Reviewer 1:

The paper presents an extension of RSM to ERSM, to manage emission reductions from multiple variables/geographical domains. In my opinion at this stage there are some issues to be modified in the paper, to improve readability and comprehension of the work.

Response: We appreciate the reviewer's comments which help us improve the readability and comprehension of our manuscript. We address the reviewer's comments below. The original comments are in black and our responses are in blue.

INTRODUCTION At page 5053 (line 20) and page 5054 (line 5) it is stated that "...number of scenarios required to build the RSM depends on the variable number via an equation of fourth or higher order...". This depends on the family of models chosen for RSM. I.e., in machine learning theory, other approaches exist with different requirements in terms of number of simulations. Please try to extend this part. Also, in surrogate modelling, various steps could be implemented to reduce the number of required simulations (apart from the family of models chosen). Please try to extend this part, commenting also this issue.

Response: We appreciate the reviewer's valuable comment. The reviewer points out two influential factors: the family of models chosen and the methods to reduce the number of required simulations for a specific family of models.

(1) We agree with the reviewer that the number of scenarios required to build RSM depends on the family of models chosen. When the RSM technique was previously applied in O_3 and $PM_{2.5}$ related studies or policy-making (U.S. Environmental Protection Agency, 2006a, b; Xing et al., 2011; Wang et al., 2011), the relationships between air pollutant concentrations and precursor emissions were established using the Maximum Likelihood Estimation - Empirical Best Linear Unbiased Predictors (MLE-EBLUPs) developed by Santner et al. (2003). When this specific statistical technique was applied, the number of scenarios required to build the RSM depends on the variable number via an equation of fourth or higher order. We have clarified this point in the revised manuscript (from Page 3, Line 32 to Page 4, Line

13 in the revised manuscript). The revised text is also shown in the paragraph after next.

(2) The number of required simulations to build the response surface using MLE-EBLUPs is affected by the sampling method to generate the emission control scenarios, the correlation family used in the prediction process, and so on. We used the Latin Hypercube Sample (LHS) method and the power exponential correlation family based on a thorough comparison conducted by Santner et al. (2003). With these configurations used, the number of scenarios required to build the RSM depends on the variable number via an equation of fourth or higher order. We also tried to distribute more samples in areas with relatively larger errors (in particular marginal area); it could slightly reduce the number of scenarios required but does not change its magnitude (unpublished results of our study).

In the revised manuscript, we revised the related description as follows:

The Response Surface Modeling (RSM) technique (denoted by "conventional RSM" technique in the following text to distinguish from the ERSM technique developed in this study), has been developed by using advanced statistical techniques to characterize the relationship between model outputs and inputs in a highly economical manner. The number of scenarios required to build RSM depends on the family of models chosen. Recently, the conventional RSM technique has been applied to O_3 and $PM_{2.5}$ related studies or policy-making in the United States (U.S. Environmental Protection Agency, 2006a, b) and China (Xing et al., 2011; Wang et al., 2011). In those applications, the relationships between air pollutant concentrations and precursor emissions were established using the Maximum Likelihood Estimation - Empirical Best Linear Unbiased Predictors (MLE-EBLUPs) developed by Santner et al. (2003). Using this group of model, the number of model scenarios required to build the RSM depends on the variable number via an equation of fourth or higher order, even if the preferable sampling method and model configurations proposed by previous studies (Santner et al., 2003) are used. (from Page 3, Line 32 to Page 4, Line 13 in the revised manuscript)

References:

Santner, T. J., Williams, B. J., and Notz, W.: The Design and Analysis of Computer Experiments, Springer Verlag, New York, U.S., 283 pp., 2003.

U.S. Environmental Protection Agency: Technical support document for the proposed PM NAAQS rule: Response Surface Modeling, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, U.S., 48, 2006a.

U.S. Environmental Protection Agency: Technical support document for the proposed mobile source air toxics rule: ozone modeling, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, U.S., 49, 2006b.

Wang, S. X., Xing, J., Jang, C. R., Zhu, Y., Fu, J. S., and Hao, J. M.: Impact assessment of ammonia emissions on inorganic aerosols in east China using response surface modeling technique, Environ. Sci. Technol., 45, 9293-9300, DOI 10.1021/Es2022347, 2011.

Xing, J., Wang, S. X., Jang, C., Zhu, Y., and Hao, J. M.: Nonlinear response of ozone to precursor emission changes in China: a modeling study using response surface methodology, Atmos. Chem. Phys., 11, 5027-5044, DOI 10.5194/acp-11-5027-2011, 2011.

At page 5054 (end of Introduction) you should clearly state the advantages of ERSM in comparison to RSM...I think pros mainly refer to the possibility to use (in comparison to RSM) an increased number of variables/geographical areas, but it is not clearly stated.

Response: We have added the following sentence in the revised manuscript:

Compared with the conventional RSM technique, ERSM is applicable for an increased number of variables and geographical regions. (Page 4, Line 23-25 in the revised manuscript)

METHODOLOGY In general, I find this part of the paper (important, because it presents the ERSM approach) quite complex/obscure. I suggest the authors to restructure this section, because now it is quite complex to follow how the model is identified, and how the various equations interact. I.e. a diagram/flow chart of the required steps/equations to be used could be quite beneficial.

Response: We appreciate the reviewer's valuable comment. In the revised manuscript, we have added an intuitional flowchart (Figure 1 in the revised manuscript), as shown below. We believe this flowchart has made it much easier for the readers to follow how the model is developed.

