JOURNAL OF ENVIRONMENTAL SCIENCES 29 (2015) 178–188

Available online at www.sciencedirect.com

ScienceDirect

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

JOURNAL OF ENVIRONMENTAL SCIENCES

www.jesc.ac.cn

Design and demonstration of a next-generation air quality attainment assessment system for $PM_{2.5}$ and O_3

Hua Wang[1], *Yun Zhu*[1,2,*], *Carey Jang*[3], *Che-Jen Lin*[1,4], *Shuxiao Wang*[2], *Joshua S. Fu*[5], *Jian Gao*[6], *Shuang Deng*[6], *Junping Xie*[1], *Dian Ding*[1], *Xuezhen Qiu*[1], *Shicheng Long*[1]

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: amy_wanghua@qq.com*
2. *State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, School of Environment, Tsinghua University, Beijing 100084, China*
3. *USEPA/Office of Air Quality Planning & Standards, RTP NC27711, USA*
4. *Department of Civil Engineering, Lamar University, Beaumont, TX 77710-0024, USA*
5. *Department of Civil & Environmental Engineering, University of Tennessee, Knoxville, TN 37996-2010, USA*
6. *Chinese Research Academy of Environmental Sciences, Beijing 100012, China*

ARTICLE INFO

Article history:
Received 11 March 2014
Revised 13 July 2014
Accepted 6 August 2014
Available online 4 February 2015

Keywords:
Air quality standards
Attainment test
Air pollution
GIS
Emission control

ABSTRACT

Due to the increasingly stringent standards, it is important to assess whether the proposed emission reduction will result in ambient concentrations that meet the standards. The Software for Model Attainment Test—Community Edition (SMAT-CE) is developed for demonstrating attainment of air quality standards of O_3 and $PM_{2.5}$. SMAT-CE improves computational efficiency and provides a number of advanced visualization and analytical functionalities on an integrated GIS platform. SMAT-CE incorporates historical measurements of air quality parameters and simulated air pollutant concentrations under a number of emission inventory scenarios to project the level of compliance to air quality standards in a targeted future year. An application case study of the software based on the U.S. National Ambient Air Quality Standards (NAAQS) shows that SMAT-CE is capable of demonstrating the air quality attainment of annual $PM_{2.5}$ and 8-hour O_3 for a proposed emission control policy.

Introduction

First-principle atmospheric models are essential tools that provide scientific data for formulating emission reduction strategy to attain a targeted air quality goal. The Community Multiscale Air Quality (CMAQ) model is one of the atmospheric models widely applied for forecasting air pollutant concentrations under future emission scenarios (Chemel et al., 2014; Shrestha et al., 2009; Xing et al., 2011). However, a first principle model such as CMAQ is complicated and computationally expensive because of the comprehensive chemical transport mechanisms parameterized in the model (Cheng et al., 2007; Gao and Zhang, 2014; Li et al., 2011; Ma et al., 2013; Sanjose et al., 2007; Zhang et al., 2012). In addition, the models typically generate large datasets on a non-conventional operating platform and in a file system that are tedious to manage. This poses a challenge in extracting model analysis for many policy makers who are not accustomed to analyze model data and therefore there is a need to develop software tool with a user-friendly interface to address analytical need of model data for policy making.

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

http://dx.doi.org/10.1016/j.jes.2014.08.023

To facilitate the analysis and visualization of atmospheric model output, we have developed a suite of software packages that support the decision making for integrated air quality management. These packages (available and free download at http://www.abacas-dss.com/) include: the Response Surface Model—Visualization Analysis Tool (RSM-VAT) to provide real-time estimates of air quality concentrations in response to emission reductions from anthropogenic sources, the Software for Model Attainment Test—Community Edition (SMAT-CE) to perform attainment assessments for the National Ambient Air Quality Standard (NAAQS) under various emission reductions, and the Environmental Benefits Mapping and Analysis Program—Community Edition (BenMAP-CE) to evaluate health and economic benefits associated with improved air quality (Fig. 1). The software suite improves the analytical efficiency of atmospheric model data under changed air emission scenarios (RSM-VAT), assists evaluation of emission reduction efficacy in relation to air quality standards (SMAT-CE), and extends air quality assessment to health and cost benefits (BenMAP-CE). In this study, we present the methodology and demonstrate the application of SMAT-CE for the design values of annual $PM_{2.5}$ and 8-hour O_3 in the U.S. The design value (DV) is a statistic of monitored pollutant concentration at a specific location used by the U.S. Environmental Protection Agency (US EPA) to determine whether or not the location attains the NAAQS (US EPA, 2008b). The DV of 8-hour O_3 for each monitored location is calculated as the 3-year average of the fourth highest daily maximum 8-hour O_3 concentration in each year (US EPA, 1997). The DV of annual $PM_{2.5}$ is the annual means averaged over a continuous three year period (US EPA, 2013). The design goal of SMAT-CE is to predict the future design value (DVF) under a given air emission scenario in a future year. Such capability allows policy maker to determine if proposed emission reductions will result in attainment of the NAAQS within a regulatory timeframe.

The regional haze and photochemical smog (mainly O_3) are two of the concerned air pollutants due to their impacts to public health and welfare (Gao et al., 2014; Sampson et al., 2013; Singh et al., 2012; Wang et al., 2014; Wei et al., 2014). Air quality standards have been set to improve the air quality (Chai et al., 2014; US EPA, 2008a, 2012b). As the next generation of USEPA MATS (Modeled Attainment Test Software) (USEPA, 2012a), SMAT-CE is developed to estimate whether targeted air quality can be attained under various emission control scenarios. The objectives of this study are to describe the functional design and calculation principles of SMAT-CE, evaluate the software performance compared to MATS, and demonstrate the application of the software for the attainment of annual $PM_{2.5}$ and 8-hour O_3.

1. Methodology

1.1. Evaluation principles

SMAT-CE statistically estimates the DVF using the monitoring and model data of a base year, and the model data of the targeted future year. The monitoring data represent the air quality of the base year. The model data of the base and future years are used for projecting the design value (DVB) in a base year to DVF. The detailed methodology and model implementation are described as follows.

The DVF near a monitoring site i is estimated as:

$$(\mathrm{DVF})_i = (\mathrm{DVB})_i \times (\mathrm{RRF})_i \quad (1)$$

where, $(DVB)_i$ (ppb or $\mu g/m^3$) is the DVB monitored at site i, $(RRF)_i$ is the dimensionless relative response factor (RRF) calculated for site i using model data, and (DVF)i (ppb or $\mu g/m^3$) is the DVF predicted at site i.

The DVs for both the monitoring sites (point estimates of DV) and for each grid cell (spatial field estimates of DV) are required to estimate DVFs. Several methodologies can be applied to calculate the DVB: (1) the design value of the designated design value period (*i.e.*, 2006–2008), (2) the design value of the design value period that straddles the baseline inventory year (*e.g.*, the 2006–2008 design value period for a base year of 2007), and (3) averaging the design values of three different periods which include the base year (*e.g.*, the 2005–2007, 2006–2008, and 2007–2009 design value periods for a base year of 2007).

