Atmos. Chem. Phys., 8, 3985–3998, 2008
www.atmos-chem-phys.net/8/3985/2008/
© Author(s) 2008. This work is distributed under
the Creative Commons Attribution 3.0 License.

Atmospheric Chemistry and Physics

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 University, Raleigh, NC 27695, USA

Received: 10 March 2008 – Published in Atmos. Chem. Phys. Discuss.: 10 April 2008
Revised: 17 June 2008 – Accepted: 17 June 2008 – Published: 24 July 2008

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; conditionally activating organic-inorganic interactions; and parameterizing the calculation of activity coefficients for organic mixtures in the hydrophilic module. The optimal speed-up method can reduce the total CPU cost by up to a factor of 31.4 from benchmark under the rural conditions with 2 ppb isoprene and by factors of 10–71 under various test conditions with 2–10 ppb isoprene and >40% relative humidity while maintaining ±15% deviation. These speed-up 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). SOA can be formed through the oxidation of the volatile organic compounds (VOCs) in the atmosphere and subsequent partitioning of their condensable products between gas- and particulate-phase. The amount of SOA formed depends on atmospheric abundance of anthropogenic and biogenic VOCs and their condensable products from photochemical oxidation reactions, as well as the chemical reactivity, solubility, and efficiency of gas/particle partitioning 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 poses 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: the Model of Aerosol Dynamics, Reaction, Ionization, and Dissolution 1 and 2 (MADRID 1 and MADRID 2) have been incorporated into the U.S. 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, they use different methods for partitioning coefficients for hydrophobic organic compounds (OC), despite the same partitioning equation. The partitioning coefficients used in MADRID 1 are obtained from the smog chamber

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

experiments. In MADRID 2, the partitioning coefficients of the hydrophobic OCs are calculated as a function of temperature and composition based on the Raoult's law following the equation derived by Pankow (1994a, 1994b). Those of the hydrophilic OCs are calculated as a function of temperature, liquid water content, and composition based on the Henry's law. The determination of chemical composition for both hydrophobic and hydrophilic OCs in MADRID 2 involves the use of the UNIversal Functional Activity Coefficient (UNIFAC) method of Fredenslund et al. (1975, 1977) for calculation of activity coefficients. Second, the hydrophilic compounds are not treated in MADRID 1 but they are treated based on the Henry's Law in MADRID 2. 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 speed-up methods while maintaining a desirable numerical accuracy. The main objective of this study is to improve the computational efficiency of the SOA module for 3-D air quality long-term simulations and real-time forecasting.

2 Model description and test conditions

2.1 Model description

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 Pun et al. (2002) and 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 the 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 ($\leq$7 carbons; or $\leq$10 carbons with three or more functional groups), high solubility ($\geq$1 g solute / 100 g water), and a high effective Henry's law constant ($\geq 1\times10^6$ M atm^{-1}). Hydrophobic condensable OCs are identified by their estimated octanol-water partitioning coefficients. If an organic species shows affinity for both aqueous and organic phases, it is chosen as either hydrophobic or hydrophilic, due to the lack of experimental data on thermodynamic properties of mixed organic-inorganic aqueous solutions (Pun et al., 2002). 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 (Pankow, 1994a, 1994b):

$$K_i = \frac{A_i / M_{\mathrm{TOM}}}{G_i} \tag{1}$$

where K_i (m$^3\mu$g^{-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_{TOM} (μg m^{-3} air) is the mass of total organic carbon (i.e., 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_{TOM}. The partitioning coefficient, K_i, can be calculated as follows (Pankow et al., 1994a, 1994b):

$$K_i = \frac{760RT}{10^6 P_i^{sat} \gamma_i \bar{M}_{om}} \tag{2}$$

where R is the ideal gas constant (8.2×10^{-5} m^3 atm mol^{-1} K^{-1}), T (K) is the temperature, P_i^{sat} (torr) is the saturation vapor pressure, γ_i is the activity coefficient in the particulate 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, γ_i can be calculated by the UNIFAC method based on initial guess of organic aerosol concentrations and composition, 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 continue until the deviation meets a convergence criterion of 10^{-4}, making the hydrophobic module in MADRID 2 computationally more expensive than the empirical absorptive partitioning module in MADRID 1.

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 the Henry's law assuming equilibrium between the bulk gas and whole particulate phase, and takes into account the activity coefficient of the molecular solutes (Pun et al., 2002):

$$HL_i = \frac{c_{wi}}{p_i} = \frac{A_i \gamma_{wi}}{M_i(\mathrm{LWC})\, p_i} \tag{3}$$

where HL_i (μM atm^{-1}) is the Henry's law constant of species i, c_{wi} is aqueous phase concentration (μmol L^{-1} of water), p_i (atm) is gas phase partial pressure, A_i (μg m^{-3}) is aqueous phase concentration, M_i is the molecular weight of the compound i (g mol^{-1}), γ_{wi} is the activity coefficient of species i in the aqueous mixture (i.e., infinite dilution reference state), and LWC is the aerosol liquid water content (l water m^{-3} air). The condensable species partitioning into the aqueous phase can be calculated through the Henry's law constant, LWC, and γ_{wi}. The Henry's law constant can be measured, or derived using a group contribution method based on the structure of the organic compound following the approach of Hine and Mookerjee (1975) and Suzuki et al. (1992). LWC is calculated from the Zdanovskii-Stokes-Robinson (ZSR) equation of Stokes and Robinson (1966) and the activity coefficients of organic solutes are calculated through UNIFAC. The calculation steps in the hydrophilic module include the determination of the type of calculation needed (saturation vs. aqueous equilibrium), the iterative solution of mass balance, Henry's law, acid dissociation constant equations (including iteration in UNIFAC to determine activity coefficients), the calculation of water uptake associated with aqueous organics, and iteration on LWC between the hydrophilic module and an inorganic module, ISORROPIA of Nenes et al. (1999). Different from γ_i calculation in the hydrophobic module, additional iteration using a Newton line search method in UNIFAC is required in the hydrophilic module to achieve aqueous-phase equilibrium based on input LWC, because γ_{wi} is a strong function of the aerosol composition that depends on the aqueous-phase dissolution/dissociation equilibria. Numerous iterations involved in the hydrophilic module make MADRID 2 computationally very demanding.

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 ISORROPIA inorganic equilibrium model 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 a number of OCs. A globally-convergent Newton/line search method is used to solve 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 ZSR equation). In the hydrophobic module, M_{TOM} 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 conditions and model inputs

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. The use of such a box model isolates gaseous and aerosol chemical kinetics from other atmospheric processes such as emissions, transport, and depositions, on which the speedup is focused in this study. 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, 1997b), 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 8 a.m. Eastern Daylight Time (EDT)), with a time step of 1 min for integration of gas-phase chemistry and aerosol processes. A diurnal photolytic rate profile is used to represent a typical summer day in the eastern U.S. 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

Table 1. The initial gas-phase species concentrations (ppb)[1].

