6.4

VERIFYING THE IMPACT OF THE FV3-GFS UPGRADE ON LAMP STATISTICAL GUIDANCE

Elizabeth Venteicher*
Valparaiso University
Department of Geography and Meteorology
Valparaiso, Indiana

Judy E. Ghirardelli
Frederick G. Samplatsky
National Weather Service, NOAA
Office of Science and Technology Integration
Meteorological Development Laboratory
Silver Spring, Maryland

1. INTRODUCTION

As part of its Next Generation Global Prediction System (NGGPS) initiative, the National Weather Service (NWS) is upgrading the Global Forecast System (GFS) from a spectral model to a Finite Volume Cubed Sphere Dynamical Core (FV3) based model (Lin 2004). The FV3-GFS upgrade is the first step towards a fully integrated and flexible modeling framework which will optimize available computing capabilities. Changes to the dynamical core do not only impact the GFS itself, but also affect the multitude of post-processed products that utilize the GFS as an input.

One such product is the Localized Aviation MOS Program (LAMP) which is run hourly (and sub-hourly for some elements), providing rapidly updating short-term guidance in the 1 to 25-hour timeframe for aviation forecasters while simultaneously supporting the National Blend of Models (NBM) (Ghirardelli et al. 2010). LAMP uses numerical model output and observations to provide guidance for weather elements that are particularly impactful for aviation, such as ceiling height and visibility. Through statistical post-processing using multiple linear regression techniques, LAMP aims to mitigate the systematic biases that are inherent in modeled forecasts. However, upgrading the GFS could degrade the LAMP guidance by altering the characteristics of the bias profile that LAMP was developed to account for.

**Corresponding author address*: Elizabeth Venteicher, Valparaiso University, Department of Geography and Meteorology, 1509 Chapel Drive Unit #690, Valparaiso, IN 46383; email: ellie.venteicher@valpo.edu

The GFS is one of a multitude of predictors for LAMP and thus the FV3 upgrade was not expected to result in drastic changes, especially in the early projections when data sources such as observations have a stronger influence. However, in the later projection hours, the value of these other data sources diminish and model results drive LAMP forecasts to a greater degree (figure 1). Thus any alterations that we observe were expected to occur in latter projections. In order to test for any such changes, a statical verification was conducted from 26 January to 30 June 2018 for a variety of predictands including temperature, dewpoint, wind speed, ceiling heights, and visibility. Verification of the convection and lightning variables continued until 31 July 2018. This date range was selected as it occurred after the 25 January 2018 LAMP implementation and spanned both cool and warm seasons. LAMP is developed using two sets of equations, delineated by season, for each predictand. Convection and lightning have an additional spring season, which is also accounted for in this range of dates. Since each season was sampled in this dataset, the seasons were used to subset the data and separate verifications were found for each set of equations.

2. METHODOLOGY

This study was designed to provide an objective evaluation of the potential impacts on LAMP products due to the changes in one of its inputs. In order to achieve an accurate assessment, two parallel datasets were needed; the operational LAMP (OPER-LAMP) and the FV3-LAMP. The FV3-LAMP dataset was gathered by simply rerunning the LAMP model for the time period of interest but utilizing a retrospective FV3-GFS dataset provided by the National Centers for Environmental Prediction (NCEP) Environmental Modeling Center (EMC) as the predictor instead of the legacy GFS. This allowed for the performance of the two forecasts relative to verifying observations to be quantitatively compared using a variety of verification statistics. The verification observations for most of the forecast elements were measured at 1444 METAR and ASOS sites throughout the country. However, convection and lightning are gridded products and thus the verification technique was slightly different as the verification occurred on a 10-km grid. The verifying observations for these two elements came from Multi-Radar/Multi-Sensor (MRMS) composite reflectivity, and total lightning flash data provided by Earth Networks, Inc. This analysis was conducted for three model cycles, namely the 0400, 1500 and 1900 UTC model runs. The exact statistics used to complete the analysis depends on the variable type. These distinctions will be discussed in the results section below.

3. RESULTS

a. *Temperature, Dewpoint, and Wind Speed*

For continuous variables such as temperature, dewpoint and wind speed, two statistical methods were used; namely, Bias and Mean Absolute Error (MAE). Generally, the MAE for all three variables were comparable between the OPER-LAMP and FV3-LAMP for all three cycles tested at all projection hours and in both seasonal subsets (not shown). The only exception to this is the slight degradation seen in the mid-range projections of the warm season temperature forecasts, illustrated by slightly higher MAE values for the FV3-LAMP over its operational counterpart (figure 2).