For point estimates, DVBs are calculated from observational data using the third methodology mentioned above. We examined the sensitivity of the years used to calculate the DVBs at the monitoring sites. The sensitivity test results demonstrate that the weighted average of design values for the three design value periods over a five year period best represents the baseline concentrations, while taking into account the year-to-year variability of emissions and

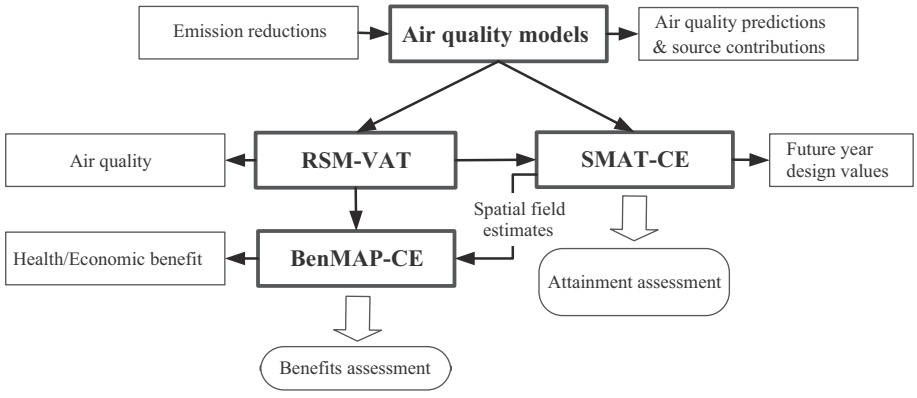

Fig. 1 – Decision support software tools including an air quality simulation tool RSM-VAT, a post analysis tool for health benefit, *i.e.*, BenMAP-CE and a modeled attainment test tool SMAT-CE.

meteorology. Additionally, the running average DV gives a lower data variability compared to the DV individually (USEPA, 2007). This methodology requires at least five years of observational data at each monitoring site. For the sites with less than 5 years of data, the DVB will be calculated using an average of two or single design value depending on the length of continuous data availability of 4 or 3 years, respectively. A monitoring site is excluded from the attainment test if there is less than three years of continuous data.

The $PM_{2.5}$ is calculated using 8 components: nitrate (NO_3^-), sulfate (SO_4^{2-}), organic carbon (OC), soil and other inorganic mass (crustal), elemental carbon (EC), ammonium (NH_4^+), salt, and particle bound water mass (PBW) attached to sulfate, nitrate, and ammonium. These concentrations are utilized to estimate the DVFs of $PM_{2.5}$. Since most of $PM_{2.5}$ monitoring sites are not co-located with a speciation site, the component-specific baseline values of each component for the $PM_{2.5}$ monitors are calculated as weighted average of observed component-specific baseline values from nearby speciation monitors by drawing Thiessen polygons (Browne, 2007) around the $PM_{2.5}$ monitoring site.

For spatial field estimates, the DVB is calculated as the weighted average of DVBs measured in adjacent monitoring sites. The weighting factor can be calculated by three different methods depending on users' choices. The default approach is calculated by the distances between the cell center of interest and surrounding monitoring sites after drawing Thiessen polygons around the cell:

$$\text{Weight}_A = \frac{\frac{1}{d_A}}{\frac{1}{d_A}+\frac{1}{d_B}+\frac{1}{d_C}+\frac{1}{d_D}} \quad (2)$$

where, d_A, d_B, d_C and d_D are the distances from a target grid cell E to monitoring sites A, B, C and D, and Weight_A is the weighing factor for site A. The second approach is to calculate the weighting factor using the square of the distances between the cell center and the surrounding monitoring sites. The third approach is to use the equal weighting factor for each neighboring monitoring site.

The DVB at site E is the sum of the weighting factor for each neighboring monitoring site multiplying the DVB as measured at the monitoring site:

$$\text{Gridcell}_{E,\text{baseline}} = \sum_{i=1}^{n} \text{Weight}_i \times \text{Monitor}_i \quad (3)$$

where, n is the number of neighboring sites, Weight_i is the weighting factor for site i, Monitor_i is the DVB at site i and $\text{Gridcell}_{E,\text{baseline}}$ is the DVB at grid cell E.

SMAT-CE also lets user choose whether to use model data for adjusting the spatial field estimates. This use of model data is referred as a gradient adjustment. The gradient adjusted DVB at grid cell E is the sum of products of the weighting factor for each monitoring site, the DVB as measured at the monitoring site and a gradient adjustment:

$$\text{Gridcell}_{E,\text{baseline}} = \sum_{i=1}^{n} \text{Weight}_i \times \text{Monitor}_i \times \text{Gradient Adjustment}_{i,E}. \quad (4)$$

A gradient adjustment is the ratio of the baseline model data for the grid cell of interest to the baseline model data for the grid cell containing the neighboring monitor:

$$\text{Gradient Adjustment}_{i,E} = \frac{\text{Model}_{E,\text{baseline}}}{\text{Model}_{i,\text{baseline}}} \quad (5)$$

where, $\text{Model}_{E,\text{baseline}}$ is the baseline model data at grid cell E, and $\text{Model}_{i,\text{baseline}}$ is the baseline model data of the grid cell containing monitor i.

The drawing of Thiessen polygons and the identification of neighboring monitors are achieved by an improved Voronoi Neighbor Averaging (VNA) algorithm following the standard VNA algorithm downloaded from the DotSpatial website (Microsoft, 2009). The improved VNA algorithm increases the computational efficiency by setting a fixed radius (7°, measured in latitude and longitude) around the cell of interest to identify neighboring monitors.

The RRF is calculated by taking the ratio of the future modeled concentration near a monitoring site (averaged over multiple days) to the baseline modeled concentration near the monitoring site (over the same days):

$$RRF_i = \frac{\text{Model}_{i,\text{future}}}{\text{Model}_{i,\text{base}}} \quad (6)$$

where, $\text{Model}_{i,\text{future}}$ is the mean of future modeled concentration over multiple days at site i, $\text{Model}_{i,\text{base}}$ is the mean of baseline concentration over the same days at site i, and RRFi is the RRF at site i.

The procedure for calculating a RRF is divided into two steps. The first step is to identify grid cells near a monitoring site. According to US EPA modeling guidelines (US EPA, 2007), the number of cells used in calculating the averages depends on the spatial resolution of the model grids (Table 1). In the case study, the model grid resolution is 12 km and therefore concentration average of 3 × 3 grid cells applies. The second step is to calculate a RRF. For 8-hour O_3, we use the highest 8-hour daily maximum concentrations in the nearby grid cells for each day both in the base and future years. The RRF value is calculated by taking the ratio of the average of the highest 8-hour daily maximum concentration of the future year to that of the base year. For annual $PM_{2.5}$, a separate RRF called component-specific RRF is calculated for each of the $PM_{2.5}$ components. The component-specific RRF is the ratio of the mean of the quarterly averaged daily predictions of the future year t2 that of the base year across the nearby grid cells.

There is uncertainty accompanying model predictions when predicting observed air quality at any given location in the base and future years. Uncertainty arises from various sources, *e.g.*, limitations in the formulation of the model and the model inputs including meteorological and other input

Table 1 – Recommendations for nearby grid cells for RRF calculation (US EPA, 2007).

Size of individual cell (km)	Size of the array of nearby cells (unitless)
4–5	7 × 7
5–8	5 × 5
8–15	3 × 3
>15	1 × 1

data base, and uncertainty in forecasting future levels of emissions. Since the RRF is a crucial parameter to predict the DVF based on the modeled concentrations in base and future years, the model results should be verified before preparing an attainment test. We recommend two steps for reducing model uncertainties. Firstly, an analysis should be preceded to verify the rationality of model input data by analyzing available air quality, meteorological and emissions data. Then, the model results should be compared with observational data to ensure the credibility of the results. In the case study, the input model files are provided and verified by U.S. EPA. As shown in Fig. 2, the CMAQ model can well capture the pattern of data variation observed at the monitoring site with a mean fractional bias (MFB, in Eq. (7)) ranging from −45.29% to 33.85% for annual $PM_{2.5}$ and from −18.81% to 19.68% for 8-hour O_3. The data used for model evaluation include those observed in 50 monitoring sites in eastern U.S. chosen randomly (Appel et al., 2007, 2008).

