Bias Removal on a Mesoscale Forecast Grid

Richard C. Steed and Clifford F. Mass
University of Washington, Seattle, Washington

1. Introduction

Mesoscale model forecasts can contain significant systematic errors for forecasts of temperature and precipitation. Because these biases are so systematic, statistical methods can be successfully used to correct for them. This work compares several methods for removing biases from a gridded mesoscale forecast. Each of these methods takes the statistical definition of bias to be the average error of a parameter at a specified time of day over a two week period. The methods compared utilize different spatial techniques for applying bias removal over all points on the mesoscale grid.

One example of such biases in mesoscale forecasts is the mean error of temperature. Results from previous work clearly show that MM5 forecasts have a significant diurnal signal in 2 meter air temperature forecast errors. The model consistently predicts a diurnal range of temperatures which is too narrow when compared with observations, as shown in Mass et al, 2002. Daytime temperatures predicted by MM5 are too cool, and nighttime temperatures too warm.

Accounting for exactly this type of error is the goal of the current work. An investigation of temperature errors in the real-time MM5 forecast system run at UW since 1995 (Mass et al, 2003) reveals that not only is there a maxima in the diurnal error at night, but that the nighttime warm bias is greatest during the early winter months. In this paper we examine the results of bias removal methods applied to nighttime temperature errors during the months of October and November.

2. Method

During the winter months of 2003-2004, forecasts from the real-time MM5 forecast system run at UW were collected along with corresponding surface weather observations. These datasets were combined to perform bias removal (using several techniques) on each day's MM5 forecast, and the resulting corrected forecasts were archived for later comparison with the uncorrected MM5 forecast and each other.

MM5 forecast errors at observing locations were computed and archived, and these were used to recompute model temperature biases each day. At each observation location, the errors during the previous two weeks were averaged to approximate the bias at that location. The two week period was chosen to best account for the seasonal change in model errors, while still being a large enough sample to eliminate noise. Also, preliminary studies showed that with a varying length of averaging period, our bias removal system produced the least error with a 2-week period compared with 1, 3, 4, and 6 weeks.

Using the bias at each observation station location, this experiment examines the performance of different methods of how to "spread" the bias removal over the entire forecast grid, including at points distant from the observations.

2a. Terrain-landuse analysis

This method takes the bias at observation locations and performs an analysis over the grid to estimate the spatial field of model forecast bias. For each model grid point, the bias computed at the nearest 5 observation stations which are within 200 m vertical elevation from the model grid point elevation and are of the same basic landuse category as the model grid point are averaged. This is repeated over all model grid points, producing a grid of model bias. The bias grid is then subtracted from the raw MM5 forecast grid to produce a "bias-removed" forecast.

2b. Rawins analysis

In a similar fashion to the terrain-landuse method, a modified Cressman scheme is employed to analyze forecast bias over the entire model grid. The effect of bias at the observation station locations is spread over the model grid using a circular weighting function, which is further constrained to only include the effect of observations on a model grid point's value which are of the same basic landuse category as the model, and also weighted to restrict the application of observations which are at similar elevations to the model. Again, the resulting gridded field of forecast bias is removed from the raw forecast grid.

2c. Domain average

In contrast to the previously described methods, this experiment takes the domain-wide average of bias at all observation locations and subtracts this from the raw forecast, equally at all grid points.

2d. RUC analysis

Using an entirely different method for estimating bias over the model grid, this method assumes the zero-hour RUC analysis (interpolated to the MM5 forecast grid) to represent the true atmospheric state. The gridded model error field is produced by subtracting the RUC analysis from the raw MM5 forecast grid. The grid of bias is simply the 2-week average of the gridded model error. This is the bias field which is subtracted from the raw forecasts to produce the "bias-removed" forecast.

3. Results

After performing bias removal using these different methods over the winter of 2003 -4, observation-point verification was performed on each of the resulting forecasts and on the raw MM5 forecasts. Statistics shown in Table 1. are the average of model errors at all stations within the model domain over all forecasts during the months of October and November. Verified here are 36 hour forecasts from all 0000 UTC initializations, all valid at 1200 UTC. Each of the experiments removed almost all of the bias compared with the raw MM5 forecasts. In addition to domain average statistics, a contingency table was constructed to measure the skill of each method at predicting nighttime temperatures below freezing.

Table 1. Domain-wide statistics for October and November 2003

Experiment name	Raw	Dom-Avg	Ter-Land	Rawins	RUC
Mean error	1.391	0.001	-0.084	-0.228	-0.366
Mean absolute error	2.862	2.683	2.764	2.652	2.780
RMSE	3.752	3.450	3.537	3.379	3.575
Correlation coeff	0.882	0.880	0.874	0.886	0.874

Table 2. Contingency table scores for October and November 2003

Experiment name	Raw	Dom-Avg	Ter-Land	Rawins	RUC
hit rate	0.715	0.811	0.809	0.820	0.821
Percent correct	0.873	0.870	0.866	0.867	0.863
inaccuracy	0.127	0.130	0.134	0.133	0.137
false alarm rate	0.047	0.096	0.102	0.106	0.113
odds ratio	51.017	40.578	37.427	38.466	35.969
bias	0.807	0.978	0.986	1.003	1.018

During the Oct-Nov, 2003 period, there were over 6000 observations within the model forecast domain. Out of these, over 2000 were below freezing, providing a large sample which is neither too prone to freezing nor was it too rare. Figures in Table 2. indicate various performance measures computed from a binary, 2x2 contingency table which uses the criteria of temperatures observed and forecast being less than 273.15 K. From these scores it is evident that each of the bias removal methods is almost equally skillful, and that they do successfully remove the systematic warm bias exhibited by the raw MM5 forecasts. Also of note is the increased hit rate and corresponding increased false alarm rate of each bias removal method compared with the raw MM5.

4. Summary

Several methods for removing model forecast bias have been shown to work, with nearly equal performance. All the methods tried increase skill as shown by the hit rate scores, however all also increase the false alarm rate. This results from the raw MM5 warm bias leading to a smaller likelihood of incorrectly forecasting freezing temperatures, while it more frequently fails to correctly forecast such events. For operational applications which depend on critical decision thresholds such as freezing temperatures, gridded forecast bias removal can be effective at improving forecast utility.

4. References

Mass, Clifford F., David Ovens, Ken Westrick, and Brian A. Colle, 2002: Does Increasing Horizontal Resolution Produce More Skillful Forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407-430

________________, and coauthors, 2003: Regional Environmental Prediction Over the Pacific Northwest. *Bull. Amer. Meteor. Soc.*, **84**, 1353-1366