Dear Editor and Reviewers,

Thank you for the comments to help improve the quality of the paper. We have revised the manuscript to address your comments. A detailed response to each comment is provided in this file with comments from referees in black, author's response in red, and author's changes in manuscript in blue.

**Report #1 by anonymous referee #2:**

The authors presented their efforts in examining the spatial and temporal distributions of air pollutants over China, with a full year simulation and local observations. Although the WRF/CMAQ model has been applied in many studies over East Asia recently, this full year simulation helps to probe into the modeling discrepancy and stability. The manuscript generally introduced the performance but lacked sufficient discussion or emphasis on the findings as a research paper. More explicit discussion and demonstration of the findings are necessary to illustrate the contribution of this manuscript. Therefore, I recommend the manuscript to be accepted with minor revision if the following questions/issues are being addressed in details by the authors.

Responses: We thank the reviewer for providing important comments that help improve the quality of manuscript. However, the comments are same as the quick review process before its publication on ACPD. We have responded to the comments and here we made additional changes together with other reviewers' comments.

(1) General comment: The manuscript spent a lot efforts describing the results of model evaluation, but lack of detailed discussion about the factors that are/may responsible for the good/bad model performance. For example, section 3.1 describes the statistics summarized in Table 1, but it would be much more informative for the research community if the authors can discuss about why T2 is overestimated in winter. Is it because of physical options in WRF or bias induced by FNL? Section 3.2 describes the performance of O3 and PM2.5 predictions at different sub-regions in China, but there some important issues remain unknown. The research domain is divided into several sub-regions, thus the authors are expected to point out why the performance varies among sub-regions, is it correlated with bias from meteorology, emission, or chem/physical mechanisms applied? These are the real important findings this paper can contribute to the research community rather than demonstrating that this specific model application meets the benchmark.

Responses: We thank the reviewer for pointing out these important issues. We responded specific comments such as meteorology performance, sub-regions etc. below. Here we want to have a general response.

We agree with the reviewer that there are important problems relating simulation of air pollution in China, from meteorology, emissions, and domain setting, to model representations of physical and chemical processes. Even observations have large limitations and China is currently lacking routine long-term measurements of $PM_{2.5}$ chemical components. These issues have to be addressed by contributions from the entire community and it is beyond the scope of study to address them all. Particularly, this paper aims to examine the capability of a state-of-art CMAQ model in reproducing severe air pollution in China by comparing with the recently public available air quality monitoring data for an entire year. By identifying the regions/seasons with good model performance, further studies investigating potential air pollution control strategies and associating human health outcomes with air

pollution exposure can be conducted with confidence using the dataset generated from this study. In the meantime, by identifying the regions/seasons with less desirable performance, we can point out which areas need more work in future. This is why we are rigorously comparing with the model application benchmarks suggested by the US EPA. Thus, as the first study to do so, this paper has its own merits.

In fact, we indeed found different model performances for different seasons and sub-regions and we pointed out the possible reasons that can be implied by this study in the discussion part. However, it is beyond the scope of this study to answer very specific questions such as whether physical options of WRF or bias associated with FNL caused the T2 overestimation in winter, or quantifying the uncertainties in the predicted concentrations due to meteorology, emission, or chemical/physical mechanisms. These specific questions are research topics that need to be addressed in separate studies. We are currently evaluating different WRF model configurations to improve our meteorology model performance. Also, we are investigating CMAQ model performance based on different emission inventories. Addressing these questions will be published in follow-up papers and are beyond the scope of this paper.

Changes in manuscript: no changes were made for this comment.

(2) General comment: Some of the discussions in section 4 lack detailed explanation thus are not persuasive to support the conclusions. For example, line#440-448 states the bias from met prediction may affect the chemistry of CMAQ, but line#451 attributes the bias of SO2 and NO2 to anthropogenic emissions. Although CO is a good indicator for estimating bias induced by anthropogenic emission, SO2 and NO2 are often affected by other factors, such as meteorology, biomass burning emission as well. How does the overestimation of T2 affect the chemistry of SO2 and NO2, is it pushing the bias induced by anthropogenic emission towards or away from the benchmarks? Are there any evidences, such as satellite products/surface monitors/field studies, that can help to identify the location and intensity of emission bias? There are some published modeling studies using nested domains with different grid resolutions, are their findings support the statements in line#478-486? As the manuscript mainly focused on model performance, it is necessary to probe deeper into at least some of these issues, to investigate the contributions from different sources of uncertainties, such as meteorology, anthropogenic emission, biomass burning/biogenic/dust emission.

Responses: Model bias is affected by the combination of bias in meteorology prediction, emissions, and air quality model algorithms. In line#451 in the original manuscript, we attribute the bias of CO, NO2 and SO2 in NW to anthropogenic emissions, not to meteorology bias because we compared the biases of air quality and meteorology predictions in NW to other regions. When checking the model performance among difference regions, we noticed larger biases in predicted air pollutant concentrations in NW than other regions, but meteorology performance among different regions was similar. In addition, unlike the NCP, YRD and PRD regions where a significant and continuous effort was undertaken to improve the accuracy of the emission estimations, the northwestern regions have not been the focus of previous investigations and it is likely that a lot of the smaller emission sources were not reported when developing the emission inventories. Biases in predicted surface temperature could directly affect the loss rate of SO2 and NO2 through reactions with OH. However, based on the current understanding of the temperature dependence of the OH reactions, one-degree Celsius bias in ambient temperature only changes the loss of SO2 and NO2 through OH reactions by less than 1%. By ruling out

other possible explanations for the significant negative biases, we believe that bias in emissions is more likely the reason for the substantial under prediction of SO2 and NO2 in the NW.

A few studies have examined the major emission sources, such as power plants, using satellite observations (Zhang et al., 2012; Liu et al., 2016) and combined with chemical transport model studies (Wang et al., 2012), building confidence in the emissions from the major point sources. However, emissions from area sources, such as residential sources, are difficult to characterize and therefore have much larger uncertainties. Recent studies indicate that residential emissions are an important but generally unrecognized source of ambient air pollution in China (Hu et al., 2005l Liu et al., 2016).

References:

Hu, J., Wu, L., Zheng, B., Zhang, Q., He, K., Chang, Q., Li, X., Yang, F., Ying, Q., Zhang, H., 2015, Source contributions and regional transport of primary particulate matter in China, Environmental Pollution, 207: 31-42

Liu, J., Mauzerall, D.L., Chen, Q., Zhang, Q., Song, Y., Peng, W., Klimont, Z., Qiu, X., Zhang, S., Hu, M., Smith, K.R., Zhu, T., 2016, Air pollutant emissions from Chinese households: A major and underappreciated ambient pollution source, Proceedings of the National Academy of Sciences of the United States of America, 113(28):7756-7761

Liu, F., Beirle, S., Zhang, Q., Dorner, S., He, K., Wagner, T., 2016, NOx lifetimes and emissions of cities and power plants in polluted background estimated by satellite observations, Atmospheric Chemistry and Physics, 16, 5283-5298.

Wang, S.W., Zhang, Q., Streets, D.G., He, K., Martin, R.V., Lamsal, L.N., Chen, D., Lei, Y., Lu, Z., 2012, Growth in NOx emissions from power plants in China: bottom-up estimates and satellite observations, Atmospheric Chemistry and Physics, 12, 4429-4447.

Zhang, Q., Geng, G., Wang, S., Richter, A., He, K., 2012, Satellite remote sensing of changes in NOx emissions over China during 1996-2010, Chinese Science Bulletin, 57(22):2857-2864.

Regarding the suggestion on nested simulations, we added the following references to support that a finer resolution will help improve model predictions for urban locations (lines 513-515 in the revised manuscript):

Fountoukis, C., Koraj, D.H., Denier van der Gon, H.A.C., Charalampidis, P.E., Pilinis, C., Pandis, S.N., 2013, Impact of grid resolution on the predicted fine PM by a regional 3-D chemical transport model, Atmospheric Environment, 68, 24-32.

Gan, C.M., Hogrefe, C., Mathur, R., Pleim, J., Xing, J., Wong, D., Gilliam, R., Pouliot, G., Wei, C., 2016, Assessment of the effects of horizontal grid resolution on long-term air quality trends using coupled WRF-CMAQ simulations, Atmospheric Environment, 132, 207-216.

Stroud, C.A., et al., 2011, Impact of model grid spacing on regional- and urban- scale air quality predictions of organic aerosol, Atmospheric Chemistry and Physics, 11, 3107-3118.

Regarding the suggestion that we took a deeper look at uncertainties caused by meteorology, anthropogenic emissions, biomass burning, biogenic, and dust emissions, we have included discussions in various relevant parts of the manuscript, such as the discussion of the underestimation of emissions in the NW region. The role of natural emission sources, such as biogenic and dust emissions on model

predictions needs further investigation. As inputs to these emission models and as well as the underlying parameterizations (e.g. the impact of soil moisture on dust and biogenic VOC emissions) have large uncertainties, a more reasonable approach is to study the modeling of these emissions and their impact on air quality in China in separate manuscripts rather than including them all into a single manuscript. Results of our study here will provide the readers in various areas in the community an opportunity to think and identify research areas within their expertise for further detailed investigation, such as the excellent suggestions recommended by the reviewer.

Changes in manuscript: In lines 513-515 in the revised manuscript, we added three more references to support that a finer resolution will help improve model predictions for urban locations.

(3) Minor comment: Please briefly describe why sub-regions are defined to evaluation model performance.
Responses: We added following description in the revised manuscript:
"Concentrations of pollutants in different regions of China exhibit large variations due to diverse climates, topography, and emission sources. Aiming to identify the model strength and weakness in different regions of China, model performance was evaluated separately for different regions."
Changes in manuscript: In lines 225-228 in the revised manuscript, we added above description.

(4) Minor comment: Table3, in Mar OBS and PRE for PM2.5 is 81.68 and 66.12 respectively, while the MNB is only 0.04, please double check this statistics as it indicates large bias but strong correlation between observation and prediction.
Responses: We checked the calculation and the numbers are correct. The small MNB but large MNE is due to model under prediction of very high PM2.5 concentrations and over prediction of very low PM2.5 concentrations, which compensates each other in the MNB calculation and leads to a small MNB value.
Changes in manuscript: No changes were needed for this comment.

(5) Minor comment: Please briefly describe how the MEIC emission is temporally/spatially allocated as CMAQ-ready inputs, since these factors have large impact on regional model performance.
Responses: The MEIC emissions are already spatially allocated into 0.25x0.25 degree grid cells before they were given to us. We re-gridded the emissions to our model domain, which uses 36 km horizontal resolution, using the Spatial Allocator program provided by the US EPA. Monthly MEIC emissions were obtained and temporal allocation of MEIC emissions were also conducted based on weekly and diurnal profiles from MEIC developers. We have cited the papers describing these methodsin the revised manuscript.
Changes in manuscript: We have cited the papers describing these methods in lines 175-176 in the revised manuscript.

(6) Minor comment: line#115, "interaction of" should be "interaction between A and B"
Responses: Corrected it.
Changes in manuscript: We corrected it in line 122 in the revised manuscript.

(7) Minor comment: How are the initial/boundary conditions generated for CMAQ ?

Responses: Initial and boundary conditions were based on the default vertical distributions of concentrations that represent clean continental conditions as provided by the CMAQ model. The impact of initial conditions was minimal as the results of the first five days of the simulation were excluded in the analyses.
Changes in manuscript: We added description of the initial/boundary conditions in lines 186-189 in the revised manuscript.

(8) Minor comment: line#175-178. The default inline dust emission module in CMAQ was reported to significantly underestimated the emission of total dust by Fu et al. (2013) and Dong et al. (2015) due to the double-count of soil moisture effect. But dust mainly dominates the coarse mode aerosol so it may not influence the performance for O3 and PM2.5 which are the focus of this study. Still I would suggest the author take a look at their PM10 results especially in spring, since PM10 is also in the criteria air pollutants.
Responses: The most recent version of the CMAQ dust module appears to have already been using the adjusted critical U* values (see DUST_EMIS.F, around line 842, in CMAQv5.0.1). For example, the U* for barren land soil was changed from 0.28 to 0.23, the recommended adjustment as mentioned in Dong et al. (2015). In general, the U* values used in our simulations agree with the ones used by Dong et al (2015). We also observed that dust emissions are likely under-estimated by our updated dust module. We calculated the PM10 results and added the results in Table 3 and Table 4 as the reviewer suggested. It is clearly that PM10 is significantly underestimated. The discussions of PM10 were added in the revised manuscript.
Changes in manuscript: We added the PM10 results in Table 3 and Table 4. We added the discussions of PM10 in line 235, lines 297-299, and line 305 in the revised manuscript.

(9) Minor comment: line#286, “Figure ?”
Responses: It is “Figure 4”.
Changes in manuscript: We have corrected it in line 306 in the revised manuscript.

(10) Minor comment: line#504: “this is the first study”. Zhao B. et al. (2013, ERL) did a full-year simulation with WRF/CMAQ in China, and there are some studies with MM5/CMAQ in China prior to 2013 too.
Responses: Although a full year air quality simulations have been conducted previously by Zhao et al. (2013, ERL) and other studies (Gao et al., 2014, Atmos. Environ.; Wang et al., 2011, Atmos. Environ.), model performance on temporal and spatial variations of air pollutants were mostly evaluated against available surface observation at a limited number of sites. In addition, the surface observations were mostly based on the MEP’s air pollution index (API) numbers which could be used to calculate the concentrations of the major pollutant of SO2, NO2 or PM10. Therefore it is still true that no studies have reported “the detailed model performance of O3 and PM2.5 for an entire year”. We modified the introduction section to include the above facts and to avoid confusion, we revised the expression to “this study reports…”
Changes in manuscript: We modified the introduction section to include the above facts in lines 103-110 in the revised manuscript, and we revised the expression in line 536 in the revised manuscript.

(11) Minor comment: line#205-206: “WRF model has acceptable”. Table 1 indicates many of the

variables failed to meet the benchmark. The Zhao B. et al. (2013, ERL) shows WRF performance all falls in the benchmark, so I would suggest the authors check their configurations of WRF namelist to either improve the WRF performance or specify the reason for relatively large bias in this study.

Responses: Many thanks for providing the reference. We checked the model configuration used in the study of Zhao B. et al. (2013, ERL) and compared to ours (list in the Table R1). We also noticed that Zhao B. et al. (2013, ERL) evaluated the WRF performance at ~380 stations, but in our study, we evaluated WRF performance at ~1200 stations. For this reason, we cannot directly compare the model performance between the two studies. Our WRF model performance is consistent with some other previous studies. In addition to the studies we compared in the manuscript, we found another study by Wang et al. (2014, ACP) reported comparable WRF performance as ours. We added both Zhao B. et al. (ERL) and Wang et al. (2014, ACP) in the revised manuscript and pointed out the different WRF performance in different studies. However, improving WRF model performance needs extended efforts and is beyond the scope of this paper.

Wang, L.T., Wei, Z., Yang, J., Zhang, Y., Zhang, F.F., Su, J., Meng, C.C., Zhang, Q., 2014, The 2013 severe haze over southern Hebei, China: model evaluation, source apportionment, and policy implications, Atmospheric Chemistry and Physics, 14, 3151-3173.

Table R1. WRF model configurations in this study and in Zhao B. et al (2013, ERL)

| **Physics** | **This Study** | **Zhao B. et al. (2013, ERL)** |
|---|---|---|
| Microphysics | New Thompson scheme | WSM 3-class scheme scheme |
| Long wave radiation | RRTM scheme | RRTM scheme |
| Shortwave radiation | Goddard shortwave | RRTM shortwave |
| Land surface | MM5 Land Surface Model | NOAH Land Surface Model |
| Planetary boundary layer | Yonsei University scheme | Mellor-Yamada-Janjic PBL scheme |
| Cumulus Parameterization | Grell-Devenyi ensemble scheme | Grell-Devenyi ensemble scheme |

Changes in manuscript: We added both Zhao B. et al. (ERL) (line 218) and Wang et al. (2013, ACP) (line 215) and pointed out the different WRF performance in different studies (lines 217-220) in the revised manuscript.

**Report #2 by anonymous referee #1:**

In this paper, the author presents a yearlong air quality simulation using a chemical transport model to provide detailed temporal and spatial distribution of O3, PM2.5 and PM2.5 chemical compositions in China. The topic is important, the method is generally sound, and the results are generally reasonable. I suggest this manuscript be accepted with revisions described below.

General comments:

(1) The main objective of this study is to provide detailed temporal and spatial distribution of O3, PM2.5 and its chemical components, which supplements the current observational network in China. The key to success is to ensure that the model well reproduces the magnitude and spatiotemporal distribution of these pollutants. However, the author only compared simulated O3 and PM2.5 concentrations with surface observations. To better evaluate the modeling results, I suggest the author also compare with satellite observations, such as AOD, NO2 column, SO2 column, and tropospheric ozone residual. Moreover, although the observations of PM2.5 chemical components are not publicly available, some data can be found in the literature. It will be very beneficial if the author can compare

the simulated chemical components with some available chemical component data, because the spatiotemporal distribution of chemical components is a major focus of this study.