The DVF is calculated by multiplying DVB with RRF. The DVF for $PM_{2.5}$ is more complicated than O_3. The component-specific baseline concentrations are multiplied by component-specific RRFs to estimate the component-specific future concentrations for each $PM_{2.5}$ component. Then the DVFs of $PM_{2.5}$ are estimated by adding the future concentrations of the eight $PM_{2.5}$ components. If the future design values are less than the concentration specified in the NAAQS, it means all the monitoring sites pass the test.

1.2. Operation environment

SMAT-CE is developed for running on Windows 7 operation system using a personal computer. Table 2 shows operational environment of the case study.

1.3. Software design

Designed with a friendly interface to provide a superior user experience, SMAT-CE consists of a configuration, an analysis and a data viewer module (Fig. 3). Configuration module defines the input parameters and operation settings. The configuration settings are saved in a configuration (CFG) file. User can run SMAT-CE with a saved CFG file or on-screen settings to perform a new attainment test. The CFG files can also be used as templates to create a series of CFG files for batch job function. Users can utilize either the GUI or a batch file to run the Analysis module. The batch file is useful for loading multiple CFG files for batch processing. The results are saved in the same file directory for the CFG files. Data Viewer module provides six types of data reporting for the attainment test results, including GIS, map, data table and chart for results, as well as the configuration settings and execution logs (Fig. 3). The six sub-modules are integrated with the Data Viewer interface.

2. Results and discussion

2.1. Comparison of model results and software performance

The performances of SMAT-CE and MATS are compared using the model output for annual $PM_{2.5}$, daily $PM_{2.5}$, O_3 and visibility in a contiguous U.S. domain in terms of data accuracy, computational efficiency and user friendliness. The same input data and settings are used for both SMAT-CE and MATS.

2.1.1. Comparison of attainment test results

SMAT-CE uses the improved VNA algorithm to enhance computation efficiency while MATS applies a modified VNA

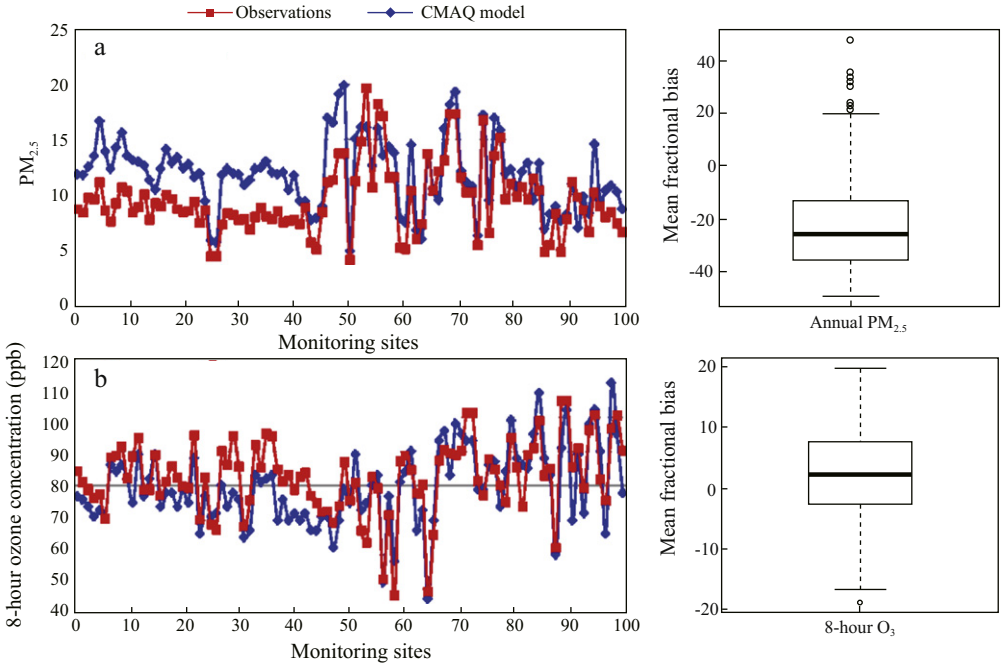

Fig. 2 – Spatial variation and mean fractional bias (MFB, in Eq. (7)) of CMAQ model and observations for annual $PM_{2.5}$ (a) and 8-hour O_3 (b) at 50 monitoring sites in the eastern U.S. chosen by random sampling.

Table 2 – Main operational environment for the case study.

Name		Description
Hardware	CPU	Inter(R) I7–950 @3.06 GHz
	Memory (RAM)	6.00 GB
	Hard disk	1T 7200 r/min
Software	Operation system	64-Bit Windows 7 Ultimate
	GIS platform	DotSpatial (Microsoft, 2009)

algorithm which is not an open-source program. To evaluate the improved VNA algorithm of SMAT-CE, we compared the neighboring monitors obtained by three VNA algorithms implemented in SMAT-CE, MATS and ArcGIS. ArcGIS (version 9.3) has been widely applied for generating Thiessen Polygon (Han, 2010; Shen and Cao, 2010). The results show that SMAT-CE replicates the neighboring monitors obtained by ArcGIS. For example, the neighboring monitors (A, B, C, D, E and F) around target monitor G are shown in Fig. 4. MATS can get similar results but miss one neighboring monitor F in Fig. 4. The agreement between SMAT-CE and ArcGIS indicates that the improved VNA algorithm of SMAT-CE is more reliable than that of MATS.

The results of the two cases in Eastern U.S. are compared to evaluate the difference between SMAT-CE and MATS. In one case (Case #1), same configuration and input files were applied for both tools. In the other case (Case #2), the effect of VNA algorithms was excluded to identify if there are other computational differences between the two. Fig. 5 shows the comparison of spatial field estimates for annual $PM_{2.5}$ in the future year between SMAT-CE and MATS for Case #1 and Case #2. The red points in the right panel of Fig. 5 are the locations of monitoring sites used for calculating the spatial field estimates of DV. The results of SMAT-CE and MATS are almost the same as shown in Fig. 5a, the main difference occurs at the boundary of model domain or at a few areas with scarce monitoring sites. Excluding the deviation caused by the two VNA algorithms, SMAT-CE replicates the calculation results of MATS (Fig. 5b).

The mean fractional bias (MFB) is calculated to compare the results of point estimates:

$$\mathrm{MFB} = \frac{1}{N}\sum_{i=1}^{N}\frac{(C_m - C_0)}{\left(\frac{C_0 + C_m}{2}\right)} \quad (7)$$