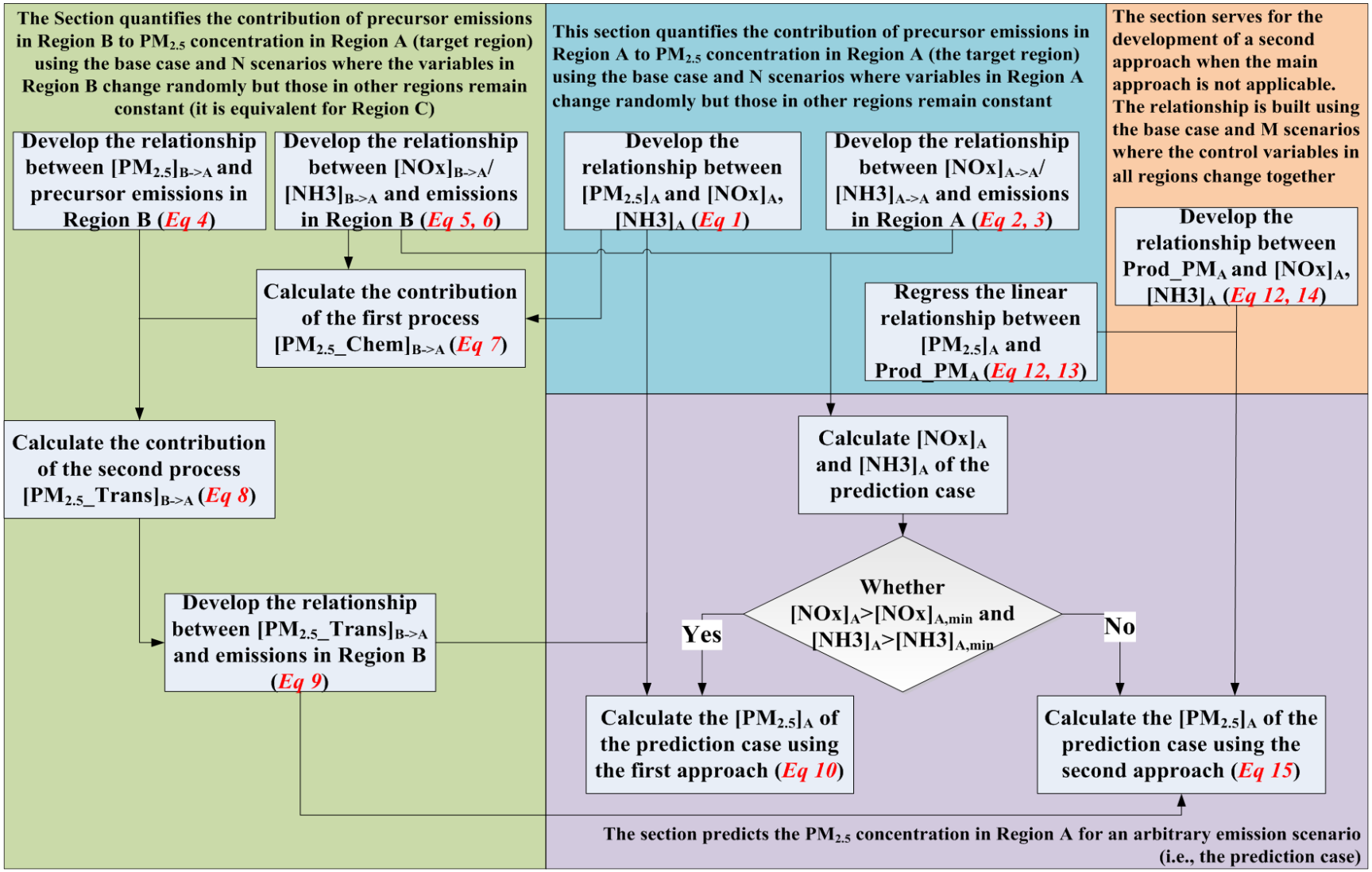

Figure 1. A flowchart illustrating the ERSM technique using the simplified case described in Sect. 2.1. Different background colors represent the procedures conducted using different groups of emission scenarios, as indicated on the top/bottom of the colored areas.

At page 5056 (line 10) the number of simulations required for RM is presented (30 and 50 scenarios) but no explanation for this is provided. There is indeed a citation, but I would suggest authors to better integrate/explain this part.

Response: We appreciate the reviewer's comments. We have explained this part in the revised manuscript, which is also shown below.

The scenario numbers N and M are determined in order that they are sufficient to accurately construct the relationship between the response variable and randomly changing control variables. Specifically, we gradually increase the scenario number and build the response surface repeatedly until the prediction performance is good enough based on the results of "out of sample" validation and 2-D isopleths validation (Xing et al., 2011; Wang et al., 2011). Based on our previous studies (Xing et al., 2011; Wang et al., 2011), the response surface for 2 and 3 variables could be built with good prediction performance (mean normalized error < 1%; correlation coefficient > 0.99) using 30 and 50 scenarios, respectively; therefore, for this simplified case, N=50, and M=30. (Page 6, Line 16-25 of the revised manuscript)

References:

Wang, S. X., Xing, J., Jang, C. R., Zhu, Y., Fu, J. S., and Hao, J. M.: Impact assessment of ammonia emissions on inorganic aerosols in east China using response surface modeling technique, Environ. Sci. Technol., 45, 9293-9300, DOI 10.1021/Es2022347, 2011.

Xing, J., Wang, S. X., Jang, C., Zhu, Y., and Hao, J. M.: Nonlinear response of ozone to precursor emission changes in China: a modeling study using response surface methodology, Atmos. Chem. Phys., 11, 5027-5044, DOI 10.5194/acp-11-5027-2011, 2011.

Some assumptions are presented (pag 5058, line 5; pag 5059, line 1-5) in the paper, to justify some of the choices done by the authors in the equation implementation; I would suggest to better explain why these assumptions are taken, and which are their implications.

