Fast Equivalent Operational Model of Tropospheric Alkane Photochemistry

S. W. Wang
Environmental and Occupational Health Sciences Institute, UMDNJ—R. W. Johnson Medical School and Rutgers University, Piscataway, NJ 08854

S. Balakrishnan
Dept. of Computer Science, Rutgers University, Piscataway, NJ 08854

P. Georgopoulos
Environmental and Occupational Health Sciences Institute, UMDNJ-R. W. Johnson Medical School and Rutgers University, Piscataway, NJ 08854

DOI 10.1002/aic.10431
Published online February 28, 2005 in Wiley InterScience (www.interscience.wiley.com).

A mathematical "kinetic lumping" approach, the direct constrained approximate lumping (DCAL) method, for developing a condensed chemical mechanism of complex tropospheric alkane photochemistry was presented in previous work. In the present work, the high dimensional model representation (HDMR) method is applied to further accelerate the chemical kinetic calculations of the Alkane/DCAL mechanism from our previous work. An efficient HDMR is based on expressing a chemical species concentration at a given reaction time as an expansion of correlated functions consisting of the initial chemical species concentrations. An HDMR expansion calculates efficiently the output species concentrations at any given reaction time, employing very rapid and stable algebraic manipulations. The HDMR estimates are shown to be almost identical to those derived from solving the full tropospheric alkane photochemistry through conventional (for example, Gear-type) chemistry solvers over wide ranges of initial conditions, while being more than 1,000 times faster. © 2005 American Institute of Chemical Engineers *AIChE J,* 51: 1297–1303, 2005

Keywords: kinetic mechanism reduction, high dimensional model representation, atmospheric chemistry mechanisms

Introduction

Gas-phase reaction mechanisms for the atmospheric organic/NO_x/O_3 chemistry are too long to be incorporated in atmospheric models intended for routine use. For this reason, reduced or lumped mechanisms (for example, CBM-IV (Gery et al., 1989), SAPRC93 (Carter, 1988), and RADM (Stockwell et al., 1990) have been developed as systematic approximations of the atmospheric chemistry and are routinely used in 3-D (three-dimensional) air quality models. The limitations of these lumped mechanisms are often associated with inaccuracies due to the fact that the lumped mechanisms have typically been optimized to fit certain atmospheric chemical conditions. In our previous article, we have shown that the Alkane/DCAL mechanisms can overcome the above limitations to provide accurate multispecies time-concentration profiles over reasonable wide ranges of atmospheric chemical conditions in a box model study of alkane photochemistry, while achieving the similar computational efficiency as the alkane portions of CBM-IV and

Correspondence concerning this article should be addressed to P. Georgopoulos at panosg@fildelio.rutgers.edu.

SAPRC93 mechanisms. In this article, we present an application of the more efficient high dimensional model representation (HDMR) method to further accelerate the chemical kinetic calculations for the Alkane/DCAL mechanism of complex alkane photochemistry.

Although the DCAL method can successfully condense the 30 explicit alkane species into 3 lumped species, the chemical kinetic equations of the Alkane/DCAL mechanism are still "stiff" to solve. To circumvent this computational difficulty, the alternative approach is to parameterize the reaction mechanisms by fitting the results obtained from solving the chemistry rate equations "off-line" with explicit algebraic equations describing the relationships between output chemical species concentrations and input variables, such as initial chemical species concentrations, temperature, and photodissociation constants (J values). The explicit algebraic equations may then be employed for the chemical kinetics component of the 3-D calculations (Klonecki and Levy, 1997; Spivakovsky et al., 1990; Turányi, 1994).

Traditionally, parameterized polynomial expansions or interpolative look-up tables have been used for the parameterization of reaction mechanisms. Spivakovsky (Spivakovsky et al., 1990) used high-order polynomials to express the input-output chemical kinetic response through least-squares fitting. Turányi (Turányi, 1994) extended the approach of Spivakovsky et al. by expressing the chemical kinetic input-output relation as an expansion in orthogonal polynomials. Klonecki and Levy (Klonecki and Levy, 1997) used standard high-dimensional look-up tables to perform chemical kinetic calculations of $CO–CH_4–NO_y–H_2O$ chemistry in 3-D global chemistry transport model (GCTM) ozone simulations. One major problem associated with these approaches is that without the possibility of simplification, the number of times the chemistry rate equations need to be solved grows exponentially with respect to the dimension of the system (that is, the number of chemical species). Therefore, attempts to pursue these traditional approaches would be prohibitively expensive in high-dimensional chemical systems.