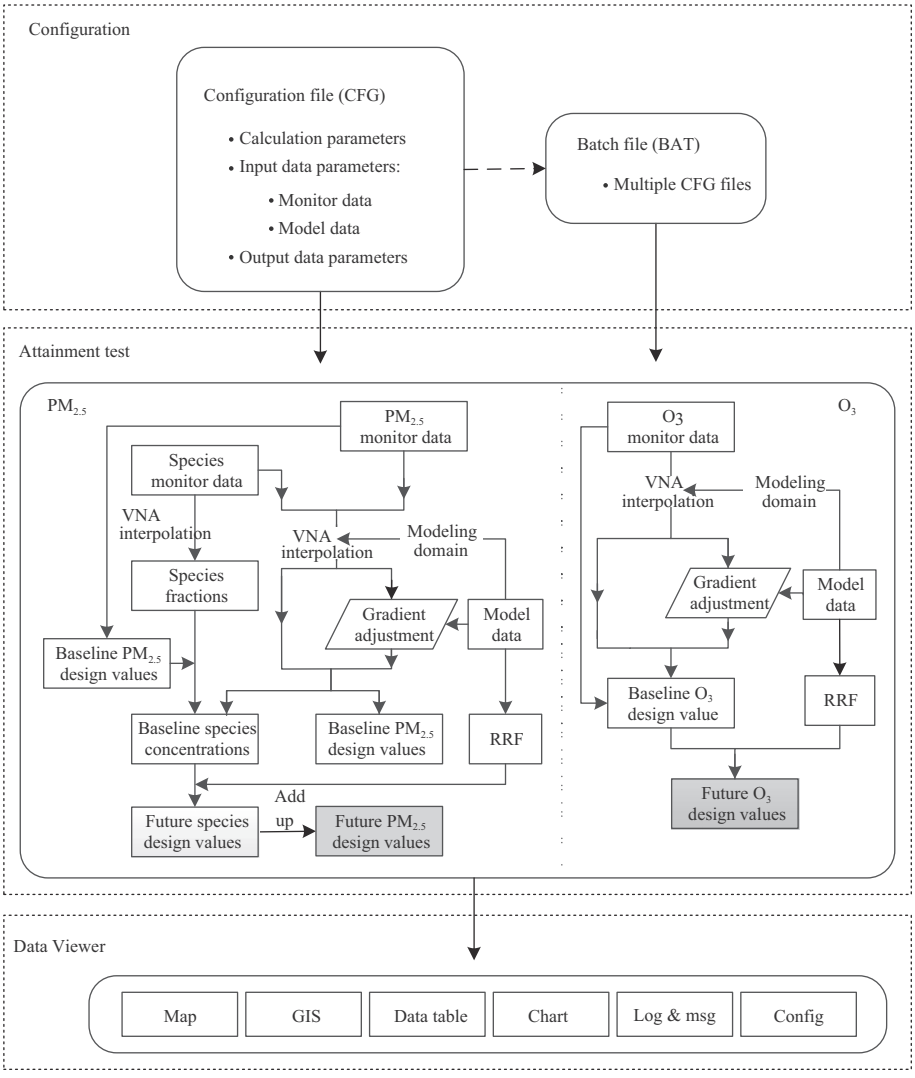

Fig. 3 – Functional framework and operational process of SMAT-CE.

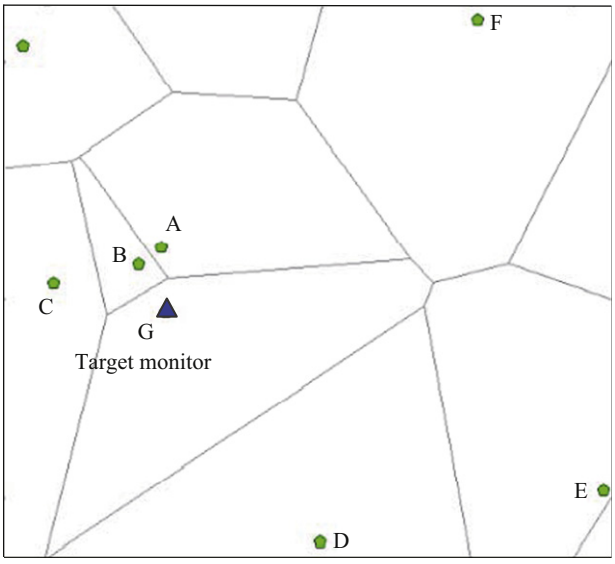

Fig. 4 – Neighboring monitors (A, B, C, D, E and F) identified by drawing a Thiessen polygon around the target point G using ArcGIS.

where, N is the number of monitoring site, C_m is the predicted DV of SMAT-CE at site i, and C_0 is the predicted DV of MATS at site i. Fig. 6 shows that the median values of MFB for both Case #1 and Case #2 are 0%. Only a few monitoring sites in Case #1 (Fig. 6a) have the MFB less than ±1%. After excluding the effects caused by the different VNA algorithms, the results of SMAT-CE and MATS are nearly identical (Fig. 6b).

2.1.2. Comparison of computational time

Three measures are adopted to improve computational efficiency in SMAT-CE while keeping the data accuracy: (1) SMAT-CE is coded with Microsoft Visual C#, which supports multithread computation (Barbhuiya and Liang, 2012) and is reported to have superior program operation efficiency comparing to JAVA and other programming languages on the Windows system (Wikipedia, 2013a), (2) the time-consuming calculation processes, such as spatial field estimates, the complicated chemical speciation calculation and the VNA algorithm, are optimized, and (3) data access mode is optimized to shorten the input data loading time during the calculation (*e.g.*, the latest method of DataReader (Wikipedia, 2013b) is used to read large size input files).

SMAT-CE and MATS were run on the same platform (Table 2) to evaluate the performance. Compared to MATS, the computational speed of SMAT-CE increases by 40%–87% for different analysis targets, which is mainly contributed by the improved VNA algorithm and data access mode (Table 3). The runtime of visibility is shortest since it calculates visibility levels for 156 specific Class I Areas (Abt, 2012) only and doesn't contain the time-consuming calculation processes. Because of the complicated chemical speciation calculation for $PM_{2.5}$ (*e.g.*, SO_4^{2-}, NO_3^-, *etc.*), the computing time of $PM_{2.5}$ is longer than that of O_3. The runtime of daily $PM_{2.5}$ is shorter than annual $PM_{2.5}$ because the spatial field estimates are not calculated for daily $PM_{2.5}$. (See Table 4.)

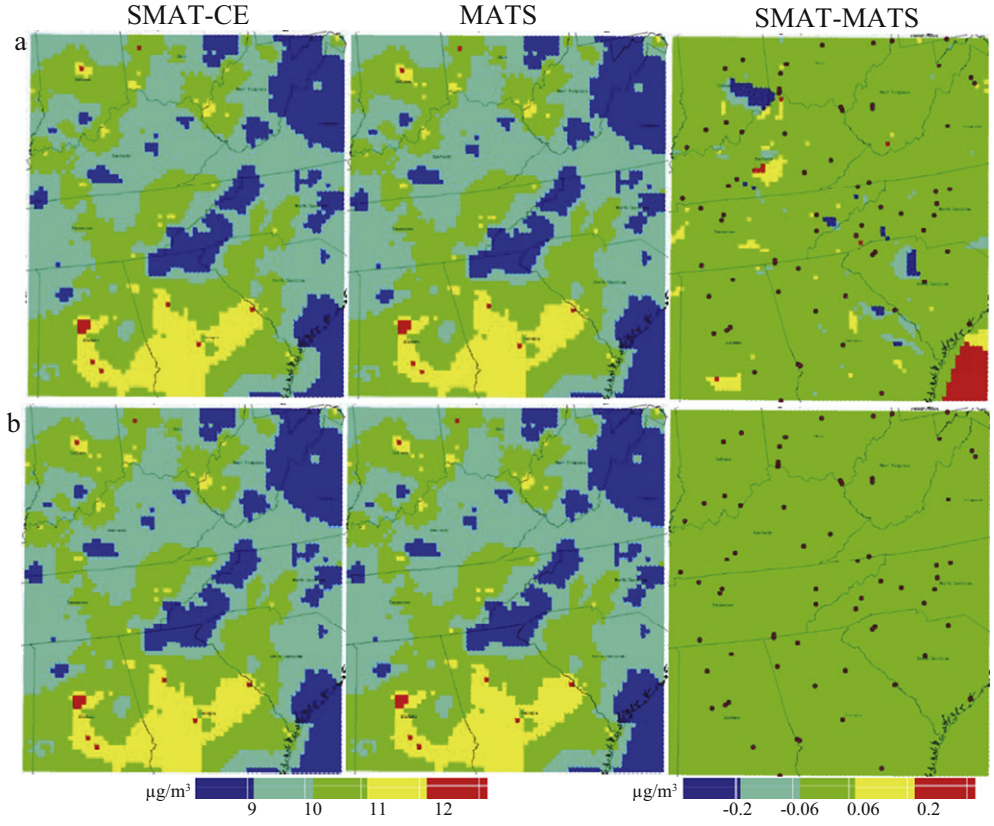

Fig. 5 – Comparison of spatial field estimates. (a) Case #1 and (b) Case #2 obtained by SMAT-CE and MATS.

