Atmospheric Environment 98 (2014) 283–289

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Two reduced form air quality modeling techniques for rapidly calculating pollutant mitigation potential across many sources, locations and precursor emission types

Kristen M. Foley*, Sergey L. Napelenok, Carey Jang, Sharon Phillips, Bryan J. Hubbell, Charles M. Fulcher

United States Environmental Protection Agency, Research Triangle Park, NC 27711, USA

HIGHLIGHTS

- Two reduced form versions of an air quality model are compared and evaluated.
- Both methods are used to estimate impacts of emission reductions from power plants.
- Important differences between the two methods are discussed.

ARTICLE INFO

Article history:
Received 2 June 2014
Received in revised form
14 August 2014
Accepted 19 August 2014
Available online 23 August 2014

Keywords:
CMAQ-DDM
Response surface model
Emission controls
$PM_{2.5}$
Ozone

ABSTRACT

Due to the computational cost of running regional-scale numerical air quality models, reduced form models (RFM) have been proposed as computationally efficient simulation tools for characterizing the pollutant response to many different types of emission reductions. The U.S. Environmental Protection Agency has developed two types of reduced form models based upon simulations of the Community Multiscale Air Quality (CMAQ) modeling system. One is based on statistical response surface modeling (RSM) techniques using a multidimensional kriging approach to approximate the nonlinear chemical and physical processes. The second approach is based on using sensitivity coefficients estimated with the Decoupled Direct Method in 3 dimensions (CMAQ-DDM-3D) in a Taylor series approximation for the nonlinear response of the pollutant concentrations to changes in emissions from specific sectors and locations. Both types of reduced form models are used to estimate the changes in O_3 and $PM_{2.5}$ across space associated with emission reductions of NO_x and SO_2 from power plants and other sectors in the eastern United States. This study provides a direct comparison of the RSM- and DDM-3D-based tools in terms of: computational cost, model performance against brute force runs, and model response to changes in emission inputs. For O_3, the DDM-3D RFM had slightly better performance on average for low to moderate emission cuts compared to the kriging-based RSM, but over-predicted O_3 disbenefits from cuts to mobile source NO_x in very urban areas. The RSM approach required more up-front computational cost and produced some spurious O_3 increases in response to reductions in power plant emissions. However the RSM provided more accurate predictions for $PM_{2.5}$ and for predictions of very large emission cuts (*e.g.* −60 to −90%). This comparison indicates that there are some important differences in the output of the two approaches that should be taken under consideration when interpreting results for a given application.

Published by Elsevier Ltd.

1. Introduction

Over the last few decades, air quality has improved substantially in most of the United States, Europe, and elsewhere. In the US, ambient air concentrations of regulated trace pollutants such as ozone (O_3), particulate matter (PM), carbon monoxide, nitrogen dioxide, sulfur dioxide (SO_2), and lead have all decreased (US-EPA,

* Corresponding author.
E-mail address: foley.kristen@epa.gov (K.M. Foley).

http://dx.doi.org/10.1016/j.atmosenv.2014.08.046
1352-2310/Published by Elsevier Ltd.

2012). This was possible from better scientific understanding of the physical and chemical processes governing the formation of these pollutants in the troposphere and the subsequent enactment of air pollution control policies of emission reductions and the emergence of cleaner technology. Still, millions of people in the US are routinely exposed to air pollutant concentrations above the levels that have been shown to be associated with increased risks for cardiovascular and respiratory disease (Lepeule et al., 2012; Atkinson, 2013). Furthermore, elevated ambient air pollutant concentrations are also found to be harmful to agricultural crops (Booker et al., 2009) and to have a dramatic impact on visibility (Liu et al., 2012). Thus, further improvements in ambient air quality would be clearly beneficial to both human health and welfare.

Sources for additional incremental reductions in air pollutant concentrations over what has already been achieved can be difficult to identify and prioritize. For this reason, numerical air quality models have been used as a testbed for quantifying the impacts of prescribed emission reductions to arrive at physically and economically feasible strategies for reducing ambient concentrations of O_3 and PM. One of these is the Community Multiscale Air Quality (CMAQ) model, which is widely employed by scientific institutes and regulatory agencies (www.cmaq-model.org). Traditionally, such models simulate the "base" level of pollution under current levels of emissions, and, then, repeat the simulation for emission levels proposed under a specific control strategy. The difference (or ratio for attainment demonstrations) between the "brute force" simulations at any receptor of interest is then assumed to be the impact of the control strategy. However, it is often extremely computationally intensive to apply state-of-the-science air quality models, such as CMAQ, to large geographic regions and long pollution episodes. Therefore, it is often resource prohibitive to test a large number of control strategies. For this reason, various source sensitivity and source apportionment techniques have been developed to reduce the size of the numerical problem (Cohan and Napelenok, 2011). Two of these techniques are developed and applied here to demonstrate their use for hypothetical emission reductions of a number of different types of sources over the eastern United States. The first is the response surface model approach (RSM; US-EPA, 2006), which aims to describe, using multidimensional kriging across the space defined by the range of potential emission reductions, the full pollutant concentration response to changes in emissions as a function of preselected control variables from a sufficiently large set of brute force simulations. The second is the Taylor series expansion of the concentration/emission function based on the model sensitivity coefficients calculated by the Decoupled Direct Method in three dimensions (DDM-3D). Both techniques are applied for a domain over the eastern United States, in order to quantify the impact on O_3 and $PM_{2.5}$ levels from hypothetical reductions to state specific emissions from electric generating units (EGUs) as well as domain-wide emission cuts to mobile, area, and other sources.

