COMMENTS TO THE AUTHOR(S)

Retrospective analysis of 2015-2017 winter-time PM2.5 in China: response to emission regulations and the role of meteorology

Manuscript ID: acp-2018-890

Authors: Chen, et al.

Reviewer 2

The manuscript by Chen et al. evaluates PM2.5 in January in 2015 - 2017 in China.

Data assimilation is used to construct a reanalysis that well matches the available observations of surface-level PM2.5 concentrations. The difference between this and a Non-forced simulation is cleverly used to parse out the separate roles of emissions from meteorology in variability of PM2.5 concentrations, which are significant across these 3 years (although they authors refer to these as trends – which I think is better described as inter annual variability). This is important and valuable work, because it helps identify reasons why emissions control strategies may or may not have an immediately visible impact. As such the topic and scope are suitable for ACP. The manuscript includes some additional analysis of AOD, but it is somewhat secondhand, not as well supported in terms of model accuracy, and doesn't particularly add to the focus of the paper, which I suggest remain on surface-level PM2.5 concentrations. This would provide space for the authors to provide more details on other aspects of their modeling and assimilation methods, which in several places are too abbreviated or presented without sufficient background or justification. The major scientific weakness in this work is likely the use of out-dated emissions (from 2010), given the rate at which species like SO2 and NOx are known to be changing since 2010 in China. Below I provide more detailed comments on these and other aspects to address prior to publication in ACP.

Response:

Thanks for the valuable and insightful suggestions! We have made several changes accordingly.

1. Removed the contents relevant with AOD comparisons.
2. Added more details on modeling and assimilation methods.
3. Rerun the assimilation experiment with looser filter criteria, in which the interannual changes were better captured.

4. Added the introduction of model deficiencies, including the imperfection of chemistry mechanism, and also the inaccuracy of 2010-MEIC EI for the study period. The impacts of the deficiencies on analysis are also discussed.

Please see the itemized responses as below. Revised manuscript is after the response letter.

Comments: The grammar needs work throughout. This might mean adding another co-author or hiring an editing service. The number of corrections needed (nearly every sentence) are far too extensive for me to detail here.

The abstract would be improved by considering some of the following suggestions:

- lead with a broader, introductory statement

- avoid jargon when possible

- try to provide a mix of qualitative and quantitative results. Presently, only qualitative results are reported

- end with a statement about the bigger impacts of this work

- reduce the overall length – currently it goes too much into details of the methods, without quantitatively summarizing the most important conclusions.

Response:

We have hired the editing service from American Journal Experts (www.aje.com) and edited the manuscript thoroughly. We also revised the abstract accordingly.

Q1. 4.22: I believe that this statement regarding the inorganic PM2.5 in this region has been known for some time.

Response:

Yes, thanks for pointing out this. We have added relevant references here and removed the repeated sentences.

Q2. 7.1-3: There is considerable debate in the literature regarding the reactions that may be leading to high concentrations in haze events. The ones included here are based on assumptions of pH that may not be correct. Other recent works have e.g. suggested

HCHO may play a role. In short, I recommend the authors review the relative literature on this topic (which should be easy to find, several very high-profile papers). Even if they decide to stick with their current

mechanism (which I think is acceptable, given their is yet to be scientific consensus on this issue), discussion is warranted, both in the introduction and consideration of sources of errors towards the end of the paper.

Response:

We have added relevant references of the recently proposed chemistry mechanisms and also discussed the potential impacts on the uncertainties of analysis.

"In addition to the uncertainties in emission inventories, deficiencies in the model chemistry can also cause model uncertainties. Increasing numbers of observations have revealed that anthropogenic emission-relevant aerosol species, such as sulfate, nitrate and ammonium (denoted as SNA), are the predominant inorganic species in the wintertime $PM_{2.5}$ in China (Wang *et al.*, 2014c; Yang *et al.*, 2015). Various reaction paths during haze events have also been proposed (e.g. Zheng *et al.*, 2015; Cheng *et al.*, 2016; Wang *et al.*, 2016; Li *et al.*, 2017; Moch *et al.*, 2018; Wang *et al.*, 2018; Shao *et al.*, 2019). For example, Moch *et al.* (2018) used a 1-D model and revealed the importance of aqueous-phase chemistry of HCHO and S(IV) in cloud droplets by forming a S(IV)-HCHO adduct, hydroxymethane sulfonate. Shao *et al.* (2019) implemented four heterogeneous sulfate formation mechanisms (via H_2O_2, O_3, NO_2, and transition metal ions on aerosols) into GEOS-Chem model which partially reduced the modeled low bias in sulfate concentrations. However, a scientific consensus regarding the importance of the reaction paths has not yet been reached partially due to the uncertainties of aerosol liquid water content, pH, and ionic strength etc. The original WRF/Chem model with either the Goddard Chemistry Aerosol Radiation and Transport (GOGART, Chin *et al.*, 2000, 2002) or the Model for Simulating Aerosol Interactions and Chemistry (MOSAIC)-4BIN aerosol scheme failed to reproduce the highest $PM_{2.5}$ concentrations; it is assumed that this failure is due to missing heterogeneous/aqueous reactions. In Chen *et al.* (2016, hereafter Chen16), we added three heterogeneous reactions (SO_2-to-H_2SO_4 and NO_2/NO_3-to-HNO_3) to the WRF/Chem model based on the MOSAIC-4BIN aerosol scheme. Although the reaction paths may still not be comprehensively understood, the new MOSAIC-4BIN aerosol scheme significantly improved the simulation of sulfate, nitrate, and ammonium on polluted days in terms of the concentrations of those species and their partitioning."

"It is worth noting that there are uncertainties in the simulation/assimilation processes. There are three sources of uncertainties in the NO_DA simulation. First, the emission inventories in the NO_DA simulations are obviously not accurate, which may introduce uncertainties into the analysis……Second, the meteorological IC/BC conditions in the NO_DA simulations, which were obtained from GFS 6-hr analysis data, also have uncertainties……Third, the deficiencies associated with the chemistry in the model also generate uncertainties, including missing reactions and the inaccurate parameterization of reactions. These three aspects all originate from the imperfections of current forward models."

Q3. Figure 2: It's not clear what this shows. What are "PM2.5 emissions"? Since many of the species contributing to PM2.5 are gas-phase precursors, which don't necessarily completely transform into aerosol, I'm quite puzzled. Perhaps it is a plot of a subset of

PM2.5 precursor emission species, but that isn't made clear.

Response:

The original figure shows the primary $PM_{2.5}$ emission, corresponding to the total of BC, OC, sulfate, nitrate and other unspecified $PM_{2.5}$ emissions. We have clarified in the figure caption. To be more comprehensive, the emission spatial distributions of $PM_{2.5}$ precursors, including SO_2, NO_x and NH_3 have also been added in Figure 2.

Q4. 8.1 and other locations (e.g., 14.11): The emissions used in this work are quite outdated. It is documented in several studies of emissions in China that SO2 emissions have been decreasing since around 2009, and NOx emissions since around 2011 or 2012. These previous studies need to be cited, and considered in the present work. It is certainly a significant source of error worth considering.

Response:

The 2010-MEIC EI is the only emission inventory that publicly available when the study was conducted. We have added the relevant references on the emission reductions of SO_2 and NO_x in the text.

"As the Chinese government has implemented strict control strategies to ensure an improved air quality during the winter season since 2013, significant reductions in emissions, including primary PM and precursor compounds (SO_2 and NO_x), in regions with the strict implementation of these policies relative to the year 2010 are expected for our simulation period. A reduction in SO_2 pollution of approximately 50% was observed from 2012-2015 for the North China Plain from OMI satellite data (Krotkov *et al.*, 2016). National anthropogenic emission reductions of approximately 67%, 17%, and 35% from 2012-2017 for SO_2, NO_x, and $PM_{2.5}$, respectively, were assumed by the bottom-up EI methodology (Zheng *et al.*, 2018). However, the expansion and relocation of the energy industry caused emission increases in northwestern China (Ling *et al.*, 2017). In addition, the uncertainties of allocated emissions in the winter season will be much larger than those in other seasons. For example, Zhi *et al.* (2017) conducted a village energy survey and revealed an enormous discrepancy in the amount of rural raw coal used for winter heating in northern China, implying an extreme underestimation of rural household coal consumption by the China Energy Statistical Yearbooks. These changes and uncertainties of emissions in the model would introduce errors into the NO_DA simulation. However, the inhomogeneous spatial changes and large uncertainties of seasonal allocations made it difficult to simply scale the original emission inventory for our study period."

"In the NO_DA case, the model results are overpredicted in SB, NCP and CC for all three months, while the overestimations are more severe in SB. The NO_DA case generally overestimates (underestimates) the surface $PM_{2.5}$ in NCP, SB and CC (XJ and FWP) in the three years, potentially indicating that the 2010 emissions are not appropriate for the 2015-2017 simulations with overestimations (underestimations). As discussed in section 2.1, the large area of overestimation is consistent with the national reductions in SO_2, NO_x

and $PM_{2.5}$ anthropogenic emissions (Zheng *et al.*, 2018); however, the underestimations in XJ and FWP also indicate the introduction of new emission sources to these two regions."

"It is worth noting that there are uncertainties in the simulation/assimilation processes. There are three sources of uncertainties in the NO_DA simulation. First, the emission inventories in the NO_DA simulations are obviously not accurate, which may introduce uncertainties into the analysis. Although the basic assumption required only that the emissions stay the same throughout the three years, emission inventory uncertainty-induced errors would be offset in the subtraction process when calculating the year-to-year differences. However it did generate uncertainties. For example, the emissions in SB, CC and PRD were generally overestimated (Fig. 3), which means that the variations in the ambient concentration might have been artificially amplified considering the meteorology impacts (Fig. 6c). In contrast, the emissions in XJ and FWP were underestimated (Fig. 3), and thus, the changes in the ambient concentrations due to meteorological conditions in these two regions might have diminished. From this point of view, if the fixed emissions are more accurate in those years, the results would be more reliable. In the case where "real" emissions are not available and the purpose is to evaluate the contribution of those emissions, uncertainties will be unavoidable and should be emphasized carefully."

Q5. 9.2: Not sure why capital pi notation is used here, as that means product, where the definition is in terms of a sum.

Response:

Changed to $M_{PM_{2.5}}$ instead.

Q6. 9.10-12: I don't understand how this works – can it be explained further? The way in which measurements of total PM2.5 are used to adjust concentrations of specific species should be clarified, even though it comes from an earlier work, if only briefly.

Response:

We have added the following explanation in the text (highlighted in blue).

"Since only surface $PM_{2.5}$ total mass observations were assimilated to analyze the 3D mass mixing ratios of 24 aerosol variables, the 3DVAR problem was underconstrained. Due to the lack of species and vertical information provided by the observations, the only mathematical solution is to utilize prior information from the model background. In the GSI system, the distribution of the analysis increments (the difference between the analysis and background) onto the different species was mostly model driven with the observation and background error covariance matrices acting as the main constraints."

"Similar to Jiang13, the background error covariance (BEC) statistics for each analysis variable required by the 3DVAR algorithm were computed by utilizing the NMC method (Parrish and Derber, 1992) based

upon the one-month WRF/Chem forecast for January 2015. No cross-correlation between different species was considered. The standard deviations and horizontal/vertical correlation length scales of the background errors (separated for each aerosol species) were calculated using the method described by Wu *et al.* (2002). These data were used as constraints for the distributions of the PM components."

Q7. 9.20: What is the origin of this assumption regarding error? The relative error component of 0.75% seems very small. Other parameters such as gamma and L, the maximum concentration threshold (500) or analysis increment (120) are similarly introduced without much explanation. I recognize these values have been used before, but still more explanation here would be appreciated.

Response:

We have added the original reference in the text. Regarding the parameter of error calculation, we have added the explanations for those parameters as below.

"The observation error covariance matrix **R** in Eq. (1) contains both measurement and representativeness errors. Pagowski *et al.* (2010) used a measurement error (ε_0) of 2 μg m^{-3}. To associate higher $PM_{2.5}$ values with larger measurement errors, S12 defined the measurement error as $\varepsilon_0 = 1.5 + 0.0075 \times M_{PM_{2.5}}$, where $M_{PM_{2.5}}$ denotes an AIRNow $PM_{2.5}$ observation and the units of each term are μg m^{-3}. According to the $PM_{2.5}$ Auto-Monitoring Instrument Technical Standard and Requirement (China National Environmental Monitoring Center, 2013), three continuous online monitoring methods, namely, a beta-ray plus dynamic heating system, a beta-ray plus dynamic heating system plus light scattering system, and a tapered element oscillating microbalance plus filter dynamic measurement system, are used at the national monitoring sites to satisfy the requirements that the display resolution should be less than 1 μg m^{-3} and the error should be less than 5 μg m^{-3}(within 24 hours). To reflect the confidence in the hourly observations, the measurement error ε_0 in this study is defined as $\varepsilon_0 = 1.0 + 0.0075 \times M_{PM_{2.5}}$, where $M_{PM_{2.5}}$ denotes a $PM_{2.5}$ observational value (unit: μg m^{-3}).

"Representativeness errors reflect the inaccuracies in the forward operator and in the interpolation from the model grid to the observation location. Elbern *et al.* (2007), Pagowski *et al.* (2010), S12 and Jiang13 defined the representativeness error (ε_r) as

$$\varepsilon_r = \gamma \varepsilon_0 \sqrt{\frac{\Delta x}{L}}, \quad (3)$$

"where γ is an adjustable parameter scaling ε_0 ($\gamma = 0.5$ was used here), Δx is the grid spacing (40 km in our case) and L is the radius of influence of an observation (set to 2 km for urban sites). These parameter settings were based on the performance of sensitivity tests."

The criteria for filter process are from two aspects, including the stability of DA optimization step and the computing efficiency. The original criteria were mostly set for operational runs. For research purpose, we

have made tests of different filter process and found that looser filter can really improve the assimilation results. Here in the revised manuscript, only $PM_{2.5}$ observations larger than 1000 μg m^{-3} (the maximum display limit of the monitoring system) were deemed unrealistic in the filter process and observations leading to deviations exceeding 500 μg m^{-3} were omitted in the assimilation process for the stability of the assimilation optimization. Besides, in the original assimilation experiment observational sites located in grids with elevation greater than 500 meter (Above Sea-Level) were not used; to better utilize those data, we chose to interpolate them to the lowest model level for assimilation.

Below shows the observed, original and updated assimilated monthly average of $PM_{2.5}$ concentrations (unit: μg m^{-3}) for January in 2015 (Left), 2016 (middle) and 2017 (right). The most prominent improvements are shown for the hotspots in Xinjiang (Region d) and Fenwei Plain (added as Region e).

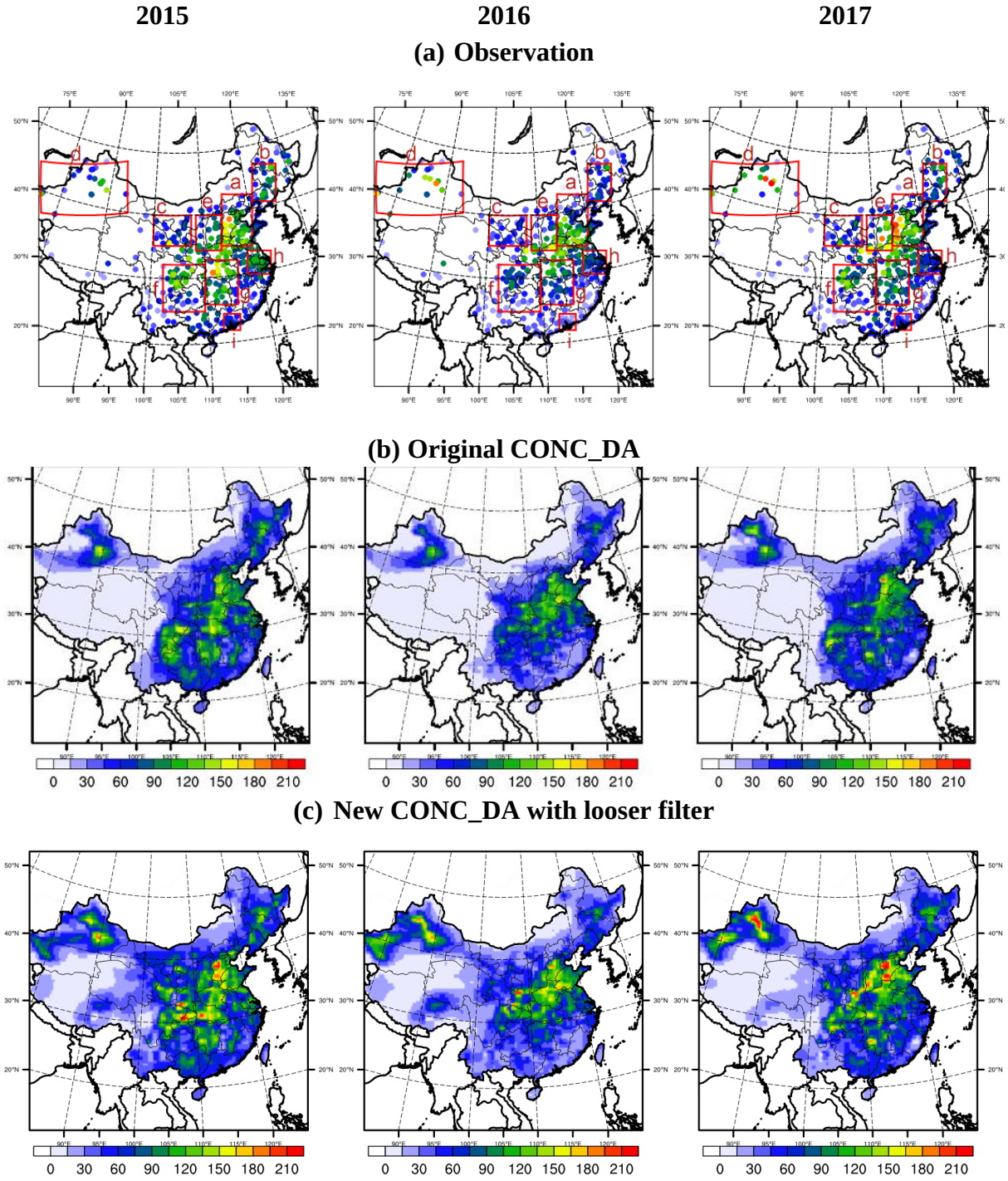


Figure S1. Observed and modeled monthly average of $PM_{2.5}$ concentrations (unit: μg m^{-3}) for January in 2015 (Left), 2016 (middle) and 2017 (right). (a) Observation, (b) Original CONC_DA, (c) New CONC_DA with looser filter. Regions defined in red rectangles are: a-NCP (North China Plain), b-NEC (Northeastern China), c- EGT (Energy Golden Triangle), d-XJ (Xinjiang), e-Fenwei Plain (FWP), f-SB (Sichuan Basin), g-CC (Central China), h-YRD (Yangtze River Delta), i-PRD (Pearl River Delta).

