A computationally-efficient secondary organic aerosol module for three-dimensional air quality models

P. Liu and Y. Zhang

Dept. of Marine, Earth, and Atmos. Sci., North Carolina State Univ., Raleigh, NC 27519, USA

Received: 10 March 2008 – Accepted: 10 March 2008 – Published: 10 April 2008

Correspondence to: Y. Zhang (yzhang9@ncsu.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Accurately simulating secondary organic aerosols (SOA) in three-dimensional (3-D) air quality models is challenging due to the complexity of the physics and chemistry involved and the high computational demand required. A computationally-efficient yet accurate SOA module is necessary in 3-D applications for long-term simulations and real-time air quality forecasting. A coupled gas and aerosol box model (i.e., 0-D CMAQ-MADRID 2) is used to optimize relevant processes in order to develop such a SOA module. Solving the partitioning equations for condensable volatile organic compounds (VOCs) and calculating their activity coefficients in the multicomponent mixtures are identified to be the most computationally-expensive processes. The two processes can be speeded up by relaxing the error tolerance levels and reducing the maximum number of iterations of the numerical solver for the partitioning equations for organic species; turning on organic-inorganic interactions only when the water content associated with organic compounds is significant; and parameterizing the calculation of activity coefficients for organic mixtures in the hydrophilic module. The optimal speedup method can reduce the total CPU cost by up to a factor of 29.7 with ±15% deviation from benchmark results. These speedup methods are applicable to other SOA modules that are based on partitioning theories.

1 Introduction

Representing secondary organic aerosols (SOA) in three dimensional (3-D) atmospheric models is very important because they constitute a sizeable fraction of fine particulate mater (PM$_{2.5}$), which has impacts on human health, visibility degradation, and climate change (Watson, 2002; Davidson et al., 2005; IPCC, 2007; Zhang et al., 2007). The formation of SOA depends on atmospheric abundance of anthropogenic and biogenic volatile organic compounds (VOCs), their chemical reactivity, solubility, and the condensable products from their photochemical oxidation, as well as gas/particle par-
tioning of the condensable products. The mechanism of SOA formation is one of the least understood research areas due to the complexities of chemical and thermodynamic properties of hundreds of VOCs (Turpin et al., 2000); therefore, simulating SOA posts a major challenge in 3-D air quality and climate modeling.

A number of modules have been developed for simulating SOA in 3-D air quality models (Strader et al., 1999; Barthelmie and Pryor, 1999; Aumont et al., 2000; Andersson-Sköld and Simpson, 2001; Schell et al., 2001; Pun et al., 2002; Griffin et al., 2002a, 2003; Tulet et al., 2006). Two aerosol modules: Model of Aerosol Dynamics, Reaction, Ionization, and Dissolution 1 and 2 (MADRID 1 and MADRID 2) have been incorporated into the US Environmental Protection Agency (EPA)'s 3-D Community Multiscale Air Quality (CMAQ) modeling system (Binkowski and Roselle, 2003) to simulate SOA (Zhang et al., 2004). MADRID 1 uses an empirical representation of SOA formation based on data obtained in smog chamber experiments (Odum et al., 1997; Griffin et al., 1999). MADRID 2 uses a mechanistic representation that simulates an external mixture of hydrophilic and hydrophobic particles (Pun et al., 2002). The two modules differ in two aspects. First, both use the same partitioning equation for hydrophobic organic compounds (OC) but with different methods for partitioning coefficients, which are obtained from the smog chamber experiments in MADRID 1 but calculated in MADRID 2 as a function of temperature and composition of the hydrophobic OCs based on Raoult’s law following the equation derived by Pankow et al. (1994). The determination of chemical composition in MADRID 2 involves the use of UNIFAC (Fredenslund et al., 1977) for activity coefficients. Second, hydrophilic compounds are treated based on the Henry’s Law in MADRID 2 but not in MADRID 1. While MADRID 2 represents a detailed treatment for SOA formation, it is more computationally-expensive than MADRID 1 (by up to a factor of 8 in 3-D simulations), which limits its application for long-term 3-D simulations and real-time air quality forecasting.

In this work, a coupled gas and aerosol box model (i.e., zero-dimensional (0-D) version of CMAQ-MADRID 2) is used to explore various speedup methods while maintaining a desirable numerical accuracy. The main objective of this study is to improve the
computational efficiency of SOA module for 3-D long-term air quality simulations and real-time forecasting.