Response: We appreciate the reviewer's valuable comment. In the revised manuscript, we explained the reason to take these assumptions and their implications in detail. The revised

texts are also shown below.

In order to quantify the contribution of the first process, we firstly use Eq. (5) and Eq. (6) to quantify the effect of the transport of gaseous precursors from Region B to Region A on the precursor concentrations in Region A. How much does the change of precursor concentrations in Region A enhance the chemical formation of secondary $PM_{2.5}$ in Region A? To answer this question, we introduce a straightforward assumption that the changes of $PM_{2.5}$ concentration owing to changes of precursor concentrations in the same region (described by Eq. (1)) are solely attributable to changes of local chemical formation. Strictly speaking, the changes of precursor concentration in Region A might affect the precursor concentrations/$PM_{2.5}$ concentrations in other regions, which might in turn affect the $PM_{2.5}$ concentrations in Region A; but this "indirect" pathway is thought to be negligible in this study. (from Page 7, Line 31 to Page 8, Line 10 in the revised manuscript)

Strictly speaking, $[PM_{2.5}\_Trans]_{B\to A}$ and $[PM_{2.5}\_Trans]_{C\to A}$ could interact with each other. In other words, the changes of precursor emissions in Region C might affect the formation of secondary $PM_{2.5}$ in Region B, which further affects the transport of secondary $PM_{2.5}$ from Region B to Region A. In this study, we assume that $[PM_{2.5}\_Trans]_{B\to A}$ depends only on the precursor emissions in Region B, and is independent of precursor emissions in other regions (as indicated by in Eq. (9) and Eq. (10)). That is, the interaction between $[PM_{2.5}\_Trans]_{B\to A}$ and $[PM_{2.5}\_Trans]_{C\to A}$ is neglected. (Page 9, Line 6-12 in the revised manuscript)

At pag 5061 (line 27) authors say that 663 scenarios are required...again, please specify how this number is computed. I think this number depends on the number of variables/geographical areas chosen. But if this is the case, it means that this ERSM approach is really in my opinion too demanding in terms of CTMs simulations (600 simulations is really a huge number, in my opinion). Also, the set of CTM simulation is quite strictly dependent on some assumptions (choice of geographical areas, choice of variables) that could be quite uncertain starting a new study...if these assumptions changes due some further

analysis, one should recompute the CTM hundreds of simulations? Please clarify this point.

Response: We have explained how the scenario number 663 was determined in the revised manuscript. The revised text is shown as follows.

We generated 663 scenarios (see Table 1) to build the response surface, following the method to create emission scenarios for the ERSM technique (the 5th paragraph of Sect. 2.1). In detail, the scenarios include (1) 1 CMAQ base case; (2) N=150 scenarios generated by applying LHS method for the control variables of gaseous precursors in Shanghai, 150 scenarios generated in the same way for Jiangsu, 150 scenarios for Zhejiang, and 150 scenarios for Others; (3) M=50 scenarios generated by applying LHS method for the total emissions of NO_X, SO_2, and NH_3 in all regions; and (4) 12 scenarios where one of the control variables of primary $PM_{2.5}$ emissions is set to 0.25 for each scenario. Here the number N=150 and M=50 are decided according to the numerical experiments conducted in our previous studies (Xing et al., 2011; Wang et al., 2011), which showed that the response surface for 6 and 3 variables could be built with good prediction performance (mean normalized error < 1%; correlation coefficient > 0.99) using 150 and 50 scenarios, respectively. (Page 12, Line 8-20 in the revised manuscript)

We agree with the reviewer that 663 scenarios are not a small number. However, theoretically over 10^5 scenarios would be required if the conventional RSM technique was used to solve same issue of this study. It can be seen that the ERSM has significantly reduced the number of scenarios compared with the conventional RSM technique. In addition, less emission scenarios are required if we allow larger prediction errors. As described above, 150 scenarios are required to build the response surface for 6 variables with the mean normalized error less than 1%; however, only about 60 scenarios are required if we allow a mean normalized error less than 5%, reducing more than half of the total scenario number (Xing et al., 2011). In future, more efforts should be made to further reduce the scenario number required while assuring the accuracy of the response surfaces. We have described this limitation in the last section of the revised manuscript. (Page 19, Line 5-9 in the revised manuscript)

The scenario number does depend on the number of geographical areas/variables chosen. We

have clearly described what and how much emission scenarios are required (Page 6, Line 13-27 in the revised manuscript), so that a user could readily calculate the required scenario number for a specific experimental design (e.g., number of geographical regions/variables) before starting CTM simulations. The user could then balance the level of detail of the control variables and the amount of calculation needed. In this way we somewhat reduced the uncertainty in starting a new study.

It is not necessary to recompute lots of CTM simulations if we make minor revision on the experimental design. For example, if one more geographical area is added, we just need to (1) add a parallel group of emission scenarios where the control variables of the added geographical area change while those of the other regions remain the base-case levels, and (2) recompute the emission scenarios where the control variables of all regions change simultaneously. Another example, if the selected emission sectors in a specific geographical area are changed, we just need to recompute the group of emission scenarios where the control variables of this geographical area change while those of the other regions remain the base-case levels. However, if the experimental design is significantly changed (e.g., change of selected pollutants, or change of selected emission sectors in all regions), most of the CTM simulations need to be recomputed. The users need to carefully design the experiment before performing the CTM simulations. We have described this limitation in the last section of the revised manuscript. (Page 19, Line 9-22 in the revised manuscript)

References:

Wang, S. X., Xing, J., Jang, C. R., Zhu, Y., Fu, J. S., and Hao, J. M.: Impact assessment of ammonia emissions on inorganic aerosols in east China using response surface modeling technique, Environ. Sci. Technol., 45, 9293-9300, DOI 10.1021/Es2022347, 2011.