The high dimensional model representation (HDMR) method (demonstrated through past applications in other areas (Rabitz and Shim, 1999; Shim and Rabitz, 1998; Shorter et al., 1999; Wang et al., 1999) has the potential to overcome the exponential sampling difficulty of the above approaches in high-dimensional system, and also provide efficient explicit expressions to accurately calculate chemical kinetics. The application of the HDMR method for chemical kinetics calculations is based on expressing kinetic output variables (for example, chemical species concentrations at a given reaction time) as expansions of correlated functions consisting of the kinetic input variables (for example, initial chemical species concentrations). Therefore, an HDMR expansion can be used to directly calculate a chemical species concentration at a later time based on the inputs of initial chemical concentrations and perhaps other variables (for example, solar intensity for photochemical reactions). The accuracy of HDMR predictions for output chemical species concentrations are comparable to conventional chemistry solvers with very significant computational savings, since the operations of an HDMR expansion only involve very rapid and stable algebraic manipulations.

High Dimensional Model Representation (HDMR) Method

The HDMR method is a family of tools (Rabitz and Alis, 1999; Rabitz et al., 1998), which prescribe systematic sampling procedures to map out the relationships between sets of input and output model variables. Let the n-dimensional vector $\boldsymbol{x} = \{x_1, x_2, \ldots, x_n\}$ represent the input variables (for example, initial concentrations of chemical species) of a chemical kinetic system, and $f(\boldsymbol{x})$ is one of the chemical species concentrations at a later time (the output variable). Since the influence of the input variables on the output variable can be independent and/or cooperative, it is natural to express the output $f(\boldsymbol{x})$ as a hierarchical correlated function expansion in terms of the input variables as follows

$$f(\boldsymbol{x}) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1\le i<j\le n} f_{ij}(x_i, x_j) + \sum_{1\le i<j<k\le n} f_{ijk}(x_i, x_j, x_k) + \cdots + f_{12\ldots n}(x_1, x_2, \ldots, x_n) \quad (1)$$

Here f_0 denotes the mean effect which is a constant. The function $f_i(x_i)$ is a first-order term giving the effect of variable x_i acting independently, although generally nonlinearly, upon the output $f(\boldsymbol{x})$. The function $f_{ij}(x_i, x_j)$ is a second-order term describing the cooperative effects of the variables x_i and x_j upon the output $f(\boldsymbol{x})$. The higher-order terms reflect the cooperative effects of increasing numbers of input variables acting together to influence the output $f(\boldsymbol{x})$. The last term $f_{12\ldots n}(x_1, x_2, \ldots, x_n)$ gives any residual dependence of all the input variables locked together in a cooperative way to influence the output $f(\boldsymbol{x})$.

After the relevant component functions in Eq. 1 are learned and suitably represented, then the expressions constitute the HDMR thereby replacing the original route to calculating $f(\boldsymbol{x})$ by the kinetic equation solver. The HDMR is developed by finding suitable expressions of the component function $f_{i_1 i_2 \ldots i_l}(x_{i_1}, x_{i_2}, \ldots, x_{i_l})$ ($l = 0, 1, \ldots$ with f_0 corresponding to $l = 0$) through minimization of the functional

$$\min_{f_{i_1 i_2 \ldots i_l}} \int_{\Omega} w_{i_1 i_2 \ldots i_l}(\hat{\boldsymbol{x}}, \mathbf{u}) \left[f(\mathbf{u}) - f_0 - \sum_{i=1}^{n} f_i(u_i) - \sum_{1\le i<j\le n} f_{ij}(u_i, u_j) - \cdots - \sum_{i_1 i_2 \ldots i_l} f_{i_1 i_2 \ldots i_l}(u_{i_1}, u_{i_2}, \ldots, u_{i_l}) \right]^2 d\mathbf{u} \quad (2)$$

where $\hat{\boldsymbol{x}} = (x_{i_1}, x_{i_2}, \ldots, x_{i_l})$, $d\mathbf{u} = du_1 du_2 \cdots du_n$, Ω is the desired domain of the input variable space, and $w_{i_1 i_2 \ldots i_l}(\hat{\boldsymbol{x}}, \mathbf{u})$ may be considered as a weight function. Different weight functions will produce distinct, but formally equivalent, HDMRs, all of the same structure as Eq. 1. The expressions of the component functions in Eq. 1 found in this way are *optimal choices* for the output $f(\boldsymbol{x})$ over Ω. Therefore, it is expected that the HDMR expansion converges very rapidly such that only low order correlations among the input variables are typically adequate in describing the output behavior. The rapid convergence of the HDMR expansion has been verified in a number of computational studies (Rabitz and Shim, 1999; Shim and

Rabitz, 1998; Shorter et al., 1999; Wang et al., 1999), and the HDMR expansions up to second-order are often sufficient to describe the outputs of many realistic systems.