Species	Species Name	Rural LBG[2]	Rural HBG[2]	Urban LBG[2]	Urban HBG[2]
NO	nitric oxide	1.5	1.5	30	30
NO_2	nitrogen oxide	0.5	0.5	15	15
HONO	nitrous acid	1.0	1.0	2.0	2.0
O_3	ozone	40	40	80	80
CO	carbon monoxide	120	120	150	150
HCHO	formaldehyde	0.2	0.2	8.3	8.3
ALD2	lumped high aldehydes	8.3×10^{-2}	8.3×10^{-2}	11.7	11.7
PAN2	peroxy acetyl nitrate	1.0	1.0	2.0	2.0
ALKL	lumped alkanes C_2–C_6	22.755	22.755	68.265	68.265
ALKM	lumped alkanes C_7–C_{12}	5.565	5.565	16.695	16.695
ALKH	lumped alkanes $>C_{12}$	7.5×10^{-2}	7.5×10^{-2}	0.225	0.225
OLEL	lumped alkenes C_3–C_6	8.2	8.2	24.6	24.6
OLEH	lumped alkenes $>C_6$	7.0×10^{-2}	7.0×10^{-2}	0.21	0.21
ETHE	ethane	2.18	2.18	6.54	6.54
AROH	lumped high SOA yield aromatic species	4.995	4.995	14.985	14.985
AROL	lumped low SOA yield aromatic species	6.615	6.615	19.845	19.845
ISOP	isoprene	0.1	2.0	0.1	2.0
BIOL	lumped low SOA yield monoterpene species	0.2	4.0	0.2	4.0
BIOH	lumped high SOA yield monoterpene species	0.2	4.0	0.2	4.0
SO_2	sulfur dioxide	2.0	2.0	20	20
NH_3	ammonia	2.0	2.0	2.0	2.0
HCl	hydrochloric acid	2.0	2.0	2.0	2.0

[1] Initial conditions are taken from Zhang et al. (1998) or otherwise noted.
[2] LBG and HBG denote the conditions with low and high biogenic VOCs.
[3] 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. (2002a) for ETHE, ALKL, ALKM, ALKH, OLEL, OLEH, AROL, AROH, ISOP, BIOL, and BIOH.
[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.

Species	Species Name	Section 1[1]	Section 2[1]	Total
Na^+	sodium ion	0.0013	2.1908	2.19
SO_4^{2-}	sulfate ion	2.0869	0.5162	2.60
NH_4^+	ammonium ion	0.7913	0.2921	1.08
NO_3^-	nitrate ion	0.4259	0.9445	1.37
Cl^-	chloride ion	0.0302	0.7887	0.82
CO_3^{2-}	carbonate ion	0.0094	0.0093	0.02
H_2O	particulate water	0.9448	0.9448	1.89
OI	other inorganic material	2.9357	6.4906	9.43
EC	elemental carbon	0.0891	0.1071	0.20
OO	primary organic carbon	0.9298	0.3993	1.33
NUM	particle number concentration	2.24×10^6	3.34×10^4	2.27×10^6

[1] Two particle sections are used in MADRID 2 simulation: Sects. 1 and 2 represent particles with the stoke diameter from 0.0215 to 2.15 μm and 2.15 to 10 μm, respectively (where a stoke diameter of 2.15 μm corresponds to an aerodynamic diameter of 2.5 μm for a density of 1.352 g cm^{-3}).

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). Following the approach of Griffin et al. (2002a) and Pun et al. (2003), the same speciation is applied to the entire simulation domain because of lack of data to distinguish speciation in rural and urban areas. 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; NTP, 2005). The only differences between rural/urban HBG and rural/urban LBG are the concentrations of ISOP and BIOL+BIOH (0.1 and 0.4 ppb, respectively, for rural and urban LBG, and 2 and 8 ppb, respectively, for rural and urban HBG). The concentration of ISOP of 2 ppb

used represents typical observations in the U.S. on a regional scale. Non-typical conditions with a high ISOP concentration of 10 ppb will be used to evaluate the improved MADRID 2 in Sect. 3.5. The initial concentrations of other inorganic gas-phase species are based on Zhang et al. (1998). The initial PM mass and number concentrations are assumed to be 20.9 μg m^{-3} and 2.27×10^6 particle cm^{-3}, respectively, for all the four conditions, which minimizes the effects of preexisting PM on the predicted SOA under different conditions. Two scenarios are considered for each condition: without PM (i.e., CACM gas-phase chemistry only) and with PM. All simulations are conducted on the Linux Cluster of IBM Blade Center of North Carolina State University, with 2.8–3.2 GHz dual Xeon compute nodes. The simulation results are written out every 30 min.

3 Improvement of computational 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. Although solving gas-phase chemistry is much less expensive than solving aqueous-phase chemistry and aerosol processes, an accurate yet computationally-efficient solver for gas-phase mechanism will yield appreciable computational benefits for 3-D simulations that cover various conditions over thousands of grid cells and a long simulation period. Such a solver, once identified, will be used to perform simulations with both gases and aerosols before and after speedup of MADRID 2. Among the three solvers implemented for CACM, SMVGEAR is widely used in atmospheric transport models (Hertel et al., 1993; Huang and Chang, 2001) and considered to be the most accurate numerical chemical solver available (U.S. 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. For example, the average percentage deviations for most species are less than 1% for both solvers under all the test conditions (e.g., $<$0.3%, 0%, $-$0.8% to 0.6%, and $<$0.5% for O_3 mixing ratios under the rural HBG/LBG and the urban HBG/LBG conditions). The maximum percentage deviations averaged over all time steps occur in the mixing ratios of HO_2 radical (i.e., 1.7% for the ROS3 solver and 2.5% for the QSSA solver) and the maximum percentage deviations at an instantaneous time step also occur for HO_2 radical (i.e., 4.2% for the ROS3 solver and 5.6% for the QSSA solver). These results are generally consistent with those of Sandu et al. (1996, 1997a, 1997b) and Hesstvedt et al. (1978).

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 seconds) (by factors of 8.6 and 3.6, respectively) than those with LBG conditions, with the highest CPU cost for the rural HBG condition. Compared with conditions with low biogenic VOCs, higher CPU costs are required under high biogenic VOCs conditions, due to the abundance of biogenic SOA precursors that are formed from the oxidation of their parent species and will partition into particulate phase. Such oxidation and partitioning are affected by concentrations of other relevant species such as anthropogenic SOA precursors and the total oxidation capacities. The rural HBG condition is more computationally expensive than the urban HBG condition, because the calculation of activity coefficients for hydrophobic and hydrophilic SOA precursors using UNIFAC requires more significant CPU and the solution to their partitioning equations takes more time steps to converge. Considering its good compromise between accuracy and computational efficiency and potential computational savings in 3-D simulations, 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, two physical processes are identified as computationally most expensive among all aerosol processes simulated: solving partitioning equations for OCs using the globally-convergent Newton/line search method and calculating activity coefficients of multicomponent mixtures through UNIFAC. They account for 2.9–4.6% and 56.1–92.3%, respectively, of total CPU cost under the four test conditions, with the CPU cost of UNIFAC much more expensive (by factors of 19.2–20.2) than that of partitioning calculation under the four test conditions. The hydrophilic module is more expensive than the hydrophobic module; it takes 65–97% of total CPU cost under the four conditions, due to its excessive iterations involving UNIFAC activity coefficient calculations and aqueous-phase equilibrium. 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 2 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 temperatures of 253–313 K and RHs of 10–95% under the four test conditions and

Table 3. The default value and function of parameters in globally-convergent Newton/line search method.

Parameter	Default value	Definition
TOLF	1×10^{-4}	The convergence criterion on species concentrations, ΔA_i
TOLMIN	1×10^{-6}	Criterion for deciding whether spurious convergence to a minimum has occurred
TOLX	1×10^{-7}	The convergence criterion on the correction to the species concentrations, δA_i
MAXITS	400	The maximum number of iterations
EPS	1×10^{-4}	The approximate square root of the machine precision
ALF	1×10^{-4}	A parameter that ensures sufficient decrease in function values (i.e., species concentrations) and ensures the algorithm's convergence

two additional non-typical conditions: the same as the rural HBG condition except with a very high mixing ratio of isoprene of 10 ppb (rather than 2 ppb) and the same as the rural HBG condition except with an extremely low nitrogen oxides (NO_x) (75 ppt nitric oxide (NO) and 25 ppt nitrogen dioxide (NO_2), rather than 1.5 ppb NO and 0.5 ppb NO_2) and a very high mixing ratio of isoprene of 10 ppb (rather than 2 ppb). They represent the rural very high HBG and the forest conditions, respectively (referred to as the rural VHBG and the forest LN_VHBG conditions, respectively).