Response: To address the reviewer's comments, we did some comparison between model predicted and satellite observed NO2 columns. An example of the comparison for August 2013 is shown in Figure R1. The spatial patterns of predicted and observed NO2 columns are consistent in general. While the satellite observations are useful in evaluating regional distribution of air pollutants, we choose not to include these in the present analysis and focus on comparison with ground-level observations in this study. A lot of assumptions were used in generating these vertical column density (VCD) products (gridded level 3 products), such as the vertical distribution of target species. To allow an "apple-to-apple" comparison of the model predictions, it is necessary to use a lower level product (e.g. level 2 (L2) products) and adjust the satellite VCD using modeled vertical distributions. The adjusted L2 products will have to be gridded to compare with model predictions (Duncan et al., 2014). The aerosol optical depth (AOD) is not directly comparable with surface PM2.5. Deducing surface PM2.5 from AOD is no trivial task due to spatial and temporal variation of the aerosol composition and weather conditions (CIESIN, 2013). We agree fully with the reviewer that comparison with satellite products is complex yet very important, and thus warrants a more detailed discussion in a separate manuscript.

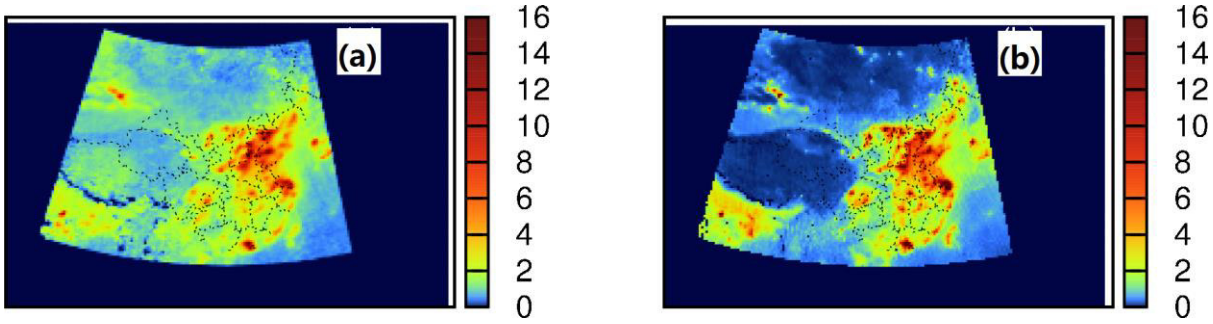



Figure R1. OMI satellite observed (a) and model predicted (b) NO2 columns for August 2013. Units are $x10^{15}$ molecules/$cm^2$

While aerosol composition data are available to us at a few locations, they are far from complete to provide a holistic view of the uncertainties in aerosol chemical composition throughout the country. Case studies to compare predicted and observed aerosol composition at selected locations will be more suitable to document these findings rather than including them in the current paper that focuses on the national level. In a published paper with similar major model settings, we compared some primary components at different locations and included the results in a previous paper (Hu et al., 2015). We cited this paper in this manuscript.

References:

Duncan, B.N., et al., Satellite data of atmospheric pollution for US air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. Atmospheric Environment, 2014. 94: p. 647-662.

Battelle Memorial Institute and Center for International Earth Science Information Network - CIESIN - Columbia University, Global Annual Average PM2.5 Grids from MODIS and MISR

Aerosol Optical Depth(AOD). 2013, Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC).

Hu, J., Wu, L., Zheng, B., Zhang, Q., He, K., Chang, Q., Li, X., Yang, F., Ying, Q. and Zhang, H., 2015. Source contributions and regional transport of primary particulate matter in China. Environmental Pollution, 207, pp.31-42.

Changes in manuscript: no changes were made for this comment.

Furthermore, the comparison results indicate that PM2.5 concentrations are underestimated significantly in some months (e.g., MFB=-48% in July), and the model performance can be quite different in different regions. The author needs to comment how these temporally and spatially variant biases affect the simulation results of spatiotemporal distribution of O3, PM2.5 and PM2.5 chemical composition.

Responses: The temporally and spatially variant biases of PM2.5 do affect the simulated spatiotemporal distribution of PM2.5 and its chemical composition. PM2.5 is more underpredicted in summer when the concentrations are lower, so the predicted seasonal variation of PM2.5 is stronger. PM2.5 is more underpredicted in NW where the concentrations are lower, so the predicted spatial difference between NW and eastern China region (i.e., NCP, YRD, etc) is likely stronger. It should also affect the spatiotemporal distribution of PM2.5 chemical compositions, but no detailed information can be obtained due to the lack of detailed PM2.5 composition observations. The biases of O3 exhibit much less variation temporally and spatially, so the predicted spatiotemporal distribution of O3 is more accurate than PM2.5.

Changes in manuscript: We added the above discussion in Section 3.3 in lines 396-404 in the revised manuscript.

(2) Introduction: The author suggests that most modeling studies focus on a specific pollution episode and extensive model performance evaluation is lacking. In fact, as far as I know, quite a few studies have been done to evaluate the model performance in China for a full year or several representative months (e.g., Gao et al., 2014; Zhang et al., 2016; Liu et al., 2016; Zhao et al., 2013; Wang et al., 2011; Liu et al., 2010), and there are more. The author should review these long-term modeling studies because they highly resemble the work presented here.

Gao, Y., Zhao, C., Liu, X. H., Zhang, M. G., and Leung, L. R.: WRF-Chem simulations of aerosols and anthropogenic aerosol radiative forcing in East Asia, Atmos Environ, 92, 250-266, DOI 10.1016/j.atmosenv.2014.04.038, 2014.

Zhang, Y., Zhang, X., Wang, L., Zhang, Q., Duan, F., and He, K.: Application of WRF/Chem over East Asia: Part I. Model evaluation and intercomparison with MM5/CMAQ, Atmos Environ, 124, 285-300, 10.1016/j.atmosenv.2015.07.022, 2016.

Liu, X. Y., Zhang, Y., Zhang, Q., and He, M. B.: Application of online-coupled WRF/Chem-MADRID in East Asia: Model evaluation and climatic effects of anthropogenic aerosols, Atmos Environ, 124, 321-336, 10.1016/j.atmosenv.2015.03.052, 2016.

Zhao, B., Wang, S. X., Wang, J. D., Fu, J. S., Liu, T. H., Xu, J. Y., Fu, X., and Hao, J. M.: Impact of

national NOx and SO2 control policies on particulate matter pollution in China, Atmos Environ, 77, 453-463, DOI 10.1016/j.atmosenv.2013.05.012, 2013.

Wang, S. X., Xing, J., Chatani, S., Hao, J. M., Klimont, Z., Cofala, J., and Amann, M.: Verification of anthropogenic emissions of China by satellite and ground observations, Atmos Environ, 45, 6347-6358, DOI 10.1016/j.atmosenv.2011.08.054, 2011.

Liu, X.-H., Zhang, Y., Cheng, S.-H., Xing, J., Zhang, Q., Streets, D. G., Jang, C., Wang, W.-X., and Hao, J.-M.: Understanding of regional air pollution over china using CMAQ, part I performance evaluation and seasonal variation, Atmos Environ, 44, 2415-2426, 10.1016/j.atmosenv.2010.03.035, 2010.

Responses: We thank the reviewer for pointing out the references. The focus of the current manuscript is to evaluate the model performance of surface level O3 and PM2.5. Although a full year or several representative months' air quality simulations have been conducted previously, model performance on temporal and spatial variations of air pollutants were mostly evaluated against available surface observation at a limited number of sites. In addition, the surface observations were mostly based on the MEP's air pollution index (API) numbers, which could be used to calculate the concentrations of the major pollutants of SO2, NO2 or PM10. Therefore, it is still true that no studies have reported "the detailed model performance of O3 and PM2.5 for an entire year". We have modified the introduction section to include the above facts and cited the above references.
Changes in manuscript: We added above discussion in Section 1 lines 103-110 in the revised manuscript.

Specific comments:
(1) Line 222-225: Why does the author filter out these data? How are the thresholds determined?
Responses: We performed quality control checks on the raw observation data, and filtered out data that are either unrealistically high or show abnormal temporal variations, which could greatly bias the model performance analysis. For the extreme values, we choose 250 ppb for hourly O3 and 1500 μg $m^{-3}$ for hourly PM2.5. These cut-off values were chosen based on past experience in regional air quality modeling. While locally, these extreme values might be possible, they were likely not representative of the regional average concentrations for a 36x36 $km^2$ grid cell. We also removed days that had standard deviation less than 5 ppb for O3 or 5 μg $m^{-3}$ for PM2.5. This is also based on the general understanding of the typical diurnal variations of O3 and PM2.5 in polluted urban areas, and examination of the data at all the monitoring stations collected for this study.
Changes in manuscript: no changes were made for this comment.

(2) Line 228: There should be a comma before "PM2.5"
Responses: Corrected it.
Changes in manuscript: add a comma before "PM2.5" in line 235.

(3) Line 284-291: The PM2.5 concentrations are underpredicted significantly in some months. The author should explain the reason for the underestimation. In addition, the author attributes the underestimation in PM10 to natural and anthropogenic dust emissions. How is wind-blown dust

emissions calculated in CMAQ (any reference)? Are there any previous studies showing that the dust emission module embedded in CMAQ underpredict wind-blown dust emissions?
Responses: PM2.5 is generally more underpredicted in the warm months (April-July) than in the cold months (November-March). We think that SOA underprediction is likely an important reason for this phenomenon and we have discussed it in lines 488-498.

Reviewer 2 also commented on the dust module (comment #8) and details about the dust module can be found in our response to that comment. Briefly we used the CMAQ inline dust module with modifications to use the land use types in MODIS rather than the BEIS, as it only works for the United States. As previous studies by Fu et al. (2013) and Dong et al. (2015) reported that the dust module significantly underestimated the emission of total dust, it is possible that dust emissions were estimated in the current study as well.
Changes in manuscript: We added the PM10 results in Table 3 and Table 4. We added the discussions of PM10 in line 235, lines 297-299, and line 305 in the revised manuscript.

(4) Line 197: The author states that the benchmarks are adapted from Emery et al. (2012). However, the author indicates that these benchmarks are from Emery et al. (2001) in the title of Table 1. The two papers seem quite different and the latter one appears the correct source. Please confirm.
Responses: It is Emery et al. (2001). We have corrected it.
Changes in manuscript: We have corrected it in line 204 in the revised manuscript.

**Report #3 by anonymous referee #3:**
Do you have any idea about why the model underestimate the O3 maximum in October or in fall over PRD region?
Responses: Figure 2 shows that the model predicted 8h maximum O3 concentrations agree well with observations in Guangzhou in March, May-September, and November-December. Model overpredicted 8h maximum O3 in April, but underpredicted 8h maximum O3 in October. The general agreement in most months indicates the O3 chemistry is well represented in the model. The meteorology performance (Table 1) indicates no obvious difference for predicted meteorology bias between the two months of April and October and the other months. By eliminating the above two reasons, we think the O3 performance in April and October was caused by emissions, although there is no direct evidence for the emissions bias in these two months compared to other months.
Changes in manuscript: No changes were made for this comment.

# One-Year Simulation of Ozone and Particulate Matter in China Using WRF/CMAQ Modeling System

Jianlin Hu[1], Jianjun Chen[2,1], Qi Ying[3,1,*], Hongliang Zhang[4,1,*]

[1]Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control, Jiangsu Engineering Technology Research Center of Environmental Cleaning Materials, Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, School of Environmental Science and Engineering, Nanjing University of Information Science & Technology, 219 Ningliu Road, Nanjing 210044, China
[2]Air Quality Planning and Science Division, California Air Resources Board, 1001 I Street, Sacramento, CA 95814, USA
[3]Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843, USA
[4]Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge LA 70803, USA

[*]Corresponding authors:
Qi Ying, Email: qying@civil.tamu.edu. Phone: +1-979-845-9709.
Hongliang Zhang, Email: hlzhang@lsu.edu. Phone: +1-225-578-0140.

## Abstract

China has been experiencing severe air pollution in recent decades. Although ambient air quality monitoring network for criteria pollutants has been constructed in over 100 cities since 2013 in China, the temporal and spatial characteristics of some important pollutants, such as particulate matter (PM) components, remain unknown, limiting further studies investigating potential air pollution control strategies to improve air quality and associating human health outcomes with air pollution exposure. In this study, a yearlong (2013) air quality simulation using the Weather Research & Forecasting model (WRF) and the Community Multi-scale Air Quality model (CMAQ) was conducted to provide detailed temporal and spatial information of ozone ($O_3$), $PM_{2.5}$ total and chemical components. Multi-resolution Emission Inventory for China (MEIC) was used for anthropogenic emissions and observation data obtained from the national air quality monitoring network were collected to validate model performance. The model successfully reproduces the $O_3$ and $PM_{2.5}$ concentrations at most cities for most months, with model performance statistics meeting the performance criteria. However, over-prediction of $O_3$ generally occurs at low concentration range while under-prediction of $PM_{2.5}$ happens at low concentration range in summer. Spatially, the model has better performance in Southern China than in Northern, Central and Sichuan basin. Strong seasonal variations of $PM_{2.5}$ exist and wind speed and direction play important roles in high $PM_{2.5}$ events. Secondary components have more boarder distribution than primary components. Sulfate ($SO_4^{2-}$), nitrate ($NO_3^-$), ammonium ($NH_4^+$), and primary organic aerosol (POA) are the most important $PM_{2.5}$ components. All components have the highest concentrations in winter except secondary organic aerosol (SOA). This study proves the ability of CMAQ model in reproducing severe air pollution in China, identifies the directions where improvements are needed, and provides information for human exposure to multiple pollutants for assessing health effects.



**Keywords:** Ozone, Particulate matter, WRF, CMAQ, MEIC, China

## 1. Introduction

Atmospheric pollutants have adverse effects on human health and ecosystems and are associated with climate change (Menon et al., 2008; Poschl, 2005; Pui et al., 2014). Developing countries usually experience severely high concentrations of air pollutants due to fast growth of population, industrialization, transportation and urbanization without prompt emission controls. As one of such countries, China started to publish real time concentration data of six criteria pollutants from the ambient air quality monitoring networks after multiple severe pollution events across the country(Sun et al., 2014; Tao et al., 2014b; Wang et al., 2014a; Zheng et al., 2015).

More than 1000 observation sites have been set up in more than 100 major cities in the country to routinely monitor hourly concentrations of six criteria pollutants, i.e., $O_3$, CO, $NO_2$, $SO_2$, $PM_{2.5}$ (PM—particulate matter), and $PM_{10}$, and to inform the public on air quality status using the air quality index (AQI). Analysis of the observation provided a general understanding of the spatial and temporal variation of the levels of air pollution (Hu et al., 2014a; Wang et al., 2014c), the roles of meteorology in air pollution (Zhang et al., 2015b), and the construction of AQI based on multiple pollutants to better inform the public about the severity of air pollution (Hu et al., 2015b). However, the monitoring system only considers criteria pollutants and the key species such as the volatile organic compounds (VOCs) and the chemical composition of PM that are needed to understand the causes of air pollution and form cost-effective emissions controls are not measured routinely. Monitoring networks focusing on the chemical composition of gaseous and particulate air pollutants, such as the Photochemical Assessment Monitoring Stations (PAMS) and the Chemical Speciation Network (CNS) in the United States, have not been established in China. Lacking of detailed chemical composition information limits our capability to understand the formation mechanisms of $O_3$ and PM, quantify the contributions of different sources, and design effective control strategies. In addition, the observation sites are mostly in highly developed urban areas but are very sparse in other suburban and rural regions which also have large population and experience high concentrations of certain pollutants, such as $O_3$. Insufficient spatial coverage in the monitoring system limits the completeness of public air pollution risk assessment for the entire country.

Chemical transport models (CTMs) are often used to reproduce past pollution events, test newly discovered atmospheric mechanisms, predict future air quality, and provide high temporal and spatial resolution data for epidemiological studies. Several modeling studies have been reported to analyze the severe air pollution events in January 2013. For example, the Community Multiscale Air Quality (CMAQ) model was updated with heterogeneous chemistry to study the formation of secondary inorganic aerosol in North China (Zheng et al., 2015). The CMAQ model was also applied to identify the contributions of both source regions and sectors to $PM_{2.5}$ in Southern Hebei during the 2013 severe haze episode with a brute force method (Wang et al., 2014b). It was found that industrial and domestic activities were the most significant local sectors while Northern Hebei province, Beijing-Tianjin city cluster, and Henan province were the major regional contributors. Using the two-way coupled Weather Research and Forecasting (WRF)/CMAQ system, Wang et al. (2014b) simulated the impacts of aerosol–meteorology interactions on the PM pollution during January 2013. They argued that enhanced planetary boundary layer (PBL) stability suppressed the dispersion of air pollutants, and resulted in higher $PM_{2.5}$ concentrations. Similar results were also reported by Zhang et al. (2015a) with the Weather Research and Forecasting/Chemistry (WRF/Chem) model. Using the Comprehensive Air Quality

Model with extensions (CAMx) and the Particulate Source Apportionment Technology (PSAT), Li et al. (2015b) determined the contributions of 7 emission categories and 11 source regions to regional air pollution in China and suggested a strong need for regional joint emission control efforts in Beijing. More recently, Hu et al. (2015a) used a tracer based technique in a source-oriented CMAQ to determine source sector/region contributions to primary PM in different seasons in 2012-2013. It was found that residential and industrial emissions from local area and the neighboring Hebei province contribute to high primary PM events in Beijing.