2. Method

All simulations for both methods were conducted using CMAQ version 4.7.1 (Foley et al., 2010) for August 2005 over the eastern United States at 12 km horizontal grid resolution and 14 vertical layers. The standard model configuration was used including the carbon bond version 2005 chemical mechanism (Sarwar et al., 2008). Boundary conditions were derived from another simulation over a larger encompassing 36 km domain. Boundary conditions for the 36 km simulation were obtained from a 2005 global GEOS-Chem simulation (http://wiki.seas.harvard.edu/geos-chem/). Emission inputs were developed using the Sparse Matrix Operator Kernel Emissions (SMOKE) processing system version 2.4 (http://www.smoke-model.org) based on the 2005 National Emissions Inventory and included year-specific data from the Continuous Emission Monitoring measurements from combustion and industrial processes. Emission summaries from the 2005 NEI for the sectors and pollutants of interest in this study are provided in Table S1 and Fig. S1 in the supplemental material. Meteorological inputs were processed using the Weather Research and Forecasting (WRF) model version 3.0 (Skamarock et al., 2008). Additional information regarding the modeling platform can be found in US-EPA (2013).

2.1. Development of a RSM-based reduced form model for CMAQ

Previous studies have demonstrated the use of response surface modeling techniques for the CMAQ model to estimate the human health benefits of reducing emissions from different sources in nine urban areas in the US (Fann et al., 2009), to compare the effectiveness of local and regional NO_x and VOC controls in three megacities in China (Xing et al., 2011), and to quantify the contribution of NH_3 emissions to fine particles in heavily developed regions of China (Wang et al., 2011). The response surface model developed here was used to estimate air quality impacts associated with emission reductions of NO_x and SO_2 from EGUs and other pollutants from other sectors in the eastern United States. The RSM was designed to provide state-specific information on the impacts of statewide shifts in emissions due to EGU policies, such as trading programs, and can also be used as a screening tool for comparing emission control strategies for attainment demonstrations.

The first step in developing an RSM is the selection of the source/emission factors that are of interest for a given application, i.e., experimental design. Here the RSM was developed to evaluate shifts in NO_x and SO_2 emissions at the state level for the EGU sector, while accounting for interactions with regional emissions of NO_x, SO_2, VOC, and NH_3 emissions from mobile, area (*e.g.*, agriculture, residential heating, dust), and non-EGU point sources (*e.g.*, industrial boilers, cement kilns). Although understanding state-specific EGU emission shifts was the focus of this study, it was not computationally feasible to designate separate emission factors for each of the states within the modeling domain. As a solution, the 42 states (including D.C.) are grouped into 15 clusters of one to three states, in order to reduce the total number of brute force simulations needed to create the RSM. States were grouped *a priori* based on zero-out simulations and SO_2 tracer model experiments such that the impact from emission cuts in any state would have minimal influence on the other states in its group. For example Group 1 consists of AR, PA and MT (the full group listings are provided in Fig. S2 in the supplemental material). In this way, although the emission factors for EGU emissions are based on state groups, it is still possible to estimate the air quality impacts of emission cuts in a single state. The final RSM experimental design consisted of 36 emission factors as shown in Table 1, mainly aimed at predicting the response of $PM_{2.5}$.

The RSM was designed to model changes in the predetermined emission factors ranging from 0% to 120% of base emission levels. This means the design space of interest was the 36 dimensional hypercube $[0, 1.2]^{36} \subset \mathbb{R}^{36}$. A two stage sampling design was used to

Table 1
Source/emission factors used in the design of the CMAQ RSM.

Sector	Geographical source	Pollutant	# emission factors
EGU	15 state cluster groups	NO_x, SO_2	30
Mobile	Domain wide	NO_x, VOC, NH_3	3
Area	Domain Wide	NH_3	1
Non-EGU	Domain Wide	NO_x, SO_2	2

efficiently sample the design space to minimize the number of brute force CMAQ simulations needed to estimate the final RSM. In the first stage, a Latin Hypercube design was used to select 141 CMAQ simulations (a base case run plus 140 control runs). Based on evaluation of an RSM using these simulations, and past experience with design of a CMAQ-based RSM, 72 additional simulations were selected to better estimate the lower bounds (50%–100% cuts) of NO_x and NH_3 source categories, to better estimate 25% and 75% cuts to SO_2 and VOC source categories, and to capture the upper bound of the design space of 20% growth. A graphical depiction of the emission factors (ranging from 0 to 1.2) used for the 213 brute force CMAQ simulations is shown in Supplemental Fig. S3.

