Fog Prediction from a Multimodel Mesoscale Ensemble Prediction System

Jun DU and Binbin ZHOU
National Centers for Environmental Prediction/NOAA

Abstract

A new multi-variable based diagnostic fog-forecasting method has been developed at NCEP. The selection of these variables, their thresholds and influence on fog forecasting are discussed. With the inclusion of the algorithm in the model post-processor, the fog forecast can now be provided centrally as direct NWP model guidance. The method can be easily adapted to other NWP models. Currently, we lack knowledge of how well fog forecasts based on operational NWP models perform. To verify the new method and assess fog forecast skill as well as to account for forecast uncertainty, this fog-forecasting algorithm is applied to a multi-model based Mesoscale Ensemble Prediction System (MEPS). This MEPS consists of 10 members using two regional models (NCEP WRF NMM and NCAR WRF ARW) with 15km horizontal resolution. Each model has five members (1 control and 4 perturbed members) using the breeding technique to perturb the initial conditions and was run once per day out to 36-hours over eastern China for seven months (Feb. – Sept. 2008). Both deterministic and probabilistic forecasts were produced based on individual member, a one-model ensemble and two-model ensembles. A case study and statistical verification, using both deterministic and probabilistic measuring scores, were performed against fog observations from 13 cities in eastern China. The verification was focused on the 12- and 36-hour forecasts.

Applying the various approaches including the new fog detection scheme, ensemble technique, multi-model approach and the increase in ensemble size, fog forecast accuracy was steadily improved in each of the approaches and dramatically: from a basically no-skill-at-all (ETS=0.063) to a skill level equivalent to that of warm-season precipitation forecasts of the current NWP models (0.334). Specifically, (1) The multi-variable based fog diagnostic method has a much higher detection capability than the LWC-only based approach. Reasons why the multi-variable approach works better than the LWC-only method were also illustrated. (2) The ensemble-based forecasts are, in general, superior to a single control forecast measured both deterministically and probabilistically. The case study also demonstrates that the ensemble approach could provide more societal value than a single forecast to end-users especially for low-probability significant events like fog. Deterministically, a forecast close to the ensemble median is particularly helpful. (3) The reliability of probabilistic forecasts can be effectively improved by using a multi-model ensemble instead of a single-model ensemble. For small ensemble such as the one in this study, the increase in ensemble size is also important in improving probabilistic forecasts although this effect is expected to decrease with the increase in ensemble size.

Key word: fog prediction, mutimodel ensemble, probabilistic forecast, evaluation

1. Introduction

Fog is frequently blamed for traffic disasters and bad air quality in poor visibility weather and has been extensively studied for more than a century. However, progress in the operational numerical weather prediction (NWP) centers has been slow due to the complexity of predicting fog and limited computing resources. For now, fog is still not a direct model guidance product produced by NWP centers but is diagnosed by local forecasters based either on statistical methods like model output statistics or on indirect model output variables. Therefore, a centrally produced and skillful NWP fog guidance is desired.

Contact info: Jun.Du@noaa.gov

At the same time, most of these fog forecasting efforts were deterministic in nature and didn't take forecast uncertainty into consideration. Given the intrinsic uncertainty of model forecasts and the fact that fog forecasting is extremely sensitive to the initial conditions and the physics schemes used in a prediction system, it is strongly desirable to have fog prediction be part of an ensemble framework. This work is one of the pioneering attempts in the trend of this new probabilistic forecast requirement from such as NextGen air traffic control system.

During the Beijing 2008 Summer Olympic Games, a subcomponent of the NCEP SREF system was reconfigured to support daily weather forecasts in China for the event as part of the WMO/WWRP Research Demonstration Project. Taking advantage of this project, a fog prediction scheme was quantitatively and objectively verified using this mesoscale ensemble with three goals. The first goal is to examine the effectiveness of a new diagnostic fog-forecasting method compared to a commonly used method; the second goal is to examine the possible forecast skill level of current operational NWP models in predicting fog; and the last goal is to compare the performances of a single-model based ensemble and multi-model based ensembles as well as to examine the impact of ensemble size on probabilistic forecasts. To the best of our knowledge, this is the first time to apply a sophisticated ensemble technique to state-of-the-art operational NWP model to centrally predict and systematically evaluate this important but difficult and complex low-probability phenomenon fog.