Despite the minimal impacts displayed in the MAE results, there was greater variability observed in the bias profiles of all three variables. Wind speeds showed mixed results with some projection hours displaying an improvement for the FV3-LAMP over the operational product while others displayed the reverse (figure 3). Meanwhile, the dewpoint bias plots for all three cycles and both seasons illustrate a larger dry bias in the FV3-LAMP than the OPER-LAMP (figure 4).

Perhaps the most significant deviation observed between the two versions of LAMP occurred on the temperature bias plots, with the FV3-LAMP consistently being cooler than the OPER-LAMP. In the cool season, OPER-LAMP had a preexisting warm bias in its temperature products. Thus the cooler FV3-LAMP was actually an improvement over the operational version (figure 5). Conversely, there was not a similar preexisting bias during the warm season temperature equations and the FV3-LAMP had a significant cool bias. The degradation from the operational product exceeded 1 degree Fahrenheit for some cycles and projection hours (figure 6).

Since MOS (Glahn et al. 1972) is a large input in LAMP equations, the statistics for the operational GFS MOS and FV3-GFS MOS were also calculated for all variables and model cycles. The results were then plotted along with the LAMP verification results. This revealed that the behavior of the LAMP products seemed to mirror that of the MOS. For example, the largest degradations in the warm season temperature bias seemed to be exacerbated at certain projection hours. To determine what could be driving this phenomenon, the full MOS verification results, provided by the MOS verification team at the Meteorological Development Laboratory, were evaluated. MOS is valid for several days and with its increased number of projections, a clear diurnal pattern was revealed (figure 7). While both the operational and FV3-GFS MOS displays this behavior, the extent of the variability is more pronounced for the FV3-GFS MOS. This is consistent with the performance of the FV3-LAMP which tracks along with the FV3-GFS MOS and thus has a larger diurnal variability in its temperature forecasts.

b. *Ceiling Height and Visibility*

Ceiling height and visibility are categorical variables and as such, MAE and bias are not the appropriate statistics to use for the verification of these predictands (Weiss et al. 2005). Instead, Threat Scores, also known as Critical Success Index (Schaefer 1990), were calculated for ceiling heights < 500 ft, < 1000 ft, and ≤ 3000 ft as well as visibilities < 0.5 mi, < 3 mi, and < 5 mi. For the cool season, neither ceilings nor visibilities displayed any discrepancies between the OPER-LAMP and FV3-LAMP (not shown). In the warm season, there were some slight degradations at the midrange projection hours for both predictands (figure 8 and 9). However, since the magnitude of the discrepancy was minimal and it only occurred for select model cycles and categories, it was not deemed to be a meaningful impact.

c. *Convection and Lightning*

Since convection and lightning are probabilistic variables, Brier Skill Scores, which illustrates a forecast's improvement over climatology (Brier 1950), were used for verification. These plots revealed only negligible differences between the FV3-LAMP and OPER-LAMP forecasts (figure 10). Interestingly, the FV3-GFS MOS and operational GFS MOS did show larger discrepancies, but these differences were still minimal.

4. SUMMARY

For the majority of the predictands, there was not an appreciable difference between OPER-LAMP forecasts and FV3-LAMP forecasts. This indicates that the FV3-GFS does not meaningfully impact the LAMP guidance for most variables and model runs. However, there were some areas in which the FV3-LAMP did show noticeable deviations from the operational products. Specifically, temperature and dewpoint showed the most significant impacts. For these predictands, the FV3-LAMP exhibited cooler and drier bias. This resulted in a more accurate temperature forecast in the cool season where OPER-LAMP had a preexisting warm bias, but resulted in a degradation in the warm season forecasts where there was not a similar bias. The impact of these degradations, however, should be mitigated as the LAMP development team is planning to redevelop the temperature and dewpoint equations.

5. ACKNOWLEDGMENTS

This work was made possible by the opportunities afforded by the NCEP Student Internship Program and the dedicated help of the LAMP and MOS development teams. In addition, datasets such as EarthNetworks, Inc.'s total lightning data and the FV3-GFS retrospective data provided by NCEP/ Environmental Modeling Center were instrumental in this verification. Archived MRMS data and HRRR model output were also provided by NOAA's National Severe Storms Laboratory and the NOAA/Earth Systems Research Laboratory/Global Systems Division, respectively.

6. RESOURCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev*., **78**, 1–3.

Charba, J. P., F .G. Samplatsky, and P.E. Shafer, 2011: Experimental LAMP 2-h convection guidance on a 20-km grid. Preprints, 24th *Conference on Weather Analysis and Forecasting/ 20th Conference on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., **J19.3**.

