5th International Conference on Fog, Fog Collection and Dew
Münster, Germany, 25–30 July 2010
FOGDEW2010-57

Forecast of Low Visibility and Fog From NCEP– Current Status and Efforts

Binbin Zhou[1], Geoff Dimego[1] and Ismail Gultepe[2]

[1] Environmental Modeling Center of NCEP, National Weather Service, NOAA, Camp Springs, Maryland, USA
[2] Cloud Physics and Severe Weather Research Section, MDR, Environmental Canada, Toronto, Ontario, Canada
Binbin.Zhou@noaa.gov

1. Abstract

This paper summarizes the performances of low visibility/fog predictions over North America using the operational National Centers for Environment Prediction (NCEP) forecast models. The evaluation of the results shows that the performances of the low visibility/fog forecasts from these models are still poor in comparison to those of operational precipitation forecasts from the same models. In order to improve the skills of the low visibility/fog prediction, several efforts have been made including implementations of a multi-rule fog detection scheme and a short range ensemble forecast system (SREF). How to apply these techniques in fog prediction is described and evaluated.

2. Introduction

Fog is an important hazardous weather event that affects aviation, transportation and marine traffic. A central guidance from NCEP on its thresholds is being considered and particularly emphasized in National Weather Service (NWS) of NOAA and in NextGen [1], a future Air Traffic Management System of Federal Aviation Administration (FAA), United States. However, fog is still not part of the documents of NCEP central guidance, due to its complexity and limitation of computational resources. Instead, it is only diagnosed locally by forecasters either through subjective visibility's forecast or through other variables from model outputs such as MOS (Model Output Statistics). Nevertheless, effort to add it to NCEP' central guidance is always considered to be important. As a step forward to echo the request from NWS and NextGen, presently, low visibility/fog forecast has been experimentally implemented, tested and validated using NCEP operational models. Currently, the visibility (Vis)-liquid water content relationship of Stoelinga [2] is used in horizontal visibility computation in all of the NCEP models. However, studies have shown that this visibility computation is of high error, particularly in situation of fog when droplet number concentration (N_d) is not considered [3]. Besides the error from Vis computation, bias in modeled liquid water content (LWC) near the surface is also another source of errors. The visibility computation error can be reduced by applying Gultepe's Vis versus LWC and N_d parameteri-zation [3]. Whereas reduction of modeled LWC error is extremely difficult due to low resolution of current operational models, lack of fog physics, low accuracy of LWC at the surface, and model bias itself, etc. To overcome these drawbacks in the operational models, we developed a rule based fog detection scheme recently and it is successfully applied for fog studies. It was used in NCEP's special SREF for 2008 Beijing Olympic Game Project (B08RDP) in China with a 15 km resolution [4]. Now, this scheme has also been applied and tested in the NCEP's SREF with 32km resolution over North America domain.

The objective of this paper is to evaluate (1) performance of fog prediction using the current Vis-LWC relationship based on 3 NCEP's regional models, and (2) improve the rule-based algorithm and ensemble forecast technique for fog detection.

3. Evaluation method and data

An evaluation of fog prediction over North America is generally difficult due to a lack of direct fog observations and the fact that model-based fog value represents a grid area that cannot be interpolated to the location of the observational sites. Here, Vis analysis obtained from the Aviation Digital Database System (ADDS; [5]) of Aviation Weather Center (AWC in Kansas City) of NCEP were used as truth for fog forecast evaluations. The observational data is from Nov. 1 2009 to Apr. 30 2010, covering 6 months. This time period is chosen because of high occurrence of the fog that is defined as Vis<1 km (WMO Manual, 2002). If Vis≤1 km in a grid point, the ADDS grid point is considered as fogy. By comparing model forecast Vis to observed one in a grid point, If Vis≤1 km in both observation and model points, this is assigned as a "hit", if forecast Vis≤1 km but observed Vis ≥ 1 km, this is assigned as a "false alarm", and forecast Vis ≥ 1 km but observed Vis≤1 km, the result is assigned as a" missed alarm". Using these statistical classifications,, forecast scores such as bias, probability of detection (POD), false alarm ratio (FAR), missing rate (MR), and equitable threat score (ETS) were derived. These scores are used to evaluate both single model (deterministic) forecast and ensemble probabilistic forecast performances [4].