2. A New Diagnostic Fog-Detection Scheme

Liquid Water Content (LWC) at the model's lowest level was commonly used to represent fog in previous studies. However, the LWC-only approach doesn't work well in an operational NWP model for the following two reasons: too coarse spatial resolution to properly resolve important physics in fog near the surface, and physics schemes or parameterizations not tailored for near-ground fog but for precipitation or clouds at higher levels. As a result, the LWC from NWP models is usually not reliable enough to represent fog and tends to seriously under-forecast fog in many cases. To better detect fog, other variables besides LWC should be considered. Considering that fog has different types with different formation mechanisms, e.g., some build from stratus-subsidence, some from advection, and some from radiation cooling near the ground, a new multi-variable based diagnostic fog-detection scheme is proposed as follows:

LWC at model lowest level ≥ 0.015 g/kg, OR (1a)
Cloud Top ≤ 400 m AND Cloud Base ≤ 50 m, OR (1b)
10-Wind Speed ≤ 2 m AND 2m-RH ≥ 90 % (1c)

The LWC rule in (1a) came from the definition of fog visibility range: LWC ≥ 0.015 g/kg is equivalent to visibility ≤ 1000 m. The cloud top threshold in (1b) follows the observations that the depth of most fogs on land is about 100 ~ 200 m and some marine fogs or advection fogs are deeper, but rarely exceed 400 m. The cloud base threshold in (1b) reflects the height of a model's lowest level. The cloud rule (1b) works quite well for large-scale deep fog but not shallow near-ground fog. The rule of large RH and weak surface wind (1c) describes shallow near ground and radiation fog. Choosing general and centralized thresholds for surface wind and *RH* over large domains in a model is more difficult than for the LWC and cloud rules because (1) ground fog is more local and (2) different models might have different *RH* and wind biases. In other words, the thresholds of rule (1c) needs some kind of tuning for a particular model. However, this study shows that such a tuning is not that hard and the 90% and 2m/s thresholds could be used as a good starting point. For the tuning details, readers are referred to the Table 2 of Zhou and Du (2010). Zhou and Du (2010) also discussed the relative contributions from the different rules of Eq. (1) in detecting fog.

3. Ensemble System and Data

In 2008, a subcomponent of the NCEP Short-Range Ensemble Forecast (SREF) system was reconfigured to the China region (Fig. 1) and ran once per day from January 29 to September 7, 2008. This subsystem is a multi-model mesoscale ensemble prediction system designed to include physics diversity, which consists of 10 members using two regional models (NCEP NMM and NCAR ARW). Each model has five members, 1 control and 4 perturbed, to also address uncertainty in the initial conditions (ICs). This system ran once per day with a forecast length of 36 hours initiating at 12:00 UTC or 20:00 BT (Beijing Time). The control IC came from the NCEP Global Data Assimilation System. IC perturbations were created using the breeding method and the lateral boundary condition perturbations were provided by the NCEP Global Ensemble System. The horizontal resolution of both models is 15km. The vertical resolutions are 52 sigma-levels for NMM and 51 sigma-levels for ARW. The lowest vertical resolutions (above the surface) for both models are about 50m. Besides the difference in the dynamic cores, the following physics are also different between the two models: convection, planetary boundary layer, surface boundary layer, long wave and short wave radiation. But the cloud microphysics and land surface schemes are the same for both models.