Ghirardelli, J. E., and B. Glahn, 2010: The Meteorological Development Laboratory's aviation weather prediction system. *Wea. Forecasting*, **25**, 1027–1051.

Glahn, H. R., and D. A. Lowry, 1972: The use of Model Output Statistics (MOS) in objective weather forecasting. *J. Appl. Meteor. Climatol*, **11**, 1203–1211.

Glahn, B., K. Gilbert, R. Cosgrove, D. P. Ruth, and K. Sheets, 2009: The gridding of MOS. *Wea. Forecasting*, **24**, 520–529.

Lin, S.-J., 2004: A "vertically Lagrangian" finite-volume dynamical core for global models. *Mon. Wea. Rev.*, **132**, 2293–2307.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575.

Weiss, M., and J. E. Ghirardelli, 2005: A summary of ceiling height and total sky cover short-term Statistical forecasts in the Localized Aviation MOS Program (LAMP). Preprints, *21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction*, Washington, DC, Amer. Meteor. Soc., **13B.6**.

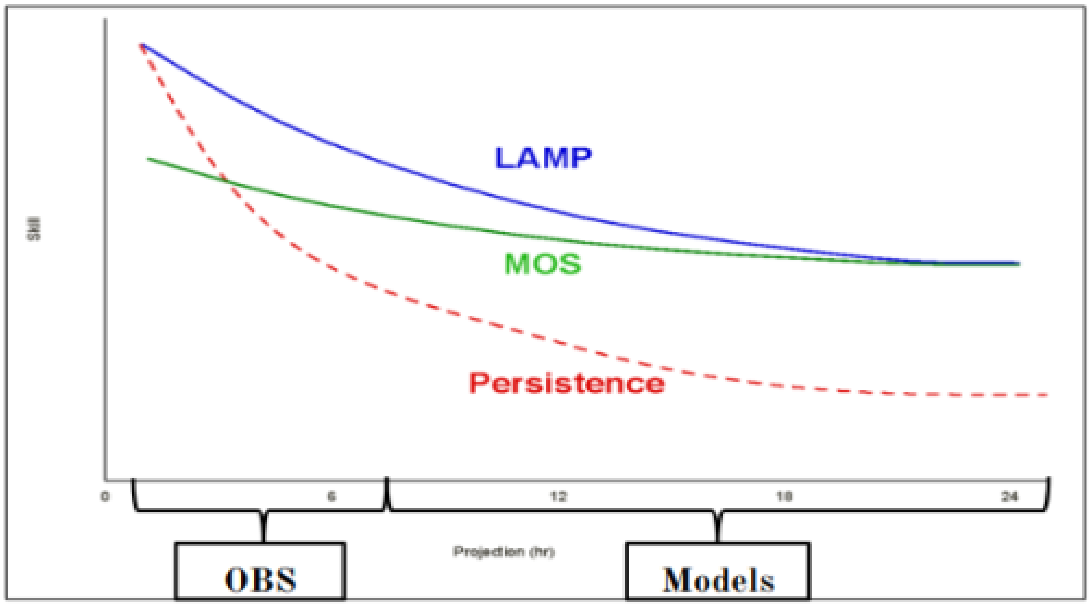

FIG. 1. Chart that illustrates the predominate predictors at given projection hours. The value of observations diminish with time as illustrated by the diminishing skill of persistence (dashed red line) at later projection hours.