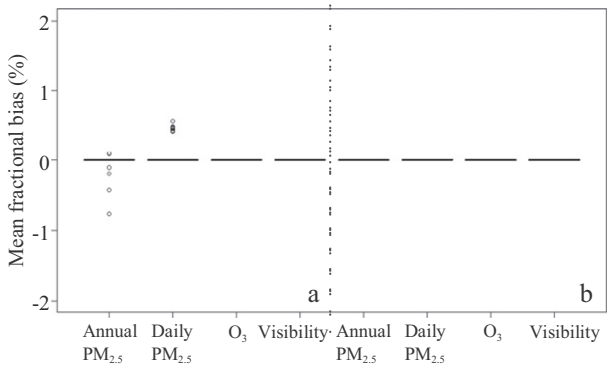

Fig. 6 – Comparison of point estimates: (a) Case #1 and (b) Case #2 by box plot for SMAT-CE and MATS (N is 186, 185, 907 and 12 for annual $PM_{2.5}$, daily $PM_{2.5}$, O_3 and visibility, respectively).

2.1.3. User-friendly design of SMAT-CE

Evolving and superior to MATS, SMAT-CE is an open-source package. Its GUI is designed to facilitate streamline operations. It provides easy and quick air quality attainment assessments. The data visualization and analysis functions are shown in Fig. 7. The legacy MATS uses a pop-up window for configuration settings and operations, which is not a user-friendly design by forcing user to a separate window. SMAT-CE is designed with an integrated functional framework to improve user experience. The configuration settings are arranged in a step-wise optional setting area in the left side of the framework. User can click the "Next"/"Back" icon or the sub-node to set the configurations sequentially. The execution logs are displayed in the right panel. Configuration indicators are utilized to remind users of the current status: green for ready, red for error, and yellow for in progress. After the computation is completed, the Data Viewer interface appears for visualization and analysis of the results. The interfaces of GIS, map, chart, data table, configuration and log are integrated into one screen. Users can view the maps, graphs or the data instantly in the right panel by clicking on the file lists on the left panel. The configuration settings can be reviewed or revised by returning to the setting area.

The Data Viewer module of MATS only allows user to load result files and view these data in GIS map or data table. Compared to MATS, SMAT-CE provides a full feature of data visualization to analysis. GIS-based map and picture-based map are two enhanced data visualization functions. The GIS-based map function allows user to perform a variety of mapping tasks, such as zooming based on different levels of shapefiles of different spatial scales. The picture-based map function is designed to improve the visualization efficiency. It provides a quick overview of air quality changes in the study domain. The speed for visualizing air quality surface by picture-based map function is 18 times faster than that of GIS-based map function (6.6 s/Map compared to 125.2 s/GIS) for the results in the U.S. domain. The Chart and Data table functions provide users with intuitive graphic and quantitative information. Chart analysis helps policy makers compare the effectiveness of air quality improvement. For example, it displays the comparison bar charts for the baseline and future values side by side to conveniently compare the air quality before and after emissions control. The Data table function allows users to examine the data specifically and output the results with CSV extension.

Table 3 – Comparison of computational time of SMAT-CE and MATS in a U.S. domain and the increasing ratio of SMAT-CE: annual $PM_{2.5}$, daily $PM_{2.5}$, O_3 and visibility.

Analyses	Computing time (min)		
	MATS	SMAT-CE	Increasing ratio (%)
Annual $PM_{2.5}$	248.87	103.80	58.29
Daily $PM_{2.5}$	103.38	13.78	86.67
O_3	29.08	17.38	40.23
Visibility	18.46	4.38	76.27

Table 4 – Input data of case study.

Data		Description
Model data	$PM_{2.5}$ and speciation	Daily or quarterly average speciated CMAQ estimates for 6 $PM_{2.5}$ species and $PM_{2.5}$
	O_3	Daily 8-hour average maximum CMAQ estimates for O_3
Monitor data	$PM_{2.5}$	Observed daily/quarterly average $PM_{2.5}$ data
	Speciation	Observed daily chemically speciated fine particle mass data
	O_3	Observed seasonal average 8-hour maximum O_3 data

2.2. Case study

The NAAQS for 8-hour O_3 and $PM_{2.5}$ are 75 ppb and 12 $\mu g/m^3$, respectively. The emission reduction measures in both nonattainment and maintenance areas are anticipated to achieve the air quality goals by 2020. SMAT-CE is applied to test if the attainment can be achieved. Annual $PM_{2.5}$ and 8-hour O_3 were selected in the case study to demonstrate the compliance of air quality at the monitoring sites in the U.S. domain under the NAAQS.

The base year is 2007 and the future year for the emissions reductions scenario is set to 2020. The 2020 control scenario for annual $PM_{2.5}$ is based on anthropogenic NO_X, anthropogenic SO_2, residential wood combustion and the direct $PM_{2.5}$ emission reductions from non-EGU (Non-Electric Generating Units) sources of 25%, 25%, 100% and 50% over the 2007 levels. The 2020 control scenario for 8-hour O_3 is based on the NO_X emission reductions from Area (NonPoint Area Sources), EGU point (Electricity Generating Unit Point Sources), nonEGU point (Non-Electricity Generating Unit Point Sources), Onroad (Onroad Mobile Sources) and Nonroad (Nonroad Mobile Sources) of 4%, 5%, 64%, 21% and 6%, as well as the VOC emission reductions from Area, nonEGU point, Onroad and Nonroad of 70%, 2%, 24%, and 4% over the 2007 levels. Table 2 lists the input data for the case study.

The base (2007) and future (2020) year attainment test results of the monitoring sites are shown in Fig. 8 (the nonattainment sites are marked in red). There are 379 nonattainment monitoring sites covering 28 different states for $PM_{2.5}$ in 2007 (Fig. 8a), mainly located in California and