All above modeling studies except Hu et al. (2015a) were focused on the formation and source apportionment of airborne PM during the severe pollution episode of January 2013 in northern China. Although additional PM formation pathways and/or emission adjustments were implemented and tuned to better predict this extreme episode, model predictions were only evaluated against a small number of measurements in and near Beijing for a relatively short period of time. A few studies have been conducted to evaluate the model performance in China for longer time periods, such as a full year or several representative months in different seasons (Gao et al., 2014; Liu et al., 2010; Liu et al., 2016; Wang et al., 2011; Zhang et al., 2016; Zhao et al., 2013b). However, due to limited ambient observation data, model performance on temporal and spatial variations of air pollutants were mostly evaluated against available surface observation at a limited number of sites. In addition, the surface observations were mostly based on the MEP's air pollution index (API) numbers, which could be used to calculate the concentrations of the major pollutants of $SO_2$, $NO_2$ or $PM_{10}$. Extensive model performance evaluation of $O_3$ and PM is urgently needed to build the confidence in the emission inventory, the predicted meteorological fields as well as the capability of the model in predicting regional $O_3$ and PM under a wide range of topographical, meteorological and emission conditions so that further modeling studies of pollutant formation mechanisms, emission control strategies, and human exposure and health risk assessment are based on a solid foundation.

In this study, a yearlong (2013) air quality simulation using a WRF/CMAQ system was conducted to provide detailed temporal and spatial distribution of $O_3$ and PM concentrations as well as $PM_{2.5}$ chemical composition in China. The publicly available observation data obtained from a total of 422 air monitoring sites in 60 major cities in China were used to provide a thorough evaluation of the model performance in the entire year. The modeled spatial and temporal concentrations of $O_3$ and $PM_{2.5}$ from this study will be used in subsequent studies to investigate the interaction between $O_3$ and PM pollution during high pollution events, the formation mechanism of secondary inorganic and organic aerosols and the population exposure and health risk.

## 2. Method

### 2.1 Model description

The CMAQ model applied in this study is based on CMAQ v5.0.1. Changes were made to the original CMAQ to improve the capability of the model in predicting secondary inorganic and organic aerosol, including 1) a modified SARPC-11 gas phase photochemical mechanism to provide more detailed treatment of isoprene oxidation chemistry (Ying et al., 2015) , 2) pathways of secondary organic aerosol (SOA) formation from surface controlled reactive uptake of dicarbonyls, isoprene epoxydiol (IEPOX) and methacrylic acid epoxide (MAE) (Li et al., 2015a; Ying et al., 2015), 3) vapor wall-loss corrected SOA yields (Zhang et al., 2014c), and 4) heterogeneous

reactions of $NO_2$ and $SO_2$ on particle surface to form secondary nitrate and sulfate (Ying et al., 2014a). More details of these changes can be found in the cited references and the references therein, thus only a short summary of the changes are provided below.

The isoprene mechanism in the original SAPRC-11 with standard lumping (Carter and Heo, 2012) was replaced by the detailed isoprene oxidation chemistry as used by Lin et al. (2013) to predict the formation of IEPOX and MAE in the gas phase. A precursor tracking scheme was implemented in the modified SAPRC-11 to track the glyoxal (GLY) and methylglyoxal (MGLY) formation from multiple biogenic and anthropogenic precursors. The surface controlled reactive uptake of SOA precursors is considered non-reversible, with constant uptake coefficients for GLY and MGLY as used by Fu et al. (2008) and an acidity dependent uptake coefficient for IEPOX and MAE as described by Li et al. (2015a). The original SOA yields for toluene and xylene under high NOx concentrations based on Ng et al. (2007) were replaced with the higher toluene yield reported by Hildebrandt et al. (2009). This update has been applied by Ying et al. (2014a) to study SOA formation in Mexico City. All SOA yields were then corrected by the average bias due to wall loss as reported in Table 1 of Zhang et al. (2014). A modeling study of SOA formation in Eastern US reported by Ying et al. (2015) shows that negative bias in predicted organic carbon (OC) concentrations reported in previous studies have been significantly reduced. Formation of sulfate and nitrate due to heterogeneous reactions on particle surface is also modeled as a reactive uptake process. The reactive surface uptake coefficients of $SO_2$ and $NO_2$ on particle surface were taken from Ying et al. (2014a) and Zheng et al. (2015), respectively.

## 2.2 Model application

The updated CMAQ model was applied to simulate $O_3$ and particulate air pollution using a 36-km horizontal resolution domain that covers China and surrounding countries in East Asia (Figure 1). The meteorological inputs were generated using WRF v3.6.1 with initial and boundary conditions from the NCEP FNL Operational Model Global Tropospheric Analyses dataset. Detailed WRF model configurations have been described by Zhang et al. (2012).

Multi-resolution Emission Inventory for China (MEIC) (0.25×0.25°) developed by Tsinghua University (http://www.meicmodel.org) was used for the monthly anthropogenic emissions from China. MEIC (V1.0) is the new version of emission inventory in China including improvements such as a unit-based emission inventory for power plants (Wang et al., 2012) and cement plants (Lei et al., 2011), a high-resolution county-level vehicle emission inventory (Zheng et al., 2014), and a non-methane VOC mapping approach for different chemical mechanisms (Li et al., 2014b). MEIC provides speciated VOC emissions for the SAPRC-07 mechanism with standard lumping (Carter, 2010). As the definitions of explicit and lumped primary VOCs have not changed from SAPRC-07 to SAPRC-11, these VOC emissions were directly used to drive SAPRC-11. Total $PM_{2.5}$ mass emissions and emissions of primary organic carbon (POC) and elemental carbon (EC) were also provided by MEIC directly. Emissions of trace metals needed by the version 6 of the aerosol module in CMAQ (AERO6) were generated using averaged speciation profiles adapted from the U.S. Environmental Protection Agency (EPA) SPECIATE database for each MEIC source category. Emissions from other countries and regions rather than China in the domain were filled with data generated from the gridded 0.25°×0.25° resolution Regional Emission inventory in ASia version 2 (REAS2) (Kurokawa et al., 2013). Details of the REAS2 emission

processing are described by Qiao et al. (2015). Detailed information about spatial and temporal allocation can also be found in the papers cited above.

Biogenic emissions were generated using the Model for Emissions of Gases and Aerosols from Nature (MEGAN) v2.1. The leaf area index (LAI) was based on the 8-day Moderate Resolution Imaging Spectroradiometer (MODIS) LAI product (MOD15A2) and the plant function types (PFTs) were based on the PFT files used in the Global Community Land Model (CLM 3.0). For more details of the biogenic emission processing, the readers are referred to Qiao et al. (2015). Open biomass burning emissions were generated from the Fire INventory from NCAR (FINN), which is based on satellite observations (Wiedinmyer et al., 2011). Dust and sea salt emissions were generated in line during the CMAQ simulations. In this updated CMAQ model, dust emission module was updated to be compatible with the 20-category MODIS land use data (Hu et al., 2015a). Initial and boundary conditions were based on the default vertical distributions of concentrations that represent clean continental conditions as provided by the CMAQ model. The impact of initial conditions was minimal as the results of the first five days of the simulation were excluded in the analyses.

## 3. Results

### 3.1 Meteorology validation

Meteorological factors are closely related to transport, transformation, and deposition of air pollutants (Hu et al., 2014b; Jacob and Winner, 2009; Tao et al., 2014a; Zhang et al., 2015b). Although the WRF model has been widely used to provide meteorological inputs for CTMs, the performance varies when applying to different domains, episodes, and with different model settings. Thus, the validation of model performance on meteorological conditions is important in assuring the accuracy of air quality predictions. Observation data from the National Climate Data Center (NCDC) was used to validate the model predictions of temperature (T2) and relative humidity (RH) at 2m above surface, and wind speed (WS) and wind direction (WD) at 10m above surface. Within the domain, there are ~1200 stations shown as purple dots in Figure 1. Model performance statistics of mean observation (OBS), mean prediction (PRE), mean bias (MB), gross error (GE) and root mean square error (RMSE) based on the observations and WRF predictions at the grid cells where the stations are located are shown in Table 1. The table also shows the benchmarks suggested by Emery et al. (2001) for the MM5 model in the East US with 4-12km grid resolution.

The WRF model predicts slightly higher T2 in winter and lower T2 in other seasons than the observations. The MB values for June, July, and September to December are within the benchmark, but the GE values of T2 are generally larger than the benchmark. The GE values of WS meet the benchmark in all months, but WS is over-predicted, as indicated by the positive MB values. The MB values meet the benchmark in January, June and August, and RMSE values are within the benchmark in June, July, and August. MB values of WD are within the benchmark of ±10 degree for four months. February, November, and December are the months with largest MB values. All GE values of WD are about 50% larger than the benchmark. RH is generally under-predicted except for July and August. The performance in this study is comparable to other studies using WRF in China (Hu et al., 2015a; Wang et al., 2010; Wang et al., 2014b; Ying et al., 2014b; Zhang et al., 2012), despite the differences in model, resolution, and study region in different

studies. Generally, the WRF model has acceptable performance on meteorological parameters. It should be noted that there is a study showing better WRF performance (Zhao et al., 2013a). However, it is difficult to compare since different model settings, simulation episodes, number of observation stations were used.

### 3.2 Model performance of $O_3$ and $PM_{2.5}$

Hourly observations of air pollutants from March to December 2013 were obtained from the publishing website of China National Environmental Monitoring Center (http://113.108.142.147:20035/emcpublish/). A total of 422 stations in 60 cities (see Figure 1 for the location of the cities) including the capital cities of all 31 provinces were obtained. Concentrations of pollutants in difference regions of China exhibit large variations due to diverse climates, topography, and emission sources. Aiming to identify the model strength and weakness in difference regions of China, model performance was evaluated separately for different regions. The regions and names of these cities are listed in Table 2. Automated quality control measures were taken to remove data points with observed $O_3$ concentrations greater than 250 ppb, $PM_{2.5}$ concentrations greater than 1500 $\mu g\ m^{-3}$, and points with standard deviation less than 5 ppb or 5 $\mu g\ m^{-3}$ in 24 hours.

#### *3.2.1 $O_3$ model performance*

Table 3 shows the model performance statistics of gaseous pollutants (1h peak $O_3$ ($O_3$-1h), 8h peak $O_3$ ($O_3$-8h), and hourly CO, $NO_2$, and $SO_2$), $PM_{2.5}$, and $PM_{10}$. Mean observations, mean predictions, mean fractional bias (MFB), mean fractional error (MFE), mean normalized bias (MNB) and mean normalized error (MNE) of hourly concentrations are calculated for each month from March to December 2013. Only $O_3$-1h or $O_3$-8h concentrations greater than 30 ppb were included in the analysis. A cutoff concentration of 40 or 60 ppb is suggested by the U.S. EPA (EPA, 2005). A lower cutoff of 30 ppb is chosen in this study considering the monitoring sites are all located in urban areas and higher $O_3$ concentrations generally occurs in downwind of urban areas. The overall model performance on $O_3$-1h and $O_3$-8h meets the model performance criteria suggested by U.S. EPA (2005) in all months, except in March and April for $O_3$-1h and June for $O_3$-8h. MNE of $O_3$-1h in June and July slightly exceeds the criteria, although MNB meets the criteria. MNB of $O_3$-8h in May exceeds the criteria, but MNE meets the criteria. The relatively small MNB/MNE and MFB/MFE in most of months indicate that $O_3$-1h and $O_3$-8h are well captured.

Model performance of $O_3$-1h and $O_3$-8h in different regions is illustrated in Table 4. Model performance meets the criteria in four regions, i.e., North China Plain (NCP), Yangtze River Delta (YRD), Pearl River Delta (PRD), and Northeast (NE). Relatively poor performance is identified in the Sichuan Basin (SCB), Central (CEN), and Northwest (NW) regions. $O_3$-1h and $O_3$-8h concentrations are slightly under-predicted in YRD and PRD, but over-predicted in all other regions. Model performance in regions other than NCP and YRD should be interpreted with care due to limited number of cities to sufficiently represent the entire region.

Figure 2 compares the predicted monthly averaged diurnal variations of $O_3$ concentrations with observations for all the 60 cities. For a city with multiple stations, observations and predictions are matched at individual station level and the averaged observations and predictions are used to

represent the concentrations for the city. Some cities, such as Beijing, exhibit substantial diurnal variations, especially in summer; and others, such as Lasa, exhibit small diurnal variations. Overall, the model successfully reproduces the monthly average diurnal variation in most cities, even though model performance among cities in the same region can be quite different. For example, in NE, the monthly averaged predictions agree well with observations in Shenyang and Changchun but are higher in Dalian, a coastal city, in summer months. In NCP, the model well predicts $O_3$ concentrations with slight over-prediction at a few cities, especially in the summer months, which agrees with the better hourly $O_3$ model performance shown in Tables 3 and 4. In YRD, the monthly diurnal variations of $O_3$ are also well predicted. Obvious under-prediction of summer peak $O_3$ at Zhoushan and Wenzhou are likely caused by underestimation of emissions in these port cities, although uncertainty in meteorology might also play a role. At PRD, $O_3$ is slightly underestimated in Guangzhou and Shenzhen for summer and fall months but well estimated in Zhuhai. In all three cities in the PRD region, $O_3$ concentrations are higher in the spring and fall months, and the model correctly captures this trend. In SCB, the model correctly predicts the higher spring $O_3$ concentrations in Chengdu but over-predicts spring $O_3$ concentrations in Chongqing. Summer $O_3$ concentrations are well predicted at both cities. For CEN, $O_3$ predictions are higher than observations in Zhengzhou and Hefei, but agree well with observations in other cities. In NW, the observed $O_3$ concentrations are much lower and are generally over-predicted all year except for Xi'an and Wulumuqi with good performance in summer.

Figure 3 shows the comparison of predicted and observed monthly averaged $O_3$-1h and $O_3$-8h concentrations at typical cities of major regions in China: Beijing for NCP, Shanghai for YRD, Guangzhou for PRD, Xi'an for NW, Shenyang for NE, and Chongqing for SCB. In Beijing, the monthly variations of both $O_3$-1h and $O_3$-8h, low in winter months and high in summer months, are well captured by model. The model slightly over-predicts $O_3$ concentrations from June to December except for August. In Shanghai, both $O_3$-1h and $O_3$-8h are underestimated by 5-10 ppb, but all observations are within the range of concentrations in the $3\times3$ grid cells surrounding the city center of Shanghai. In Guangzhou, $O_3$ concentrations vary slightly over months. $O_3$-1h is under-predicted especially in summer and fall months. $O_3$-8h predictions are closer to the observations. In Xi'an, the model well predicts the $O_3$-1h and $O_3$-8h concentrations in July, August, and September while over-predicts all other months by up to 20 ppb. In Shenyang, the trend of $O_3$-1h and $O_3$-8h are well reproduced with less than 5ppb differences for all the months. In Chongqing, over-prediction occurs in spring, fall, and winter while under-prediction occurs in summer.

### *3.2.2 $PM_{2.5}$ model performance*

$PM_{2.5}$ model performance in different months and regions are also illustrated in Table 3 and Table 4, respectively. The model performance statistics of MFB and MFE of hourly $PM_{2.5}$ concentrations meet the US EPA criteria in all months. Negative MFB is found in all months, indicating the model under-predicts the $PM_{2.5}$ concentrations. Model performance is better in March, September, November and December, with MFB less than 0.3. The bias is relatively larger in April, May, June, July and October, with MFB over 0.4. $PM_{10}$ is largely underestimated and is very likely to due to underestimation of dust emissions from both natural sources as well as human activities.

Model performance of $PM_{2.5}$ in different regions is also different. The model significantly under-predicts $PM_{2.5}$ in the NW and the Other (mostly Southwest cities) regions. Especially in the NW region, MFB value is -0.75 and MFE value is 0.88. $PM_{2.5}$ in all the other regions meets the performance criteria. Although most regions meet the model performance criteria in this study, under-prediction of $PM_{2.5}$ concentrations are found in all regions (except SCB), as indicated by the large negative MFB values. $PM_{10}$ has similar performance in various regions.

Figure 4 illustrates the comparison of predicted and observed monthly averaged $PM_{2.5}$ concentrations for all the 60 cities. In NE, the predictions agree well with observations in summer months. Concentrations in fall and winter months are under-predicted, except for Dalian, where the all values are well reproduced. In NCP, the annual trends at most cities are well captured. The model trends to under-predict spring and summer concentrations and over predict December concentrations. The coastal city, Qingdao, is unique with under-prediction in summer and good estimation in other months. In YRD, the model well produces $PM_{2.5}$ for all the months at most sites except in coastal cities (Zhoushan and Wenzhou) and mountainous cities (Quzhou and Lishui). In SCB, the model underestimates concentrations in the winter months in Chongqing but well estimates the concentrations in Chengdu except for March and April. In CEN, the seasonal trend is well captured at all cities but most cities show over-predicted concentrations in December. In NE, $PM_{2.5}$ is uniformly under-predicted. For Other regions, predictions agree with observations at the coastal cities (Fuzhou and Haikou) but concentrations in Lasa are largely under-predicted. The values closest to the observations in the $3\times3$ surrounding grid cells are similar to the predictions at city centers for most months with clear differences in October, November, and December at several cities. It indicates the higher contributions of primary PM, which has steeper concentration gradients than secondary PM, in winter months than in summer months.

Generally, the WRF/CMAQ modeling system with MEIC inventory well reproduces the $O_3$ and $PM_{2.5}$ concentrations in most regions for most months. Over-prediction of $O_3$ occurs at low concentrations in winter while under-prediction of $PM_{2.5}$ happens at low concentration range in summer and in cities in the NW region. The model performance on CO, $NO_2$, and $SO_2$ are also calculated and listed in Tables 3 and 4. There are no performance criteria for these pollutants, but the model performance are in the same ranges as compared to other studies in other countries/regions (Tao et al., 2014a). The model performance at different regions differs due to the differences in emission, topography, and meteorological conditions. The performance on these species can be used as indicator for emission uncertainties. The possible uncertainties are discussed in the *Discussion* section.