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., MAXITS). 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=1×10^{-3}, relaxing TOLX from 1×10^{-6} to 1×10^{-3} results in 6.2 to 47.1% CPU reduction. At TOLF=1×10^{-2}, relaxing TOLX from 1×10^{-6} to 1×10^{-3} results in 15.3 to 48.2% CPU reduction. As shown in Fig. 1b, reducing MAXITS 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 1×10^{-4} to 1×10^{-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 three combinations of speed-up parameters, where solver_1, solver_2, and solver_3 correspond to a combination of TOLF=1×10^{-3}, TOLX=1×10^{-5}, MAXITS=100, EPS=1×10^{-4}, and ALF=1×10^{-4}; the same with solver_1 but with TOLX=1×10^{-3}; and a combination of TOLF=1×10^{-2}, TOLX=1×10^{-3}, MAXITS=5, EPS=1×10^{-3}, and ALF=1×10^{-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 maximum instantaneous 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, the other two combinations, in particular, solver_2, 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 on not only the amounts and properties of the compound, but also 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-MADRID 2, 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 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., LWCORG) and the total water content from inorganics (i.e., RWATER) may vary significantly (e.g., 10^{-4} to 2 and 3–15 $\mu g\,m^{-3}$, respectively), their ratios (i.e., LWCORG/RWATER) 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 LWCORG/RWATER 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 LWCORG/RWATER 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 are reduced by 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 maximum instantaneous percentage deviations are −16.7 to 49.2% for major gas-phase species and

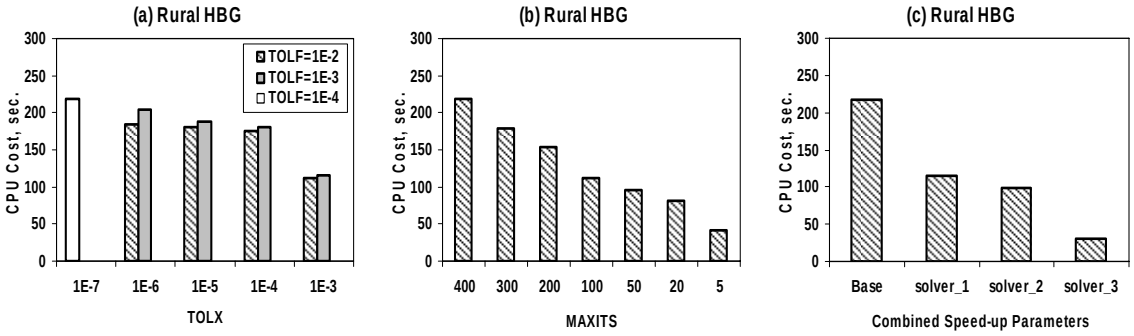

Fig. 1. The CPU time cost as a function of **(a)** various tolerance levels of TOLX by fixing values of TOLF; **(b)** various maximum iteration numbers; and **(c)** various combination of speed-up parameters, where solver_1, solver_2, and solver_3 correspond to TOLF=1×10^{-3}, TOLX=1×10^{-5}, MAXITS=100, EPS=1×10^{-4}, and ALF=1×10^{-4}; TOLF=1×10^{-3}, TOLX=1×10^{-3}, MAXITS=100, EPS=1×10^{-4}, and ALF=1×10^{-4}; and TOLF=1×10^{-2}, TOLX=1×10^{-3}, MAXITS=5, EPS=1×10^{-3}, and ALF=1×10^{-3}, respectively.

−21.2 to 17.7% for total $PM_{2.5}$ for a cutoff value of 20% for LWCORG/RWATER. The cutoff value of 0.5% can reduce CPU cost by 7% with relatively small maximum instantaneous 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 LWCORG/RWATER 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 LWCORG/RWATER >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 the relative amount of water content associated with organics are smaller and LWCORG/RWATER 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 LWCORG/RWATER between the two conditions.

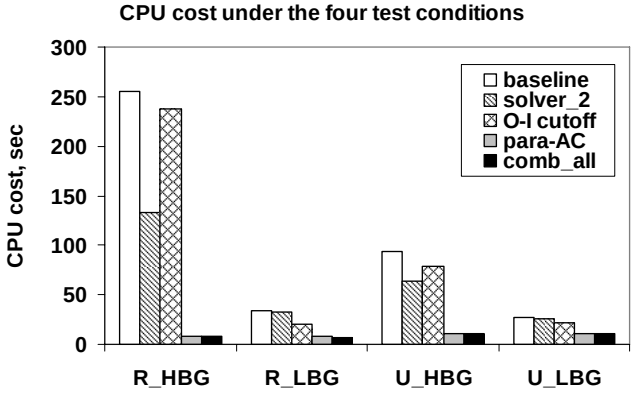

Fig. 2. The CPU cost for various scenarios, baseline (benchmark), solver_2 (i.e., TOLF=1×10^{-3}, TOLX=1×10^{-3}, MAXITS=100, EPS=1×10^{-4}, and ALF=1×10^{-4}), using a cutoff value in organic-inorganic interactions (O-I cutoff), parameterization of activity coefficients calculation (para-AC) , and a combination of all these methods (comb_all) under four conditions using the ROS3 solver.

3.4 Parameterization of activity coefficient calculation

The accuracy of the partitioning coefficients highly depends on the reliability of the calculated activity coefficients in aerosol simulation. The current approaches used for activity coefficient calculations in aerosol modeling range from the simplest method by assuming an ideal solution with a value of 1 for all SOA precursors (e.g., CMAQ, Schell et al., 2001; Binkowski and Roselle, 2003) to simple parameterizations such as the multicomponent Wilson equation or the NonRandom Two-Liquid (NRTL) equation of Prausnitz et al. (1999) (e.g., Bowman and Karamalegos, 2002), or to the most sophisticated methods such as the use of UNIFAC (e.g., CMAQ-MADRID, Pun et al., 2002; Zhang et al., 2004). UNIFAC is a semi-empirical system based on a group contribution method that has been used widely to calculate non-electrolyte activity coefficients in non-ideal mixtures. It treats solution as a mixture of a number of functional groups and accounts for interactions among these groups. Among these alternative approaches for activity coefficient calculations, UNIFAC, despite its significant computational burden, is the most accurate method and shows better performance than other activity coefficient models for multicomponent aerosol mixtures against laboratory measurements (Bowman and Melton, 2004). UNIFAC can be used to calculate activity coefficients at any solute concentrations and temperatures (Lyman et al., 1990) and has been extensively evaluated under a temperature range of 275 to 400 K (e.g., Fredenslund et al., 1975; Bowman and Melton, 2004). As expected, the calculation of activity coefficients using UNIFAC in aqueous mixtures in the hydrophilic module in MADRID 2 is identified to be the most computationally-expensive process. In this study, the UNIFAC-calculated activity coefficients are parameterized using a multiple linear regression analysis based on the least squares approach (referred to as

Table 4. Polynomially-fitted coefficients derived in the parameterization of activity coefficient in the hydrophilic module.

