JOURNAL OF ENVIRONMENTAL SCIENCES 27 (2015) 97–107

Available online at www.sciencedirect.com

ScienceDirect

www.journals.elsevier.com/journal-of-environmental-sciences

JOURNAL OF ENVIRONMENTAL SCIENCES

www.jesc.ac.cn

Development and case study of a science-based software platform to support policy making on air quality

Yun Zhu[1,2,*], *Yanwen Lao*[3], *Carey Jang*[4], *Chen-Jen Lin*[1,5], *Jia Xing*[2], *Shuxiao Wang*[2], *Joshua S. Fu*[6], *Shuang Deng*[7], *Junping Xie*[3], *Shicheng Long*[3]

1. *Guangdong Provincial Key Laboratory of Atmospheric Environment and Pollution Control, College of Environmental and Energy, South China University of Technology, Guangzhou Higher Education Mega Center, Guangzhou 510006, China. E-mail: zhuyun@scut.edu.cn*
2. *State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, School of Environment, Tsinghua University, Beijing 100084, China*
3. *College of Environmental and Energy, South China University of Technology, Guangzhou Higher Education Mega Center, Guangzhou 510006, China*
4. *USEPA/Office of Air Quality Planning & Standards, RTP, NC 27711, USA*
5. *Department of Civil Engineering, Lamar University, Beaumont, TX 77710-0024, USA*
6. *Department of Civil & Environmental Engineering, University of Tennessee, Knoxville, TN 37996-2010, USA*
7. *Chinese Research Academy of Environmental Sciences, Beijing 100012, China*

ARTICLE INFO

Article history:
Received 3 February 2014
Revised 15 July 2014
Accepted 28 July 2014
Available online 26 November 2014

Keywords:
Air quality
Policy making
Response surface modeling
Emission control scenarios
Data visualization

ABSTRACT

This article describes the development and implementations of a novel software platform that supports real-time, science-based policy making on air quality through a user-friendly interface. The software, RSM-VAT, uses a response surface modeling (RSM) methodology and serves as a visualization and analysis tool (VAT) for three-dimensional air quality data obtained by atmospheric models. The software features a number of powerful and intuitive data visualization functions for illustrating the complex nonlinear relationship between emission reductions and air quality benefits. The case study of contiguous U.S. demonstrates that the enhanced RSM-VAT is capable of reproducing the air quality model results with Normalized Mean Bias <2% and assisting in air quality policy making in near real time.

Introduction

Continued economic development has led to air quality degradation in urban areas throughout the world. Air pollution is a result of complex interactions among air pollutant emissions, meteorological conditions and a wide variety of atmospheric processes including transport, chemical transformation and deposition. To improve air quality without impairing economic development, strategies of air pollutant emission control must be carefully formulated by assessing the fate of air pollutant emissions in the atmosphere. Typically, such assessments are achieved by air quality modeling using science-based atmospheric models, such as the Community Multi-scale Air Quality (CMAQ) model (Byun and Schere, 2006; Wang et al.,

* *Corresponding author.*
E-mail address: zhuyun@scut.edu.cn (Y. Zhu).

http://dx.doi.org/10.1016/j.jes.2014.08.016
1001-0742/

2014a, 2014b), for multiple emission scenarios. However, the high computational cost associated with the modeling effort often becomes the bottleneck of the assessment process. In addition, most atmospheric models used in air quality management lack a user-friendly interface that can synthesize the model results produced from different air pollutant emission inputs, and do not offer adequate data visualization for supporting policy making.

To improve the efficiency of data synthesis in air quality management, it is necessary to develop a tool to translate the massive amount of model data to policy-relevant visualization. Ideally, such policy decision support tools should be efficient, easy-to-use, and capable of estimating air quality changes for given emission reduction scenarios. In addition, the tools should offer intuitive policy-related visualizations. Some earlier prototypes, such as: the IMPAQT — Integrated Modular Program for Air Quality Tools (Lim et al., 2005) and the GIS-CALPUFF coupled air quality decision support system (Fraser et al., 2013), have demonstrated the benefits of a decision support tool for urban air quality assessment with limited visualization features. To our knowledge, there has not been a comprehensive package that has been demonstrated to support air quality management at both urban and regional scales. Since many air quality problems are of regional scale (Wang et al., 2014a, 2014b), for example regional haze and trans-boundary transport, an efficient and powerful tool for supporting decision making in regional air quality is in great need.

To address this need, we rewrote the code of the response surface model (RSM) (Lao et al., 2012) with C# using the high-dimensional Kriging techniques (Mohammadi et al., 2012; Xiao et al., 2009) to interpolate the model results of CMAQ for policy making support. The RSM serves as a statistical metamodeling interface for rapid assessment of the massive data of atmospheric model results (Ashok et al., 2013). Following the data assessment, it is essential to couple the model results in a graphical user interface such as USEPA's Visual Policy Analyzer (VPA), which allows using slider bars and a mouse to change emission factors for visualizing the air quality response (U.S.EPA, 2006). However, VPA has incurred user-friendliness and data accuracy issues. At this moment, the development and support of VPA have been discontinued.

In this work, an enhanced response surface model-visualization and analysis tool (RSM-VAT) is developed for better understanding and analyzing the established statistical relationship (RSM) between air pollution emission reductions and air pollutant concentrations. The statistically generalized RSM surface by the method of MPerK (MATLAB Parametric Empirical Kriging) reported in our previous work has demonstrated that the nonlinear relationship between emission sources and air quality concentration of $PM_{2.5}$ (Wang et al., 2011) or O_3 (Xing et al., 2011) can be well analyzed by RSM. The established nonlinear relationship is not only a foundation of decision making, but also critical for formulating emission control strategies. The development of enhanced RSM-VAT is the extension of our previous work, it aims to provide a series advanced visualization and analysis functions for validating RSM prediction results and for examining the nonlinear interactions among multiple air pollutant precursors, in addition to instantly generate the air pollutant concentration surface response to air pollutant emission changes.

1. Methodology

Fig. 1 shows the block diagram of the development and application of the RSM-VAT. The modules of control matrix, RSM, and quality assurance (QA) are designed to guide the users through the RSM development and accuracy validation. Most end users (policy makers) would directly use the RSM data analysis module for policy decisions. It greatly improves the user interface and visualization accuracy of VPA for the RSM results. The design of control matrix, response surface modeling, quality assurance, and RSM data analysis are described here.

1.1. *Design of emission control matrix*

RSM aggregates the results of pre-specified CMAQ simulations into a multi-dimensional "response surface" that also incorporates the acceptable uncertainties in environmental decision-making. This allows a rapid assessment of air quality impacts caused by different combinations of emission sources. These sources are selected as emission control factors as shown in Table 1. The CMAQ simulations are performed at a defined set of distributed emission inventories (control scenarios) in a high-dimensional experimental design space. For example, as one of the source targets subject to emission control, Electricity Generating Units (EGUs) NO_x emission is set at 0–120% of the base-year level. The response surface of the CMAQ results attempts to maintain the model accuracy while minimizing the computationally expensive CMAQ run at a given emission scenario within the experimental design space. The RSM method uses statistical techniques to relate a response variable (*e.g.*, the annual $PM_{2.5}$ at multiple U.S. receptor sites) to its influencing factors (*e.g.*, the emission of a pollutant precursor, NO_x) from local and regional emission sources.