The key to utilizing the HDMR technique is the ability to rapidly compute the expansion terms shown in Eq. 1. In this article, the Cut-HDMR procedure will be used to compute the expansion terms. With the Cut-HDMR method, first a reference point $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ is defined in the variable space. When taken to convergence the Cut-HDMR is invariant to the choice of reference point $\bar{\boldsymbol{x}}$. In practical circumstances it can be wise to choose $\bar{\boldsymbol{x}}$ within the neighborhood of interest in the input space. The expansion functions are determined by evaluating the input-output responses of the system relative to the defined reference point $\bar{\boldsymbol{x}}$ along associated lines, surfaces, subvolumes, and so on (that is, cuts) in the input variable space. This process reduces to the following relationship for the component functions in Eq. 1

$$f_0 = f(\bar{\boldsymbol{x}}) \tag{3}$$

$$f_i(x_i) = f(x_i, \bar{\boldsymbol{x}}^i) - f_0 \tag{4}$$

$$f_{ij}(x_i, x_j) = f(x_i, x_j, \bar{\boldsymbol{x}}^{ij}) - f_i(x_i) - f_j(x_j) - f_0 \tag{5}$$

where the terminology $f(x_i, \bar{\boldsymbol{x}}^i) \equiv f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \ldots, \bar{x}_n)$ means that all the input variables are at their reference point values except x_i, and so on. The functions $f(\bar{\boldsymbol{x}})$, $f(x_i, \bar{\boldsymbol{x}}^i)$, $f(x_i, x_j, \bar{\boldsymbol{x}}^{ij})$ and so on, are obtained as outputs from the original kinetics code. The f_0 term is the output response of the system evaluated at the reference point $\bar{\boldsymbol{x}}$. The higher-order terms are evaluated as cuts in the input variable space through the reference point. Therefore, each first-order term $f_i(x_i)$ is evaluated along its variable axis through the reference point. Each second-order term $f_{ij}(x_i, x_j)$ is evaluated in a plane defined by the binary set of input variables x_i, x_j through the reference point, and so on. The process of subtracting off the lower-order expansion functions removes their dependence to assure a unique contribution from the new expansion function. Thus, the expansion functions only contain information of the specified level of interaction and they satisfy the following orthogonality conditions

$$f_{i_1 i_2 \ldots i_l}(x_{i_1}, x_{i_2}, \ldots, x_{i_l})\big|_{x_{i_s}=\bar{x}_{i_s}} = 0, \text{ for } s \in \tag{6}$$

This property permits the independent sequential evaluations of the HDMR functions in Eqs. 3–5.

In practice, each of the HDMR expansion functions is numerically represented as a low-dimensional look-up table over its variables. Note that by virtue of Eq. 6, the HDMR in Eq. 1 is exact along any of the cuts. Then, the output response $f(\boldsymbol{x})$ at a point $\boldsymbol{x}$ off of the cuts can be obtained by the following procedure: (1) interpolate each of the low-dimensional HDMR expansion terms in the look-up tables with respect to the input values of the point $\boldsymbol{x}$, and (2) sum the interpolated values of the HDMR terms from zeroth order to the highest order retained in keeping with the desired accuracy.

Application to Alkane Photochemistry

In our previous article, we have formulated a "model" explicit mechanism which considers detailed alkane/NO_x/O_3 photochemistry involving 68 reactions and 52 species for a photochemical box-model study. Then, the Alkane/DCAL mechanism was built by condensing the "model" explicit mechanism through lumping the 30 nonmethane alkane species into three lumped species, while the 21 inorganic species and methane are constrained to remain as individual species. The photochemical box-model from our previous article (Wang et al., 1998), which employed this Alkane/DCAL mechanism and the LSODE routine (Hindmarsh, 1983), is used here to simulate alkane photochemistry. The building of the HDMR requires a series of photochemical box-model integrations to capture the input-output relationships of chemical kinetics into the HDMR expansion terms. Before performing the box-model runs to construct the HDMR expansion terms, the dynamic ranges of the 25 chemical species in the Alkane/DCAL mechanism need to be determined for covering the appropriate domain of the input variable space.