Species i	c_i	α_i	$\beta_{1,j}$	$\beta_{2,j}$	$\beta_{3,j}$	$\beta_{4,j}$	$\beta_{5,j}$	$\beta_{6,j}$
1	6.3348E-01	−1.9681E-04	2.9703E-01	1.1274E+00	1.5841E+00	4.7219E-01	4.7393E-01	6.4553E-01
2	3.8327E-01	1.5469E-03	5.9834E-01	2.3660E-01	2.0210E-02	−3.8780E-01	−3.8173E-01	5.8548E-01
3	4.1680E+00	−2.5780E-03	−7.5121E-02	−2.3780E+00	−3.5164E+00	−2.9379E+00	−2.5049E+00	−5.7167E-01
4	1.2160E+00	1.6760E-03	−8.6314E-01	−1.3003E+00	−1.1568E+00	−8.2380E-01	−1.1698E+00	−1.1147E+00
5	1.4972E+00	2.9238E-03	−1.1053E+00	−1.8553E+00	−1.5649E+00	−1.7719E+00	−1.7171E+00	−1.5855E+00
6	5.6509E-01	2.7920E-04	4.2956E-01	9.1606E-01	1.1775E+00	1.2487E-01	4.8657E-02	3.6076E-01
7	1.0734E+00	1.4214E-03	−6.7259E-01	−7.0190E-01	−1.6368E-01	−2.1953E-01	1.2504E-01	−5.8113E-01

Note: Species 1–7 represent 5 hydrophilic surrogates, butandioic acid, and water. The surrogates include propandioic acid (C2), diendioic acid with an aldehyde branch (C8), hydroxy-dien-dial (C8), hydroxy-carbonyl acid with one double bond (C9), and hydroxy-carbonyl aldehyde (C10). The sum of the mole fractions of the 7 species must be equal to 1. The mole fraction of water is therefore not considered as an independent species in the polynomial equation.

Table 5. Performance statistics of MADRID 2_Fast against MADRID 2 benchmark under the rural HBG condition[1].

Species name	Species symbol	Bench-mark	MADRID 2_Fast	Corr.	MB[2]	RMSE[2]	NMB (%)	NME (%)	% deviation[1] Min	Max	Ave
nitrogen dioxide	NO_2	0.03	0.03	1.00	−4.5E–4	2.9E–3	−1.30	1.30	−8.94	4.94	−2.00
nitric oxide	NO	0.03	0.03	1.00	−4.9E–6	2.0E–5	−0.02	0.02	−5.74	14.58	4.42
ozone	O_3	40.47	40.50	1.00	3.3E–2	5.7E–2	0.08	0.08	0.00	0.27	0.13
nitrate radical	NO_3	0.00	0.00	1.00	−1.5E–8	6.5E–8	−3.16	3.18	−10.11	4.74	−2.69
hydroxyl radical	OH	0.00	0.00	1.00	1.9E–7	6.9E–7	0.71	1.34	−4.58	6.64	1.03
hydroperoxy radical	HO_2	0.01	0.01	1.00	3.0E-4	5.8E-4	3.56	4.49	−5.49	11.49	3.00
carbon monoxide	CO	133.60	133.60	1.00	2.0E-2	1.4E-1	0.02	0.02	0.00	0.72	0.36
hydrogen peroxide	H_2O_2	3.07	3.02	1.00	−5.3E–2	5.6E–2	−1.74	1.74	−5.41	−1.12	−3.27
sulfur dioxide	SO_2	1.91	1.91	1.00	−4.9E–9	2.0E–3	0	0.02	−0.53	0.51	−0.01
methyl peroxy radical from oxidation of CH_4	RO21	0.02	0.01	0.99	−1.7E–3	2.2E–3	−10.57	10.99	−14.86	4.57	−5.15
formaldehyde	HCHO	2.79	2.79	1.00	−3.5E–3	1.8E–2	−0.12	0.30	−2.86	0.51	−1.17
ethane	ETHE	1.43	1.43	1.00	−1.2E–3	4.9E-3	−0.09	0.17	−0.88	0.57	−0.16
hydroxy alkyl peroxy radical <C6 from oxidation of ETHE, ETOH, OLEL, and ALCH (C4, 1-peroxy, 2-hydroxy)	RO22	0.05	0.05	1.00	−1.7E–3	3.2E–3	−3.34	5.21	−10.45	6.54	−1.96
lumped higher aldehydes (n-pentanal)	ALD2	6.62	6.62	1.00	−7.1E–3	1.8E–2	−0.11	0.16	−3.18	0.15	−1.52
lumped alkenes C3-C6 (1-pentene)	OLEL	2.02	2.02	1.00	6.1E–4	1.2E–2	0.03	0.44	−0.97	0.86	−0.05
lumped alkanes C2-C6 (2-methyl-butane)	ALKL	19.27	19.26	1.00	−4.1E–3	2.9E–2	−0.02	0.04	−0.55	0.47	−0.04
isoprene	ISOP	0.15	0.15	1.00	1.1E–3	3.7E–3	0.72	0.79	−3.18	3.27	0.04
lumped alkenes >C6 (4-methyl-1-octene)	OLEH	0.02	0.02	1.00	6.5E–6	1.1E–4	0.04	0.44	−1.21	1.26	0.02
lumped alkanes C7-C12 (3,5-dimethyl-heptane)	ALKM	3.60	3.59	1.00	−4.3E–3	8.7E–3	−0.12	0.20	−0.34	0.40	0.03
branched alkyl peroxy radical >C6 from oxidation of OLEH and ALKM (C7 chain, 3-methyl, 4-peroxy)	RO220	0.03	0.02	0.99	−1.4E–3	1.8E–3	−5.20	6.07	−9.89	6.10	−1.89
lumped low SOA yield aromatic species (1,2,3-trimethyl-benzene)	AROL	1.95	1.95	1.00	−8.6E–4	1.2E–2	−0.04	0.44	−1.01	1.05	0.02
lumped high SOA yield aromatic species (3-n-propyl-toluene)	AROH	2.66	2.66	1.00	−2.2E–3	8.5E–3	−0.08	0.25	−0.51	0.47	−0.02
lumped low SOA yield monoterpene species (a-terpineol)	BIOL	0.13	0.13	1.00	4.7E–4	2.5E–3	0.37	0.37	−7.27	5.33	−0.97
lumped high SOA yield monoterpene species (g-terpinene)	BIOH	0.14	0.14	1.00	7.3E–4	3.6E–3	0.52	0.52	−6.03	5.71	−0.16
lumped alkanes >C12 (n-hexadecane)	ALKH	0.03	0.03	1.00	−4.5E–5	1.3E–4	−0.13	0.30	−0.90	0.65	−0.12
particles with aerodynamic diameter ≤2.5 μm	$PM_{2.5}$	59.41	52.07	0.93	−7.3E+0	8.3E+0	−12.36	12.57	−15.02	3.86	−5.58

1. corr - correlation coefficient, MB - mean bias, RMSE - root mean square error, NMB - normalized mean bias, and NME- normalized mean error, and Min, Max, and Ave % deviations – minimum, maximum, and average instantaneous percentage deviations at a specific time step during the 24-h simulation time period.
2. The units for MB and RMSE are ppb for gas-phase species and $\mu g\,m^{-3}$ for $PM_{2.5}$.

para-AC) to increase computational efficiency. Compared with other alternative methods for activity coefficient calculations, the accuracy of UNIFAC provides a basis for its selection for parameterization. The parameterization is conducted for a temperature range of 253–313 K and an RH range of 10–95%, based on typical surface temperatures and RHs observed over the continental U.S.