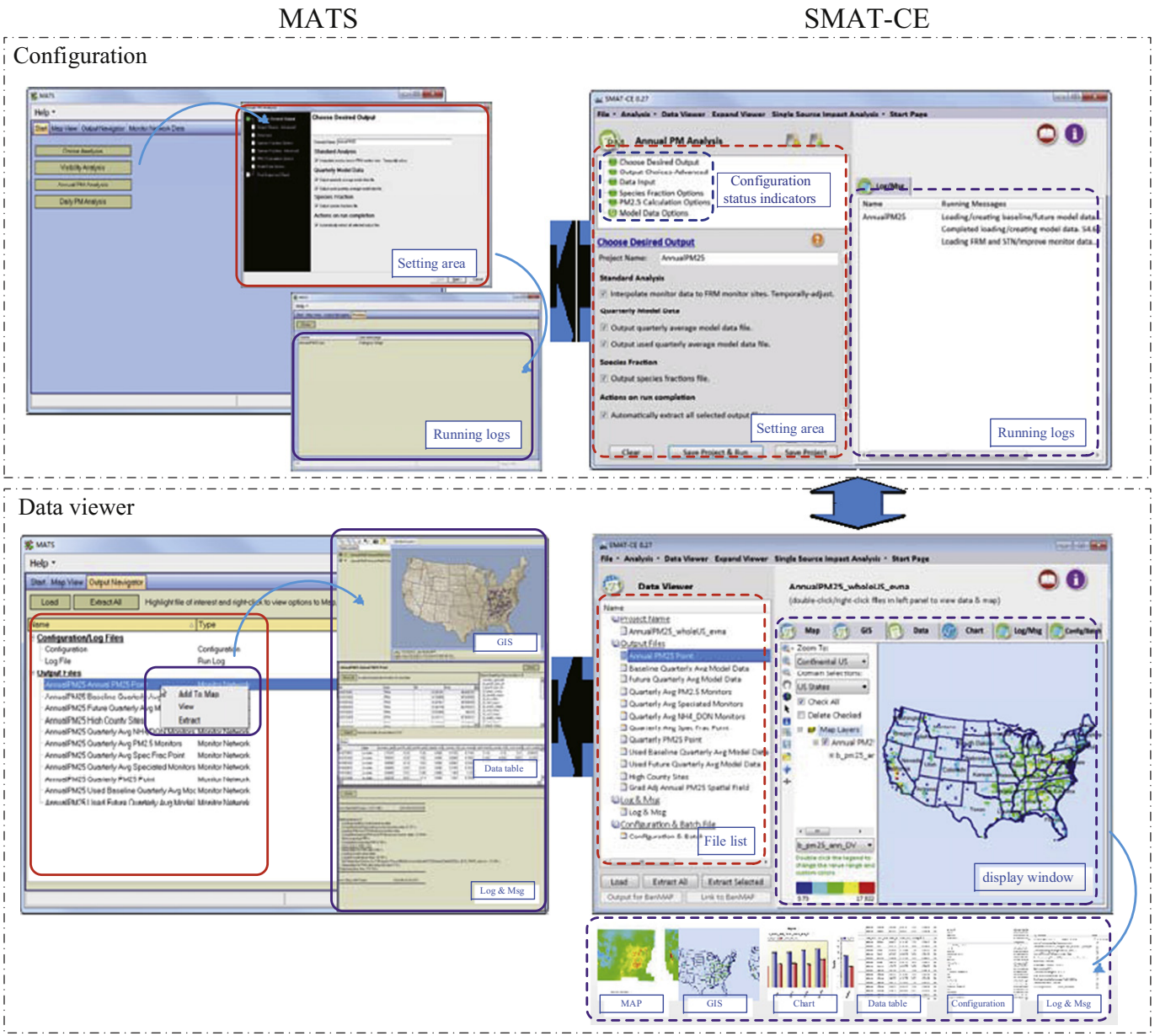

Fig. 7 – Comparison of operations and user-friendly designs between MATS and SMAT-CE.

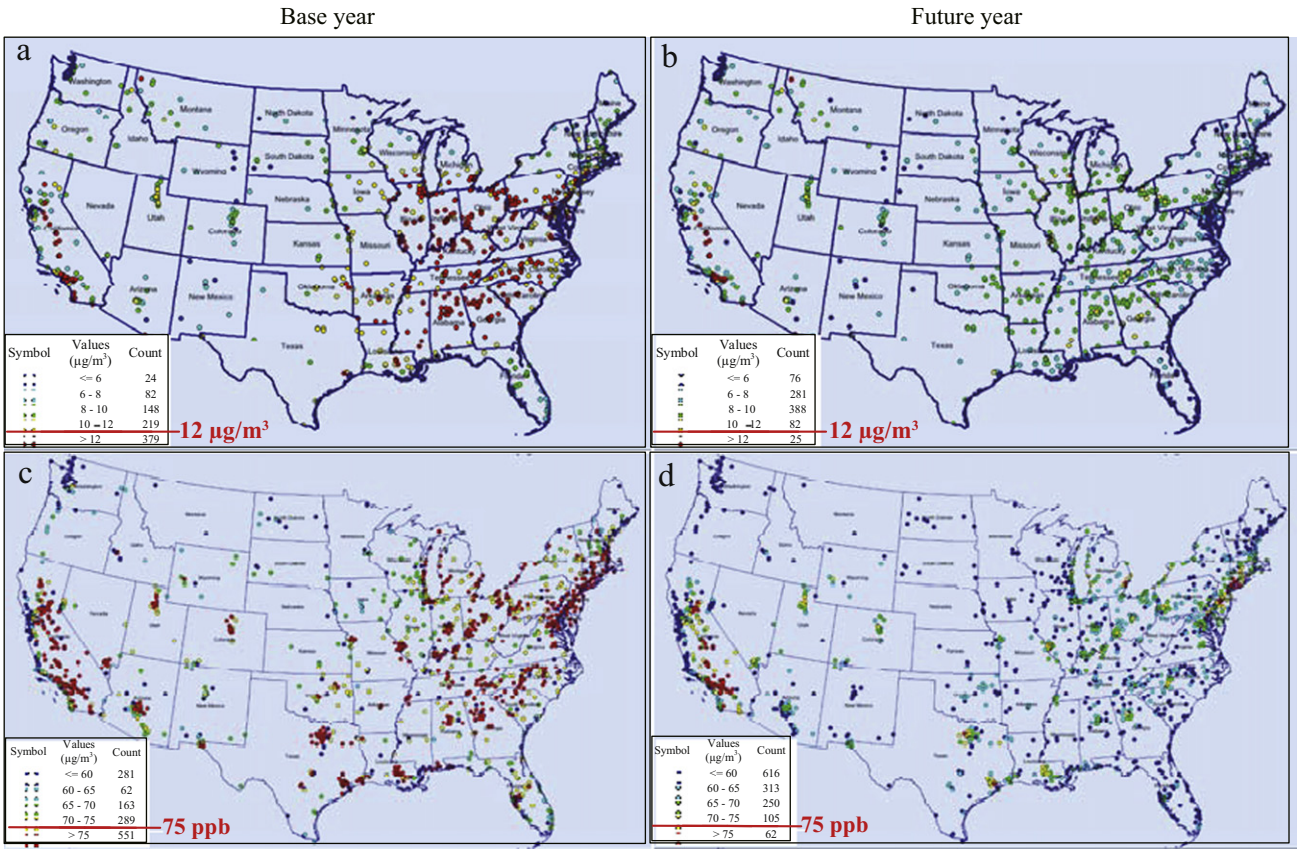

Fig. 8 – Attainment test results for demonstrating the 8-hour O_3 (a and b) and annual $PM_{2.5}$ (c and d) compliance with the NAAQS under the proposed emission scenarios, respectively. The base year is 2007 and the future scenario is set to 2020; the 2020 control scenario for annual $PM_{2.5}$ is based on anthropogenic NO_X, anthropogenic SO_2, residential wood combustion and the direct $PM_{2.5}$ emission reductions from non-EGU (Non-Electric Generating Units) sources of 25%, 25%, 100% and 50% over 2007 levels; the 2020 control scenario for 8-hour O_3 is based on the NO_X emission reductions from Area (NonPoint Area Sources), EGU point (Electricity Generating Unit Point Sources), nonEGU point (Non-Electricity Generating Unit Point Sources), Onroad (Onroad Mobile Sources) and Nonroad (Nonroad Mobile Sources) of 4%, 5%, 64%, 21% and 6%, as well as the VOC emission reductions from Area, nonEGU point, Onroad and Nonroad of 70%, 2%, 24%, and 4% over 2007 levels.

eastern U.S. Under the emissions control scenario in 2020, the number of nonattainment sites dramatically will reduce to 25, including 22 in California, 1 in Montana, 1 in Arizona and 1 in Wisconsin (Fig. 8b). The result indicates that the emission control of anthropogenic NO_X, anthropogenic SO_2 and $PM_{2.5}$ from non-EGU sources is effective in accordance with the current $PM_{2.5}$ standard as specified in the NAAQS. States that have the nonattainment sites are required to implement additional emission control in 2020.