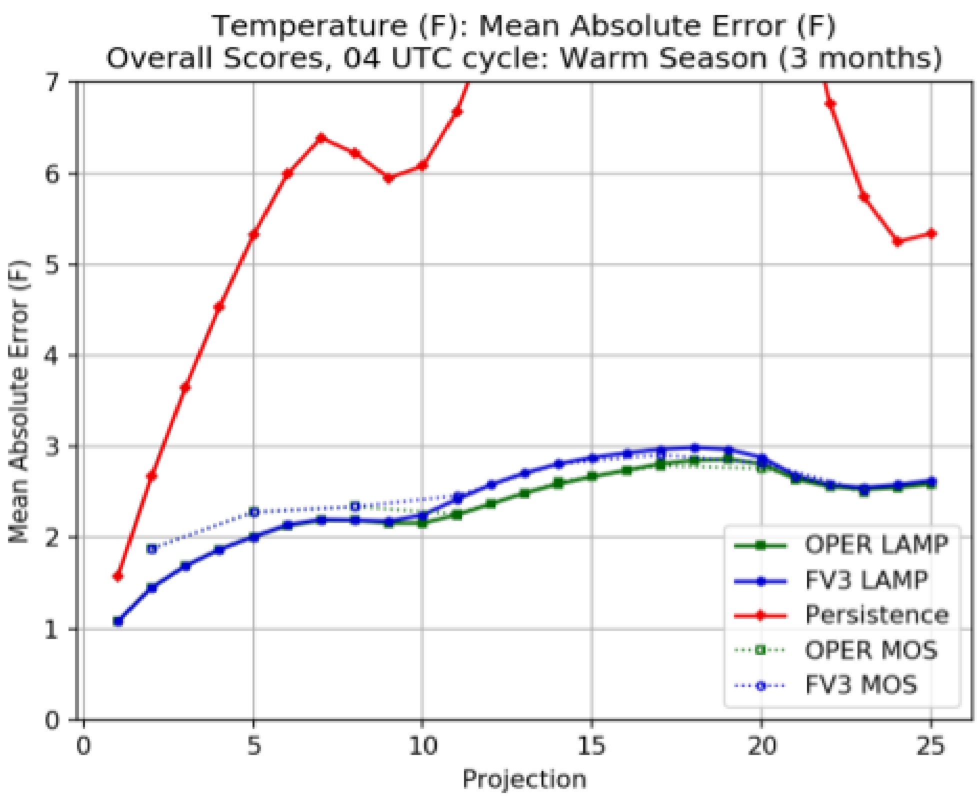

FIG. 2. Mean Absolute Error (in Fahrenheit) in warm season temperature forecasts including OPER-LAMP (filled green squares), OPER-MOS (open green squares), FV3-LAMP (filled blue circles), FV3-MOS (open blue circles) , and Persistence (red diamonds). The graph above is valid for the 0400 UTC verification but the degradation at midranges was also seen in the 1500 and 1900 UTC verifications.

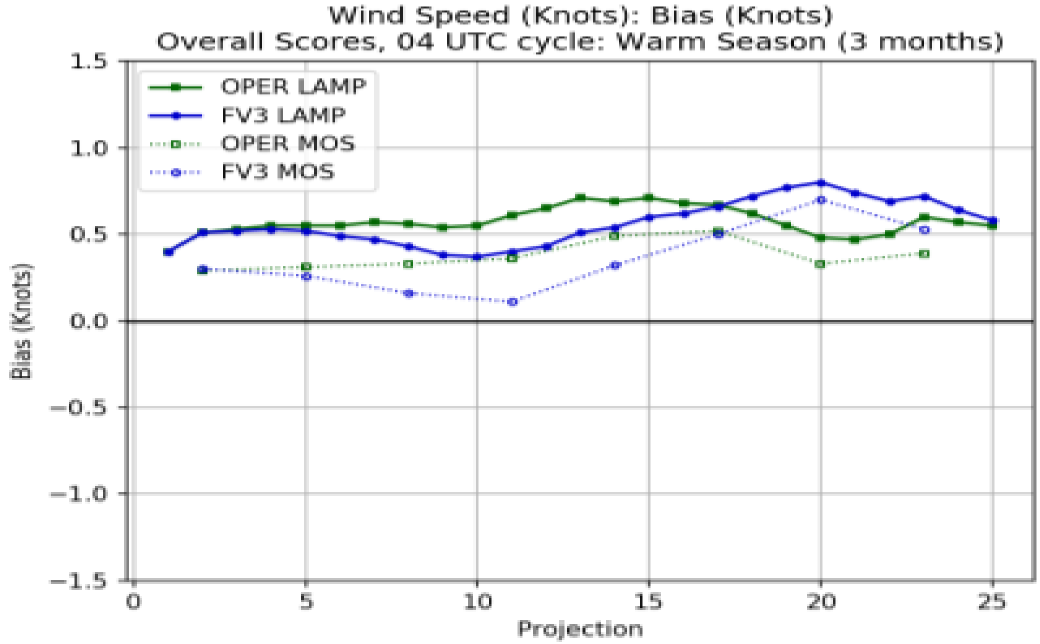

FIG. 3. Bias plots for warm season wind speed forecasts including OPER-LAMP (filled green squares), OPER-MOS (open green squares), FV3-LAMP (filled blue circles), and FV3-MOS (open blue circles). This chart is valid for the 0400 UTC warm season verification but is representative of all other wind speed bias results for the model cycles tested.