2 Model description and test conditions

2.1 Model treatment

CMAQ-MADRID 2 simulates detailed gas-phase and aqueous-phase chemistry, and important aerosol microphysical processes that govern the chemical composition and size distribution of PM using a sectional size representation of particles within the CMAQ model framework (Zhang et al., 2004; Pun et al., 2005). All other atmospheric processes that govern the sources and fates of chemical species such as emissions, transport, and removal are either the same as or similar to those of CMAQ. The major aerosol microphysical processes include chemical thermodynamic equilibrium for inorganic species, secondary aerosol formation, new particle formation, condensational growth (or shrinkage by volatilization), mass transfer between the bulk gas phase and particulate phase, and aerosol activation by cloud droplets. The relevant processes involved in SOA treatments in MADRID 2 are described below. A detailed description can be found in Zhang et al. (2004).

CMAQ-MADRID 2 uses the Caltech Atmospheric Chemical Mechanism (CACM) gas-phase chemistry (Griffin et al., 2002b) to provide the rate of formation of condensable products for SOA module. CACM contains 361 reactions of 191 species and provides detailed descriptions of several generations of products from alkanes (3 classes), alkenes (2 classes), aromatics (2 classes), alcohols (3 classes), isoprene, and terpenes (2 classes). This mechanism is uniquely suitable for simulating SOA formation because it explicitly treats 42 condensable second- and third-generation products. The MADRID 2 SOA module includes 10 surrogate compounds, grouped according to their affinity for water (5 surrogate species for 28 explicit hydrophobic OCs and 5 surrogate species for 14 hydrophilic OCs), origin (anthropogenic vs. biogenic), size (number of
carbons), volatility, and dissociation properties (Pun et al., 2002). Hydrophilic OCs include those with a short carbon chain (≤7 carbons; or ≤10 carbons with three or more functional groups), high solubility (≥1 g solute/100 g water), and a high effective Henry’s law constant (≥1 × 10^6 M atm⁻¹). Hydrophobic condensable OCs are identified by their estimated octanol-water partitioning coefficients. Each surrogate compound assumes the characteristic size and functional groups of the corresponding explicit compounds. Due to the paucity of property data for complex OCs, partitioning parameters, such as Henry’s law constants and saturation vapor pressures, are estimated using group contribution methods (Pun et al., 2002). Dissociation constants and deliquescence relative humidities (DRH) are assigned based on analogy to simpler compounds.

The partitioning of hydrophobic condensable OCs into particulate phase through absorption and gas-particle equilibrium is described as follows (Pun et al., 2002; 2003; Zhang et al., 2004):

\[
K_i = \frac{A_i}{G_i} / \frac{M_{om}}{G_i} \tag{1}
\]

where \( K_i \) (m³ µg⁻¹) is the partitioning coefficient, \( A_i \) and \( G_i \) (µg m⁻³ air) are the mass concentrations of species \( i \) in the particulate- and gas-phase, respectively, and \( M_{om} \) (µg m⁻³ air) is the sum of primary (non-volatile) and secondary organic carbon (semi-volatile) in the particulate phase that serve as the organic absorbing medium. \( A_i \) can be calculated based on \( K_i \), \( G_i \), and \( M_{om} \). The partitioning coefficient, \( K_i \), can be calculated as follows (Pun et al., 2003):

\[
K_i = \frac{760RT}{10^6 P_{i}^{sat} \gamma_i \bar{M}_{om}} \tag{2}
\]

where \( R \) is the ideal gas constant (8.2×10⁻⁵ m³ atm mol⁻¹ K⁻¹), \( T \) (K) is the temperature, \( P_{i}^{sat} \) (torr) is the saturation vapor pressure, \( \gamma_i \) is the activity coefficient in the liquid phase, and \( \bar{M}_{om} \) is the average molecular weight of the organic absorbing medium. Given an initial guess of \( A_i, \bar{M}_{om} \) can be estimated, \( \gamma_i \) can be calculated by UNIFAC.

7089
and $K_i$ can be solved through Eq. (2). $K_i$ can then be substituted into Eq. (1) to calculate a new $A_i$, and the deviation between the two $A_i$ values. Such iterative calculations will stop until the deviation meets a convergence criterion of $10^{-4}$, making this iteration computationally very expensive.

In the absence of liquid water, hydrophilic condensable products undergo absorption that is similar to that of hydrophobic condensable OCs. In the presence of liquid water, the partitioning of hydrophilic condensable OCs into existing aqueous particles is governed by Henry's law assuming equilibrium between the bulk gas and whole particulate phase, and takes into account the activity coefficient of the molecular solute. The additional water uptake can occur in the presence of hydrophilic SOA when the ambient relative humidity (RH) is above the DRH of any individual OC in the particulate phase (Saxena and Hildemann, 1996, 1997). The addition of organic ions changes the pH and total water content of aqueous particles, and consequently, affects the partitioning of inorganic compounds, such as nitrate and ammonium (Saxena et al., 1995; Cruz and Pandis, 2000; Pun et al., 2003). The role of SOA in total aerosol water and in transferring of nitrate into the aerosol phase is significant, especially at low RH (<50%) and high SOA mass fractions (>20% of total PM$_{2.5}$) (Ansari and Pandis, 2000). The above effects are also species-specific (Choi and Chan, 2002), and their modeling is difficult due to the lack of fundamental thermodynamic data for OCs and a suitable theoretical approach (Clegg et al., 2001). These organic-inorganic (O-I) interactions are simulated in CMAQ-MADRID 2 by coupling the hydrophilic SOA module with an inorganic equilibrium model, ISORROPIA of Nenes et al. (1999), to relate water and ions between organic and inorganic species.