For 8-hour O_3 standard, there are 41 states containing 551 nonattainment sites in 2007 (Fig. 8c). Under the 2020 emission scenario, many regions will attain the current O_3 standard. There will be 62 nonattainment sites in 2020, including 48 in California, 2 in Texas, 1 in Michigan, 2 in Pennsylvania, 1 in Maryland, 1 in New Jersey, 3 in New York and 4 in Connecticut (Fig. 8d). This suggests that the additional emission reduction for NO_X and VOC in 2020 is required to attain the current O_3 standard of the NAAQS.

The function of spatial field estimates provides an overview of air quality in the modeling domain. The spatial fields of predicted air quality also allow an assessment of whether the NAAQS is met at the locations without air monitoring stations. Fig. 9a displays the annual $PM_{2.5}$ DVFs ("f_pm25_ann_DV", $\mu g/m^3$) of spatial field estimates. It shows that the air quality in most areas will attain current $PM_{2.5}$ standard under the 2020 emission scenario. The nonattainment areas are concentrated in the western U.S. The results of the spatial field estimates can be exported to BenMAP-CE (Fig. 1) for evaluating the health and economic benefits of emission reduction.

User can also view the detailed information of the interested grid cell on clicks. For example, the annual $PM_{2.5}$ DVB ("b_pm25_ann_DV", $\mu g/m^3$) and DVF can be shown in the left bottom of the display area by clicking a cell (Fig. 9a). The DV of the grid cell is calculated as the weighted average of DVs of the neighboring monitoring sites based on Eqs. (1), (3) and (7). The Chart function can examine the air quality change of the neighboring monitoring sites around the target grid cell. Fig. 9b shows the comparison bar charts for the annual $PM_{2.5}$ DVBs and DVFs of the eight neighboring monitoring sites around a grid cell. Most of the air quality data at these neighboring monitoring sites are classified as nonattainment in 2007. Under the 2020 emission scenario, attainment of the current $PM_{2.5}$ standard is achieved (Fig. 9b).

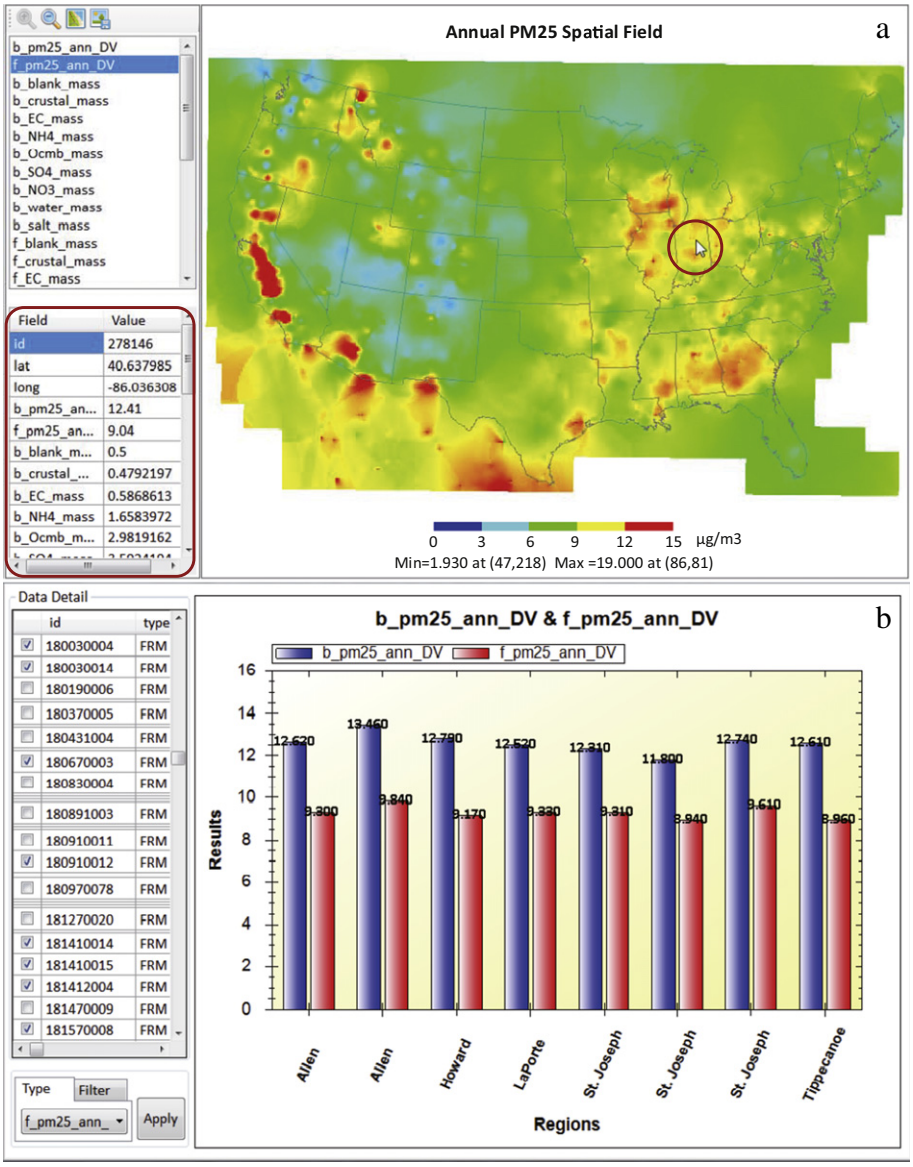

Fig. 9 – Picture-based Map function (a) of annual $PM_{2.5}$ DVFs (f_pm25_ann_DV) of spatial field estimates for the specific grid cell in Illinois, and Chart function (b) of annual $PM_{2.5}$ DVBs (b_pm25_ann_DV) and DVFs (f_pm25_ann_DV) for the eight neighboring monitoring sites in Illinois around the target grid cell marked within the red circle (The grid cell ID is 278146, and the eight neighboring monitoring sites are 180030004, 180030014, 180670003, 180910012, 181410014, 181410015, 181412004 and 181570008).

3. Conclusions

This article describes the development and application of an attainment test tool for air quality standards, SMAT-CE. The package uses spatial statistical method to combine observational data with model data for air quality attainment assessments aiming to provide a better scientific support for air quality management. SMAT-CE significantly enhances the performance and functionalities over MATS while retains the calculation accuracy and core functions. The redesigned framework of SMAT-CE greatly improves the computational speed, user-friendliness, visualization and analysis capability. The case study in a U.S. domain demonstrates that SMAT-CE can effectively perform the attainment tests for $PM_{2.5}$ and O_3, and assist in policy decision making by visualizing the attainment tests results.

The evaluation principles and methodology implemented in SMAT-CE are universally applicable in other nations or regions. Users are advised to follow the format of U.S. case included in the installation file to prepare the input data and use SMAT-CE internationally.

Acknowledgments

Financial support and data for this work are provided by the U.S. Environmental Protection Agency (Subcontract Number OR13810-001.04 A10-0223-S001-A04). This work is also partly supported by the funding of Guangdong Provincial Key

Laboratory of Atmospheric Environment and Pollution Control (No. 2011A060901011), the funding of State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex (No. SCAPC201308), and the project of Atmospheric Haze Collaborative Control System Design (No. XDB05030400) from Chinese Academy of Sciences.

REFERENCES

Abt, A., 2012. Modeled Attainment Test Software User's Manual. Available at:, http://www.epa.gov/ttn/scram/guidance/guide/MATS-2-5-1_manual.pdf (Date accessed: October 2012).