Within these 25 chemical species, there are 14 radical species whose concentrations are set as zeros for the initial concentrations since the initial time is midnight (see later), and 5 chemical species (N_2, O_2, H_2O, CO_2, and CH_4) whose concentrations remain constant during the simulation time. Therefore, only 6 chemical species (O_3, NO_2, NO, and 3 lumped alkane species) initial concentrations will be used as input variables to construct the HDMR expansion terms. Since the 3 lumped alkane species are linear combinations of the 30 explicit alkane species through the transformation of the DCAL lumping matrix, their dynamic ranges need to be determined based on the original input variable space of 30 explicit alkane species. The appropriate dynamic ranges of the 30 explicit alkane species and 3 inorganic species (O_3, NO_2, and NO) are given in Table 1. Then, the dynamic ranges for each of the 3 lumped alkane species are determined by the following procedure: (1) take ten thousand random sampling points from the 30-D hypercube, consisting of 30 explicit alkane species ranges in Table 1, (2) then transform the ten thousand random sampling points from the 30-D hypercube space to the 3-D lumped space through the DCAL lumping matrix, and (3) find the maximum and minimum values for each axis of the 3-D lumped space among the ten thousand transformed points.

In this article, we applied the HDMR method to construct the function $f(\bar{\boldsymbol{x}}_0, t_j)$ representing one of the chemical species concentrations after a total time t_j has elapsed, for example, 14 h. Here, $\bar{\boldsymbol{x}}_0$ is the vector of initial concentrations at time 0, the beginning of the 24-h simulation. Thus, 24 HDMR expansions were generated corresponding to $t_j = 1$ h, $t_j = 2$ h, ..., $t_j = 24$ h. The effects of the temperature and light intensity on the alkane photochemistry are also taken into account by constructing different HDMR expansions for different hours of the day. We select the midpoint in each of the 6 dynamic ranges of the input variables as the reference point $\bar{\boldsymbol{x}}$ in Eqs. 3–5 to construct an HDMR at a given hour. Within the 6 specified dynamic ranges of the input variables, the sampling grid points were taken as 10 equally spaced grids along each input variable axis tabulated in Table 2. According to Table 2, it should be pointed out that the dynamic ranges of the 2nd and 3rd lumped alkane species cover negative values, since the DCAL-lumped

Table 1. Dynamic Ranges (ppb) for the 33 Input Chemical Species Concentrations

No.	Chemical Species	Range	No.	Chemical Species	Range
1	Ozone	1–200	18	Methylcyclopentane	4–18
2	Nitrogen dioxide	1–50	19	Heptane	60–275
3	Nitrogen oxide	1–100	20	3-Methylhexane	5–12
4	Ethane	5–24	21	2,4-Dimethylpentane	2–16
5	Propane	6–25	22	2,3-Dimethylpentane	3–17
6	Butane	7–38	23	Methylcyclohexane	2–16
7	Iso-butane	3–17	24	Octane	7–428
8	Pentane	2–20	25	4-Methylheptane	2–11
9	Iso-pentane	1–4	26	2,2,4-Trimethylpentane	2–11
10	Neo-pentane	1–3	27	Ethylcyclohexane	2–11
11	Cyclopentane	1–3	28	Nonane	3–37
12	Hexane	24–280	29	4-Ethylheptane	1–10
13	2-Methylpentane	3–11	30	Decane	1–9
14	3-Methylpentane	3–11	31	4-Propylheptane	1–8
15	2,2-Dimethylbutane	4–12	32	Undecane	1–7
16	2,3-Dimethylbutane	3–17	33	Dodecane	1–7
17	Cyclohexane	3–16			

"species" are not intended to represent concentrations but linear combinations. So it could represent the difference of the concentrations of two species and, thus, assume negative, though still physically meaningful values.

According to the formulas specified in Eqs. 3–5, the individual HDMR expansion terms were obtained by algebraic manipulations of the box model outputs. Special care is necessary to ensure that the relevant portion of the input variable space is covered by the chosen sampling points for the box model runs. The calculated HDMR expansion terms from zeroth order to the highest desired order were saved as look-up tables with respect to the chosen sampling points. Thus, the output $f(\boldsymbol{x})$ evaluated at a point $\boldsymbol{x}$ off of those sampling points can be obtained by interpolation and algebraic manipulation of the HDMR tables. The goal is to employ the HDMRs for calculating diurnal multispecies time-concentration profiles, based on the inputs of the initial chemical species concentrations.