In CMAQ-MADRID 2, both the hydrophilic and hydrophobic SOA modules require the simultaneous solution of partitioning equations for several OCs. A globally-convergent Newton/line search method is used to solve for the partitioning equations simultaneously. Some simplifications are provided as alternatives to detailed treatments to improve computational efficiency in 3-D applications (Pun et al., 2003). For example, the amount of water associated with hydrophilic OCs is calculated by assuming that the
water associated with each hydrophilic compound in a binary solution is additive (i.e., using the Zdanovskii, Stokes, and Robinson (ZSR) equation (Stokes and Robinson, 1966)). In the hydrophobic module, $M_{om}$ in Eq. (1) is calculated based on the concentration and composition of PM at the beginning of the time step in each grid cell, and held constant for the partitioning calculation during the same time step.

2.2 Test condition and model input

The 0-D CMAQ-MADRID 2 used in this study simulates the CACM gas-phase chemistry and all of aforementioned aerosol processes except aerosol activation. Other atmospheric processes such as emissions, dilution, transport, removal, and aqueous-phase chemistry are not included in the box model. Three numerical solvers are implemented to solve CACM in the box model, the quasi steady state approximation solver (QSSA) (Hestveldt et al., 1978), the RosenBrock solver (ROS3) (Sandu et al., 1997a, b), and the sparse-matrix vectorized Gear’s solver (SMVGEAR) (Jacobson and Turco, 1994). The box model is set up for a one-day simulation starting 12:00 GMT (corresponds to 08:00 a.m. Eastern Daylight Time (EDT)). A diurnal photolytic rate profile is used to represent a typical summer day in the eastern US. Four hypothetical atmospheric conditions are used to represent the rural and urban conditions with low and high biogenic VOCs (rural LBG/HBG, and urban LBG/HBG). The initial gas-phase and particulate concentrations under the four conditions are given in Tables 1 and 2. The typical ambient non-methane hydrocarbon (NMHC) concentrations range from a few to 2000 ppbC (Finlayson-Pitts and Pitts, 1986; Seinfeld and Pandis, 1998). In the baseline simulation, the total NMHC concentrations are assumed to be 50 and 150 ppbC for the rural and urban conditions, respectively, which represent the low limits under such conditions. The speciation of NMHCs for all the initial CACM organic species except biogenic VOCs (i.e., isoprene (ISOP), and monoterpenes (BIOL + BIOH)) is based on the specific distribution factors of Griffin et al. (2002a). The concentrations of isoprene and monoterpenes in the ambient air are typically in the range of 1 to 10 ppbC (Finlayson-Pitts and Pitts, 1986). The only differences between rural/urban HBG and
Improve the efficiency of MADRID 2

3.1 Methodology and simulation design for computational efficiency improvement

Baseline simulations without aerosol treatments are first conducted to identify the most appropriate solver for gas-phase chemistry. Among the three solvers implemented for CACM, the SMVGEAR is widely used in atmospheric transport model (Hertel et al., 1993; Huang and Chang, 2001) and considered as the most accurate numerical chemical solver available (US EPA, 2004). The results from SMVGEAR with the most stringent error tolerances for convergence (a relative error tolerance (RTOL) of $10^{-5}$ and an absolute error tolerance (ATOL) of $10^{-13}$ ppm) are therefore used as a benchmark for accuracy. The results with the ROS3 and the QSSA solvers at the same most stringent error tolerance are quite similar to those with the SMVGEAR solver (the deviations in mixing ratios are within 4.2% for the ROS3 solver and 5.6% for the QSSA solver for all major gas-phase species). For all conditions, the QSSA solver is the fastest among the three solvers, while the ROS3 solver is slightly faster than the SMVGEAR solver for most simulations. Regardless of the solvers used, simulations under rural and urban HBG conditions take much more CPU time (in second) (by factors of 8.6 and 3.6,
respectively) than those with LBG conditions, with the highest CPU cost for the rural HBG condition. Considering accuracy and computational efficiency, the ROS3 solver is selected for the speed-up test experiments for all simulations with aerosols.