3. Performance of current models

In this section, first, evaluation results are given for low visibility/fog forecasts from each of the three regional models. These models include a 12 km resolution North American Mesoscale Model (NAM-12 km or NMM-12 km), a 13 km Rapid Updated Cycle Model (RUC-13km) Model, and a 32 km Nonhydrostatic Mesoscale Model (NMM-13km). The NAM is used to provide regular weather guidance for NWS, the RUC is used as aviation weather guidance, and the NMM is a NCEP version of WRF model based on which other systems, such as NAM, are built.

The performance scores for all three models Vis forecasts are illustrated in Fig. 1, from which the performance scores for fog (Vis<1 km) can be examined. It can be seen that the general performances worsen as visibility threshold decreases. For the Vis threshold of fog, the POD is about 25% for RUC-13 km, 10% for NAM-12 km and only 5% for NMM-32 km. Since NAM is based on NMM-WRF regional model, it can be expected that coarse resolution model NMM-32 km has a lower hit rate (POD) than that of higher resolution (12 km) of the same model in fog prediction.

Another feature shown in Fig. 1 is that the POD for dense fog (Vis<0.5 km) is lower than that of medium fog intensity (Vis>0.5 km but <1 km). In other words, dense fog events are more difficult to detect by these operational models in fog prediction. Fig. 1b shows significant high biases for fog predictions by all 3 models (where bias ~1 means no bias). A positive bias implies an over-prediction or a false alarm of fog forecast. For shallow fog (Vis<1 km), the highest bias is 3 (or 300%) for RUC-13 km. The bias for dense fog (visibility <0.5 km) prediction is even higher. Such high positive biases for all models indicate that low visibility or fog from all NCEP regional models are highly overpredicted. The very low POD with very high bias leads to very poor performances indicated by ETS (Fig. 1c), where the ETS values for all 3 models are less than 5%. For dense fog, their ETS values become much smaller. To compare the ETS values for fog prediction to that for precipitation prediction, the average precipitation forecast ETS (~35%) from current NCEP regional models is also depicted in Fig. 1c, showing that the ETS for fog prediction is much lower than that for precipitation prediction. This means that in order to catch up the performance of precipitation forecast at NCEP we have to devote tremendous efforts to improve fog forecast. Very low POD is shown in Fig. 1a and a high bias shown in Fig. 1b and that implies the current low Vis or fog prediction from these models missed most of the fog events over North America and it overpredicted (false alarm) the fog events. To explain this point, results from an east coast regional fog event occurred on Nov.16, 2009 are shown in Fig. 2:

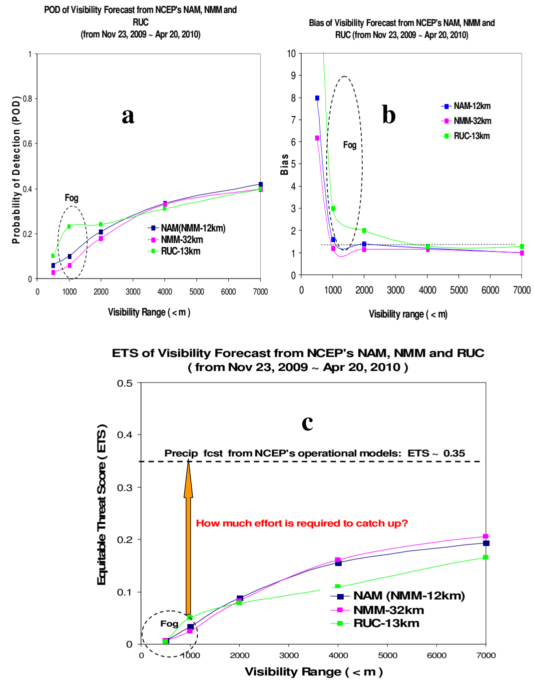

Fig. 1: The tests for Vis: POD (a), Bias (b) and ETTS (c) for each of the 3 forecast models.