Appel, K.W., Gilliland, A.B., Sarwar, G., Gilliam, R.C., 2007. Evaluation of the community multiscale air quality (CMAQ) model version 4.5: sensitivities impacting model performance; part I—ozone. Atmos. Environ. 41 (40), 9603–9615.

Appel, K.W., Bhave, P.V., Gilliland, A.B., Sarwar, G., Roselle, S.J., 2008. Evaluation of the community multiscale air quality (CMAQ) model version 4.5: sensitivities impacting model performance; part II: particulate matter. Atmos. Environ. 42 (24), 6057–6066.

Barbhuiya, S., Liang, Y., 2012. A multi-threaded programming strategy for parallel Weather Forecast Model using C#. 2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing. Parallel, Distributed and Grid Computing PDGC, Solan, Himachal Pradesh, INDIA, Dec 06–08.

Browne, M., 2007. A geometric approach to non-parametric density estimation. Pattern Recog. 40 (1), 134–140.

Chai, F.H., Gao, J., Chen, Z.X., Wang, S.L., Zhang, Y.C., Zhang, J.Q., et al., 2014. Spatial and temporal variation of particulate matter and gaseous pollutants in 26 cities in China. J. Environ. Sci. 26 (1), 75–82.

Chemel, C., Fisher, B.E.A., Kong, X., Francis, X.V., Sokhi, R.S., Good, N., et al., 2014. Application of chemical transport model CMAQ to policy decisions regarding PM2.5 in the UK. Atmos. Environ. 82, 410–417.

Cheng, Y.L., Bai, Y.H., Li, J.L., Liu, Z.R., 2007. Modeling of air quality with a modified two-dimensional Eulerian model: a case study in the Pearl River Delta (PRD) region of China. J. Environ. Sci. 19 (5), 572–577.

Gao, Y., Zhang, M.G., 2014. Modeling study on seasonal variation in aerosol extinction properties over China. J. Environ. Sci. 26 (1), 97–109.

Gao, J., Zhang, Y.C., Zhang, M., Zhang, J.Q., Wang, S.L., Tao, J., et al., 2014. Photochemical properties and source of pollutants during continuous pollution episodes in Beijing, October, 2011. J. Environ. Sci. 26 (1), 44–53.

Han, J., 2010. Streamflow Analysis Using ArcGIS and HEC-GeoHMS. Available at:, https://ceprofs.civil.tamu.edu/folivera/TxAgGIS/Fall2010/Jeongwoo%20Han.pdf (Data accessed: 06 December 2010).

Li, L., Chen, C.H., Huang, C., Huang, H.Y., Zhang, G.F., Wang, Y.J., et al., 2011. Ozone sensitivity analysis with the MM5-CMAQ modeling system for Shanghai. J. Environ. Sci. 23 (7), 1150–1157.

Ma, J.Y., Yi, H.H., Tang, X.L., Zhang, Y., Xiang, Y., Pu, L., 2013. Application of AERMOD on near future air quality simulation under the latest national emission control policy of China: a case study on an industrial city. J. Environ. Sci. 25 (8), 1608–1617.

Microsoft, 2009. DotSpatial. Available at:, http://dotspatial.codeplex.com/SourceControl/latest#DotSpatial.Analysis/Voronoi.cs (Date accessed: 26 August 2009).

Sampson, P.D., Richards, M., Szpiro, A.A., Bergen, S., Sheppard, L., Larson, T.V., et al., 2013. A regionalized national universal kriging model using Partial Least Squares regression for estimating annual $PM_{2.5}$ concentrations in epidemiology. Atmos. Environ. 75, 383–392.

Sanjose, R., Perez, J., Gonzalez, R., 2007. An operational real-time air quality modelling system for industrial plants. Environ. Model Softw. 22 (3), 297–307.

Shen, Y.Y., Cao, B.Y., 2010. Design and implementation of a parallel Thiessen polygon generator. 2010 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010). IACSIT Press, Singapore, Nov 16–18.

Shrestha, K.L., Kondo, A., Kaga, A., Inoue, Y., 2009. High-resolution modeling and evaluation of ozone air quality of Osaka using MM5-CMAQ system. J. Environ. Sci. 21 (6), 782–789.

Singh, H.B., Cai, C., Kaduwela, A., Weinheimer, A., Wisthaler, A., 2012. Interactions of fire emissions and urban pollution over California: ozone formation and air quality simulations. Atmos. Environ. 56, 45–51.

U.S. EPA, 1997. Electronic Code of Federal Regulations: Protection of Environment, 40 CFR Part 50, Appendix I. Available at:, http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=52efd88f50c2a07c11b054c47c7c3652&n=40y2.0.1.1.1&r=PART&ty=HTML#40:2.0.1.1.1.0.1.19.10 (Date accessed: 18 July 1997).

U.S. EPA, 2007. Guidance on the Use of Models and Other Analyses for Demonstrating Attainment of Air Quality Goals for Ozone, PM2.5, and Regional Haze. Available at:, http://www.epa.gov/ttn/scram/guidance/guide/final-03-pm-rh-guidance.pdf (Date accessed: April 2007).

US EPA, 2008a. Final Ozone NAAQS Regulatory Impact Analysis. Available at:, http://www.epa.gov/ttnecas1/regdata/RIAs/452_R_08_003.pdf (Date accessed: March 2008).

US EPA, 2008b. Glossary of Terms and Acronyms. Available at:, http://www.mwcog.org/environment/air/downloads/pmp/GLOSSARY%20OF%20TERMS%20AND%20ACRONYMS%203.07.08.pdf (Date accessed: 07 March 2008).

US EPA, 2012a. Photochemical Modeling Tools — MATS 2.5.1 Installer Package. Available at:, http://www.epa.gov/ttn/scram/modelingapps_mats.htm (Date accessed: 28 December 2012).

US EPA, 2012b. Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter. Available at:, http://www.epa.gov/ttn/ecas/regdata/RIAs/finalria.pdf (Date accessed: December 2012).

US EPA, 2013. Electronic Code of Federal Regulations: Protection of Environment, 40 CFR Part 50, Appendix N. Available at:, http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=52efd88f50c2a07c11b054c47c7c3652&n=40y2.0.1.1.1&r=PART&ty=HTML#40:2.0.1.1.1.0.1.19.15 (Date accessed: 15 January 2013).

Wang, S.L., Gao, J., Zhang, Y.C., Zhang, J.Q., Cha, F.H., Wang, T., et al., 2014. 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.

Wei, W., Cheng, S.Y., Li, G.H., Wang, G., Wang, H.Y., 2014. Characteristics of ozone and ozone precursors (VOCs and NOx) around a petroleum refinery in Beijing, China. J. Environ. Sci. 26 (2), 332–342.

Wikipedia, 2013a. Comparison of C Sharp and Java. Available at:, http://en.wikipedia.org/wiki/Comparison_of_C_Sharp_and_Java (Date accessed: 04 June 2014).

Wikipedia, 2013b. DataReader. Available at:, http://en.wikipedia.org/wiki/DataReader (Data accessed: 15 May 2014).

Xing, J., Zhang, Y., Wang, S.X., Liu, X.H., Cheng, S.H., Zhang, Q., et al., 2011. Modeling study on the air quality impacts from emission reductions and atypical meteorological conditions during the 2008 Beijing Olympics. Atmos. Environ. 45 (10), 1786–1798.

Zhang, Q.J., Laurent, B., Velay-Lasry, F., Ngo, R., Derognat, C., Marticorena, B., et al., 2012. An air quality forecasting system in Beijing—application to the study of dust storm events in China in May 2008. J. Environ. Sci. 24 (1), 102–111.