Solving aerosol processes takes much more CPU time than solving gas-phase chemistry (e.g., 2.1 vs. 199.3 s using the ROS3 solver under the rural HBG condition). Under the test conditions, among all aerosol processes simulated, two computationally most expensive physical processes are identified: solving partitioning equations for OCs using the globally-convergent Newton/line search method and calculating activity coefficients of multicomponent mixtures through UNIFAC. The speed-up test experiments therefore focus on the two physical processes under the rural HBG condition. Speed-up methods that are tested include: speeding up the numerical solver for the SOA formation; conditionally activating organic-inorganic interactions; and parameterizing the UNIFAC calculation of activity coefficients. These methods are first tested in the box MADRID with a constant temperature of 298 K and an RH of 70% under the four test conditions. The computationally-efficient MADRID 2 (referred to as MADRID 2.Fast hereafter) using the optimal speed-up methods is further tested for atmospheric conditions with temperatures of 273–313 K and RHs of 10–95%.

3.2 Numerical solver for partitioning of organic compounds

The globally-convergent Newton/line search method is developed from Newton’s method and can solve the simultaneous equations of the form \( F(x) = 0 \). It can converge to a solution from any starting point (i.e., initial guess value) (Press et al., 1997). Six parameters are tested for potential speedup including three tolerance control parameters (i.e., TOLF, TOLMIN, and TOLX), two internal parameters (i.e., EPS and ALF) and one external parameter (i.e., MAXIT). Their default values and function are given in Table 3. Relaxing TOLF and TOLX results in small-to-significant speed-up. As shown in Fig. 1a, under the rural HBG condition, at TOLF = \( 1 \times 10^{-3} \), relaxing TOLX from \( 1 \times 10^{-6} \) to \( 1 \times 10^{-3} \) results in 6.2 to 47.1% CPU reduction. At TOLF = \( 1 \times 10^{-2} \), relaxing TOLX from \( 1 \times 10^{-6} \) to \( 1 \times 10^{-3} \) results in 15.3 to 48.2% CPU reduction. As shown in Fig. 1b,
reducing MAXIT from 400 (the default value) to 5 results in 18.3 to 80.9% of CPU reduction. Relaxing EPS and ALF from the default value of $1 \times 10^{-4}$ to $1 \times 10^{-3}$ can result in 28% and 13.5% speed-up, respectively.

Multiple parameters are relaxed in one simulation to obtain an optimal speed-up. Figure 1c shows various combinations of speed-up parameters, where solver_1, solver_2, and solver_3 correspond to a combination of TOLF=$1 \times 10^{-3}$, TOLX=$1 \times 10^{-5}$, MAXITS=100, EPS=$1 \times 10^{-4}$, and ALF=$1 \times 10^{-4}$; the same with solver_1 but with TOLX=$1 \times 10^{-3}$; and a combination of TOLF=$1 \times 10^{-2}$, TOLX=$1 \times 10^{-3}$, MAXITS=5, EPS=$1 \times 10^{-3}$, and ALF=$1 \times 10^{-3}$, respectively. The speed-up is much more significant as compared with relaxing single parameter in one simulation. Solver_1, solver_2, and solver_3 reduce the CPU cost by 45%, 54.5%, and 86.1%, respectively. Compared with the benchmark results (i.e., using the SMVGEAR solver for gas-phase chemistry without any speed-up in aerosol simulation), the ranges of percentage deviations for simulated major gas-phase species and total PM$_{2.5}$ with solver_1, solver_2, and solver_3 are $-2.7$ to $1.4\%$, $-5.5$ to $1.7\%$, and $-10.6$ to $39.1\%$, respectively. While the speed-up with solver_3 is at the expense of reducing accuracy, solver_2 and solver_1 provide a good compromise between the CPU cost and the numerical accuracy.