In the updated assimilation experiment, the interannual changes are also better captured as shown below. Those improvements make our analysis more solid, especially for the Xinjiang region and Fenwei Plain. The updated figures and discussions are highlighted in blue in the revised manuscripts.

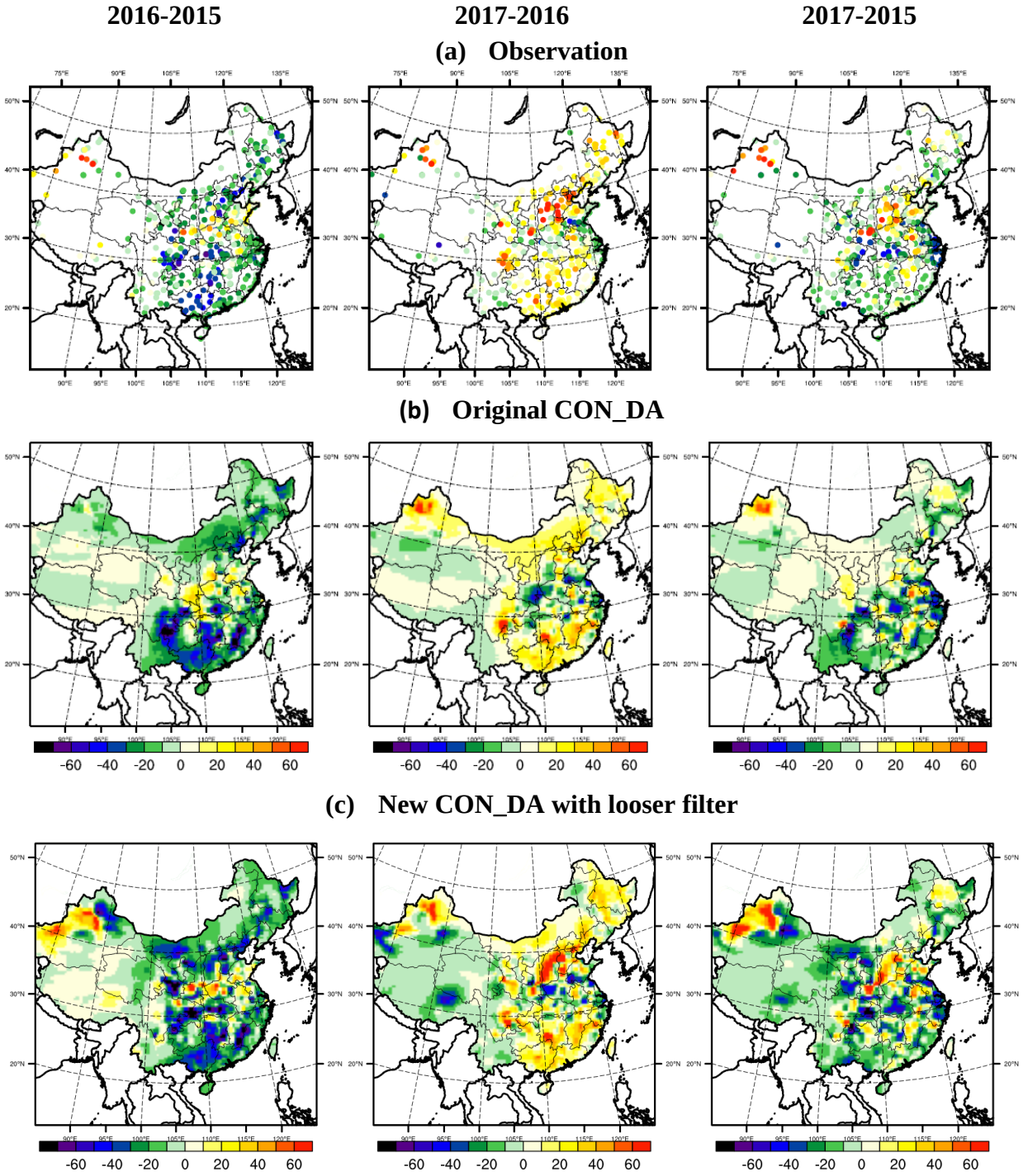


Figure S2. Observed and modeled $PM_{2.5}$ ambient concentration changes for 2016-2015 (left), 2017-2016 (middle) and 2017-2015 (right). (a) Observation, (b) Original CONC_DA, (c) New CONC_DA with looser filter. (Unit: μg m^{-3})

Q8. 10.12: Omission of cross-correlation between species like ammonium and nitrate seems critical – how does this impact the results?

Response:

Actually, it would be critical if the individual species observations (such as ammonium, sulfate, and nitrate) were assimilated while the correlation between them were not considered. However, due to the lack of species observation available, the adjustments of species were entirely from the prior information form the model background. As the species partitioning were reasonably considered in the updated chemistry scheme, it's expected that the correlation among the species are well presented and the omission of cross-correlation would bring very small impacts to the assimilation results. In the future studies, with more species observations, the comparisons of individual species should be addressed.

Q9. Section 2.5: The method used for estimating the impact of meteorology separately from emissions seems sound; I'm not sure the extended explanation (by way of comparison to radiative forcing calculations, etc.) is needed and suggest simply removing the first paragraph of Section 2.5 and jumping directly into the statement of how this study was carried out.

Response:

Thanks for the suggestion. We have removed most contents of the first paragraph of section 2.5.

Q10. Fig 4: The model performance seems good after the assimilation. One small question though – it seems like the residual bias in the CONC_DA case is most often negative.

Is there a reason for this? I would have expected that, given the initial simulation is biased high, the analysis would not necessarily be unbiased, but would similarly have a small slight high bias (owing to the constraint term in the cost function).

Response:

Thanks for pointing out this! Actually I'm thinking that's due to the data filter process in which high values of observations were not assimilated; but all the data with observations available (including high values) are compared and shown in the figure.

As answered in Q6, we have chosen looser filter criteria that only $PM_{2.5}$ observations larger than 1000 μg m^{-3} were deemed unrealistic and observations leading to deviations exceeding 500 μg m^{-3} were omitted. Below the updated figure shows that the analysis is slightly high bias as expected.

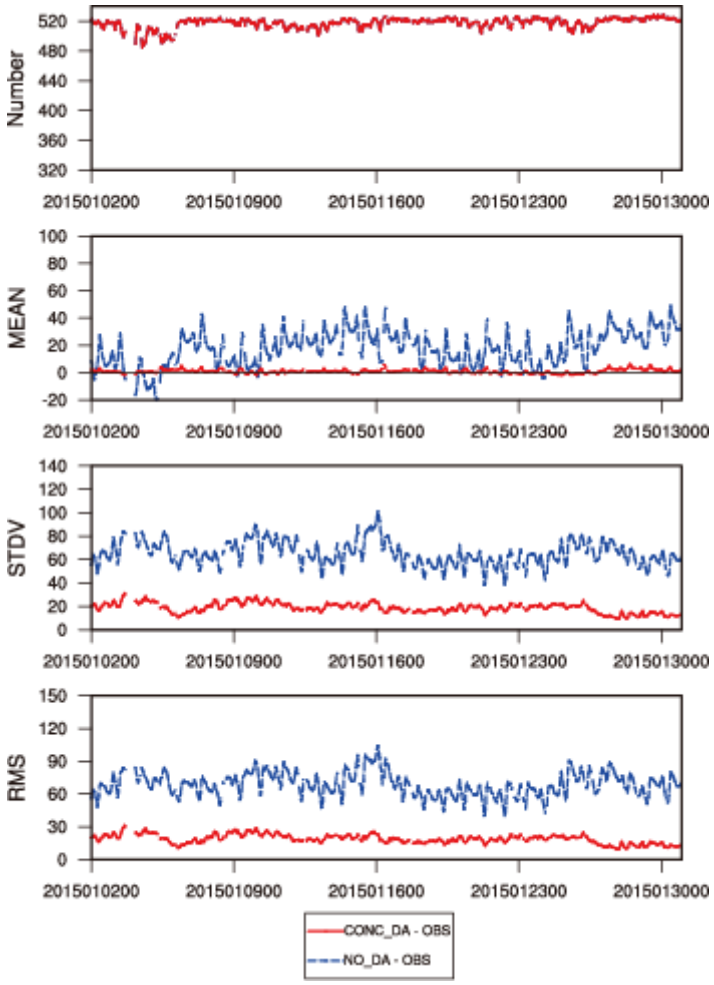


Figure S3. Time series of the statistics between the model simulations and observations. Red lines-CONC_DA minus observations, blue lines –NO_DA minus observations. Statistics include the number of data pairs for comparison, the MEAN-mean bias, the STDV- standard deviation, and the RMS-root mean square error. Left-2015, middle-2016, right-2017. (units are μg m^{-3} for MEAN, STDV and RMS).

Q11. Fig 5: Figs 5b and 5c could be omitted or moved to supporting information.

Response:

Thanks for the suggestion. Figs 5b and 5c were moved to the supporting information.

Q12. 15.24: Why not use level 2 data?

16.7: I didn't follow nor agree with the logic behind this statement. Why not include statistics of the AOD comparison in a table, or directly on the figures themselves (there is room in the white space). Also, it seems that DA only serve to decrease the AOD, not increase it, even in areas.

16.14: Let's be honest – the DA didn't "didn't correct the bias" – it made the bias even worse, for 6 out of the 9 sites. This won't go un-noticed by the readers, so it might was well be presented fairly.

Overall, the AOD analysis was on the weaker side, the connections to the policy and haze questions not as clear, and the model performance not as good (especially for

AERONET). I would suggest the authors consider dropping AOD related content entirely, unless some more satisfactory explanations can be included.

Response:

Thanks for the suggestion! Actually it's difficult to make judge of the assimilation experiment improvements by using MODIS/AERONET AOD data, as the vertical profiles and the assumptions of optical properties in the model can't be evaluated at this stage. According to your suggestion, we decided to remove the entire session relevant with MODIS/AERONET AOD.

Q13. Section 4: Evaluation of concentrations in January alone of 3 consecutive years is not sufficient enough time range for a "trends" analysis. We do get some sense of interannual variability though, which is interesting. It is just mislabeled. For example, every place that says something like "decreasing trend from 2015 to 2016" should say "decrease from 2015 to 2016", as the long-term trend hasn't been determined. If the authors really do wish to study trends, they should have considered years such as 2005, 2010 and 2015. Or if they can't go back that far, owing to data availability,

2012, 2014, and 2017. That would start to be close to enough years to make a trend analysis. My overall suggestion would actually be to remove section 4.1 entirely, at least the second paragraph on AOD. I also wonder if any of these years happened to be impacted by PM2.5 transport more than others, e.g. from fires?

Response:

Thanks for the suggestion! The second paragraph on AOD was removed. We have changed the word "trends" to "interannual changes".

Actually Chinese government has implemented strict control strategies on crop/biomass burning, thus the local and central monitoring teams were mandatory on site to prevent fire events. Besides, no big natural forest fires occurred during our study period. For these reasons, we assumed that the interannual changes of $PM_{2.5}$ are mostly driven by meteorology and anthropogenic emissions.

Q14. 18.13: I don't understand what is meant by the sentence beginning "Thus only. . ."

Response:

Revised the sentence to “The ambient response from 2015-2017 is contradictory if considering only the reductions in emissions and omitting the changes in meteorological conditions.”

Q15. Fig 8: Suggest removing this figure; it is barely discussed, and doesn’t add much. The analysis of PM2.5 surface concentrations is sufficient and also more convincing, since the model performance is better.

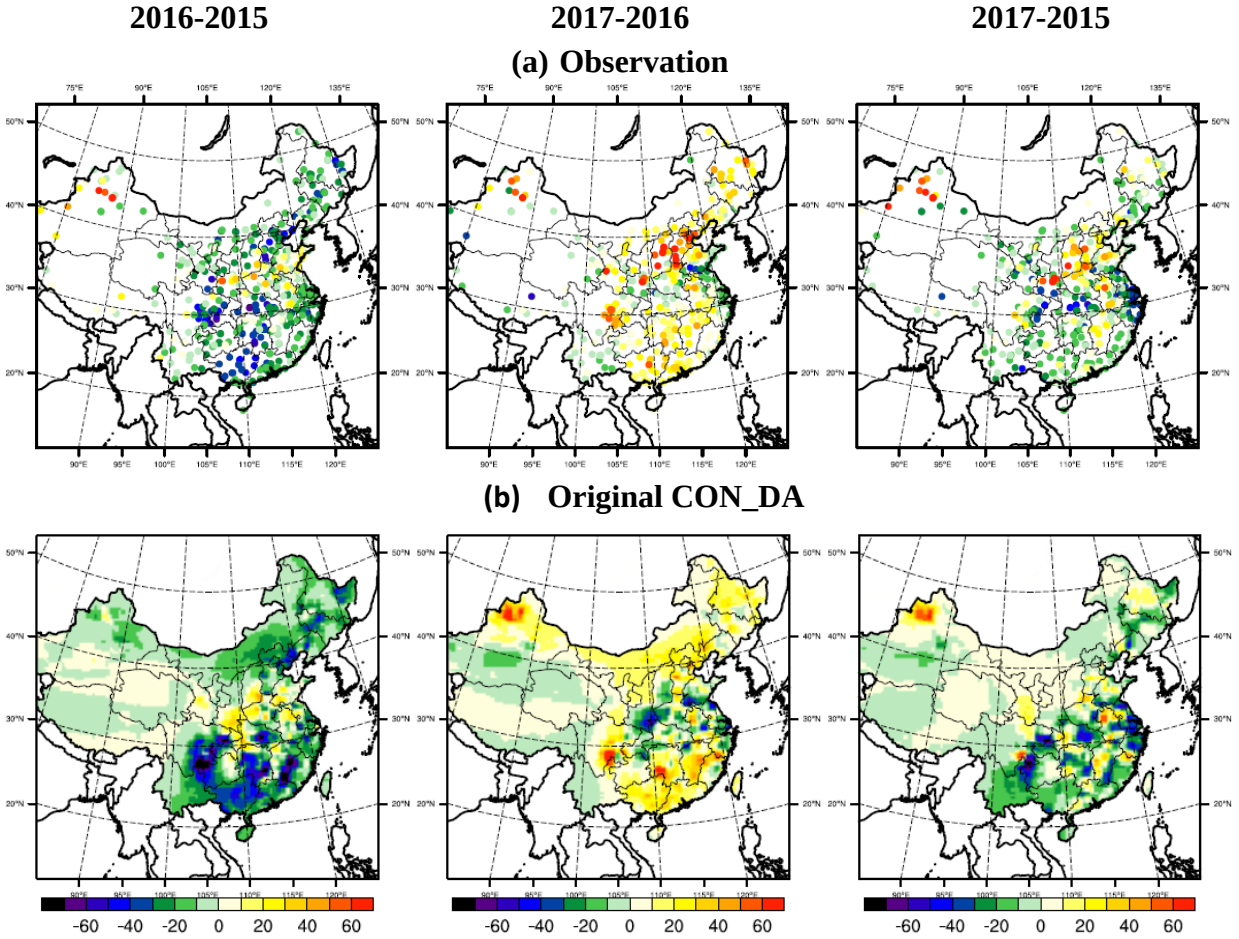
Response:

Thanks! Figure 8 had been removed.

Q16. 20.1: I think this is an important point – if the authors are using this approach to separate emissions impacts from meteorology in the observed dataset, then it is critical the relative changes in the total assimilation experiment (row b) match the observations.

Response:

Thanks for pointing out this! We have rerun the assimilation experiment and the updated results did reproduce the interannual changes (as shown below).