### 3.3 Seasonal variations and regional distribution of $O_3$ and $PM_{2.5}$

Figure 5 shows the predicted regional distribution of seasonal averaged $O_3$-1h and $O_3$-8h. In spring, highest $O_3$-1h concentration (~100 ppb) occurs in South Asia due to higher temperature, solar radiation and significant amount of emissions from open biomass burning activities (Kondo et al., 2004). Southern China has higher concentrations (~70 ppb) than Northern China (~50 ppb). However, in summer, NCP has the highest concentration of 80ppb while Southern China (and other regions) has lower concentrations of 50-60 ppb. In fall, most of the regions in China have $O_3$-1h concentrations of 50-60 ppb. In winter, NE China and NCP have $O_3$-1h concentrations lower than 30ppb while Southern China has the concentrations of 40-50 ppb. In addition to NCP in the summer, SCB is also another hot spot for ozone with high summer and wintertime $O_3$-1h

of ~100 ppb and 60-70 ppb, respectively. $O_3$-8h has similar spatial distribution patterns as $O_3$-1h for all seasons with lower concentrations (by 5~10 ppb).

Figure 6 shows the spatial distribution of seasonal averaged $PM_{2.5}$ concentrations together with the averaged wind vectors as the regional distribution of $PM_{2.5}$ is significantly influenced by wind patterns. In spring, the $PM_{2.5}$ concentrations in China reach approximately 50-70 μg $m^{-3}$ in Northern, Eastern, and Southern China except coastal provinces of Zhejiang, Fujian, and Guangdong. It is evident that the high concentrations are related to low wind speed. In summer, the areas with high $PM_{2.5}$ concentrations of ~50 μg $m^{-3}$ are limited to NCP and SCB while all other regions have concentrations of < 30 μg $m^{-3}$. Emissions brought to the NCP by the southerly wind, blockage of dispersion due to mountain ranges to the north and west, and secondary organic aerosol formed due to strong solar radiation are contributing factors for higher summer $PM_{2.5}$ in NCP. In fall, the high concentration regions are similar to those in spring but with higher concentrations of up to 100 μg $m^{-3}$ in NCP, YRD, CEN and SCB. In winter, high $PM_{2.5}$ concentrations are located in the NE, NCP, YRD, CEN and SCB regions. Seasonal average concentrations of more than 200 μg $m^{-3}$ occur in large portions of NCP, CEN, and SCB due to low wind speed and mixing height. Strong gradient exists between the high concentration regions and surrounding areas where wind is more lenient to pollutant dispersion.

Figure 7 shows the spatial distribution of seasonal averaged $PM_{2.5}$ components. All components show clear seasonal variations. For secondary inorganic components and anthropogenic primary components (EC and POA), concentrations are usually highest in winter and lowest in summer. Spring and fall concentrations are similar with slightly higher concentrations in fall. For EC and POA, this seasonal variation is largely driven by large increase in the emissions from residential sources in winter, as well as reduced ventilation that is often associated with winter stagnant conditions. For secondary inorganic components, gas phase formation rate of $HNO_3$ and $H_2SO_4$ decreases as temperature and solar radiation intensity decreases in fall and winter, leading to decrease in their formation from the homogeneous pathways. However, the amount of secondary $NO_3^-$ and $SO_4^{2-}$ from surface heterogeneous reactions of $SO_2$ and $NO_2$ increases as their emissions increases, and more particle surface area becomes available due to increase in primary PM concentrations. In addition, ammonium nitrate is preferentially partitioned into the particle phase under colder temperatures (Aw and Kleeman, 2003). In most regions with high concentrations, wintertime $NO_3^-$ concentrations are 150-200% higher than annual average concentrations, while $SO_4^{2-}$ and $NH_4^+$ concentrations are approximately 100-150% higher (see Figure 8). POA concentrations in winter are also approximately 100-150% higher in winter than the annual average, especially in northern part of China where residential heating is a significant source of $PM_{2.5}$ emissions. In provinces in southern China with warm temperature, winter POA is not significantly deviated from the annual mean (see Figure 8). Maximum concentrations of $NO_3^-$ and $SO_4^{2-}$ increase to beyond 50 μg $m^{-3}$ and $NH_4^+$ as high as 40 μg $m^{-3}$ in portions of NCP, CEN, YRD and SCB. This suggests that in large areas, secondary inorganic PM is the most significant contributor to elevated wintertime $PM_{2.5}$ concentrations. EC has limited spatial distribution since it is only directly emitted. Highest EC concentrations are in NCP, CEN and SCB. The EC concentrations are 10-15 μg $m^{-3}$ in winter but lower than 5 μg $m^{-3}$ in other seasons. POA concentrations are highly season dependent with the highest concentrations of ~30 μg $m^{-3}$ in NCP, CEN, SCB and NE occurring in winter.

SOA shows different seasonal variations from the secondary inorganic aerosol and anthropogenic primary PM components. In CEN and Eastern China, higher seasonal average SOA concentrations of 10-15 μg $m^{-3}$ occur in summer and winter, while in southern China similar levels of SOA occur in spring. The spring and summer high SOA concentrations are dominantly formed from biogenic isoprene emissions but winter SOA is mainly formed from semi-volatile oxidation products of anthropogenic aromatic compounds. Details of SOA formation and composition will be discussed in a separate paper. "Other" components are primary $PM_{2.5}$ including most part of dust. The concentrations are high in spring, fall and winter. In summary, secondary components have more boarder distribution than primary components. $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and POA are the most important aerosol components based on their absolute concentrations.

It should be noted that the simulated spatiotemporal distribution of $PM_{2.5}$ and its chemical composition is affected by the temporally and spatially variant biases of $PM_{2.5}$. In summer $PM_{2.5}$ is more under-predicted when the concentrations are lower, therefore the actual seasonal variation of $PM_{2.5}$ is likely weaker the predictions. $PM_{2.5}$ is more under-predicted in NW where the concentrations are lower, therefore the actual spatial difference between NW and eastern China region (i.e., NCP, YRD, etc.) is also likely weaker. The spatiotemporal distribution of $PM_{2.5}$ chemical composition is expected to be affected similarly but needed to be confirmed with detailed $PM_{2.5}$ composition observations. The biases of $O_3$ exhibit much less variation temporally and spatially, so the predicted spatiotemporal distribution of $O_3$ is more accurate than $PM_{2.5}$.

### 3.4 Temporal variation of $PM_{2.5}$ components in representative cities

Temporal variations of $PM_{2.5}$ components are also shown at typical cities in different regions as in Figure 9. The total $PM_{2.5}$ concentrations in Beijing are high in winter and low in summer with the peak of ~150 μg $m^{-3}$ in January. EC contributions are ~5-10% in winter but less than 5% in other seasons. POA has similar pattern as EC but contributions can be ~35% in winter and ~20% in summer. SOA contributions are high in summer with the peak of ~30% in August and very low in winter. $SO_4^{2-}$ and $NO_3^-$ are the top two largest contributors with comparable contributions all the time. $NH_4^+$ can be as high as ~20% in January and only ~10% in summer. Other components ("Other", mostly oxides of crustal elements and other trace metals) contribute up to 15% in some months. In Shanghai, the monthly averaged concentrations are highest in winter and decrease gradually from spring to fall. Five out of the 12 months are over the Chinese Ambient Air Quality Standards (CAAQS) Grade II standard for 24-hour average $PM_{2.5}$ (75 μg $m^{-3}$, simply Grade II standard hereafter). EC and POA have similar pattern with a total contribution of 20% in most months. $SO_4^{2-}$, $NO_3^-$, and $NH_4^+$ contribute to more than 70% from November to June and less than 50% in other months, while the contribution of SOA increases significantly to as much as 40% in the summer months. The relative contributions of the "Other" components are about 2 times of those in Beijing (15% to 30%). In Guangzhou, the $PM_{2.5}$ concentrations are lower than Beijing and Shanghai. Predicted $PM_{2.5}$ concentrations are all within the Grade II standard in China. Although the contribution of SOA is higher, $SO_4^{2-}$, $NO_3^-$, and $NH_4^+$ are still the major components with more than 60% contribution all over the year.

In Xi'an, the largest city in NW, the differences in $PM_{2.5}$ at winter and other months are significant. In winter, the total $PM_{2.5}$ concentrations are 150-180 μg $m^{-3}$ with POA, $SO_4^{2-}$, $NO_3^-$, and

$NH_4^+$ as major components. In Shenyang, a NE city, the $PM_{2.5}$ concentrations are ~250 μg $m^{-3}$ in January followed by ~200 μg $m^{-3}$ in February and ~150 μg $m^{-3}$ in December. The extremely high concentrations are related to winter residential heating or uncontrolled open biomass (such as straw) burning as can be indicated by the elevated emissions from residential sources. For other seasons, contributions of other components are much lower but contribution of SOA increases to more than 20% (~10 μg $m^{-3}$) in June, likely due to increased biogenic emissions in the densely forested regions in the NE. In Chongqing, located in Sichuan basin, monthly average reaches as high as 230 μg $m^{-3}$ in January due to increased atmospheric stability. Spring, summer and fall months have much lower $PM_{2.5}$ concentrations especially for July, when the $PM_{2.5}$ is lower than 50 μg $m^{-3}$.

One of the questions that remain unclear is whether secondary PM formation is enhanced during the high pollution days or high pollution events are simply caused by enhanced emissions and reduced dilution due to stagnant conditions. As an attempt to address this question, Figure 10 shows the comparison of relative contributions of $PM_{2.5}$ components in episode days (>= the Grade II standard of 75 μg $m^{-3}$) and non-episode days. In Guangzhou, there are no episode days predicted, thus only Beijing, Shanghai, Xi'an, Shenyang and Chongqing are included in Figure 10. In all cities, the minimum episode-day averaged concentration occurs in summer while the maximum concentration occurs in winter. In most cities and in most seasons, episode days have larger contributions of secondary components (SOA, $SO_4^{2-}$, $NO_3^-$, and $NH_4^+$, 69.8% on episode days vs. 59.9% on non-episode days) and lower contributions of primary components (EC, POA and Other, 30.2% on episode days vs 40.1% on non-episode days). Some cities show much drastic differences in secondary PM contributions between episode and non-episode days. For example, contribution of secondary PM in Xi'an increases from 40% on non-episode days to more than 60% on episode days in winter. Other cities, such as Chongqing, show less difference in the relative contributions of secondary PM between episode and non-episode days. While most of the secondary PM increase is due to enhanced formation of secondary inorganic components, the contribution of SOA to total PM is significantly higher than that on non-episode days in summer Beijing. This suggests that enhanced SOA formation could also play a significant role in summer PM pollution events of urban areas. In conclusion, in most cities in most seasons, episode days have more rapid formation of secondary PM components than accumulation of primary pollutants due to unfavorable weather conditions. This also suggests that controlling the emissions of secondary PM precursors needs to be considered in designing emission control strategies as in many conditions it can be more effective in reducing PM concentrations.

## 4. Discussion

Model predicted concentrations of $O_3$ and $PM_{2.5}$ are evaluated by comparing to ground-level observations at 422 stations in 60 cities in China for ten months in 2013. Predicted concentrations generally agree well with observations, with the model performance statistics meeting the criteria in most of the regions and months. Relatively large bias in model predicted concentrations is found in certain regions in certain months/episodes. Model bias is mainly attributed to uncertainties associated with meteorological fields, emissions, model treatment and configurations. Further studies are still needed to continue improving the model capability in accurately predicting air quality in China.

The WRF model performance in this study is comparable to other studies (Hu et al., 2015a; Wang et al., 2010; Wang et al., 2014b; Ying et al., 2014b; Zhang et al., 2012), but a better WRF performance was reported in Zhao et al. (2013a). Mesoscale meteorological modeling studies are also needed to improve the WRF model capability in China. In this study, some meteorological parameters are biased, for example ground-level wind speed is consistently over-predicted and RH is more biased low in winter months (Table 1). A previous study has revealed that air pollution levels are associated with these parameters in highly polluted regions in China (Wang et al., 2014c). It is also demonstrated that bias in predicted meteorological parameters by WRF contributes to bias in $PM_{2.5}$ prediction (Hu et al., 2015c; Zhang et al., 2014a; Zhang et al., 2014b). A companion study is undergoing to evaluate the sensitivity of predictions to meteorological fields.

Uncertainties associated with emission inventory often are the major factor leading to bias in model predictions. The overall good model performance in most regions indicates general accuracy of the MEIC inventory. However, larger negative bias in CO, $NO_2$, and $SO_2$ in NW (Table 4) suggests that anthropogenic emissions, including primary $PM_{2.5}$ are severely under-estimated in this region. Similarly, under-predictions of $PM_{2.5}$ in Lasa are also likely due to under-predictions of anthropogenic emissions, mostly likely those from residential sources. Studies have suggested that dust contributes significantly to $PM_{2.5}$ in NW (Li et al., 2014a; Shen et al., 2009). The current estimation of dust from wind erosion of natural soil surfaces in the NW is approximately 20 $\mu g\ m^{-3}$ in spring and lower than 10 $\mu g\ m^{-3}$ in other seasons. This relatively low estimation of $PM_{2.5}$ in the NW of China generally agrees with the most recent global long term $PM_{2.5}$ estimation based on satellite AOD measurements (Battelle Memorial Institute and Center for International Earth Science Information Network - CIESIN - Columbia University, 2013; de Sherbinin et al., 2014). Emissions of dust from other sources in the urban/rural areas, such as paved and unpaved road and construction activities could be a more important factor that leads to under-predictions of mineral PM components in the NW cities. Both activity data and emission factors used to generate these area emissions should be examined carefully. Source apportionment studies based on receptor-oriented techniques should be used to differentiate the contributions from these different dust sources to further constrain the uncertainties in dust emissions.

Another important source of under-prediction of $PM_{2.5}$ is SOA, especially in the summer when the biases in $PM_{2.5}$ predictions are larger and more SOA is expected to form due to higher VOCs emissions and higher atmospheric reactivity. While significant progresses have been made to improve model predictions and the SOA module used in the current study has incorporated many of the newly found SOA formation pathways, the understanding of both gas phase and particle phase chemistry that lead to SOA formation is still very limited, and many experimental findings have yet been incorporated by the modeling community. To constrain the uncertainties in SOA predictions, speciated measurements of SOA tracers and gas phase VOC precursors are needed along with models with detailed chemical mechanisms to represent the species. While some VOC speciation data are available, more data in different regions and episodes are needed to improve both estimation of VOC emissions (Zhang and Ying, 2011) and model predictions of SOA.

Model grid resolution also contributes to the bias in predictions. The emissions are instantly mixed into $36 \times 36\ km^2$ grids after being released from sources. Some of the monitoring stations are located in urban areas near emission sources, such as traffic and industrial facilities, which could imply negative prediction biases when compared with modeled concentrations which represent average concentrations in a grid cell. Higher resolution modeling studies are believed to

more accurately capture the concentrations and to reveal finer scale spatial distribution of pollutants (Fountoukis et al., 2013; Gan et al., 2016; Joe et al., 2014; Stroud et al., 2011). The grid dilution effect theoretically has larger impact on CO and $SO_2$ than on $O_3$ and $PM_{2.5}$, because $O_3$ and secondary $PM_{2.5}$ components are often formed regionally and consequently have a more uniform spatial distribution.

## 5. Conclusion

In this study, $O_3$ and $PM_{2.5}$ in China during the entire year of 2013 is simulated using an updated WRF/CMAQ model system and anthropogenic emissions from MEIC. The WRF model predicts reasonable meteorological inputs for the CMAQ model. The comparison of predicted and observed hourly $O_3$, peak hour $O_3$, and daily and monthly averaged $PM_{2.5}$ at 60 cities shows that the current model can successfully reproduces the $O_3$ and $PM_{2.5}$ concentrations at most cities for most months of the year. Over-prediction of $O_3$ occurs at low concentration range in winter while under-prediction of $PM_{2.5}$ happens at low concentration range in summer. Spatially, the model has better performance in NE, NCP, Central YRD and SCB but significant under-prediction biases exist for the cities in the NW region. Strong seasonal variations of $PM_{2.5}$ exist and wind speed and direction play important roles in high $PM_{2.5}$ events. Secondary components have more boarder distribution than primary components. Contributions of secondary PM components increase during high PM events in a number of urban areas, suggesting that secondary PM formation rates are enhanced more than the accumulation rate of primary pollutants. Overall, $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and POA are the most important $PM_{2.5}$ components. All components have the highest concentrations in winter except SOA. NCP, CEN and SCB have more severe $PM_{2.5}$ levels than YRD and PRD.

This study reports the detailed model performance of $O_3$ and $PM_{2.5}$ in China for an entire year with the public available observations nationwide in China. Although much needs to be done to improve the model performance, this study shows the capability of the model with MEIC emission in reproducing severe air pollution. The concentrations of $O_3$, $PM_{2.5}$ total mass and its chemical components from this study will be used in future studies to understand formation mechanisms of severe air pollution episodes, investigate the effectiveness of emission control strategies, and estimate human exposure to multiple pollutants for assessing health burden of air pollution in China.