3.3 Organic-Inorganic (O-I) interactions

The extent of the partitioning for any individual OC between the gas and particulate phases depends not only on the amounts and properties of the compound, but also on the amount of water present in the atmosphere. Meanwhile, the water uptake and ions associated with OCs will affect the partitioning of inorganic aerosols. In CMAQ-MADRID2, such O-I interactions are simulated by coupling of a module solving dissolution of water soluble OCs with an inorganic gas-aerosol equilibrium module (i.e., ISORROPIA). Due to the lack of experimental data on the equilibrium in mixed inorganic/organic/aqueous system, the inorganics and organics are treated separately and coupled only through the liquid water content (LWC) and pH of the aerosols (Pun et al., 2002). Simulating O-I interactions requires iterations that may be computationally-expensive. Since the
absolute values of water uptake by OCs (i.e., LWORG) and the total water content from inorganics (i.e., RWA) may vary significantly (e.g., $10^{-4}$ to 2 and 3–15 µg m$^{-3}$, respectively), their ratios (i.e., LWORG/RWA) provide a good indicator on the relative importance of water uptake by OCs in the system, which is chosen as a cutoff value to turn on/off the O-I interactions. If the value of LWORG/RWA is less than the cutoff value, the water uptake and ion added by OCs will not be treated in ISORROPIA (i.e., no O-I interactions). Several cutoff values of LWORG/RWA are tested at an RH of 70% in terms of CPU costs and numerical accuracy. With the increased cutoff values (e.g., from 0.1 to 20%), the CPU costs reduce up to 64% under all test conditions. The CPU cost is reduced significantly when choosing a large cutoff value, but the tradeoff is that the percentage deviations from the benchmark results are also large. For example, the percent deviations are $-16.7$ to $49.2\%$ for major gas-phase species and $-21.2$ to $17.7\%$ for total PM$_{2.5}$ for a cutoff value of 20%. The cutoff value of 0.5\% can reduce CPU cost 7\% with relatively small percent deviations, i.e., $-7.2$ to $2.6\%$ for major gas-phase species and $-4.2$ to $7.3\%$ for total PM$_{2.5}$. The value of LWORG/RWA of 0.5\% is thus selected as the cutoff value to turn on/off the O-I interactions. As shown in Fig. 2, activating O-I interactions only when LWORG/RWA >0.5\% can reduce CPU cost by 7.1\% and 15.6\% for the rural and urban HBG conditions, and 39.3\% and 18.6\% for the rural and urban LBG conditions. The reduction in CPU cost is greater under the rural LBG than rural HBG conditions, because relative amount of water content associated with organics are smaller and LWORG/RWA is less than the cutoff value under the rural LBG condition. However, the CPU reductions are similar under urban LBG and HBG conditions due to smaller differences of LWORG/RWA between the two conditions.

3.4 Parameterization of activity coefficient calculation

Multiple linear regression analysis based on the least squares approach is applied for parameterization of activity coefficient calculation (para-AC) based on UNIFAC in the hydrophilic SOA module that is identified to be the most computationally-expensive

7095
process. Six molecules are treated for hydrophilic OC partitioning in CMAQ-MADRID 2: propandioic acid, C8 dien-dioic acid with an aldehyde branch, C8 hydroxy-dien-dial, C9 hydroxy-carbonyl acid with one double bond, C10 hydroxy-carbonyl aldehyde, and butandioic acid. The molecule fractions of the above species and water content and temperature are treated as independent variables. Given the range of each independent variable (i.e., temperature ranges from 253 to 313 K; mole fraction ranges from 0 to 1), a total of $1.2 \times 10^8$ simulations are conducted (120 temperatures $\times 10^6$ mole fractions of 6 hydrophilic species). The following parameterization expression of activity coefficient is then derived and applied in box MADRID 2:

$$Y = c + \alpha T + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$$

where $Y$ is activity coefficient, $c$ is the intercept (i.e., constant term), $\alpha$ and $\beta$ are vectors representing the polynomial-fitted coefficients, $T$ and $X$ are the matrixes representing temperature in K and molecule fractions of six hydrophilic species and water content. In Eq. (3), the values of $c$ are all zeros, and the values of $\alpha$, and $\beta$ are shown in Table 4. The correlation coefficients between the calculated activity coefficient by the parameterization and the UNIFAC range from 0.6 to 0.9.

The above parameterization reduces the CPU cost by 90%, 54%, 82%, and 52% (corresponding to speed up factors of 10.5, 2.2, 5.6, and 2.1) for rural HBG/LBG and urban HBG/LBG conditions (see Fig. 2), respectively. The optimal speed-up method that combines solver.2, O-I cutoff, and para-AC (i.e., comb_all) can reduce the total CPU cost by 97%, 78%, 89%, and 61% (corresponding to speed up factors of 29.7, 4.5, 8.9, and 2.5) for rural HBG/LBG and urban HBG/LBG conditions, respectively. Compared with the baseline without speed-up, the ranges of percentage deviations in species concentrations are $-14.9$ to $14.6\%$ for major gas-phase species and $-16.0$ to $3.9\%$ for total PM$_{2.5}$ under the rural HBG condition, and $-1.9$ to $0.5\%$ for major gas-phase species and $-0.8$ to $0.4\%$ for total PM$_{2.5}$ under the rural LBG condition.
3.5 Sensitivity test of MADRID 2_Fast