Since the visibility computations in the three models are uniquely from fog LWC given at the surface, the green colors (dark, regular and light green) indicate the LWC amount (or intensity), and locations of the fog. Comparing the observed fog location and its intensity (Fig.2a) with the fog Vis obtained from NAM (Fig 2b), one can easily check that the NAM forecast missed most of the fog events in Virginia and North Carolina although it captured some of fog locations in Maryland and Delaware. But the results issued false alarms for half of Pennsylvania and New York states, and most of other northeast states as well as some regions of Canada. Fig. 2c shows that NMM-32 km forecast almost missed all of fog event locations in east coast. The differences in RUC-13 km forecasts for this case can be seen by comparing Figs. 2a and 2d. This run also missed most of locations in Maryland, Virginia, North Carolina and South Carolina, and overpredicted the fog in Pennsylvania and New York states as well as over most of the other northeast states and some regions of Canada, similar to NAM runs. This case, again, indicates much worse performance based on the lower resolution NMM-32 km model than that of the higher resolution NMM-12 km model (or NAM-12 km). Furthermore, the "large-false-alarm" feature of current forecasts of low visibility or fog from these models are also estimated. This feature reminds us that incorrectly predicted location and amount of grid-scale fog LWC at surface may bring some difficulties to applying more precise visibility formulations in the current operational models.

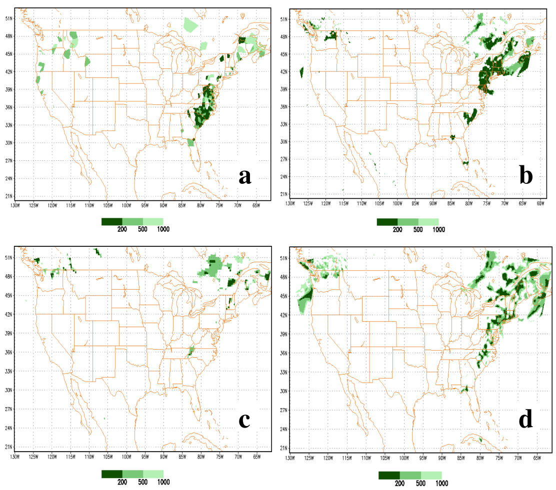

Fig. 2: Nov. 16, 2009 fog visibility observations from ADDS at 1200 UTC in east coast (a) and their forecasts from NAM (NMM-12 km) (b), NMM-32 km (c) and RUC-13 km (d) at the same time. Dark green is for visibility < 0.2 km, regular green for <0.5 km and light green for <1.0 km.

4. Model improvements

Three efforts have been dedicated to improving the performance of experimental central forecast for fog at NCEP recently. The first is testing the "multi-rule" fog diagnostic scheme in the three models, the second is conducting ensemble fog prediction in NCEP's 21-member Short Range Ensemble Forecast System (SREF), and the third is a combination of the multi-rule fog detection and the ensemble technique.