## Acknowledgement

This project is partly funded by the Natural Science Foundation of Jiangsu Province (BK20150904 and BK20151041), Jiangsu Distinguished Professor Project (2191071503201), Jiangsu Six Major Talent Peak Project (2191071502101), the Startup Fund for Talent at NUIST (2243141501008) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), Jiangsu Key Laboratory of Atmospheric Environment Monitoring and Pollution Control of Nanjing University of Information Science and Technology, and Jiangsu Province Innovation Platform for Superiority Subject of Environmental Science and Engineering (No. KHK1201). We would like to thank the computation resources from the Texas A&M Supercomputing Facility (http://sc.tamu.edu/) for completing some of the model simulations reported in this study.

## References


Aw, J. and Kleeman, M.J., 2003. Evaluating the first-order effect of intraannual temperature variability on urban air pollution. Journal of Geophysical Research: Atmospheres, 108(D12): 4365.

Battelle Memorial Institute and Center for International Earth Science Information Network - CIESIN - Columbia University, 2013. Global Annual Average PM2.5 Grids from MODIS and MISR Aerosol Optical Depth (AOD). NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY.

Carter, W.P.L., 2010. Development of the SAPRC-07 chemical mechanism. Atmos. Environ., 44(3): 5324-5335.

Carter, W.P.L. and Heo, G., 2012. Development of revised SAPRC aromatics mechanisms. Final Report to the California Air Resources Board, Contracts No. 07-730 and 08-326, April 12, 2012. .

de Sherbinin, A., Levy, M., Zell, E., Weber, S. and Jaiteh, M., 2014. Using Satellite Data to Develop Environmental Indicators. Environmental Research Letters, 9(8).

Emery, C., Tai, E. and Yarwood, G., 2001. Enhanced meteorological modeling and performance evaluation for two texas episodes, Novato, CA.

EPA, U.S., 2005. Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-hour Ozone NAAQS. EPA-454/R-05-002.

Fountoukis, C. et al., 2013. Impact of grid resolution on the predicted fine PM by a regional 3-D chemical transport model. Atmospheric Environment, 68: 24-32.

Fu, T.M. et al., 2008. Global budgets of atmospheric glyoxal and methylglyoxal, and implications for formation of secondary organic aerosols. J Geophys Res-Atmos, 113(D15).

Gan, C.-M. et al., 2016. Assessment of the effects of horizontal grid resolution on long-term air quality trends using coupled WRF-CMAQ simulations. Atmospheric Environment, 132: 207-216.

Gao, Y., Zhao, C., Liu, X.H., Zhang, M.G. and Leung, L.R., 2014. WRF-Chem simulations of aerosols and anthropogenic aerosol radiative forcing in East Asia. Atmos Environ, 92: 250-266.

Hildebrandt, L., Donahue, N.M. and Pandis, S.N., 2009. High formation of secondary organic aerosol from the photo-oxidation of toluene. Atmos. Chem. Phys., 9(9): 2973-2986.

Hu, J., Wang, Y., Ying, Q. and Zhang, H., 2014a. Spatial and temporal variability of PM2.5 and PM10 over the North China Plain and the Yangtze River Delta, China. Atmospheric Environment, 95(0): 598-609.

Hu, J. et al., 2015a. Source contributions and regional transport of primary particulate matter in China. Environmental Pollution, 207: 31-42.

Hu, J., Ying, Q., Wang, Y. and Zhang, H., 2015b. Characterizing multi-pollutant air pollution in China: Comparison of three air quality indices. Environ Int, 84: 17-25.

Hu, J. et al., 2015c. Long-term particulate matter modeling for health effect studies in California - Part I: model performance on temporal and spatial variations. Atmos Chem Phys, 15: 3445-3461.

Hu, X.-M. et al., 2014b. Impact of the Loess Plateau on the atmospheric boundary layer structure and air quality in the North China Plain: A case study. Science of The Total Environment, 499: 228-237.

Jacob, D.J. and Winner, D.A., 2009. Effect of climate change on air quality. Atmospheric Environment, 43(1): 51-63.

Joe, D.K. et al., 2014. Implementation of a high-resolution Source-Oriented WRF/Chem model at the Port of Oakland. Atmospheric Environment, 82(0): 351-363.

Kondo, Y. et al., 2004. Impacts of biomass burning in Southeast Asia on ozone and reactive nitrogen over the western Pacific in spring. Journal of Geophysical Research: Atmospheres, 109(D15): n/a-n/a.

Kurokawa, J. et al., 2013. Emissions of air pollutants and greenhouse gases over Asian regions during 2000–2008: Regional Emission inventory in ASia (REAS) version 2. Atmos. Chem. Phys., 13(21): 11019-11058.

Lei, Y., Zhang, Q., Nielsen, C. and He, K., 2011. An inventory of primary air pollutants and CO2 emissions from cement production in China, 1990–2020. Atmospheric Environment, 45(1): 147-154.

603 Li, J. et al., 2015a. Modeling regional secondary organic aerosol using the Master Chemical Mechanism.
604 Atmos Environ, 102: 52-61.
605 Li, J. et al., 2014a. Comparison of abundances, compositions and sources of elements, inorganic ions and
606 organic compounds in atmospheric aerosols from Xi'an and New Delhi, two megacities in China
607 and India. Science of The Total Environment, 476–477(0): 485-495.
608 Li, M. et al., 2014b. Mapping Asian anthropogenic emissions of non-methane volatile organic
609 compounds to multiple chemical mechanisms. Atmos. Chem. Phys., 14(11): 5617-5638.
610 Li, X. et al., 2015b. Source contributions of urban PM2.5 in the Beijing–Tianjin–Hebei region: Changes
611 between 2006 and 2013 and relative impacts of emissions and meteorology. Atmos Environ, In
612 press.
613 Lin, Y.-H. et al., 2013. Epoxide as a precursor to secondary organic aerosol formation from isoprene
614 photooxidation in the presence of nitrogen oxides. Proceedings of the National Academy of
615 Sciences, 110(17): 6718-6723.
616 Liu, X.H. et al., 2010. Understanding of regional air pollution over China using CMAQ, part I performance
617 evaluation and seasonal variation. Atmos Environ, 44(20): 2415-2426.
618 Liu, X.Y., Zhang, Y., Zhang, Q. and He, M.B., 2016. Application of online-coupled WRF/Chem-MADRID in
619 East Asia: Model evaluation and climatic effects of anthropogenic aerosols. Atmos Environ, 124:
620 321-336.
621 Menon, S. et al., 2008. Aerosol climate effects and air quality impacts from 1980 to 2030. Environmental
622 Research Letters, 3(2): 024004.
623 Ng, N.L. et al., 2007. Secondary organic aerosol formation from m-xylene, toluene, and benzene.
624 Atmospheric Chemistry and Physics, 7: 3909-3922.
625 Poschl, U., 2005. Atmospheric aerosols: Composition, transformation, climate and health effects.
626 Angewandte Chemie-International Edition, 44(46): 7520-7540.
627 Pui, D.Y.H., Chen, S.-C. and Zuo, Z., 2014. PM2.5 in China: Measurements, sources, visibility and health
628 effects, and mitigation. Particuology, 13(0): 1-26.
629 Qiao, X. et al., 2015. Modeling dry and wet deposition of sulfate, nitrate, and ammonium ions in
630 Jiuzhaigou National Nature Reserve, China using a source-oriented CMAQ model: Part I. Base
631 case model results. Sci Total Environ, 532: 831-839.
632 Shen, Z. et al., 2009. Ionic composition of TSP and PM2.5 during dust storms and air pollution episodes
633 at Xi'an, China. Atmospheric Environment, 43(18): 2911-2918.
634 Stroud, C.A. et al., 2011. Impact of model grid spacing on regional- and urban- scale air quality
635 predictions of organic aerosol. Atmos. Chem. Phys., 11(7): 3107-3118.
636 Sun, Y. et al., 2014. Investigation of the sources and evolution processes of severe haze pollution in
637 Beijing in January 2013. Journal of Geophysical Research: Atmospheres, 119(7): 2014JD021641.
638 Tao, J. et al., 2014a. PM2.5 pollution in a megacity of southwest China: source apportionment and
639 implication. Atmos. Chem. Phys., 14(16): 8679-8699.
640 Tao, M. et al., 2014b. Formation process of the widespread extreme haze pollution over northern China
641 in January 2013: Implications for regional air quality and climate. Atmospheric Environment,
642 98(0): 417-425.
643 Wang, D. et al., 2014a. Source contributions to primary and secondary inorganic particulate matter
644 during a severe wintertime PM2.5 pollution episode in Xi'an, China. Atmospheric Environment,
645 97(0): 182-194.
646 Wang, L.T. et al., 2010. Assessment of air quality benefits from national air pollution control policies in
647 China. Part I: Background, emission scenarios and evaluation of meteorological predictions.
648 Atmospheric Environment, 44: 3442-3448.
649 Wang, L.T. et al., 2014b. The 2013 severe haze over southern Hebei, China: model evaluation, source
650 apportionment, and policy implications. Atmos Chem Phys, 14(6): 3151-3173.

651 Wang, S.W. et al., 2012. Growth in NOx emissions from power plants in China: bottom-up estimates and
652 satellite observations. Atmos. Chem. Phys., 12(10): 4429-4447.
653 Wang, S.X. et al., 2011. Verification of anthropogenic emissions of China by satellite and ground
654 observations. Atmos Environ, 45(35): 6347-6358.
655 Wang, Y., Ying, Q., Hu, J. and Zhang, H., 2014c. Spatial and temporal variations of six criteria air
656 pollutants in 31 provincial capital cities in China during 2013–2014. Environment International,
657 73(0): 413-422.
658 Wiedinmyer, C. et al., 2011. The Fire INventory from NCAR (FINN): a high resolution global model to
659 estimate the emissions from open burning. Geoscientific Model Development, 4: 625-641.
660 Ying, Q. et al., 2014a. Impacts of Stabilized Criegee Intermediates, surface uptake processes and higher
661 aromatic secondary organic aerosol yields on predicted PM2.5 concentrations in the Mexico City
662 Metropolitan Zone. Atmos Environ, 94(0): 438-447.
663 Ying, Q., Li, J. and Kota, S.H., 2015. Significant Contributions of Isoprene to Summertime Secondary
664 Organic Aerosol in Eastern United States. Environmental Science & Technology, 49(13): 7834-
665 7842.
666 Ying, Q., Wu, L. and Zhang, H., 2014b. Local and inter-regional contributions to PM2.5 nitrate and sulfate
667 in China. Atmospheric Environment, 94(0): 582-592.
668 Zhang, B., Wang, Y. and Hao, J., 2015a. Simulating aerosol–radiation–cloud feedbacks on meteorology
669 and air quality over eastern China under severe haze conditionsin winter. Atmos. Chem. Phys.,
670 15(5): 2387-2404.
671 Zhang, H. et al., 2014a. Evaluation of a seven-year air quality simulation using the Weather Research and
672 Forecasting (WRF)/Community Multiscale Air Quality (CMAQ) models in the eastern United
673 States. Science of The Total Environment, 473–474(0): 275-285.
674 Zhang, H., Hu, J., Kleeman, M. and Ying, Q., 2014b. Source apportionment of sulfate and nitrate
675 particulate matter in the Eastern United States and effectiveness of emission control programs.
676 Science of The Total Environment, 490(0): 171-181.
677 Zhang, H. et al., 2012. Source apportionment of PM2.5 nitrate and sulfate in China using a source-
678 oriented chemical transport model. Atmos. Environ., 62(0): 228-242.
679 Zhang, H., Wang, Y., Hu, J., Ying, Q. and Hu, X.-M., 2015b. Relationships between meteorological
680 parameters and criteria air pollutants in three megacities in China. Environmental Research,
681 140(0): 242-254.
682 Zhang, H. and Ying, Q., 2011. Secondary Organic Aerosol Formation and Source Apportionment in
683 Southeast Texas. Atmospheric Environment, 45(19): 3217-3227.
684 Zhang, X. et al., 2014c. Influence of vapor wall loss in laboratory chambers on yields of secondary
685 organic aerosol. Proceedings of the National Academy of Sciences, 111(16): 5802-5807.
686 Zhang, Y. et al., 2016. Application of WRF/Chem over East Asia: Part I. Model evaluation and
687 intercomparison with MM5/CMAQ. Atmos Environ, 124: 285-300.
688 Zhao, B. et al., 2013a. Environmental effects of the recent emission changes in China: implications for
689 particulate matter pollution and soil acidification. Environmental Research Letters, 8(2): 024031.
690 Zhao, B. et al., 2013b. Impact of national NOx and SO2 control policies on particulate matter pollution in
691 China. Atmos Environ, 77: 453-463.
692 Zheng, B. et al., 2014. High-resolution mapping of vehicle emissions in China in 2008. Atmos. Chem.
693 Phys., 14(18): 9787-9805.
694 Zheng, B. et al., 2015. Heterogeneous chemistry: a mechanism missing in current models to explain
695 secondary inorganic aerosol formation during the January 2013 haze episode in North China.
696 Atmos. Chem. Phys., 15(4): 2031-2049.

697

Table 1. Meteorology performance in all the months in 2013 (OBS, mean observation; PRE, mean prediction; MB, mean bias; GE, gross error; and RMSE, root mean square error). The benchmarks are suggested by Emery et al. (2001) for the MM5 model in the East US with 4-12km grid resolution. The values that do not meet the criteria are shaded.

| | | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | Bench-mark |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T2 (K) | OBS | 267.3 | 270.4 | 277.5 | 282.7 | 289.3 | 293.9 | 297.0 | 297.1 | 292.1 | 286.0 | 278.1 | 272.8 | |
| | PRE | 266.1 | 268.9 | 276.2 | 281.8 | 288.7 | 293.6 | 296.5 | 296.5 | 291.9 | 286.0 | 278.4 | 273.1 | |
| | MB | 1.2 | -1.4 | -1.3 | -0.8 | -0.7 | -0.3 | -0.5 | -0.6 | -0.2 | 0.0 | 0.3 | 0.3 | ≤ ±0.5 |
| | GE | 3.7 | 3.3 | 3.0 | 2.7 | 2.7 | 2.7 | 2.6 | 2.5 | 2.4 | 2.5 | 2.7 | 2.8 | ≤ 2.0 |
| | RMSE | 4.7 | 4.5 | 4.0 | 3.6 | 3.5 | 3.6 | 3.5 | 3.3 | 3.2 | 3.3 | 3.5 | 3.8 | |
| WS ($ms^{-1}$) | OBS | 3.0 | 3.5 | 3.7 | 3.8 | 3.6 | 3.3 | 3.4 | 3.2 | 3.3 | 3.4 | 3.5 | 3.5 | |
| | PRE | 3.2 | 4.8 | 4.8 | 4.8 | 4.4 | 3.8 | 4.0 | 3.8 | 4.0 | 4.4 | 4.6 | 4.7 | |
| | MB | 0.2 | 1.3 | 1.1 | 1.0 | 0.7 | 0.5 | 0.6 | 0.5 | 0.7 | 1.0 | 1.1 | 1.2 | ≤ ±0.5 |
| | GE | 1.3 | 2.0 | 1.9 | 1.9 | 1.7 | 1.53 | 1.6 | 1.5 | 1.6 | 1.7 | 1.9 | 1..9 | ≤ 2.0 |
| | RMSE | 2.6 | 2.6 | 2.5 | 2.4 | 2.2 | 2.0 | 2.0 | 1.9 | 2.1 | 2.3 | 2.4 | 2.5 | ≤ 2.0 |
| WD (°) | OBS | 187.5 | 212.0 | 205.0 | 202.4 | 187.3 | 171.2 | 187.0 | 190.6 | 174.8 | 183.0 | 216.0 | 216.4 | |
| | PRE | 209.9 | 229.1 | 220.4 | 216.8 | 198.5 | 175.8 | 200.8 | 203.4 | 171.4 | 182.1 | 236.5 | 234.0 | |
| | MB | 10.5 | 17.1 | 15.4 | 14.4 | 11.2 | 4.6 | 13.8 | 12.9 | -3.4 | -0.9 | 20.5 | 17.7 | ≤ ±10 |
| | GE | 46.3 | 47.7 | 46.7 | 44.8 | 46.2 | 49.4 | 46.6 | 47.4 | 47.5 | 45.6 | 44.8 | 46.6 | ≤ ±30 |
| | RMSE | 66.3 | 65.1 | 64.1 | 62.1 | 63.4 | 66.4 | 63.5 | 64.4 | 65.0 | 62.9 | 61.8 | 63.8 | |
| RH (%) | OBS | 64.9 | 78.9 | 69.5 | 67.1 | 64.3 | 68.7 | 70.8 | 70.4 | 6938 | 71.7 | 72.2 | 75.3 | |
| | PRE | 63.6 | 73.4 | 68.4 | 65.3 | 64.0 | 68.1 | 72.0 | 72.1 | 69.2 | 71.0 | 68.9 | 68.7 | |
| | MB | -1.4 | -5.6 | -1.1 | -1.8 | -0.3 | -0.5 | 1.2 | 1.7 | -0.6 | -0.7 | -3.3 | -6.5 | |
| | GE | 19.2 | 14.1 | 15.4 | 14.9 | 14.5 | 13.4 | 13.5 | 13.0 | 12.6 | 13.5 | 14.1 | 14.8 | |
| | RMSE | 21.2 | 18.3 | 19.4 | 18.9 | 18.6 | 17.4 | 17.3 | 16.6 | 16.3 | 17.4 | 18.4 | 19.8 | |