MADRID 2_Fast is further tested under the aforementioned four conditions with 5 temperatures (273, 283, 293, 303, and 313 K) and 5 RHs (10, 40, 60, 80, and 95%), with a total of 100 simulations (=5 temperatures×5 RHs×4 conditions). Figure 3 shows the average percentage deviations from the benchmark for major gas-phase species and total PM$_{2.5}$ under the rural HBG condition at an RH of 10% or 95% and temperatures of 273–313 K and at temperatures of 273 K or 313 K and RHs of 10–95%. For most species, the average deviations are within ±10%. Larger deviations occur for nitrogen dioxide (NO$_2$), nitric oxide (NO), and nitrate radical (NO$_3$) at 273 K and RH≥80% or at RH=95% and temperature ≤283 K, due to their very small concentrations in the baseline simulation. For example, at T=273 K, RH=80%, the mixing ratios of NO decrease from 1.24×10$^{-6}$ ppb at 01:00 p.m. to 6.87×10$^{-19}$ ppb at 11:30 p.m. in the baseline simulation, those in the sensitivity simulation with MADRID 2_Fast decrease from 2.19×10$^{-6}$ at 01:00 p.m. to 1.8×10$^{-19}$ ppb at 11:30 p.m., such changes from the small baseline values at each time step result in a large average percentage deviation of 92.5% (see Fig. 3c), although the absolute changes are negligible. Figure 4 shows a fairly good correlation of simulated AROH and BIOH, total SOA, and total PM$_{2.5}$ between simulations with MADRID 2_Fast and MADRID 2 under the four test conditions during the 24 simulation period, with correlation coefficients of 0.99 to 1.0 and >86% of data points within ±15% deviations. These results illustrate that MADRID 2_Fast provides an overall good compromise between computational efficiency and numerical accuracy under most atmospheric conditions.

4 Summary and future work

A box model, 0-D CMAQ-MADRID 2, is applied to explore various methods in improving computational efficiency of the SOA module. Solving the partitioning equations for condensable VOCs and estimating their activity coefficients with the UNIFAC are iden-
ified as the most computationally-expensive processes in simulating SOA. Potential speedup methods tested include relaxing the error tolerance levels, reducing the maximum number of iterations, and adjusting several internal parameters of the numerical solver; turning off O-I interactions when water content associated organics is relatively small, and parameterizing the UNIFAC calculation of activity coefficients using multiple linear regression analysis at various temperatures for the mixtures with different mole fractions. A combination of several speed-up parameters related to the numerical solver gives 47–55% CPU reductions (speedup by 1.9–2.2) with percentage deviations of −5.5 to 1.7%. The percentage of water associated with organics over the total water from inorganics is selected as an indicator of O-I interactions with an optimal cutoff value of 0.5%. Using the cutoff value to turn on/off O-I interactions can reduce CPU by 7 to 39.3% under the test conditions with the percentage deviations of −7.2 to 7.3% for most species. Among all the methods tested, the parameterization of the activity coefficient calculation in the hydrophilic module gives the most effective CPU reduction (reduced by 51.9 to 90.5%, or speedup by factors of 2.1–10.5) with percentage deviations of −14.9 to 14.6%. The optimal speed-up method that combines all above methods can significantly reduce the CPU cost by 60.5 to 96.6% (speedup by factors of 2.5–29.7) with percentage deviations of −14.9 to 14.6% from the benchmark. Under conditions with typical ranges of temperatures and RHs, the MADRID 2_Fast can reproduce the benchmark results with acceptable accuracy (within ±15%) and computational efficiency for most species. The methods used to develop the MADRID 2_Fast in this study are generic to other SOA modules that employ the UNIFAC calculation for activity coefficients of OCs and/or that use a numerical solver to solve the partitioning equations for OCs. The MADRID 2_Fast is being implemented in 3-D CMAQ-MADRID of Zhang et al. (2004) and the Weather Research and Forecast Model with Chemistry and MADRID (WRF/Chem-MADRID) (Zhang et al., 2005; Hu et al., 2006). It will be further tested with representative episodes in the US and abroad for long-term applications and real-time air quality forecasting.
Acknowledgements. This work is supported by the NSF Award No. Atm-0348819 and under the Memorandum of Understanding between the US EPA and the US Department of Commerce’s NOAA and under agreement number DW13921548. The authors thank G. L. Gipson, a former employee at the US EPA for providing a gas-phase box model of CBM-IV mechanism with the QSSA, SMVGEAR, and ROS3 solvers as a starting point for the implementation of the three solvers for CACM gas-phase mechanism performed in this work. Thanks are also due B. Pun, Atmospheric and Environmental Research (AER), Inc., for helpful discussions regarding MADRID2. This work constitutes a contribution to the NOAA Air Quality Program.

References


Pun, B., Karamchandani, P., Vijayaraghavan, K., Chen, S.-Y., and Seigneur, C.: Models-3/Community Multiscale Air Quality Model (CMAQ) user’s guide to alternative modules: Model of Aerosol Dynamics, Reaction, Ionization, and Dissolution (MADRID), Mercury (Hg), and Advanced Plume Treatment (APT), Electric Power Research Institute, Palo Alto, California, CP194-05-1, 2005.


Table 1. The initial gas-phase species concentrations (ppb)\(^1,2\).