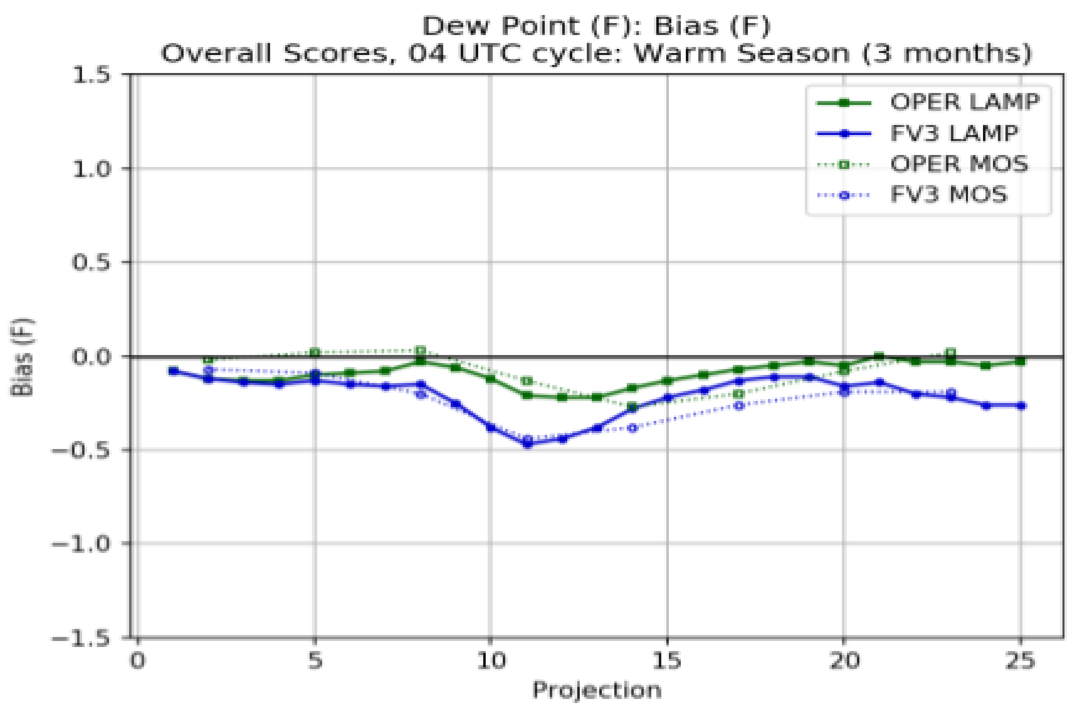

FIG. 4. Same as Fig. 3, but for dew point forecasts.

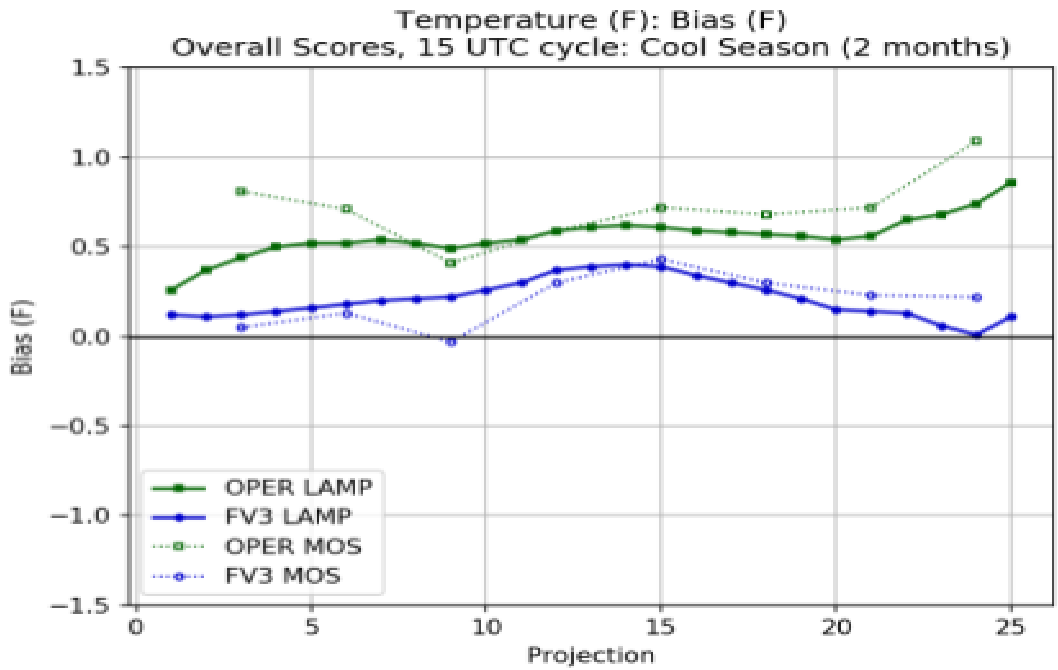

FIG. 5. Same as Fig. 3, but for cool season temperature forecasts.

