Geosci. Model Dev. Discuss., 2, C624–C639, 2010
www.geosci-model-dev-discuss.net/2/C624/2010/

Geoscientific
Model Development
Discussions

Interactive comment on "Assessment of bias-adjusted $PM_{2.5}$ air quality forecasts over the continental United States during 2007" *by* D. Kang et al.

D. Kang et al.

kang.daiwen@epa.gov

Received and published: 5 April 2010

GENERAL COMMENTS This paper evaluates a post-processing method that combines recent and new hourly PM2.5 predictions from a source-oriented air-quality modeling system (WRF-NMMCMAQ) with recent hourly PM2.5 measurements from a network of monitoring stations in order to produce improved forecasts at the locations of those monitoring stations (though not elsewhere). A set of model PM2.5 forecasts and measurements from the North American AIRNow meta-network for a one-year period have been used to evaluate this method. The same Kalman Filter Predictor bias-adjustment approach has been used previously for ozone forecasts and was shown to

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

improve forecast skill. However, due to a number of differences between PM2.5 and ozone, there was no reason to expect the same result a priori for PM2.5, but the results presented in this paper suggest that this technique does improve skill for PM2.5 forecasts as well on average. This is a well-written and useful paper. It extends previous work on ozone forecasting and shows that the KF bias-adjustment method can add skill for forecasting of PM2.and could be implemented operationally in conjunction with the near-real-time PM2.5 measurements that are available for much of North America from the AIRNow meta-network. I have made a number of specific comments and suggestions below that I would ask the authors to consider, as I believe that addressing them would strengthen the paper further. My recommendation is to accept this paper for publication in Geoscientific Model Development conditional upon minor revisions.

RESPONSE: We thank the reviewer for the overall positive assessment of the manuscript and for the numerous constructive suggestions.

SPECIFIC COMMENTS 1. It is not mentioned in Section 1 whether the U.S. Air Quality Index is defined in terms of hourly values or 8-hourly values or daily maximum values or other quantities. One reason for raising this point is that the analyses described in the manuscript are restricted to observed and forecasted daily mean PM2.5 concentrations even though the measurement data set used had hourly time resolution and both raw and KF bias adjusted model predictions were also available every hour. It seems that an opportunity to look at diurnal variations in model error, the hour-specific performance of the KF bias-adjustment technique, and any improvement offered by post-processing to the prediction of daily maximum PM2.5 values was missed.

RESPONSE: In the revised manuscript, we have now specified that the U.S. Air Quality Index is defined using daily maximum 8-hr O3 and daily (24-hr) mean PM2.5 concentrations. The diurnal variations in model errors have previously been examined (e.g., Eder et al., 2005; Mathur et al., 2009). Even though we could investigate the diurnal variations for the performance of the KF bias-adjustment technique, we believe that the errors are manifested collectively in the analysis for the daily mean quantity which

GMDD
2, C624–C639, 2010

is the air quality standard. Further, this study focuses on improving AQI forecasts, and hence, the selected metrics. The reviewer however makes a good point that the ability of the KF method in improving the representation of the diurnal variation in PM forecasts should also be examined; this will be pursued in subsequent analysis.

2. There seems to be a disconnect between the Eta model references that are given on p. 1377, l. 8 and the use of the WRF-NMM model noted on p. 1379, l. 14. References for the latter would seem to be more appropriate if that is the meteorological model on which the present study was based.

RESPONSE: Even though the meteorological model was switched from Eta to WRF, the scientific core of the modeling system remained unchanged, that is the North American Mesoscale (NAM) model which was described in the references. Eta and WRF only differ in their modeling frame work, i.e., the structure of the modeling system. In addition we have now provided an online reference for the WRF model (http://www.dtcenter.org/wrf-nmm/users/).

3. On p. 1378, l. 5, it would strengthen this statement if substantiation were provided by referencing several publications as examples of the use of post-processing bias adjustment techniques with NWP model forecasts.

RESPONSE: We have now added a reference and the online link of the operational post-processing bias-adjustment for NWP model forecasts as the reviewer suggested.