Due to a current lack of fog analysis data, it's impossible to verify fog forecasts grid point by grid point over the entire domain. Instead, 13 big metropolitan areas are chosen for verification and tuning in this study (Fig. 1). The verification data for this study were available from daily fog reports issued by local weather services or airports in 13 cities over eastern China from February to September of 2008. Since the observational data were reported only for morning fogs, the verification had to be done on both the 12-h and 36-h forecasts of a particular cycle which correspond, respectively, to 8:00am BT on the first and the second days after the model initiation time (12:00 UTC or 20:00 BT). The foggy days in the 13 cities during the seven months are summarized in Table 1 and include both dense fog (visibility < 500m) and light fog (visibility between 500 to 1000m) events. There were a total of 242 foggy days during the verification period. Fog is a local weather phenomenon which is strongly influenced by factors such as terrain, local flows and local surface boundary layer conditions and, therefore, may well be a sub-grid-scale event on many occasions. As a result of this, a model with coarser horizontal resolution may not capture all fog events well. Thus, a robust assessment of the systematic performance of fog forecasts can only be gotten from verifications over a long period of time or over a large number of cases. In this study, a total of about 5,460 forecasts [7 months x 30days x 2 forecasts per day (f12h and f36h) x 13 cities] were used in the verification to make sure that the verification results are representative. Considering that each forecast was actually predicted by 10 ensemble members, the total number of forecasts reached 54,600 in this study.

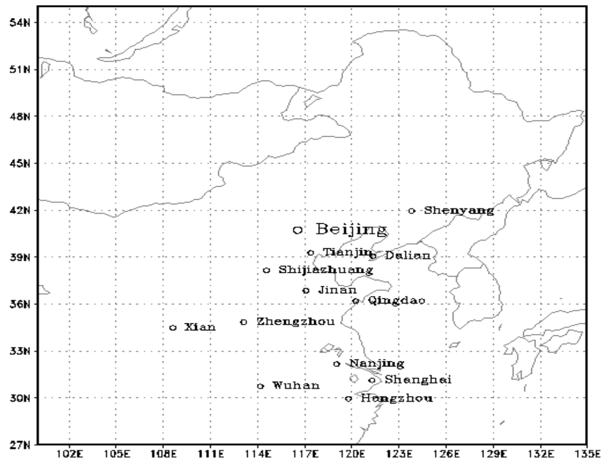


Fig 1. Ensemble forecast model domain and the locations of the 13 fog verification cities.

Table 1: Number of fog observed days for the 13 cities in eastern China from February to August 2008.

City	Feb	Mar	Apr	May	Jun	Jul	Aug	Total
1. Shenyang (SY)	0	2	1	2	4	5	5	19
2. Beijing (BJ)	0	1	1	2	4	5	4	17
3. Tianjin (TJ)	2	2	1	4	1	12	4	26
4. Shijiazhuang (SJZ)	0	3	3	4	2	4	3	19
5. Zhengzhou (ZZ)	1	1	4	0	0	6	4	18
6. Dalian (DL)	0	2	3	4	4	7	3	23
7. Qingdao (QD)	1	1	3	5	9	8	2	29
8. Jinan (JN)	0	1	0	1	2	2	2	8
9. Xian (XA)	1	2	2	1	2	1	0	9
10. Nanjing (NJ)	4	2	3	0	5	3	2	19
11. Shanghai (SH)	5	1	5	3	4	1	3	22
12. Hangzhou (HZ)	5	2	4	3	4	2	2	22
13. Wuhan (WH)	3	2	4	1	1	1	1	13
total	22	22	34	30	42	57	35	**242**

4. Deterministic Verification

Figure 2 shows all the *ETS*es from the various approaches used in this study including the new fog detection method, ensemble technique, multi-model approach and the increase in ensemble size. We can see that a steady improvement was made through each of those steps (1 through 5), with two big jumps, one associated with the use of the new multivariable fog detection method (a 205% increase in ETS) and the other associated with the combing of the two single-model ensembles (a mixed contribution of the multi-model approach and the ensemble size increase, a 26.5% increase). The overall improvement was impressive and dramatic: from a basically no-skill-at-all (ETS=0.063 at System 1) to a skill level equivalent to that of warm-season precipitation forecasts of the current NWP models (0.334 at System 5).