(c) New CON_DA with looser filter

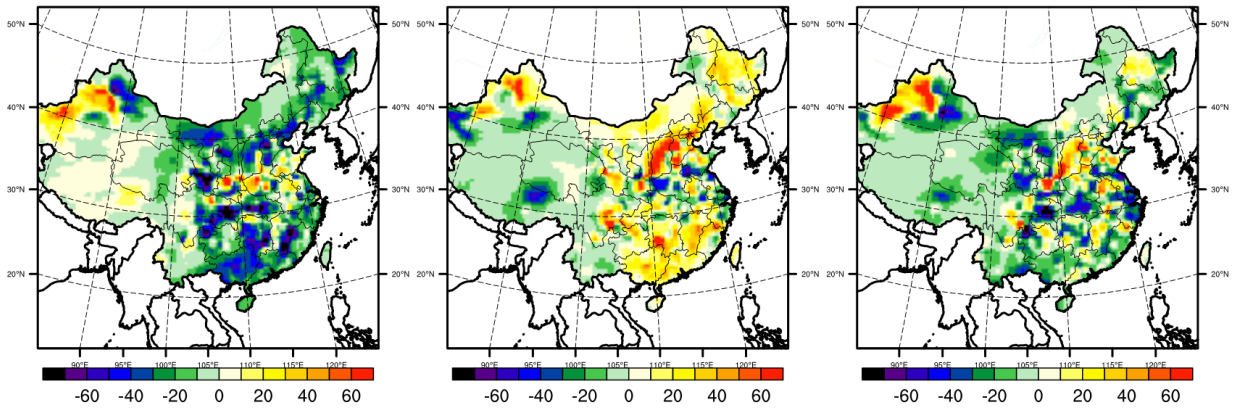


Figure S2. Observed and modeled $PM_{2.5}$ ambient concentration changes for January 2016-2015 (left), 2017-2016 (middle) and 2017-2015 (right). (Unit: μg m^{-3}).

Q17. 21.14: Change "verified" to "evaluated", here and throughout, since in the strict sense of the word, the model has certainly not been verified.

Response:

Corrected in the text.

Q18. 22.2: I note the authors stop short of including anything regarding the AERONET evaluation in the conclusions – an indication that this part of the manuscript could be removed without impact.

Response:

Yes, the contents relevant with MODIS AOD and AERONET AOD have been removed.

Retrospective analysis of 2015-2017 winter-time $PM_{2.5}$ in China: response to emission regulations and the role of meteorology

Dan Chen[1*], Zhiquan Liu[2*], Junmei Ban[2], Pusheng Zhao[1], and Min Chen[1]

[1]Institute of Urban Meteorology, China Meteorological Administration, Beijing, 100089, China
[2]National Center for Atmospheric Research, Boulder, CO, 80301, USA

February, 2019

* *Corresponding author: Dr. Zhiquan Liu (liuz@ucar.edu) and Dr. Dan Chen (dchen@ium.cn)*

Abstract

To better characterize anthropogenic emission-relevant aerosol species, the GSI-WRF/Chem data assimilation system was updated from the GOCART aerosol scheme to the MOSAIC-4BIN scheme. Three years (2015-2017) of wintertime (January) surface $PM_{2.5}$ observations from 1600+ sites were assimilated hourly using the updated 3DVAR system. In the control experiment (without assimilation) using 2010_MEIC emissions, the modeled January averaged $PM_{2.5}$ concentrations were severely overestimated in the Sichuan Basin, Central China, Yangtze River Delta, and Pearl River Delta by 98-134, 46-101, 32-59, and 19-60 μg m^{-3}, respectively, indicating that the emissions for 2010 are not appropriate for 2015-2017, as strict emission control strategies were implemented in recent years. Meanwhile, underestimations of 11-12, 53-96, and 22-40 μg m^{-3} were observed in northeastern China, Xinjiang and the Energy Golden Triangle, respectively. The assimilation experiment significantly reduced both high and low biases to within ± 5 μg m^{-3}.

The observations and the reanalysis data from the assimilation experiment were used to investigate the year-to-year changes and the driving factors. The role of emissions was obtained by subtracting the meteorological impacts (by control experiments) from the total combined differences (by assimilation experiments). The results show a reduction in $PM_{2.5}$ of approximately 15 μg m^{-3} for the month of January from 2015 to 2016 in the North China Plain (NCP), but meteorology played the dominant role (contributing a reduction of approximately 12 μg m^{-3}). The change (for January) from 2016 to 2017 in NCP was different; meteorology caused an increase in $PM_{2.5}$ of approximately 23 μg m^{-3}, while emission control measures caused a decrease of 8 μg m^{-3}, and the combined effects still showed a $PM_{2.5}$ increase for that region. The analysis confirmed that emission control strategies were indeed implemented and emissions were reduced in both years. Using a data assimilation approach, this study helps identify the reasons why emission control strategies may or may not have an immediately visible impact. There are still large uncertainties in this approach, especially the inaccurate emission inputs, and neglecting aerosol-meteorology feedbacks in the model can generate large uncertainties in the analysis as well.

1. Introduction

Anthropogenic $PM_{2.5}$ (fine particulate matter with an aerodynamic diameter smaller than 2.5 μm) is known as a robust indicator of mortality and other negative health effects associated with ambient air pollution.

$PM_{2.5}$ components are originate not only from primary emissions but also from secondary formations through various precursors (e.g., SO_2, NO_x, and VOCs). Regional haze with extremely high $PM_{2.5}$ concentrations (exceeding the WHO standard tenfold) has become the primary air quality concern in China, especially over northern China (e.g., Wang *et al.* 2014a, 2014b; Han *et al.* 2015; Sun *et al.* 2015). To control $PM_{2.5}$ pollution and improve the overall air quality, a series of strict pollution control strategies have been implemented by the government since 2010, including the *Guiding Options on Promoting the Joint Prevention and Control of Air Pollution to Improve Regional Air Quality* (The Central Government of the People's Republic of China, 2010) and the *Atmospheric Pollution Prevention and Control Action Plan* (The Central Government of the People's Republic of China, 2013), in which the government stated that environmental-related equipment (for flue-gas desulfurization, selective catalyst reduction, exhaust dust removal, etc.) are mandatory for both industries and vehicles. In addition to long-term pollution control strategies, different emergency measures under different pollution alerts were also implemented occasionally. For example, the production of large industrial sources (coal burning and cement) was limited to reduce emissions, construction sites were restricted to prevent fugitive dust pollution, and traffic restrictions were implemented on even- and odd-numbered license plates. These emission control strategies were stricter and implemented more often in northern China than anywhere else in winter, when haze events occur more frequently. These control strategies were expected to reduce both the concentrations of significant precursors (e.g., SO_2, NO_x) and the emissions of $PM_{2.5}$.

Despite these strict emission control strategies, the ambient $PM_{2.5}$ concentrations in major cities still fluctuated during the wintertime from year to year. For example, the overall January $PM_{2.5}$ concentrations in 74 cities generally decreased from 2015 to 2016, but the concentrations in January 2017 were still higher than those in 2016 (*Ambient Air Quality Monthly Report 2015-01/2016-01/2017-01*, http://www.cnemc.cn/kqzlzkbgyb2092938.jhtml). While annual emission reduction trends were expected from 2015 to 2017, the overall increase in the surface concentrations observed in January 2017 contradicted these expectations, thereby indicating that other factors (especially meteorological conditions) in addition to emissions may play important roles. Some studies have attempted to investigate the variability of air pollution

and the effects of climate change on wintertime air pollution by using statistical data. Li *et al.* (2016) indicated that wintertime fog-haze days across central and eastern China have a close relationship with the East Asian winter monsoon. Zuo *et al.* (2015) concluded that the significant weakening and strengthening of the Siberian high and East Asian trough, respectively, are the two main factors for the occurrence of extreme cold and extreme warm events over China in winter, when warm air boosts air pollution. In addition to utilizing statistical methodology, it is necessary to distinguish the roles of emissions and meteorology to further investigate the driving factors of interannual air pollution changes.

Regional air quality models are important tools for scientifically understanding the formation of haze events, technically constructing forecasts, and evaluating the effects of control strategies. For regional modeling studies, emission inventories are important for reflecting the emission inputs into the atmosphere. Generally, an emission inventory is based on a "bottom-up" methodology, thereby relying on the statistics of energy activity and emission factors, etc. However, uncertainties in energy statistics can cause variations in the emission estimates (Zhao *et al.*, 2017; Hong *et al.*, 2017; Zhi *et al.*, 2017). For regional modeling applications, the total emissions based on statistics are spatially and temporally distributed according to relevant factors (He, 2012). Nevertheless, the occasional emission control strategies implemented in winter can cause large uncertainties in not only total emission estimations but also spatial and temporal allocations, which would lead to large biases in the model simulations.

In addition to the uncertainties in emission inventories, deficiencies in the model chemistry can also cause model uncertainties. Increasing numbers of observations have revealed that anthropogenic emission-relevant aerosol species, such as sulfate, nitrate and ammonium (denoted as SNA), are the predominant inorganic species in the wintertime $PM_{2.5}$ in China (Wang *et al.*, 2014c; Yang *et al.*, 2015). Various reaction paths during haze events have also been proposed (e.g. Zheng *et al.*, 2015; Cheng *et al.*, 2016; Wang *et al.*, 2016; Li *et al.*, 2017; Moch *et al.*, 2018; Wang *et al.*, 2018; Shao *et al.*, 2019). For example, Moch *et al.* (2018) used a 1-D model and revealed the importance of aqueous-phase chemistry of HCHO and S(IV) in cloud droplets by forming a S(IV)-HCHO adduct, hydroxymethane sulfonate. Shao *et al.* (2019) implemented four

heterogeneous sulfate formation mechanisms (via H_2O_2, O_3, NO_2, and transition metal ions on aerosols) into GEOS-Chem model which partially reduced the modeled low bias in sulfate concentrations. However, a scientific consensus regarding the importance of the reaction paths has not yet been reached partially due to the uncertainties of aerosol liquid water content, pH, and ionic strength etc. The original WRF/Chem model with either the Goddard Chemistry Aerosol Radiation and Transport (GOGART, Chin *et al.*, 2000, 2002) or the Model for Simulating Aerosol Interactions and Chemistry (MOSAIC)-4BIN aerosol scheme failed to reproduce the highest $PM_{2.5}$ concentrations; it is assumed that this failure is due to missing heterogeneous/aqueous reactions. In Chen *et al.* (2016, hereafter Chen16), we added three heterogeneous reactions (SO_2-to-H_2SO_4 and NO_2/NO_3-to-HNO_3) to the WRF/Chem model based on the MOSAIC-4BIN aerosol scheme. Although the reaction paths may still not be comprehensively understood, the new MOSAIC-4BIN aerosol scheme significantly improved the simulation of sulfate, nitrate, and ammonium on polluted days in terms of the concentrations of those species and their partitioning.

Data assimilation (DA), that is, the combination of observations with numerical model output, has proven to be skillful at improving aerosol forecasts (e.g., Collins *et al.*, 2001; Pagowski *et al.*, 2010; Liu *et al.*, 2011; Liu *et al.*, 2016; Zhang et al., 2016). Liu *et al.* (2011, hereafter Liu11) implemented DA on AOD estimates within the National Centers for Environmental Prediction (NCEP) gridpoint statistical interpolation (GSI) three-dimensional variational (3DVAR) DA system coupled with the GOCART aerosol scheme within the Weather Research and Forecasting/Chemistry (WRF/Chem) model (Grell *et al.*, 2005). Schwartz *et al.* (2012, hereafter S12) and Jiang *et al.* (2013, hereafter Jiang13) extended the above system to assimilate surface $PM_{2.5}$ and PM_{10}. The evaluation results demonstrated improved aerosol forecasts from the DA system in studies over East Asia and the United States.

Following Liu11, S12 and Chen16, we updated the GSI-WRF/Chem system by changing from the GOCART aerosol scheme to the MOSAIC-4BIN aerosol scheme to better characterize the complex $PM_{2.5}$ pollution in China. We applied the updated system to assimilate $PM_{2.5}$ concentrations of January 2015, 2016 and 2017 for two purposes: 1) to reproduce the $PM_{2.5}$ output by the DA system and 2) to investigate the

different impacts of meteorological conditions and emissions on the $PM_{2.5}$ pollution in different years. In this paper, section 2 provides descriptions of the model, observations and methodology and addresses the updated GSI-WRF/Chem-coupled DA system with the MOSAIC-4BIN aerosol scheme. In section 3, the assimilation results for the $PM_{2.5}$ concentrations from January 2015, 2016 and 2017 are presented and compared with surface observations ($PM_{2.5}$ total mass) to evaluate the DA system. In contrast to previous applications emphasizing the forecast skill improvement achieved by the DA system, we fully utilized reanalysis data to investigate the driving factors of pollution and to differentiate the roles played by meteorological conditions and emissions in different years by analyzing the reanalysis data and model simulations. The results are given in section 4, and the conclusions are given in section 5.

2. Model description, observations and methodology

The WRF/Chem settings are very similar to those of Chen16, although Chen16 focused on the SNA aerosols in the North China Plain during October 2014; in addition, several heterogeneous reactions were newly added to the original chemistry modules to improve the SNA simulation performance. The DA system used herein was based upon the NCEP GSI system extended by Liu11 and S12. We assimilated surface $PM_{2.5}$ observations, and the only difference is that the MOSAIC-4Bin aerosol scheme (32 PM species) was chosen for the WRF/Chem model instead of the GOCART aerosol scheme. Thus, the 3-D mass mixing ratios of those MOSAIC species at each grid point composed the analysis (or control) variables in the GSI 3DVAR minimization process.

Here, only a brief summary of the WRF/Chem configuration is provided below prior to a description of the updated GSI DA system and the settings used in this work. The most important differences are noted, e.g., the forward operator for observations in the GSI system.

2.1 WRF/Chem model and emissions

As in Chen16, version 3.6.1 of the WRF/Chem model was used in this study (Grell *et al.,* 2005; Fast *et al.,* 2006). The physical parameterizations employed in the WRF/Chem model were identical to those of

Chen16, and they are listed in Table 1. The Carbon-Bond Mechanism version Z (CBMZ) and the Model for Simulating Aerosol Interactions and Chemistry (MOSAIC) were used as the gas phase and aerosol chemical mechanisms, respectively, in this study. The aerosol species in MOSAIC are defined as black carbon (BC), organic compounds (OC), sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), sodium (NA), chloride (CL) and other inorganic compounds (OIN). We used 4 size bins with aerosol diameters ranging from 0.039-0.1, 0.1-1.0, 1.0-2.5, and 2.5-10 μm. The 24 variables in the first three bins (8 species times 3 bins) consist of the $PM_{2.5}$ total. The newly added relative humidity (RH)-dependent SO_2-to-H_2SO_4 and NO_2/NO_3-to-HNO_3 heterogeneous reactions (details are provided in Chen16) were also applied in the simulations.

The model domain with a 40-km horizontal grid spacing covers most of China and the surrounding regions (Fig. 2), and there are 57 vertical levels extending from the surface to 10 hPa. The simulation started from Dec. 20 of the previous year; the first eleven days were treated as a spin-up period and were not used in our analyses.

Table 1. WRF/Chem model configuration.

Aerosol scheme	MOSAIC (4 bins) (Zaveri *et al.*, 2008)
Photolysis scheme	Fast-J (Wild *et al.*, 2000)
Gas phase chemistry	CBM-Z (Zavier *et al.*, 1999)
Cumulus parameterization	Grell 3D scheme
Short-wave radiation	Goddard Space Flight Center Shortwave radiation scheme (Chou and Suarez, 1994)
Long-wave radiation	RRTM (Mlawer *et al.*, 1997)
Microphysics	Single-Moment 6-class scheme (Grell and Devenyi, 2002)
Land-surface model	NOAH LSM (Chen and Dudhia, 2001)
Boundary layer scheme	YSU (Hong *et al.*, 2006)
Meteorology initial and boundary conditions	GFS analysis and forecast every 6 hour
Initial condition for chemical species	11-day spin-up
Boundary conditions for chemical species	averages of mid-latitude aircraft profiles (McKeen *et al.*, 2002)
Dust and sea salt Emissions	GOCART

As in Chen16, the Multi-resolution Emission Inventory for China (MEIC) (Zhang *et al.*, 2009; Lei *et al.*, 2011; He 2012; Li *et al.*, 2014) for January 2010 was used as the emission input, as it is the only emission

inventory that was publicly available when the study was conducted. The original grid spacing of the MEIC is 0.25° × 0.25°, and this inventory has been processed to match the model grid spacing (40 km). The spatial distributions of primary $PM_{2.5}$, SO_2, NO_x and NH_3 emissions are shown in Fig. 2. The MEIC-2010 emission inventory has already been applied in other studies (e.g., Wang *et al.*, 2014a; Zheng *et al.*, 2015) for simulations over China in the past few years; these recent studies found that the MEIC provides reasonable estimates of total emissions but is subject to uncertainties in the spatial allocations of these emissions over small spatial scales. For our simulations, uncertainties may also arise from two other sources: the difference between the emission base year (2010) and our simulation period (2015 through 2017) and the monthly allocations. As the Chinese government has implemented strict control strategies to ensure an improved air quality during the winter season since 2013, significant reductions in emissions, including primary PM and precursor compounds (SO_2 and NO_x), in regions with the strict implementation of these policies relative to the year 2010 are expected for our simulation period. A reduction in SO_2 pollution of approximately 50% was observed from 2012-2015 for the North China Plain from OMI satellite data (Krotkov *et al.*, 2016). National anthropogenic emission reductions of approximately 67%, 17%, and 35% from 2012-2017 for SO_2, NO_x, and $PM_{2.5}$, respectively, were assumed by the bottom-up EI methodology (Zheng *et al.*, 2018). However, the expansion and relocation of the energy industry caused emission increases in northwestern China (Ling *et al.*, 2017). In addition, the uncertainties of allocated emissions in the winter season will be much larger than those in other seasons. For example, Zhi *et al.* (2017) conducted a village energy survey and revealed an enormous discrepancy in the amount of rural raw coal used for winter heating in northern China, implying an extreme underestimation of rural household coal consumption by the China Energy Statistical Yearbooks. These changes and uncertainties of emissions in the model would introduce errors into the NO_DA simulation. However, the inhomogeneous spatial changes and large uncertainties of seasonal allocations made it difficult to simply scale the original emission inventory for our study period.