With the brute force simulations complete, a statistical model was used to estimate the nonlinear relationships between a given pollutant and the set of 36 emission factors for a given time period. The statistical model is estimated using maximum likelihood estimation techniques using PROC MIXED in SAS based on monthly average CMAQ output for O_3 and $PM_{2.5}$. The RSM estimate for the concentration of species i at location $\boldsymbol{s}$, based on fractional emission levels defined by a vector$\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{36}) \in [0, 1.2]^{36}$, is:

$$C_{i,\boldsymbol{\phi}}(\boldsymbol{s}) = \widehat{\beta}_i + w_{i,\boldsymbol{\phi}}\left(y_i(\boldsymbol{s}) - \widehat{\beta}_i\right), \quad (1)$$

where $\widehat{\beta}_i$ is the estimated mean response across all grid cells, $w_{i,\boldsymbol{\phi}}$ is a transposed vector of weights and $y_i(\boldsymbol{s})$ is the vector of pollutant concentrations from the 213 RSM simulations for location $\boldsymbol{s}$. The weights, $w_{i,\boldsymbol{\phi}}$, are determined by estimating a correlation model to quantify the correlation between the pollutant field under emission perturbation $\boldsymbol{\phi}$ and the pollutant fields from the 213 brute force runs. Here a two parameter Gaussian correlation model is used: $R(\boldsymbol{\phi}_j, \boldsymbol{\phi}_k) = \sigma^2 e^{-(d_{jk}/\theta)}$, where d_{jk} is the Euclidean distance between the vectors of perturbations, $\boldsymbol{\phi}_j$ and $\boldsymbol{\phi}_k$, associated with design points (*e.g.* runs) j and k, and σ^2 and θ are estimated during model fitting based on pollutant concentrations from all grid cells. As a result, the brute force simulation that is closest, in terms of Euclidean distance within the design space, to the new emission scenario of interest receives the greatest weight. For further details on the development and estimation of the statistical model see US-EPA, (2006). Thus, using the formula in Equation (1), it was possible to estimate the O_3 and $PM_{2.5}$ levels under emission cuts to any combination of the 36 emission factors without the computational burden of re-running the entire CMAQ model.

2.2. Development of a DDM-3D-based reduced form model for CMAQ

While the RSM method described in the previous section has been shown to be an efficient and accurate approach for exploring a large set of emission control scenarios (*e.g.* Wang et al., 2011; Xing et al., 2011), it still required a great deal of up-front computational cost to develop for this application. A reduced form version of CMAQ based on the sensitivity coefficients calculated by DDM-3D was also developed to determine if such an approach would be competitive with the RSM in terms of accuracy and computational cost. CMAQ-DDM-3D model version 4.7.1 was used in this application, which shares the same chemical and physical modules as the standard CMAQ model release used for the development of RSM. The DDM-3D implementation is described in detail elsewhere (Napelenok et al., 2008). The DDM-3D-based reduced form model was constructed with particular effort to maintain consistency with the RSM approach in order to facilitate comparison between the two methods. Therefore, the state clustering approach for EGU emissions was maintained identically as described above. First- and second-order DDM-3D sensitivity coefficients represent the slope and the curvature of the change in the pollutant concentration for a given species with respect to changes to the input parameters of interest. First-order sensitivity coefficients of $PM_{2.5}$ and first- and second-order sensitivity coefficients of O_3 were calculated to elevated EGU emissions of NO_x and SO_2 separately for each of the 15 state groups. Additionally, first-order sensitivity coefficients of $PM_{2.5}$ and first- and second-order sensitivity coefficients of O_3 were also calculated to domain-wide emissions of non-EGU point sources of NO_x and SO_2, to domain-wide emissions of area sources of ammonia (NH_3), and to domain-wide emissions of mobile sources of NO_x, VOCs, and NH_3. All 36 possible first- and second-order pairs of sensitivities of O_3 to the emission factors in Table 1 were examined. Many of the second-order sensitivities were found to be close to zero and were dropped from the application. Second-order "cross" sensitivities (*e.g.*, sensitivity to simultaneous perturbation of both EGU SO_2 and EGU NO_x) were also not considered to reduce the size of the computational burden. At the time of simulation, second-order sensitivity capability for PM species was not yet available in the CMAQ model. In total, 36 first order sensitivities were used for $PM_{2.5}$ and 36 first-order and 18 second-order sensitivities were used for O_3 for this application. Taylor series expansion was then used to construct a reduced form model of each modeled species as a function of these sensitivity coefficients:

$$C_{i,\boldsymbol{\phi}}(\boldsymbol{s},t) = C_{i,0}(\boldsymbol{s},t) + \sum_{j=1}^{36}\left(\Delta\varepsilon_j \cdot S_{i,j}^{(1)}(\boldsymbol{s},t) + \frac{1}{2}\Delta\varepsilon_j^2 \cdot S_{i,j}^{(2)}(\boldsymbol{s},t)\right), \quad (2)$$

where $C_{i,\boldsymbol{\phi}}(\boldsymbol{s},t)$ is the concentration of species i at location $\mathbf{s}$ and time t under emission perturbation $\boldsymbol{\phi}$, $C_{i,0}(\boldsymbol{s},t)$ is the base model condition, $S_{i,j}^{(.)}(\boldsymbol{s},t)$ is the first- (1) or second-order (2) sensitivity to a perturbation in emission source j, and $\Delta\varepsilon_j = \phi_j - 1$ (*e.g.*, for a 20% cut in emission source j, $\phi_j = .8$, and $\Delta\varepsilon_j = -0.2$).

The DDM-3D-based model in Equation (2) provided hourly concentration information. While the RSM approach could be extended to also provide hourly output by estimating separate models for each time step, the model in this study is for monthly-average results since this was the $PM_{2.5}$ metric of interest. For the remainder of the discussion, the hourly DDM-3D concentrations for the month are averaged in order to be consistent with the RSM output. However, the flexibility of hourly output from either model could be used to provide additional information on the impacts of emission cuts on different air quality metrics such as maximum 8 h average or peak daily ozone (*e.g.*, Simon et al., 2012).