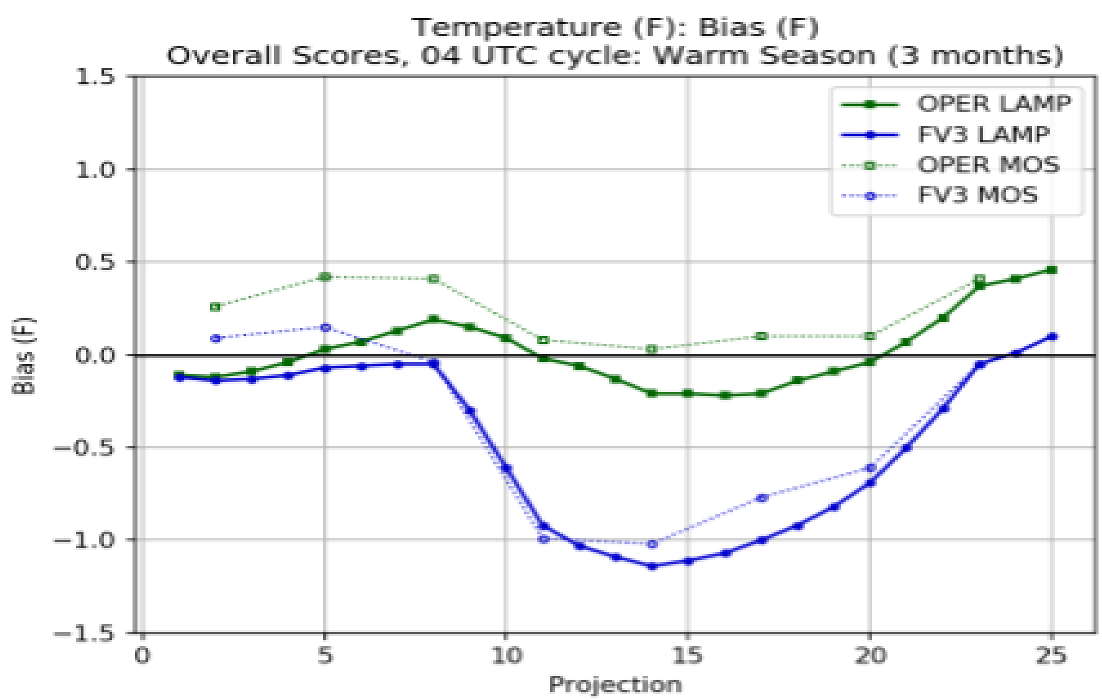

FIG. 6. Same as Fig. 3, but for warm season temperature forecasts.

