2005 CMAS Workshop, Session 5: Urban/Regional Model Applications and Evaluation

INCORPORATING SUB-GRID VARIABILITY CONCENTRATION DISTRIBUTIONS WITH CMAQ

Jason Ching* [(1)], Vlad Isakov [(1)], Jerry Herwehe[(1)], Mohammed A Majeed[(2)]

[1]Atmospheric Sciences Modeling Division, ARL, NOAA, RTP, NC
[2]Delaware Dept. of Natural Resources & Environmental Control, New Castle, DE

1. INTRODUCTION

Advanced air quality models (e.g., Community Multiscale Air Quality model, CMAQ) are formulated to predict outcomes in terms of gridded concentration fields derived from gridded pollutant emission sources into a dispersive and chemically reactive atmospheric environment. Such models recognize and are able to treat and track the evolution of each and all pollutant species in an environment that is in constant state of change. It is also a property of such models that their quantitative outcomes are dependent on the selection of the model grid size. However, methods are available to refine the spatial definitions of the modeled outcome range using adaptable grid meshes or of "nested" sub domains with decreasing grid sizes. Methods implemented are dependent on the nature of the application and computational resources. Unfortunately, these methodologies have both computational and physical constraints that limit the spatial refinements in grid size to about 1 km grid size. While grid size can be chosen to be smaller than one km, the stationarity requirements for satisfying turbulence closure assumption for scales less that one km are unsatisfied for typical boundary layers (Lumley and Panofsky, 1964). Irregardless of the finest grid size chosen for the simulation, the nature of grid modeling is that there is in principle, always within-grid or sub-grid spatial variability (hereinafter, SGV) Ching et al., 2004a. In general, grid modeling has been performed without explicit model provisions and representation of SGV. This paper provides simple means by which SGV concentration information is introduced along with the gridded model fields.

The first step is to model the SGV. To do this, we recognize a variety of processes that can contribute to SGV in CMAQ. These include: (a) unresolved meteorological processes, (b) primary source contributions from within and from neighboring grids that have not achieved fully dispersed status as is required in the emissions processing of CMAQ, and (c) from heterogeneities arising from within-cell photochemistry and turbulent interactions. For primary sources, our approach utilized local scale modeling methods set up to generate the gridded SGVs. Preliminary results for several statistical parameters are presented for the Philadelphia domain. For more reactive species, we utilized fine (1 km grid size) scale simulations to derive the SGV statistics for 12 km grid size. The treatment of SGV at grid sizes smaller than 1 km is in principle possible using finer scale modeling methods such as a coupled LES-photochemistry model (Herwehe, 2000). However, since the purpose of this paper is to illustrate introducing SGV to grid models, we limit the SGV from photochemical processes to the 1 km CMAQ simulations.

*In partnership with U.S. Environmental Protection Agency, AMD/NERL, 109 TW Alexander Dr, RTP, NC 27709 (MD E243-04). Corresponding author's email address: ching.jason@epa.gov

2. RATIONALE

There are a variety of issues for which adjunct SGV information to models such as CMAQ are both pertinent and useful. This include air toxics exposure assessment and model evaluation studies. For air toxics exposure assessments, adding SGVs to CMAQ can provide additional information about the distribution of the possible range of concentrations that at risk and susceptible populations within any grid cell might be exposed to. Currently, the USEPA is developing an air toxics version of CMAQ (Luecken et. al., 2005) and implementing its linkage to the HAPEM5 (Hazardous Air Pollutant Exposure Model, Version 5) www.epa.gov/ttn/atw/nata/modelexp.html. Simulations of air toxics have been performed (Ching et al., 2004) to provide annual concentration fields for several grid sizes, including 36, 12 and 4 km. These results were used for subsequent exposure assessments. However, studies have already indicated that for many toxic species, SGV below 4 km grid sizes can be very large.

This situation could seriously underestimate human exposure given significant peak within-grid concentration values. Additionally, SGV may also provide a basis for constructing valid statistical design for comparing what is simulated to what is observed (Ching et al., 2005). We now present a template for introducing SGV with the gridded CMAQ outputs.

