



Using CMAQ for Exposure Modeling and Characterizing the Subgrid Variability for Exposure Estimates

VLAD ISAKOV*

NOAA/Atmospheric Sciences Modeling Division, Research Triangle Park, North Carolina

JOHN S. IRWIN

John S. Irwin and Associates, Raleigh, North Carolina

JASON CHING*

NOAA/Atmospheric Sciences Modeling Division, Research Triangle Park, North Carolina

(Manuscript received 31 January 2006, in final form 31 August 2006)

ABSTRACT

Atmospheric processes and the associated transport and dispersion of atmospheric pollutants are known to be highly variable in time and space. Current air-quality models that characterize atmospheric chemistry effects, for example, the Community Multiscale Air Quality model (CMAQ), provide volume-averaged concentration values for each grid cell in the modeling domain given the stated conditions. Given the assumptions made and the limited set of processes included in any model's implementation, there are many sources of "unresolved" subgrid variability. This raises the question of the importance of the unresolved subgrid variations on exposure assessment results if such models were to be used to assess air toxics exposure. In this study, the Hazardous Air Pollutant Exposure Model (HAPEM) is applied to estimate benzene and formaldehyde inhalation exposure using ambient annually averaged concentrations predicted by CMAQ to investigate how within-grid variability can affect exposure estimates. An urban plume dispersion model was used to estimate the subgrid variability of annually averaged benzene concentration values within CMAQ grid cells for a modeling domain centered on Philadelphia, Pennsylvania. Significant (greater than a factor of 2) increases in maximum exposure impacts were seen in the exposure estimates in comparison with exposure estimates generated using CMAQ grid-averaged concentration values. These results consider only one source of subgrid variability, namely, the discrete location and distribution of emissions, but they do suggest the importance and value of developing improved characterizations of subgrid concentration variability for use in air toxics exposure assessments.

1. Introduction

Developing a strategy for mitigating undesirable health consequences routinely includes an assessment of the extent to which human exposure to airborne pollutants will be mitigated. Human exposure assessment involves tracking different individuals or targeted

groups who encounter different levels of exposure resulting from differences in activity patterns (e.g., the amount of time spent in leisure activities, at the office, at home, or in an automobile), and physical characteristics (e.g., breathing rates). Different activities may bring individuals into very close proximity to the actual sources of pollutants (e.g., emissions from automobiles while driving, emissions from cleaning products in a home or work environment). Thus, to a certain degree, human exposure assessments can be strongly affected by the level of detail provided in resolving the time and spatial variation of airborne pollutant concentration values (Jerrett et al. 2005; Spicer et al. 1996). We will be using the Hazardous Air Pollutant Exposure Model (HAPEM; U.S. Environmental Protection Agency

* Additional affiliation: U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

Corresponding author address: Vlad Isakov, U.S. Environmental Protection Agency, AMD/NERL, 109 TW Alexander Dr, Research Triangle Park, NC 27709.
E-mail: isakov.vlad@epa.gov

2005) in this study. HAPEM uses annually averaged distributions of ambient concentrations adjusted for a diurnal temporal pattern by averaging the model results over 3-h blocks (i.e., 0000–0300, 0300–0600, etc.) for the entire year. The model accounts for behavior of demographic groups and provides estimates of “potential” inhalation exposures for the population.

Although air-quality models are used to provide the possible extent and severity of exposure of airborne pollutants on humans, their true value resides in the ability 1) to provide local-level details on the concentration distribution that would be prohibitive to acquire through monitoring, and 2) to provide quantitative estimates of the effectiveness of alternative strategies for mitigating undesirable health consequences. The diffusion of gases or particulate matter in the atmosphere is the result of a “chaotic system” (turbulence) whose individual effects on a particular release by definition cannot be simulated in a deterministic manner. Implicit in current air-quality models is the assumption that the hourly observed time and spatial variations in the concentration values can be envisioned as being partly deterministic and partly stochastic. For specified boundary conditions, the deterministic part of the time and spatial concentration variations are the ensemble-averaged hourly concentrations to be seen at each location in the modeling domain. Current models attempt to simulate the ensemble averages, but uncertainties arise because of limitations in our understanding of atmospheric processes and imperfect input data (meteorological, emissions, terrain, buildings, etc.). Thus, the observed scatter of observations about model predictions is a combination of naturally occurring stochastic variations that are impossible for any model to ever explicitly simulate and variations (“uncertainties”) arising from limitations in our knowledge and imperfect input data. Those processes not explicitly addressed and the simplification assumptions made in a model’s implementation provide many sources for “unresolved” (subscale) variability. For instance, the Community Multiscale Air Quality model (CMAQ; Byun and Schere 2006) provides volume-averaged concentration values for each cell in the modeling domain for a given set of stated conditions. Emissions are assumed to be instantaneously well mixed within the cell in which they are emitted. These model implementation assumptions were made to simplify the problem of simulating the atmospheric chemistry and the fate of pollutant species that are formed (secondarily from “primary” emissions that are directly emitted to the atmosphere) through photochemical reactions and aerosol processes. The CMAQ is designed to provide gridded concentration

information on hourly bases; their outputs are grid size dependent. Of course, the finer the grid size, as determined by using nesting approaches (Odman and Russell 1991), the finer the spatial features that are resolved. CMAQ and other such models can be used to augment ambient monitors for the purpose of conducting air pollution exposure assessments (Georgopoulos and Liou 2006). It is important to recognize that while grid models can provide refined information for exposure assessments, there are limitations. Typically, the smallest grid sizes applied in such models is on the order of 1 km or greater. For many pollutants, there is evidence of significant spatial variability at scales smaller than 1 km (Weijers et al. 2004). Even modeling assumptions regarding subgrid processes can affect the model outcomes (Miao et al. 2006).

While grid models are the model platform of choice for the simulation of atmospheric chemistry and fate of airborne pollutants, there are various transport and diffusion models (often called dispersion models) that have been developed to simulate the fate of those airborne pollutants that are relatively chemically inert. Not having to treat the atmospheric chemistry of mixtures, dispersion models can provide detailed resolution of the spatial variations in hourly averaged concentrations of airborne pollutants. Examples of such dispersion models include the American Meteorological Society (AMS)–Environmental Protection Agency (EPA) Regulatory Model Improvement Committee (AERMIC) modeling system (AERMOD; Cimorelli et al. 2005), the Assessment System for Population Exposure Nationwide (ASPEN; U.S. Environmental Protection Agency 2000), and the Industrial Source Complex Short Term (ISCST) model (U.S. Environmental Protection Agency 1995). To date, local-scale dispersion models have been relied upon to provide the desired detailed description of the concentration pattern. However, local-scale dispersion models cannot properly treat photochemical effects. Most of the air toxic pollutants listed in the National Toxics Assessment (U.S. Environmental Protection Agency 2000) are identified as having a photochemical origin or being strongly affected by photochemical processes. Also, an estimate of the background concentration levels must be provided to local-scale models, whereas such levels are provided directly by CMAQ. It would be desirable to combine the capabilities of chemical grid and dispersion models into one model, but this is a yet-evolving area of research and development.

We might well expect that a more explicit treatment of the dispersion of the emissions, plus a more discrete characterization of the concentration values within

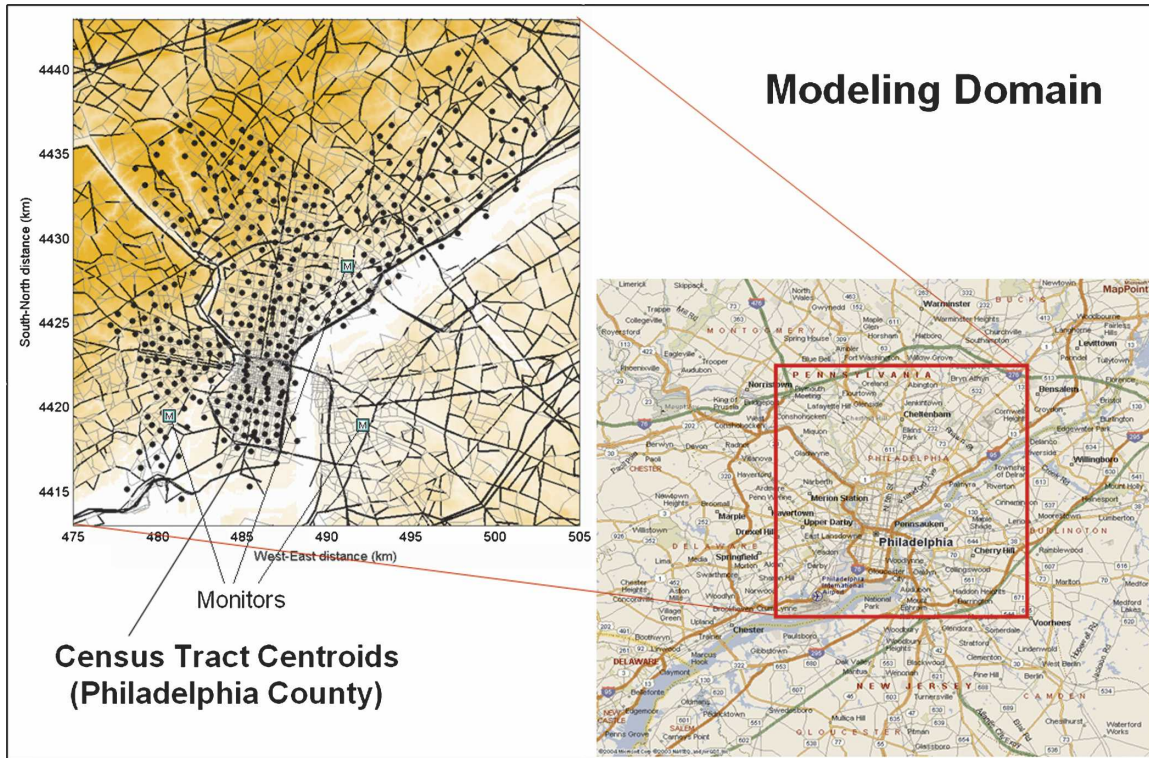


FIG. 1. Philadelphia study area.

each cell, would provide a more variable concentration distribution than what a standard run of CMAQ would provide. This raises the question of how large the unresolved subgrid variability might be and how this unresolved variability might be of consequence in human exposure assessments that thrive on detailed resolution of the temporal and spatial variations to be seen in concentration values.