<table>
<thead>
<tr>
<th>Species</th>
<th>Species Name</th>
<th>Rural LBG(^3)</th>
<th>Rural HBG(^3)</th>
<th>Urban LBG(^3)</th>
<th>Urban HBG(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>nitric oxide</td>
<td>1.5</td>
<td>1.5</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>NO(_2)</td>
<td>nitrogen oxide</td>
<td>0.5</td>
<td>0.5</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>HONO</td>
<td>nitrous acid</td>
<td>1.0</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>O(_3)</td>
<td>ozone</td>
<td>40</td>
<td>40</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>CO</td>
<td>carbon monoxide</td>
<td>120</td>
<td>120</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>HCHO</td>
<td>formaldehyde</td>
<td>0.2</td>
<td>0.2</td>
<td>8.3</td>
<td>8.3</td>
</tr>
<tr>
<td>ALD2</td>
<td>lumped high aldehydes</td>
<td>(8.3 \times 10^{-2})</td>
<td>(8.3 \times 10^{-2})</td>
<td>11.7</td>
<td>11.7</td>
</tr>
<tr>
<td>PAN2</td>
<td>peroxy acetyl nitrate</td>
<td>1.0</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>ALKL</td>
<td>lumped alkanes C(_2)–C(_6)</td>
<td>22.755</td>
<td>22.755</td>
<td>68.265</td>
<td>68.265</td>
</tr>
<tr>
<td>ALKM</td>
<td>lumped alkanes C(<em>7)–C(</em>{12})</td>
<td>5.565</td>
<td>5.565</td>
<td>16.695</td>
<td>16.695</td>
</tr>
<tr>
<td>ALKH</td>
<td>lumped alkanes C(<em>{12})–C(</em>{6})</td>
<td>(7.5 \times 10^{-2})</td>
<td>(7.5 \times 10^{-2})</td>
<td>0.225</td>
<td>0.225</td>
</tr>
<tr>
<td>OLEL</td>
<td>lumped alkenes C(_3)–C(_6)</td>
<td>8.2</td>
<td>8.2</td>
<td>24.6</td>
<td>24.6</td>
</tr>
<tr>
<td>OLEH</td>
<td>lumped alkenes C(_{6})–C(_6)</td>
<td>(7.0 \times 10^{-2})</td>
<td>(7.0 \times 10^{-2})</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>ETHE</td>
<td>ethane</td>
<td>2.18</td>
<td>2.18</td>
<td>6.54</td>
<td>6.54</td>
</tr>
<tr>
<td>ISOP(^4)</td>
<td>isoprene</td>
<td>0.1</td>
<td>2.0</td>
<td>0.1</td>
<td>2.0</td>
</tr>
<tr>
<td>BIOL(^4)</td>
<td>lumped low SOA yield monoterpenes species</td>
<td>0.2</td>
<td>4.0</td>
<td>0.2</td>
<td>4.0</td>
</tr>
<tr>
<td>BIOH(^4)</td>
<td>lumped high SOA yield monoterpenes species</td>
<td>0.2</td>
<td>4.0</td>
<td>0.2</td>
<td>4.0</td>
</tr>
<tr>
<td>SO(_2)</td>
<td>sulfur dioxide</td>
<td>2.0</td>
<td>2.0</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>NH(_3)</td>
<td>ammonia</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

\(^1\) Initial conditions are taken from Zhang et al. (1998) or otherwise noted.
\(^2\) The total NMHC concentrations are assumed to be 50 and 150 ppbC for rural and urban conditions with speciation distribution factor of Griffin et al. (2002).
\(^3\) LBG and HBG denote the conditions with low and high biogenic VOCs.
\(^4\) The concentrations of ISOP, BIOL, and BIOH are assumed based on reported measurements in Finlayson-Pitts and Pitts (1986).
Table 2. The initial PM mass (µg m\(^{-3}\)) and number (particles cm\(^{-3}\)) concentrations.