3. APPROACH

There is little, if any, experience with linking SGV to grid model output. We now describe a simple method for introducing SGV to CMAQ outputs:

Let C_g and C_{SGV} be defined as the CMAQ gridded simulations and SGV outputs, respectively. Thus

$$C_j = C_g + C_{SGV} \quad (1)$$

where C_j describes the joint fields (in principle, negative values for C_{SGV} could also be considered). We now assume the distribution of SGV to be represented by some distribution function, or parameter, DF. Subsequently, we designate DF as a normalized form of SGV, e.g.,

$$DF = C_{SGV} / C_g \quad (2)$$

Then:

$$C_j = C_g(1+DF) \quad (3)$$

This yields the SGV Adjusted Concentration (SAC) or

$$SAC_j = C_j / C_g = 1+DF \quad (4)$$

With this nomenclature, j can refer to different descriptors. One might chose to describe SAC in statistical terms, such as the standard deviation to its grid mean (or in normalized form, the coefficient of variation, COV) or by some arbitrary percentile (e.g., 90^{th} , 95^{th}, etc) values. One can also explore using some normalized parameters of the distribution, e.g., peak value or range of values to its grid mean. The next section will illustrate by examples this means of introducing SGVs with the CMAQ in this manner.

While we have begun to produce and study the distributions from histograms that describe SGVs, we do not as yet have a convenient method for introducing these distribution functions directly with CMAQ. We anticipate that if descriptions of SGV (Herwehe et.al., 2004) were possible and available as generalized distribution functions, it would be operationally feasible to add the robust descriptor of SGVs to the CMAQ simulations.

4. PRELIMINARY RESULTS

In this section, we illustrate using a variety of examples results applying the method described above using different statistical measures for DF. The first case is for exposure modeling using annual averaged CMAQ results and results are shown for benzene, a relatively inert pollutant; the second case is pertinent to acute exposure assessments which involve CMAQ simulations for more reactive pollutant species (e.g., formaldehyde) on monthly and episodic periods.

4.1 CASE 1: Air toxics exposure assessment is typically performed on an annual basis. As a pilot study CMAQ was run for the year 2001, and results introduced into the HAPEM5 for the exposure assessment (Ching et al., (2004). The annual results for CMAQ are shown in Figure 1a (only grids for Philadelphia County are displayed). We chose benzene for this illustration; the mobile emissions are the primary source of this pollutant in this application. The roadway network is indicated as part of Figure 1. Clearly, the distribution of the roadways varies greatly in density and pattern across the county. Local scale modeling was applied to resolve concentrations fields near roadways. Such fine scale analysis requires more detailed information on emission sources. In this study, we used a fine-scale emissions inventory, when emissions from mobile sources were allocated to individual road links (road segments). We used the EPA's regulatory dispersion model ISCST3 to simulate annual average benzene concentrations at 200 m receptor resolution. This information was used to estimate statistical properties of concentration distributions within CMAQ grid cells. Figure 1b and c shows the standard deviation and the highest value (maximum or peak). The 95^{th} percentile of the distribution for each 4 km cell were obtained from this set of results and is the type of information that the HAPEM5 accepts to represent sub-grid variability. The results and descriptive statistics on the SGV pertinent to this paper are shown as 1+DF or SAC (eq. 4) were prepared for each of the CMAQ grid cells. Figure 1 (d,e and f) shows results for SGV represented by its standard deviation (StdDev), peak (maximum) and the 95^{th} percentile value of the SGV distribution. The values of mean and both indicators of SGV were greatest in the urban center and along the major interstate highways as expected. On an annual basis, the SGV as represented by peak values were typically almost an order of magnitude greater than the StdDev values. When displayed as SACs, we observe that in general, each of the various SACs were (a) qualitatively similar across the County and (b) the SAC based on peak was again nearly an order of magnitude larger than the SAC for StdDev. However, SAC values were

also noticeably larger in areas both to the northwest and south of the downtown areas. Additionally, we also notice the broad area of low SACs values in the more suburban areas to the north of the city center. SAC represented by StdDev and the 95 percentile were also relatively large along the major SW to NE highway corridor; this was not observed for SACs for peak concentrations. While not shown, we have observed that the shape of the histogram for benzene varies from cell to cell, providing evidence of the heterogeneity in the distribution of emission from roadways and road links. We discern that the spatial differences in the SACs are a reflection of the fact that the characteristic shapes of the histogram distributions differ across the modeling domain. This is consistent with modeling results reported earlier by Herwehe et al., 2004.