In this study, we are attempting to provide an explicit characterization of the variations to be seen in the ensemble-averaged concentration values within each grid, resulting from the lack of treatment by the grid model for the subgrid variations in the average emissions. Various subgrid “adjustments” to the CMAQ results were used as input to HAPEM. The objective of this study is to evaluate the sensitivity of human exposure modeling results to subgrid variations in concentration values arising from the level of detail used to characterize pollutant emissions and the level of detail used to characterize the transport and diffusion of the pollutants. If we could develop a means to confidently characterize the subgrid variations to be seen in the ensemble-averaged concentration values, we provide exposure assessments as a means for assessing whether these subgrid variations in the ensemble-averaged concentration values are of any significance (i.e., Do we

need this greater detail in model output and its consequent requirement for more inputs and simulation time?).

2. Philadelphia study

a. Study description

To evaluate the sensitivity of human exposure modeling results to subgrid variations in the modeled concentrations, we focus on a $100 \text{ km} \times 100 \text{ km}$ area that included Philadelphia County and several surrounding counties (Fig. 1). The figure displays a roadway network and locations of 380 census-tract centroids in Philadelphia County.

An air toxics version of CMAQ was used to simulate ambient concentrations of various pollutants (Luecken et al. 2006). The CMAQ modeling system was run for an annual period in a nested mode at 36-, 12-, and 4-km horizontal grid dimensions using the 1999 National Emission Inventory (U.S. Environmental Protection Agency 2004) and meteorological outputs from 2001 simulations with the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Grell et al. 1995). The 36-km grid mesh encompasses the continental United States, the 12-km grid mesh covers Pennsylvania, Mary-

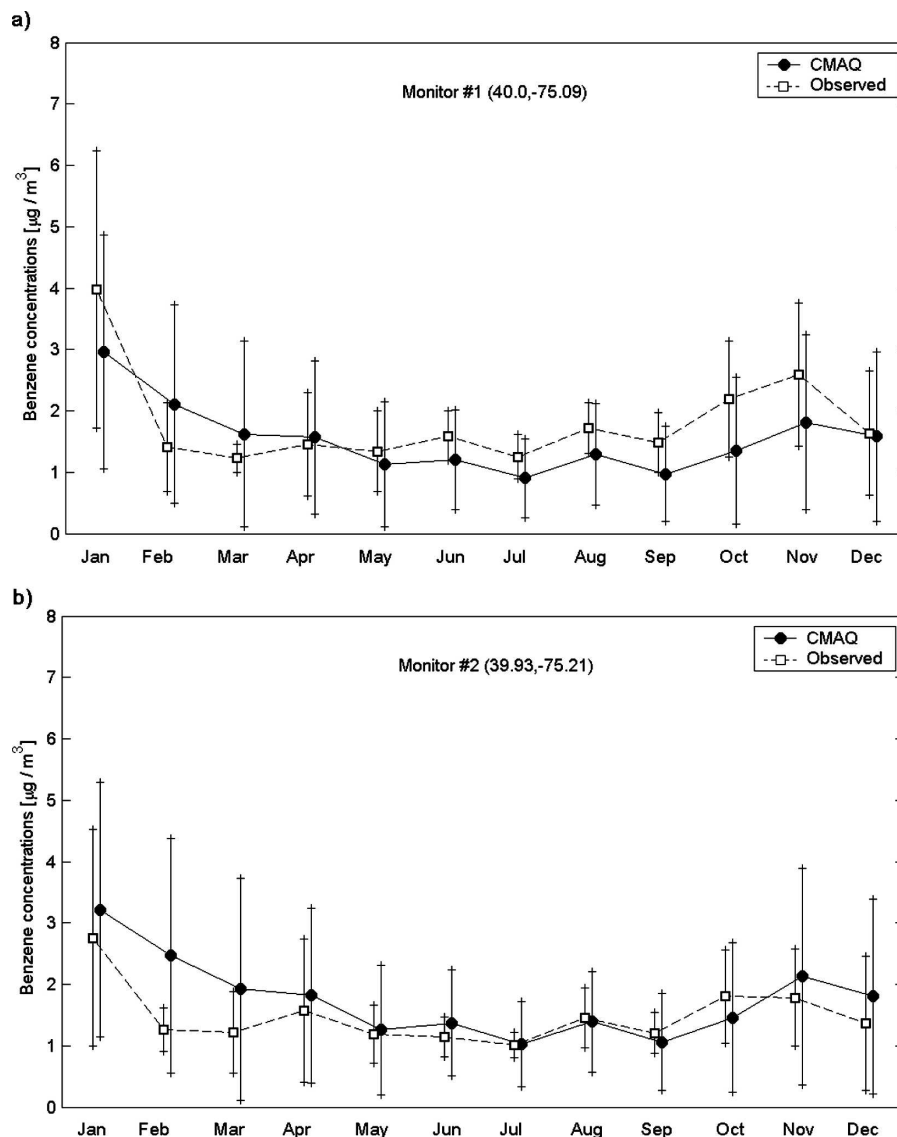


FIG. 2. Time series of observed and modeled benzene concentrations in Philadelphia at two monitor locations: (a) 40.0°N, 75.09°W and (b) 39.93°N, 75.21°W. Vertical lines indicate one standard deviation range of variability in monthly averaged concentrations.

land, New Jersey, Massachusetts, most of New York, and Virginia, and the 4-km grid mesh encompass Philadelphia, Pennsylvania, and Delaware. Luecken et al. (2006) compared CMAQ results at 36-km resolution with observed concentrations from 35 monitors in eight cities. Although detailed field observations for the Philadelphia modeling domain are limited, air toxics concentration measurements are available from two monitors in Philadelphia County. The locations of monitoring sites are indicated in Fig. 1. Details of the model runs and preliminary results have been reported in Ching et al. (2004). In that study, the temporal variability over the year for 3-hourly diurnal time blocks

was introduced into HAPEM, version 5 (HAPEM5). In general, the modeled mean values compared reasonably well to the observed values. While this limited comparison is encouraging, to properly assess and evaluate model performance will require observational data from more than two monitor locations. To evaluate the model's ability to reproduce the seasonal trend of pollutant concentrations, time series of monthly averaged modeled and observed concentrations for benzene and formaldehyde are shown in Figs. 2 and 3, respectively. As we can see from the figures, the model reasonably simulates seasonal patterns for both pollutants.

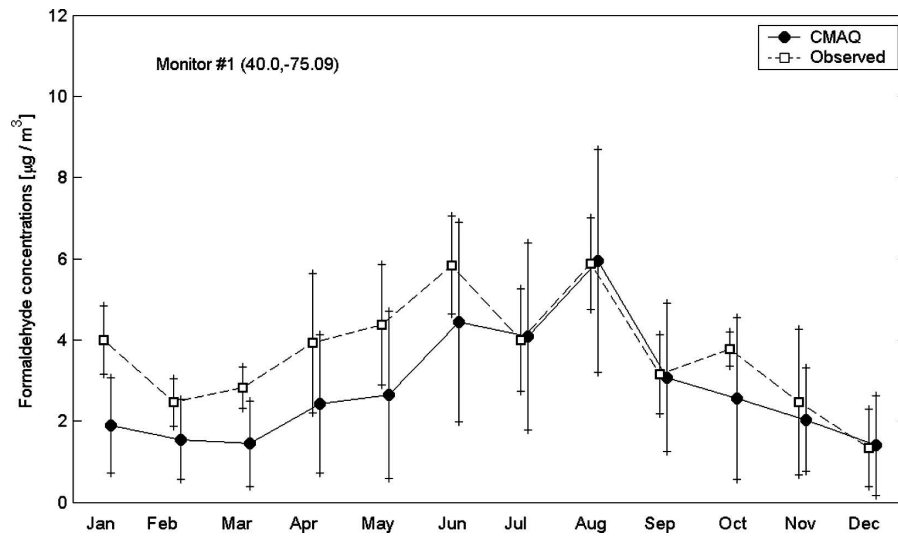


FIG. 3. Time series of observed and modeled formaldehyde concentrations in Philadelphia at monitor 1 (40.0°N, 75.09°W). Vertical lines indicate one standard deviation range of variability in monthly averaged concentrations.

Exposure to air pollutants is defined as contact over time and space between persons and pollutant species; exposure assessments attempt to define, characterize, and quantify that exposure in human populations; these assessments are typically performed on a census-tract level because this is the most detailed information available on population distribution. CMAQ provides average concentrations for grid volumes, which do

provide the finescale details in the concentration pattern that might occur within a grid volume. Locations of census-tract centroids in the 4-km CMAQ grid are shown in Fig. 4. As can be seen from the figure, the number of census tracts in a 4 km × 4 km CMAQ grid cell ranges from a few tracts in suburban areas to up to 30 tracts in the downtown area. This suggests that both population density and concentration values will likely

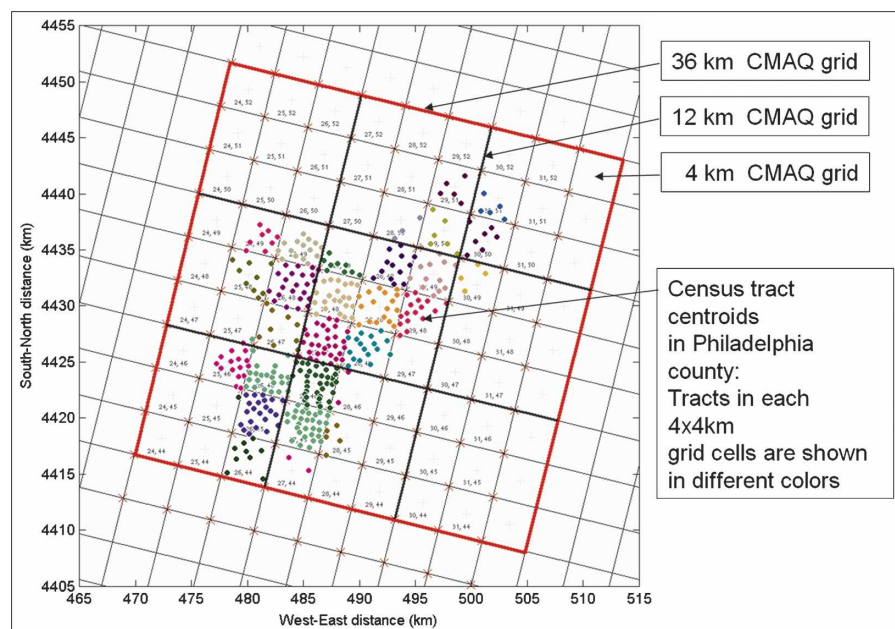


FIG. 4. Philadelphia modeling domain: CMAQ grid and census-tract centroids.