Table 2
Evaluation of RSM and DDM predictions against 10 out-of-sample brute force simulations. Summary statistics are calculated based on model output at all grid cells falling within the US boundary for all of the out-of-sample simulations (n = 10*number US grid cells = 267140).

	$PM_{2.5}$ (μg/m^3)		O_3 (ppb)	
	RSM	1st order DDM	RSM	2nd order DDM
Min bias	−1.11	−1.49	−2.81	−7.76
25th % bias	−0.01	0.11	−0.21	−0.011
Median bias	0.01	0.30	−0.03	0.00
75th % bias	0.04	0.57	0.14	0.26
Max bias	1.31	3.91	3.57	11.30
RMSE	0.09	0.52	0.46	0.62
R^2	0.999	0.95	0.99	0.96

2.3. Computational cost and out of sample validation

The computational cost of the RSM and DDM-3D based tools is compared for a one month simulation. Using 72 Intel Xeon x5550 processors, a single base CMAQ simulation had a run time

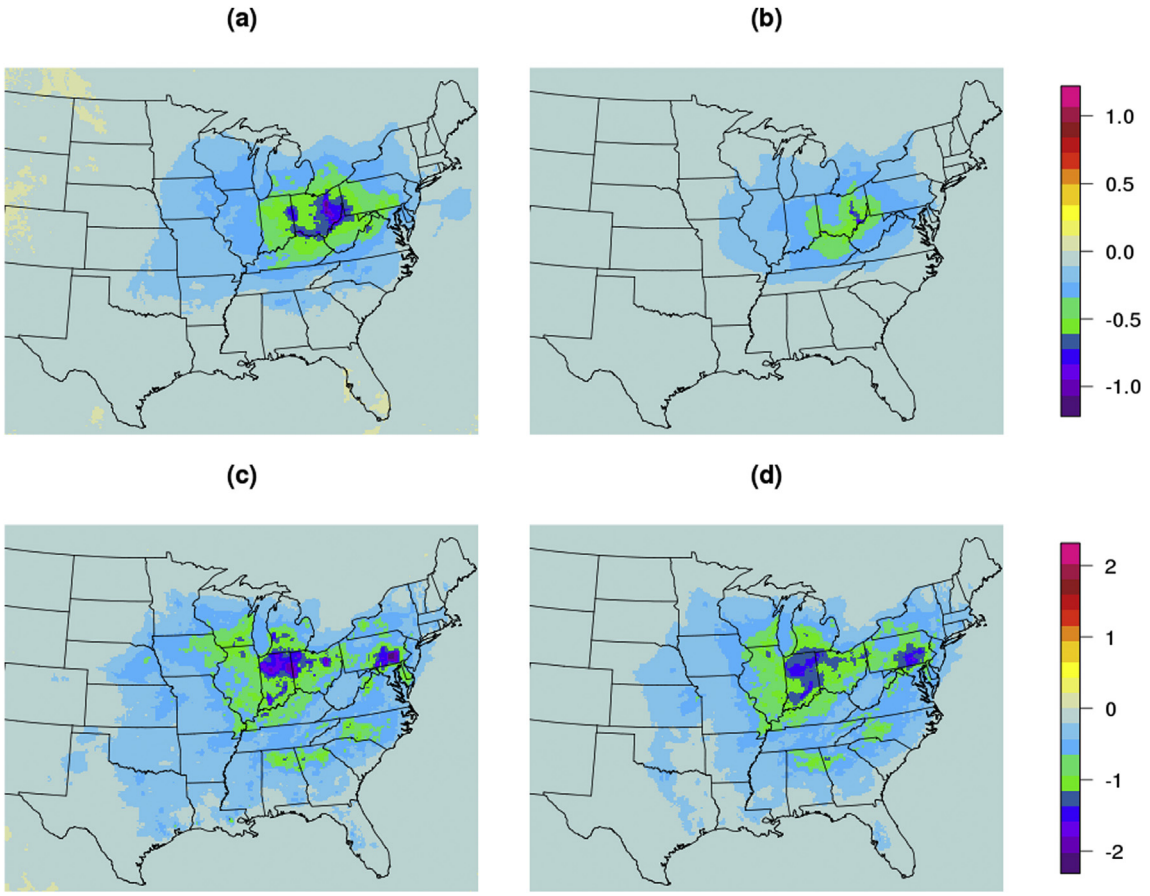

Fig. 1. Change in monthly-average $PM_{2.5}$ ($\mu g/m^3$) after (a) RSM: 50% cut to EGU emissions in OH (b) DDM: 50% cut to EGU emissions in OH (c) RSM: 50% cut to NH_3 emissions from agriculture sources (d) DDM: 50% cut to NH_3 emissions from agriculture sources.

requirement of 0.42 days. The 213 simulations used to build the RSM required 89.5 days. All 90 DDM-3D sensitivities were calculated in a single model simulation requiring 10 days of run time, offering a clear advantage over the RSM approach in terms of the initial computational cost. Once the brute force and DDM-3D runs are complete, both tools can provide near-instantaneous estimates of the pollutant response to user-defined changes to the 36 emission factors.

