exceeded 50F, b) RMSE and ME across the CONUS, where the 2-m T exceeded 50F *and the FFWI exceeded 40* to include only areas with fire weather concerns, and c) contingency-table statistics via a performance diagram (Roebber 2009) for FFWI values of 40, 50, and 60, where the 2-m T exceeded 50F.

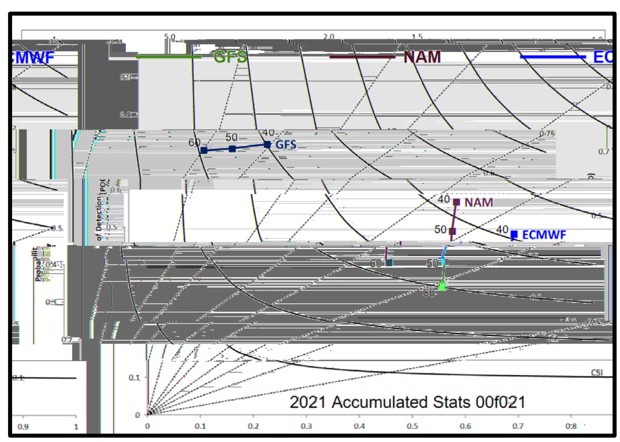
3.1 Objective Verification of Coarse Models

When verifying the FFWI for all grid points across the CONUS with 2-m T >50F, the ECMWF stands out as having the lowest (i.e., best) RMSE and best ME (near 0) while the GFS has the highest (i.e. worst) RMSE and ME (Fig. 2). However, if the verification is restricted to locations where the analyzed FFWI values are above 40 (i.e., where fire weather concerns exist), the results look much different (Fig. 3). The first thing to notice is that the errors are larger and vary more on a daily basis, owing to the smaller sample size. When only examining operationally meaningful values of the FFWI, the GFS has the lowest (i.e., best) RMS and best ME (closest to 0) while the ECMWF now has the worst RMSE and ME (Fig. 3). These objective verification statistics, when limited to areas with FFWI values exceeding 40, agree with subjective assessments of model usefulness by SPC forecasters when issuing Fire Weather Outlooks.



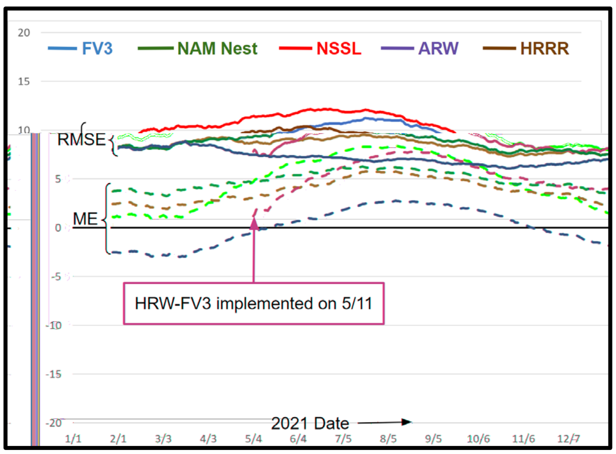


For another statistical perspective, the performance diagram highlights different model attributes (Fig. 4). Despite very similar overall CSI values for all three models, they have very different performance characteristics (Fig. 4). The ECMWF has a low bias (0.5-0.7), POD (~0.4), and FAR while the GFS has a high bias (1.5-2.0), POD (~0.7), and FAR for fire weather forecasting. The performance attributes of the NAM fall in between the ECMWF and GFS with a tendency toward a low bias and low POD. It is worth noting that for high-impact weather forecasting, like fire-weather forecasting, POD is a very important forecast attribute even if it comes with a high bias. Thus, given the similar CSI values among the GFS, NAM, and ECMWF, SPC forecasters find the GFS more useful for fire-weather forecasting given the much higher POD, followed by the NAM and then ECMWF in terms of usefulness for fire-weather forecasting.