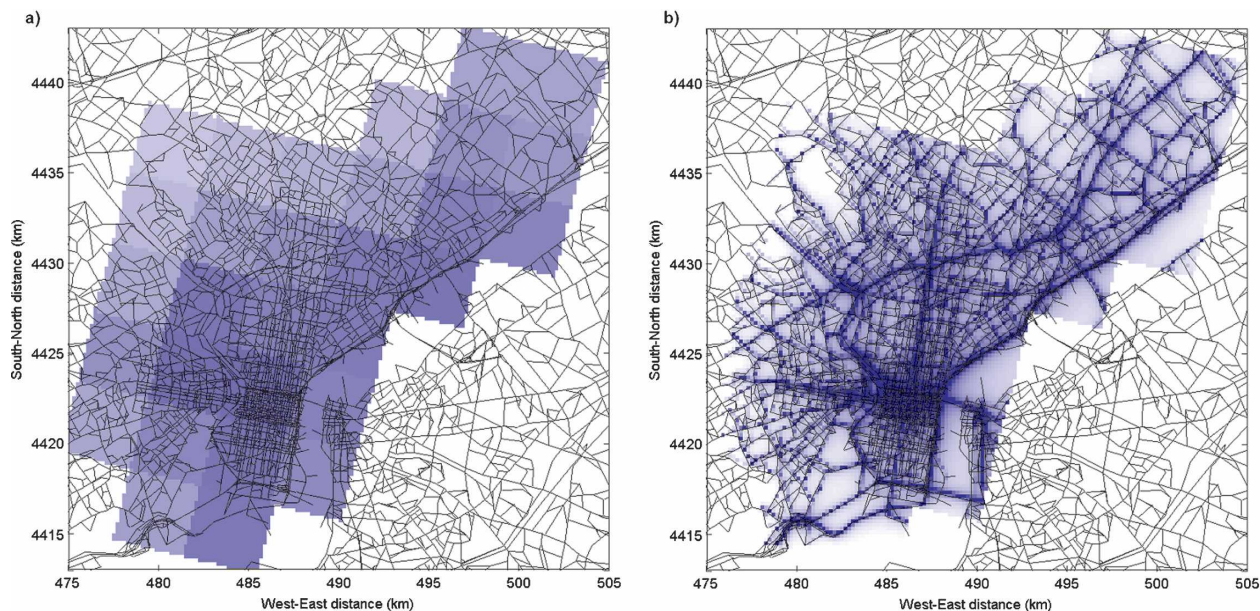


FIG. 5. An example of benzene concentrations in Philadelphia County (a) modeled at $4 \text{ km} \times 4 \text{ km}$ resolution using CMAQ and (b) modeled at finer scale using ISCST3. A darker shade indicates higher concentrations, and a lighter shade indicates lower concentrations.

vary within a typical CMAQ grid cell, such that if the variations in concentration are not adequately addressed, an exposure assessment may be compromised.

To examine the effects of the subgrid concentration variability, we analyzed benzene emission from mobile sources in Philadelphia. Benzene is primarily a mobile source emission, and thus road location and traffic data can be used as a basis for investigating subgrid variability of roadway emissions. The roadway network is indicated in Fig. 5. The distribution of the roadways varies greatly in density and pattern across the county. Benzene emissions from mobile sources were allocated to individual road links (road segments) to provide the finescale details on mobile source emissions. The ISCST, version 3, (ISCST3) model was used to simulate annually averaged benzene concentrations at 200-m receptor resolution, and this output provided the statistical properties of concentration distributions for individual CMAQ 4-km grid cells.

b. A method to account for local details

Figure 5 compares the spatial distribution of benzene concentrations modeled by a CMAQ, regional photochemical model at 4-km resolution (Fig. 5, left panel), and modeled at a finer scale using ISCST (Fig. 5, right panel). In this color-coded figure, a darker shade indicates higher concentrations, and a lighter shade indicates lower concentrations. As expected, the dispersion modeling results, conducted at a finer scale, reveal high

gradients near the sources, which are not predicted by CMAQ. These local features can be resolved by dispersion models (such as AERMOD, ASPEN, and ISCST), but these models cannot handle chemical transformations. Chemical transport models (such as CMAQ) can resolve temporal variations resulting from photochemistry, but not spatial details on this finescale. Therefore, we seek a method that allows CMAQ to provide regional background concentration values and contributions from chemically reactive pollutants, and local details in relatively inert pollutants concentrations to be provided by a dispersion model.

We propose the following approach for combining CMAQ results with local-scale dispersion modeling results: for each modeled pollutant, total concentration C^{Total} ($\mu\text{g m}^{-3}$) can be represented as a combination of its primary and its secondary contribution,

$$C^{\text{Total}} = C_{\text{Primary}}^{\text{Total}} + C_{\text{Secondary}}^{\text{Total}}, \quad (1)$$

where $C_{\text{Primary}}^{\text{Total}}$ ($\mu\text{g m}^{-3}$) is CMAQ concentration resulting from the contributions of primary emission sources and $C_{\text{Secondary}}^{\text{Total}}$ ($\mu\text{g m}^{-3}$) is CMAQ concentration resulting from the contribution of secondary production.

Furthermore, concentrations can be represented as a combination of the volume-averaged concentration C and subgrid variations ΔC ,

$$C = (C_{\text{Primary}} + \Delta C_{\text{Primary}}) + (C_{\text{Secondary}} + \Delta C_{\text{Secondary}}). \quad (2)$$

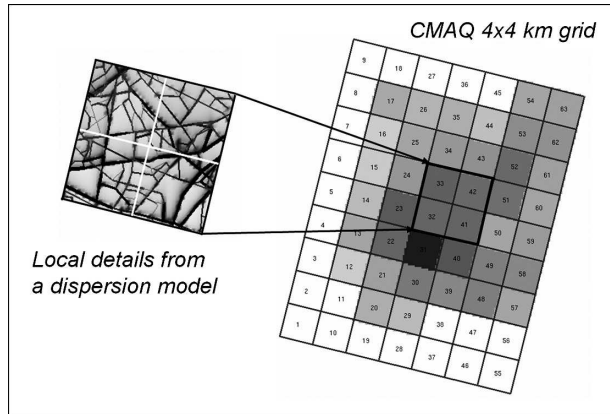


FIG. 6. Combined local-scale and regional modeling results.

The subgrid variations ΔC arise from the contribution of subgrid emission sources (Ching et al. 2006a), and spatial and temporal variability arise from chemical and turbulent interactions (Herwehe 2000). Subgrid variations for primary contributions can be estimated by resolving the local scale using dispersion models (Ching et al. 2006b; Isakov and Venkatram 2006; Touma et al. 2006), but for secondary contributions, methods are yet to be developed [although employing large-eddy simulation (LES) of coupled pollutant photochemistry–dynamics techniques is a promising approach for situations when chemical reaction time scales are on the order of the turbulent eddy time scales (Herwehe 2000)]. In this study, we are using a dispersion model to estimate $\Delta C_{\text{Primary}}$,

$$\Delta C_{\text{Primary}} = (C_{\text{Local}} - \overline{C_{\text{Local}}}), \quad (3)$$

where C_{Local} is the concentration at any location within the grid obtained from the dispersion model and $\overline{C_{\text{Local}}}$ is the average of all of the C_{Local} concentrations in the grid cell. Therefore, in order to adjust CMAQ results for local-scale variability, for each grid cell we simply combine CMAQ concentrations ($C_{\text{CMAQ}} = C_{\text{Primary}} + C_{\text{Secondary}}$), and differences between an average value from a dispersion model for all receptors within a grid cell and actual modeled concentrations at every receptor in a grid cell,

$$C_{\text{Adjusted}} = C_{\text{CMAQ}} + (C_{\text{Local}} - \overline{C_{\text{Local}}}), \quad (4)$$

where C_{Adjusted} is the new CMAQ concentration ($\mu\text{g m}^{-3}$) adjusted for local-scale variability.

Figure 6 illustrates the approach. In this figure, CMAQ at $4 \text{ km} \times 4 \text{ km}$ resolution provides a general pattern of benzene concentrations in Philadelphia County, and local subgrid details are obtained from a dispersion model such as AERMOD, ASPEN, or ISCST.

This approach was applied to Philadelphia using CMAQ as the chemical grid model and ASPEN as the dispersion model. Two pollutants were treated—benzene and formaldehyde. The CMAQ concentrations, enhanced with local details from ASPEN for benzene and formaldehyde, are shown in Fig. 7. Benzene is considered an inert pollutant, while formaldehyde is chemically reactive. For chemically reactive species, the adjustments provided by the dispersion model (ASPEN in this case) reflect the effects of direct emissions (of formaldehyde in this example). It was found in this example that the formaldehyde concentrations are dominated by secondary formed contributions from CMAQ. For both chemicals, local details are important, and the approach provides a “texture” on top of CMAQ results. Table 1 shows a comparison between modeled annually averaged concentrations from CMAQ adjusted for local details from ASPEN and observations.

To investigate how reasonable the assumption about grid-averaged concentrations is, we compared grid-averaged concentrations from CMAQ and ASPEN. The results of this comparison are shown in Fig. 8 for benzene (Fig. 8a) and formaldehyde (Fig. 8b). As can be seen from the figure, grid-averaged values from two models are similar (within a factor of 2), except for three outliers. In these cases, ASPEN receptors were at the edge of the county boundary and were few in number. Thus, using this method might introduce an error resulting from the limited sample size. In addition, special attention is needed to check the emissions prior to applying the method based on Eqs. (1)–(3).

c. Linkage between air-quality and exposure modeling

While ambient concentrations for a given time period at a given location can be estimated by various means, people move from location to location (e.g., from home to work, or from home to school), and they spend portions of their time in nonambient settings, such as at home, in the workplace, in school, or in a vehicle. Because air-quality concentrations in these various microenvironments can be very different than those in the outdoor (ambient) environment, human exposure models generally consider microenvironmental factors in predicting the “potential” inhalation exposure.