The two methods were evaluated using brute force simulations. A random sample of ten of the available simulations generated by the Latin hypercube sample were withheld from the development of the RSM (see Supplemental Fig. S3), and the RSM- and DDM-3D based models were used to predict pollutant concentrations for these out-of-sample runs. Summary statistics of the prediction error at all grid cells within the US boundary are given in Table 2. The RSM predictions for changes in $PM_{2.5}$ due to emission reductions had very low bias and error, whereas the DDM-based predictions tended to be biased high by 0.1–0.6 $\mu g/m^3$. This implies that the reduced form model will under predict the decrease in $PM_{2.5}$ concentrations resulting from a reduction in emissions. This will be further discussed in Sections 3 and 4.

The bias in predicting O_3 was more similar between the two methods with 75% of the prediction error falling within about ±0.25 ppb for both methods. The DDM-3D-based model did produce a few very large errors (less than 0.1% of the grid cells had an absolute error greater than 5 ppb). This is because the Taylor-expansion approach was not expected to accurately estimate very large changes in emission inputs that were used in these out-of-sample simulations (*e.g.*, cuts up to 70–99% of base levels). Previous studies have demonstrated the accuracy of DDM-3D for emission perturbations within ±50% (Cohan et al., 2005; Simon et al., 2012). This behavior was also evident in the validation analysis, where errors in the DDM-3D-based predictions for O_3 tended to increase with increasingly large cuts to NO_x emissions from mobile sources. For cuts within ±50% the bias in the second order DDM-3D-based predictions actually tended to be less than the RSM prediction bias (see Supplemental Fig. S4).

3. Application: comparison of local versus regional emission controls in the eastern US

The advantage of developing a reduced form version of the CMAQ model is the ability to rapidly assess the impacts of many different emission reduction strategies. For demonstration of this capability, the impact of the differences between the RSM- and DDM-3D-based methods shown in Section 2 on a specific application were investigated. Specifically, the following questions were explored: How do state-specific EGU controls compare to regional controls (*e.g.*, mobile or agricultural emissions) for a given state? How do state-specific EGU controls impact states adjacent or near the state of interest?

Figs. 1 and 2 show example output from the RSM- and DDM-3D-based reduced form models. Fig. 1 shows the predicted change in $PM_{2.5}$ resulting from a 50% cut to NO_x and SO_2 emissions from EGUs in the State of Ohio (OH) and the change after a domain-wide 50% cut in NH_3 emissions from area sources. Both methods show that