2.2 Updated GSI 3DVAR DA system

The NCEP's GSI 3DVAR DA system was used to assimilate surface $PM_{2.5}$ observations. The GSI 3DVAR

DA system calculates a best-fit analysis considering the observations (hourly surface $PM_{2.5}$ concentrations in our case) and background fields (a 1-hr short-term WRF/Chem forecast in our case) weighted by their error characteristics. The GSI 3DVAR DA system produces an analysis in a model grid space through the minimization of a scalar objective function J(x) given by

$$J(\mathrm{x}) = \frac{1}{2}(\mathrm{x} - \mathrm{x_b})^{\mathrm{T}}\mathrm{B}^{-1}(\mathrm{x} - \mathrm{x_b}) + \frac{1}{2}[H(\mathrm{x}) - \mathrm{y}]^{\mathrm{T}}\mathrm{R}^{-1}[H(\mathrm{x}) - \mathrm{y}], \quad (1)$$

where $\mathbf{x_b}$ denotes the background vector (with dimension m), **y** is a vector of observations (with dimension p), and **B** and **R** represent the background and observation error covariance matrices of dimensions $\mathrm{m} \times \mathrm{m}$ and $\mathrm{p} \times \mathrm{p}$, respectively. The covariance matrices determine the relative contributions of the background and observation terms to the final analysis. *H* is the potentially nonlinear "observation operator" that interpolates the model grid point values into observation spaces and converts model-predicted variables into observed quantities.

2.2.1 $PM_{2.5}$ observation operator

In our updated DA system, GSI was used to assimilate surface $PM_{2.5}$ total mass observations, whereas the WRF/Chem model predicts the $PM_{2.5}$ total mass as different prognostic variables depending on the chosen aerosol scheme. As we chose the MOSAIC-4Bin aerosol scheme, the analysis variables here were the 3D mass mixing ratios of the 24 MOSAIC aerosol variables at each grid point. The model-simulated $PM_{2.5}$ observations $M_{PM_{2.5}}$ were computed by summing the 24 species as

$$M_{PM_{2.5}} = \sum_{i=1}^{3}[BC\_i + OC\_i + SO4\_i + NO3\_i + NH4\_i + CL\_i + NA\_i + OIN\_i], \quad (2)$$

where *i* denotes the bin number in the MOSAIC aerosol scheme, where the first three bins consist of the $PM_{2.5}$ total, and BC, OC, SO_4, NO_3, NH_4, NA, CL, and OIN denote black carbon, organic compounds, sulfate, nitrate, ammonium, sodium, chloride and other inorganic compounds, respectively. This formula is identical to that used in the WRF/Chem MOSAIC scheme to diagnose $PM_{2.5}$. The WRF-Chem-simulated aerosol mixing ratios of the species listed inside the brackets of Eq. 2 are in units of $\mu g\, kg^{-1}$, and thus, the dry air density ρ_d is multiplied to convert the units into $\mu g\, m^{-3}$ for consistency with the observations.

Since only surface $PM_{2.5}$ total mass observations were assimilated to analyze the 3D mass mixing ratios of 24 aerosol variables, the 3DVAR problem was underconstrained. Due to the lack of species and vertical information provided by the observations, the only mathematical solution is to utilize prior information from the model background. In the GSI system, the distribution of the analysis increments (the difference between the analysis and background) onto the different species was mostly model driven with the observation and background error covariance matrices acting as the main constraints. This speciated approach to aerosol DA within a variational system was introduced by Liu11 and further applied by S12 and Jiang13. By using individual aerosol species as the control variables, no assumptions were made regarding the contribution of each species' mass to the total aerosol mass or to the shapes of the vertical profiles.

2.2.2 $PM_{2.5}$ observations and errors

Hourly surface $PM_{2.5}$ observations for January 2015-2017 were obtained from the China National Environmental Monitoring Center (CNEMC). There are 1600+ sites in our modeling domain. As the 1600+ monitoring sites fall into 531 model grids, all observations within the same grid are averaged (as well as the latitude and longitude) for the purpose of performing statistical calculations and evaluation. The observation sites (Fig. 3) span mostly northern, central and eastern China, while the sites are relatively sparse in western China.

The observation error covariance matrix **R** in Eq. (1) contains both measurement and representativeness errors. Pagowski *et al.* (2010) used a measurement error (ε_0) of 2 $\mu g\ m^{-3}$. To associate higher $PM_{2.5}$ values with larger measurement errors, S12 defined the measurement error as $\varepsilon_0 = 1.5 + 0.0075 \times M_{PM_{2.5}}$, where $M_{PM_{2.5}}$ denotes an AIRNow $PM_{2.5}$ observation and the units of each term are $\mu g\ m^{-3}$. According to the $PM_{2.5}$ Auto-Monitoring Instrument Technical Standard and Requirement (China National Environmental Monitoring Center, 2013), three continuous online monitoring methods, namely, a beta-ray plus dynamic heating system, a beta-ray plus dynamic heating system plus light scattering system, and a tapered element oscillating microbalance plus filter dynamic measurement system, are used at the national monitoring sites to satisfy the

requirements that the display resolution should be less than 1 $\mu g\ m^{-3}$ and the error should be less than 5 $\mu g\ m^{-3}$(within 24 hours). To reflect the confidence in the hourly observations, the measurement error ε_0 in this study is defined as $\varepsilon_0 = 1.0 + 0.0075 \times M_{PM_{2.5}}$, where $M_{PM_{2.5}}$ denotes a $PM_{2.5}$ observational value (unit: $\mu g\ m^{-3}$).

Representativeness errors reflect the inaccuracies in the forward operator and in the interpolation from the model grid to the observation location. Elbern *et al.* (2007), Pagowski *et al.* (2010), S12 and Jiang13 defined the representativeness error (ε_r) as

$$\varepsilon_r = \gamma \varepsilon_0 \sqrt{\frac{\Delta x}{L}}, \quad (3)$$

where γ is an adjustable parameter scaling ε_0 ($\gamma = 0.5$ was used here), Δx is the grid spacing (40 km in our case) and L is the radius of influence of an observation (set to 2 km for urban sites). These parameter settings were based on the performance of sensitivity tests. The total $PM_{2.5}$ error ($\varepsilon_{PM2.5}$) is defined as

$$\varepsilon_{PM2.5} = \sqrt{\varepsilon_0^2 + \varepsilon_r^2}, \quad (4)$$

which constituted the diagonal elements in the **R** matrix. The $PM_{2.5}$ data were provided in near-real time without any data quality control. To ensure the data quality before DA, $PM_{2.5}$ observational values larger than 1000 $\mu g\ m^{-3}$ (the maximum display limit of the monitoring system) were deemed unrealistic in the filter process and thus were not assimilated. In addition, observations leading to innovations/deviations (observations minus the model-simulated values determined from the first-guess fields) exceeding 500 $\mu g\ m^{-3}$ were also omitted for the stability of the DA optimization step.

2.2.3 Background error covariance

Similar to Jiang13, the background error covariance (BEC) statistics for each analysis variable required by the 3DVAR algorithm were computed by utilizing the NMC method (Parrish and Derber, 1992) based upon the one-month WRF/Chem forecast for January 2015. No cross-correlation between different species was considered. The standard deviations and horizontal/vertical correlation length scales of the background errors (separated for each aerosol species) were calculated using the method described by Wu *et al.* (2002). These

data were used as constraints for the distributions of the PM components. It is important to have phenomena-specific background error statistics to allow for an appropriate adjustment of individual species. The domain-averaged standard deviations of the background errors for 6 aerosol species (BC, OC, SO_4, NO_3, NH_4, and OIN) in the first three size bins are shown in Fig. 1 as a function of the vertical model level; CL and NA are not shown here because they are excessively small relative to the other PM species. By using the MOSAIC aerosol scheme, the characteristics of different aerosol species in different size bins are more appropriate for the China region in the model. As shown in Fig. 1, the standard deviations of different aerosol species errors are different in the three size bins; the errors of NO_3, OIN and SO_4 are relatively larger than those of the other species in the three size bins; OC is also important, especially in the second (0.1-1.0 μm) and third (1.0-2.5 μm) size bins. The larger background errors of those species allowed the field to be better adjusted, which was crucial for the aerosol analyses in this study.

2.3 Experimental design

We conducted two sets of experiments (NO_DA and CONC_DA) for January 2015, 2016 and 2017. In both cases, the MEIC_2010 emission inventory was used. The NO_DA experiment initialized a new WRF/Chem forecast every 6 hr starting at 00 UTC on 20 December of the previous year to spin up the aerosol fields and was run through 23 UTC on 31 January. Only the simulations in January were used for the analysis. In the NO_DA experiment, the chemical/aerosol fields were simply carried over from cycle to cycle (similar to a continuous aerosol forecast), while the meteorological IC/BC were updated from GFS analysis data every 6 hr to prevent the meteorological simulation from drifting. For CONC_DA, the GSI 3DVAR system updated the MOSAIC aerosol variables every hour starting from 00 UTC on 1 January. The background of the first cycle at 00 UTC on 1 January was obtained from the NO_DA experiment, and all subsequent cycle were derived from the previous cycle's 1-hr forecast. In CONC_DA, the GFS analysis data were interpolated from a 6-hr frequency to a 1-hr frequency and were then used to update the meteorological IC/BC in each 1-hr cycle. The newly added heterogeneous reactions were activated in both sets of experiments.

2.4 Distinguishing the impacts of meteorological conditions and emissions

As introduced in section 1, interannual air quality changes are strongly influenced by both emissions and meteorological conditions. It is challenging to distinguish and quantify the impacts of these two aspects solely based on observations or modeling. In our case, the impacts of meteorological conditions are diagnosed by analyzing the differences between two sets of modeling simulations (with the same emission inventory but different meteorology conditions). For NO_DA, the emission inputs for January of the three years (2015-2017) were all from the MEIC_2010 emission inventory, and the only differences among the simulations of these three months were the meteorological conditions, which were acquired from the GFS 6-hr analysis data. Therefore, we can assume that the differences in the simulated NO_DA $PM_{2.5}$ concentrations among the three months were driven purely by differences in the meteorological conditions (similar to Xu *et al.* 2017). However, it is difficult to distinguish the impacts of emissions by using the same approach. As temporary emission control measures were applied according to the pollution severity (alarm level), the emission reduction ratios actually continued to change during the winter season, and thus, no exact emission reduction ratios were provided for those days. Nevertheless, the simulation approach with different emission scenarios is simply impossible when lacking exact emission reduction ratios. Instead, we subtracted the meteorological effects from the total effects by utilizing the reanalysis data and pure model simulations. The CONC_DA result, in which the hourly surface $PM_{2.5}$ observations from 531 groups of sites were utilized, can be treated as a reanalysis dataset that reflects the actual conditions (very close to the observations). Therefore, the differences in the assimilated CONC_DA $PM_{2.5}$ concentrations among the three months reflect the combined effects of both meteorological conditions and emissions. As the two experiments were generated on gridded aerosol fields, we can separate the effects of emissions from the collective effect by subtracting the NO_DA differences from the CONC_DA differences. Hence, we can better comprehend how meteorological conditions and emissions play different roles in driving the changes among the three years. Table 2 illustrates this approach by taking 2015 and 2016 as an example. However, some uncertainties might be associated with this approach, as will be discussed in detail in section 4.2.

Table 2. The approach used to distinguish the different impacts of meteorological conditions and emissions by calculating them from different scenarios (taking 2015 and 2016 as an example).

A. Assimilated total changes	CONC_DA_2016-CONC_DA_2015	Reflecting the combined effect of all driving factors, e.g., emissions and meteorological conditions, from 2015 to 2016
B. Simulated changes due to meteorological differences	NO_DA_2016-NO_DA_2015	As NO_DA_2015 and NO_DA_2016 were conducted with same emissions but different meteorological conditions, the differences reflect the effects due to meteorological differences from 2015 to 2016
C. Calculated changes due to emission differences = (A-B)	(CONC_DA_2016-CONC_DA_2015) - (NO_DA_2016-NO_DA_2015)	Mostly reflecting the effects from emission differences between 2015 and 2016

3. Evaluation of the assimilated $PM_{2.5}$

This section presents the results from the NO_DA and assimilation experiments outlined above. In slight contrast to S12 and Jiang13, our purpose was to reproduce the spatial-temporal variations in the surface $PM_{2.5}$ within the reanalysis dataset rather than to provide the IC of aerosol fields for improving forecasts.

Figure 3 shows the observed and modeled monthly averages of the surface $PM_{2.5}$ for January 2015, 2016 and 2017. Nine regions are illustrated as rectangles in the figure: North China Plain (NEC), northeastern China (NEC), Energy Golden Triangle (EGT), Xinjiang (XJ), Fenwei Plain (FWP), Sichuan Basin (SB), Central China (CC), Yangtze River Delta (YRD), and Pearl River Delta (YRD). Both the observations and the model show that high values are mostly observed in NCP, FWP, SB and CC. In the NO_DA case, the model results are overpredicted in SB, NCP and CC for all three months, while the overestimations are more severe in SB. The NO_DA case generally overestimates (underestimates) the surface $PM_{2.5}$ in NCP, SB and CC (XJ and FWP) in the three years, potentially indicating that the 2010 emissions are not appropriate for the 2015-2017 simulations with overestimations (underestimations). As discussed in section 2.1, the large area of overestimation is consistent with the national reductions in SO_2, NO_x and $PM_{2.5}$ anthropogenic emissions (Zheng *et al.*, 2018); however, the underestimations in XJ and FWP also indicate the introduction of new emission sources to these two regions.

Compared to the NO_DA case, the CONC_DA assimilation experiment effectively reproduces the spatial distribution of surface $PM_{2.5}$ for the three months in terms of the relatively higher values observed in NCP, SB and CC and in some "hot spots" (NEC, FWP, and XJ), which are closer to the observations.

Basic statistical measures, including the bias (BIAS), standard deviation (STDV), root-mean-square error (RMSE) and correlation coefficient (CORR), were applied to evaluate the experiments. Figure 4 shows the time series of the BIAS, STDV and RMSE for all the data used in the entire domain. The statistics were calculated for each 1-hr DA cycle. After quality control, the number of $PM_{2.5}$ observations used in the DA process differed; the number of observations was normally approximately 500-520 but reached a minimum of 320-450 occasionally due to the data availability. From the time series, we can see that the BIAS, STDV and RMSE are greatly improved in the CONC_DA case. The maximum BIAS values are approximately 50 $\mu g\ m^{-3}$ for January 2015 and approximately 80 $\mu g\ m^{-3}$ for 2016 and 2017 in NO_DA, while they are reduced to approximately $\pm 5\ \mu g\ m^{-3}$ in CONC_DA. The STDV and RMSE are also reduced by at least 50% most of the time.

Figure 5 shows the spatial distributions of the error statistics (BIAS, RMSE and CORR) at each observational site (with more than 2/3 valid data in the month) in January 2015, 2016 and 2017. We start with 2015 and then address the differences with comparisons in 2016 and 2017. In 2015 in the NO_DA case, the surface $PM_{2.5}$ concentrations are generally overestimated by 20-60 $\mu g\ m^{-3}$ in eastern China (NCP, SB, CC, PRD and YRD) but are underestimated in NEC, FWP, EGT and especially XJ. The high/low BIAS values in eastern/western China are greatly corrected in CONC_DA. Consistent with the BIAS changes in CONC_DA, the RMSE and CORR distributions in eastern China and NEC are also greatly improved; the RMSE is reduced by at least 50%, and the CORR increases to almost above 0.8-0.9. The inhomogeneous distributions of the BIAS in NO_DA in 2016 and 2017 are very similar to that in 2015 (overestimated in eastern China but underestimated in NEC, EGT and XJ). However, the high biases in CC and PRD and the low biases in XJ are even larger in 2016 and 2017. Similar to the comparisons between NO_DA and CONC_DA for the year 2015, improvements are generally achieved for almost all the regions in both 2016 and 2017. The statistics for the 9

regions are listed in Table 3.

Table 3. Statistics of the observed and model-simulated surface $PM_{2.5}$ for January 2015, 2016 and 2017 in 9 regions (units are μg m^{-3} for BIAS and RMSE).