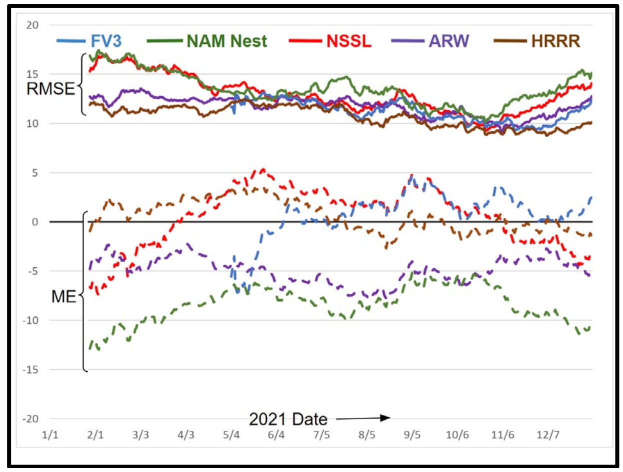


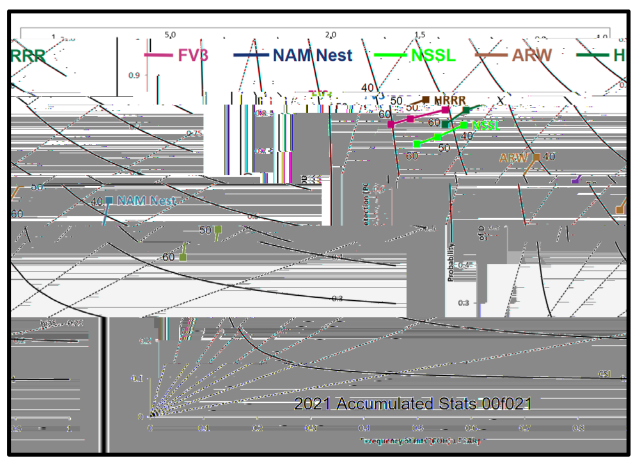
3.2 Objective Verification of High-Resolution Models

For all grid points across the CONUS with 2-m T >50F, the NAM Nest stands out as having the lowest (i.e., best) RMSE and best ME (near 0) of the FFWI while the NSSL has the highest (i.e. worst) RMSE and ME (Fig. 5). It is interesting to note that the high-resolution models actually have larger RMSE and ME than the coarse models for this analysis (cf. Figs. 2 and 5). Again, the results look much different if the verification is restricted to where fire weather concerns may exist (i.e., FFWI values above 40). First of all, the errors are slightly larger overall and vary more on a daily basis with this restriction, but are now smaller than those of the coarse models (cf. Figs. 3 and 6). When only examining operationally meaningful values of the FFWI, the HRRR has the lowest (i.e., best) RMS and best ME (closest to 0) while the NAM Nest now has the worst RMSE and ME (Fig. 6). This is consistent with the coarse-model results, where models with a low bias for fire weather conditions have the lowest errors when considering all grid points. Objective results from limiting the verification to areas with FFWI values exceeding 40 agree better with subjective assessments by SPC forecasters than the results for all grid points.