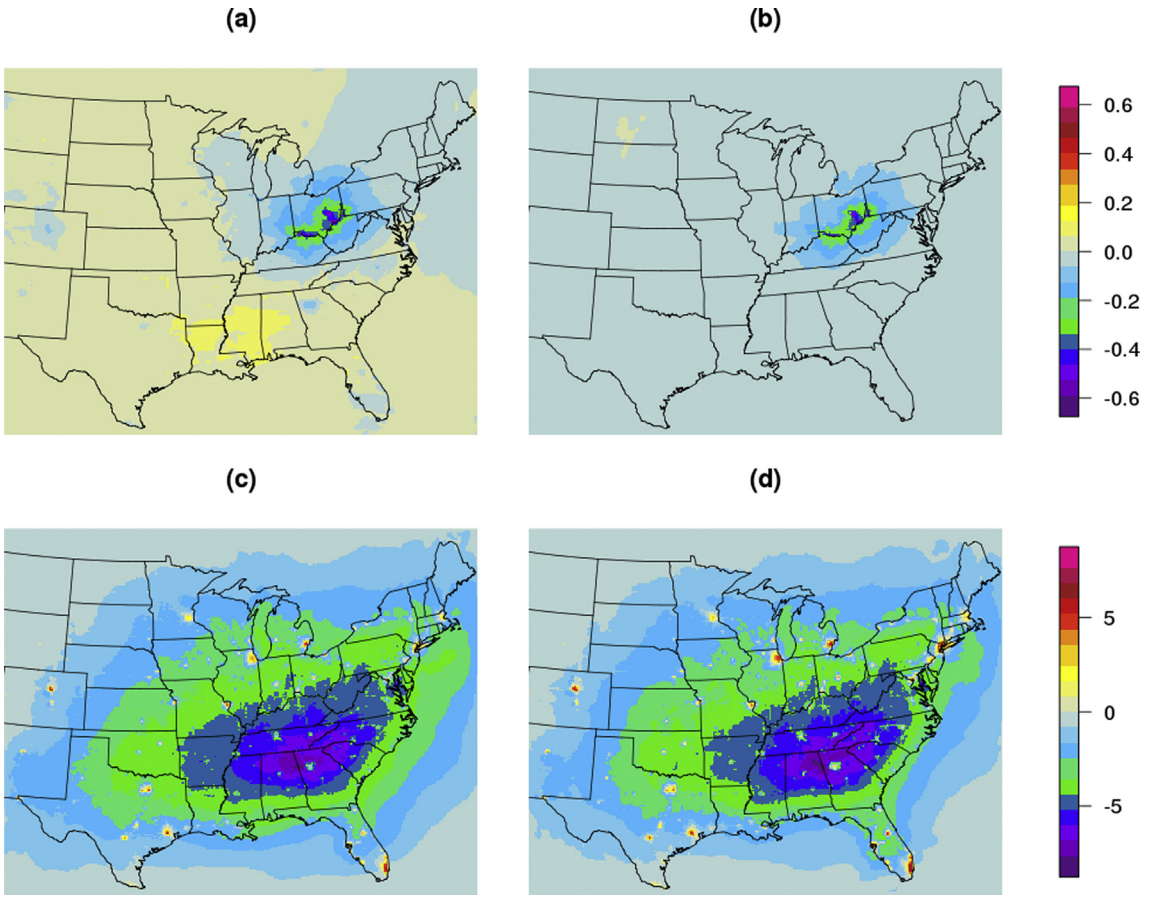

Fig. 2. Change in monthly-average O_3 (ppb) after (a) RSM: 50% cut to EGU emissions in OH (b) DDM: 50% cut to EGU emissions in OH (c) RSM: 50% cut to NO_x emissions from mobile sources (d) DDM: 50% cut to NO_x emissions from mobile sources.

the OH EGU controls impact surrounding states in the region. Consistent with the evaluation in Section 2, the DDM-3D-based model predicts less of a decrease in $PM_{2.5}$ for both scenarios, compared to the RSM predictions. The domain-wide average change in $PM_{2.5}$ concentrations after the emission cut to OH EGU emissions was -0.05 μg/m^3 using the DDM-3D approach compared to -0.08 μg/m^3 using the RSM. The maximum difference between the two methods was in a grid cell in OH with a DDM-3D based predicted change of -0.57 μg/m^3 and a RSM prediction of -1.01 μg/m^3. For the emission cut to NH_3 emissions, the DDM-3D approach predicted a domain-wide average change in $PM_{2.5}$ of -0.20 μg/m^3 compared to an average change of -0.22 μg/m^3 using the RSM. The maximum difference between the two methods was in a grid cell at the border of OH an IN that had a DDM-3D based predicted change of -0.95 μg/m^3 and a RSM prediction of -2.02 μg/m^3.

Fig. 2 shows the predicted change in O_3 after a 50% cut to OH EGU emissions and the change after a domain-wide 50% cut to NO_x emissions from mobile sources. The main difference in the O_3 response to EGU emission cuts is that the RSM approach predicts small O_3 increases away from the actual emission change. Such increases are unrealistic and reflect a spurious result in the statistical model that was not seen in the underlying brute force simulations. Additional brute force simulations would be needed to better capture the nonlinear relationship between O_3 and emission inputs to resolve this issue. Separate experimental designs may need to be considered for O_3 and $PM_{2.5}$ for future complex multi-region applications.

The O_3 response to a 50% cut to NO_x emissions from mobile sources is similar for both methods for most of the region, with the exception of urban areas. The DDM-3D-based model predicts larger increases in O_3 in very urban areas compared to the RSM approach. These increases represent O_3 disbenefits due to decreased O_3 titration with decreasing NO_x emissions. Evaluation against brute force simulations suggests that the DDM-3D predictions are 1.5–2 ppb too high in these locations while the disbenefits predicted by the RSM approach are more in-line with the underlying brute force results.

A summary of the pollutant response to 50% emission cuts in different sectors for seven states is shown in Fig. 3. The seven states were selected as an illustration and had the largest state-wide average $PM_{2.5}$ concentrations in the domain in the base simulation. Results for the other states in the domain are provided in Tables S2 and S3 in the supplemental material. The barplots depict population-weighted state-wide average pollutant changes. The weighting, based on census tract population data from the 2000 census (see Supplemental Fig. S5), was used to focus on the benefits of emission controls on human exposure levels, the emphasis of this application.

For $PM_{2.5}$ controls four sectors were compared: (1) NO_x and SO_2 emissions from EGUs in the listed state (Local EGU); (2) NO_x and SO_2 emissions from EGUs in all other states in the domain (Outside EGU); (3) domain-wide NO_x emissions from mobile sources (Mobile NO_x); (4) domain wide NH_3 emissions from area sources (Area NH_3). Due to the state-clustering used in the design of this RSM it was not feasible to estimate the response in a given state to