Statistics	Sites	Pairs of data	BIAS		RMSE		CORR	
			NO_DA	CONC_DA	NO_DA	CONC_DA	NO_DA	CONC_DA
2015								
NCP	67	46699	19.38	2.08	68.09	24.26	0.72	0.96
NEC	30	20910	-11.94	-1.04	49.47	21.11	0.59	0.93
EGT	28	19516	-40.43	5.28	60.62	19.45	0.37	0.90
XJ	19	13243	-53.76	4.16	71.69	19.74	0.40	0.94
FWP	27	18819	4.05	1.75	56.71	23.05	0.63	0.93
SB	48	33456	98.02	0.61	125.76	20.76	0.55	0.94
CC	49	34153	46.94	-0.38	81.31	21.18	0.46	0.93
YRD	34	23698	32.22	-0.43	59.90	15.14	0.73	0.96
PRD	20	13940	19.36	-0.03	47.81	9.10	0.24	0.95
2016								
NCP	67	46699	20.90	1.41	57.77	20.74	0.78	0.96
NEC	30	20910	-11.05	0.04	40.91	16.08	0.57	0.94
EGT	28	19516	-22.55	0.69	39.63	13.75	0.42	0.90
XJ	19	13243	-72.92	0.25	98.19	27.16	0.51	0.96
FWP	27	18819	-3.51	1.51	62.04	26.01	0.76	0.94
SB	48	33456	134.63	2.77	165.38	15.49	0.51	0.92
CC	49	34153	86.28	1.89	109.09	18.76	0.46	0.92
YRD	34	23698	46.13	1.03	62.11	13.40	0.73	0.95
PRD	20	13940	59.79	2.05	74.76	6.51	0.04	0.91
2017								
NCP	67	46699	25.75	2.35	82.31	28.91	0.74	0.95
NEC	30	20910	-11.38	0.01	53.38	21.35	0.64	0.94
EGT	28	19516	-26.88	1.40	48.83	16.96	0.41	0.90
XJ	19	13243	-95.92	3.82	125.09	35.65	0.51	0.96
FWP	27	18819	-6.78	-1.02	89.26	31.69	0.65	0.94
SB	48	33456	122.82	2.33	149.08	20.08	0.56	0.93
CC	49	34153	101.22	3.49	132.97	19.50	0.23	0.92
YRD	34	23698	59.31	2.40	78.02	12.32	0.63	0.93
PRD	20	13940	35.01	0.04	61.84	9.55	-0.16	0.94

4. Interannual changes during 2015 through 2017

Given reliable $PM_{2.5}$ reanalysis fields produced by assimilating surface $PM_{2.5}$ (CONC_DA), the change trends among the three years can be analyzed for not only scattered observational sites but also different

regions. To distinguish the roles of meteorological conditions and emissions in driving these changes, an analysis based on the NO_DA and CONC_DA simulations is performed. As assumed in section 2.4, meteorology-driven changes can be analyzed in the NO_DA simulations with different meteorological conditions but the same emission inventory for different years; however, the changes in the reanalysis data among different years are actually the combination of all the driving forces, including meteorological conditions and emissions. By analyzing both sets of simulations, we can attempt to distinguish the roles of meteorology and emissions in determining these changes.

4.1 Spatial distribution

The monthly mean $PM_{2.5}$ differences for January in the three years (2015-2017) are shown in Fig. 6 in terms of the surface concentrations measured at observational sites (Fig. 6a) and those from assimilation experiments (Fig. 6b). The surface observations are mostly reduced from 2015 to 2016 except for a few sites in the southern parts of NCP and FWP and in XJ. For the changes from 2016 to 2017, the surface observations increase at almost all the sites, especially the sites in the southern part of NCP; the only exceptions are the sites along the coastline in YRD. The assimilated (CONC_DA) differences are consistent with the surface observations insomuch that the decreasing trend from 2015 to 2016 and the increasing trend from 2016 to 2017 for most of the regions are reproduced. However, for the changes in Tibet, EGT and XJ, where observational sites are sparse, some "cold spots" were artificially generated by CONC_DA due to the scarcity of data and the horizontal length scale set in the assimilation. As already shown in Fig. 3 and indicated here again, January 2016 is the cleanest month among the three years.

4.2 The roles of meteorological conditions and emissions

The surface $PM_{2.5}$ concentrations from both the observations and the assimilation experiments show a decreasing trend from 2015 to 2016 but an increasing trend from 2016 to 2017 for most of the regions in eastern China (Fig. 6). The Chinese government has implemented a strict emission control strategy since 2013, especially in northern China, and thus, emission reductions are expected for each year following 2013. The

ambient response from 2015-2017 is contradictory if considering only the reductions in emissions and omitting the changes in meteorological conditions. There are two possible assumptions: the first is that the emission reduction target was not achieved from 2016 to 2017, and the second is that other factors in addition to emissions played more important roles.

The NO_DA differences among the different years are shown in Fig. 6c, which reflects the effect of meteorological condition changes (section 2.4). The effect due to emissions (the other major factor in addition to meteorological conditions) is given by subtracting the NO_DA differences from the CONC_DA differences (Fig. 6d). We can clearly see that the meteorology played two different roles from 2016 to 2017. It caused a decrease in the ambient concentrations for northern China (NCP and NEC) from 2015 to 2016 but induced a large increase for northern and central China (CC) from 2016 to 2017. This indicates that the meteorological conditions might have differed from 2016 to 2017. After considering the impacts of meteorological conditions, those of emission reductions are still confirmed for these two regions from 2016 to 2017. The contributions from both meteorological conditions and emissions in the 9 regions (defined in Fig. 3) were calculated, and the results are listed in Table 4. The calculations show a reduction of approximately 15-20 μg m^{-3} in $PM_{2.5}$ for the month of January from 2015 to 2016 in northern China (NCP and NEC), but the meteorology played a dominant role (contributing a reduction of approximately 12-21 μg m^{-3} in $PM_{2.5}$). The changes from 2016 to 2017 in NCP and NEC are completely different; meteorological conditions caused an increase in $PM_{2.5}$ of approximately 12-23 μg m^{-3}, and emission control measures caused a decrease of 1-8 μg m^{-3} in $PM_{2.5}$, while the combined effects still showed a $PM_{2.5}$ increase for that region. It is reasonable to say that emissions were indeed reduced for the northern regions from 2016 to 2017. However, the meteorology played an important role in offsetting those emission reductions and leading to an increase in surface concentrations in 2017.

Table 4. Modeled ambient $PM_{2.5}$ concentration changes for 2016-2015, 2017-2016 and 2017-2015 in 9 regions and the contributions of the meteorology (MET) and emissions (EMIS) calculated according to Table 2. Units: μg m^{-3}.

	2016-2015	2017-2016	2017-2015

	Total	MET	EMIS	Total	MET	EMIS	Total	MET	EMIS
NCP	-15.23	-12.52	-2.71	+14.91	+23.16	-8.25	-0.31	+10.65	-10.96
NEC	-20.09	-21.23	+1.14	+11.44	+12.61	-1.18	-8.66	-8.62	-0.04
EGT	-21.69	1.68	-23.37	+4.86	+3.81	+1.05	-16.83	+5.48	-22.31
XJ	+3.69	+0.07	+3.63	+1.85	+0.28	+1.57	+5.54	+0.34	+5.20
FWP	-7.05	-10.19	+3.13	+22.95	+25.62	-2.66	+15.90	+15.43	+0.47
SB	-18.75	+8.72	-27.48	+10.31	+4.02	+6.29	-8.45	+12.74	-21.19
CC	-21.80	+14.73	-36.54	+9.35	+19.36	-10.01	-12.45	+34.09	-46.54
YRD	-10.43	-3.03	-7.40	-11.45	-2.93	-8.52	-21.88	-5.96	-15.92
PRD	-23.48	13.02	-36.50	+12.71	-6.12	+18.83	-10.77	+6.90	-17.67

It is worth noting that there are uncertainties in the simulation/assimilation processes. There are three sources of uncertainties in the NO_DA simulation. First, the emission inventories in the NO_DA simulations are obviously not accurate, which may introduce uncertainties into the analysis. Although the basic assumption required only that the emissions stay the same throughout the three years, emission inventory uncertainty-induced errors would be offset in the subtraction process when calculating the year-to-year differences. However it did generate uncertainties. For example, the emissions in SB, CC and PRD were generally overestimated (Fig. 3), which means that the variations in the ambient concentration might have been artificially amplified considering the meteorology impacts (Fig. 6c). In contrast, the emissions in XJ and FWP were underestimated (Fig. 3), and thus, the changes in the ambient concentrations due to meteorological conditions in these two regions might have diminished. From this point of view, if the fixed emissions are more accurate in those years, the results would be more reliable. In the case where "real" emissions are not available and the purpose is to evaluate the contribution of those emissions, uncertainties will be unavoidable and should be emphasized carefully. Second, the meteorological IC/BC conditions in the NO_DA simulations, which were obtained from GFS 6-hr analysis data, also have uncertainties. The biases in meteorological conditions might lead to uncertainties in the $PM_{2.5}$ analysis. Third, the deficiencies associated with the chemistry in the model also generate uncertainties, including missing reactions and the inaccurate parameterization of reactions. These three aspects all originate from the imperfections of current forward models. From another perspective, the accuracy of the CONC_DA assimilation experiment also affects the analysis. For example, the assimilation artificially made some "code spots" in Tibet, EGT and XJ, where observational sites are sparse; this could also induce biases. Finally, the contribution of aerosol-meteorology

feedback was not considered in our calculations. As noted by Gao *et al.* (2017), reduced aerosol feedbacks due to emission reductions accounted for approximately 10.9% of the total decrease in $PM_{2.5}$ concentrations in urban Beijing in their APEC study. In our current approach, this effect is integrated into the emissions in the subtracting process.

4.3 Meteorological changes in 2016 and 2017

It is interesting to see that meteorology played different roles in each of the three years. Here, we compared some meteorological parameters to explain the impacts of the meteorology. Differences in the monthly mean planetary boundary layer height (PBLH), surface pressure (PSFC), 2-meter temperature (T2), 2-meter relative humidity (RH2) and 10-meter wind speed in different years are given in Fig. 7. The statistics of the differences in these parameters in the 9 regions are listed in Table 5, which shows that the changes in the PSFC and T2 for the periods 2015-2016 and 2016-2017 are different over the whole region. Comparing the parameters between 2015 and 2016, the pressure system is stronger, the temperature is lower, and the wind speed is larger in most regions in the latter; these conditions are favorable for the dispersion of pollution. However, there are some unfavorable conditions, including a lower PBLH and a higher RH (and thus, more heterogeneous reactions with the high RH) in northern and southern China, which may offset the impacts of high pressure systems and low temperatures. Therefore, the combined impacts of these meteorological parameters caused a decrease in the ambient concentration in northern China and an increase in southern China from 2015 to 2016, as shown in Fig. 6. The meteorological changes are different from 2016 to 2017 with a weaker pressure system, higher temperature, smaller wind speed, and lower PBLH in most regions, which caused the pollution to accumulate. As suggested by recent studies, climate change has had important impacts on extreme haze events in northern China based on historical statistical approaches or climate models. Those studies (e.g., Li *et al.*, 2015, Zuo *et al.*, 2015) revealed that wintertime fog-haze days across central and eastern China have a close relationship with the East Asian winter monsoon; in addition, significant weakening (strengthening) of the Siberian high and East Asian trough are the two main factors for extreme cold events and extreme warm events throughout China in winter, while warmth boosts air pollution. Consistent with our

study, Zhao *et al.* (2018) noted that a stronger Siberian high period in January 2016 produced a significant decrease in $PM_{2.5}$ concentrations relative to those during weaker periods in other years. The abovementioned studies emphasized that climate change factors and the impacts of emission changes are still difficult to evaluate. Our study used the DA technique in combination with regional models and surface observations to distinguish the impacts of emissions and meteorological conditions to further investigate the year-to-year changes at the regional scale.

Table 5. Statistics of the meteorological differences by region for January 2015, 2016 and 2017.

	PBLH (meter)			PSFC (Pa)			T2 (degree)			RH2 (%)			WS10 (m/s)		
	2016 - 2015	2017 - 2016	2017 - 2015	2016 - 2015	2017 - 2016	2017 - 2015	2016 - 2015	2017 - 2016	2017 - 2015	2016 - 2015	2017 - 2016	2017 - 2015	2016 - 2015	2017 - 2016	2017 - 2015
NCP	27.9	-26.7	1.2	138.5	-30.2	108.4	-4.9	3.3	-1.6	3.0	5.1	8.1	1.15	-0.78	0.37
NEC	22.7	35.3	58.0	117.0	-58.7	58.3	-4.9	4.4	-0.5	-5.7	3.1	-2.6	0.96	-0.38	0.57
EGT	13.6	1.1	14.7	28.0	-8.4	19.7	-4.0	4.0	0.0	10.0	-14.9	-4.9	0.14	-0.50	-0.36
XJ	-0.9	-13.8	-14.7	151.3	-43.1	108.1	-1.3	-0.8	-2.1	5.5	-2.1	3.4	0.36	-0.14	0.22
FWP	67.7	-51.6	16.1	64.6	-12.2	52.4	-3.8	3.4	-0.4	2.8	-0.8	2.0	1.05	-1.00	0.06
SB	9.8	-13.2	-3.4	-15.9	15.9	0.1	-2.4	2.5	0.2	3.9	-1.8	2.0	0.43	-0.02	0.41
CC	34.8	-56.6	-21.9	82.8	-53.2	29.6	-2.5	2.1	-0.3	10.8	0.7	11.5	0.60	-0.07	0.53
YRD	64.7	-22.0	42.7	77.1	-27.8	49.2	-1.7	1.9	0.2	7.8	2.5	10.3	0.89	-0.40	0.49
PRD	-36.1	8.2	-27.9	-16.2	-60.1	-76.3	-0.5	2.4	1.9	11.9	-8.7	3.2	0.94	-0.48	0.46

5. Conclusions

To analyze the complex $PM_{2.5}$ pollution in China, the GSI-WRF/Chem aerosol data assimilation system was updated from the GOCART aerosol scheme to the MOSAIC-4BIN scheme, which is more appropriate for characterizing anthropogenic emission-relevant aerosol species. Three years (2015-2017) of wintertime (January) surface $PM_{2.5}$ observations from 1600+ sites were assimilated hourly using the updated 3DVAR system in the CONC_DA assimilation experiment. A parallel control experiment that did not employ DA (NO_DA) was also performed.

Both the control and the assimilation experiments were evaluated against the surface $PM_{2.5}$ observations. In the NO_DA experiment, in which the 2010_MEIC emission inventory was used, the modeled $PM_{2.5}$ were severely overestimated in the Sichuan Basin (SB), Central China (CC), Yangtze River Delta (YRD), and Pearl

River Delta (PRD) by 98-134, 46-101, 32-59, and 19-60 μg m^{-3}, respectively, which indicated that the emission estimates for 2010 are not appropriate for 2015-2017, as strict emission control strategies were implemented in recent years. Meanwhile, underestimations of 11-12, 53-96, and 22-40 μg m^{-3} were observed in northeastern China (NEC), Xinjiang (XJ) and the Energy Golden Triangle (EGT), respectively. The assimilation experiment significantly reduced the high biases of surface $PM_{2.5}$ in SB, CC, YRD, and PRD and the low biases in NEC and XJ with biases within ±5 μg m^{-3}.

Both the observation and the assimilation experiments showed decreasing ambient concentrations from 2015 to 2016 but increasing concentrations from 2016 to 2017 for most of the regions. To distinguish the important factors driving these changes, the reanalysis data from the assimilation experiment and the modeling results from the control experiment were analyzed. The results showed a reduction in $PM_{2.5}$ of approximately 15-20 μg m^{-3} for the month of January from 2015 to 2016 in northern China (NCP and NEC), but meteorology played the dominant role (contributing approximately 12-21 μg m^{-3} of the $PM_{2.5}$ reduction). The changes from 2016 to 2017 in NCP and NEC were different; meteorological conditions caused an increase in $PM_{2.5}$ of approximately 12-23 μg m^{-3}, while emission control measures caused a decrease of 1-8 μg m^{-3}, and the combined effects still showed a $PM_{2.5}$ increase for that region. The analysis confirmed that meteorology played different roles in 2016 and 2017: the higher pressure system, lower temperatures and higher PBLH in 2016 (compared with 2015) were favorable for pollution dispersion, whereas the situation was almost the opposite in 2017 (compared with 2016) and led to an increased $PM_{2.5}$ from 2016 to 2017, although emission control strategies were implemented in both years. After considering the impacts of the meteorology, the analysis showed that emissions were indeed reduced from 2015 to 2016 and 2017, especially in NCP for the year 2017 (although the surface concentrations increased that year). The analysis also showed that emissions increased in XJ and FWP.

There are still large uncertainties in this approach, such as the deficiencies of forward models (including inaccurate emission inputs, uncertainties in the meteorological IC/BC, and the chemistry mechanism) and the assimilation process, and the imperfection of the aerosol-meteorology feedbacks in the model simulation

generated large biases in the analysis. The most straightforward approach is thus to directly estimate the emissions by data assimilation, which will be the topic of a separate study.

Acknowledgement

This work was supported by the National Key R&D Program on Monitoring, Early Warning and Prevention of Major Natural Disasters under grant (2017YFC1501406), the National Natural Science Foundation of China (Grant No. 41807312) and Basic R&D special fund for central scientific research institutes (IUMKYSZHJ201701). NCAR is sponsored by US National Science Foundation.

References

Chen, D., Liu, Z. Q., Fast, J., and Ban, J. M.: Simulations of sulfate-nitrate-ammonium (SNA) aerosols during the extreme haze events over northern China in October 2014, Atmos. Chem. Phys., 16, 10707-10724, 10.5194/acp-16-10707-2016, 2016.

Chen, F., and Dudhia, J.: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Mon. Weather Rev., 129, 569-585, 2001.

Cheng, Y., Zheng, G., Wei, C., Mu, Q., Zheng, B., Wang, Z., Gao, M., Zhang, Q., He, K., Carmichael, G., Poschl, U., and Su, H.: Reactive nitrogen chemistry in aerosol water as a source of sulfate during haze events in China, Sci. Adv., 2, e1601530, doi:10.1126/sciadv.1601530, 2016.

China National Environmental Monitoring Center, Ambient Air Quality Monthly Report 2015-01/2016-01/2017-01; http://www.cnemc.cn/kqzlzkbgyb2092938.jhtml.

China National Environmental Monitoring Center, PM2.5 Auto-Monitoring Instrument Technical Standard and Requirement, Beijing, 2013.