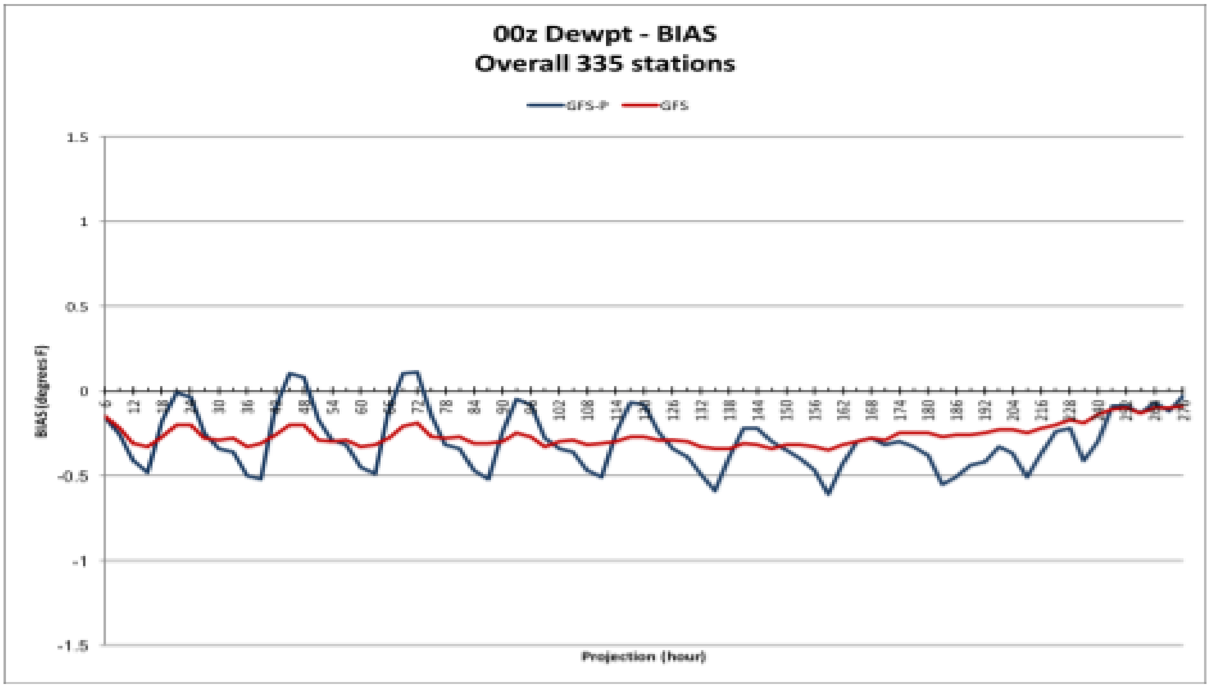

FIG. 7. MOS Verification provided by the MOS verification team at MDL. A clear diurnal variability can be seen in the FV3-MOS (blue line). A similar pattern is exhibited by the OPER-MOS (red line) but not to the same extent.

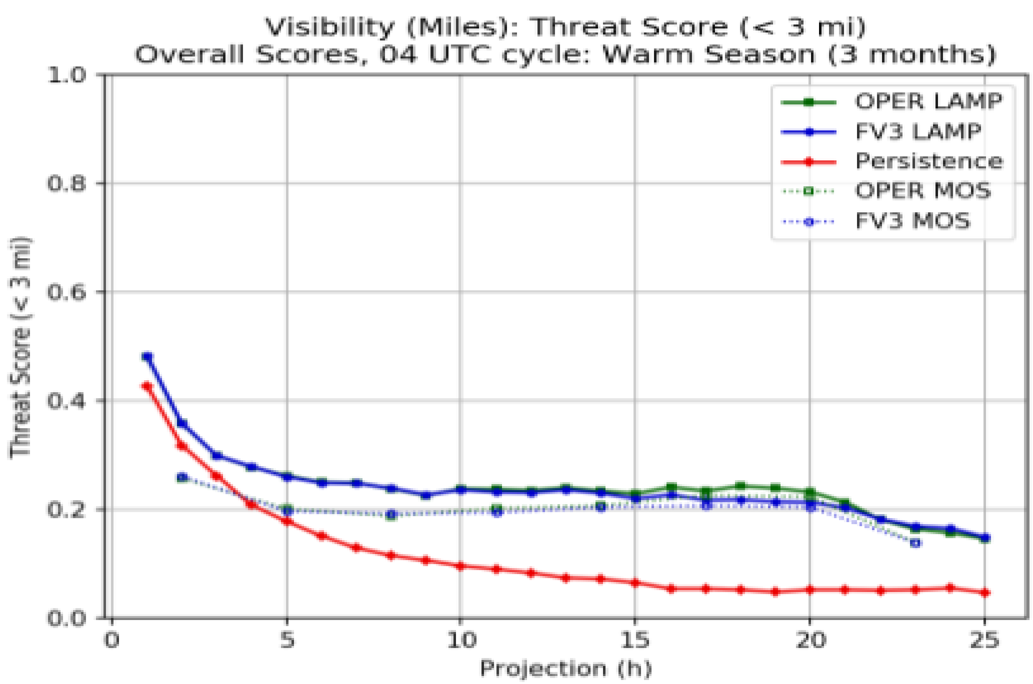

FIG. 8. Threat Score for visibility < 3 miles for OPER-LAMP (filled green squares), OPER-MOS (open green squares), FV3-LAMP (filled blue circles), FV3-MOS (open blue circles), and Persistence (red diamonds). This chart is valid for the 0400 UTC warm season verification but is representative of the visibility < 5 miles result.