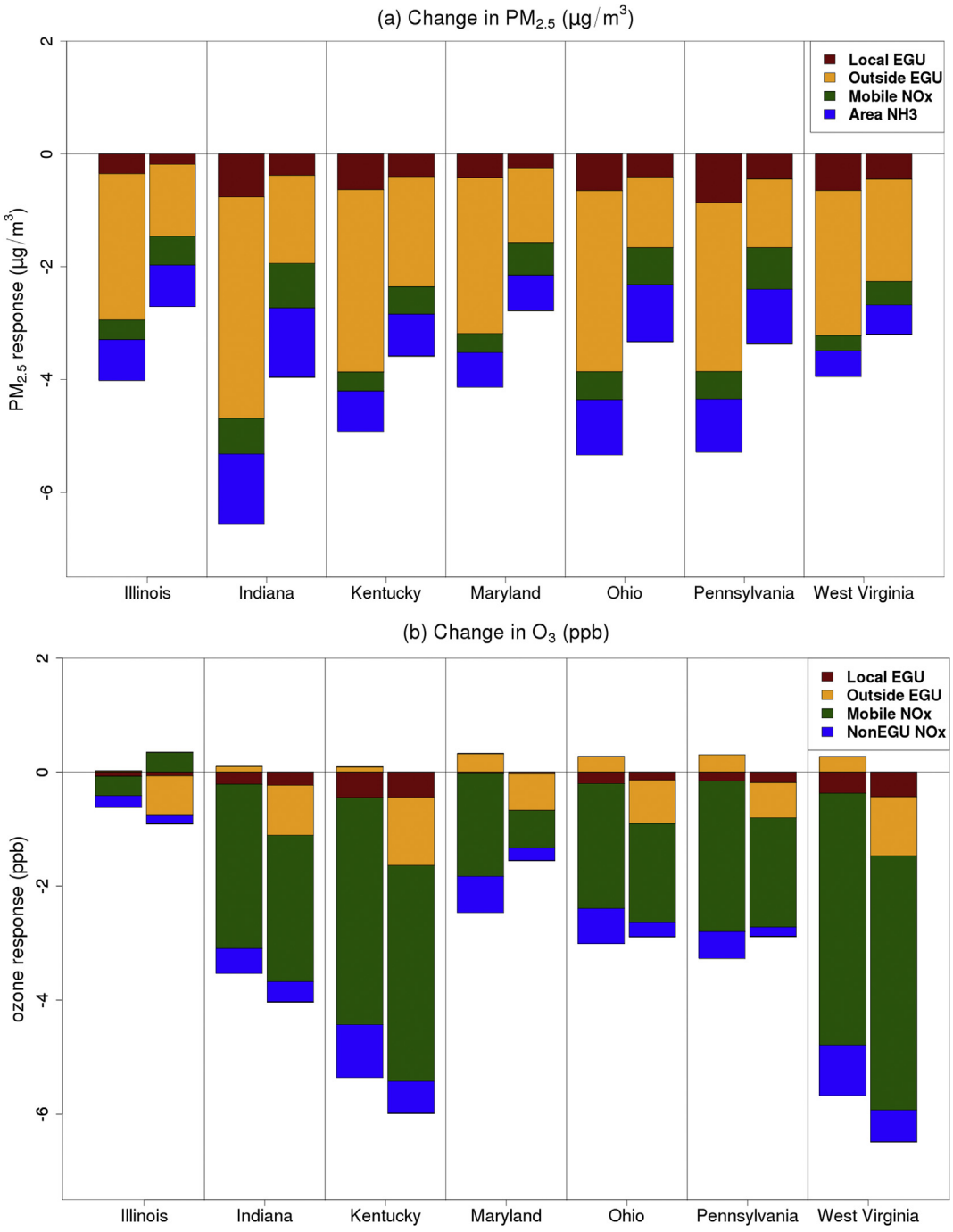

Fig. 3. Population-weighted state-wide average change in (a) $PM_{2.5}$ and (b) O_3 due to 50% emission cuts in different sectors. Estimates based on the RSM are shown on the left for each state; DDM-3D-based results are on the right. Note that results for "Maryland" are estimates for all grid cells falling within Maryland, Delaware and D.C. (Note that Fig. S6 in the Supplemental Material provides a version of these figures without the population-weighting.)

simultaneous EGU cuts in all surrounding states. Instead, to approximate the impact from outside EGU controls, the impacts on a given state to cuts in the other states in the domain were estimated separately and then summed together. This approximation is likely slightly larger in magnitude than a simultaneous shift due to nonlinearities in the relationship between $PM_{2.5}$ (or O_3) and emissions. However, this information is still valuable for understanding the role of outside emission sources on local pollutant levels.

The DDM-3D-based estimates for the population-weighted change in $PM_{2.5}$ due to changes in region-wide Area NH_3 were found to be very similar to the RSM results. This shows that some of the spatial differences appearing in Fig. 1(c) and (d) are less important when focus is on highly-populated grid cells. Consistent with the previous evaluation, the DDM-3D-based results for both Local and Outside EGU controls were smaller in magnitude compared to the RSM results, providing estimates of control impacts that are too conservative. This bias could lead policy makers

to under-estimate the efficacy of EGU controls compared to the other regional controls.

For O_3, controls for Local EGU, Outside EGU, Mobile NO_x, as well as NO_x emissions from non-EGU sources (NonEGU NO_x) were compared. Both methods predicted the largest decreases in O_3 levels result from cuts to mobile NO_x emissions, rather than EGU controls. This reflects the fact that a 50% cut to precursor NO_x emissions from mobile sources represents a much larger emission reduction in terms of total tonnage, compared to a 50% reduction in NO_x emissions from EGUs in this domain (see Supplemental Table S1). Estimates for the impacts of Local EGU controls were very similar for the RSM- and DDM-3D-based approaches. However, for Outside EGU controls, the RSM-based model predicted O_3 disbenefits in most of the states due to the spurious O_3 increases mentioned earlier. In contrast, the DDM-3D-based method predicted Outside EGU controls to lead to O_3 decreases of 0.6–1.2 ppb.

This set of emission reductions was chosen for illustrative purposes. Both RSM and DDM tools offer the flexibility to explore a much wider range of reduction scenarios in terms of the size and source of the reduction. Here, it was shown that there are some important differences in the output of the two methods that should be taken under consideration when interpreting results for a given application.