Given that HAPEM5 can incorporate SGVs as SACs (e.g., 95th percentiles), the SAC values shown in Figure 1f range from 1 to 5 times the gridded concentration for benzene. The significance of this result is that the exposure assessment using HAPEM5 will increase the modeled population risk to exposure to benzene by the SAC factor. The approach taken here can provide information showing the location of areas where the adjusted enhanced exposures are located.

4.2 CASE 2: Preliminary results are shown in Figure 2 for formaldehyde, a photochemically reactive species. Short term exposure modeling requires CMAQ simulations on seasonal and episodic periods. Whereas, for the inert species (Case 1) where local scale modeling can be used, information from finer-scale CMAQ simulations is utilized to derive the SGV within a coarse grid. For illustrative purposes we utilize information from 1 km grid-scale runs to define the SGV within a 12 km grid scale modeling domain. In other words, information from 144 1-km grid cells will represent the SGV in a 12 km grid cell.

Preliminary results are from a set of CMAQ runs for July 2001, the modeling at 1 km grid size for the domain covering the state of Delaware were nested from 4 and 12 km runs used for Case 1 discussed earlier. For this discussion, we have limited the examination of SACs to the cell's coefficient of variation (COV). For this example, the normalization of the hourly StdDev for each grid is with the monthly averaged mean concentrations for each grid (Figure 2a shows the monthly average formaldehyde concentrations for the Delaware portion of the 12 km Philadelphia modeling domain). Alternatively the normalization could have been chosen to be the corresponding hour value of the StdDev. The SGV is obtained for each hour for any 12 km grid and is determined from 144 cells of the 1 km grid-scale CMAQ simulations. Figure 2b show the SACs for an arbitrary hour and typical summer day; 20 UTZ on July 15, 2001 for formaldehyde.

The monthly mean concentration fields are relatively smooth. The highest value of HCHO is in the northwest part of the domain which is mostly urbanized and along the major SW to NE traffic corridor connecting Baltimore and Philadelphia. The range of SAC values (based on the COV indicator) from Figure 2b is from 1 to 2, not unlike the results for benzene. However, and interestingly, the lowest SAC value for HCHO was in the more urbanized area in the NW section of the domain. This result is being investigated further.

5. SUMMARY AND DISCUSSION

We have shown examples to illustrate a simple means to introduce SGV as outputs with CMAQ results. This is embodied in equation (1). However, it should be noted that the utility of (1) in actual applications of the SGV information is not addressed here. The preliminary results shown include pollutants ranging from chemically inert to reactive ones. The SGV for the relatively inert species were obtained with the use of local scale modeling technique. For the reactive species the fine scale details were obtained using CMAQ at fine grid mesh size; however, we note that these results, in principle, will underestimate the variability since additional SGV for such species exist at even finer resolutions. We have utilized various statistical descriptors of the SGV distributions including their normalized standard deviation, peak, and percentile values. We find that these simple ratios vary greatly across a modeling domain in urban areas.

These early findings are already of significance. Results using percentiles obtained from the SGV distributions show SAC values considerably larger than the grid simulations and its immediate impact will be for a greater risk potential to populations exposed to air pollutants.

SGV modeling is, itself, at the research development stage. Much more model research and community dialogue will be needed to formulate the use of SGVs in practical modeling applications. In terms of development, future investigations are being planned along several lines. The modeling of SGV for reactive species will require modeling methods such as coupled Large-Eddy Simulations with photochemistry models (e.g., LESCHEM, Herwehe, 2000) to provide a greater range of SGVs than can be obtained using fine scale CMAQ modeling alone.

We also note that SGV is present and applicable to model simulations of meteorology and transport fields as well. For this, a model (QUIC) (Williams, 2004) is being implemented to support modeling studies for

Wilmington, DE to generate the SGV transport and wind fields.

Also, given the ability to model the SGVs, we feel it is important to pursue the goal of deriving parameterizations from gridded distributions to provide a fuller expression of DF and SACs for practical applications (Herwehe, et al., 2004). Alternatively, investigations are underway using statistical techniques to determining SGVs, which are based upon developing conditional representations of fine scale modeling given coarse grid resolution outputs.

Disclaimer: *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) and 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.*

REFERENCES

Ching, J., S. DuPont, J. Herwehe, T. Otte, A. Lacser, D. Byun, and R. Tang, 2004a: Air quality modeling at coarse-to-fine scales in urban areas. *Preprints, Sixth Conference on Atmospheric Chemistry: Air Quality in Megacities, January 11-15, 2004, Seattle, Washington*. American Meteorological Society, Boston.