In this study, we used HAPEM5, a screening-level exposure model designed to predict the potential inhalation exposure for the general population, or specific subpopulations, over spatial scales ranging from urban environments to nationwide. HAPEM5 tracks representatives of specified demographic groups as they move among indoor and outdoor microenvironments

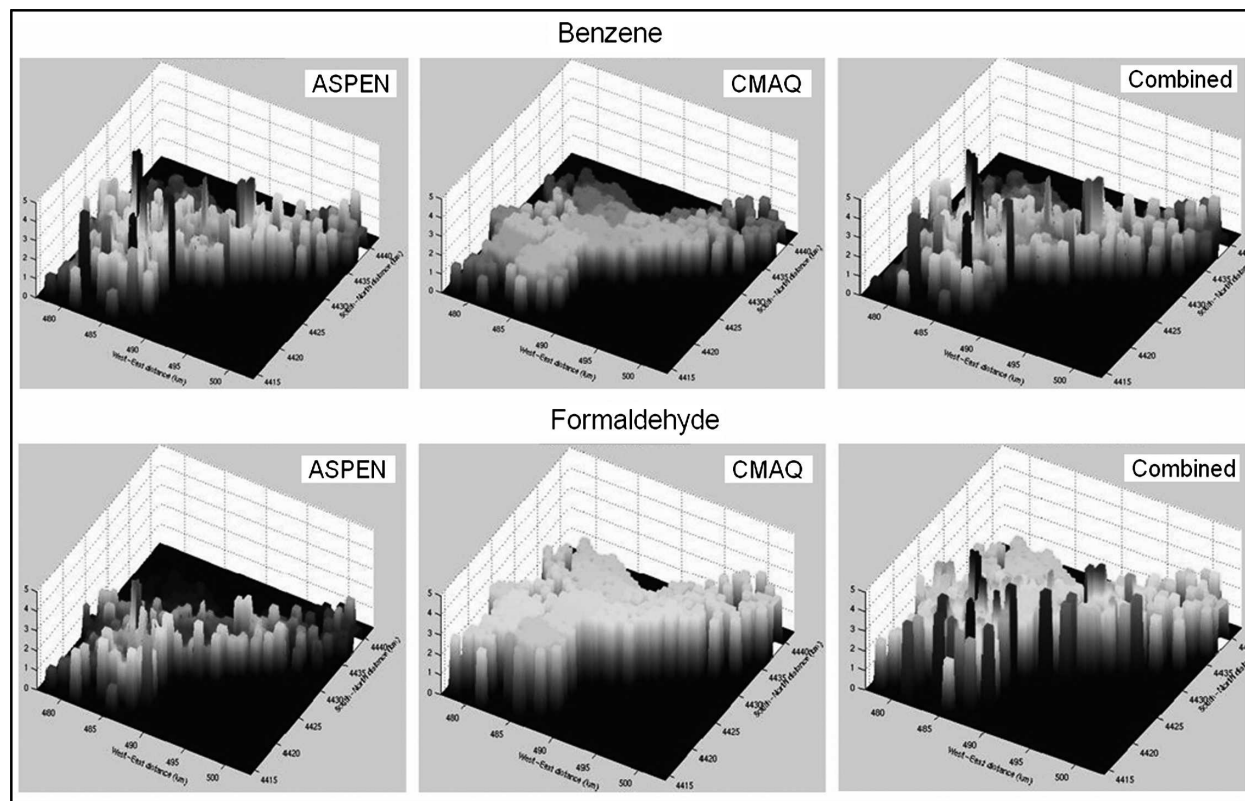


FIG. 7. Annually averaged concentrations ($\mu\text{g m}^{-3}$) of (left) ASPEN, (middle) CMAQ, and (right) combined—ASPEN + CMAQ—local- and regional-scale model results for two pollutants: (top) benzene and (bottom) formaldehyde.

and among geographic locations. Four major inputs are used in HAPEM: population data from the U.S. Census, population activity data from human diary data, microenvironmental data, and air-quality data. In this study, air-quality data are provided by hourly CMAQ concentrations enhanced with local details from ASPEN to build a diurnal temporal pattern of the concentration values by averaging the model results over 3-h blocks for the entire year.

A series of HAPEM5 simulations were made using the CMAQ results that were enhanced by using ASPEN results for benzene and formaldehyde. An example of simulated exposure for benzene is shown in Fig. 9. While the exposures are different from ambient concentrations, both maps reveal similar patterns, depicting the road net-

work in Philadelphia County. This is not surprising, because most of the benzene emissions in Philadelphia County come from mobile sources. We also found that for benzene, both concentrations and exposures are not correlated with population density, as illustrated in Fig. 9.

d. Sensitivity tests to investigate the importance of local details for exposure assessments

In this study, we use various subgrid adjustments to the CMAQ results as input to HAPEM. We also evaluate the sensitivity of human exposure modeling results to subgrid variations in concentration values. It is a property of the CMAQ model that its quantitative outcome is dependent on the selection of the grid cell size. Regardless of the finest grid size chosen, model output

TABLE 1. Comparison of annually averaged benzene and formaldehyde concentrations with observations.

	Benzene ($\mu\text{g m}^{-3}$)		Formaldehyde ($\mu\text{g m}^{-3}$)	
	Monitor 1	Monitor 2	Monitor 1	Monitor 2
	40.0°N, 75.09°W	39.93°N, 75.21°W	40.0°N, 75.09°W	39.93°N, 75.21°W
CMAQ	1.54	1.74	2.80	2.97
ASPEN + CMAQ	1.45	1.36	3.80	3.89
Obs	1.29	1.54	3.61	—

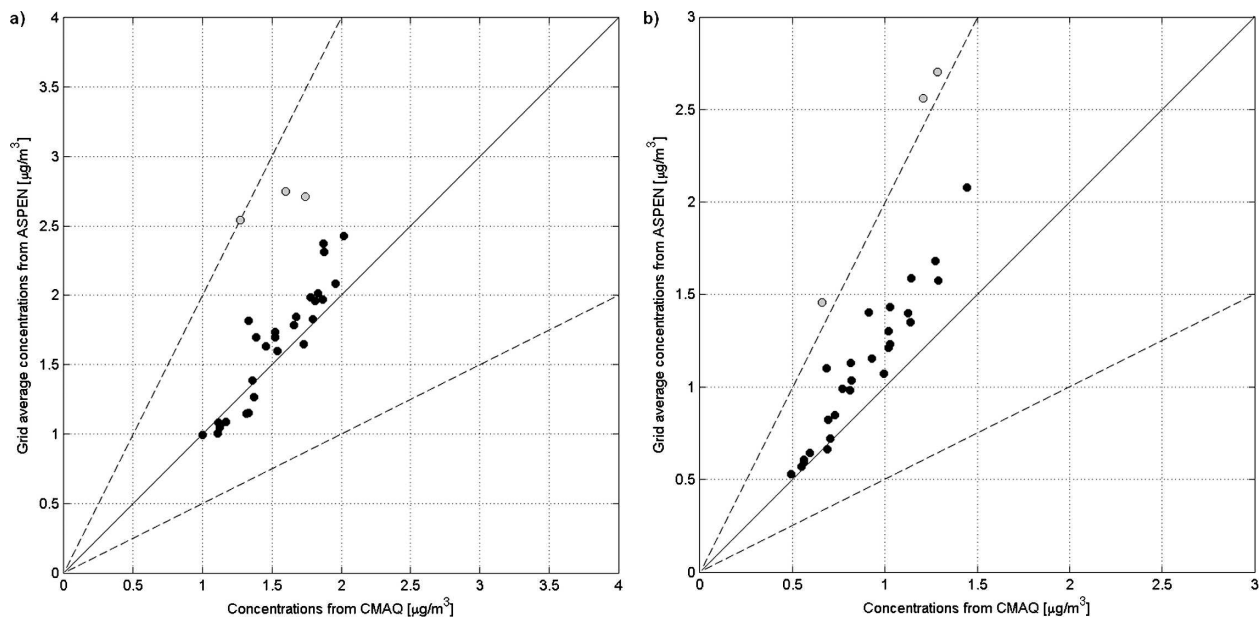


FIG. 8. Comparison of average concentrations ($\mu\text{g m}^{-3}$) from ASPEN and CMAQ for all grid cells for (a) benzene and (b) formaldehyde. Dashed diagonal lines indicate a factor-of-2 difference, and the gray dots indicate three outliers.

is a single value for each grid cell for the hour of simulation. In this sensitivity, we investigate how important it is to enhance CMAQ with local details from a dispersion model and how important spatial resolution in CMAQ (i.e., 4, 12, or 36 km) is for exposure modeling. A series of HAPEM5 simulations were made using the different inputs from CMAQ: standard CMAQ output and that enhanced with local details from ASPEN, using various spatial resolutions (4, 12, and 36 km). Results of HAPEM5 simulations for benzene and formaldehyde are shown in Fig. 10. The figure compares exposure from HAPEM using CMAQ only and CMAQ enhanced with local details from ASPEN. As can be seen from the figure, exposure has a significantly wider range when using CMAQ enhanced with local details as compared with a standard CMAQ output, even with an increased spatial resolution. This is more significant for benzene than for formaldehyde, because the largest fraction of the formaldehyde concentration values formed through chemical reactions and is provided by CMAQ, which at best is providing a 4-km average value in these simulations.

3. Analysis of subgrid variability

Results shown above indicate current operational feasibility for running CMAQ on an annual basis at grid sizes of 4 km or so for potential use in performing exposure model assessments.

We have noted that exposure models are designed to

assess exposure at census-tract levels. We have further seen that the more highly populated sections of urban areas generally contain multiple census-tract centroid locations within even this relatively fine grid (4-km resolution) CMAQ. By treating emissions with local-scale modeling, we have seen resulting concentrations at the census-tract level to exhibit added spatial variability to that using CMAQ alone. Indeed, potential additional variability exists even within census tracts because of subgrid or even sub-census-tract emission patterns and complex photochemistry. To illustrate this, we compare results of a) subgrid variability (SGV) estimated using data from finescale modeling at 380 tract centroids (ASPEN) and b) SGV estimated using data from finescale modeling at 16 000 receptors (ISCST3; 200-m resolution), as shown in Fig. 11.