703 Table 2. List of the cities in different regions with available observations.

| **Region** | **City list** |
|---|---|
| Northeast (4 cities) | 1. Harbin, 2. Changchun, 3. Shenyang, 4. Dalian |
| North China Plain (NCP) (14) | 5. Chengde, 6. Beijing, 7. Qinhuangdao, 8. Tangshan, 9. Langfang, 10. Tianjin, 11. Baoding, 12. Cangzhou, 13. Shijiazhuang, 14. Hengshui, 15. Handan, 16. Jinan, 17. Qingdao, 28. Huhehaote |
| Yangtze River Delta (YRD) (20) | 21. Lianyungang, 22. Suqian, 23. Xuzhou, 24. Huai'an, 25. Taizhou, 26. Yangzhou, 27. Nanjing, 29. Nantong, 30. Suzhou, 31. Wuxi, 32. Shanghai, 33. Huzhou, 34. Hangzhou, 35. Jiaxing, 36. Shaoxing, 37. Zhoushan, 38. Wenzhou, 39. Jinhua, 40. Quzhou, 41. Lishui |
| Pearl River Delta (PRD) (3) | 46. Guangzhou, 47. Zhuhai, 60. Shenzhen |
| Central China (6) | 18. Taiyuan, 19. Zhengzhou, 20. Hefei, 43. Wuhan, 44. Nanchang, 45. Changsha |
| Northwest (5) | 54. Xi'an, 55. Yinchuan, 56. Lanzhou, 57. Xining, 58. Wulumuqi |
| Sichuan basin (SCB) (2) | 52. Chongqing, 53. Chengdu |
| Southwest+Other (6) | 42. Fuzhou, 48. Haikou, 49. Nanning, 50. Kunming, 51. Guiyang, 59. Lasa |

704

Table 3. Model performance on $O_3$-1h, $O_3$-8h, $PM_{2.5}$, $PM_{10}$, CO, $NO_2$, and $SO_2$ in March to December 2013 (OBS, mean observation; PRE, mean prediction; MFB, mean fractional bias; MFE, mean fractional error; MNB, mean normalized bias; MNE, mean normalized error). The performance criteria for $PM_{2.5}$ are suggested by EPA (2007), and the performance criteria for $O_3$ are suggested by EPA (2005). The values that do not meet the criteria are shaded.

| | | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | Criteria |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $O_3$-1h (ppb) | OBS | 53.96 | 57.73 | 65.37 | 67.72 | 65.7 | 68.3 | 60.73 | 57.97 | 49.18 | 46.53 | |
| | PRE | 58.09 | 61.76 | 66.91 | 67.82 | 63.23 | 66.47 | 59.5 | 54.92 | 45.66 | 42.09 | |
| | MFB | 0.08 | 0.09 | 0.05 | 0.01 | -0.01 | -0.01 | 0.01 | -0.03 | -0.05 | -0.09 | |
| | MFE | 0.29 | 0.27 | 0.25 | 0.3 | 0.29 | 0.28 | 0.27 | 0.26 | 0.27 | 0.32 | |
| | MNB | 0.16 | 0.17 | 0.11 | 0.1 | 0.06 | 0.06 | 0.07 | 0.03 | 0.01 | -0.01 | $\leqslant \pm 0.15$ |
| | MNE | 0.34 | 0.32 | 0.28 | 0.33 | 0.31 | 0.3 | 0.29 | 0.26 | 0.26 | 0.28 | $\leqslant 0.3$ |
| $O_3$-8h (ppb) | OBS | 50.4 | 47.44 | 52.59 | 54.36 | 51.79 | 54.03 | 48.63 | 48.03 | 40.31 | 38.92 | |
| | PRE | 48.81 | 51.49 | 57.86 | 59.58 | 54.05 | 58.07 | 50.64 | 48.48 | 40.6 | 40.7 | |
| | MFB | -0.05 | 0.07 | 0.1 | 0.08 | 0.03 | 0.06 | 0.04 | 0.01 | -0.01 | 0.01 | |
| | MFE | 0.29 | 0.24 | 0.24 | 0.28 | 0.26 | 0.26 | 0.25 | 0.24 | 0.25 | 0.27 | |
| | MNB | 0.03 | 0.13 | 0.16 | 0.16 | 0.09 | 0.12 | 0.1 | 0.06 | 0.03 | 0.07 | $\leqslant \pm 0.15$ |
| | MNE | 0.29 | 0.28 | 0.28 | 0.32 | 0.28 | 0.29 | 0.27 | 0.25 | 0.24 | 0.27 | $\leqslant 0.3$ |
| $PM_{2.5}$ ($\mu g\ m^{-3}$) | OBS | 81.68 | 62.07 | 60.12 | 60.83 | 45.52 | 47.1 | 56.08 | 85.69 | 88.93 | 123.73 | |
| | PRE | 66.12 | 43.24 | 39.28 | 41.6 | 31.31 | 39.07 | 52.24 | 56.09 | 80.21 | 126.83 | |
| | MFB | -0.24 | -0.4 | -0.47 | -0.41 | -0.48 | -0.31 | -0.21 | -0.42 | -0.17 | -0.07 | $\leqslant \pm 0.6$ |
| | MFE | 0.59 | 0.63 | 0.68 | 0.69 | 0.72 | 0.65 | 0.62 | 0.64 | 0.6 | 0.59 | $\leqslant 0.75$ |
| | MNB | 0.04 | -0.16 | -0.19 | -0.09 | -0.17 | -0.01 | 0.11 | -0.16 | 0.17 | 0.3 | |
| | MNE | 0.61 | 0.54 | 0.58 | 0.63 | 0.63 | 0.64 | 0.68 | 0.56 | 0.7 | 0.75 | |
| $PM_{10}$ ($\mu g\ m^{-3}$) | OBS | 151.39 | 121.56 | 111.90 | 96.95 | 79.90 | 85.04 | 98.27 | 136.02 | 150.27 | 178.78 | |
| | PRE | 74.72 | 52.48 | 45.37 | 46.58 | 35.59 | 44.63 | 57.53 | 65.12 | 90.22 | 136.26 | |
| | MFB | -0.59 | -0.73 | -0.79 | -0.68 | -0.78 | -0.65 | -0.54 | -0.65 | -0.48 | -0.34 | |
| | MFE | 0.74 | 0.83 | 0.89 | 0.82 | 0.88 | 0.79 | 0.73 | 0.77 | 0.72 | 0.63 | |
| | MNB | -0.31 | -0.43 | -0.45 | -0.35 | -0.44 | -0.35 | -0.24 | -0.36 | -0.16 | -0.04 | |
| | MNE | 0.56 | 0.58 | 0.62 | 0.62 | 0.63 | 0.59 | 0.60 | 0.59 | 0.64 | 0.62 | |
| CO (ppm) | OBS | 1.17 | 0.94 | 0.86 | 0.8 | 0.73 | 0.75 | 0.85 | 1.09 | 1.16 | 1.48 | |
| | PRE | 0.37 | 0.26 | 0.25 | 0.26 | 0.23 | 0.25 | 0.29 | 0.31 | 0.41 | 0.59 | |
| | MFB | -0.89 | -0.97 | -0.97 | -0.91 | -0.95 | -0.92 | -0.9 | -0.98 | -0.88 | -0.8 | |
| | MFE | 0.95 | 1.01 | 1 | 0.95 | 0.99 | 0.96 | 0.95 | 1.02 | 0.92 | 0.86 | |
| | MNB | -0.54 | -0.6 | -0.6 | -0.56 | -0.58 | -0.56 | -0.56 | -0.61 | -0.54 | -0.49 | |
| | MNE | 0.63 | 0.65 | 0.65 | 0.63 | 0.64 | 0.63 | 0.63 | 0.66 | 0.62 | 0.59 | |
| $NO_2$ (ppb) | OBS | 23.33 | 21.26 | 19.83 | 18.11 | 16.34 | 16.5 | 19.74 | 24.82 | 27.41 | 31.41 | |
| | PRE | 10.11 | 8.87 | 8.51 | 8.74 | 8.12 | 8.77 | 10.45 | 11.85 | 13.45 | 13.87 | |
| | MFB | -0.83 | -0.88 | -0.86 | -0.79 | -0.79 | -0.73 | -0.71 | -0.76 | -0.7 | -0.77 | |
| | MFE | 0.94 | 0.99 | 0.99 | 0.95 | 0.95 | 0.91 | 0.89 | 0.91 | 0.85 | 0.87 | |
| | MNB | -0.45 | -0.48 | -0.46 | -0.4 | -0.4 | -0.35 | -0.35 | -0.39 | -0.37 | -0.44 | |
| | MNE | 0.65 | 0.67 | 0.68 | 0.68 | 0.68 | 0.67 | 0.66 | 0.65 | 0.62 | 0.61 | |
| $SO_2$ (ppb) | OBS | 19.1 | 15.8 | 15.25 | 12.93 | 12.32 | 12.96 | 13.24 | 15.53 | 21.74 | 27.88 | |
| | PRE | 11.64 | 8.87 | 8.31 | 8.61 | 7.09 | 8.88 | 11.94 | 14.25 | 17.91 | 23.32 | |
| | MFB | -0.61 | -0.66 | -0.68 | -0.59 | -0.73 | -0.56 | -0.39 | -0.29 | -0.31 | -0.32 | |
| | MFE | 0.89 | 0.9 | 0.91 | 0.89 | 0.98 | 0.89 | 0.84 | 0.78 | 0.82 | 0.83 | |
| | MNB | -0.14 | -0.23 | -0.23 | -0.11 | -0.22 | -0.08 | 0.23 | 0.25 | 0.29 | 0.31 | |
| | MNE | 0.79 | 0.74 | 0.76 | 0.8 | 0.81 | 0.82 | 1 | 0.95 | 1.01 | 1.03 | |

711 Table 4. Model performance on $O_3$-1h, $O_3$-8h, $PM_{2.5}$, $PM_{10}$, CO, $NO_2$, and $SO_2$ in different re-
712 gions during March to December, 2013. The values that do not meet the criteria are shaded.

| | | NCP | YRD | PRD | SCB | NE | CEN | NW | Other |
|---|---|---|---|---|---|---|---|---|---|
| $O_3$-1h (ppb) | OBS | 65.18 | 63.84 | 65.7 | 67.85 | 53.37 | 63.1 | 54.5 | 54.21 |
| | PRE | 65.84 | 59.02 | 56.6 | 71.36 | 57.9 | 62.79 | 60.5 | 55.37 |
| | MFB | 0.03 | -0.07 | -0.13 | 0.08 | 0.09 | 0.03 | 0.14 | 0.05 |
| | MFE | 0.27 | 0.27 | 0.3 | 0.31 | 0.24 | 0.31 | 0.28 | 0.28 |
| | MNB | 0.1 | -0.01 | -0.06 | 0.18 | 0.14 | 0.12 | 0.22 | 0.13 |
| | MNE | 0.3 | 0.26 | 0.29 | 0.36 | 0.27 | 0.34 | 0.33 | 0.3 |
| $O_3$-8h (ppb) | OBS | 53.38 | 52.96 | 51.25 | 53.48 | 46.73 | 49.88 | 44.26 | 45 |
| | PRE | 57.51 | 51.72 | 46.13 | 59.04 | 52.18 | 54.33 | 52.67 | 49.94 |
| | MFB | 0.06 | -0.03 | -0.11 | 0.1 | 0.1 | 0.08 | 0.18 | 0.1 |
| | MFE | 0.26 | 0.26 | 0.26 | 0.26 | 0.23 | 0.26 | 0.28 | 0.24 |
| | MNB | 0.13 | 0.02 | -0.06 | 0.17 | 0.15 | 0.15 | 0.25 | 0.16 |
| | MNE | 0.3 | 0.26 | 0.24 | 0.3 | 0.26 | 0.3 | 0.33 | 0.28 |
| $PM_{2.5}$ ($\mu g\ m^{-3}$) | OBS | 90.85 | 65.55 | 49.28 | 65.61 | 60.93 | 77.74 | 70.13 | 42.7 |
| | PRE | 65.5 | 55.55 | 29.19 | 78.83 | 48.57 | 74.95 | 33.84 | 33.55 |
| | MFB | -0.33 | -0.27 | -0.56 | 0.05 | -0.26 | -0.16 | -0.75 | -0.53 |
| | MFE | 0.64 | 0.57 | 0.68 | 0.57 | 0.62 | 0.57 | 0.88 | 0.77 |
| | MNB | -0.01 | -0.04 | -0.33 | 0.47 | 0.03 | 0.15 | -0.39 | -0.2 |
| | MNE | 0.65 | 0.54 | 0.52 | 0.84 | 0.63 | 0.66 | 0.65 | 0.63 |
| $PM_{10}$ ($\mu g\ m^{-3}$) | OBS | 164.80 | 104.94 | 69.85 | 104.79 | 99.08 | 122.64 | 143.95 | 68.67 |
| | PRE | 73.69 | 63.47 | 34.20 | 86.70 | 52.80 | 80.44 | 44.25 | 35.63 |
| | MFB | -0.71 | -0.55 | -0.69 | -0.25 | -0.62 | -0.49 | -0.98 | -0.76 |
| | MFE | 0.84 | 0.70 | 0.77 | 0.62 | 0.78 | 0.70 | 1.05 | 0.87 |
| | MNB | -0.37 | -0.30 | -0.43 | 0.07 | -0.32 | -0.20 | -0.56 | -0.42 |
| | MNE | 0.63 | 0.54 | 0.55 | 0.68 | 0.60 | 0.60 | 0.69 | 0.62 |
| CO (ppm) | OBS | 1.22 | 0.8 | 0.81 | 0.82 | 0.79 | 1.11 | 1.13 | 0.75 |
| | PRE | 0.37 | 0.29 | 0.22 | 0.41 | 0.25 | 0.4 | 0.23 | 0.22 |
| | MFB | -0.89 | -0.86 | -1.11 | -0.62 | -0.93 | -0.87 | -1.21 | -1.04 |
| | MFE | 0.95 | 0.9 | 1.12 | 0.71 | 0.96 | 0.93 | 1.22 | 1.07 |
| | MNB | -0.54 | -0.55 | -0.69 | -0.39 | -0.58 | -0.52 | -0.72 | -0.63 |
| | MNE | 0.63 | 0.6 | 0.7 | 0.52 | 0.63 | 0.62 | 0.74 | 0.68 |
| $NO_2$ (ppb) | OBS | 24.28 | 21.42 | 23.12 | 21.2 | 21.09 | 21.01 | 22.23 | 16.2 |
| | PRE | 11.26 | 11.77 | 10.71 | 12.53 | 6.37 | 12.03 | 8.4 | 4.29 |
| | MFB | -0.72 | -0.65 | -0.7 | -0.56 | -1.09 | -0.62 | -0.95 | -1.24 |
| | MFE | 0.85 | 0.83 | 0.83 | 0.78 | 1.15 | 0.83 | 1.05 | 1.28 |
| | MNB | -0.39 | -0.31 | -0.39 | -0.24 | -0.61 | -0.27 | -0.52 | -0.7 |
| | MNE | 0.62 | 0.63 | 0.6 | 0.62 | 0.73 | 0.66 | 0.69 | 0.75 |
| $SO_2$ (ppb) | OBS | 22.31 | 14.07 | 10.41 | 12.83 | 21.06 | 17.26 | 16.66 | 11.81 |
| | PRE | 12.24 | 8.66 | 8.07 | 25.77 | 5.13 | 18.55 | 11.58 | 10.28 |
| | MFB | -0.57 | -0.62 | -0.45 | 0.34 | -1.14 | -0.24 | -0.6 | -0.63 |
| | MFE | 0.8 | 0.87 | 0.77 | 0.73 | 1.21 | 0.8 | 0.95 | 1 |
| | MNB | -0.21 | -0.22 | -0.1 | 1.5 | -0.61 | 0.46 | -0.07 | -0.02 |
| | MNE | 0.66 | 0.71 | 0.69 | 1.78 | 0.76 | 1.13 | 0.86 | 0.94 |

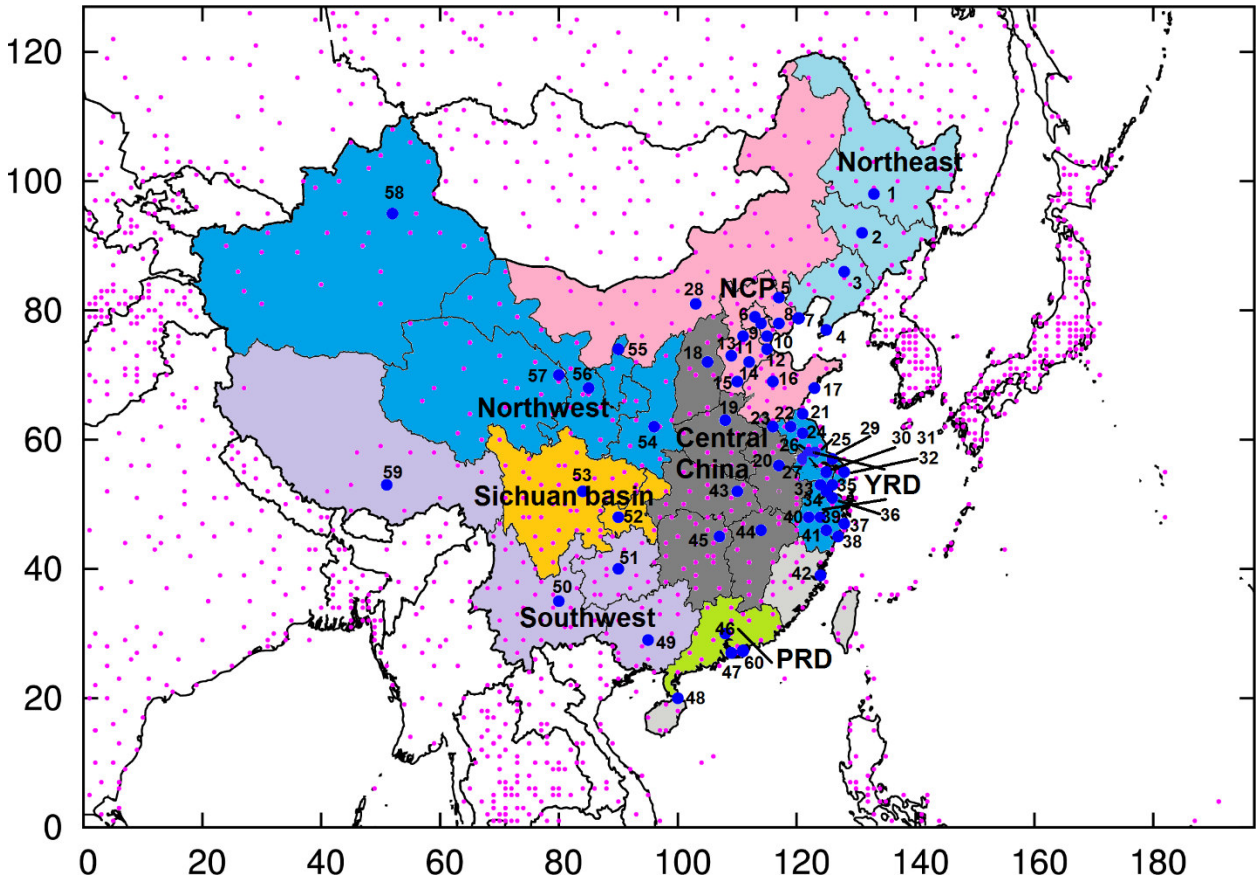



Figure 1. Model domain. The axes are the number of grid cells. Blue filled circles show the locations of cities with air quality observations (see Table 2). The purple dots show the locations of meteorological stations. The figure also shows the regions discussed in the text for better understanding. NCP represents North China Plain, YRD represents Yangtze River Delta, and PRD represents Pearl River Delta.