Ching, J.K.S., T.Pierce**,** T.Palma, W.Hutzell, *R.Tang*, A.Cimorelli and J.Herwehe, (2004b: Application of fine scale air toxics modeling with CMAQ to HAPEM5, Preprints, *3rd Annual Models-3 User's Conference*, October 18-20, 2004, Chapel Hill, NC Community Modeling and Analysis System, CD ROM 5.1(2004).

Ching, J., J.Herwehe and J.Swall, 2005: On joint deterministic grid modeling and sub-grid variability conceptual framework for model evaluation, Submitted to Atmos. Env .

Herwehe, J.A., 2000: A numerical study of the effects of large eddies on trace gas measurements and photochemistry in the convective boundary layer. PhD dissertation, Department of Atmospheric Sciences, University of Alabama in Huntsville

Herwehe, J.A., J.K.S.Ching and J. L. Swall (2004): Stochastic description of subgrid pollutant variability in CMAQ. Preprints, *3rd Annual Models-3 User's Conference*, October 18-20, 2004, Chapel Hill, NC Community Modeling and Analysis System, CD ROM 5.3 (2004).

Luecken, D.J., W.T. Hutzell, and G.J. Gipson, 2005: Development and analysis of air quality modeling simulations for hazardous air pollutants. Atmospheric Environment (in review).

Lumley, J.L., and H.A. Panofsky, 1964: The Structure of Atmospheric Turbulence. Monographs and Texts in Physics and Astronomy, Vol XII. Interscience Publishers, John Wiley and Sons, New York, 239 pg.

Williams, M. D., 2004: The quick urban & industrial complex QUIC model for urban dispersion applications. LANL report LA-UR-04-3275, http://library.lanl.gov/cgi-bin/newrepslist.pl?20050228