Chin, M., Savoie, D. L., Huebert, B. J., Bandy, A. R., Thornton, D. C., Bates, T. S., Quinn, P. K., Saltzman, E. S., and De Bruyn, W. J.: Atmospheric sulfur cycle simulated in the global model GOCART: Comparison with field observations and regional budgets, J. Geophys. Res.-Atmos., 105, 24689-24712, doi:10.1029/2000jd900385, 2000.

Chin, M., Ginoux, P., Kinne, S., Torres, O., Holben, B. N., Duncan, B. N., Martin, R. V., Logan, J. A., Higurashi, A., and Nakajima, T.: Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and Sun photometer measurements, J. Atmos. Sci., 59, 461-483, doi:10.1175/1520-0469(2002)059<0461:Taotft>2.0.Co;2, 2002.

Chou, M.-D., and M.J. Suarez.: An efficient thermal infrared radiation parameterization for use in general circulation models, NASA Tech. Memo., TM 104606, vol. 3, 25 pp., NASA Goddard Space Flight Cent., Greenbelt, Md. 1994

Collins, W. D., Rasch, P. J., Eaton, B. E., Khattatov, B. V., Lamarque, J. F., and Zender, C. S.: Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: Methodology for INDOEX, J. Geophys. Res.-Atmos., 106, 7313-7336, doi:10.1029/2000jd900507, 2001.

Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-

dimensional variational inversion, Atmos Chem Phys, 7, 3749-3769, DOI 10.5194/acp-7-3749-2007, 2007.

Fast, J. D., Gustafson, W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model, J. Geophys. Res.-Atmos., 111, D21305, doi:10.1029/2005jd006721, 2006.

Gao, M., Liu, Z., Wang, Y., Lu, X., Ji, D., Wang, L., Li, M., Wang, Z., Zhang, Q., Carmichael., G. R.: Distinguishing the roles of meteorology, emission control measures, regional transport, and co-benefits of reduced aerosol feedbacks in "APEC Blue". Atmos. Environ., 167, 476-486, doi:10.1016/j.atmosenv.2017.08.054, 2017.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957-6975, doi:10.1016/j.atmosenv.2005.04.027, 2005.

Grell, G. A., and Devenyi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophys. Res. Lett., 29, Artn 1693.Doi 10.1029/2002gl015311, 2002.

Han, B., Zhang, R., Yang, W., Bai, Z., Ma, Z., and Zhang, W.: Heavy air pollution episodes in Beijing during January 2013: inorganic ion chemistry and source analysis using Highly Time-Resolved Measurements in an urban site, Atmos. Chem. Phys. Discuss., 15, 11111-11141, doi:10.5194/acpd-15-11111-2015, 2015.

He, K.B., Multi-resolution emission Inventory for China (MEIC): model framework and 1990-2010 anthropogenic emissions. In: Presented on the international Global Atmospheric Chemistry Conference, September 17-21, Beijing, China 2012

Hong, C. P., Zhang, Q., He, K. B., Guan, D. B., Li, M., Liu, F., and Zheng, B.: Variations of China's emission estimates: response to uncertainties in energy statistics, Atmos. Chem. Phys., 17, 1227-1239, 10.5194/acp-17-1227-2017, 2017.

Hong, S. Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Mon. Weather Rev., 134, 2318-2341, 2006.

Jiang, Z., Z. Liu, T. Wang, C. S. Schwartz, H.-C. Lin, and F. Jiang.: Probing into the impact of 3DVAR assimilation of surface PM10 observations over China using process analysis, J. Geophys. Res. Atmos., 118, 6738–6749, doi:10.1002/jgrd.50495, 2013.

Krotkov, N. A., McLinden, C. A., Li, C., Lamsal, L. N., Celarier, E. A., Marchenko, S. V., Swartz, W. H., Bucsela, E. J., Joiner, J., Duncan, B. N., Boersma, K. F., Veefkind, J. P., Levelt, P. F., Fioletov, V. E., Dickerson, R. R., He, H., Lu, Z. F., and Streets, D. G.: Aura OMI observations of regional SO2 and NO2 pollution changes from 2005 to 2015, Atmos Chem Phys, 16, 4605-4629, 10.5194/acp-16-4605-2016, 2016.

Liu, J., Han, Y., Tang, X., Zhu, J., Zhu, T.: Estimating adult mortality attributable to PM2.5 exposure in China with assimilated PM2.5 concentrations based on a ground monitoring network. Sci. Total Environ., 568, 1253-1262, 2016.

Liu, Z. Q., Liu, Q. H., Lin, H. C., Schwartz, C. S., Lee, Y. H., and Wang, T. J.: Three-dimensional variational assimilation of MODIS aerosol optical depth: Implementation and application to a dust storm over East Asia, J. Geophys. Res.-Atmos., 116, Artn D23206. doi:10.1029/2011jd016159, 2011.

Lei, Y., Zhang, Q., He, K. B., and Streets, D. G.: Primary anthropogenic aerosol emission trends for China, 1990-2005, Atmos. Chem. Phys., 11, 931-954, 10.5194/acp-11-931-2011, 2011.

Li, M., Zhang, Q., Streets, D. G., He, K. B., Cheng, Y. F., Emmons, L. K., Huo, H., Kang, S. C., Lu, Z., Shao, M., Su, H., Yu, X., and Zhang, Y.: Mapping Asian anthropogenic emissions of non-methane volatile organic compounds to multiple chemical mechanisms, Atmos. Chem. Phys., 14, 5617-5638, 10.5194/acp-14-5617-2014, 2014.

Li, Q., Zhang, R. H., and Wang, Y.: Interannual variation of the wintertime fog-haze days across central and eastern China and its relation with East Asian winter monsoon, Int. J. Climatol., 36, 346-354, 10.1002/joc.4350, 2016.

Li, G., Bei, N., Cao, J., Huang, R., Wu, J., Feng, T., Wang, Y., Liu, S., Zhang, Q., Tie, X., and Molina, L. T.: A possible pathway for rapid growth of sulfate during haze days in China, Atmos. Chem. Phys., 17, 3301-3316, doi:10.5194/acp-17-3301-2017, 2017.

McKeen, S. A., Wotawa, G., Parrish, D. D., Holloway, J. S., Buhr, M. P., Hubler, G., Fehsenfeld, F. C., and Meagher, J. F.: Ozone production from Canadian wildfires during June and July of 1995, J. Geophys. Res.-Atmos., 107, 4192, doi:10.1029/2001jd000697, 2002.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, J. Geophys. Res.-Atmos., 102, 16663-16682, 1997.

Moch, J. M., Dovrou, E., Mickley, L. J., Keutsch, F. N., Cheng, Y., Jacob, D. J., Jiang, J. K., Li, M., Munger, J. W., Qiao, X. H., and Zhang, Q.: Contribution of Hydroxymethane Sulfonate to Ambient Particulate Matter: A Potential Explanation for High Particulate Sulfur During Severe Winter Haze in Beijing, Geophys Res Lett, 45, 11969-11979, 10.1029/2018gl079309, 2018.

Pagowski, M., Grell, G. A., McKeen, S. A., Peckham, S. E., and Devenyi D.: Three dimensional variational data assimilation of ozone and fine particulate matter observations: some results using the Weather Research and Forecasting-Chemistry model and Grid-point Statistical Interpolation, Q. J. R. Meteorol. Soc., 136(653), 2010.

Sun, Y. L., Wang, Z. F., Du, W., Zhang, Q., Wang, Q. Q., Fu, P. Q., Pan, X. L., Li, J., Jayne, J., and Worsnop, D. R.: Long-term real-time measurements of aerosol particle composition in Beijing, China: seasonal variations, meteorological effects, and source analysis, Atmos. Chem. Phys., 15, 10149-10165, 10.5194/acp-15-10149-2015, 2015.

Schwartz, C. S., Z. Liu, H. C. Lin, and S. A. McKeen.: Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth, J. Geophys. Res.-Atmos., 117, 2012.

Shao, J., Chen, Q., Wang, Y., Lu, X., He, P., Sun, Y., Shah, V., Martin, R. V., Philip, S., Song, S., Zhao, Y., Xie, Z., Zhang, L., and Alexander, B.: Heterogeneous sulfate aerosol formation mechanisms during wintertime Chinese haze events: Air quality model assessment using observations of sulfate oxygen isotopes in Beijing, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1352, in review, 2019.

The Central People's Government of the People Republic of China. Guiding Opinions on Promoting the Joint Prevention and Control of Air Pollution to Improve Regional Air Quality, Beijing. 2010.

The Central People's Government of the People Republic of China. Atmospheric Pollution Prevention and Control Action Plan. 2013.

Wang, L. T., Wei, Z., Yang, J., Zhang, Y., Zhang, F. F., Su, J., Meng, C. C., and Zhang, Q.: The 2013 severe haze over southern Hebei, China: model evaluation, source apportionment, and policy implications, Atmos. Chem. Phys., 14, 3151-3173, 10.5194/acp-14-3151-2014, 2014a.

Wang, W., Maenhaut, W., Yang, W., Liu, X. D., Bai, Z. P., Zhang, T., Claeys, M., Cachier, H., Dong, S. P., and Wang, Y. L.: One-year aerosol characterization study for $PM_{2.5}$ and PM10 in Beijing, Atmos. Pollut. Res., 5, 554-562, 10.5094/Apr.2014.064, 2014b.

Wang, Y. S., Yao, L., Wang, L. L., Liu, Z. R., Ji, D. S., Tang, G. Q., Zhang, J. K., Sun, Y., Hu, B., and Xin, J. Y.: Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China, Sci. China Earth Sci., 57, 14-25, 10.1007/s11430-013-4773-4, 2014c.

Wang, G., Zhang, R., Gomez, M. E., Yang, L., Levy Zamora, M., Hu, M., Lin, Y., Peng, J., Guo, S., Meng, J., Li, J., Cheng, C., Hu, T., Ren, Y., Wang, Y., Gao, J., Cao, J., An, Z., Zhou, W., Li, G., Wang, J., Tian, P., Marrero-Ortiz, W., Secrest, J., Du, Z., Zheng, J., Shang, D., Zeng, L., Shao, M., Wang, W., Huang, Y., Wang, Y., Zhu, Y., Li, Y., Hu, J., Pan, B., Cai, L., Cheng, Y., Ji, Y., Zhang, F., Rosenfeld, D., Liss, P. S., Duce, R. A., Kolb, C. E., and Molina, M. J.: Persistent sulfate formation from London Fog to Chinese haze, Proc. Natl. Acad. Sci., 113, 13630-13635, doi:10.1073/pnas.1616540113, 2016.

Wang, G., Zhang, F., Peng, J., Duan, L., Ji, Y., Marrero-Ortiz, W., Wang, J., Li, J., Wu, C., Cao, C., Wang, Y., Zheng, J., Secrest, J., Li, Y., Wang, Y., Li, H., Li, N., and Zhang, R.: Particle acidity and sulfate production during severe haze events in China cannot be reliably inferred by assuming a mixture of inorganic salts, Atmos. Chem. Phys., 18, 10123-10132, doi:10.5194/acp-18-10123-2018, 2018

Wild, O., Zhu, X., and Prather, M. J.: Fast-j: Accurate simulation of in- and below-cloud photolysis in tropospheric chemical models, J. Atmos. Chem., 37, 245-282, 2000.

Xu, J., Chang, L., Yan, F., and He, J.: Role of climate anomalies on decadal variation in the occurrence of wintertime haze in the Yangtze River Delta, China. Science of the Total Environment,599-600, 918-925, doi:10.1016/j.scitotenv.2017.05.015, 2017.

Yang, Y. R., Liu, X. G., Qu, Y., An, J. L., Jiang, R., Zhang, Y. H., Sun, Y. L., Wu, Z. J., Zhang, F., Xu, W. Q., and Ma, Q. X.: Characteristics and formation mechanism of continuous hazes in China: a case study during the autumn of 2014 in the North China Plain, Atmos. Chem. Phys., 15, 8165-8178, 10.5194/acp-15-8165-2015, 2015.

Zhang, L., Shao, J., Lu, X., Zhao, Y., Hu, Y., Henze, D. K., Liao, H., Gong, S., Zhang, Q.: Sources and Processes Affecting Fine Particulate Matter Pollution over North China: An Adjoint Analysis of the Beijing APEC Period. Environ. Sci. Technol., 50, 8731-8740, 2016.

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., Klimont, Z., Park, I. S., Reddy, S., Fu, J. S., Chen, D., Duan, L., Lei, Y., Wang, L. T., and Yao, Z. L.: Asian emissions in 2006 for the NASA INTEX-B mission, Atmos. Chem. Phys., 9, 5131-5153, 2009.

Zaveri, R. A., and Peters, L. K.: A new lumped structure photochemical mechanism for large-scale applications, J. Geophys. Res.-Atmos., 104, 30387-30415, 1999.

Zaveri, R. A., Easter, R. C., Fast, J. D., and Peters, L. K.: Model for Simulating Aerosol Interactions and Chemistry (MOSAIC), J. Geophys. Res.-Atmos., 113, D13204, doi:10.1029/2007jd008782, 2008.

Zhao, S., Feng, T., Tie, X., Long, X., Li, G., Cao, J., Zhou, W., and An, Z.: Impact of Climate Change on Siberian High and Wintertime Air Pollution in China in Past Two Decades, Earth's Future, 6, 118–133, https://doi.org/10.1002/2017EF000682, 2018.

Zhao, Y., Zhou, Y. D., Qiu, L. P., and Zhang, J.: Quantifying the uncertainties of China's emission inventory for industrial sources: From national to provincial and city scales, Atmos. Environ., 165, 207-221, 10.1016/j.atmosenv.2017.06.045, 2017.

Zheng, B., Zhang, Q., Zhang, Y., He, K. B., Wang, K., Zheng, G. J., Duan, F. K., Ma, Y. L., and Kimoto, T.: Heterogeneous chemistry: a mechanism missing in current models to explain secondary inorganic aerosol formation during the January 2013 haze episode in North China, Atmos. Chem. Phys., 15, 2031-2049, 10.5194/acp-15-2031-2015, 2015.

Zheng, B., Tong, D., Li, M., Liu, F., Hong, C. P., Geng, G. N., Li, H. Y., Li, X., Peng, L. Q., Qi, J., Yan, L., Zhang, Y. X., Zhao, H. Y., Zheng, Y. X., He, K. B., and Zhang, Q.: Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions, Atmos Chem Phys, 18, 14095-14111, 10.5194/acp-18-14095-2018, 2018.

Zhi, G. R., Zhang, Y. Y., Sun, J. Z., Cheng, M. M., Dang, H. Y., Liu, S. J., Yang, J. C., Zhang, Y. Z., Xue, Z.

G., Li, S. Y., and Meng, F.: Village energy survey reveals missing rural raw coal in northern China: Significance in science and policy, Environ. Pollut., 223, 705-712, 10.1016/j.envpol.2017.02.009, 2017.

Zuo, Z. Y., Zhang, R. H., Huang, Y., Xiao, D., and Guo, D.: Extreme cold and warm events over China in wintertime, Int. J. Climatol., 35, 3568-3581, 10.1002/joc.4229, 2015.

Tables and Figures

Table 1. WRF/Chem model configuration.

Table 2. The approach used to distinguish the different impacts of meteorological conditions and emissions by calculating them from different scenarios (taking 2015 and 2016 as an example).

Table 3. Statistics of the observed and model-simulated surface $PM_{2.5}$ for January 2015, 2016 and 2017 in 9 regions (units are μg m^{-3} for BIAS and RMSE).

Table 4. Modeled ambient $PM_{2.5}$ concentration changes for 2016-2015, 2017-2016 and 2017-2015 in 9 regions and the contributions of the meteorology (MET) and emissions (EMIS) calculated according to Table 2. Units: μg m^{-3}.

Table 5. Statistics of the meteorological differences by region for January 2015, 2016 and 2017.

Figure 1. Domain-averaged standard deviations of the background errors (μg kg^{-1}) as a function of the height for each aerosol variable in three bins: (a) Bin-01: 0.039-0.1 μm; (b) Bin-02: 0.1-1.0 μm; (c) Bin-03: 1.0-2.5 μm.

Figure 2. Spatial distribution of primary $PM_{2.5}$ (the sum of BC, OC, sulfate, nitrate and other unspecified $PM_{2.5}$ emissions), SO_2, NO_x and NH_3 emissions (units are μg m^{-2} S^{-1} for $PM_{2.5}$ and mol km^{-2} hr^{-1} for the other three) used in this study.

Figure 3. Observed and modeled monthly average $PM_{2.5}$ concentrations (unit: μg m^{-3}) for January 2015 (left), 2016 (middle) and 2017 (right). Regions defined in red rectangles are as follows: a-NCP (North China Plain), b-NEC (northeastern China), c-EGT (Energy Golden Triangle), d-XJ (Xinjiang), e-SB (Sichuan Basin), f-CC (Central China), g-YRD (Yangtze River Delta), and h-PRD (Pearl River Delta).

Figure 4. Time series of the statistics between the model simulations and observations. Red lines-CONC_DA minus observations, blue lines –NO_DA minus observations. Statistics include the number of data pairs for comparison, the MEAN-mean bias, the STDV- standard deviation, and the RMS-root mean square error. Left-2015, middle-2016, right-2017. (units are μg m^{-3} for MEAN, STDV and RMS).

Figure 5. Spatial distributions of the statistics between the model simulations and observations for January 2015. Top: NO_DA vs. observations, bottom: CONC_DA vs. observations. BIAS-model minus observation, RMSE-root mean square error, CORR-correlation coefficient. (units are μg m^{-3} for BIAS and RMSE).