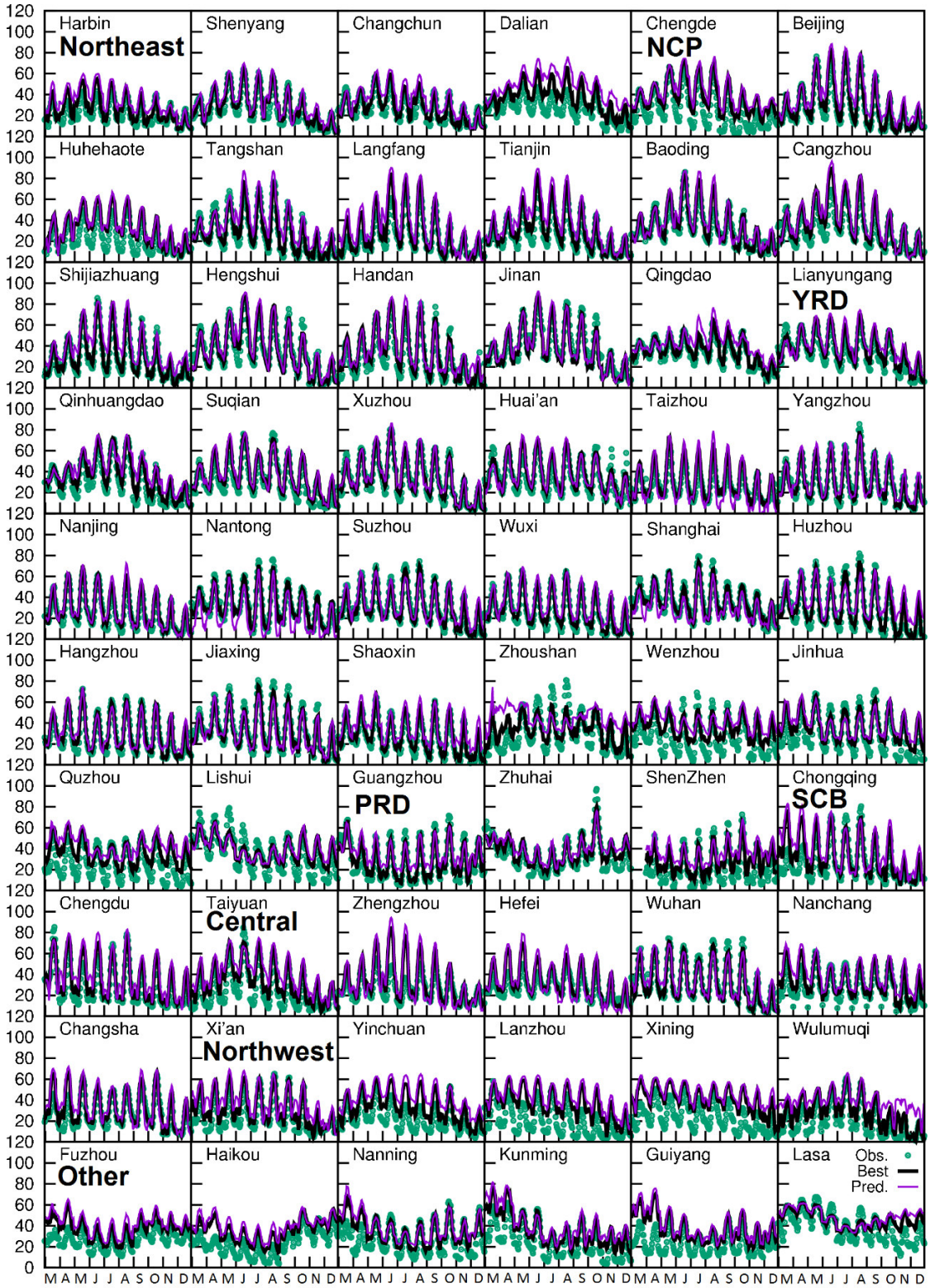



Figure 2. Comparison of monthly averaged diurnal variations of $O_3$ concentrations from March to December, 2013. Pred. are the values predicted at the grid cell each city center located while Best are the values predicted closest to the observations within 3 by 3 grid cell regions that surround the observation. Units are ppb.

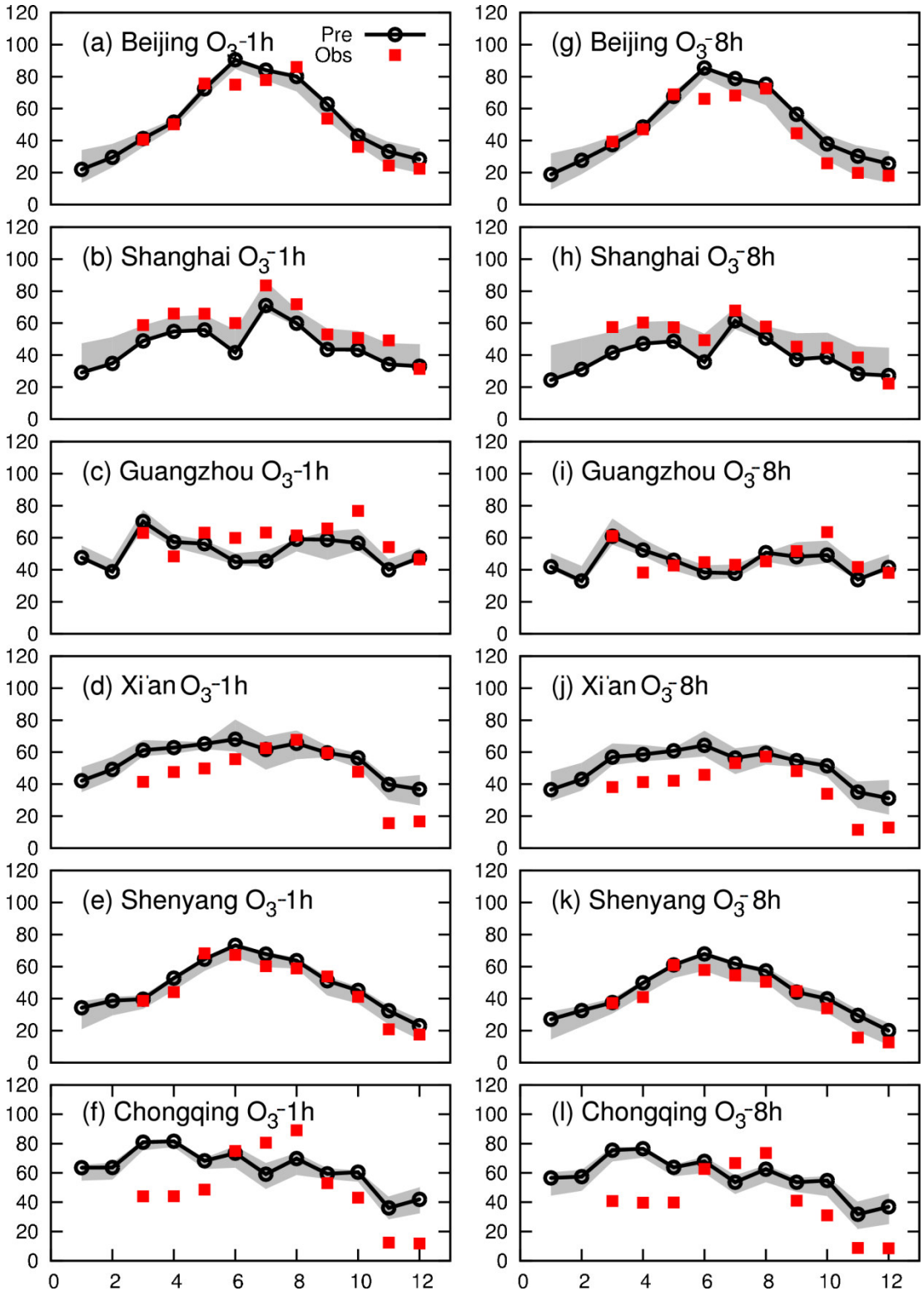



Figure 2. Comparison of predicted and observed $O_3$-1h and $O_3$-8h concentrations at Beijing, Shanghai, Guangzhou, Xi'an, Shenyang, and Chongqing. Grey areas represent ranges in model predictions within 3x3 grid cells surrounding the observation Units are ppb.

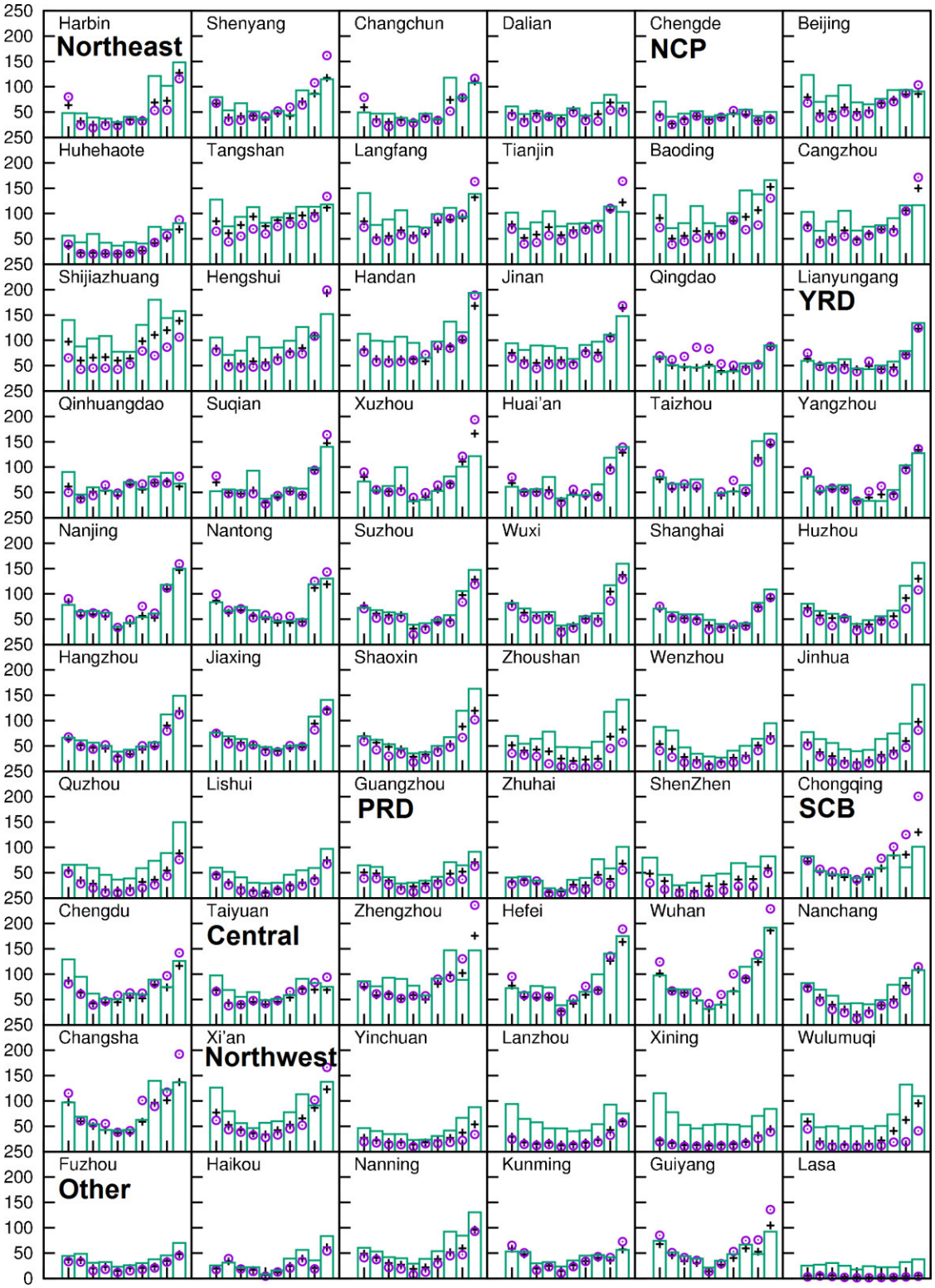
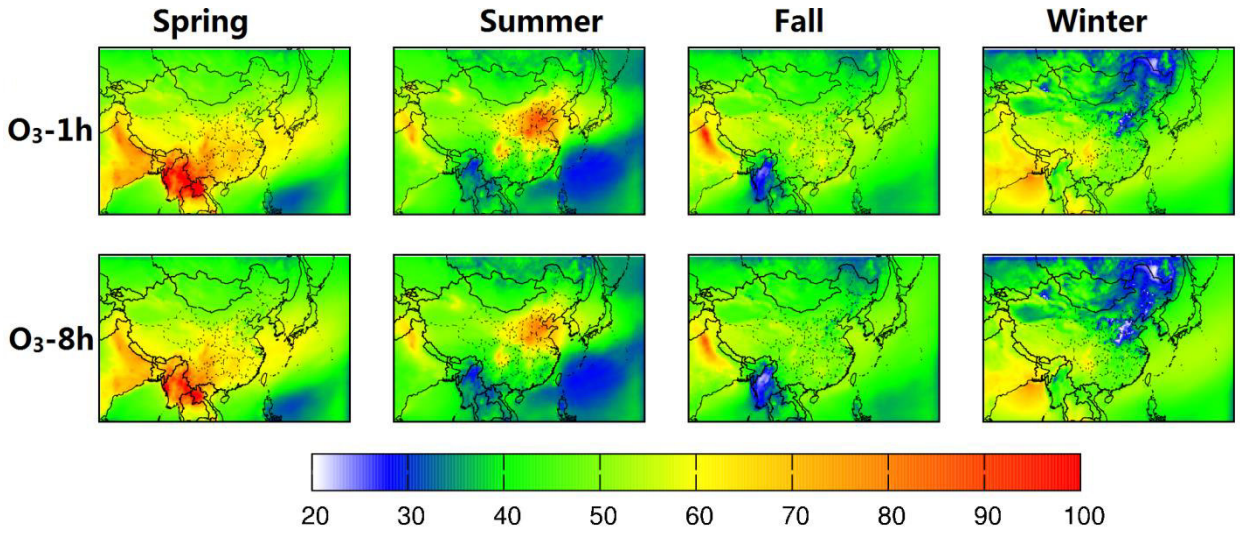



Figure 4. Comparison of predicted (in column) and observed (in circle) monthly averaged $PM_{2.5}$ concentrations for March to December, 2013. The "Best" lines (in "+") represent predictions closest to the hourly observations within a 3×3 grid cell region with the grid cell where the monitoring sites are located at the center. Units are $\mu g\ m^{-3}$.

736




737
738 Figure 5. Seasonal variations of predicted regional distribution of $O_3$-1h and $O_3$-8h. Units are ppb.