The control matrix module (Fig. 1) is to create the experiment design of emission control matrix. The matrix is created by sampling the control factors in the design space (*e.g.*, from 0.00 to 1.20) using Latin Hypercube Sampling — LHS (Hirabayashi et al., 2011). LHS is often applied to generate the experimental design for a Kriging model (Kleijnen, 2009). It is a statistical method for generating a sample of plausible collections of parameter values from a multidimensional distribution which is more efficient than random sampling for a large number of factors. Table 2 shows an example of emission control matrix created by LHS. The base-year emission inventory is factored by the weight shown in the matrix (Table 2), forming 181 emission control scenarios for CMAQ model runs. These emission inventory factors, together with the identical meteorology and other model input, are applied in CMAQ simulations to provide the annual $PM_{2.5}$ concentration estimates for the 181 emission control scenarios. In order to reduce computational time for such a large number of annual model runs, the results of 4-monthly (February, April, July, and October 2001) CMAQ runs over the contiguous US are utilized in the response surface modeling. These months were chosen based on greatest predictability of the quarterly mean in US. The emission projections for three future years (2010, 2015 and 2020) together with 2001 meteorology data (U.S.EPA, 2005) are used in the CMAQ runs (U.S.EPA, 2006).

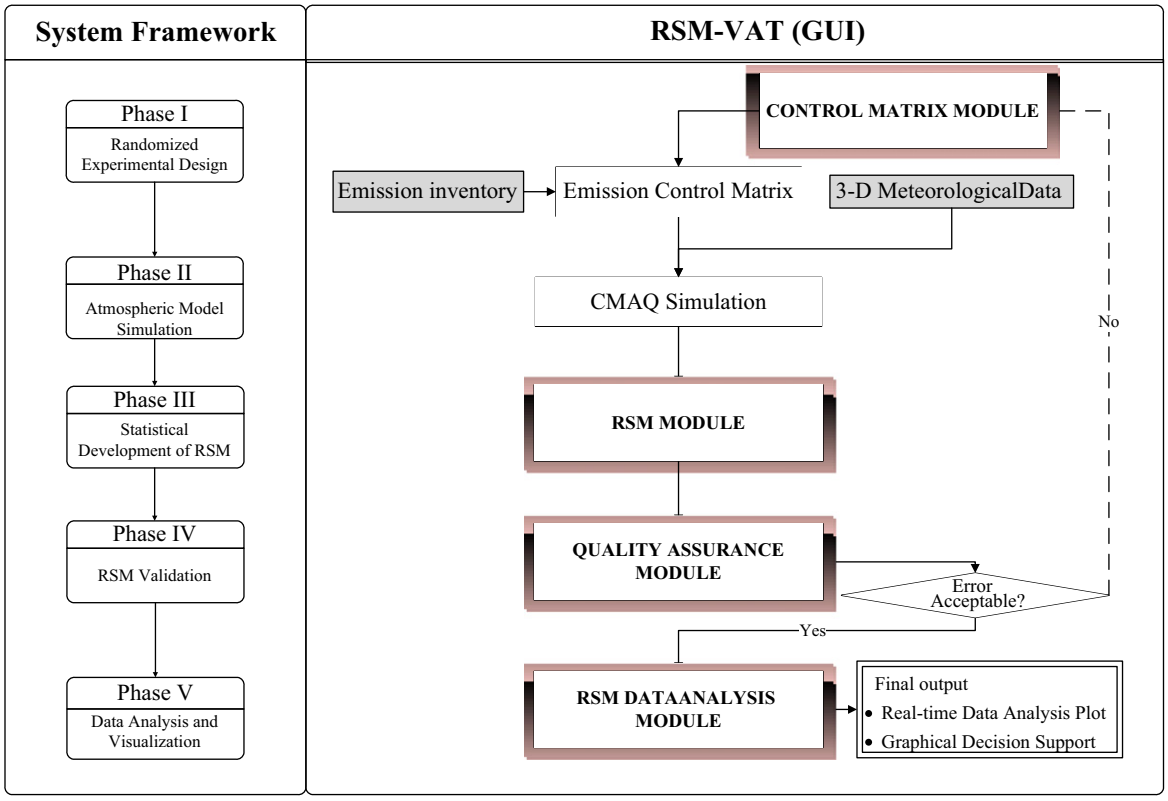


Fig. 1 – The primary components and the typical application process of the developed response surface model-visualization and analysis tool (RSM-VAT).

Table 1 – Twelve emissions control factors selected for response surface modeling (RSM) modeling.

Control factors	Factor description
NO_x/EGU	NO_x IPM (Integrated Planning Model) EGU (Electricity Generation Unit) point source emissions
NO_x/ NonEGU + area	NO_x IPM non-EGU point source, area source, and agricultural source emissions
NO_x/mobile	NO_x nonroad source and mobile source emissions
SO_x/EGU	SO_x IPM EGU point source emissions
SO_x/ NonEGU_Point	SO_x IPM Non-EGU point source emissions
SO_x/area	SO_x area source and agricultural source emissions
VOC/all	Volatile organic carbon IPM EGU point source, IPM Non-EGU point source, area source, agricultural source, nonroad source, and mobile source emissions
NH_3/area	Ammonia area source and agricultural source emissions
NH_3/mobile	Ammonia non-road source and mobile source sources
POC & PEC/ EGU + NonEGU	Elemental carbon and organic carbon IPM EGU point source and IPM Non-EGU point source emissions
POC & PEC/ mobile	Elemental carbon and organic carbon nonroad source and mobile source emissions
POC & PEC/area	Elemental carbon and organic carbon area source and agricultural source emissions

Twelve emission sources, which are treated separately in local and regional area, are selected for the RSM model development (Table 1). All the urban emissions are controlled using the same control factor. Under this circumstance, the within-city emission is considered "local" and the emissions in other cities are considered "regional" such that the emission in one city does not significantly influence the air quality in another city. This is a reasonable assumption as long as the selected urban areas are sufficiently distant from each other (Fig. 2). The 8 urban areas include New York/Philadelphia (combined), Chicago, San Joaquin, Atlanta, Salt Lake City, Phoenix, Seattle, and Dallas.

The emission control factors range from 0 to 120%. A zero control factor means that the emission of interest is reduced to zero while a 120% factor indicates that the emission is increased by 20%. The LHS method ensures a uniformly distributed sampling in the high-dimensional experimental design space such that it can capture the linear and nonlinear interactions among pollutants. The emission control factors in the emission control matrix (Tables 1, 2) are selected based on emission types and source categories. This allows the generated RSM to evaluate air quality changes resulting from adjusting each of the twelve emission sources (Table 1) on urban or regional basis. This experimental design ensures the representativeness of the data used in the response surface regression.

1.2. Development of response surface modeling

A high-dimensional Kriging approach is applied for the response surface regression of the CMAQ simulation results

Table 2 – Emission control matrix created by control matrix module using Latin Hypercube Sampling (LHS) method.