4. It is not made clear in the manuscript that the approach being discussed is only applicable at locations where near-real-time PM2.5 measurements are available. I have made a few suggestions in the Technical Corrections section of places in the manuscript where some clarification could be added. Neither is there any discussion of the limitations and inconsistencies introduced by this approach. By that I mean (a) that this technique cannot provide any guidance for model predictions away from monitoring station locations and (b) that it introduces implicit inconsistencies between model predictions in between monitoring stations and the bias-adjusted predictions at

GMDD
2, C624–C639, 2010

those station locations that a forecaster would have to be aware of. A few groups have suggested objective-analysis approaches (e.g., Blond et al., 2003, J. Geophys. Res., 108, doi:10.1029/2003JD003679) that it might also be possible to apply to PM2.5 forecasts to address this second problem by modifying the gridded forecast based on the bias-adjusted point-specific forecasts.

RESPONSE: Following the reviewer's suggestion, we revised the summary to read: "In this study the KF bias-adjustment technique is only applied at discrete points, i.e., at location of the monitors. Further research is needed to extend this technique for the development of bias-adjusted spatial maps (i.e., also at location where no monitor information is available) for surface-level PM2.5 distributions. Since surface-level PM2.5 distributions are influenced by local forcing associated with several meteorological drivers and spatially heterogeneous emissions, information on the spatial representativeness of the individual measurements and, consequently, the adjusted bias is critical to the extension of the method presented here to develop bias-adjusted spatial maps of PM2.5 forecast." We thank the reviewer for the suggestions to clarify this point.

5. I acknowledge the ever-present tension between brevity and completeness in describing a study, but there are some details missing from Section 2.1 that I would have liked to have seen: âĂć What were the horizontal domain, map projection, model top, vertical coordinate, and horizontal and vertical grid spacings used by the WRF-NMM model in this application?

RESPONSE: The WRF-NMM covers 1/3 of northern Hemisphere with central latitude-longitude at 52 oN, 106 oW (southern central Canada) using 12km horizontal grid spacing and rotated latitude-longitude projection with Arakawa E-grid-staggering. There are 60 vertical layers with lowest interface at 38m and the model top is set at 2mb. This information has now been incorporated into the revised manuscript.

âĂć What was the CMAQ vertical coordinate and model top that was used?

RESPONSE: CMAQ employed the same hybrid vertical coordinate system as WRF-

GMDD
2, C624–C639, 2010

NMM. The vertical extent in CMAQ ranged from the surface to 100mb which was discretized using 22 layers of variable thickness. . Following the reviewer's suggestion, this information has now been included in the revised manuscript.

âĂć What was the base year for the U.S. EPA national emissions inventory that was used? Which emissions inventories if any were used to account for Canadian and Mexican emissions?

RESPONSE: Area source emissions were based on the 2001 National Emission Inventory (NEI) version 3. Wild fire emissions were based on an average of estimates for the years 1996-2002. Emission estimates for all Canadian provinces were based on the 1995 Canadian emission inventory, while estimates for point sources in Mexico were derived from the 1999 BRAVO inventory. For the EGU point sources, Continuous Emission Monitoring (CEM) data from 2005 are updated using projections from the Department of Energy's Annual Energy Outlook released in January of 2007. For mobile source emissions, Vehicle Miles Traveled (VMT) data are projected to the forecast year (2007) and are used with forecast year fleet data to initialize EPA's MOBILE6 model (U.S. EPA, 2003). Biogenic emissions of VOCs and NO are processed using Biogenic Emission Inventory System (BEIS) version 3.13

âĂć Were biogenic emissions considered, and if so, how?

RESPONSE: The biogenic emissions were processed using Biogenic Emission Inventory System (BEIS) version 3.13. This information has now been incorporated into the manuscript.

âĂć Were any intermittent, natural PM emissions sources, such as sea salt from wave breaking, wildfires, and wind-blown dust, considered? If not, this should be noted along with the implication of an expected negative bias for PM concentrations as a consequence.