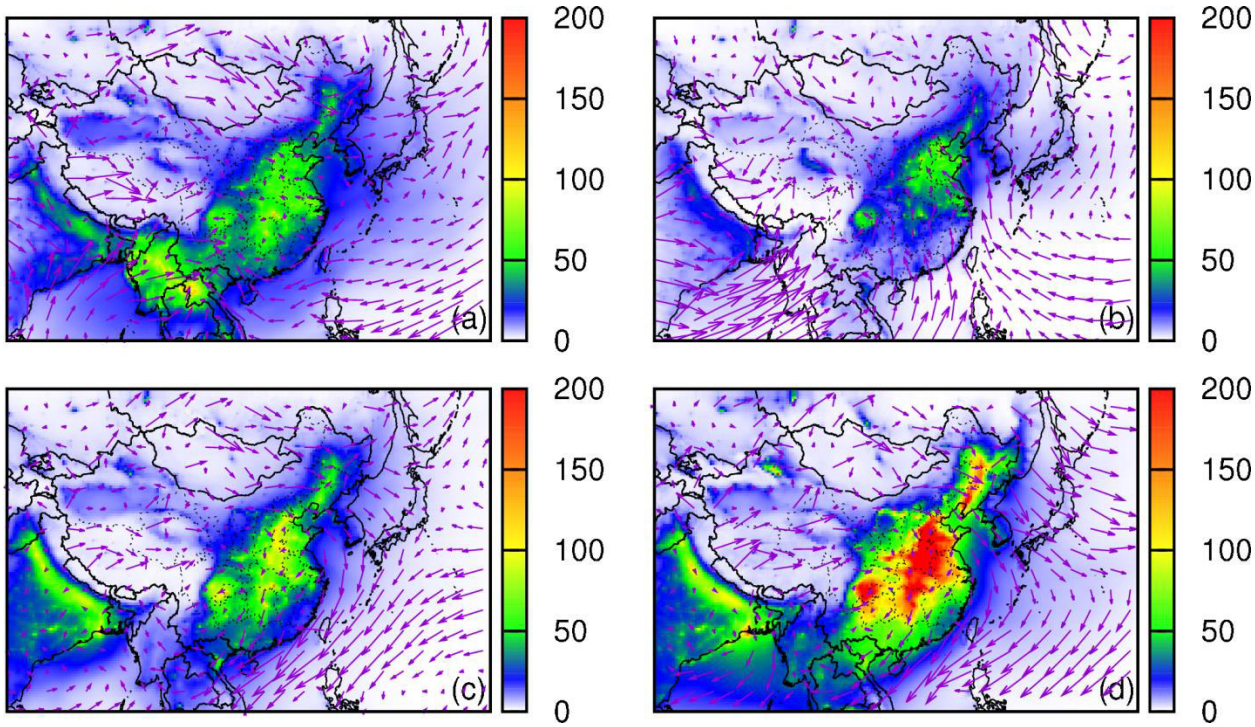



Figure 6. Seasonal variation of predicted $PM_{2.5}$ and wind vectors: (a) spring, (b) summer, (c) fall, and (d) winter. Units are $\mu g\ m^{-3}$.

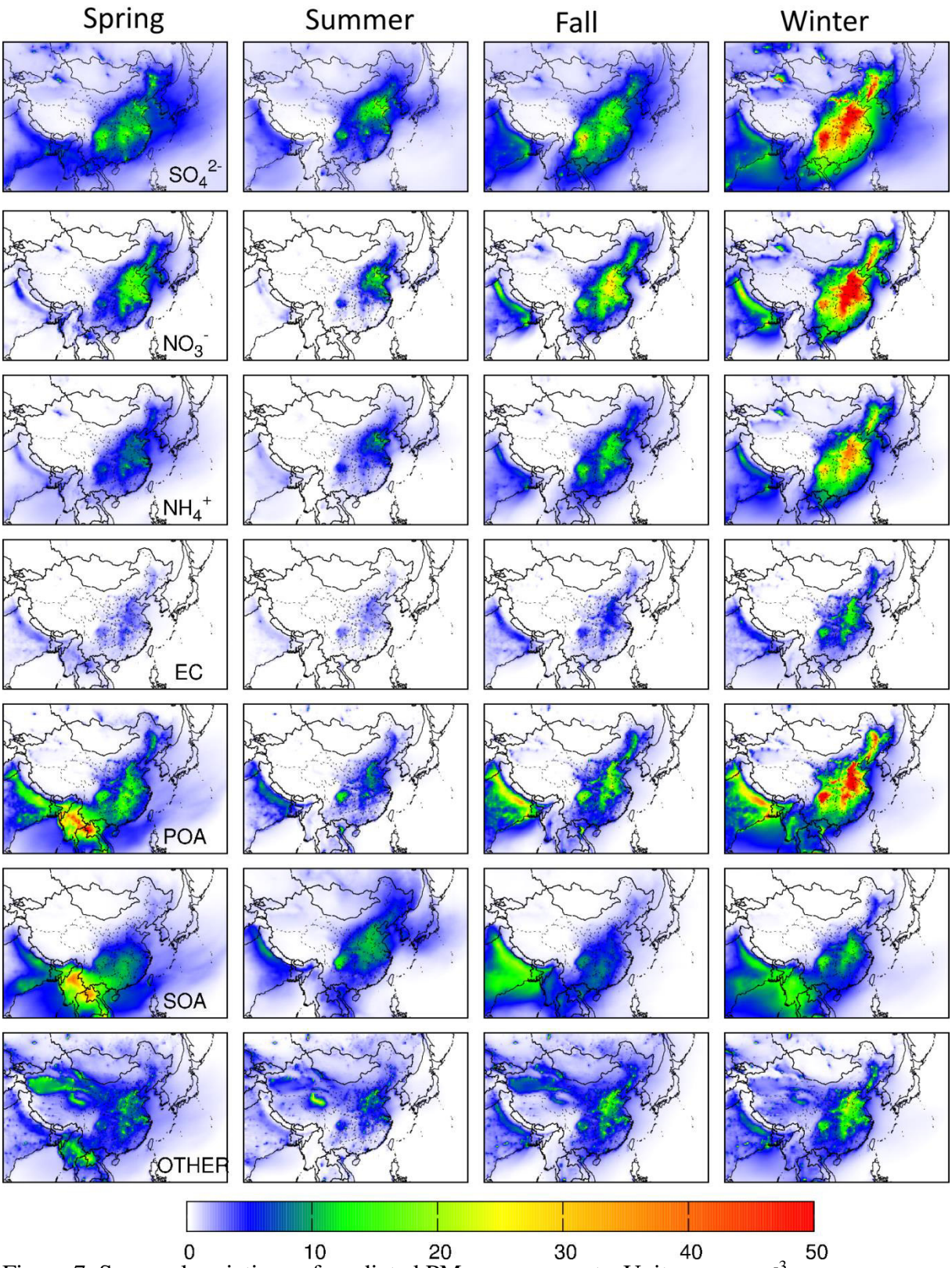
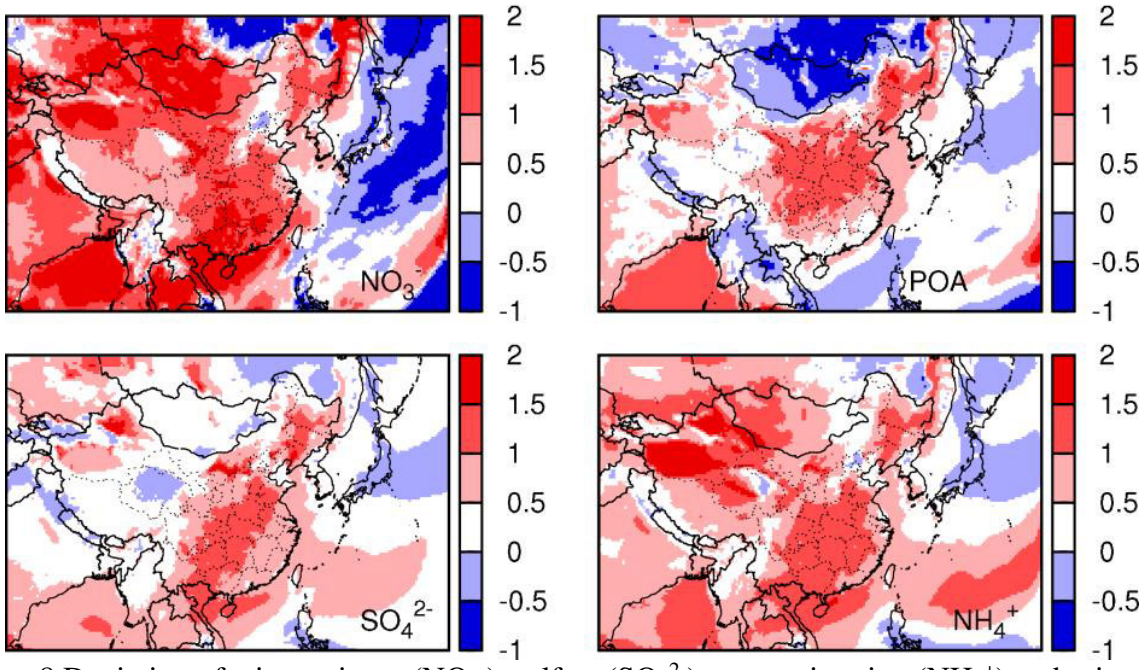



Figure 7. Seasonal variations of predicted $PM_{2.5}$ components. Units are $\mu g\ m^{-3}$.




744
745 Figure 8 Deviation of winter nitrate ($NO_3^-$), sulfate ($SO_4^{2-}$), ammonium ion ($NH_4^+$) and primary
746 organic aerosol (POA) from annual average, as calculated by (W-A)/A, where W and A are win-
747 ter and annual concentrations, respectively.
748

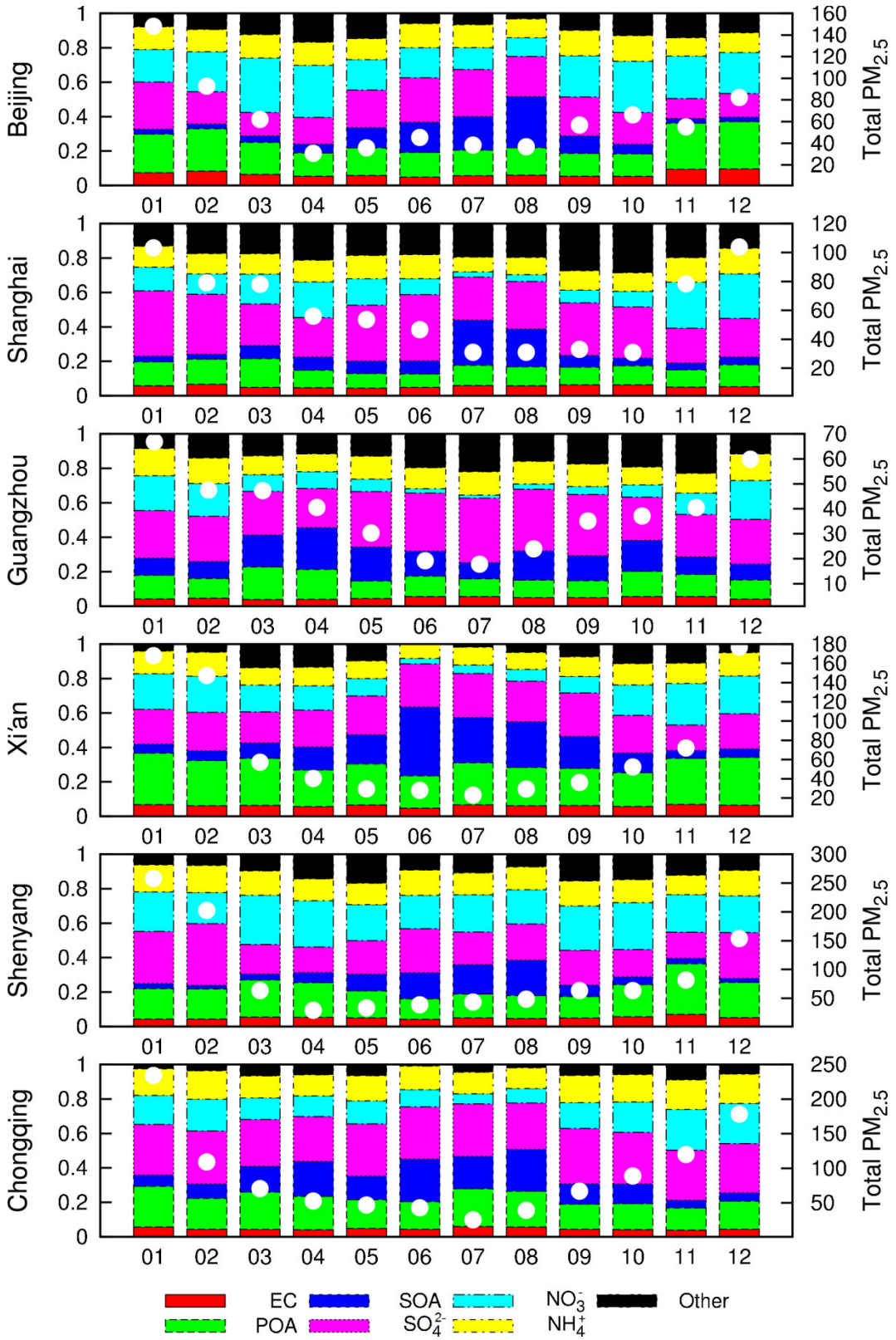



Figure 9. Contributions of different components to monthly averaged $PM_{2.5}$ concentrations at selected cities in China. White circles are absolute concentrations according to right y-axis with unit of $\mu g\ m^{-3}$.

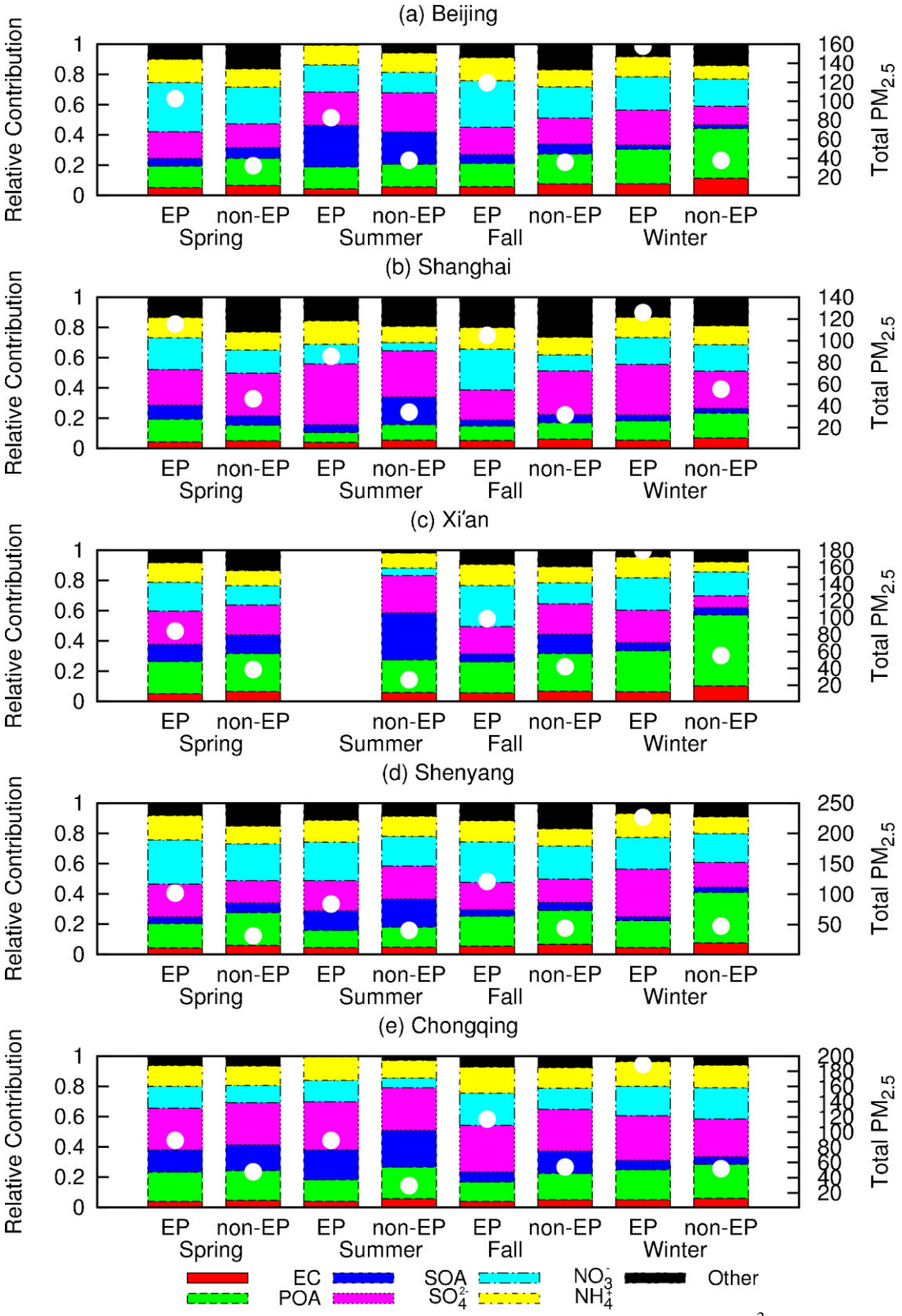



Figure 10. Comparison of $PM_{2.5}$ components at episode days (Ep, >=75 μg $m^{-3}$) and non-episode days (non-EP, <75 μg $m^{-3}$). White circles are absolute concentrations according to right y-axis with unit of μg $m^{-3}$. Note Xi'an does not have episode days in summer.