This example demonstrates that assessments based on CMAQ modeling at 4 km, and perhaps even at finer grid sizes of the order of 1 km, cannot accurately characterize “hot spots” and will typically underestimate maximum potential exposure levels. We now review features and characteristics of SGV in air-quality grid models and outline a strategy for their incorporation for operational exposure modeling implementation.

In general, air-quality grid models like CMAQ predict concentration fields from pollutant emission sources into an evolving dispersive and chemically reactive atmospheric environment. By decreasing the grid size, or by using adaptable grid meshes or “nested” subdomains with decreasing grid sizes, the concentra-

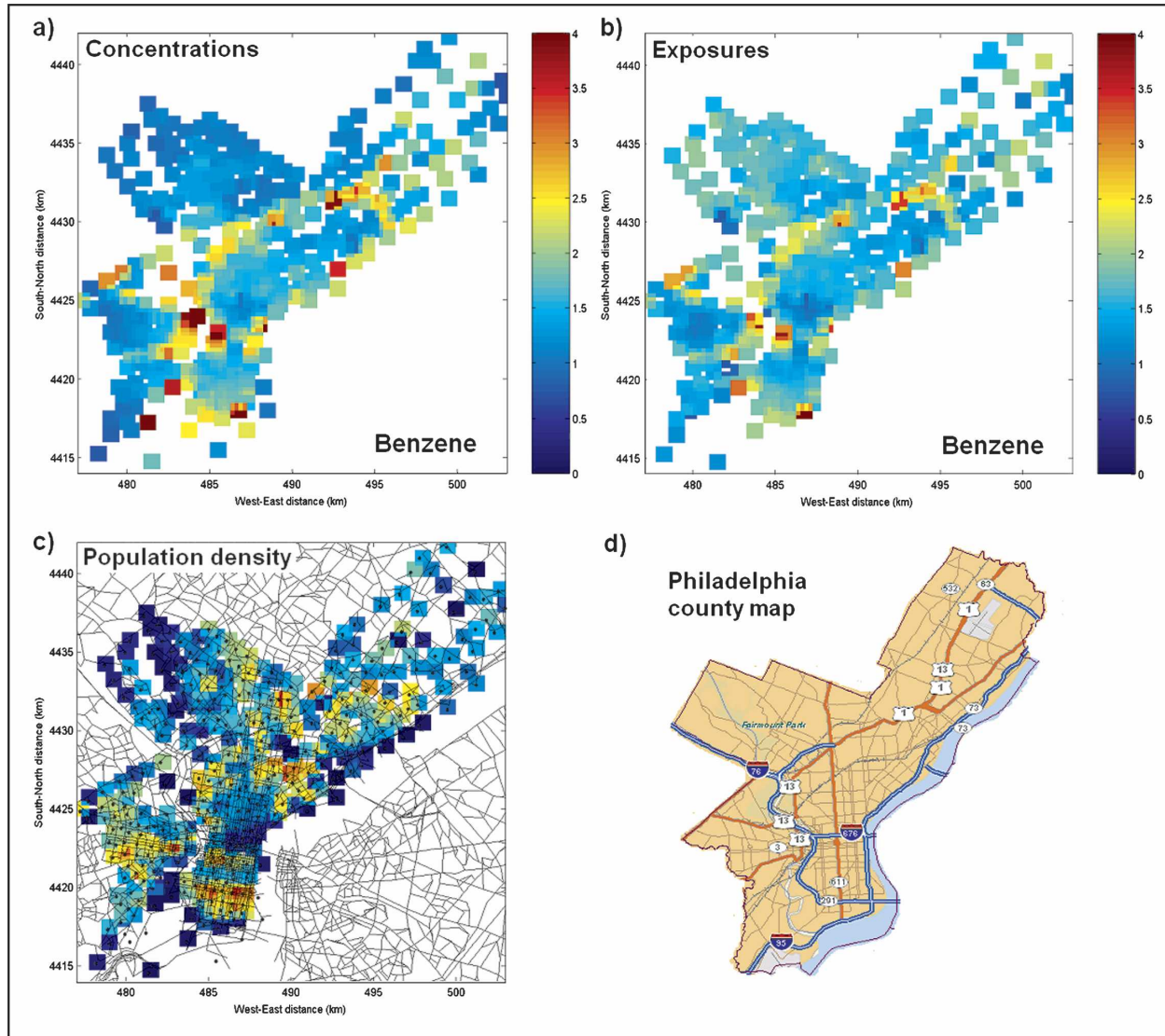


FIG. 9. Spatial maps of (a) annually averaged benzene concentrations ($\mu\text{g m}^{-3}$), (b) inhalation exposure ($\mu\text{g m}^{-3}$), (c) population density [color scale indicates normalized population density from low (blue) to high (red)], and (d) road locations in Philadelphia County.

tion fields can be enhanced. However, while grid size can be user specified to be smaller than 1 km, requirements satisfying turbulence closure assumptions for scales less than 1 km are unsatisfied for typical boundary layers (Lumley and Panofsky 1964), thus imposing 1 km as representing a lower limit to such models. There are further limitations to finescale modeling in that not only must the physical processes be appropriately characterized, but also the requisite model inputs (e.g., emissions, surface conditions, roughness, moisture availability, deposition properties). In principle, each grid cell will have a unique characteristic subgrid concentration distribution. This distribution arises from a multitude of contributing processes at smaller grid sizes

not explicitly treated, including the coupling of turbulent motions with chemical mixtures of pollutant constituents that have varying photoreactivity, in grids with differing source configurations, chemical composition, and spatial distributions. The magnitude of such variations can extend over a wide range. For achieving operational capability, it would be desirable to derive and parameterize such distributions.

A variety of approaches have been investigated to describe and parameterize SGV characteristics in CMAQ: 1) running CMAQ at sub-4-km grid sizes (order of 1 km; Ching et al. 2004), 2) modeling large buoyant stack plumes with plume-in-grid (PinG) formulations (Karamchandani et al. 2002), 3) refined local-scale

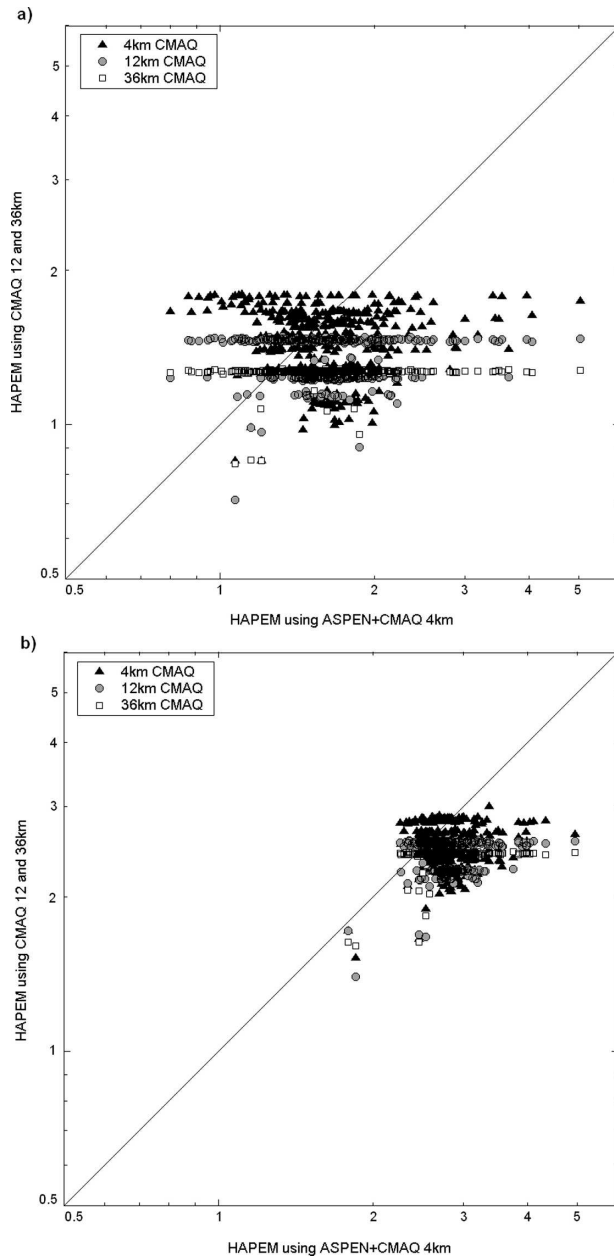


FIG. 10. Comparison of exposure concentrations ($\mu\text{g m}^{-3}$) in Philadelphia County from HAPEM using only CMAQ (12- and 36-km resolution) as an input and CMAQ (4-km resolution) enhanced with local details from ASPEN: (a) benzene and (b) formaldehyde.

dispersion modeling with fine spatially resolved emission data (Isakov and Venkatram 2006), and 4) application of coupled turbulence and photochemistry model such as LES with chemistry (LESChem; Herwehe et al. 2004) to determine the contributions from subgrid turbulence and chemistry fields.

To illustrate the development process of parameter-

izing SGV, we examine benzene, a relatively slow reacting species, and for which the major contribution to SGV is from mobile sources. In this instance, local road details and traffic activity are important. Figure 12 shows benzene concentrations for Philadelphia County simulated with ISCLT3 using annually averaged benzene emissions allocated at individual road links. Two $4 \text{ km} \times 4 \text{ km}$ grid cells, one for downtown and one for a suburban section of Philadelphia County, are highlighted. Clearly, the characteristic of the road distributions differ between the two grids. The normalized distributions of emission and concentrations for each of these two cells are shown in Fig. 13. Despite the great differences between the cells' road patterns, both cells shared a characteristically and qualitatively similar shape for their emission distributions, and also for the corresponding concentration distributions. For this case, it would appear possible to derive parametric formulations of SGV concentration distribution based on emissions distribution.