RESPONSE: Emissions from these intermittent and natural sources were not consid-

GMDD

2, C624–C639, 2010

ered in the applications discussed here. This information has now been incorporated into the revised manuscript.

âĂć What were the chemical lateral boundary conditions used by CMAQ? Did they vary at all in time and in space? Would trans-Pacific transport episodes be represented?

RESPONSE: A fixed vertical profile (Byun and Ching, 1999), representing “clean” background conditions for each transported species was used to specify the chemical lateral boundary conditions. The impacts of episodic transport from outside the domain (such as episodic trans-pacific events) are thus not captured.

âĂć Which model versions of WRF-NMM and CMAQ (and PREMAQ) were used in this study?

RESPONSE: WRF-NMM was version 2.0 and CMAQ version was 4. .6; the aerosol represention was based on the “AERO3” module configuration. PREMAQ was specifically designed for this forecasting project, and hence, there is no version number. This information is added in the manuscript.

âĂć (p. 1380, l. 14) Was just one 48-hour CMAQ forecast made per day, and if so, why was the 06 UTC time chosen when the meteorological forecasts starting at 00 UTC and 12 UTC are based on more meteorological observations?

RESPONSE: The PM2.5 forecasts in 2007 were in the developmental mode. Even though the meteorological forecasts were issued 4 times per day at 00, 06, 12, and 18 UTC, the PM2.5 forecasts were initialized once daily at 06 UTC only. We have now further clarified this point in the manuscript.

6. The discussion in Section 2.2 of the PM2.5 observations misses a number of issues. First, it is not mentioned but the TEOM measurement bias is known to vary with ambient temperature and hence with season, so that wintertime biases are considerably larger than summertime ones. Second, were the TEOM measurements that were reported to AIRNow bias-corrected or not before transmission to AIRNow (this may also

GMDD

2, C624–C639, 2010

vary by network)? This is a fundamental issue because if the measurements themselves are biased, then the bias-adjustment technique will adjust model predictions towards the measurement bias. Third, some beta attenuation measurement (BAM) instruments are also employed in North America to measure real-time hourly PM2.5 concentration, and these instruments have different error characteristics than TEOMs; were any measurements from these instruments considered? And fourth, it might be worth noting that although other PM2.5 concentration measurements are made and are available retrospectively (e.g., IMPROVE network), they are not near-real-time measurements nor are they available through AIRNow, which means that they are not useful for bias adjustment of operational forecasts.

RESPONSE: We agree with the reviewer that there are biases and uncertainties associated with the TEOM measurements as was briefly mentioned in the manuscript "It should be recognized that TEOM measurements are somewhat uncertain and are believed to be lower limits to a "true" value because of volatilization of semivolatile material in the drying stages of the measurement". Detailed information about TEOM's measurements can be found in the references given therein. As a bias-adjustment technique study, we can only rely on what is available in the observations and treat the observations as the "true" representation of the real world. Currently, the TEOM measurements are the only hourly near real-time PM2.5 measurement data available in the U.S. that can be employed in any real-time air quality forecast for the bias-adjustment study; thus, our study is based solely on data from this measurement network. Therefore, it is true that the proposed bias-adjustment technique will adjust model prediction towards biased measurements as the reviewer stated.

7. Figure 1 has some problems. First, the caption suggests that it shows the forecast domain whereas it only shows that the domain includes 48 states of the U.S.A. plus the District of Columbia. It does not show where the lateral boundaries of the domain are actually located, which would be valuable information. Second, the figure (and Figure 7) shows that PM2.5 measurements are also available from southern Canada,

GMDD
2, C624–C639, 2010

but nowhere is it mentioned (e.g., Section 2.4) whether the Canadian measurements are used or not in the calculation of the continental and subregional statistics (i.e., Tables 1 and 2, Figures 2, 4, 5a, 6, 8, 9, 10, 11). MB was calculated for these stations for Figure 7.