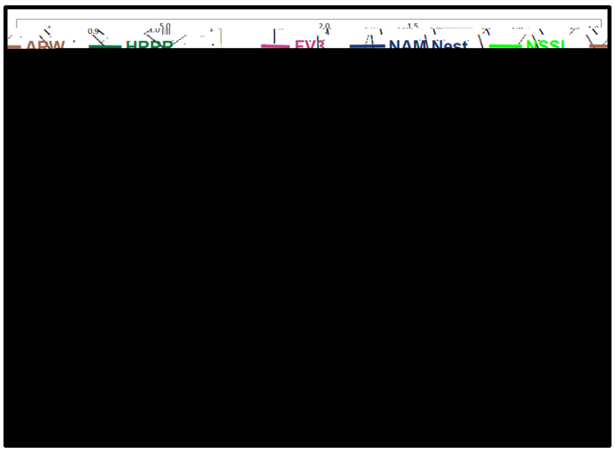


The performance diagram highlights some interesting differences among the operational high-resolution models (Fig. 7). The NAM Nest stands out as the model with the lowest bias (0.7-0.9), POD (~0.5), and FAR. There are three models (FV3, NSSL, and HRRR) clustered in the upper-left portion of the diagram with similar characteristics: high bias (>2.0) and high POD (0.7-0.8). Of these three, however, the HRRR has notably lower FAR. Lastly, the ARW has average characteristics that fall between the HRRR and NAM Nest. Even though the NAM Nest, ARW and HRRR have similar overall CSI values (perhaps even slightly higher for the NAM Nest), the importance of POD for high-impact events for fire weather forecasting suggests that the HRRR is the most useful operational high-resolution model for forecasting fire weather conditions, followed by the ARW and NAM Nest. The NSSL and FV3 lag behind the other three models in terms of overall usefulness owing to the higher bias and FAR. These performance characteristics and overall rankings are generally consistent with subjective perspectives of SPC forecasters who use these models on a daily basis for generating Fire Weather Outlooks.





While the focus of this study was on the performance and characteristics of deterministic operational models, the High-Resolution Ensemble Forecast (HREF; Roberts et al. 2019) system was examined to quantify the value of ensembles for fire-weather forecasting. The HREF is a 10-member time-lagged ensemble consisting of the five runs examined herein along with their respective time-lagged members. The mean of the FFWI from all ten HREF members was verified alongside the individual members. The HREF results in improved performance (i.e., higher CSI) for fire weather forecasting over any of the individual members (Fig. 8). Not surprisingly, the characteristics (POD, FAR, bias) of the HREF lie in the middle of the characteristics of the individual HREF members with the most similarity to the deterministic ARW.



4. SUMMARY AND CONCLUSIONS

The performance of coarse and high-resolution 0000-UTC operational models was examined for fire weather forecasting by verifying the Fosberg Fire Weather Index (FFWI) valid at 2100 UTC for all days during 2021. One important takeaway is that the RMSE and ME for fire-weather forecasting are very sensitive to whether all grid points are examined or whether only grid points with operationally meaningful FFWI values (i.e. FFWI>40) are examined. Examining and verifying only the meaningful areas are better aligned with subjective impressions of SPC forecasters regarding the characteristics and utility of the models for fire-weather forecasting.

Regarding the coarse operational models examined, the GFS is better for fire-weather forecasting given its lower RMSE and ME and higher POD than the NAM and ECMWF for areas where fire weather concerns may exist. The ECMWF has an especially notable low bias and negative ME for fire weather conditions. These biases and characteristics can be seen in the historic wildfire event on 15 December 2021 across portions of the Central and Southern Plains (Fig. 9).

Regarding the high-resolution operational models examined, the HRRR, FV3, and NSSL all have a high bias and POD in forecasting fire weather conditions, but, of these, the HRRR has lower RMSE, ME, and FAR. The NAM Nest has a low bias and POD for fire weather conditions while the ARW has characteristics that fall between the HRRR and NAM Nest. Considering the importance of POD for fire weather forecasting, the HRRR is the best deterministic operational model for forecasting fire weather conditions followed by the ARW, NAM Nest, NSSL, and FV3 in order of usefulness to operational forecasters. Overall, the HREF performs better on average than any individual member for fire-weather forecasting, indicating the utility and usefulness of an ensemble for fire-weather forecasting. Many of these characteristics and biases can be seen in the forecasts for the historic wildfire event on 15 December 2021 across portions of the Central and Southern Plains (Fig. 10).

REFERENCES

Bothwell, P.D., J.A. Hart, and R.L. Thompson, 2002: An integrated three-dimensional objective analysis scheme, *Preprints*, 21st Conf. on Severe Local Storms, San Antonio, TX, Amer. Meteor. Soc., 18.2.

Fosberg, M.A. 1978. Weather in wildland fire management: the fire weather index. In Proceedings of the Conference on Sierra Nevada Meteorology, Lake Tahoe, CA.

Roberts, B., I. Jirak, A. Clark, S. Weiss, and J. 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.

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, https://doi.org/10.1175/ 2008WAF2222159.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences: An Introduction*. 2nd ed. Academic Press, 467 pp.