Seven molecules are treated for hydrophilic OC partitioning in CMAQ-MADRID 2: propandioic acid (C2), diendioic acid with an aldehyde branch (C8), hydroxy-dien-dial (C8), hydroxy-carbonyl acid with one double bond (C9), hydroxy-carbonyl aldehyde (C10), butandioic acid, and water. To apply multiple linear regression analysis for parameterization, the molecule fractions of the above species

(excluding water) and temperature are treated as independent variables. Because the sum of the mole fractions of the seven species must be equal to 1, the mole fraction of water depends on the sum of the six species, making it a dependent variable. Given the range of each independent variable (i.e., temperatures of 253 to 313 K with an interval of 0.5 K; mole fractions of 0 to 1 with an interval of 0.1), a total of $\sim 1.16\times10^8$ simulations are conducted (120 temperatures $\times$ 968 968 mole fractions of 6 hydrophilic species). The following parameterization expression of activity coefficient is then derived using UNIFAC-calculated activity coefficients and applied in box MADRID 2:

$$Y_j = c + \alpha T + \beta_{1,j}X_1 + \beta_{2,j}X_2 + \beta_{3,j}X_3 + \beta_{4,j}X_4 + \beta_{5,j}X_5 + \beta_{6,j}X_6 \qquad j = 1,7 \tag{4}$$

where Y_j is activity coefficient of species j, c is the intercept (i.e., constant term), α and $\beta_{i,j}$ are vectors representing the polynomial-fitted coefficients of species i for activity coefficient of species j, T and X_i are the matrixes representing temperature in K and molecule fractions of six hydrophilic species. The values of α and $\beta_{i,j}$ are shown in Table 4. The correlation coefficients (r) between the calculated activity coefficient by the parameterization and the UNIFAC range from 0.5 to 1.0. Similar parameterization has been tested for the hydrophobic SOA module, but the results give large deviations as compared with benchmark. This could be attributed to several reasons. First, the hydrophobic module treats a larger number of species and functional groups than the hydrophilic module (10 and 16 vs. 7 and 10, respectively), which increases the uncertainties in the parameterization using the linear multiple regression approach. Second, the R program used has a size limit for the input data, which limits the use of a very fine temperature interval for the hydrophobic module (i.e., every 1 K for the hydrophobic module vs. every 0.5 K for the hydrophilic module between 253–313 K). Since the hydrophilic module takes most CPU cost under the four conditions, the parameterization of activity coefficients in the hydrophilic module will provide a major speedup. While the parameterization of activity coefficients in hydrophobic module may provide additional speedup, it is subject to the aforementioned uncertainties and introduces larger errors in the activity coefficient calculation. The speed-up of hydrophobic module using other methods warrants further study in the future.

The above parameterization for the hydrophilic module reduces the CPU cost by 97%, 76%, 88%, and 62% (corresponding to speed up factors of 30.7, 4.1, 8.6, and 2.6) 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 63% (corresponding to speed up factors of 31.4, 4.5, 8.8, and 2.7) for rural HBG/LBG and urban HBG/LBG conditions, respectively. Compared with the baseline simulation without speed-up, the ranges of maximum instantaneous and average percentage deviations in species concentrations are −15.0 to 14.6% and −5.58 to 0.02% for major gas-phase species and total $PM_{2.5}$ (see Table 5) under the rural HBG condition. A more quantitative evaluation is conducted using typical statistical measures including correlation coefficient (corr.), mean bias (MB), root mean square error (RMSE), normalized mean bias (NMB) and error (NME), as shown in Table 5. The correlation coefficients between simulated concentrations of MADRID 2_Fast and MADRID 2 for 26 major species are 0.99–1.0. The MB and RMSE values are -7.1×10^{-3} to 3.3×10^{-2} ppb and 6.5×10^{-8} to 1.4×10^{-1} ppb, respectively, for gas-phase species and −7.3 and 8.3 μg m^{-3}, respectively, for total $PM_{2.5}$ concentrations. The NMB and NME values are −10.6 to 3.6% and 0.02–11%, respectively, for gas-phase species and −12.4% and 12.6% for total $PM_{2.5}$ concentrations. Given relatively large uncertainties in current 3-D air quality simulations (e.g., the NMBs for O_3 and $PM_{2.5}$ are typically in the range of 10–30% and 30–50%, see e.g., Seigneur, 2001; Zhang et al., 2004, 2006), MADRID 2_Fast is considered to gain computational benefits without a significant compromise in accuracy.

3.5 Sensitivity test of MADRID 2_Fast

MADRID 2_Fast that uses the optimal speed-up method (i.e., comb_all) is further tested under the aforementioned four and additional two non-typical conditions (i.e., the rural VHBG and the forest LN_VHBG conditions) with 7 temperatures (253, 263, 273, 283, 293, 303, and 313 K) and 5 RHs (10, 40, 60, 80, and 95%), with a total of 210 simulations (=7 temperatures $\times$ 5 RHs $\times$ 6 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 and rural VHBG conditions at a low RH of 10% or a high RH of 95% and a range of temperatures of 253–313 K and those at a low temperature of 253 K or a high temperature of 313 K and a range of RHs of 10–95%. For most species, the average deviations are within ±10%. Larger deviations occur for NO, nitrate radical (NO_3), BIOH, and BIOL at 253 K and RHs ≥80% or at RH=95% and temperatures ≤283 K or ≥303 K, due to their very small concentrations in the baseline simulation. For example, at T=263 K, RH=95%, the mixing ratios of NO decrease from 4.16×10^{-10} ppb at 11:30 p.m. in day 1 to 1.3×10^{-20} ppb at 00:00 a.m. in day 2 in the baseline simulation, those in the sensitivity simulation with MADRID 2_Fast decrease from 5.76×10^{-10} ppb at 11:30 a.m. to 2.42×10^{-20} ppb at 00:00 a.m., such changes from the small baseline values at each time step result in a large average percentage deviation of 34.0% (see Fig. 3b), although the absolute changes are negligible. In Fig. 3d and h, the large PM deviations of −15% occur at a high RH of 80% and a high temperature of 313 K. They are attributed to large deviations of both hydrophilic SOA and inorganic aerosols such as NH_4^+ (with deviations of −34.1%