RESPONSE: Figure 1 illustrates the sub-regions used in the analysis. The reviewer is correct that the horizontal extent of the model domain is slightly larger than that illustrated in Figure 1. We have modified the figure caption and the figure to remove the ambiguity noted by the reviewer. Measurements from southern Canada shown in Figure 1 were included in the estimation of the domain mean statistics but not included in the sub-regional statistics. We have now further clarified this point in the manuscript.

8. The second paragraph of Section 2.4 describes a subregional division of the continental U.S. into seven regions based on analysis of an O3 climatology, but the authors do not indicate whether they believe that this division is similarly relevant for PM2.5, given its different characteristics from O3. Figure 1 also shows only six regions, not seven?

RESPONSE: In addition to O3, other atmospheric constituents which are compositions of PM2.5 also displayed similar characteristics as shown by Gego et al. (2005), therefore we assume that PM2.5 would also present similar characteristics. We have now revised the manuscript and referenced the paper by Gego et al. (2005). There are only six regions and the typographical error has been corrected; we thank the reviewer for pointing out the inconsistency.

9. I think the sentence at the end of the third paragraph of Section 2.4 sows confusion. First, no explanation is given as to why the cool season was divided into two parts (perhaps in recognition of the significant change to CMAQ noted in Section 2.1 that was made in mid-September 2007, perhaps not). And second, none of the analyses presented in the figures and tables thereafter mention these two subseasons; instead, they all seem to be based on measurements and predictions from Jan. 1 to mid-April

GMDD
2, C624–C639, 2010

and from September to December 2007 (see Table 1 and Figures 5, 7, 8, 10, and 11). There is also only one subsequent reference to a first or second cool season thereafter in the text that I noticed (p. 1383, l. 4). If the authors wish to retain the terms "first cool season" and "second cool season", perhaps they should also refer to the "full cool season" in the appropriate table and figure captions.

RESPONSE: We thank the reviewer for pointing out the confusion. We have now removed the separation of the cool season in the revised manuscript.

10. Based on the formulas in Kang et al. (2005), the calculation of the two categorical statistics in Section 3.4 of this paper appears to have been done correctly. However, different practitioners use different names for these two statistics, and, as noted recently by Barnes et al. (2009, Wea. Forecasting, 24, 1452-1454), there has been much confusion in the literature between false alarm rate (F) and false alarm ratio (FAR). As a service to the reader, it would be helpful if alternate names for these two statistics could also be mentioned and the Barnes et al. (2009) note referenced. For example, perhaps p. 1388, l. 22-23 of the manuscript could be reworded as follows: "Figure 11 displays the false alarm ratio (FAR; also known as probability of false alarm) and hit rate (H; also known as probability of detection)(see Kang et al., 2005; Barnes et al., 2009) for the raw model and ..."]. It would also help a non-specialist reader if the possible range and the interpretation of extreme values were given for FAR and H in this section.

RESPONSE: We have now incorporated the suggestions into the manuscript by giving the alternate names of the metrics and included the suggested reference. The possible range and the extreme values along with the definitions are already detailed in Kang et al. (2005) when these metrics were defined in more detail; interested readers are referred to Kang et al(2005) for this information.

11. The discussion of Figure 11 near the end of Section 3.4 states that the KF forecasts "increased the H values for all the sub-regions except for the LM and RM". Would it be more correct to append the phrase "in the warm season and the UM in the cool

GMDD

2, C624–C639, 2010

season”?

RESPONSE: We thank the reviewer for this more precise expression; the suggestion has now been incorporated into the manuscript.

12. Two references cited on p. 1378 (McKeen et al. 2007 and Appel et al., 2008) aremissing from the References section.

RESPONSE: The two references have now been added in the reference.

13. Following the example of Table 3 of Eder et al. (2009, Atmos. Environ., 43, 2312-2320), it would be useful to add three columns to both Tables 1 and 2 to provide sample size N, observed mean O, and modeled mean M for each subregion.

RESPONSE: Following the reviewer’s suggestion, we have now provided the size N, observed mean, and modeled mean values for the domain and each subregion in Tables 1 and 2.