Run [a]	R-1 (base)	…	R-178	R-180	R-181
(1) Local/NO_x/EGU	1	…	0.110483	1.022551	1.087 341
(2) Local/NO_x/ NonEGU + area	1	…	0.027116	0.06448	0.109 682
(3) Local/NO_x/mobile	1	…	0.043133	1.049465	0.029 364
(4) Local/SO_x/EGU	1	…	1.048884	1.008878	1.023 087
(5) Local/SO_x/ NonEGU_Point	1	…	1.086225	0.555295	1.024 059
(6) Local/So_x/area	1	…	0.108373	1.099252	0.353 902
(7) Local/VOC/all	1	…	1.021591	0.036443	1.091 375
(8) Local/NH_3/area	1	…	1.049548	1.082221	1.008 858
(9) Local/NH_3/mobile	1	…	1.080622	0.521056	0.357 506
(10) Local/POC & PEC/ EGU + NonEGU	1	…	1.084612	0.357084	0.637 807
(11) Local/POC & PEC/mobile	1	…	0.965029	0.106736	0.522 265
(12) Local/POC & PEC/area	1	…	1.084325	0.034402	0.281 576
(13) Region/NO_x/EGU	1	…	0.762628	1.087254	0.718 672
(14) Region/NO_x/ NonEGU + area	1	…	0.765136	0.283111	0.392 529
(15) Region/NO_x/ mobile	1	…	0.930531	0.318877	0.085 269
(16) Region/SO_x/EGU	1	…	1.155187	1.086086	0.763 455
(17) Region/SO_x/ NonEGU_Point	1	…	0.717255	0.260321	0.418 828
(18) Region/SO_x/area	1	…	1.013104	0.035435	1.089 675
(19) Region/VOC/all	1	…	0.389028	0.608322	0.266 133
(20) Region/NH_3/area	1	…	1.184101	0.032021	0.062 598
(21) Region/NH_3/ mobile	1	…	0.060786	0.714711	0.260 525
(22) Region/POC & PEC/ EGU + NonEGU	1	…	0.107094	0.556018	0.036 105
(23) Region/POC & PEC/mobile	1	…	1.084976	0.635975	0.285 701
(24) Region/POC & PEC/area	1	…	1.009483	0.634597	1.089 613

[a] Run #1–171 are used for creating RSM and run #172–181 are used for out-of-sample validation.

(Ginsbourger et al., 2013; Jia and Taflanidis, 2013). Kriging is a geostatistical method based on an exponentially weighted sum of the sample runs results, which is reported superior to deterministic methods (Eldrandaly and Abu-Zaid, 2011). It can approximate highly nonlinear surfaces as long as they are locally continuous. The predicted changes in the ambient concentration of $PM_{2.5}$ in each CMAQ grid cell are modeled as a function of the weighted average of the modeled responses among the experimental design space. The weight assigned to a model output depends on the Euclidean distance between the factor levels covering the emission control ranges and the factor levels defining the CMAQ experimental runs. A Maximum Likelihood Estimation–Experimental Best Linear Unbiased Predictors (MLE–EBLUPs) Kriging Interpolation is adopted. We recoded the Parametric Empirical Kriging algorithm (Kleijnen, 2009; Peng et al., 2006) using C# language and accelerated the calculation efficiency by multi-threading technology. The RSM predicted value $\vec{Y}(x_0)$ can be expressed as:

$$\vec{Y}(x_0) = \sum_{j=1}^{d} f_j(x)\vec{\beta}_j + z(x) \equiv f_0^T\vec{\beta} + \vec{\gamma}_0^T\vec{R}^{-1}\left(Y^n - F\vec{\beta}\right) \quad (1)$$

where, f_0 is a $d \times 1$ vector of the regression functions for $\vec{Y}(x_0)$, and $\vec{\beta}$ is a $d \times 1$ vector for the unknown regression coefficients; γ_0 is the $n \times 1$ vector of correlations of Y^n with $\vec{Y}(x_0)$; $\vec{R}$ is a $n \times n$ matrix of correlations among the Y^n; F is a $n \times d$ matrix of regression functions for the training data; $z(x)$ is a zero-mean Gaussian stochastic process with correlation function $R(h|\xi)$ and correlation parameters ξ. R can be estimated by the product power exponential correlation function:

$$R(h|\xi) = \prod_{i=1}^{d} \exp[-\theta_i|h_i|^{p_i}] \quad (2)$$

where, $\xi = (\theta, p) = (\theta_1, \ldots, \theta_d, p_1, \ldots, p_d)$ with $\theta_i \geq 0$ and $0 \leq p_i \leq 2$, and the correlation parameter ξ is obtained by maximum likelihood estimation (Li and Sudjianto, 2005; Xing et al., 2011). Since the RSM is built from CMAQ simulation results, it has the same strengths and limitations of the CMAQ modeling system as well as its input data driving the model runs.

1.3. Quality assurance of RSM

The developed RSM is validated using a number of model performance evaluation metrics as shown in Table 1. Individual grad cell scatterplot and visual examination of RSM prediction maps is also conducted to ensure that the spatial distribution of air quality obtained by the RSM is consistent with the original CMAQ model results. Cross validation (CV) was performed. During each CV routine, one of the experimental runs is left out for creating the RSM (CV-RSM). The CV-RSM predicted data are then verified against the CMAQ results that are left out. This comparison process is sequentially carried out for 171 times. It is found that the predictions of CV-RSM and RSM do not have observable differences. Out-of-sample validation (OOS) is also conducted to evaluate the RSM performance, which compares RSM-predicted values to those of CMAQ for a set of model runs not used in developing the RSM. Ten OOS model runs are evaluated in the case study. The same performance evaluation metrics (Table 3) were used for both CV and OOS methods.

The users also have the options of using add-on sampling to generate additional runs for CMAQ simulations if the RSM prediction bias does not meet the desired tolerance. Combined with the original model runs, a new sampling matrix and additional CMAQ simulation results can be applied to create a new response relationship. The new RSM can be examined through the validation process until the desired accuracy is achieved. Theoretically, the more experimental runs are used for creating the RSM, the more accurate the RSM results will be. However, the improvement of accuracy and the computational cost of CMAQ simulations should be balanced. The examples of input data file, including: CMAQ model output concentration files, emission factor information file, state

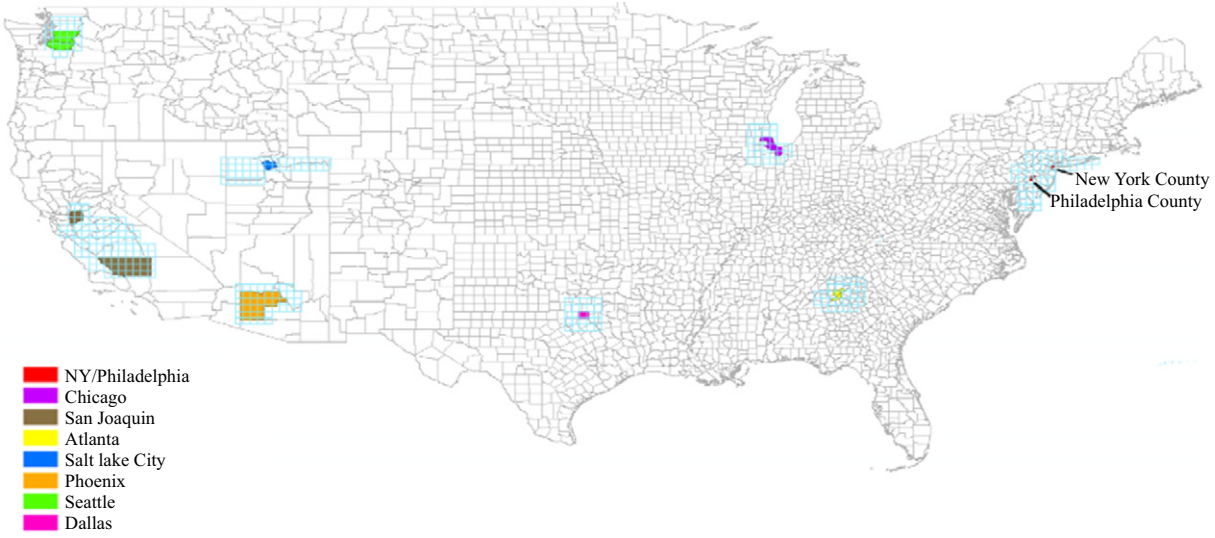


Fig. 2 – Study domain and 8 cities of interest in 36 km grid.

shape file (region boundary map file), area information file (target area or region file) and control matrix file are provided by U.S.EPA (2006) and can be downloaded by visiting http://www.abacas-dss.com/. Those data files are utilized to create the RSM and to support the visualization and analysis tool in the case study.