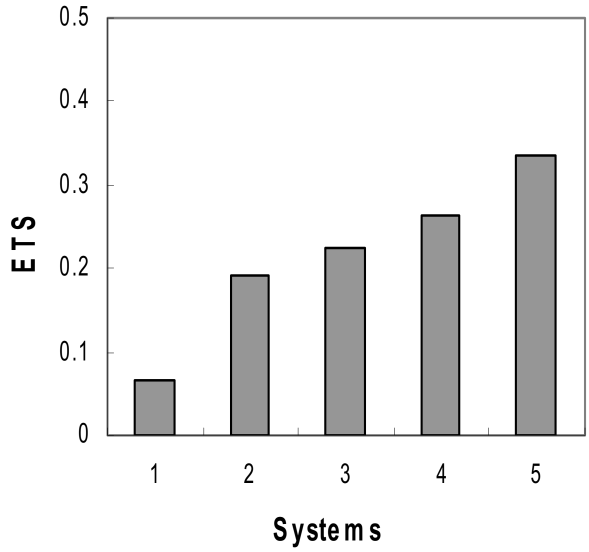


Fig 2. ETSes (averaged of the ARW and NMM over the 7-month period at 12- and 36-h forecast lengths) from the various forecast systems: 1) the single control runs based on the LWC-only approach (ETS=0.063), 2) the single control runs but based on the new multivariable fog diagnosis (0.192; a 205% improvement over the previous step), 3) the 40% probability forecasts based on the 5-member single model ensembles (0.225; 17.2%), 4) the 40% probability forecast based on the 5-member multimodel NMM-ARW ensemble (0.264;17.3%), and 5) the 40% probability forecast based on the 10-member multimodel SREF-B08RDP ensemble (0.334;26.5%).

Figure 3 demonstrates a case of a major fog episode. Fog was almost completely missed by a single forecast (NMM control run as an example) with the LWC-only approach (Fig. 3b) comparing to the observation (Fig. 3a), while it is much improved with the new multivariable fog scheme (Fig. 3c). Furthermore, the 10-member NMM/ARW multimodel ensemble (Fig. 3d) obviously depicts the most complete picture of this episode by considering the uncertainties from both initial condition and model physics.

5. Probabilistic Verification

Figure 4a shows the Brier Skill Score (*BSS*) over each month for both the NMM-ensemble based and ARW-ensemble based probabilistic forecasts. Clearly, both ensembles show skill over the single control forecasts for the entire verification period from February to August. The mean *BSS* averaged over all seven months is shown in Fig. 4b. The improvement from single-model ensembles to multi-model ensemble is obvious. This can be further confirmed from the reliability diagram (Fig. 5). By decomposing Brier Score into *reliability, resolution and uncertainty* three terms, it's found that the major improvement from multimodel approach is in *reliability* (see Table 4 of the Zhou and Du 2010). Increase in ensemble membership also plays an important role when ensemble size is small such as in this study (from 5 to 10 members).

6. Summary

A new multivariable based diagnostic fog-forecasting method has been developed at NCEP. Since all the five base variables used for the diagnosis are direct outputs from a model, this fog diagnostic algorithm can be included as part of the model post-processor and, therefore, the fog forecast can now be provided conveniently and centrally as a direct NWP model guidance.

Applying the various approaches including the new fog detection scheme, ensemble technique, multi-model approach and the increase in ensemble size, the improvement in fog forecast accuracy was steady in each of the approaches and dramatic: from a basically no-skill-at-all (Equitable Threat Score=0.063) to a skill level equivalent to that of warm-season precipitation forecasts of the current state-of-the-art NWP models (0.334). In specific,

(1) The multivariable based fog diagnostic method has a much higher detection capability than LWC-only based approach (a commonly used method in current practice). The latter has a very low detection rate and tends to miss almost 90% of fog events; the former can greatly improve the fog detection rate;

(2) The ensemble-based forecasts are, in general, superior to a single forecast measured both deterministically and probabilistically. The case study also demonstrates that ensemble approach could provide more societal value than a single forecast to end-users especially for low-probability significant events like fog. Deterministically, a forecast close to the ensemble median (50% probability) is particularly helpful;

(3) The reliability of probabilistic forecasts can be effectively improved by using a multi-model ensemble instead of a single-model ensemble. For small-size ensemble such as the one in this study, the increase in ensemble size is also important in improving probabilistic forecasts although such an importance is expected to decrease with the increase in ensemble size (Du et. al, 1997).