Results and Discussion

At the first stage of the HDMR development, we construct the HDMR expansion only up to first-order to test if it is sufficient to accurately predict the time-concentration profiles of chemical species concentrations. The performance of the first-order HDMR predictions are evaluated by comparing the diurnal profiles of ozone, NO_2, NO, and OH concentrations predicted by the first-order HDMR expansions with those simulated from the photochemical box-model runs using the "model" explicit mechanism, the Alkane/DCAL mechanism, and the alkane portions of the CB4 and SAPRC93 mechanisms (shown in Figure 1). Initial species concentrations for the photochemical box-model simulation (see Table 3) are typical of polluted urban air for the species included in the "model" explicit mechanism (Eq. 15). According to Figure 1, the first-order HDMR expansion produces predictions almost identical to those obtained from the full model, as well as the Alkane/DCAL mechanism, while the alkane portions of the CB4 and SAPRC93 mechanisms appear to produce deviations from the full model calculations for reaction times exceeding 500 min.

Table 2. Dynamic Ranges (ppb) for the 6 Input Variables in Alkane/DCAL Mechanism

No.	Chemical Species	Range
1	Ozone	$1, \ldots, 200$
2	Nitrogen dioxide	$1, \ldots, 50$
3	Nitrogen oxide	$1, \ldots, 100$
4	1st Lumped alkane species	$30, \ldots, 180$
5	2nd Lumped alkane species	$-5.5, \ldots, 41$
6	3rd Lumped alkane species	$-140, \ldots, -36$

The number of box-model runs required to determine the HDMR expansion depends on the number of input variables and the number of sampling grid points for each input variable. If s grid points are used for each input variable, then $(s-1)$ box-model runs are required to specify each first-order expansion term $f_i(x_i)$. The model run at the reference point is not required since the value of the first-order expansion is zero at that point by virtue of Eqs. 3–5. In an analogous manner, $(s-1)^2$ model runs are required to determine each second-order term $f_{ij}(x_i, x_j)$. In the present case there is one f_0 term, and 6 f_i terms determined for the HDMR up to first-order. Therefore,

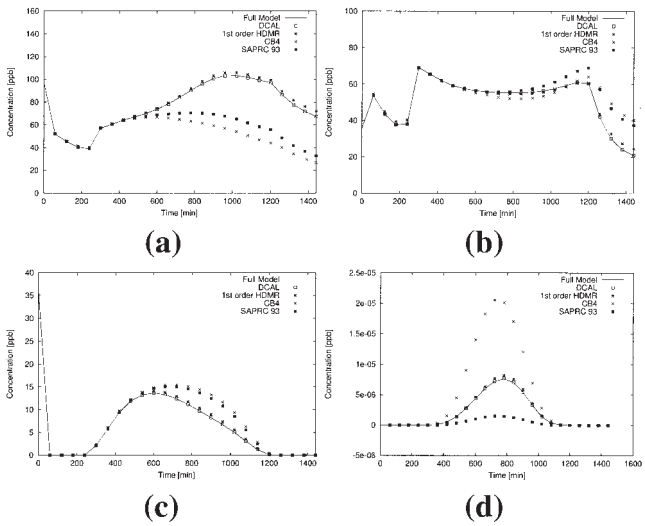

Figure 1. Diurnal concentration-time profiles for (a) O_3, (b) NO_2, (c) NO, and (d) OH predicted by the HDMR and the full model.

Table 3. Initial Conditions (ppb) for the 33 Input Chemical Species Concentrations

No.	Chemical Species	Value	No.	Chemical Species	Value
1	Ozone	97.4	18	Methylcyclopentane	7.3
2	Nitrogen dioxide	35.7	19	Heptane	159
3	Nitrogen oxide	35.7	20	3-Methylhexane	6.4
4	Ethane	13.6	21	2,4-Dimethylpentane	8.0
5	Propane	18.9	22	2,3-Dimethylpentane	13.8
6	Butane	20.1	23	Methylcyclohexane	7.6
7	Iso-butane	12.1	24	Octane	320
8	Pentane	9.0	25	4-Methylheptane	5.3
9	Iso-pentane	2.9	26	2,2,4-Trimethylpentane	8.4
10	Neo-pentane	1.7	27	Ethylcyclohexane	5.0
11	Cyclopentane	2.2	28	Nonane	26.2
12	Hexane	108	29	4-Ethylheptane	3.8
13	2-Methylpentane	7.4	30	Decane	6.2
14	3-Methylpentane	5.4	31	4-Propylheptane	2.9
15	2,2-Dimethylbutane	8.2	32	Undecane	4.7
16	2,3-Dimethylbutane	6.7	33	Dodecane	2.5
17	Cyclohexane	9.4			