Table 3 – Evaluation metrics of RSM (Response Surface Model).

Performance metric	Equation [a]
Mean Bias (MB, μg/m^3)	$MB = \frac{1}{N}\sum_{i=1}^{N}(Cri - Cci)$
Mean Error (ME, μg/m^3)	$ME = \frac{1}{N}\sum_{i=1}^{N}\lvert Cri - Cci\rvert$
Mean Normalized Bias (MNB, −100% to +∞)	$MNB = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Cri - Cci}{Cci}\right)$
Mean Normalized Error (MNE, 0% to +∞)	$MNE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{Cri - Cci}{Cci}\right\rvert$
Mean Fractional Bias (MFB, −200% to +200%)	$MFB = \frac{1}{N}\sum_{i=1}^{N}\frac{(Cri - Cci)}{\left(\frac{Cri+Cci}{2}\right)}$
Mean Fractional Error (MFE, 0% to +200%)	$MFE = \frac{1}{N}\sum_{i=1}^{N}\frac{\lvert Cri - Cci\rvert}{\left(\frac{Cri+Cci}{2}\right)}$
Normalized Mean Bias (NMB, −100% to + ∞)	$NMB = \frac{\sum_{i=1}^{N}(Cri - Cci)}{\sum_{i=1}^{N} Cci}$
Normalized Mean Error (NME, 0% to + ∞)	$NME = \frac{\sum_{i=1}^{N}\lvert Cri - Cci\rvert}{\sum_{i=1}^{N} Cci}$

[a] N is the number of grid cells, Cri is the RSM predicted value in grid i, and Cci is the CMAQ value in grid i.

2. Results & discussion

2.1. Performance of the RSM

The $PM_{2.5}$ results are used as an example for illustrating the RSM model performance. Comparison of RSM predictions to the CMAQ simulation results shows good agreement in both the 171 CV cases and the 10 OOS cases (Fig. 3). The median values of NMB, NME, MFB and MFE for the OOS cases (0.18–0.85%) in the box-plot are slightly higher than those of CV cases (−0.31%–0.87%). We choose Case #178 whose MFB is the highest at 3.69% and Case #180 whose MFB is the lowest at −0.11% (Fig. 3, line chart) to further analyze the model accuracy. Shown as Fig. 4a, the RSM accurately reproduces the results of Case #180 and slightly over-predicts those of Case #178. The over-predictions of Case #178 can be viewed clearly when we compare the base case concentration reduction in Fig. 4b because of the amplified coordinate axis. Fig. 5 shows that the concentration distribution map of RSM predicted values can well coordinate with that of the CMAQ simulation results. The slight biases seem more easily occurred in the regions with high $PM_{2.5}$ concentration because of the greater emission reduction (Fig. 5, RSM–CMAQ), while the biases of these regions are less than 0.25 μg/m^3 compared to their ambient $PM_{2.5}$ concentration from 11.58 to 16.54 μg/m^3 using San Joaquin and New York/Philadelphia as examples. Comparing the LHS sampling values between the control factors of Case #178 and Case #180 listed in Table 2, the control factor of "region/NH_3/Area" in Case #178 is 1.184101 and only 0.032021 in Case #180. Additional experiment runs for the emission control factors having a significant impact on

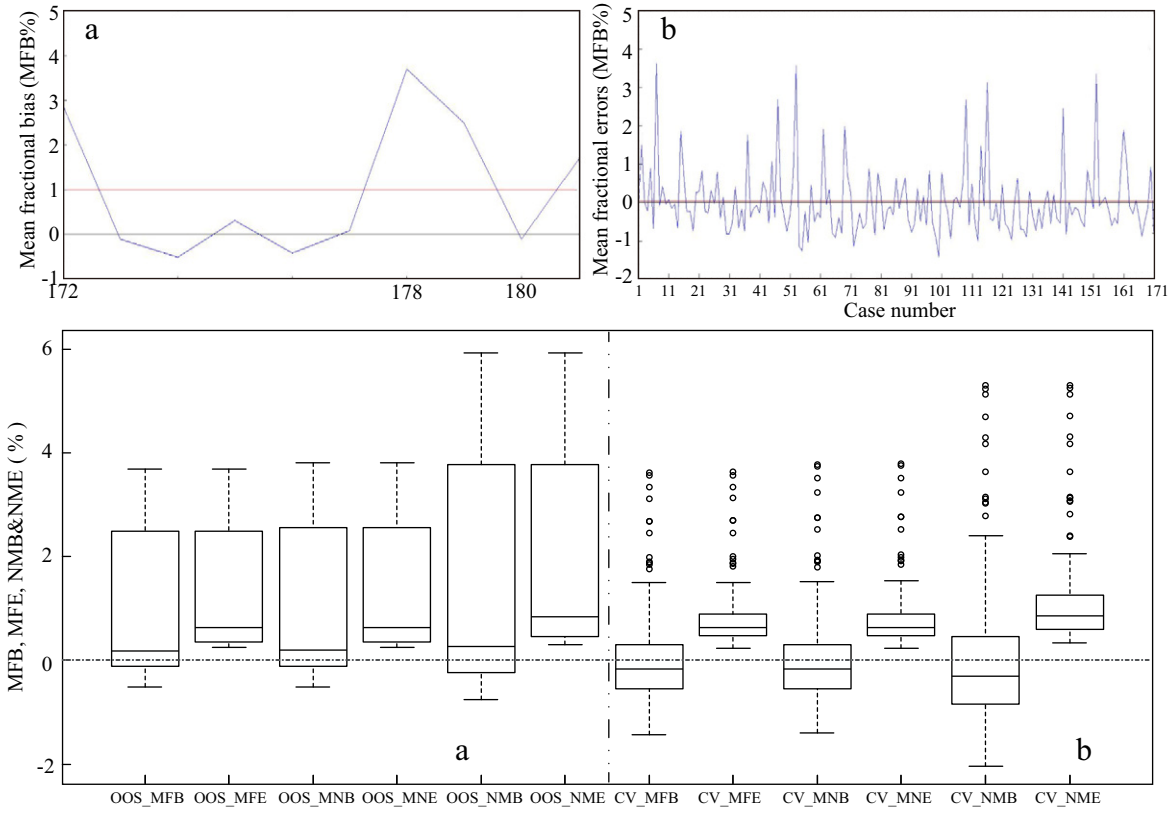


Fig. 3 – Normalized Mean Bias (NMB), Normalized Mean Error (NME), Mean Fractional Bias (MFB) and Mean Fractional Error (MFE) over (a) the 10 outside of the experimental design runs by out-of-sample (OOS) validation, and (b) the 171 control scenarios used to create RSM by cross-validation (CV); the line chart evaluates the comparison percentage values (Y) with individual case number (X), and the box-plot demonstrates the statistical analysis of the bias for all of the 10 OOS or the 171 CV cases.