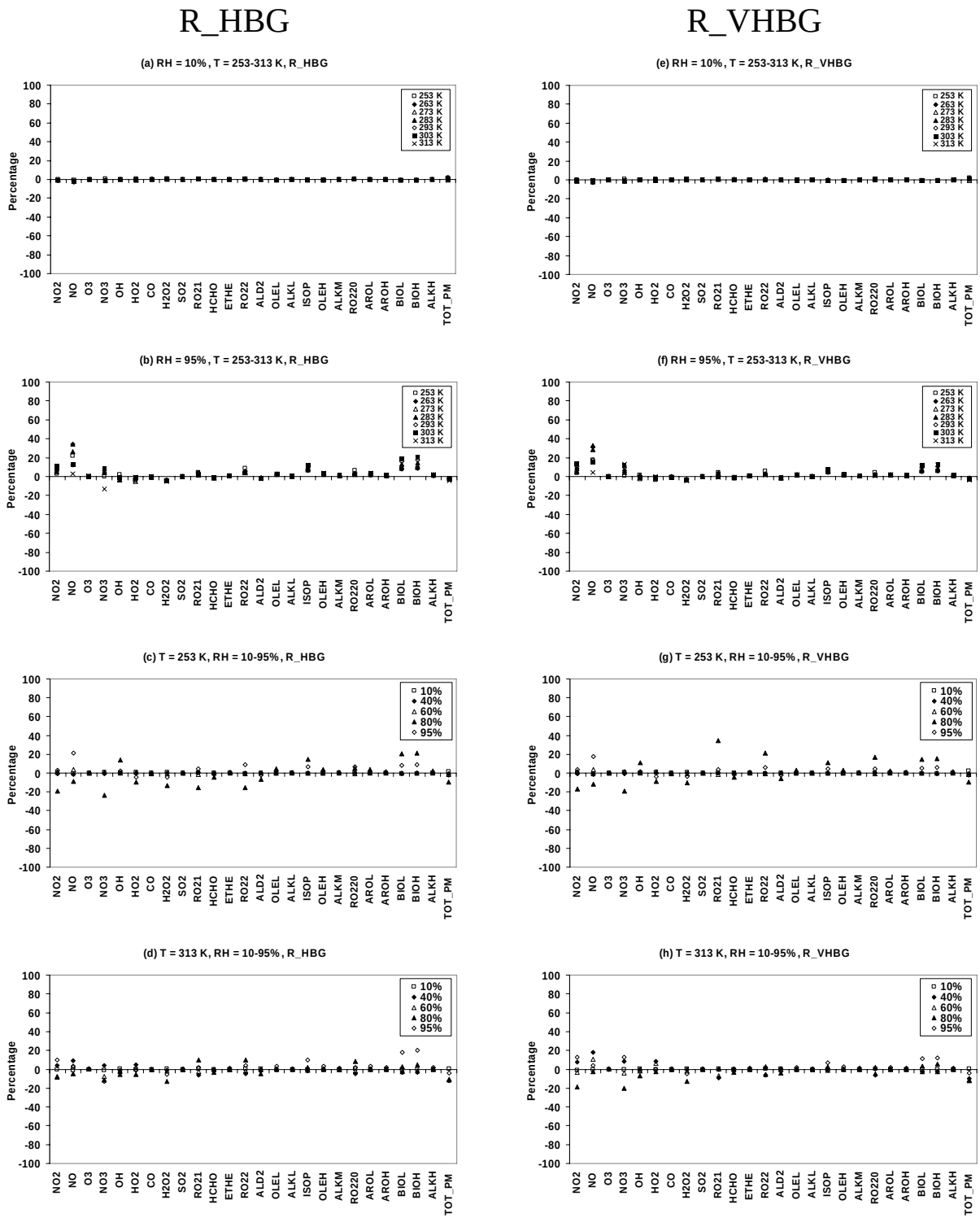

Fig. 3. Average percentage deviations from benchmark (i.e., the results from SMVGEAR with the most stringent error tolerances.) of major gas-phase species and total $PM_{2.5}$ during the 24-h simulation period using MADRID 2_Fast under the rural high biogenic VOC (R_HBG) (**(a)–(d)**) and the rural very high biogenic VOC (R_VHBG) condition (**(e)–(h)**) at RH=10%, T=253-313 K ((a) and (e)), RH=95%, T=253–313 K ((b) and (f)), T=253 K, RH=10–95% ((c) and (g)), and T=313 K, RH=10–95% ((d) and (h)) using the improved SOA module. The symbols for simulations at a constant temperature and a range of RHs are: square – 253 K; solid diamond – 263 K; triangle – 273 K; solid triangle – 283 K; diamond – 293 K; solid square – 303 K; and cross – 313 K. The symbols for simulations at a constant RH and a range of temperatures are: square – 10%; solid diamond – 40%; triangle – 60%; solid triangle – 80%; and diamond – 95%.

and −30.6%, respectively). Higher RHs give a higher effective Henry's law constant that subsequently favors the hydrophilic SOA formation. The extra water associated with hydrophilic OCs causes additional partitioning of inorganic species to the aqueous phase (Pun et al., 2002). The use of the O-I cutoff under such conditions can cause the deviations in both hydrophilic SOA and NH_4^+. The percentage deviations under the other four conditions are smaller than those shown in Fig. 3. In addition to the box model test simulation, one-day 3-D test simulations with CMAQ-MADRID 2_Fast against CMAQ-MADRID 2 is conducted and the results also show small deviations, with absolute differences for O_3 concentrations of −0.02 to 0.02 ppb (−0.07 to 0.1%), and those for $PM_{2.5}$ concentrations of −0.60 to 0.06 $\mu g\,m^{-3}$ (−6.1 to 1.5%). Figure 4 shows the computational speedup factors vs. average percentage deviations from benchmark for concentrations of O_3 and $PM_{2.5}$ using MADRID 2_Fast from the 210 test simulations. The average percentage deviations are within ±0.6% for O_3 (mostly within ±0.2%) and ±15% for $PM_{2.5}$ (mostly within ±10%). The computational speed up factor is greater than 2, 5, and 10 for more than 77%, 56%, and 37% of simulations, respectively. As expected, the highest speedup occurs under conditions with high biogenic VOCs and high RHs. The small speedup occurs under conditions with RH=10%. In such cases, the concentrations of solute species in aqueous-phase are relatively low and hydrophobic OCs dominate, the computational savings are dominated by speedup methods other than the parameterized activity coefficients for hydrophilic OCs. Figure 5 shows a fairly good correlation of simulated AROH, BIOH, total SOA, and total $PM_{2.5}$ between simulations with MADRID 2_Fast and MADRID 2 from a total of 210 simulations under the six test conditions during the 24-h simulation period, with correlation coefficients of 0.99 to 1.0 and >92% 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.

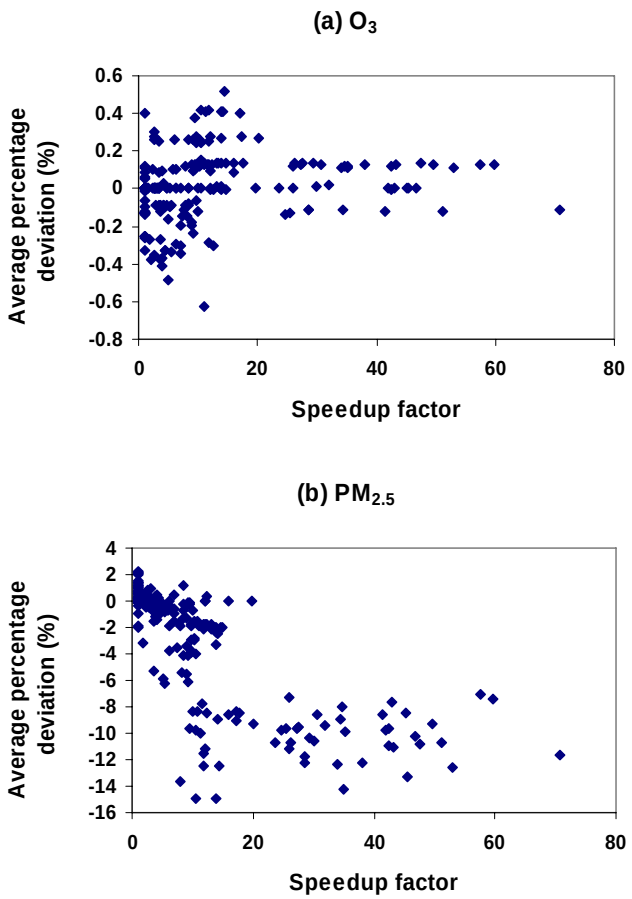

Fig. 4. Computational speedup factors vs. average percentage deviations from benchmark (i.e., the results from SMVGEAR with the most stringent error tolerances.) for concentrations of **(a)** O_3 and **(b)** $PM_{2.5}$ during the 24-h simulation period using MADRID 2_Fast from 210 test simulations.

4 Summary and future work

A box model, 0-D CMAQ-MADRID 2, is applied in this study to explore various methods in improving computational efficiency of the SOA module. Estimating their activity coefficients with the UNIFAC and solving partitioning equations for condensable VOCs are identified to be the most computationally-expensive processes in simulating SOA. Potential speed-up 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 for hydrophilic OCs using multiple linear regression analysis at various temperatures for the mixtures at various solute concentrations. 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 61.7 to 96.7%, or speedup by factors of 2.6–30.7) with the maximum instantaneous percentage deviations of −15.0 to 14.6%. The optimal speed-up method that combines all above methods can significantly reduce the CPU cost by 62.5 to 96.8% (speedup by factors of 2.7–31.4) with maximum instantaneous percentage deviations of −15.0 to 14.6% from the benchmark.