The multi-rule fog diagnostic method has been extensively evaluated in B08RDP in China, showing that the fog prediction ETS from two 15 km resolution WRF models (NCEP and NCAR WRF) was significantly raised from 0.063 with Stoelinga method to 0.192 with the multi-rule diagnosis. The multi-rule fog detection contains 3 sub-rules: (i) LWC rule, which detects fog by modeled surface LWC <0.015 g/kg, (equivalent to visibility <1 km), (ii) cloud rule, which detects fog by modeled cloud base <50 m and cloud top <0.4 km at same time, and (iii) RH-wind rule, which detects fog by RH>90% at 2 m and 10-wind speed < 2 m/s at same time. In B08RDP, the evaluation shows that the RH-wind rule has at least 50% contribution to the total ETS, while the LWC and cloud rules has 30 and 20 % contributions, respectively. This implies that radiation fog is the most frequent fog type in China since RH and calm air are two critical conditions for radiation fog. Without the RH-wind rule, the models would miss at least 50% of fog events. Recently the multi-rule fog diagnosis has also been tested in NAM, NMM-32 km and RUC-13 km over North America. The LWC and cloud rules are kept same but RH-wind rule was retuned as RH>95% and wind speed <1 m/s to better detect radiation fog over North America.

The second effort is conducting the ensemble fog forecast from the NCEP-SREF system. The ensemble forecast technique was suggested and developed recently to deal with model uncertainties and errors in numerical weather prediction (NWP) [6] and has shown many advantages over single model forecast not only for regular weather conditions but also for fog [4][7]. For example, an ensemble forecast provides not only more precise forecast but also the forecast error range and its confidence in advance. Furthermore, with an ensemble probability forecast of a specific weather, users have more freedom to determine or make decision on their own cost/loss economic benefits. Computation of a fog occurrence probability at a grid point from the NCEP SREF is relatively simple: first step is counting how many members predict fog at this grid point, and then dividing the count by total ensemble members to get the fog probability at this grid. After the fog probabilities at all of grid points are calculated, the fog probability distribution over entire domain can be obtained. Obviously, such a fog probability distribution is given in a grid-scale since the computation is grid-wide. The ensemble fog forecast was also applied and evaluated in B08RDP with a 10-member SREF and showed further improvements in forecast performance in addition to the multi-rule application [4].

To use the traditional measures in evaluation of an ensemble forecast, the SREF fog probabilistic forecast needs to be transferred to a deterministic forecast. To do this, a probability threshold, e.g. 50%, can be selected. If the fog probability at a grid point exceeds 50%, or half of ensemble members predict fog, the SREF predicts fog at this point. Different users may select different thresholds. The predictions with each threshold should be evaluated. In the analysis, 10 probability thresholds, from 10% to 100% at 10% bins, were evaluated. Then, the evaluated scores against the 10 probability thresholds are shown in Fig.3.

From Fig. 3, we can see that for different ensemble probability thresholds, the SREF ensemble forecast for fog visibility has different performance values. For a smaller probability threshold, the ensemble forecast gives a higher POD with a large bias. To decrease the bias, a larger probability threshold should be used. In this case; however, the ensemble forecast POD decreases significantly. Therefore, how to select an appropriate ensemble probability threshold in fog forecast is a trade-off. Usually, the threshold is selected as an intermediate value; not very small and large, but around 30~50%, where the ensemble forecast usually has a best performance [4].

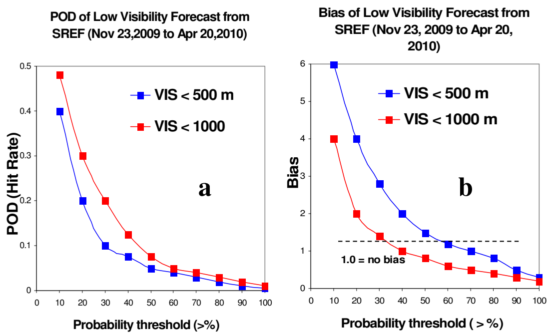

Fig. 3: SREF's fog prediction POD (a) and Bias (b) under different ensemble probability thresholds as in x-axis

The third step is to combine the first and the second efforts. This work is very straightforward and has been tested in ensemble models in the NCEP SREF. The performance of the first, second and the third efforts in comparison to the predictions without these efforts can be examined in Fig. 4. The improvements in the multi-rule effort can be expected by comparing POD or ETS between with multi-rule and without multi-rule (i.e. LWC-only) for