Further, given that spatial concentration distribution is dependent on distance from sources, which is certainly true in the case of roadways, we would expect to find parameterizations to describe the relationships between annually averaged concentrations C ($\mu\text{g m}^{-3}$) and two major parameters: emissions Q (kg) and source-receptor distances r (km). For example, the following empirical expression relates C and Q/r^2 ,

$$C = f\left[\sum(Q/r^2)\right]. \quad (5)$$

An analysis of model outputs for concentration at receptors has been performed at two different resolutions for this set of simulations yielding the following results:

- 1) For census-tract resolution, with parameterization based on the ASPEN model,

$$C = a_1x_1 + b_1, \quad (6)$$

where $x_1 = Q/(r^2R_t)$ or $x_1 = Q/R_t$ for very small distances (i.e., $r < 100 \text{ m}$). In this equation, C is the annually averaged benzene concentration ($\mu\text{g m}^{-3}$) at census-tract centroids, Q represents annual total benzene emissions for each source in tons (10^3 kg), r is the source-receptor distance in kilometers, R_t is an effective radius of a census tract in kilometers, a_1 is a nondimensional coefficient = 0.0753, and b_1 is a constant parameter = 0.271 ($\mu\text{g m}^{-3}$). These empirical coefficients need to be determined by the user for specific applications.

- 2) For finer resolution (200-m spacing), with parameterization based on ISCLT3 model results,

$$C = a_2x_2 + b_2, \quad (7)$$

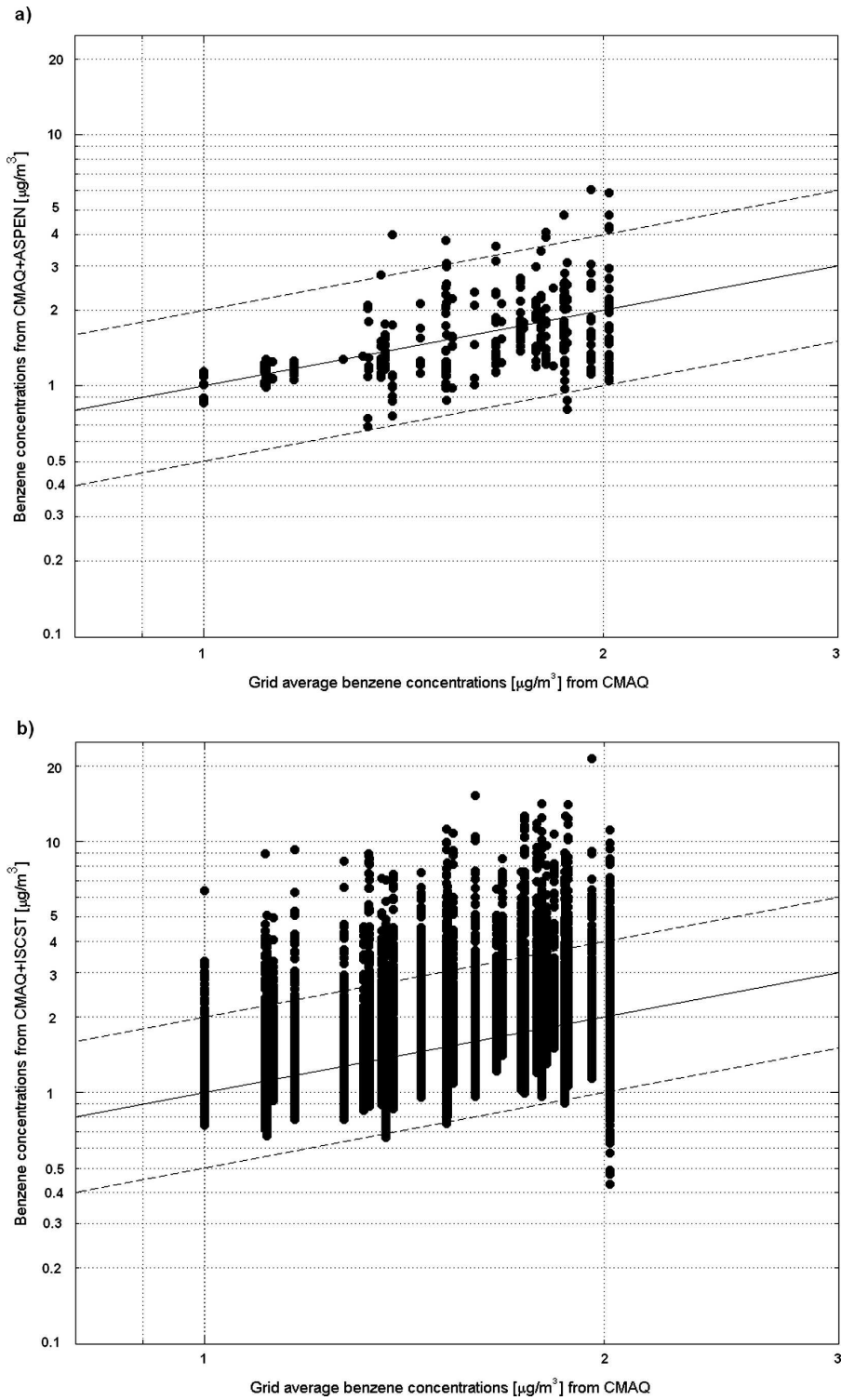


FIG. 11. Variability of benzene concentrations at $4 \text{ km} \times 4 \text{ km}$ grid cells derived from CMAQ alone plus (a) ASPEN and (b) ISCST3.

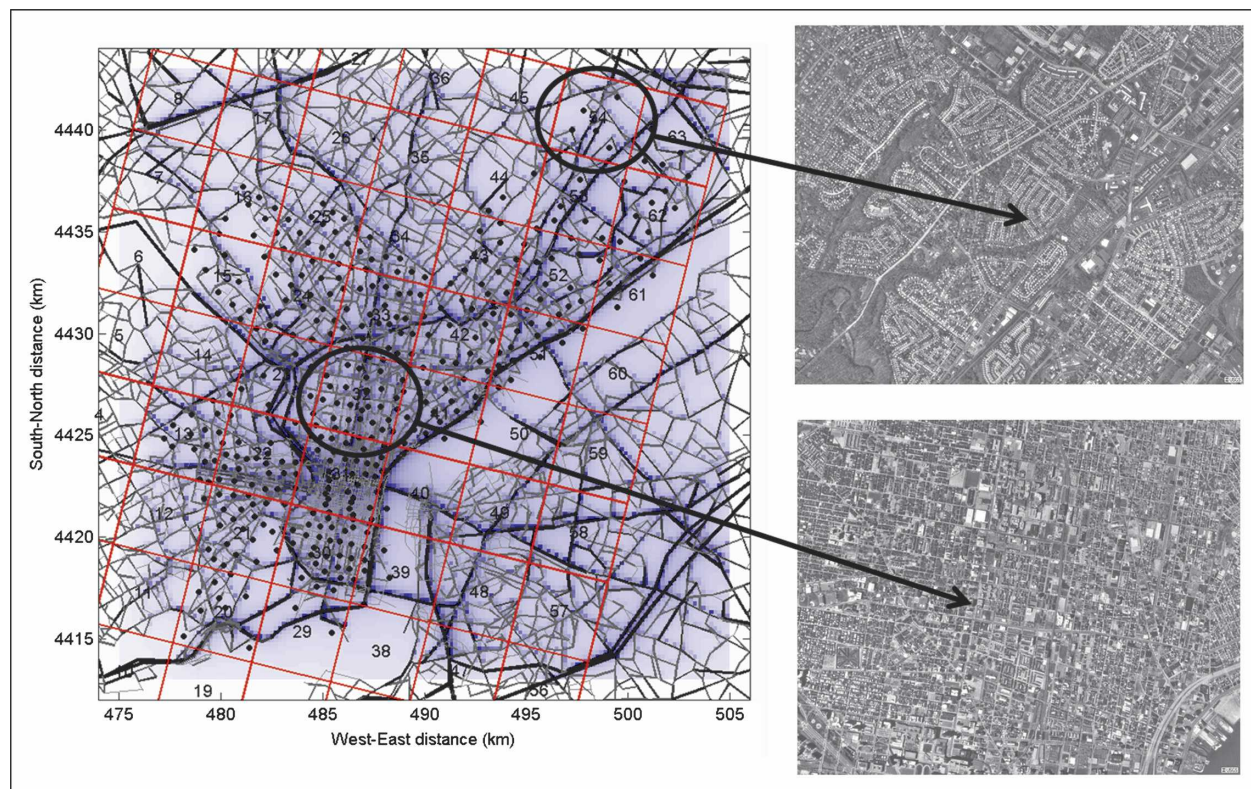


FIG. 12. Philadelphia (left) benzene concentrations [color scale indicates normalized concentrations from low (light blue) to high (dark blue)] and (right) aerial photos for (bottom) downtown and (top) suburbs.

where $x_2 = Q/r^2$ for distances larger than 50 m, or r is set to a default value of 50 m if $r < 50$ m. In this equation, C is the annually averaged benzene concentration ($\mu\text{g m}^{-3}$) at each receptor, Q represents annual total benzene emissions for each source in tons (10^3 kg), r is the source–receptor distance in kilometers, a_2 is a coefficient = 0.0204 with dimensions of length $^{-1}$, and b_2 is another parameter = 0.0204 with dimensions of micrograms per cubic meter.

The comparison of concentrations computed using the modeled versus the parameterized form given above is shown in Fig. 14 for ASPEN and Fig. 15 for ISCLT3. It is clear that both sets of parameterizations compare favorably to each of the different models, and for each of the two levels of resolution in receptor density.

Although further study is needed to support the development of parametric formulations, it seems, at least for benzene, that the possibility of deriving relatively simple parameterizations for the SGV component to augment background values from grid models such as CMAQ is encouraging. However, we are aware that the desired parameterizations for more reactive pollutant

species or for more complex source contributions may be more complex than shown above. Exploratory investigations have shown that the SGV distributions for more reactive species may be relatively complex, and the parametric formulations that describe the SGV for such chemical species remains a goal (Herwehe et al. 2004).

Given the ability to model (or parameterize) the SGV distribution, there is the subsequent and final step needed to appropriately incorporate SGV information with the gridded fields. A preliminary methodology has been proposed by Ching et al. (2006b). The approach takes the product of the modeled concentration by some weighting function of the SGV for each individual grid cell, yielding outputs that represent both lower and upper bounds to the gridded concentration fields. It is anticipated that the weighting function to be applied will be application dependent. In the case of exposure modeling, the weighting function might be a 95th percentile of the SGV distribution for population exposure if the purpose is for the determination of the risk to the most susceptible portions of the population. The choice of such a metric has cost–benefit consequences, and we suggest that it will need the input and consideration of the health sciences and impacted communities.

