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.

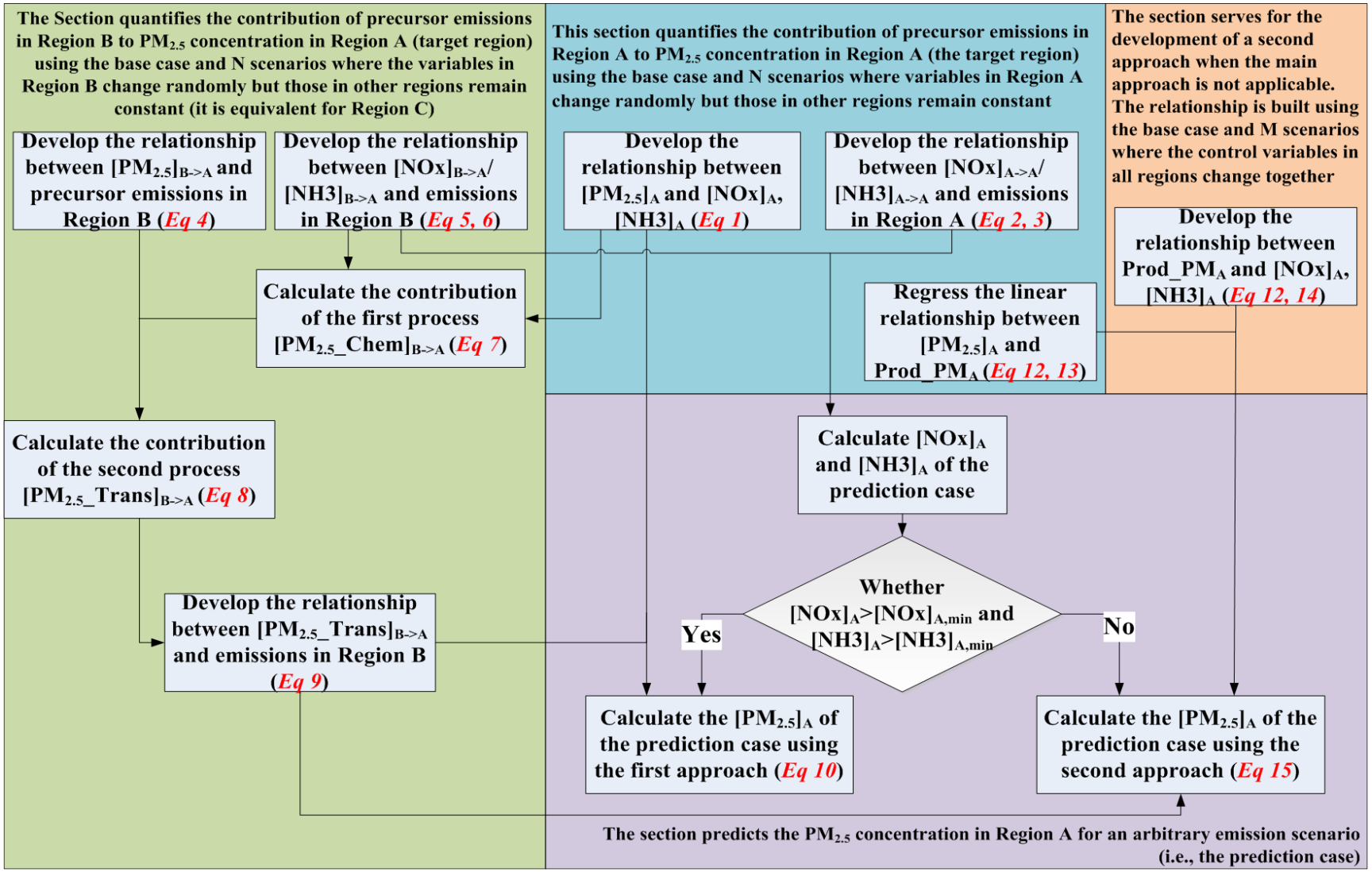

Figure 1. A flowchart illustrating the ERSM technique using the simplified case described in Sect. 2.1. Different background colors represent the procedures conducted using different groups of emission scenarios, as indicated on the top/bottom of the colored areas.

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, $[PM_{2.5}_Trans]_{B\rightarrow A}$ and $[PM_{2.5}_Trans]_{C\rightarrow A}$ 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 $[PM_{2.5}_Trans]_{B\rightarrow A}$ 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 $[PM_{2.5}_Trans]_{B\rightarrow A}$ and $[PM_{2.5}_Trans]_{C\rightarrow A}$ 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.