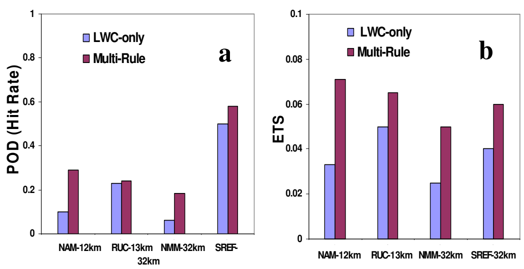

Fig. 4: POD (a) and ETS (b) of fog prediction from NAM (NMM-12 km), RUC-13 km, NMM-32 km and SREF-3 km with multi-rule fog detection and without multi-rule detection scheme (LWC-only).

detection from NAM (or NMM-12 km), NMM-32 km, and RUC-13 km. Although these improvements are not significant as those shown in B08RDP, the improvements are still obvious for NAM and NMM-32 km models, and results about 50% increases in both POD and ETS. The significant performance improvements from the ensemble technique can be observed by comparing NMM-32 km and SREF-32 km (since NMM-32 km is one of the reference models in the SREF system). The further improvements of multi-rule and ensemble combination can be seen from comparing SREF-32 km model with multi-rule and without multi-rule (i.e. LWC only). There is additional 20-30% increase in POD or ETS by using multi-rule fog detection scheme in the SREF system.

5. Conclusion

Through verifications of low visibility (<1 km) operational forecasts from NCEP's three regional models, NAM-12 km, RUC-13 km and WRF-NMM-32 km, over North America against ADDS data over 6 months time period, the performance of fog predictability were evaluated. This shows that the performances of fog predictions from the models still need improvements. The reason may be that these models are unable to predict correct locations and intensities of fog events due probably to too-coarse model resolutions, missing appropriate fog physics in the models, and model numerical bias.

Three efforts have been made to improve fog prediction at NCEP, including an application of a "multi-rule" fog detection scheme, an application of ensemble technique in SREF, and a combination of these two applications. The verifications with the same ADDS data over North America have shown obvious increases in prediction performances, suggesting that these techniques can be further applied in future North American High-Resolution Rapid Refresh Ensemble (HRRRE) Forecast System at NCEP requested by the NextGen of FAA.

Acknowledgement

This work was fully funded by FAA. Our special appreciation is given to Aviation Weather Center for providing the ADDS data to support this study.

Reference

[1]Souders, C. G., et al., 2010: NextGen weather requirements: an update. Preprint, 14th *Conf. on Aviation, Range, and Aerospace Meteorology*, Atlanta, GA, Amer. Meteor. Soc.

[2]Stoelinga, T. G. and T. T. Warner, 1999: Nonhtdrostattic, mesobeta-scale model simulations of cloud ceiling and visibility for an east coast winter precipitation event. *J. Apply. Meteor.*, 38, 385-404.

[3]Gultepe, I., M. D. Müller, and Z. Boybeyi, 2006: A new visibility parameterization for warm fog applications in numerical weather prediction models. *J. Appl. Meteor.* Clim., 45, 1469-1480.

[4]Zhou, B. and J. Du, 2010: Fog prediction from a multimodel mesoscale ensemble prediction system. *Wea Forecasting*, 25, 303-322.

[5]http://aviationweather.gov/adds

[6]Toth, Z. and E. Kalnay, 1993: Ensemble forecasting in NMC: The generationof perturbations, *Bull. Amer. Meteor. Soc.*, 74, 2317-2330.

[7]Roquelaure, D. S. and T. Bergot, 2008: a local ensemble prediction system (L-EPS) for fog and low clouds: Construction, Bayesian model averaging calibration, and validation. *J. Appl. Meteor. Clim.,* 47, 3072-30.