14. For Figure 10, would it be possible to indicate the number of samples for each bin somehow either in the text or on the figure itself?

RESPONSE: We have now provided the sample sizes for each bin in the figure caption.

TECHNICAL CORRECTIONS p. 1376, l. 2. Change “particular” to “particulate”.

RESPONSE: The typographical error has now been corrected.

p. 1376, l. 15. Capitalize “pacific coast”.

RESPONSE: It has been corrected.

p. 1376, l. 18. Change “systematical” to “systematic”.

RESPONSE: It has been changed.

p. 1377, l. 12. “integral(?) data set”

RESPONSE: We have now changed it to “unique data set”.

GMDD
2, C624–C639, 2010

p. 1377, l. 20. Perhaps “emissions and physical, chemical, and ...”.

RESPONSE: The change has now been made.

p. 1377, l. 22. Insert comma before “resulting”.

RESPONSE: The comma has been inserted.

p. 1377, l. 23. Change “poses” to “pose”.

RESPONSE: The typographical error has now been corrected.

p. 1377, l. 29. (Also p. 1378, l. 6) Insert hyphen between “bias” and “adjustment”.

RESPONSE: The hyphen has been inserted.

p. 1378, l. 5. Change “model” to “models”.

RESPONSE: The change has now been made.

p. 1378, l. 7. Would suggest inserting the phrase “at locations with PM2.5 monitors” before “is warranted” [see Specific Comment 4].

RESPONSE: As we have mentioned earlier [in response to Specific Comment 4], even though the KF bias-adjustment technique is only applied to the locations with PM2.5 monitors in this study, it is possible to extend this technique to cover locations where no monitors are available by generating the bias-adjusted forecast map, but further research is needed before it can be operational.

p. 1378, l. 11. Change “refer” to “referred”.

RESPONSE: This has been changed.

p. 1378, l. 16. Perhaps “It was not clear whether they would be readily applicable for PM forecasts ...”.

RESPONSE: We have revised the sentence as the reviewer suggested.

GMDD

2, C624–C639, 2010

p. 1379, l. 9. This is where it might be expected, but there is no “bridge” here to the rest of the manuscript via a brief description of the structure of the rest of the manuscript.

RESPONSE: We have now revised the manuscript as the reviewer suggested.

p. 1379, l. 15. This is the first use of AQF but this acronym is not defined (same comment for “PDFs” on p. 1385, l. 3).

RESPONSE: We have now defined the acronym AQF. PDFs was defined earlier on p. 1384, l. 27.

p. 1379, l. 16. Perhaps “... model, which simulates the transport ...”.

RESPONSE: The revision has now been made.

p. 1379, l. 17. Change “substance” to “substances”.

RESPONSE: The change has been made.

p. 1379, l. 21. Change “spacings” to “spacing”.

RESPONSE: The change has been made.

p. 1380, l. 4. Perhaps “... can improve point-specific forecast results over the raw model ...” [cf. Specific Comment 4].

RESPONSE: As we mentioned earlier, even though the KF bias-adjustment technique is only applied to the locations with PM2.5 monitors in this study, one can extend this technique to cover locations where no monitors are available by generating the bias-adjusted forecast map. But, the robustness of spatial mapping techniques need to be evaluated before such methods are used in operational air quality forecasting.

p. 1381, l. 24. Perhaps “from then on” in place of “further on”.

RESPONSE: The change has been made.

p. 1382, l. 4. Change “Normalize Mean Error” to “Normalized Mean Error”.

GMDD

2, C624–C639, 2010

RESPONSE: The change has been made.

p. 1382, l. 5. IOA is not included in this list but is used in Figure 8.

RESPONSE: We have now included IOA in the verification statistical metrics.

p. 1382, l. 7. Perhaps “Two categorical metrics, False Alarm Ratio (FAR) and Hit Rate (H), are used ...”.

RESPONSE: We have revised the sentence as the reviewer suggested.

p. 1382, l. 16. “... time series are ...”.