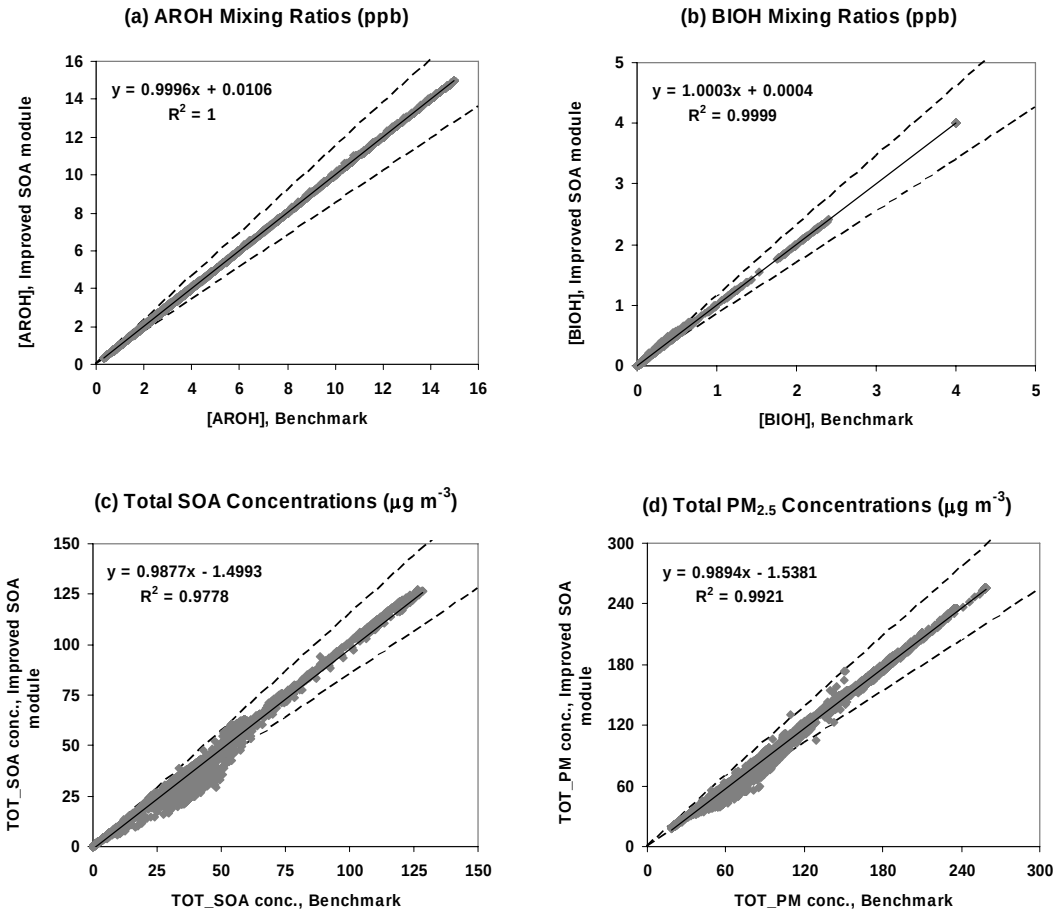

Fig. 5. The correlations of **(a)** AROH, **(b)** BIOH, **(c)** total SOA, and **(d)** total $PM_{2.5}$ between improved SOA module and benchmark results (i.e., the results from SMVGEAR with the most stringent error tolerances.) under the R_HBG, R_LBG, U_HBG, U_LBG, R_VHBG, and LN_VHBG conditions with various temperatures (i.e., 253, 263, 273, 283, 293, 303, 313 K) and RHs (i.e., 10, 40, 60, 80, 95%). The dashlines indicate deviation ranges of ±15%.

Under typical and non-typical atmospheric conditions with ranges of temperatures of 253–313 K and RHs of 10–95%, MADRID 2_Fast can improve computational efficiency by more than a factor of 2 for 77% of the 210 test simulations while reproducing the benchmark results with acceptable accuracy (with average percentage deviations within ±15%) for most species. The methods used to develop 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. 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 U.S. 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 U.S. EPA and the U.S. Department of Commerce's NOAA and under agreement number DW13921548. The authors thank G. L. Gipson, a former employee at the U.S. 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 to B. K. Pun, Atmospheric and Environmental Research (AER), Inc., for helpful discussions regarding MADRID 2. This work constitutes a contribution to the NOAA Air Quality Program.

Edited by: J. Seinfeld

References

Andersson-Sköld, Y. and Simpson, D.: Secondary organic aerosol formation in northern Europe: A model study, J. Geophys. Res., 106(D7), 7357–7374, 2001.

Ansari, A. S. and Pandis, S. N.: Water absorption by secondary organic aerosol and its effect on inorganic aerosol behavior, Environ. Sci. Technol., 34, 71–77, 2000.

Aumont, B., Madronich, S., Bey, I., and Tyndall, G. S.: Contribution of secondary VOC to the composition of aqueous atmospheric particles: A modeling approach, J. Atmos. Chem., 35, 59–75, 2000.

Barthelmie, R. J. and Pryor, S. C.: A model mechanism to describe oxidation of monoterpenes leading to secondary organic aerosol 1. α-pinene and β-pinene, J. Geophys. Res., 104(D19), 23 657–23 670, 1999.

Binkowski, F. S. and Roselle, S. J.: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description, J. Geophys. Res., 108(D6), 4183, doi:10.1029/2001JD001409, 2003.

Bowman, F. M. and Karamalegos, A. M.: Estimated effects of composition on secondary organic aerosol mass concentrations, Environ. Sci. Technol., 36, 2701–2707, 2002.

Bowman, F. M. and Melton, J. A.: Effect of activity coefficient models on predictions of secondary organic aerosol partitioning, J. Aerosol Sci., 35, 1415–1438, 2004.

Choi, M. Y. and Chan, C. K.: The effects of organic species on the hygroscopic behaviors of inorganic aerosols, Environ. Sci. Technol., 36, 2422–2428, 2002.

Clegg, S. L., Seinfeld, J. H., and Brimblecombe, P.: Thermodynamic modeling of aqueous aerosols containing electrolytes and dissolved organic compounds, J. Aerosol Sci., 32, 713–738, 2001.

Cruz, C. N. and Pandis, S. N.: Deliquescence and hygroscopic growth of mixed inorganic-organic atmospheric aerosol, Environ. Sci. Technol., 34, 4313–4319, 2000.

Davidson, C. I., Phalen, R. F., and Solomon, P. A.: Airborne particulate matter and human health: A review, Aerosol Sci. Tech., 39, 737–749, 2005.

Finlayson-Pitts, B. J. and Pitts Jr., J. N.: Atmospheric Chemistry: Fundamentals and Experimental Techniques, John Wiley & Sons, Inc., USA, 1986.

Fredenslund, A., Jones, R. L., and Prausnitz, J. M.: Groupcontribution estimation of activity coefficients in nonideal liquid mixtures, Aiche J., 21, 1086–1099, 1975.

Fredenslund, A., Gmehling, J., and Rasmussen, P.: Vapor–Liquid equilibria using UNIFAC: a group-contribution method, Elsevier Scientific Publishing, New York, 1977.

Griffin, R. J., Cocker III, D. R., Flagan, R. C., and Seinfeld, J. H.: Organic aerosol formation from the oxidation of biogenic hydrocarbons, J. Geophys. Res., 104(D3), 3555–3567, 1999.

Griffin, R. J., Dabdub, D., Kleeman, M. J., Fraser, M. P., Cass, G. R., and Seinfeld, J. H.: Secondary organic aerosol 3. Urban/regional scale model of size- and composition-resolved aerosols, J. Geophys. Res., 107(D17), 4334, doi:10.1029/2001JD000544, 2002a.

Griffin, R. J., Dabdub, D., and Seinfeld, J. H.: Secondary organic aerosol 1. Atmospheric chemical mechanism for production of molecular constituents, J. Geophys. Res., 107(D17), 4332, doi:10.1029/2001JD000541, 2002b.

Griffin, R. J., Nguyen, K., Dabdub, D., and Seinfeld, J. H.: A coupled hydrophobic-hydrophilic model for predicting secondary organic aerosol formation, J. Atmos. Chem., 44, 171–190, 2003.

Hertel, O., Berkowicz, R., Christensen, J., and Hov, Ø.: Test of two numerical schemes for use in atmospheric transport-chemistry models, Atmos. Environ., 27A, 2591–2611, 1993.

Hesstvedt, E., Hov, Ö., and Isaksen, I. S. A.: Quasi-steady-state approximations in air pollution modeling: a comparison of two numerical schemes for oxidant prediction, Int. J. Chem. Kinet., 10, 971–994, 1978.