Figure 6. Observed and modeled ambient $PM_{2.5}$ concentration changes for January 2016-2015 (left), 2017-2016 (middle) and 2017-2015 (right). (a) Observations, (b) assimilated total changes, (c) modeled changes due to meteorological conditions, (d) calculated changes due to emissions. (Units: μg m^{-3})

Figure 7. Modeled meteorological changes for 2016-2015 (left), 2017-2016 (middle) and 2017-2015 (right). (a) PBLH, (b) PSFC, (c) T2, (d) RH2 and (e) 10-m wind speed.

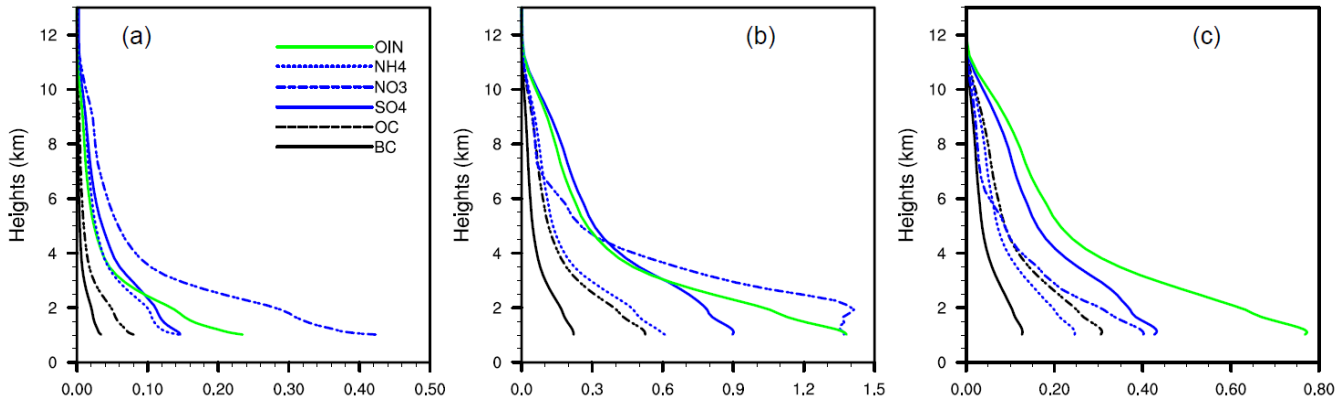


Figure 1. Domain-averaged standard deviations of the background errors (μg kg^{-1}) as a function of the height for each aerosol variable in three bins: (a) Bin-01: 0.039-0.1 μm; (b) Bin-02: 0.1-1.0 μm; (c) Bin-03: 1.0-2.5 μm.

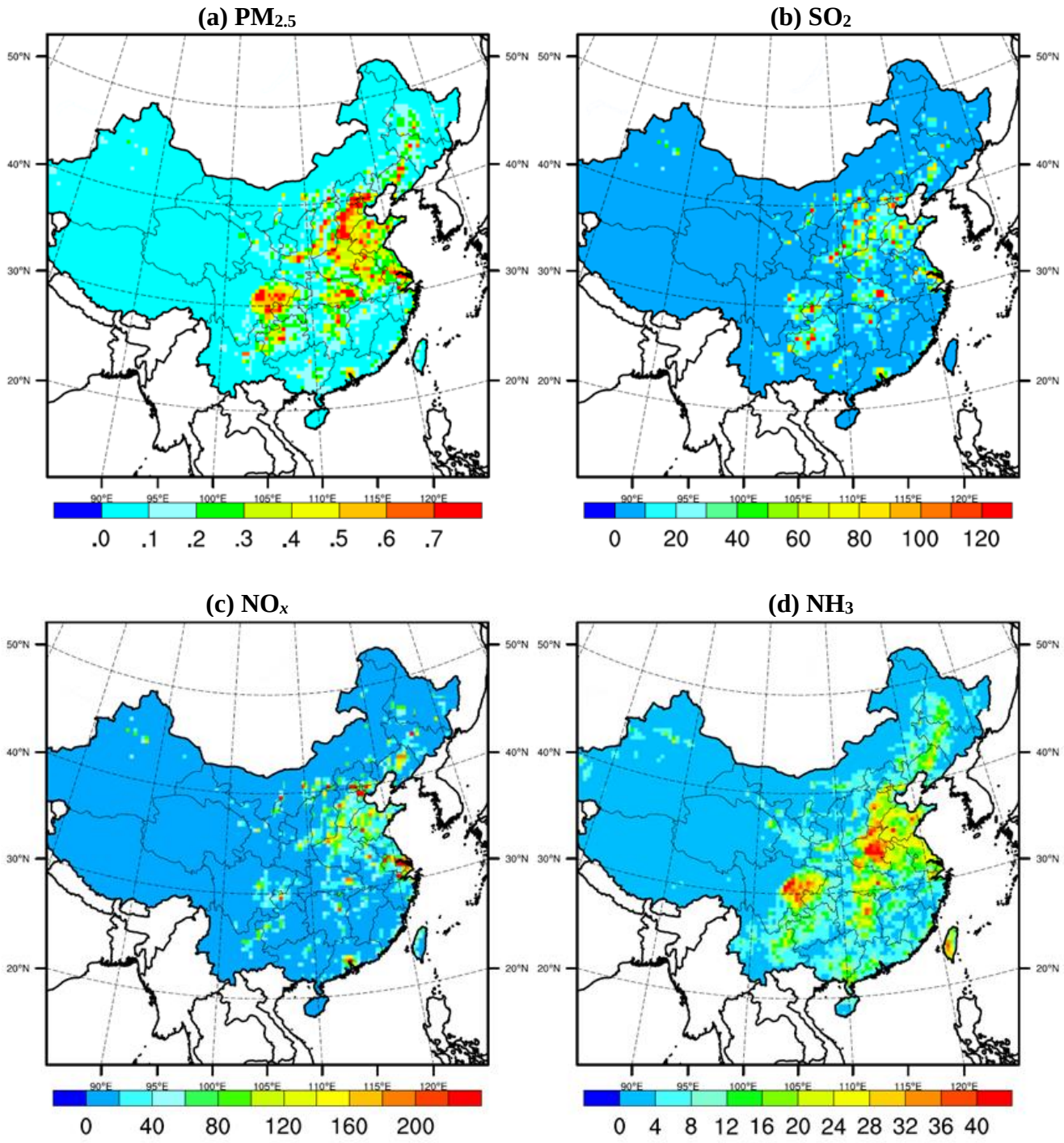


Figure 2. Spatial distribution of primary $PM_{2.5}$ (the sum of BC, OC, sulfate, nitrate and other unspecified $PM_{2.5}$ emissions), SO_2, NO_x and NH_3 emissions (units are μg m^{-2} S^{-1} for $PM_{2.5}$ and mol km^{-2} hr^{-1} for the other three) used in this study.

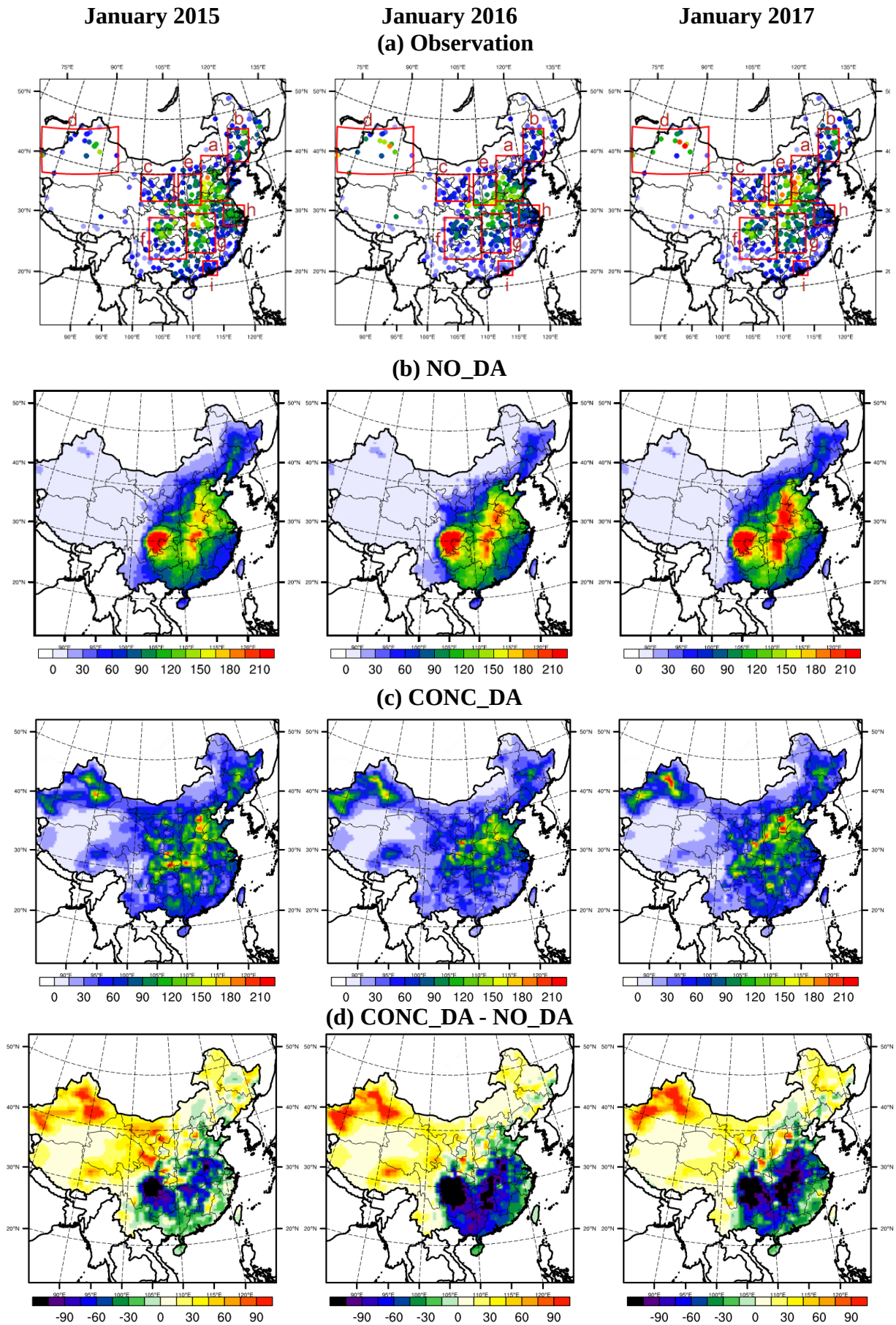


Figure 3. Observed and modeled monthly average $PM_{2.5}$ concentrations (unit: μg m^{-3}) for January 2015 (left), 2016 (middle) and 2017 (right). Regions defined in red rectangles are as follows: a-NCP (North China Plain), b-NEC (northeastern China), c-EGT (Energy Golden Triangle), d-XJ (Xinjiang), e-SB (Sichuan Basin), f-CC (Central China), g-YRD (Yangtze River Delta), and h-PRD (Pearl River Delta).

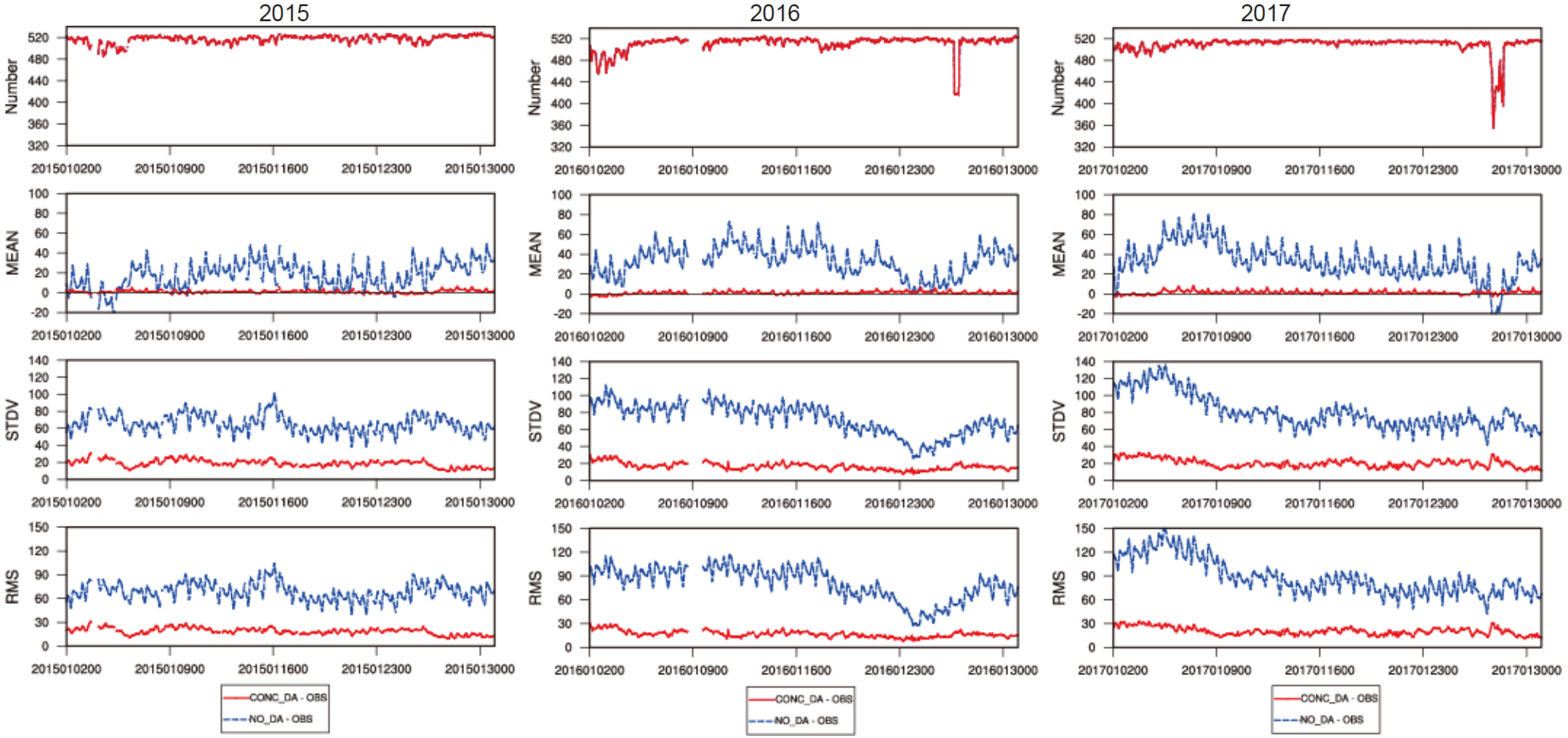


Figure 4. Time series of the statistics between the model simulations and observations. Red lines- CONC_DA minus observations, blue lines –NO_DA minus observations. Statistics include the number of data pairs for comparison, the MEAN-mean bias, the STDV- standard deviation, and the RMS-root mean square error. Left-2015, middle-2016, right-2017. (units are μg m^{-3} for MEAN, STDV and RMS).

(a). January 2015 - NO_DA (top) and CONC_DA (bottom)

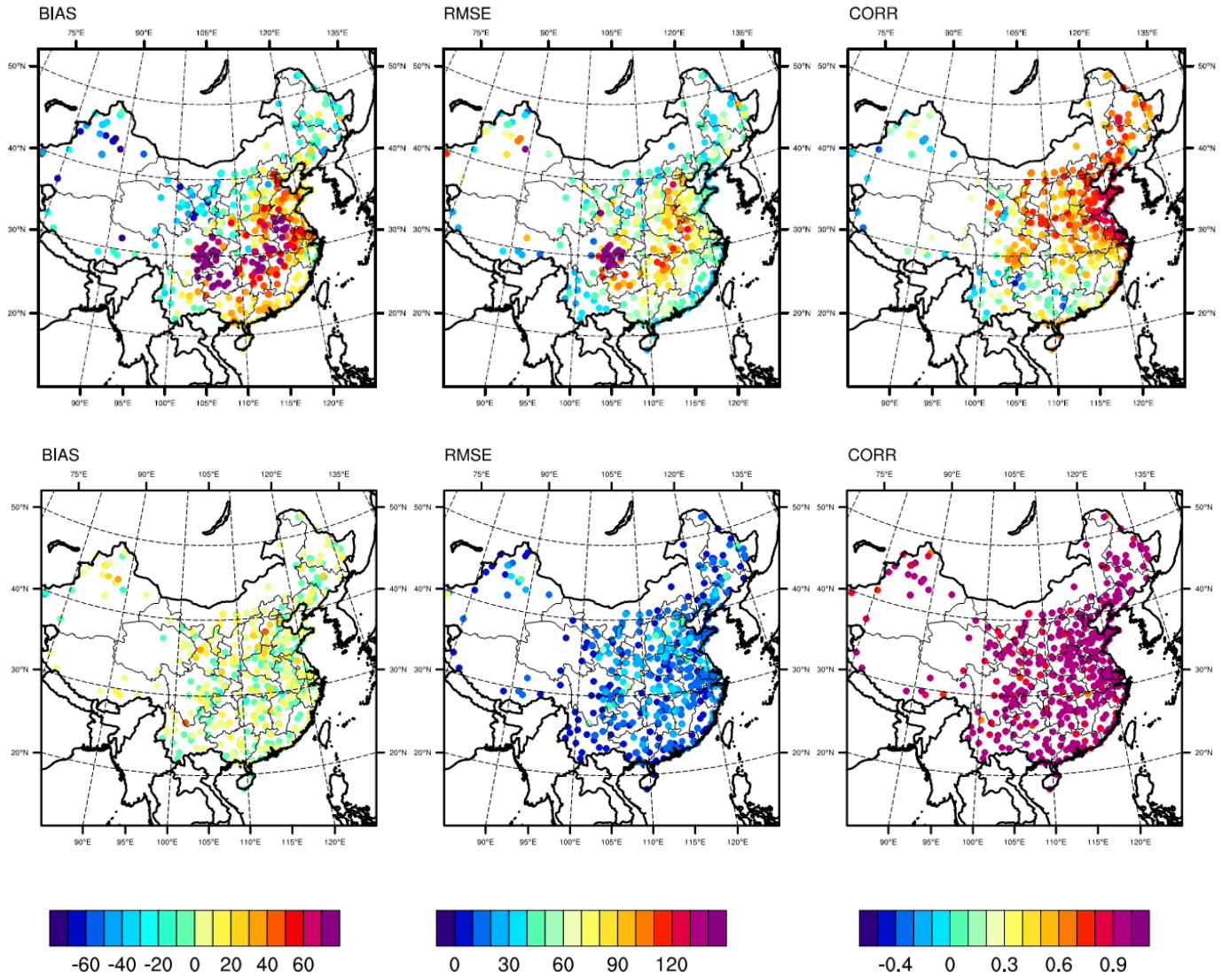


Figure 5. Spatial distributions of the statistics between the model simulations and observations for January 2015. Top: NO_DA vs. observations, bottom: CONC_DA vs. observations. BIAS-model minus observation, RMSE-root mean square error, CORR-correlation coefficient. (units are μg m^{-3} for BIAS and RMSE).