Xing, J., Wang, S. X., Jang, C., Zhu, Y., and Hao, J. M.: Nonlinear response of ozone to precursor emission changes in China: a modeling study using response surface methodology, Atmos. Chem. Phys., 11, 5027-5044, DOI 10.5194/acp-11-5027-2011, 2011.

RESULTS AND DISCUSSION It seems to me the paper deals with analysis on January and

August periods...please explain how one (if possible) could extend this analysis to the full year (i.e. to analyses possible structural emission reductions).

Response: The most rigorous way to extend this analysis to a full year is to finish the CMAQ simulations for a full year and build the response surfaces following the same procedure. Alternatively, the relationship for a full year can be roughly estimated using the average values of January and August. Another approach is to finish the simulations for an additional month in Spring and Autumn, respectively, and represent the situation of a full year with the average values of the four typical months. We have added this explanation in the Methodology section of the revised manuscript. (Page 11, Line 19-25 in the revised manuscript)

CONCLUSIONS I would discuss (in the conclusions, or somewhere in the paper) issues about - how to deal with meteorology variability in this approach - how uncertain is the model identification - as already said, the limitations of this approach (need of hundreds of simulations, that depend on assumptions on geographical areas/variables, etc...).

Response: We appreciate the reviewer's valuable comment. We have discussed these issues in the last section of the revised manuscript. The amended texts are shown as follows.

The ERSM technique still has several limitations. Firstly, the technique currently does not consider the variability of meteorological conditions. Secondly, although the ERSM technique represents an essential improvement compared with the conventional RSM technique, it usually needs over 500 emission scenarios for a medium-size problem. Future studies should be done to further reduce the number of scenarios required while assuring the accuracy of the response surfaces. Thirdly, the emission scenarios required to build the response surface depends strictly on the experimental design (e.g., selection of geographical regions and control variables). It is not necessary to recompute lots of CTM simulations if we make minor revision on the experimental design. For example, if one more geographical area is added, we

just need to (1) add a parallel group of emission scenarios where the control variables of the added geographical area change while those of the other regions remain the base-case levels, and (2) recompute the emission scenarios where the control variables of all regions change simultaneously. Another example, if the selected emission sectors in a specific geographical area are changed, we just need to recompute the group of emission scenarios where the control variables of this geographical area change while those of the other regions remain the base-case levels. However, if the experimental design is significantly changed (e.g., change of selected pollutants, or change of selected emission sectors in all regions), most of the CTM simulations need to be recomputed. The users need to carefully design the experiment before performing the CTM simulations. (Page 19, Line 4-22 in the revised manuscript)

FIGURE I would keep, as an example, only Figure 3 (Figure 4-5 in my opinion are not needed).

Response: We thank the reviewer for this comment. In the revised manuscript we keep only Figure 3 in the main text (Figure 4 in the revised manuscript) and move Figure 4 and Figure 5 to the Supporting Information.

Reviewer 2:

1 General comments

This is an interesting paper which develops an advanced modelling technique for particulate matter concentration responses to changes in emissions of primary particles and precursor gases. The authors apply the model to analyse the sensitivity of PM concentrations in Chinese cities to the emissions from different economic sectors and spatial origin. Given the recurring episodes of high pollution in Chinese cities, this is a very important and timely topic and fits well to this journal. Considering the still rather high number of CTM simulations needed I am not fully convinced that this is an easily applicable technique for source response modelling in general, but the example provided in the paper demonstrates feasibility of the approach for a limited set of source regions (4) and sectors.

The article is generally well written, although the methodology section is in parts rather difficult to follow. I recommend the paper to be published after addressing the comments below, mostly regarding clarifications in the methodology.

Response: We thank the reviewer for supporting the publication of our manuscript. We also appreciate his/her comments which help us improve the quality of our manuscript. We address the reviewer's comments below. The original comments are in black and our responses are in blue.

2 Detailed comments

Sect. 2.1, general question: What is the time scale for the ERSM development? Do the authors aim to develop annual mean coefficients, monthly coefficients, summer/winter coefficients? Would full-year coefficients be feasible?

Response: The ERSM technique is applicable for various time scales, ranging from a single day to several years. As a case study in the Yangtze River Delta, we developed the response surfaces for monthly mean $PM_{2.5}$ concentrations in January and August, representing winter and summer, respectively. The analysis could be extended to a full year as long as we finish

the CMAQ simulations for a full year and build the response surfaces following the same procedure. We have added the explanation accordingly in the revised manuscript. (Page 11, Line 17-21 in the revised manuscript)

p. 5055 l. 7ff: “The ERSM technique first quantifies the relationship...” - What is missing in the general description of the methodology here is a simple description of how this functional relationship is quantified. This should be explained briefly to readers who are not proficient with the RSM technique and the MPerK program.

Response: We thank the reviewer for this comment. We have added description of how the relationship is quantified in the revised manuscript. The amended texts are shown as follows.