a total of 55 box-model runs ($1 + 9 \times 6$) are required to construct the first-order HDMR look-up table in this work. The computational effort for constructing the HDMR look-up table of a fixed order, only scales polynomially with the system dimension n and the number of grid points s rather than the conventional exponential scaling.

Testing for different initial conditions

The earlier results were obtained by simulating only one set of initial conditions for chemical species concentrations. In order to test the robustness of the HDMR expansions, various sets of initial conditions were also simulated to study the agreement between the full model calculations and the HDMR predictions. For thoroughly studying the performance of the HDMR predictions over the full input variable space, we consecutively take every one thousand random sampling points from the expanding hypercubes, consisting of 6 input variable axes by gradually increasing the length of the axis from 10% to 100% of the full dynamic ranges centered at the reference point. Then, the quality of the HDMR predictions is represented by calculating the relative error (*R.E.*) for each comparison defined as

$$R.E. = \left| \frac{f_t - f_r}{f_r} \right| \tag{7}$$

where f_t is the HDMR prediction of chemical species concentration, and f_r is the corresponding box-model output. We present the general quality of the HDMR predictions for ozone and NO_2 concentrations at the time of noon.

To summarize the comparison results, we plot the percentages of 1,000 random comparisons for first-order HDMR expansion that are within 5% *R.E.* against the size of the sampling hypercube at the input variable space in Figure 2 for ozone and NO_2 concentrations. The behavior shown in Figure 2 is typical of the overall results for other species and times. According to Figure 2, it is shown that more than 85% of the 1,000 random comparisons are within 5% *R.E.* by taking the random initial inputs up to 40% of the full dynamic ranges of the 6 input variables. When we expand the sampling hypercube to 100% of the full dynamic ranges, this quantity drops to 40%. This

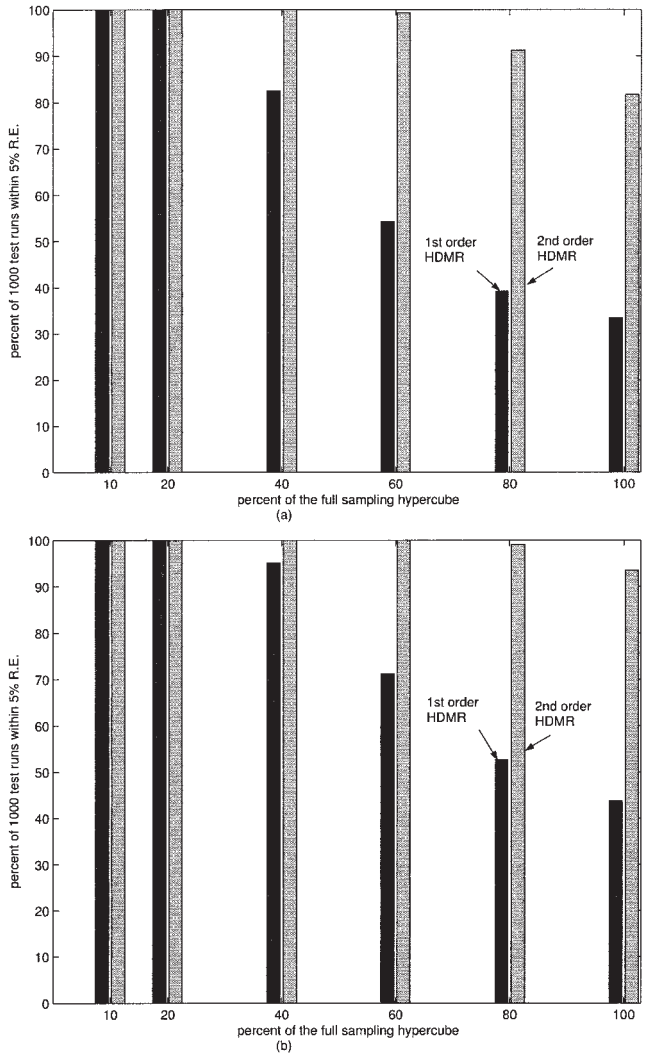

Figure 2. Error analyses for (a) O_3 and (b) NO_2 concentrations of the HDMR predictions by calculating the percentage of 1,000 random tests that are within 5% relative error for each size of the sampling range.