Hine, J. and Mookerjee, P. K.: The intrinsic hydrophilic character of organic compounds: correlations in terms of structural contributions, J. Org. Chem., 40, 292–298, 1975.

Hu, X.-M. and Zhang, Y.: Implementation and testing of a new aerosol module in WRF/Chem, Proceedings of the 86th AMS Annual Meeting/the 8th Conference on atmospheric chemistry, Atlanta, 2006.

Huang, H.-C. and Chang, J. S.: On the performance of numerical solvers for a chemistry submodel in three-dimensional air quality models: 1. Box model simulations, J. Geophys. Res., 106(D17), 20 175–20 188, 2001.

IPCC: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp., 2007.

Jacobson, M. Z. and Turco, R. P.: SMVGEAR: A sparse-matrix, vectorized Gear code for atmospheric models, Atmos. Environ., 28, 273–284, 1994.

Lyman, W. J., Reehl, W. F., and Rossnblatt, D. H.: Handbook of Chemical Property Estimation Methods, American Chemical Society, Washington, DC, 11–22, 1990.

National Toxicology Program (NTP), Public Health Service, U.S. Department of Health and Human Services: the 11th report on carcinogens, http://ntp.niehs.nih.gov/index.cfm?objectid=32BA9724-F1F6-975E-7FCE50709CB4C932, 2005.

Nenes, A., Pandis, S. N., and Pilinis, C.: Continued development and testing of a new thermodynamic aerosol module for urban and regional air quality models, Atmos. Environ., 33, 1553–1560, 1999.

Odum, J. R., Jungkamp, T. P. W., Griffin, R. J., Forstner, H. J. L., Flagan, R. C., and Seinfeld, J. H.: Aromatics, reformulated gasoline, and atmospheric organic aerosol formation, Environ. Sci. Technol., 31, 1890–1897, 1997.

Pankow, J. F.: An absorption model of gas-particle partitioning of organic compounds in the atmosphere, Atmos. Environ., 28, 185–188, 1994a.

Pankow, J. F.: An absorption model of the gas/aerosol partitioning involved in the formation of secondary organic aerosol, Atmos. Environ., 28, 189–193, 1994b.

Prausnitz, J. M., Lichtenthaler, R. N., and Gomes de Azevedo, E.: Molecular thermodynamics of fluid-phase equilibria, Upper Saddle River, NJ, Prentice-Hall, 1999.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P.: Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, New York, 1997.

Pun, B. K., Griffin, R. J., Seigneur, C., and Seinfeld, J. H.: Secondary organic aerosol, 2. Thermodynamic model for gas/particle partitioning of molecular constituents, J. Geophys. Res., 107(D17), 4333, doi:10.1029/2001JD000542, 2002.

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.

Pun, B. K., Wu, S.-Y., Seigneur, C., Seinfeld, J. H., Griffin, R. J., and Pandis, S. N.: Uncertainties in modeling secondary organic aerosols: Three-dimensional modeling studies in Nashville/Western Tennessee, Environ. Sci. Technol., 37, 3647–3661, 2003.

Sandu, A., Potra, F. A., Carmichael, G. R., and Damian, V.: Efficient implementation of fully implicit methods for atmospheric chemical kinetics, J. Comput. Phys., 129, 101–110, 1996.

Sandu, A., Verwer, J. G., Loon, M. V., Carmichael, G. R., Potra, F. A., Dabdub, D., and Seinfeld, J. H.: Benchmarking stiff ODE solvers for atmospheric chemistry problems-I. Implicit vs explicit, Atmos. Environ., 31, 3151–3166, 1997a.

Sandu, A., Verwer, J. G., Blom, J. G., Spee, E. J., Carmichael, G. R., and Potra, F. A.: Benchmarking stiff ODE solvers for atmospheric chemistry problems-II. Rosenbrock solvers, Atmos. Environ., 31, 3459–3472, 1997b.

Saxena, P., Hildemann, L. M., McMurry, P. H., and Seinfeld, J. H.: Organics alter hygroscopic behavior of atmospheric particles, J. Geophys. Res., 100(D9), 18 755–18 770, 1995.

Saxena, P. and Hildemann, L.: Water-soluble organics in atmospheric particles: A critical review of the literature and application of thermodynamics to identify candidate compounds, J. Atmos. Chem., 24, 57–109, 1996.

Saxena, P. and Hildemann, L. M.: Water absorption by organics: survey of laboratory evidence and evaluation of UNIFAC for estimating water activity, Environ. Sci. Technol., 31, 3318–3324, 1997.

Schell, B., Ackermann, I. J., Hass, H., Binkowski, F. S., and Ebel, A.: Modeling the formation of secondary organic aerosol within a comprehensive air quality model system, J. Geophys. Res., 106(D22), 28 275–28 294, 2001.

Seigneur, C.: Current status of air quality models for particulate matter, J. Air Waste Ma., 51, 1508–1521, 2001.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, John Wiley & Sons, Inc., New York, 1998.

Stokes, R. H. and Robinson, R. A.: Interactions in aqueous nonelectrolyte solutions. I. Solute-solvent equilibria, J. Phys. Chem., 70, 2126–2130, 1966.

Strader, R., Lurmann, F., and Pandis, S. N.: Evaluation of secondary organic aerosol formation in winter, Atmos. Environ., 33, 4849–4863, 1999.

Suzuki, T., Ohtaguchi, K., and Koide, K.: Application of principal components analysis to calculate Henry's constant from molecular structure, Comput. Chem., 16, 41–52, 1992.

Tulet, P., Grini, A., Griffin, R. J., and Petitcol, S.: ORILAM-SOA: A computationally efficient model for predicting secondary organic aerosols in three-dimensional atmospheric models, J. Geophys. Res., 111, D23208, doi:10.1029/2006JD007152, 2006.

Turpin, B. J., Saxena, P., and Andrews, E.: Measuring and simulating particulate organics in the atmosphere: Problems and prospects, Atmos. Environ., 34, 2983–3013, 2000.

U.S. Environmental Protection Agency: CHEM_SOLVER_NOTES. CMAQ release notes, Research Triangle Park, NC, 2004.

Watson, J. G.: Visibility: science and regulation, J. Air Waste Ma., 52, 628–713, 2002.

Zhang, Q, Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T., Jayne, J. T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y., Zhang, Y. M., and Worsnop, D. R.: Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitude, Geophys. Res. Lett., 34, L13801, doi:10.1029/2007GL029979, 2007.

Zhang, Y., Bischof, C. H., Easter, R. C., and Wu, P.-T.: Sensitivity analysis of a mixed-phase chemical mechanism using automatic differentiation, J. Geophys. Res., 103(D15), 18 593-18 980, 1998.

Zhang, Y., Hu, X.-M., Howell, G. W., Sills, E., Fast, J. D., Gustafson Jr., W. I., Zaveri, R. A., Grell, G. A., Peckham, S. E., and McKeen, S. A.: Modeling atmospheric aerosols in WRF/CHEM, Proceeding of the 2005 Joint WRF/MM5 User's Workshop, Boulder, 27–30 June, 2005.

Zhang, Y., Liu, P., Queen, A., Misenis, C., Pun, B., Seigneur, C., and Wu, S.-Y.: A comprehensive performance evaluation of MM5-CMAQ for the summer 1999 Southern Oxidants Study Episode, Part-II. Gas and aerosol predictions, Atmos. Environ., 40, 4839–4855, 2006.

Zhang, Y., Pun, B. K., Vijayaraghavan, K., Wu, S.-Y., Seigneur, C., Pandis, S. N., Jacobson, M., Nenes, A., and Seinfeld, J. H.: Development and application of the Model of Aerosol Dynamics, Reaction, Ionization, and Dissolution (MADRID), J. Geophys. Res., 109, D01202, doi:10.1029/2003JD003501, 2004.