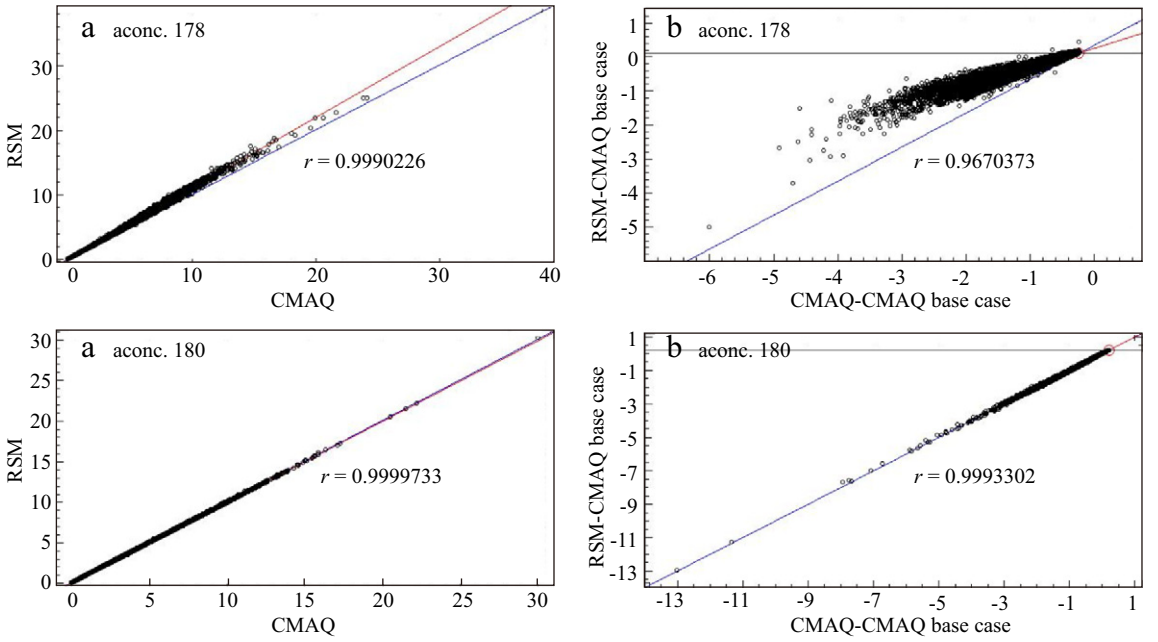


Fig. 4 – X–Y scatterplot for the concentrations of $PM_{2.5}$ ($\mu g/m^3$) simulated by CMAQ (X, Community Multi-scale Air Quality) and predicted by RSM (Y, Response Surface Model). (a) RSM *vs.* CMAQ, (b) the CMAQ base case concentration reduction (predicted or simulated value subtracts the concentration of CMAQ base case) of RSM *vs.* CMAQ.

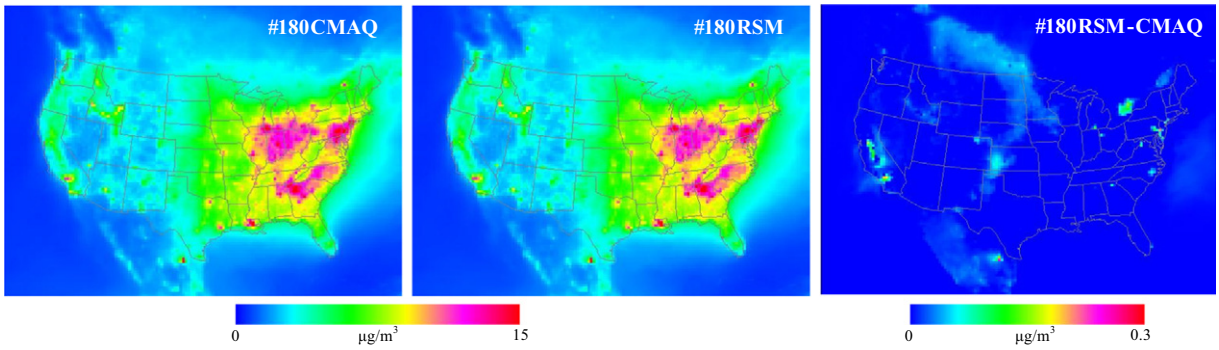


Fig. 5 – Concentration and different (RSM subtract CMAQ) maps for CMAQ simulated results and RSM predicted values created by the function of "comparison plot".

$PM_{2.5}$ are necessary if the RSM model accuracy needs to be further improved.

Validation & QA module also summarizes the values of the performance metrics for the RSM (Table 4). The mean bias is <2% and the maximum bias is <6% in the test cases. Increasing the sampling density will further improve the RSM prediction accuracy (Xing et al., 2011). The add-on sampling function of the control matrix module allows users to choose the factors, and to set the number of samples and the factor range. The module uses LHS to generate a sub-matrix supplementing the existing-matrix for additional sampling runs. For example, to further reduce the difference between RSM and CMAQ results, one can use the function of "add-on sampling" to generate additional 20 experiment runs. The CMAQ simulation results of the 20 additional runs together are then used to recreate the RSM.

2.2. RSM data analysis and visualization

The developed RSM-VAT provides a series of visualization and analytical functions for supporting decision making. Although the quality assurance module also provides visualized verifications for the RSM prediction against CMAQ results (Fig. 5) and for different CMAQ experiment runs, decision support is mainly provided by the RSM data analysis module. Fig. 6 lists the different input data files or configurations for different function modules, and shows the screenshots of the instant/real-time visualization and analysis features of RSM-VAT, which include 2D and 3D maps (Fig. 6a, b) for instant overview of air quality response, bar charts (Fig. 6c, d, e, f) for quantitatively analyzing and comparing the effectiveness of emission reduction, and contour plots (Fig. 6g, h) for dynamically description of the non-liner or liner relationship between two single/groups of emission control factors.

The RSM-VAT allows instantaneous visualization of the response surface of CMAQ outputs corresponding to emission changes. For example, users can change any emission factors and view the impact on the ambient concentration of $PM_{2.5}$ instantly. Fig. 7 displays the 2D and 3D map examples of the RSM-VAT. The 3D map has an instant three-dimensional visual effect for the air quality changes, while 2D map can identify these changes in different administrative regions intuitively. Although VPA also provides the visualization of 2D and 3D maps, it does not have the other visualization functions to support quantitative evaluation of the interrelationship among emission control factors. The bar charts and contour plots are important features for analyzing the complex relationship between local/regional emission reduction and their air quality effect.

2.3. Case study for decision making support

A test case is employed to illustrate how the RSM-VAT can be used to assist in air quality decision making support. Consider that a policy maker attempts to know: (1) what emission reduction is more effective for achieving the targeted air quality, (2) which emission source has a greater impact on air quality, and (3) what is the most effective emission reduction strategy for improving air quality? The answers can be provided with the RSM-VAT as demonstrated.

All of the emission sources (control factors) listed in Table 2 has different degrees of contribution to $PM_{2.5}$ which is a public-focused complex pollutant (Zhuang et al., 2014). The relative contributions to $PM_{2.5}$ of regional and urban

Table 4 – Results of cross- and out-of-sample validations generated by the function listed in "evaluation metrics of RSM".

Performance metric	Cross-validation (171 scenarios)			Out-of-sample validation (10 scenarios)		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
MB (μg/m^3)	0.0005	−0.0587	0.1513	0.0448	−0.0209	0.1933
ME (μg/m^3)	0.0328	0.0098	0.1517	0.0588	0.0084	0.1933
MNB (%)	0.03%	−1.41%	3.77%	1.03%	−0.51%	3.81%
MNE (%)	0.80%	0.22%	3.78%	1.39%	0.25%	3.81%
MFB (%)	0.02%	−1.43%	3.62%	0.99%	−0.51%	3.69%
MFE (%)	0.79%	0.23%	3.63%	1.35%	0.25%	3.69%
NMB (%)	0.03%	−2.04%	5.31%	1.52%	−0.75%	5.93%
NME (%)	1.13%	0.33%	5.31%	2.00%	0.30%	5.93%