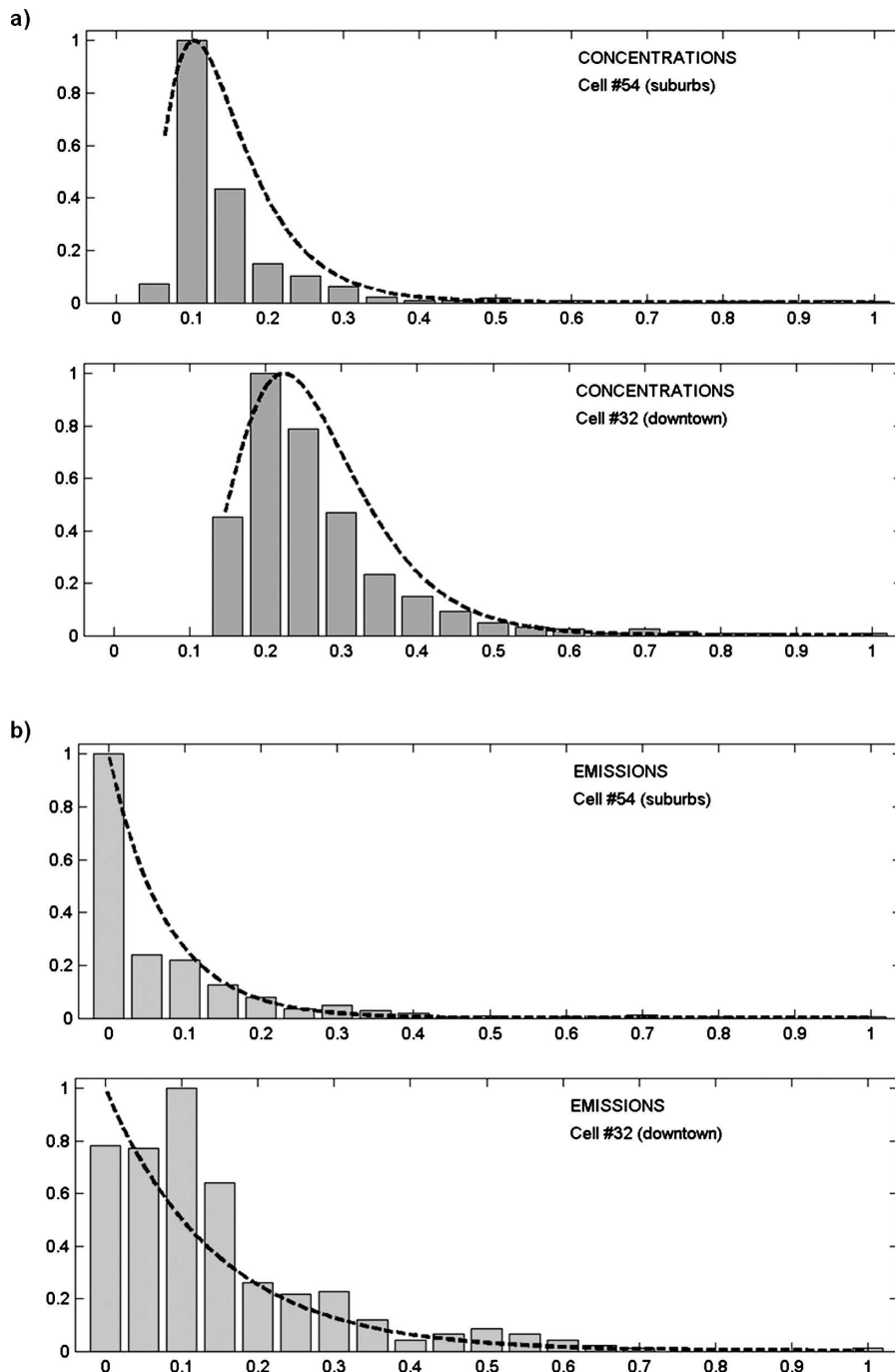


FIG. 13. Normalized distributions of benzene (a) concentrations and (b) emissions for two different grid cells: (top) suburbs and (bottom) downtown.

4. Summary

We used HAPEM to estimate benzene and formaldehyde inhalation exposures using ambient annually averaged concentrations predicted by CMAQ to investigate how within-grid variability due to emissions char-

acterizations affects exposure estimates. It is a CMAQ property that its quantitative outcome is dependent on the selection of the grid cell size. Regardless of the finest grid size chosen, model output is a single value for each grid cell for the hour of simulation. There are additional but unresolved subgrid variability details.

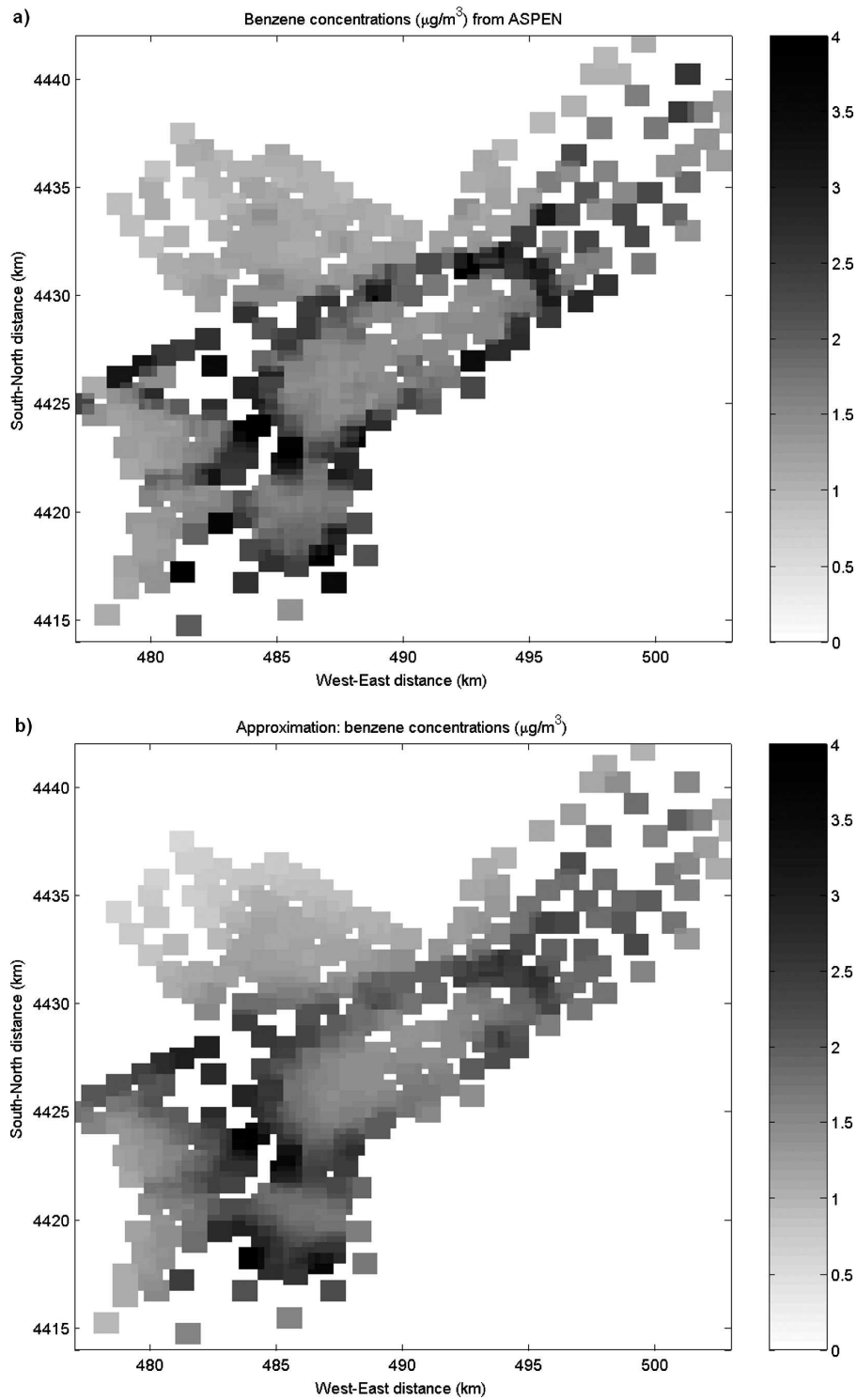


FIG. 14. Annually averaged benzene concentrations (a) from ASPEN and (b) using parametric formulations.