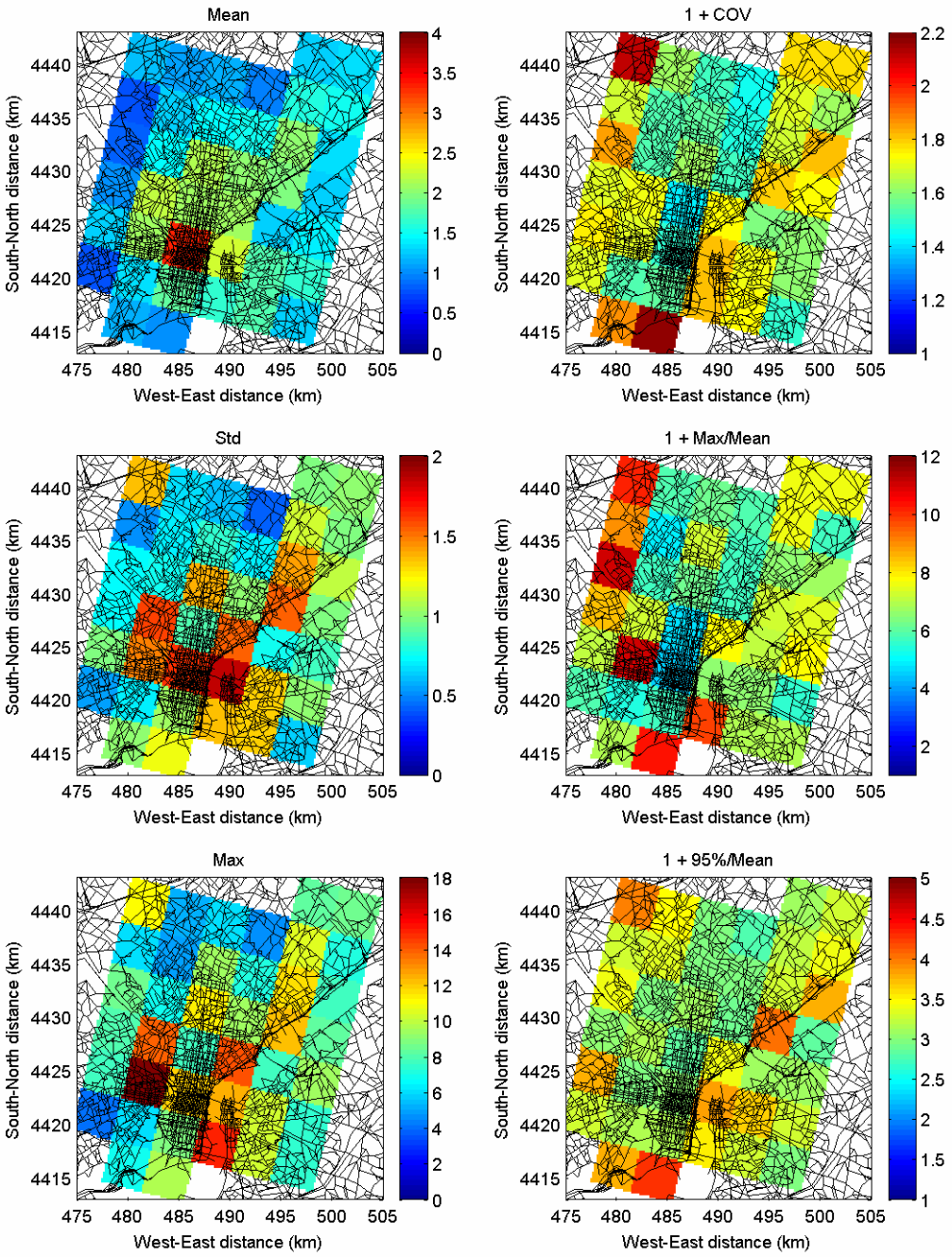

Figure 1. Benzene simulations from CMAQ for annual (2001) simulations using 4 km grid size. The SACs (Std (is the standard deviation), peak and 95th percentile) values are derived from ISCST3. The concentration units are in ug/m3. The SACs are dimensionless. The road links are shown as background.

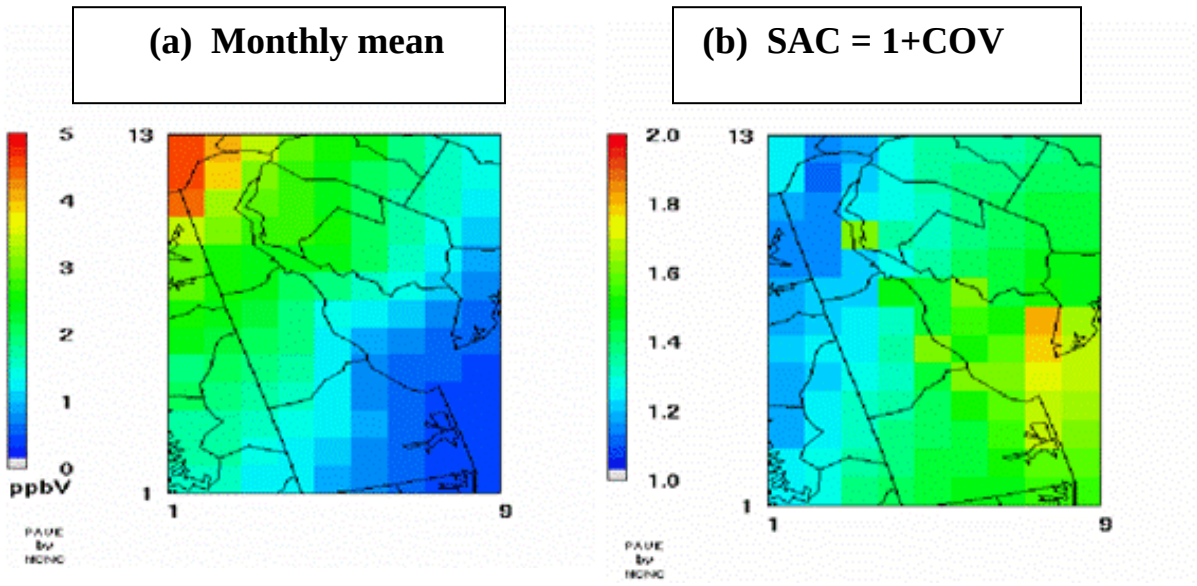

Figure 2: CMAQ simulation results for at 12 km grid sizes formaldehyde. Monthly averages for July, 2001 are shown in Figure 2a. SACs, in terms of 1+ COV are shown in Figure 2b for July 15, 2001 at 2000 UTC. The SAC results for each 12 km grid were derived from 144 1 km cells.