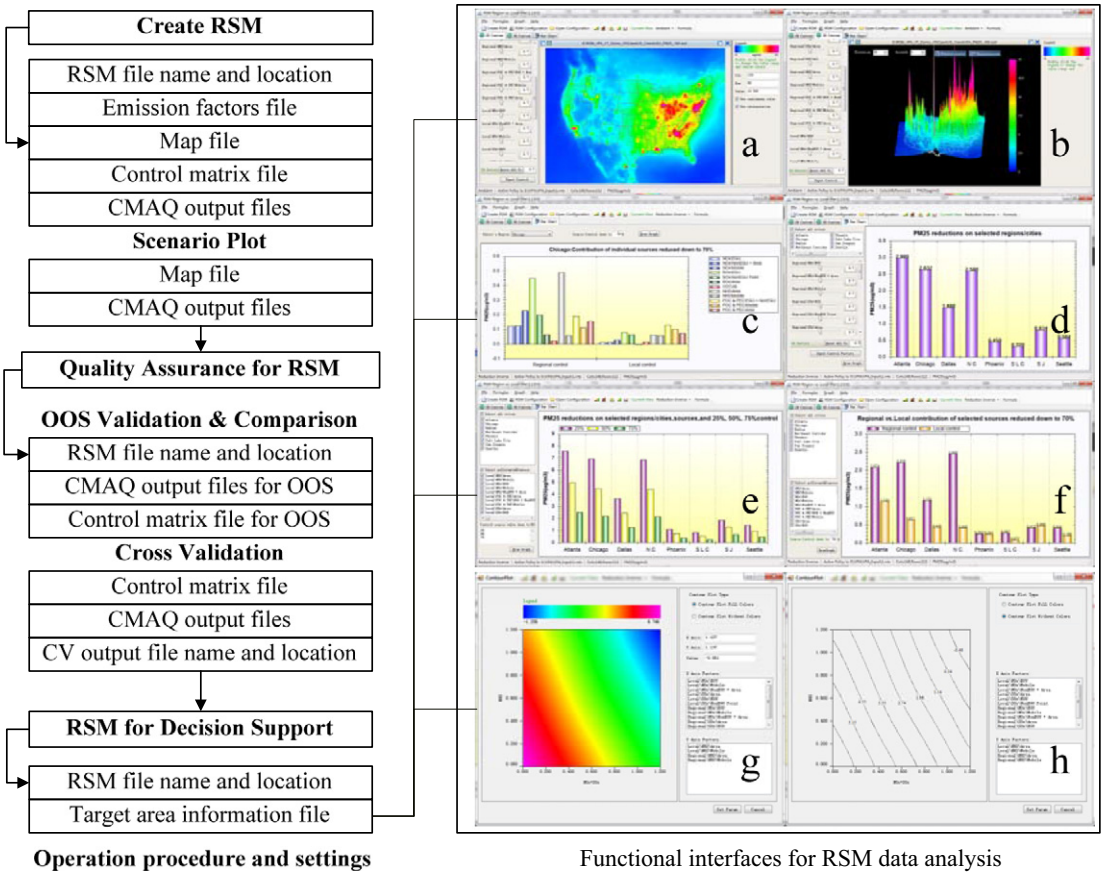
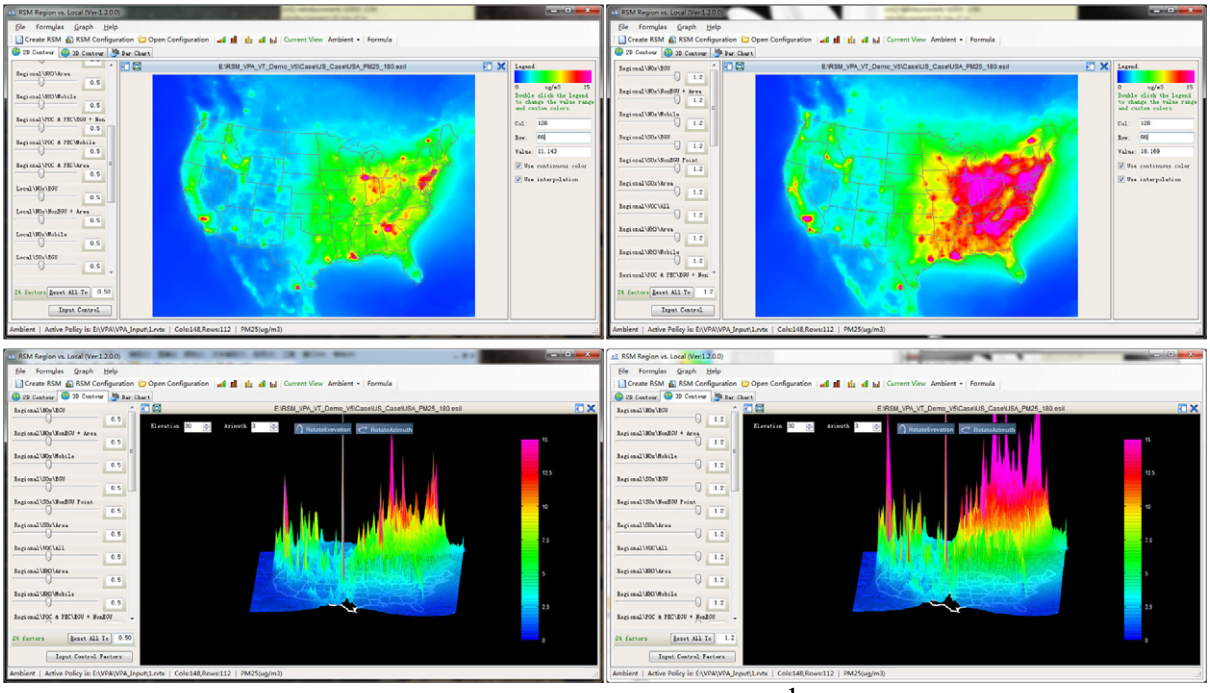


Fig. 6 – Visualization and analysis features of the RSM-VAT. The main functional modules are: (a) 2-D map, (b) 3-D map, (c) regional and local emission sources control effects, (d) air quality improvement on selected regions/cities, (e) comparison of multi-regions/cities for multi-emission–reduction-ratios, (f) regional and local emission sources control effects for chosen emission sources, (g) 2-D color filled contour, (h) 2-D contour.

emission sources for the eight urban areas (Fig. 2) can be obtained by using the regional and local emission control effect comparison plot function shown in Fig. 6f. The comparison result shows that the regional emissions have a greater air quality impact in Chicago while the urban emissions in San Joaquin are more important than the regional emissions. Additional information can be obtained by investigating the response curves of $PM_{2.5}$ caused by the changes of regional and local emissions using the contour plot. For example, Fig. 8a shows that reducing regional emissions is more effective in controlling $PM_{2.5}$ compared to reducing local emissions in Chicago. In contrast, the benefit caused by the local emission reduction in San Joaquin is slightly greater than that of regional emissions (Fig. 8b). Analyses can also be carried out by source categories and emission types for a subset of the geographic areas. The functions shown in Fig. 9 allows user to compare the cumulative impact and the different source contributions to $PM_{2.5}$ decrease caused by regional and local emission reduction under different emission control scenarios. The results in Fig. 9 show that the regional area-NH_3 emission and EGU-SO_x emission have a greater impact on $PM_{2.5}$ in Chicago, while the local emission of area-NH_3 and regional emission of mobile NO_x have a greater impact on $PM_{2.5}$ in San Joaquin. Therefore, we can use a combination of the contour plot (Fig. 8) and bar chart (Fig. 9) to provide sound and quantitative air quality improvement suggestions for different cities; for example, if we want to lower the concentration of annual $PM_{2.5}$ around 1.5 $\mu g/m^3$ in Chicago and San Joaquin, both of the regional and local emission sources of Chicago will be reduced down to 83.5% (Figs. 8a, 9), while those of San Joaquin will be reduced down to 52.8% (Figs. 8b, 9); in addition, the regional EGU-SO_x and area-NH_3 for Chicago, and local area-NH_3, regional and local mobile NO_x for San Joaquin should be focused on (Fig. 9). An additional CMAQ simulation for the emission reduction scenario of Chicago down to 83.5% was run to further investigate the accuracy of RSM. The comparison results show that RSM can well reproduce the simulation results of CMAQ (Fig. 10a) with the deviation of $PM_{2.5}$ concentrations less than 0.3 $\mu g/m^3$ (Fig. 10b) and MFB = 0.072%.