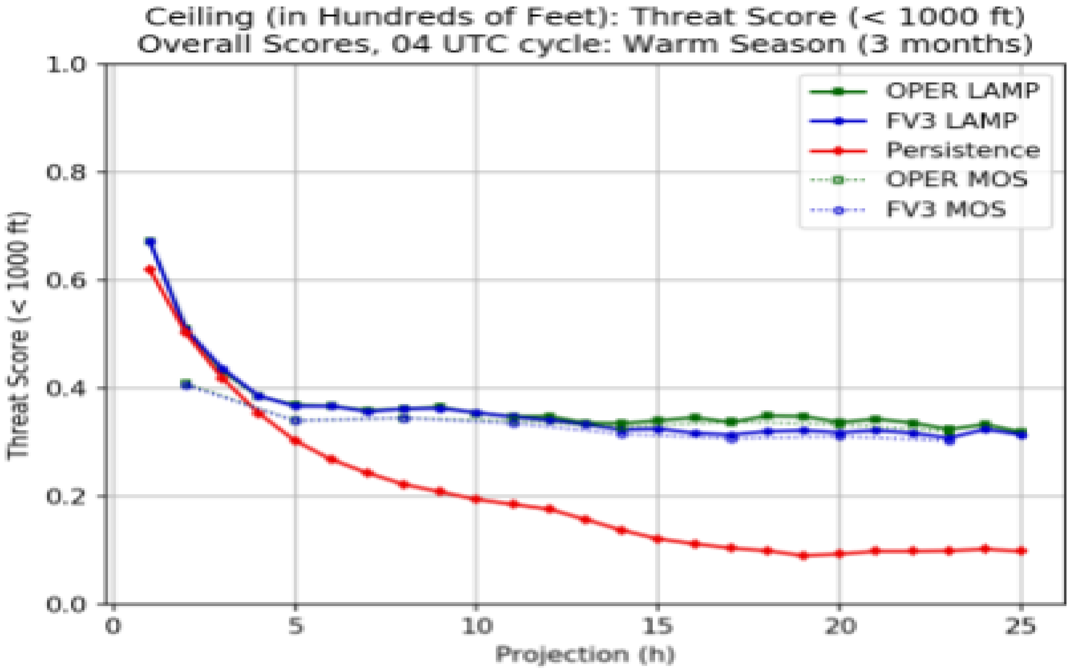

FIG. 9. Same as Fig. 8 but for ceiling heights < 1000 ft. This is also representative of behavior exhibited on the < 500 ft verification during the 0400 and 1500 UTC cycles.

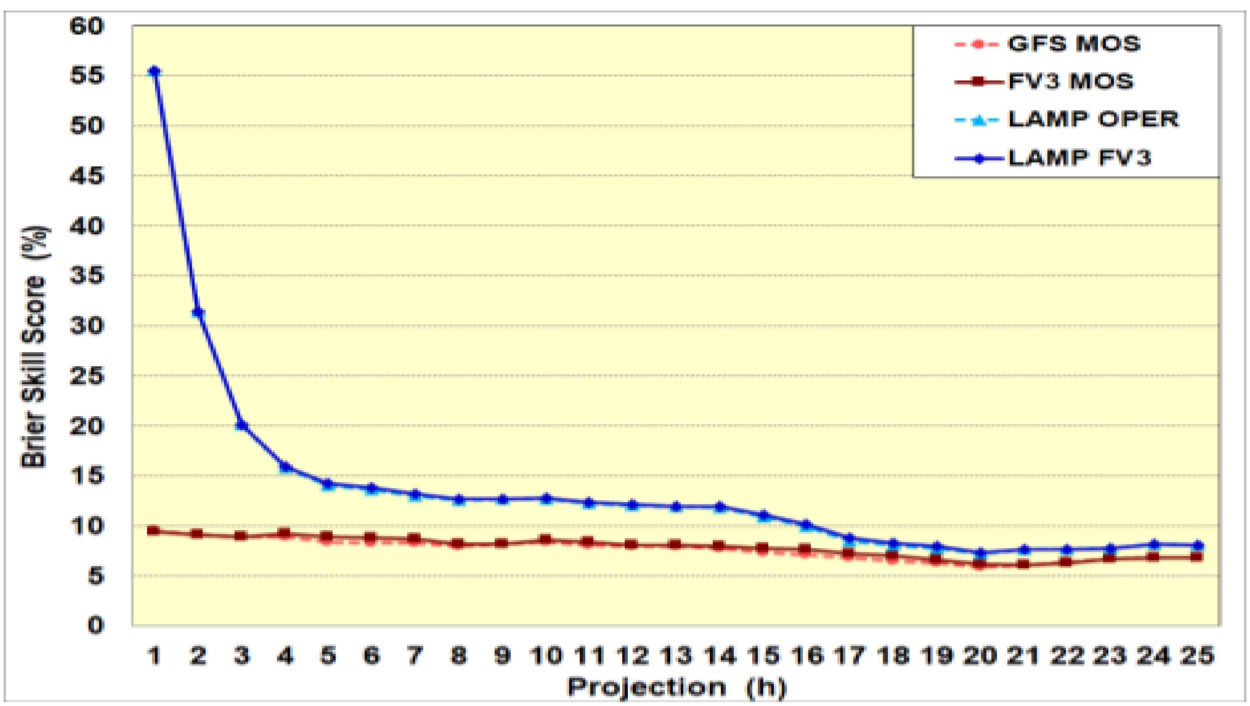

Fig. 10. Brier Skill Score for lightning (all seasons) including OPER-LAMP (light blue triangles), OPER-MOS (light red circles), FV3-LAMP (dark blue diamonds), and FV3-MOS (dark red squares). This chart is also representative of the convection verification results.