implies that the first-order HDMR expansion is not sufficient for the full dynamic ranges, and higher-order HDMR expansion terms are needed. Therefore, we include the second-order HDMR expansion terms in the predictions to increase this accuracy. In this case study, there are 15 $(6 \times 5)/2)$ f_{ij} terms. According to the formula given earlier, additional 1,215 (15×9^2) box-model runs are required to construct the second-order HDMR look-up table. According to Figure 2, it is shown that the contributions from these 15 f_{ij} terms can make up the deficiencies of the first-order HDMR predictions for those 60% offset points.

Testing for computational efficiency

The CPU times for performing one-thousand simulations, corresponding to various initial conditions, on a SUN ULTRA SPARC-1 170 MHz workstation, for the full alkane mechanism, for the alkane portions of CB4 and SAPRC93 mechanisms, for the Alkane/DCAL mechanism with three lumped species, as well as for the first- and second-order HDMR expansions are 2464, 1256, 1343, 1305, 1.3, and 2.6 s, respectively. The second-order HDMR predictions can achieve excellent accuracy for the full dynamic ranges, while being about 1,000 times faster than the full mechanism and 500 times faster than the Alkane/DCAL mechanism.

Conclusions and Future Work

In recent years, Photochemical Air Quality Simulation Models (PAQSMs) have evolved into complex systems that simulate the dynamics of multiple co-occurring atmospheric pollutants, such as tropospheric ozone, fine particulate matter (PM2.5), and air toxics, simultaneously, by taking into account their physical and chemical interactions; this evolution has substantially increased the computational requirements of PAQSMs. Various approaches, such as parallel computing and efficient numerical techniques (for example, the sparse matrix vectorized Gear method (Jacobson and Turco, 1994), and the modified euler backward iterative (MEBI) method (Hertel et al., 1993)), have been introduced to reduce simulation time.

This study demonstrated that the HDMR expansion can generate estimates of chemical species concentrations as accurate as those obtained from solving tropospheric alkane photochemistry through conventional (for example, Gear-type) chemistry solvers, over wide ranges of initial conditions, while performing about 1,000 times faster. The speed of computations utilizing the HDMR expansion provides excellent potential for resolving the computational burden associated with simulations employing state-of-the-art comprehensive regional and multiscale PAQSMs. Such models are now routinely used for regulatory purposes, in order to study and establish emission source control strategies necessary to meet the National Ambient Air Quality Standards (NAAQS) for photochemical pollutants. The computational requirements are actually becoming heavier, as applications now often include multiple and long-term (for example, seasonal or annual) simulations at multiple scale (that is, local, urban, regional) resolutions. Also, recently, The National Oceanic and Atmospheric Administration (NOAA), in association with the U.S. Environmental Protection Agency (USEPA) initiated the National Air Quality Forcasting (NAQF) program. The vision of the NAQF program is to provide ozone, PM2.5, and other pollutant forecasts in a sufficiently accurate and timely manner, so as to take action for preventing or reducing adverse health effects.

Fast Equivalent Operational (FEOM) versions of PAQSMs that employ the HDMR approach for the chemical kinetics code significantly facilitate both scientific and regulatory applications. For scientific purposes, they can be used to perform comprehensive uncertainty characterization/propagation for providing insights into model parameters/inputs (for example, meteorology, emissions, chemistry, and so on) that have the highest impact on the model outputs, so that resources can be focused on reducing uncertainties where it is most appropriate. For regulatory purposes, they can be used to perform multiple and long-term simulations at multiple resolutions in a timely manner, for developing rational air quality management strategies.

Ongoing work is focusing on extending the current application of the HDMR method by (a) merging the chemistry time-step as one of the input variables for constructing the HDMR expansion, so that only one HDMR expansion will be needed for use in 3-D models, and (b) employing the HDMR approach to the analysis and reduction of more comprehensive atmospheric chemistry mechanisms (such as CB4 or SAPRC93) for developing a "faster" PAQSM. When applied to accelerate the chemical kinetic calculations for these comprehensive atmospheric chemistry mechanisms, which account for a large portion of PAQSM computation, the HDMR method could provide an effective solution for reducing the huge computational requirements by developing an "accurate and efficient" alternative to the original models.