3. Conclusions

This article describes the development and application of a response surface modeling-visualization and analysis tool



a Emission control factors 0.50 b Emission control factors 1.20

Fig. 7 – Visualization effects of 2D and 3D contour for $PM_{2.5}$ concentration at emission control factors set to (a) 0.50 and (b) 1.2.

(RSM-VAT) for supporting the decision making of air quality policy. The RSM is based on a new approach that aggregates pre-specified air quality modeling results into a multi-dimensional "response surface". A sophisticated, high-dimensional Kriging interpolation approach is implemented in the development. It is shown that the RSM can reproduce the results from individual CMAQ simulations accurately.

The RSM-VAT is implemented on Windows operation system and has a user-friendly interface. The U.S. case study shown that users can easily obtain the graphic decision supports by entering emission control factors and/or clicking the mouse. The decision support features include air quality benefits resulted from emission reductions, determination of the relative importance of regional and local emission sources, and identification of important emission sources. RSM-VAT can also be used

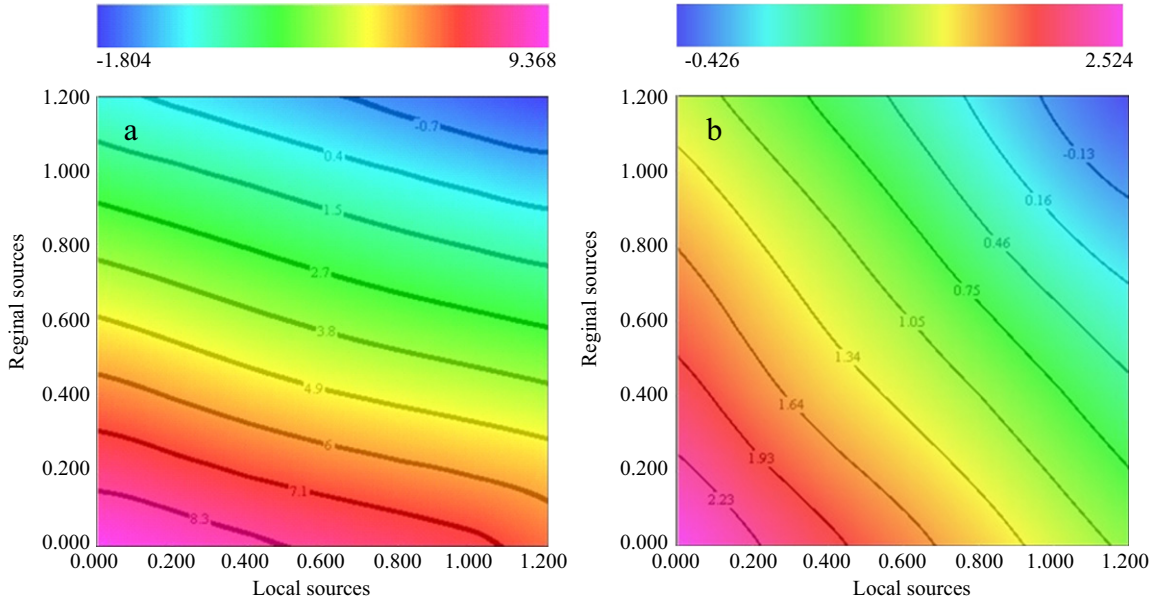


Fig. 8 – Contour curves for annual average $PM_{2.5}$ in (a) Chicago and (b) San Joaquin area response on the control of regional (Y) and local (X) emission sources.

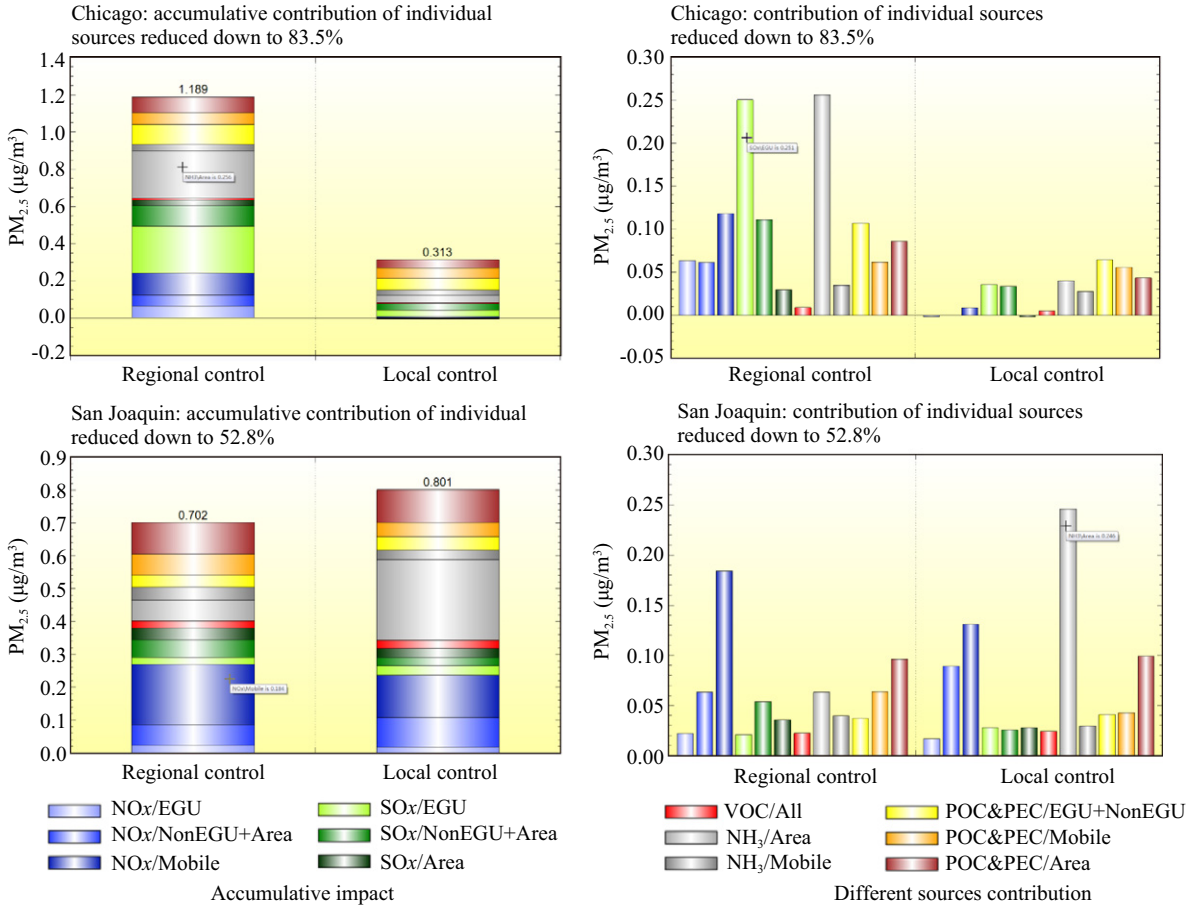


Fig. 9 – Accumulative impact and different sources contribution comparison of regional and local emission sources reduced down to 83.5% in Chicago and 52.8% in San Joaquin for lowering 1.5 μg/m^3 annual $PM_{2.5}$.

internationally for various model domains after redesigning the emission control factors and the experimental model runs. We are trying to use RSM-VAT to assist evaluation of the emission control impacts on air quality among the cites/areas within Yangtze River Delta (YRD), China. The current YRD experiment results for two months (January and August, 2010) can be downloaded from our website of http://www.abacas-dss.com/.