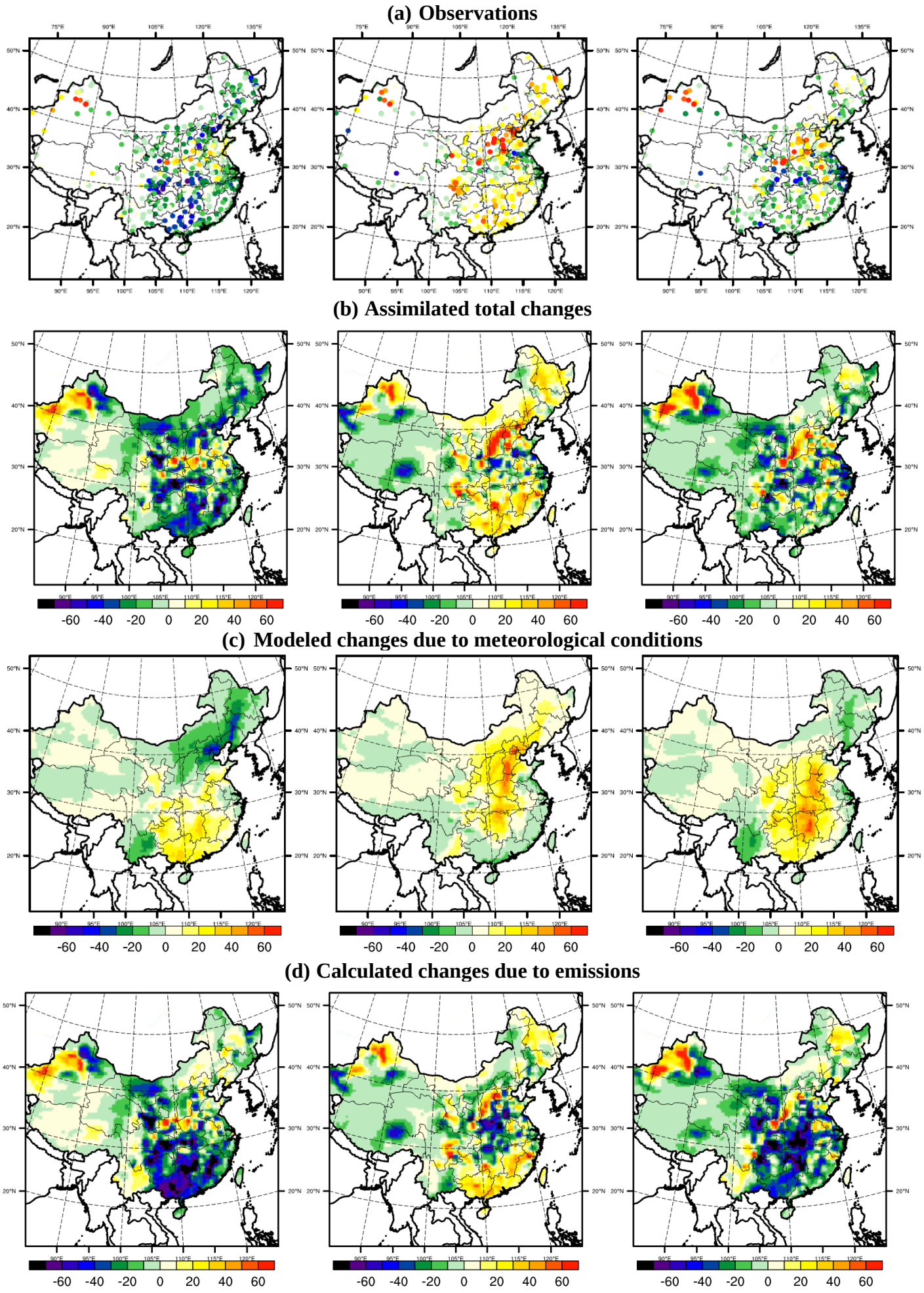

January 2016-2015
January 2017-2016
January 2017-2015
(a) Observations
(b) Assimilated total changes
(c) Modeled changes due to meteorological conditions
(d) Calculated changes due to emissions
-60 -40 -20 0 20 40 60

Figure 6. Observed and modeled ambient $PM_{2.5}$ concentration changes for 2016-2015 (left), 2017-2016 (middle) and 2017-2015 (right). (a) Observations, (b) assimilated total changes, (c) modeled changes due to meteorological conditions, (d) calculated changes due to emissions. (Units: $\mu g\ m^{-3}$)

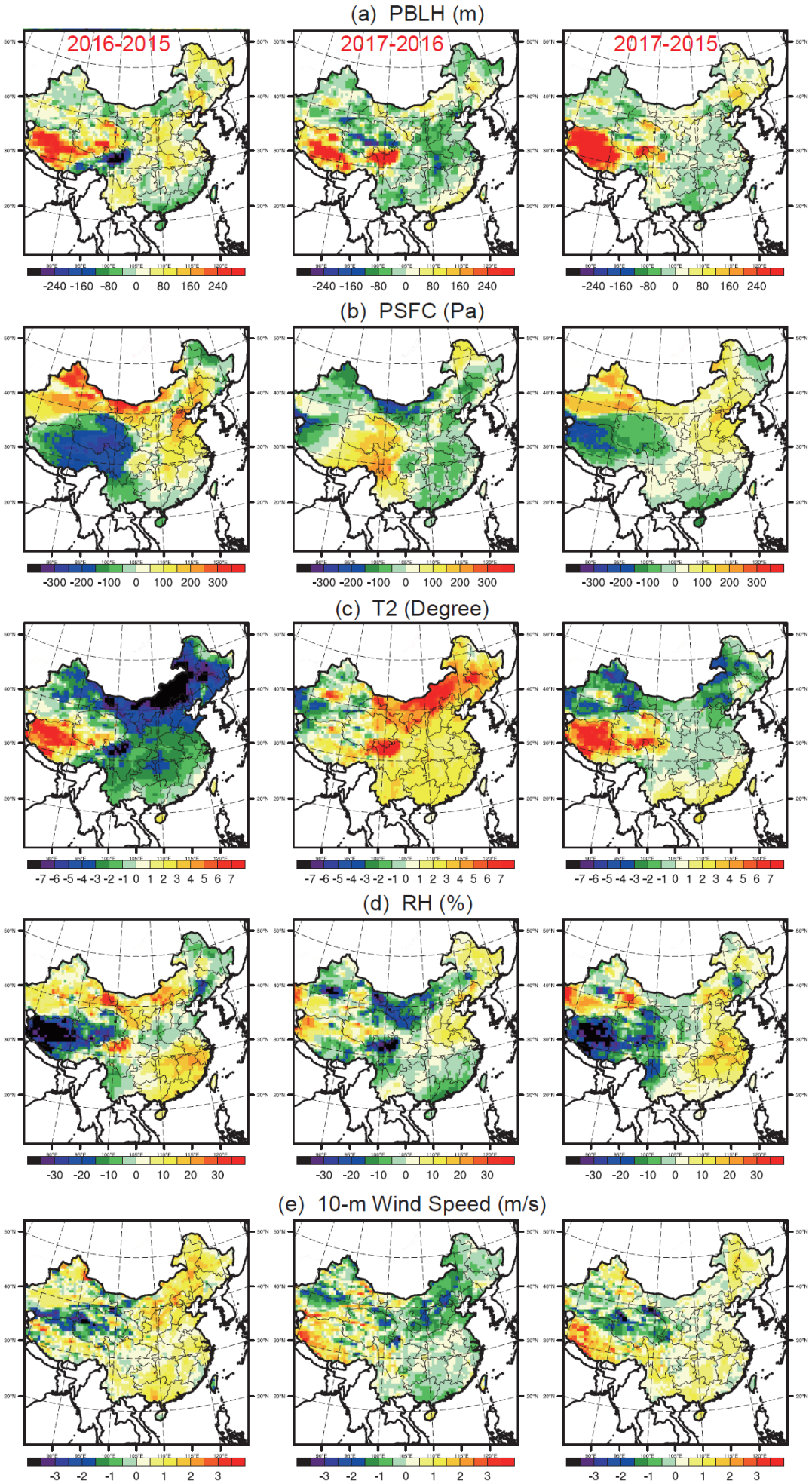
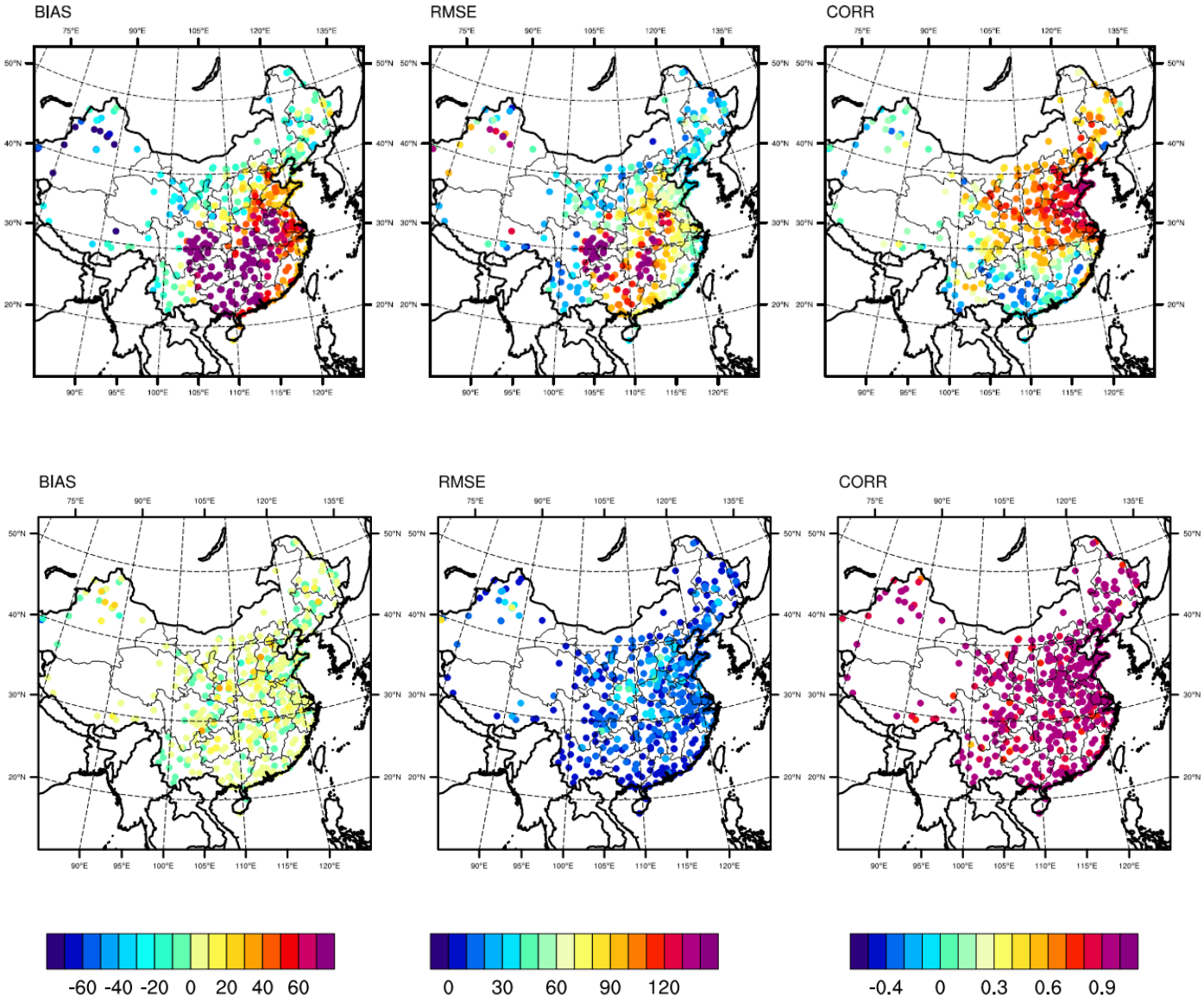
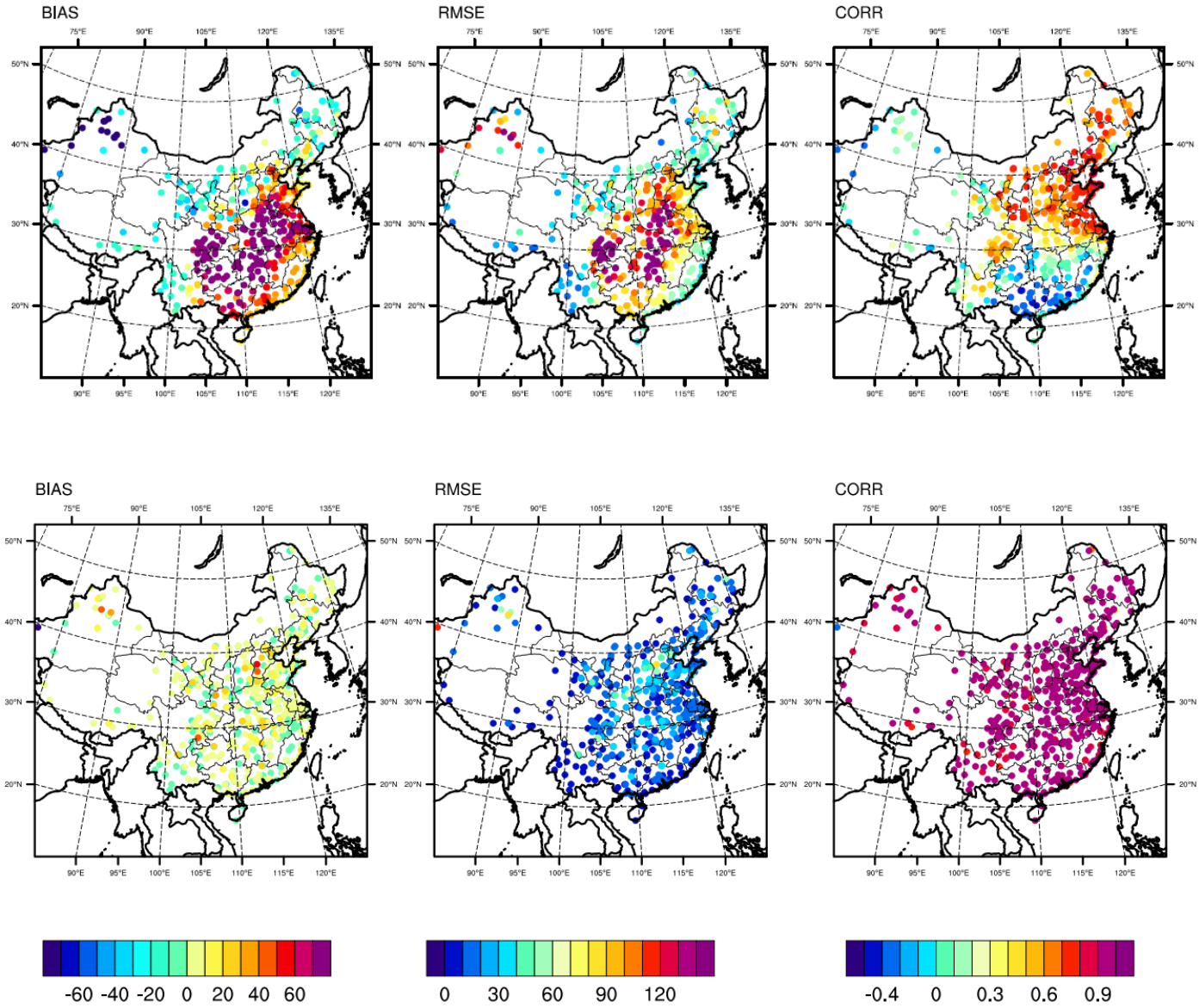


Figure 7. Modeled meteorological changes for 2016-2015 (left), 2017-2016 (middle) and 2017-2015 (right). (a) PBLH, (b) PSFC, (c) T2, (d) RH2 and (e) 10-m wind speed.



Supplemental Figure 1. Same as Figure 5 but for January 2016.

January 2017 - NO_DA (top) and CONC_DA (bottom)



Supplemental Figure 2. Same as Figure 5 but for January 2017.