The ERSM technique is developed starting from the conventional RSM technique; the latter characterizes the relationships between a response variable (e.g., $PM_{2.5}$ concentration) and a set of control variables (i.e., emissions of particular precursors from particular sources) following the procedures described in our previous paper (Xing et al., 2011). First, a number of emission control scenarios are generated with the Latin Hypercube Sample (LHS) method (Iman et al., 1980), a widely-used sampling method which ensures that the ensemble of random samples is representative of actual variability. Then the $PM_{2.5}$ concentration for each emission scenario is calculated with a regional CTM, and finally the RSM prediction system is developed using a MPerK (MATLAB Parametric Empirical Kriging) program (Santner et al., 2003) based on MLE-EBLUPs. The robustness of the conventional RSM technique has been validated through leave-one-out cross validation, out of sample validation and 2-D isopleths validation, as documented in our previous papers (Xing et al., 2011; Wang et al., 2011). (Page 5, Line 3-15 in the revised manuscript)

References:

Iman, R. L., Davenport, J. M., and Zeigler, D. K.: Latin Hypercube Sampling (Program User's Guide), Sandia National Laboratories, Albuquerque, NM, U.S.Technical Report SAND79-1473, 78, 1980.

Santner, T. J., Williams, B. J., and Notz, W.: The Design and Analysis of Computer Experiments, Springer Verlag, New York, U.S., 283 pp., 2003.
Wang, S. X., Xing, J., Jang, C. R., Zhu, Y., Fu, J. S., and Hao, J. M.: Impact assessment of ammonia emissions on inorganic aerosols in east China using response surface modeling technique, Environ. Sci. Technol., 45, 9293-9300, DOI 10.1021/Es2022347, 2011.
Xing, J., Wang, S. X., Jang, C., Zhu, Y., and Hao, J. M.: Nonlinear response of ozone to precursor emission changes in China: a modeling study using response surface methodology, Atmos. Chem. Phys., 11, 5027-5044, DOI 10.5194/acp-11-5027-2011, 2011.

p. 5055 l. 12ff: How about formation of secondary PM in the source region followed by transport to the target region? Is this process missed or is just the formulation unclear at this point? If it is missed, do the authors have an estimate of its contribution? How about interactions of precursor emissions from different source regions?

Response: This process is not missed but is not clearly described in our original manuscript. This process is calculated using Eq. (8). To make it clear, we have revised the original descriptions as follows:

In order to quantify the interaction among regions, we make a key assumption that the emissions of gaseous precursors in the source region affect $PM_{2.5}$ concentrations in the target region through two major processes: (1) the inter-regional transport of gaseous precursors enhancing the chemical formation of secondary $PM_{2.5}$ in the target region; (2) the formation of secondary $PM_{2.5}$ in the source region followed by transport to the target region. (Page 5, Line 20-25 of the revised manuscript)

The texts that describe how we quantify the contribution of the second process are as follows:

The contribution of the second process to $PM_{2.5}$ concentration in Region A (denoted by $[PM_{2.5}\_Trans]_{B\rightarrow A}$ defined below) is then calculated by extracting the contribution of the first process (Eq. (7)) from the total (Eq. (4)), as expressed by Eq. (8).

$$[PM_{2.5}\_Trans]_{B\rightarrow A}=[PM_{2.5}]_{B\rightarrow A}-[PM_{2.5}\_Chem]_{B\rightarrow A} \quad (8)$$

where $[PM_{2.5}\_Trans]_{B\rightarrow A}$ is the change of $PM_{2.5}$ concentration in Region A affected by the changes of precursor emissions in Region B through the transport of secondary $PM_{2.5}$ (the second process). (Page 8, Line 15-21 of the revised manuscript)

For your last question, the interaction of precursor emissions from different source regions could affect the $PM_{2.5}$ concentrations in the target region in two different ways: (1) by affecting the inter-regional transport of gaseous precursors from the source regions to the target region; (2) by affecting the inter-regional transport of secondary $PM_{2.5}$ from the source regions to the target region. The first effect is implicitly considered in Eq. (10), which sums up the changes of precursor concentrations in the target region induced by the changes of precursor emissions in every source region. The second effect is neglected, which is described in Page 9, Line 6-12 of the revised manuscript.

p. 5058 l. 4ff: Similar to above: Is something missed due to this assumption? In fact, as it is formulated it reads rather straightforward and I don't quite understand why this assumption is needed. How should PM2.5 in A be influenced by precursor gas concentration changes in A other than through local chemical formation?

Response: We thank the reviewer for this comment. The changes of precursor concentration in Region A might affect the precursor concentrations/$PM_{2.5}$ concentrations in other regions, which might in turn affect the $PM_{2.5}$ concentrations in Region A. This "indirect" pathway is neglected with the above-mentioned assumption.

We have explained this assumption in detail in the revised manuscript. The revised texts are also given below.

In order to quantify the contribution of the first process, we firstly use Eq. (5) and Eq. (6) to quantify the effect of the transport of gaseous precursors from Region B to Region A on the precursor concentrations in Region A. How much does the change of precursor concentrations in Region A enhance the chemical formation of secondary $PM_{2.5}$ in Region A? To answer this question, we introduce a straightforward assumption that the changes of $PM_{2.5}$ concentration owing to changes of precursor concentrations in the same region (described by Eq. (1)) are solely attributable to changes of local chemical formation. Strictly speaking, the changes of

precursor concentration in Region A might affect the precursor concentrations/$PM_{2.5}$ concentrations in other regions, which might in turn affect the $PM_{2.5}$ concentrations in Region A; but this "indirect" pathway is thought to be negligible in this study. (from Page 7, Line 31 to Page 8, Line 10 in the revised manuscript)

p. 5058, l. 10ff, Eqns 8 and 9: This paragraph is a bit unclear to me. If $[PM_{2.5}\_Trans]_{B\rightarrow A}$ is calculated through Eq. 8, how is Eq. 9 used? Or is Eq. 8 merely the definition and Eq. 9 is in fact used for calculating the transport contribution from B to A? This should be clarified.