Further improvement on the accuracy of RSM can be achieved by additional sampling model runs, although the added computational cost should also be considered. The RSM methodology and the visualization tool demonstrated in this work represent a step forward to effective and efficient science-based decision support.

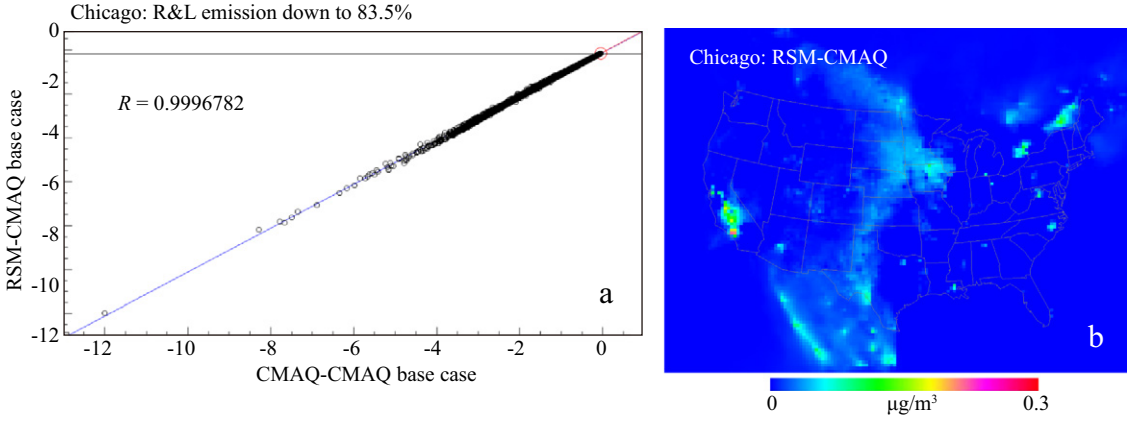


Fig. 10 – Comparison of the $PM_{2.5}$ (μg/m^3) concentrations simulated by CMAQ and predicted by RSM when reducing emission sources inside and outside Chicago down to 83.5%. (a) X–Y scatterplot for CMAQ base case concentration reduction of RSM *vs.* CMAQ, and (b) the concentration difference (RSM subtract CMAQ) maps for CMAQ and RSM predicted values.

Acknowledgments

Financial and data support for this work is provided by the U.S. Environmental Protection Agency (No. GS-10F-0205T). This work is also partly supported by the funding of Guangdong Provincial Key Laboratory of Atmospheric Environment and Pollution Control (No. h2xjD612004III), the funding of State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex (No. SCAPC201308), and the project of Atmospheric Haze Collaboration Control Technology Design (No. XDB05030400) from Chinese Academy of Sciences.

REFERENCES

Ashok, A., Lee, I.H., Arunachalam, S., Waitz, I.A., Yim, S.H.L., Barrett, S.R.H., 2013. Development of a response surface model of aviation's air quality impacts in the United States. Atmos. Environ. 77, 445–452.

Byun, D., Schere, K.L., 2006. Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. Appl. Mech. Rev. 59 (1–6), 51–76.

Eldrandaly, K.A., Abu-Zaid, M.S., 2011. Comparison of six GIS-based spatial interpolation methods for estimating air temperature in Western Saudi Arabia. J. Environ. Inform. 18 (1), 38–45.

Fraser, S., Marceau, D.J., Visscher, A.D., Roth, S.H., 2013. Estimating exposure by loose-coupling an air dispersion model and a geospatial information system. J. Environ. Inform. 21 (2), 84–92.

Ginsbourger, D., Rosspopoff, B., Pirot, G., Durrande, N., Renard, P., 2013. Distance-based kriging relying on proxy simulations for inverse conditioning. Adv. Water Resour. 52, 275–291.

Hirabayashi, S., Kroll, C.N., Nowak, D.J., 2011. Component-based development and sensitivity analyses of an air pollutant dry deposition model. Environ. Model Softw. 26 (6), 804–816.

Jia, G., Taflanidis, A.A., 2013. Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment. Comput. Methods Appl. Mech. Eng. 261–262, 24–38.

Kleijnen, J.P.C., 2009. Kriging metamodeling in simulation: a review. Eur. J. Oper. Res. 192 (3), 707–716.

Lao, Y.W., Zhu, Y., Jang, C., Lin, C.J., Xing, J., Chen, Z.R., Xie, J.P., et al., 2012. Research and development of regional air pollution control decision support tool based on response surface model. Acta Sci. Circum. 32 (8), 1913–1922.

Li, R., Sudjianto, A., 2005. Analysis of computer experiments using penalized likelihood in Gaussian kriging models. Technometrics 47 (2), 111–120.

Lim, L.L., Hughes, S.J., Hellawell, E.E., 2005. Integrated decision support system for urban air quality assessment. Environ. Model Softw. 20 (7), 947–954.

Mohammadi, H., Seifi, A., Foroud, T., 2012. A robust Kriging model for predicting accumulative outflow from a mature reservoir considering a new horizontal well. J. Petrol. Sci. Eng. 82–83, 113–119.

Peng, X.L., Yin, H., Li, R., Fang, K.T., 2006. The application of Kriging and empirical Kriging based on the variables selected by SCAD. Anal. Chim. Acta. 578 (2), 178–185.

U.S.EPA, 2005. Clean air interstate rule emissions inventory technical support document. Available at:, http://www.epa.gov/interstateairquality/pdfs/finaltech01.pdf (Date accessed: 4 March 2005).

U.S.EPA, 2006. Technical support document for the proposed PM NAAQS rule: response surface modeling. Available at:, http://www.epa.gov/ttn/scram/reports/pmnaaqs_tsd_rsm_all_021606.pdf (Date accessed: February, 2006).

Wang, S.L., Gao, J., Zhang, Y.C., Zhang, J.Q., Cha, F., Wang, T., et al., 2014a. Impact of emission control on regional air quality: an observational study of air pollutants before, during and after the Beijing Olympic Games. J. Environ. Sci. 26 (1), 175–180.

Wang, S.X., Xing, J., Jang, C., Zhu, Y., Fu, J.S., Hao, J.M., 2011. Impact assessment of ammonia emissions on inorganic aerosols in East China using response surface modeling technique. Environ. Sci. Technol. 45 (21), 9293–9300.

Wang, S.X., Xing, J., Zhao, B., Jang, C., Hao, J.M., 2014b. Effectiveness of national air pollution control policies on the air quality in metropolitan areas of China. J. Environ. Sci. 26 (1), 13–22.

Xiao, M., Breitkopf, P., Filomeno Coelho, R., Knopf-Lenoir, C., Sidorkiewicz, M., Villon, P., 2009. Model reduction by CPOD and Kriging. Struct. Multidiscip. Optim. 41 (4), 555–574.

Xing, J., Wang, S.X., Jang, C., Zhu, Y., Hao, J.M., 2011. Nonlinear response of ozone to precursor emission changes in China: a modeling study using response surface methodology. Atmos. Chem. Phys. 11 (10), 5027–5044.

Zhuang, X., Wang, Y., He, H., Liu, J., Wang, X., Zhu, T., et al., 2014. Haze insights and mitigation in China: an overview. J. Environ. Sci. 26 (1), 2–12.