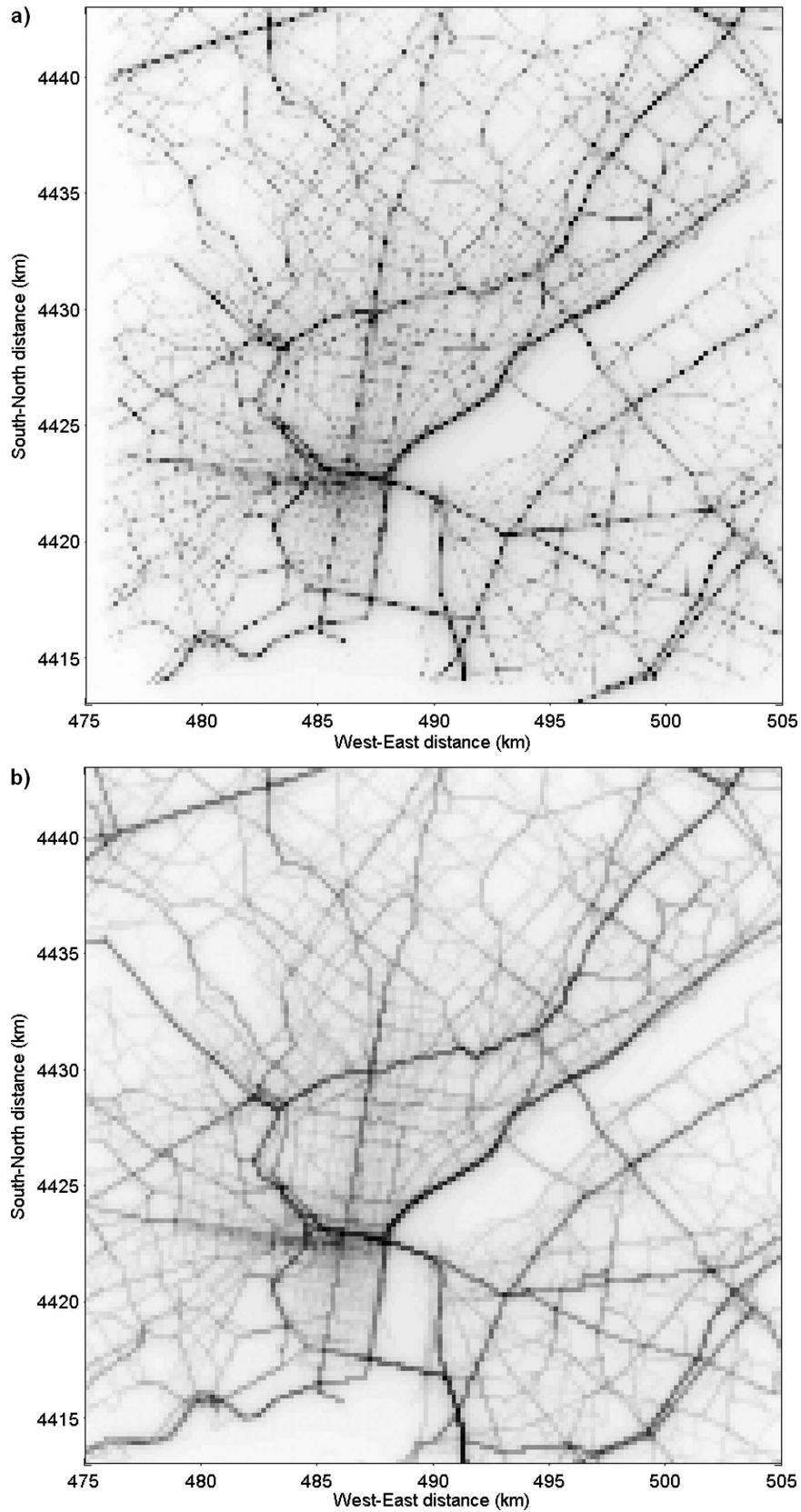


FIG. 15. Annually averaged benzene concentrations from ISCST3 using (a) link-based inventory and (b) parametric formulations. The scale indicates normalized concentrations from low (light gray) to high (dark gray).

Obtaining this information requires introducing other techniques, such as modeling large buoyant stack plumes with plume-in-grid formulations (Karamchandani et al. 2002) or combining a regional grid model with local plume/puff models (Touma et al. 2006).

In this study, we used local-scale plume dispersion models to estimate the subgrid variability of annually averaged benzene and formaldehyde concentration values within CMAQ grid cells for a modeling domain centered on Philadelphia, Pennsylvania. In this approach, local details can be provided by dispersion models such as AERMOD, ASPEN, or ISCST, and regional background and contributions from chemically reactive pollutants can be obtained from chemical transport models (e.g., CMAQ).

Significant increases (more than a factor of 2) in maximum exposure impacts were seen in the exposure estimates in comparison with exposure estimates based on grid-averaged concentration values. Although results examine only one source of subgrid variability, namely, the discrete location and distribution of emissions, they do suggest the importance and value of developing improved characterizations of subgrid concentration variability for use in air toxics exposure assessments.

Given the significant impact of SGV on estimating maximum exposures, we proposed a two-step process to incorporate SGV into the grid model simulations. Because SGV can be described in terms of concentration distributions (e.g., Herwehe et al. 2004; Ching et al. 2006a), the first step involved the derivation and ultimately the development of parametric forms of the distribution for each individual pollutant species. Subsequently, a method for incorporating the distribution or parameterizations with CMAQ must be introduced. Preliminary results for a special case of benzene primarily from mobile sources were presented, and the outcomes were encouraging. Further investigations and additional developments along these lines are necessary to handle the more general situations, for example, SGVs from the complex distribution of a variety of source sectors as well as for the more reactive pollutant species. With this capability, operational links to other exposure models, including the Stochastic Human Exposure to Dose Simulation (SHEDS; Graham and Burke 2003) model, are readily possible. SHEDS requires complementary information on the variability of the input concentration fields as stochastic inputs. The parametric form of the SGV distributions can easily be formatted for that purpose.

Acknowledgments. The research presented here was performed under the Memorandum of Understanding

between the U.S. Environmental Protection Agency (EPA) and the U.S. Department of Commerce's National Oceanic and Atmospheric Administration (NOAA) under Agreement Number DW13921548. This work constitutes a contribution to the NOAA Air Quality Program. Although it has been reviewed by EPA and NOAA and approved for publication, it does not necessarily reflect their policies or views. The authors thank Mr. Al Cimorelli of the U.S. EPA Region 3 for helpful discussions and valuable comments.

REFERENCES

- Byun, D. W., and K. Schere, 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**, 51–77.
- Ching, J., T. Pierce, T. Palma, W. Hutzell, R. Tang, A. Cimorelli, and J. Herwehe, 2004: Linking air toxic concentrations from CMAQ to the HAPEM5 exposure model at neighborhood scales for the Philadelphia area. Preprints, *Fifth Symp. on the Urban Environment*, Vancouver, BC, Canada, Amer. Meteor. Soc., CD-ROM, J4.4.
- , J. Herwehe, and J. Swall, 2006a: On joint deterministic grid modeling and sub grid variability conceptual framework for model evaluation. *Atmos. Environ.*, **40**, 4936–4945.
- , V. Isakov, M. Majeed, and J. S. Irwin, 2006b: An approach for incorporating sub-grid variability information into air quality. Preprints, *14th Joint Conf. on the Applications of Air Pollution Meteorology*, Atlanta, GA, Amer. Meteor. Soc., CD-ROM, 2.1.
- Cimorelli, A. J., and Coauthors, 2005: AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. *J. Appl. Meteor.*, **44**, 682–693.
- Georgopoulos, P. G., and P. J. Lioy, 2006: From theoretical aspects of human exposure and dose assessment to computational model implementation: The Modeling Environment for Total Risk studies (MENTOR). *J. Toxicol. Environ. Health*, **9B**, 457–483.
- Graham, S. E., and J. M. Burke, 2003: Microenvironmental exposures to benzene: A critical review and probabilistic model input distribution development. Environmental Protection Agency Rep. EPA/600/J-03/008, Office of Research and Development, Research Triangle Park, NC, 54 pp.
- Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech. Note NCAR/TN-398, 122 pp.
- Herwehe, J. A., 2000: A numerical study of the effects of large eddies on the trace gas measurements and photochemistry in the convective boundary layer. Ph.D. dissertation, University of Alabama in Huntsville, 242 pp. [Available from University of Alabama in Huntsville, Atmospheric Science Department, Huntsville, AL 35899.]
- , J. K. S. Ching, and J. Swall, 2004: Quantifying subgrid pollutant variability in Eulerian air quality models. Preprints, *13th Conf. on the Applications of Air Pollution Meteorology with the Air & Waste Management Association*, Vancouver, BC, Canada, Amer. Meteor. Soc., CD-ROM, 7.5.
- Isakov, V., and A. Venkatram, 2006: Resolving neighborhood

- scale in air toxics modeling: A case study in Wilmington, California. *J. Air Waste Manage. Assoc.*, **56**, 559–568.
- Jerrett, M., A. Arain, P. Kanaroglou, B. Becjerman, D. Potoglou, T. Sahsuvaroglu, J. Morrison, and C. Giovis, 2005: A review and evaluation of intraurban air pollution exposure models. *J. Expo. Anal. Environ. Epidemiol.*, **15**, 185–204.
- Karamchandani, P., C. Seigneur, K. Vijayaraghavan, and S.-Y. Wu, 2002: Development and application of a state-of-the-science plume-in-grid model. *J. Geophys. Res.*, **107**, 4403, doi:10.1029/2002JD002123.
- Luecken, D. J., W. T. Hutzell, and G. J. Gipson, 2006: Development and analysis of air quality modeling simulations for hazardous air pollutants. *Atmos. Environ.*, **40**, 5087–5096.
- Lumley, J. L., and H. A. Panofsky, 1964: *The Structure of Atmospheric Turbulence*. Monographs and Texts in Physics and Astronomy, Vol. 12, Interscience, 239 pp.
- Miao, J. F., D. Chen, and K. Wyser, 2006: Modelling subgrid scale dry deposition velocity of O₃ over the Swedish west coast with MM5-PX model. *Atmos. Environ.*, **40**, 415–429.
- Odman, M. T., and A. G. Russell, 1991: Multiscale modeling of pollutant transport and chemistry. *J. Geophys. Res.*, **96**, 7363–7370.
- Spicer, C. W., and Coauthors, 1996: Variability of hazardous air pollutants in an urban area. *Atmos. Environ.*, **30**, 3443–3456.
- Touma, J. S., V. Isakov, J. Ching, and C. Seigneur, 2006: Air quality modeling of hazardous pollutants: Current status and future directions. *J. Air Waste Manage. Assoc.*, **56**, 547–558.
- U.S. Environmental Protection Agency, 1995: User's guide for the Industrial Source Complex (ISC3) dispersion models. Office of Air Quality Planning and Standards Rep. EPA-454/B-95-003b, Vol. 2, Office of Air Quality Planning and Standards, Research Triangle Park, NC, 120 pp.
- , 2000: User's guide for the Assessment System for Population Exposure Nationwide (ASPEN, version 1.1) Model. Office of Air Quality Planning and Standards Rep. EPA-454/R-00-017, 108 pp.
- , cited 2004: 1999 national emission inventory documentation and data—Final version 3.0. [Available online at <http://www.epa.gov/ttn/chief/net/>.]
- , cited 2005: Human exposure modeling—Hazardous Air Pollutant Exposure Model (HAPEM). [Available online at http://www.epa.gov/ttn/fera/human_hapem.html.]
- Weijers, E. P., A. Y. Khlystov, G. P. A. Kos, and J. W. Erisman, 2004: Variability of particulate matter concentrations along roads and motorways determined by a moving measurement unit. *Atmos. Environ.*, **38**, 2993–3002.