Supporting Information

The code for the HDMR method, used in the current study to generate the reduced mechanism, is available at http://ccl.rutgers.edu/hdmr.htm (on the Computational Chemodynamics Laboratory website). Detailed instruction for using the code, along with a demonstration of the case study discussed in the current article, are available at this site.

Acknowledgments

This work was conducted under the auspices of the Ozone Research Center (ORC); base funding for the ORC is provided by the New Jersey Department of Environmental Protection (NJDEP); additional funding was provided by U.S. EPA's National Exposure Research Laboratory (NERL) through an EPA—University Partnership Agreement (EPA-UPA).

Literature Cited

Carter, W. P. L., "Development and Implementation of an Up-to-Date Photochemical Mechanism for Use in Airshed Modeling," Technical report, Statewide Air Pollution Research Center, University of California, Riverside (1988).

Gery, M., G. Whitten, J. Killus, and M. Dodge, "A Photochemical Kinetics Mechanism for Urban and Regional Scale Computer Modeling," *J. Geophys. Res.,* **94**, 12,925 (1989).

Hertel, O., R. Berkwicz, and J. Christensen, "Test of Two Numerical Schemes for Use in Atmospheric Transport-Chemistry Models," *Atmos. Environ.,* **27A**, 2591 (1993).

Hindmarsh, A. C., ODEPACK, *A Systematized Collection of ODE Solvers, in Scientific Computation,* chapter; Livermore Solver for Ordinary Differential Equations, 55–64, North-Holland, Amsterdam (1983).

Jacobson, M. Z., and R. P. Turco, "SMV Gear: A Sparse Matrix Vectorized Gear Code for Atmospheric Models," *Atmos. Environ.,* **28A**, 273 (1994).

Klonecki, A., and H. I. Levy, "Tropospheric Chemical Ozone Tendencies

in CO–CH_4–NO_y–H_2O System: Their Sensitivity to Variations in Environmental Parameters and Their Application to a Global Chemistry Transport Model Study," *J. Geophys. Res.,* **102**, 21,221 (1997).
Rabitz, H., and O. Alis, "General Foundations of High Dimensional Model Representations," *J. Math. Chem.,* **25**, 197 (1999).
Rabitz, H., O. Alis, J. Shorter, and K. Shim, "Efficient Input-Output Model Representations," *Comp. Phys. Comm.,* **115**, 1 (1998).
Rabitz, H., and K. Shim, "Multicomponent Semiconductor Material Discovery Using a Generalized Correlated Function Expansion," *J. Chem. Phys.,* **111**, 10,640 (1999).
Shim, K., and H. Rabitz, "Independent and Correlated Composition Behavior of the Energy Band Gaps for the ga_α Alloys," *Phys. Rev. B,* **58**, 1940 (1998).
Shorter, J., P. Ip, and H. Rabitz, "An Efficient Chemical Kinetics Solver Using High Dimensional Model Representations," *J. Phys. Chem. A,* **103**, 7192 (1999).
Spivakovsky, C., S. Wofsy, and M. Prather, "A Numerical Method for Parameterization of Atmospheric Chemistry: Computation of Tropospheric OH," *J. Geophys. Res.,* **95**, 18,433 (1990).
Stockwell, W., P. Middleton, J. S. Chang, and X. Tang, "The Second Generation Regional Acid Deposition Model Chemical Mechanism for Regional Air Quality Modeling," *J. of Geophysical Res.,* **95**, 16,343 (1990).
Turányi, T., "Parameterization of Reaction Mechanisms Using Orthogonal Polynomials," *Computers Chem.,* **18**, 45 (1994).
Wang, S., P. Georgopoulos, G. Li, and H. Rabitz, "Condensing Complex Atmospheric Chemistry Mechanisms—1: The Direct Constrained Approximate Lumping (DCAL) Method Applied to Alkane Photochemistry," *Environ. Sci. Technol.,* **32**, 2018 (1998).
Wang, S., H. Levy, II, G. Li, and H. Rabitz, "Fully Equivalent Operational Models for Atmospheric Chemical Kinetics within Global Chemistry-Transport Models," *J. Geophys. Res.,* **104**, 30,417 (1999).

Manuscript received Aug. 6, 2003, revision received Aug. 16, 2004, and final revision received Nov. 5, 2004.