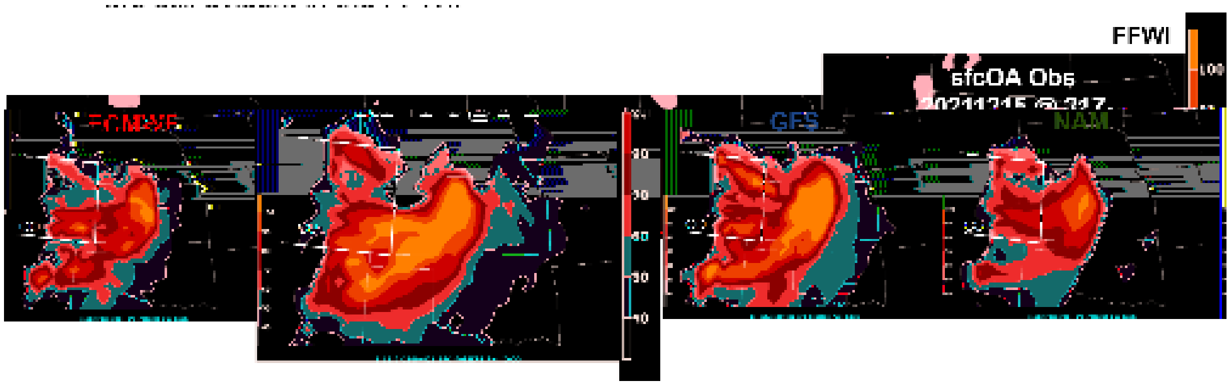


Figure 9. Forecasts of the FFWI for the GFS, NAM, ECMWF, and analysis (from left to right) valid at 2100 UTC on 15 December 2021 for the historic wildfire event across portions of the Central and Southern Plains.

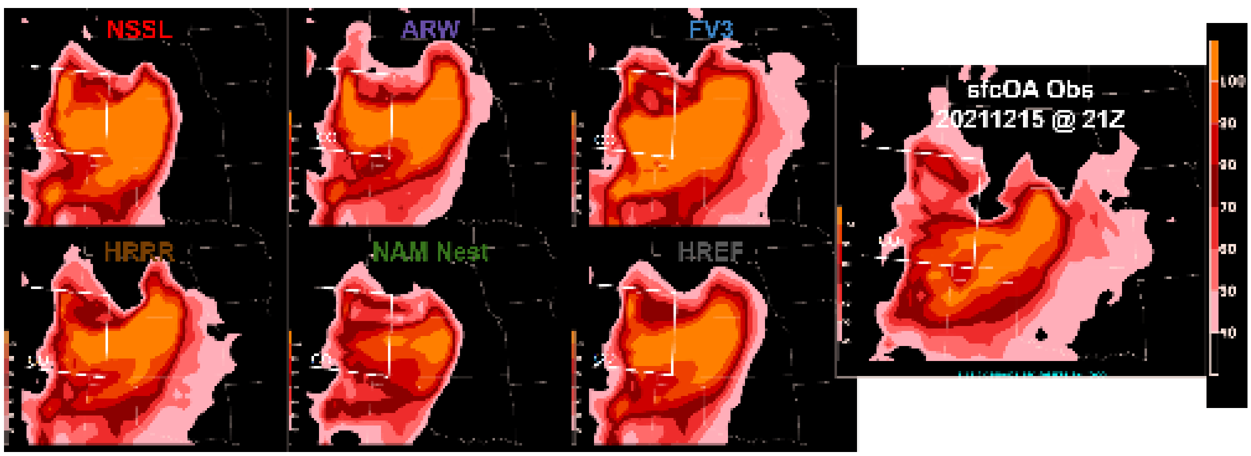


Figure 10. Forecasts of the FFWI for the NSSL (top left), ARW (top middle), FV3 (top right), HRRR (bottom left), NAM Nest (bottom middle), and HREF (bottom right) with the analysis (far right) valid at 2100 UTC on 15 December 2021 for the historic wildfire event across portions of the Central and Southern Plains.