4. Conclusions

The statistical Response Surface Model and the sensitivity-based DDM-3D approach for approximating CMAQ output were shown to have different advantages and disadvantages for estimating the impact of emission reductions from different sources and locations. The RSM required roughly 9 times the up-front computational resources, but provided more accurate predictions for very large emission cuts (*e.g.* −70 to −99%). The DDM-3D approach is not designed to handle such large emission changes. For $PM_{2.5}$, the DDM-3D-based results are biased high (under-responsive for emission reductions) compared to the RSM results. Second order DDM for $PM_{2.5}$ is currently scheduled for release in the next version of the CMAQ model (v5.0.2) and should address some of the bias seen in this study. Additional cross-sensitivities could also be included in the DDM formulation to better capture potential interactions between NO_x, SO_2 and NH_3 across multiple sources, although this would necessarily increase the computational cost. The RSM approach is designed to model such interactions through careful selection of the underlying brute force simulations.

For O_3, DDM-3D had slightly better performance on average for low to moderate emission cuts but over-predicted O_3 disbenefits from cuts to mobile source NO_x in very urban areas. The RSM produced some spurious O_3 increases far away from the location of the EGU emission cuts. Separate experimental designs may need to be developed for O_3 and $PM_{2.5}$ for future applications involving many different source regions.

Reduced form methods such as the RSM- and DDM-3D-based approaches offer computationally efficient screening methods for rapidly comparing a large number of emission reduction options while still leveraging the complexity of the physical and chemical processes represented within the underlying state of the science air quality model. Information from these types of tools can be used to then select a smaller set of more refined brute force simulations to run for an attainment demonstration or other regulatory modeling needs. Future work in this area includes extending the RSM and DDM-3D methods to estimating single source impacts, *e.g.*, impact of a single power plant for permitting purposes.

Disclaimer

Although this work was reviewed by EPA and approved for publication, it may not necessarily reflect official Agency policy.

Acknowledgments

The authors would like to thank Christian Hogrefe and Rohit Mathur from the U.S. EPA as well as the two anonymous reviewers for their constructive and thoughtful comments on this analysis which greatly improved the final manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.atmosenv.2014.08.046.

References

Atkinson, R.W., 2013. Long-term exposure to outdoor air pollution and incidence of cardiovascular diseases (vol 24, pg 44, 2013). Epidemiology 24 (2), 339.

Booker, F.L., Muntifering, R., McGrath, M., Burkey, K.O., Decoteau, D., Fiscus, E.L., Manning, W., Krupa, S., Chappelka, A., Grantz, D.A., 2009. The ozone component of global change: potential effects on agricultural and horticultural plant yield, product quality and interactions with invasive species. J. Integr. Plant Biol. 51, 337–351.

Cohan, D.S., Hakami, A., Hu, Y.T., Russell, A.G., 2005. Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis. Environ. Sci. Technol. 39 (17), 6739–6748.

Cohan, D.S., Napelenok, S., 2011. Air quality response modeling for decision support. Atmosphere 2, 19.

Fann, N., Fulcher, C., Hubbell, B., 2009. The influence of location, source, and emission type in estimates of the human health benefits of reducing a ton of air pollution. Air Qual. Atmos. Health 2 (3), 169–176.

Foley, K.M., Roselle, S.J., Appel, K.W., Bhave, P.V., Pleim, J.E., Otte, T.L., Mathur, R., Sarwar, G., Young, J.O., Gilliam, R.C., Nolte, C.G., Kelly, J.T., Gilliland, A.B., Bash, J.O., 2010. Incremental testing of the Community Multiscale Air Quality (CMAQ) modeling system version 4.7. Geosci. Model Dev. 3 (1), 205–226.

Lepeule, J., Laden, F., Dockery, D., Schwartz, J., 2012. Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009. Environ. Health Perspect. 120 (7), 965–970.

Liu, M., Tao, J., Chan, C.-Y., Cao, J.-J., Zhang, Z.-S., Zhu, L.-H., Zhang, R.-J., 2012. Regression analyses between recent air quality and visibility changes in megacities at four haze regions in China. Aerosol Air Qual. Res. 12, 1049–1061.

Napelenok, S.L., Cohan, D.S., Odman, M.T., Tonse, S., 2008. Extension and evaluation of sensitivity analysis capabilities in a photochemical model. Environ. Model. Softw. 23 (8), 994–999.

Sarwar, G., Luecken, D., Yarwood, G., Whitten, G.Z., Carter, W.P.L., 2008. Impact of an updated carbon bond mechanism on predictions from the CMAQ modeling system: preliminary assessment. J. Appl. Meteorol. Clim. 47 (1), 3–14.

Simon, H., Baker, K.R., Akhtar, F., Napelenok, S.L., Possiel, N., Wells, B., Timin, B., 2012. A direct sensitivity approach to predict hourly ozone resulting from compliance with the National Ambient Air Quality Standard. Environ. Sci. Technol. 47 (5), 2304–2313.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.Y., Wang, W., Powers, J.G., 2008. A Description of the Advanced Research WRF Version 3. National Center for Atmospheric Research, Boulder, Colorado. NCAR/TN-475.

US-EPA, 2006. Technical Support Document for the Proposed PM NAAQS Rule: Response Surface Modeling. OAQPS, RTP, NC. http://www.epa.gov/scram001/reports/pmnaaqs_tsd_rsm_all_021606.pdf.

US-EPA, 2012. Our Nation's Air Status and Trends Through 2010. EPA-454/R-12-001.

US-EPA, 2013. Air Quality Modelig Technical Support Document: Proposed Tier 3 Emission Standards. A. Q. A. D. Office of Air Quality and Standards. EPA-454/R-13-006.

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

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. Discuss. 11, 5027–5044.