Response: We appreciate the reviewer's comment. When the response surface is built, $[PM_{2.5}\_Trans]_{B\rightarrow A}$ is calculated using Eq. 8. But we also need to know the relationship between $[PM_{2.5}\_Trans]_{B\rightarrow A}$ and the precursor emissions in Region B, which is quantified using conventional RSM technique, and is described as Eq. 9.

When we predict the $PM_{2.5}$ concentration for an arbitrary scenario using Eq. 10 or Eq. 15, $[PM_{2.5}\_Trans]_{B\rightarrow A}$ is calculated using Eq. 9. In this case, it cannot be calculated using Eq. 8 because Eq. 8 holds only if the emissions in the regions other than Region B remain at the base-case levels.

In the revised manuscript, this issue has been clarified. (from Page 8, Line 15 to Page 9, Line 5 in the revised manuscript)

p 5059, l. 4ff and Eq. 11: It would be helpful if $[NOx]_{A,\,min}$ were defined explicitly. Is it certain that $[NOx]_{A\rightarrow A,\,min} = RSM_{A\rightarrow A}^{NOx}(0,\,0,\,0)$? Is the case (0,0,0) i.e. all precursor emissions equal 0 covered in any case? From a random draw it could even be higher?

Response: We have defined $[NOx]_{A,\,min}$ explicitly in the revised manuscript: $[NOx]_{A,\,min}$ is defined as the minimum NO_X concentration in Region A when the emissions from Region A change arbitrarily and those in other regions remain the base-case level. (Page 9, Line 20-22 in the revised manuscript)

We concluded $[NOx]_{A\rightarrow A,\,min} = RSM_{A\rightarrow A}^{NOx}(0,\,0,\,0)$ based on a series of numerical experiments.

In detail, we have constructed the relationship between $[NOx]_{A\to A}$ and the precursor emissions in Region A in our manuscript, as described by Eq. (2).

$$[NOx]_{A\to A}=RSM^{NOx}_{A\to A}\left(Emis\_NOx\_1_A, Emis\_NOx\_2_A, Emis\_NH3_A\right) \quad (2)$$

Similar to Fig. 4 of the revised manuscript, we used Eq. (2) to plot the 2-D isopleths of $[NOx]_{A\to A}$ in response to any two of the control variables. From those isopleths we concluded that $[NOx]_{A\to A}$ is minimal when all precursor emissions equal zero.

p. 5060 l. 15 ff: δ as defined in l. 16 is not an interval. I assume the actual transition interval is $([NOx]_{A,\,min},\ [NOx]_{A,\,min}+\delta_{NOx})$ for NOx and equivalent for NH3? If so, this should be clarified in the text.

Response: Yes, this is what we meant. We have revised the texts as follows:

To assure the consistency between Eq. (10) and Eq. (15), we introduce "transition intervals" of $([NOx]_{A,\,min}, [NOx]_{A,\,min}+\delta_{NOx})$ and $([NH3]_{A,\,min}, [NH3]_{A,\,min}+\delta_{NH3})$, where $\delta_{NOx}=0.1*[NOx]_{A0}$ and $\delta_{NH3}=0.1*[NH3]_{A0}$. Eq. (10) is applied for $[NOx]_A \geq [NOx]_{A,\,min}+\delta_{NOx}$ and $[NH3]_A \geq [NH3]_{A,\,min}+\delta_{NH3}$, and we linearly interpolate between Eq. (10) and Eq. (15) for the transitional range. (from Page 10, Line 27 to Page 11, Line 2 in the revised manuscript)

I do not completely understand why the physical transport versus chemical production diagnostics module is used in the second approach outside the minimal precursor concentrations, but not within the first approach. For example, could it not be useful for distinguishing chemical and transport contributions as in Eq. 7 and 8? This should be better motivated by the authors.

Given that two different approaches are used here, the authors should comment on how smooth the transition is between the two regimes considered, i.e. how large the deviations are in the overlap interval.

Response: We thank the reviewer for this comment. Using the first approach, we could

distinguish the contributions of chemical formation and physical transport without this diagnostic module (see Eq. 7 and Eq. 8). If this module was used, we would need to develop the relationship between the chemically formed $PM_{2.5}$ and the $PM_{2.5}$ concentration, which was an extra step compared with the first approach and added to the complexity. In the second approach, we are not able to distinguish the chemical and transport contributions without the diagnostic module; this explains why that module is only used in the second approach.

Based on the case study in the Yangtze River Delta, the discrepancy between the two approaches is 1-8% in the overlap interval.

We have explained these two issues accordingly in the revised manuscript. (from Page 10, Line 20 to Page 11, Line 2 of the revised manuscript)

p. 5062 l. 15: “the cases where all control variables are controlled stringently” – If I have not overlooked something, these cases are mentioned here for the first time. They should be mentioned in the text describing the scenario generation in the last paragraph of Sect. 2.2.

Response: We agree with the reviewer’s comment and have added a brief description of the cases used for out-of-sample validation in Sect. 2.2. The added text is also given as follows.

Finally, we generated 40 independent scenarios for out-of-sample validation, as described in detail in Sect. 3.1. (Page 12, Line 20-21 of the revised manuscript)

The detailed description in Sect. 3.1 is also revised more clearly as follows:

These 40 out-of-sample scenarios include 32 cases (case 1-32) where the control variables of gaseous precursors change but those of primary $PM_{2.5}$ stay the same as the base case, 4 cases (case 33-36) the other way around, and 4 cases (case 37-40) where control variables of gaseous precursors and primary $PM_{2.5}$ change simultaneously. Most cases are generated randomly with the LHS method (case 4-6, 10-12, 16-18, 22-24, 28-40), and some cases are designed where all control variables are subject to large emission changes (case 1-3, 7-9, 13-15, 19-21, 25-27). A more detailed description of the out-of-sample control scenarios is

given in Table S3. (from Page 12, Line 29 to Page 13, Line 4 in the revised manuscript)

p. 5063 l. 13 and 17: changes of total emissions: in which regions? Does this refer to all regions reducing at the same time? This should also be mentioned in the caption to Fig. 3.