References:

Zhou, B., and J. Du, 2010: The same title, *Weather and Forecasting*, **25**, 303-322.

Du, J., S.L. Mullen, and F. Sanders, 1997: Short-range ensemble forecasting of quantitative precipitation. *Mon. Wea. Rev.*, **125**, 2427-2459.

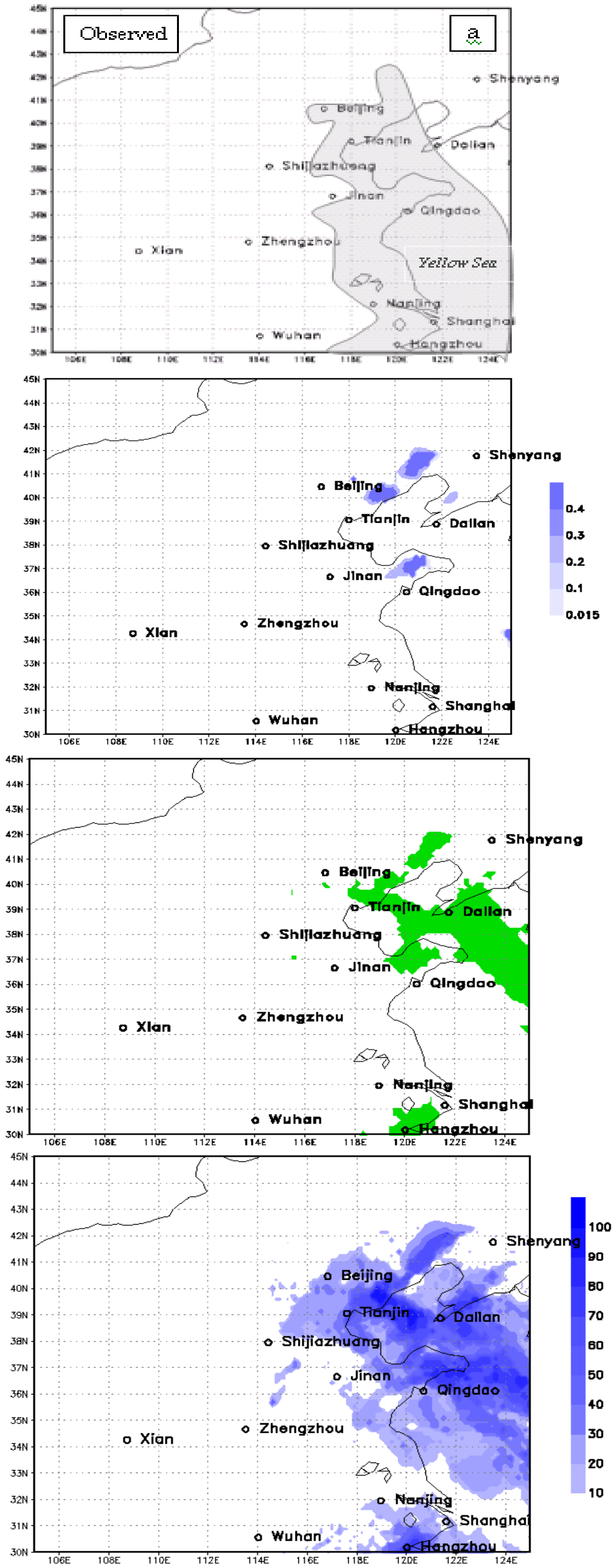
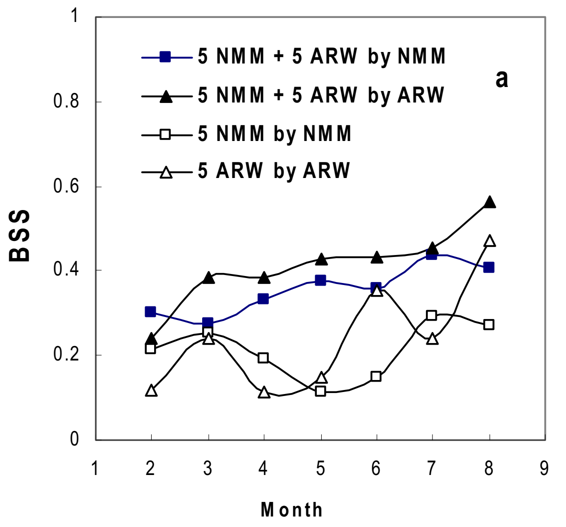