RESPONSE: The change has been made.

p. 1382, l. 18. “... during the cool season ...”, “... during the warm season ...” (see also p. 1385, l. 2 and l. 12, p. 1386, l. 17 and l. 20, etc.).

RESPONSE: We have made all the changes in the manuscript.

p. 1382, l. 20. “... the time series are divided into ...”.

RESPONSE: The change has been made.

p. 1382, l. 22. Change “Further more” to “Furthermore”.

RESPONSE: Following the reviewer’s suggestion, this sentence has been removed.

p. 1383, l. 3. Perhaps “... overestimated PM2.5 concentrations on average during the ...”. p. 1383, l. 20. “... Coast presents a completely different story, ...”.

RESPONSE: All the suggested changes have now been made.

p. 1383, l. 26 Change “attributed” to “attributable”.

RESPONSE: The change has been made.

p. 1384, l. 23. Don’t need hyphen here in “bias adjustment”, since it is not being used as a compound adjective. Same comment on p. 1385, l. 21, p. 1386, l. 10, etc.).

GMDD

2, C624–C639, 2010

RESPONSE: All the suggested changes have been made.

p. 1386, l. 23. Perhaps “reduce” instead of “rectify”.

RESPONSE: The change has been made.

p. 1388, l. 25. Perhaps “An exceedance threshold value of ...”.

RESPONSE: It has been revised.

p. 1389, l. 13. Suggest “... during the year of 2007 for locations with hourly PM2.5 monitors” [see Specific Comment 4].

RESPONSE: We have revised the sentence as suggested.

p. 1389, l. 21. Perhaps “the warm season; in contrast the opposite is true ...”.

RESPONSE: We have revised the sentence as suggested.

p. 1389, l. 26. Suggest “... significantly improved the PM2.5 forecasts for locations with hourly PM2.5 monitors as revealed by reductions ...” [see Specific Comment 4].

RESPONSE: The revision has been made as suggested.

p. 1390, l. 3. Perhaps “... transition of seasons or model changes”.

RESPONSE: The addition has been made.

p. 1390, l. 29. Perhaps “... in issuance of air-quality-degradation-related health advisories”.

RESPONSE: The change has been made.

p. 1392, l. 21. Change “Mckeen” to “McKeen”.

RESPONSE: The typo has been corrected.

Fig. 1. In caption change “AIRNOW” to “AIRNow”.

GMDD

2, C624–C639, 2010

RESPONSE: The change has been made.

Fig. 2. Add to caption that time series are daily means over all stations.

RESPONSE: The suggested change has been made.

Fig. 3. Add “subregions” at end of caption.

RESPONSE: The word has been added.

Fig. 4. Revise caption to indicate that scatterplots are based on forecasts and measurements for all stations and all days.

RESPONSE: The caption has been revised as suggested.

Fig. 10. Third line of figure caption seems out of place; e.g., why is Figure 9 referred to?

RESPONSE: This typographical error has been corrected.

Fig. 11. Why not continue to use the abbreviation “MOD” instead of “MD” (cf. Tables 1 and 2 and Figures 2, 3, 7, and 8)?

RESPONSE: The change has been made.

One general comment is that GMD is an international journal and it should be clear from the article text that a air quality modeling system for the U.S. is being discussed. There are several places in the text where this clarification could be made: p. 1376, l. 2. Insert “for the US” after “forecasts”. p. 1376, l. 25. Insert “in the US” before “to compute the Air Quality Index”. p. 1377, l. 3. Insert “in the US” after “publicly available” (otherwise, reference list is parochial: e.g., see http://www.esrl.noaa.gov/csd/events/iwaqfr/ for other groups and agencies involved in real-time ozone and PM forecasting). p. 1388, l. 25. Perhaps “... based on the U.S. National Ambient ...”. p. 1389, l. 16. “the entire domain” or “the continental U.S. portion of the domain”?

RESPONSE: All suggested changes have been incorporated into the manuscript.

GMDD
2, C624–C639, 2010

Interactive comment on Geosci. Model Dev. Discuss., 2, 1375, 2009.

GMDD

2, C624–C639, 2010