<table>
<thead>
<tr>
<th>Species</th>
<th>Species Name</th>
<th>Sect. 1</th>
<th>Sect. 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na(^+)</td>
<td>sodium ion</td>
<td>0.0013</td>
<td>2.1908</td>
<td>2.19</td>
</tr>
<tr>
<td>SO(_4^{2-})</td>
<td>sulfate ion</td>
<td>2.0869</td>
<td>0.5162</td>
<td>2.60</td>
</tr>
<tr>
<td>NH(_4^+)</td>
<td>ammonium ion</td>
<td>0.7913</td>
<td>0.2921</td>
<td>1.08</td>
</tr>
<tr>
<td>NO(_3^-)</td>
<td>nitrate ion</td>
<td>0.4259</td>
<td>0.9445</td>
<td>1.37</td>
</tr>
<tr>
<td>Cl(^-)</td>
<td>chloride ion</td>
<td>0.0302</td>
<td>0.7887</td>
<td>0.82</td>
</tr>
<tr>
<td>CO(_3^{2-})</td>
<td>carbonate ion</td>
<td>0.0094</td>
<td>0.0093</td>
<td>0.02</td>
</tr>
<tr>
<td>H(_2)O</td>
<td>particulate water</td>
<td>0.9448</td>
<td>0.9448</td>
<td>1.89</td>
</tr>
<tr>
<td>OI</td>
<td>other inorganic material</td>
<td>2.9357</td>
<td>6.4906</td>
<td>9.43</td>
</tr>
<tr>
<td>EC</td>
<td>elemental carbon</td>
<td>0.0891</td>
<td>0.1071</td>
<td>0.20</td>
</tr>
<tr>
<td>OO</td>
<td>primary organic carbon</td>
<td>0.9298</td>
<td>0.3993</td>
<td>1.33</td>
</tr>
<tr>
<td>NUM</td>
<td>particle number concentration</td>
<td>2.24×10(^6)</td>
<td>3.34×10(^4)</td>
<td>2.27×10(^6)</td>
</tr>
</tbody>
</table>
Table 3. The default value and function of parameters in globally-convergent Newton/line search method.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default value</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOLF</td>
<td>$1 \times 10^{-4}$</td>
<td>The convergence criterion on species concentrations, $\Delta x$.</td>
</tr>
<tr>
<td>TOLMIN</td>
<td>$1 \times 10^{-6}$</td>
<td>Criterion for deciding whether spurious convergence to a minimum has occurred.</td>
</tr>
<tr>
<td>TOLX</td>
<td>$1 \times 10^{-7}$</td>
<td>The convergence criterion on the correction to the species concentrations, $\delta x$.</td>
</tr>
<tr>
<td>MAXIT</td>
<td>400</td>
<td>The maximum number of iterations.</td>
</tr>
<tr>
<td>EPS</td>
<td>$1 \times 10^{-4}$</td>
<td>The approximate square root of the machine precision.</td>
</tr>
<tr>
<td>ALF</td>
<td>$1 \times 10^{-4}$</td>
<td>A parameter that ensures sufficient decrease in species concentrations.</td>
</tr>
</tbody>
</table>
Table 4. Polynomial-fitted coefficients derived in the parameterization of activity coefficient.

<table>
<thead>
<tr>
<th>Species</th>
<th>$\alpha$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\beta_5$</th>
<th>$\beta_6$</th>
<th>$\beta_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.9681E–04</td>
<td>9.3051E–01</td>
<td>1.7609E+00</td>
<td>2.2175E+00</td>
<td>1.1057E+00</td>
<td>1.1074E+00</td>
<td>1.2790E+00</td>
<td>6.3348E–01</td>
</tr>
<tr>
<td>3</td>
<td>-2.5780E–03</td>
<td>4.0929E+00</td>
<td>1.7900E+00</td>
<td>6.5160E–01</td>
<td>1.2301E+00</td>
<td>1.6631E+00</td>
<td>3.5964E+00</td>
<td>4.1680E+00</td>
</tr>
</tbody>
</table>
Fig. 1. The CPU time cost as a function of (a) various tolerance levels of TOLX by fixing values of TOLF; (b) various maximum iteration numbers; and (c) various combination of speed-up parameters, where solver_1, solver_2, and solver_3 correspond to TOLF = $1 \times 10^{-3}$, TOLX = $1 \times 10^{-5}$, MAXITS = 100, EPS = $1 \times 10^{-4}$, and ALF = $1 \times 10^{-4}$; TOLF = $1 \times 10^{-3}$, TOLX = $1 \times 10^{-3}$, MAXITS = 100, EPS = $1 \times 10^{-4}$, and ALF = $1 \times 10^{-4}$; and TOLF = $1 \times 10^{-2}$, TOLX = $1 \times 10^{-3}$, MAXITS = 5, EPS = $1 \times 10^{-3}$, and ALF = $1 \times 10^{-3}$, respectively.
**Fig. 2.** The CPU cost for various scenarios, baseline (benchmark), solver_2 (i.e., TOLF=1×10^-3, TOLX=1×10^-3, MAXITS=100, EPS=1×10^-4, and ALF=1×10^-4), using a cutoff value in organic-inorganic interactions (O-I cutoff), parameterization of activity coefficients calculation (para-AC), and a combination of all these methods (comb_all) under four conditions using the ROS3 solver.
**Fig. 3.** Average percentage deviations from benchmark (i.e., the results from SMVGEAR with the most stringent error tolerances.) of major gas-phase species and total PM during the 24-h simulation period at (a) RH=10%, T=273–313 K, (b) RH=95%, T=273–313 K, (c) T=273 K, RH=10–95%, and (d) T=313 K, RH=10–95% under the rural high biogenic VOC condition (R_HBG) using the improved SOA module.