Fig. 3. (a) Observed fog episode at 0500BT 7 April 2008; (b) 9-h fog forecast based on LWC-only approach from NMM control run; (c) 9-h fog forecast based on the new multivariable approach from NMM control run; and (d) 9-h fog forecast(prob>=10%) based on 10-member NMM/ARW multimodel ensemble.



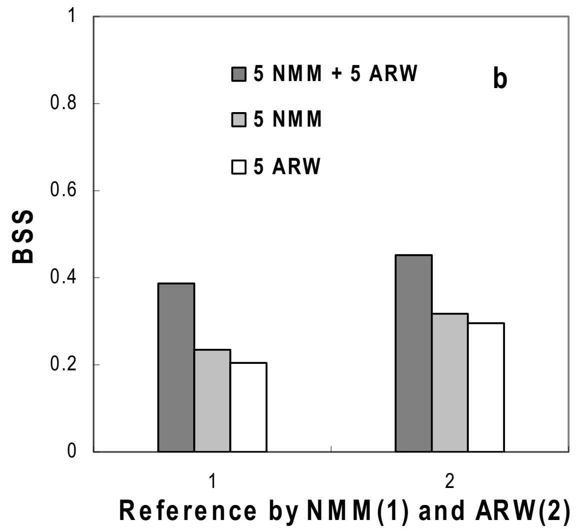


Fig. 4. (a) Monthly BSS scores of probabilistic forecasts based on 5-member NMM (using NMM control as a reference), 5-member ARW ensemble (using ARW control as a reference), and 10-member multimodel ensemble (using both NMM and ARW control as references); (b) BSS averaged over the 7-months for the same three ensembles but with both NMM and ARW control as references.

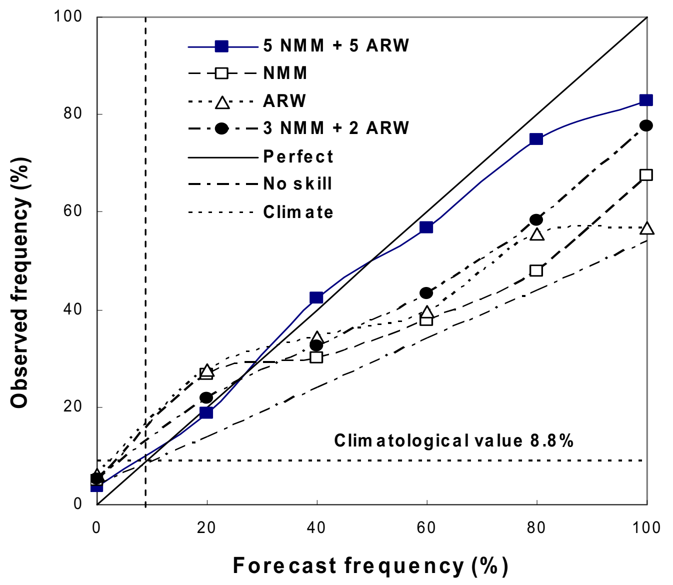


Fig. 5. Reliability diagram of probabilistic forecasts based on 5- member NMM and ARW single model ensembles, 5-member multimodel ensemble, and 10-member multi-model ensemble.