Response: Yes, this refers to reducing emissions in all regions at the same time. We have revised this sentence as well as the caption of Fig. 3 (Fig. 4 in the revised manuscript) to make it clear. The revised texts are as follows.

We further evaluated the performance of the ERSM technique by comparing the 2D-isopleths of $PM_{2.5}$ concentrations in response to the simultaneous changes of $NO_X/SO_2/NH_3$ emissions in all regions derived from both the conventional RSM and the ERSM technique.

Figure 4. Comparison of the 2-D isopleths of $PM_{2.5}$ concentrations in Shanghai in response to the simultaneous changes of precursor emissions in all regions derived from the conventional RSM technique and the ERSM technique in Shanghai. The X- and Y-axis shows the emission ratio, defined as the ratios of the changed emissions to the emissions in the base case. The different colors represent different $PM_{2.5}$ concentrations.

Figures 3-5: If the authors would like to show the comparison for all three regions, then they should discuss differences and characteristics – otherwise one example would be sufficient.

Response: We agree with the reviewer that one example is sufficient. In the revised manuscript we keep only Figure 3 in the main text (Figure 4 in the revised manuscript) and move Figure 4 and Figure 5 to the Supporting Information.

Figures 6-7: Is there an easy explanation for the sometimes significant differences between effects of reducing individual pollutants, and reducing all of them together? In most cases the combined effect seems lower than the sum of individual effects, which might be explained by overlap effects of reductions in both species involved in the formation of ammonium sulfate and ammonium nitrate, but also the opposite is the case? It would be interesting if the authors

could comment on this.

Response: We appreciate the reviewer's valuable comment. The reviewer is correct that the combined effect is lower than the sum of individual effects in most cases because of the overlap effects of reductions in both species involved in the formation of ammonium sulfate and ammonium nitrate. However, it is sometimes the other way around in January, as shown in the left part of Figure 7 (Figure 6 of the revised manuscript, as shown below). This is because, in January, the response of $PM_{2.5}$ to NO_X emissions is negative for relatively small reductions (< 40-70%), but becomes positive for large reductions (> 40-70%). The reduction of NO_X emissions of a single emission sector leads to the increase of $PM_{2.5}$ even at large reduction ratio since the emission reduction is small compared with the total NO_X emissions. Therefore, when the NO_X emissions from each sector are reduced individually (the bars), we sum up the negative effects. In contrast, when all pollutants from all sectors are reduced simultaneously (the red dotted line), the NO_X emission reduction at large ratio could have positive effect on $PM_{2.5}$ reduction. This is why the combined effect sometimes exceeds the sum of individual effects in January. We have added the explanations accordingly in the revised manuscript. (from Page 15, Line 24 to Page 16, Line 3 in the revised manuscript)

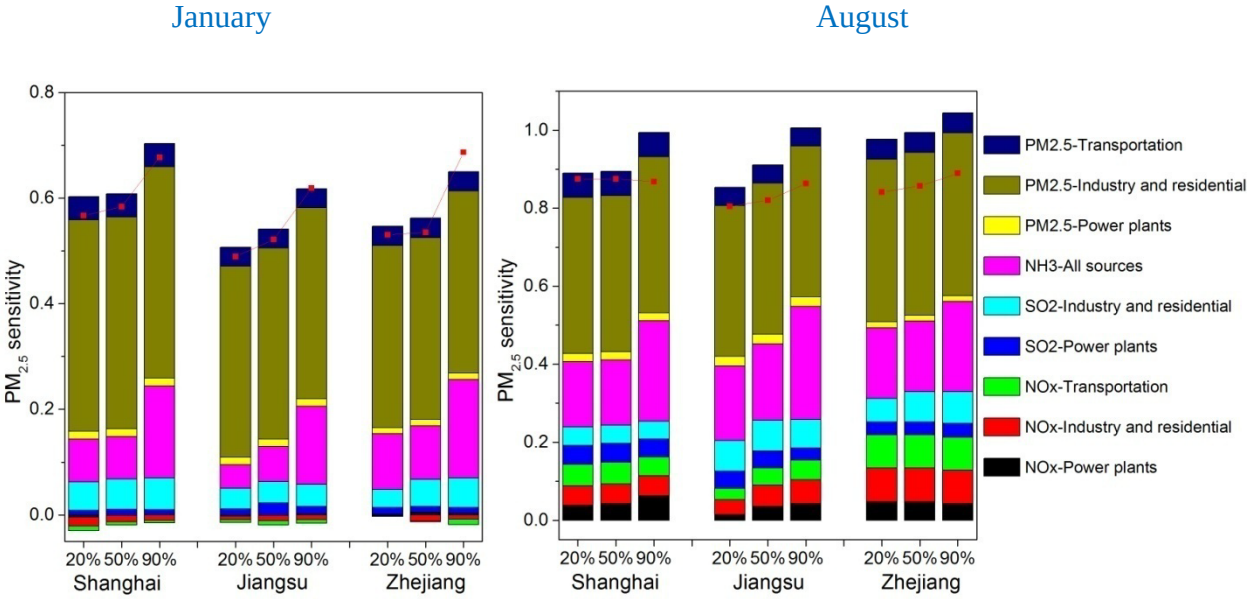

Figure 7 (Figure 6 of the revised manuscript). Sensitivity of $PM_{2.5}$ concentrations to the stepped control of individual air pollutants from individual sectors. The X-axis shows the reduction ratio (= 1 – emission ratio). The Y-axis shows $PM_{2.5}$ sensitivity, which is defined as the change ratio of concentration divided by the reduction ratio of emissions. The colored bars denote the $PM_{2.5}$ sensitivities when a particular emission source is controlled while the others stay the same as the base case; the red dotted line denotes the $PM_{2.5}$ sensitivity when all emission sources are controlled simultaneously.

3 Technical corrections

p. 5056 l. 29: “the same as NH3”: I assume what is meant is “equivalent for NH3”.

Response: Revision has been made. (Page 7, Line 8 of the revised manuscript)

p. 5058 l. 21: “changes”: change

Response: Revision has been made. (Page 8, Line 27 of the revised manuscript)

p. 5059, l. 9: “the same as NH3”: I assume what is meant is “equivalent for NH3”.

Response: Revision has been made. (Page 9, Line 18 of the revised manuscript)

p. 5060 l. 17: “interpolates”: interpolate

Response: Revision has been made. (Page 10, Line 30 of the revised manuscript)

Figures 3-5: Axes font sizes are very small and hardly readable, I suggest to increase them. Also, I wonder whether there is a reason for the upside-down colorbars, otherwise I’d suggest inverting them.

Response: We have increased the font sizes and inverted the colorbars. The revised Figure 3 (Figure 4 in the revised manuscript) is shown below.

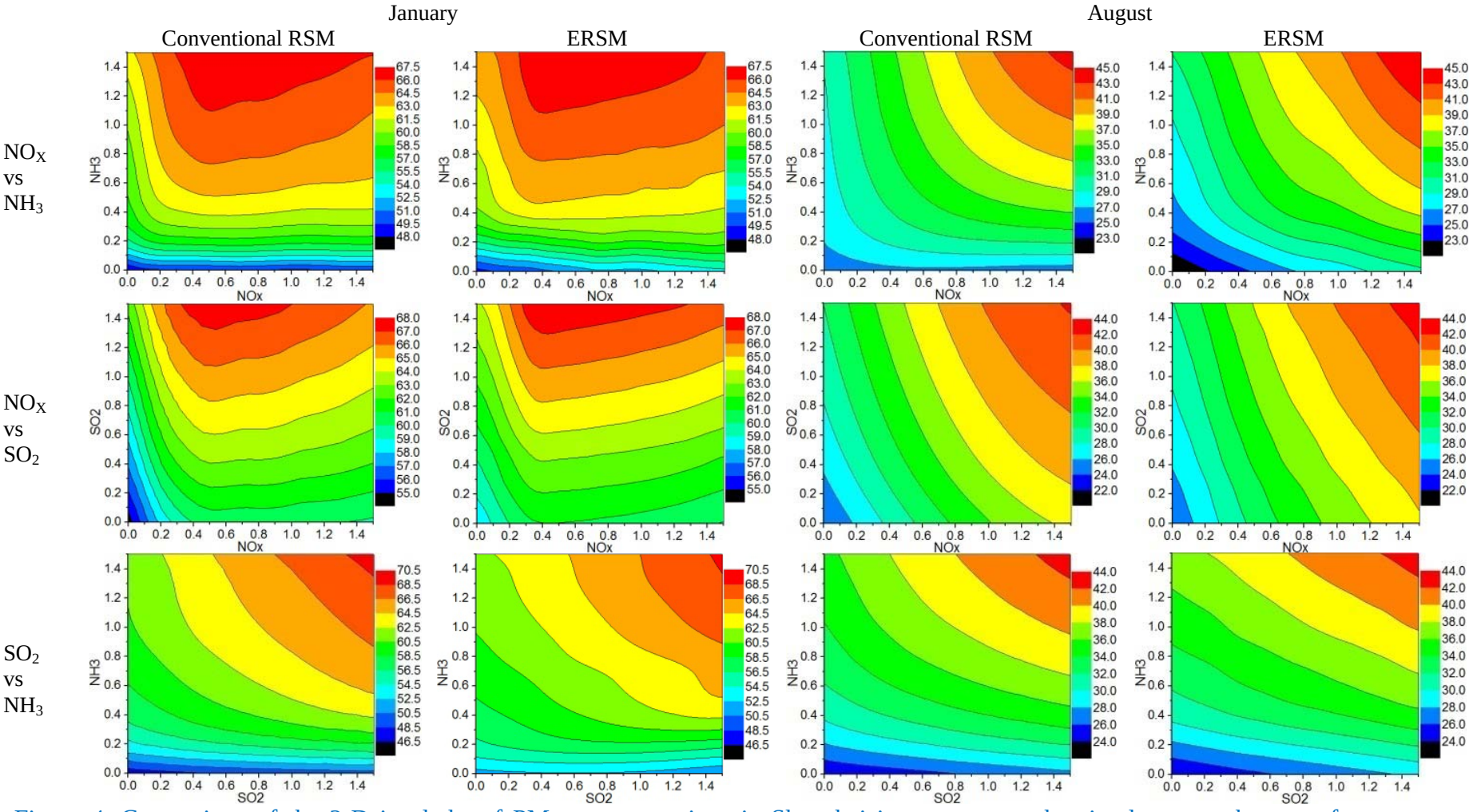

Figure 4. Comparison of the 2-D isopleths of $PM_{2.5}$ concentrations in Shanghai in response to the simultaneous changes of precursor emissions in all regions derived from the conventional RSM technique and the ERSM technique. The X- and Y-axis shows the emission ratio, defined as the ratios of the changed emissions to the emissions in the base case. The different colors represent different $PM_{2.5}$ concentrations (unit: $\mu g\ m^{-3}$).