Available online at www.sciencedirect.com
SCIENCE DIRECT®
Atmospheric Environment 39 (2005) 6957–6975
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Fully coupled "online" chemistry within the WRF model

Georg A. Grell[a,*], Steven E. Peckham[a], Rainer Schmitz[c], Stuart A. McKeen[b], Gregory Frost[b], William C. Skamarock[d], Brian Eder[e]

[a]*Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado/NOAA Research—Forecast Systems Laboratory, Boulder, CO, USA*
[b]*Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado/NOAA Research—Aeronomy Laboratory, Boulder, CO, USA*
[c]*Department of Geophysics, University of Chile, Santiago, Chile Institute for Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Forschungszentrum Karlsruhe, Garmisch-Partenkirchen, Germany*
[d]*National Center for Atmospheric Research, Boulder, CO, USA*
[e]*NOAA Air Resources Laboratory, Boulder, CO, USA*

Received 11 November 2004; accepted 24 March 2005

Abstract

A fully coupled "online" Weather Research and Forecasting/Chemistry (WRF/Chem) model has been developed. The air quality component of the model is fully consistent with the meteorological component; both components use the same transport scheme (mass and scalar preserving), the same grid (horizontal and vertical components), and the same physics schemes for subgrid-scale transport. The components also use the same timestep, hence no temporal interpolation is needed. The chemistry package consists of dry deposition ("flux-resistance" method), biogenic emission as in [Simpson et al., 1995. Journal of Geophysical Research 100D, 22875–22890; Guenther et al., 1994. Atmospheric Environment 28, 1197–1210], the chemical mechanism from RADM2, a complex photolysis scheme (Madronich scheme coupled with hydrometeors), and a state of the art aerosol module (MADE/SORGAM aerosol parameterization).

The WRF/Chem model is statistically evaluated and compared to MM5/Chem and to detailed photochemical data collected during the summer 2002 NEAQS field study. It is shown that the WRF/Chem model is statistically better skilled in forecasting O_3 than MM5/Chem, with no appreciable differences between models in terms of bias with the observations. Furthermore, the WRF/Chem model consistently exhibits better skill at forecasting the O_3 precursors CO and NO_y at all of the surface sites. However, the WRF/Chem model biases of these precursors and of other gas-phase species are persistently higher than for MM5/Chem, and are most often biased high compared to observations. Finally, we show that the impact of other basic model assumptions on these same statistics can be much larger than the differences caused by model differences. An example showing the sensitivity of various statistical measures with respect to the treatment of biogenic volatile organic compounds emissions illustrates this impact.

Keywords: Urban and regional pollution; Urban and regional air quality modeling; Air quality forecasting; Aerosols and particles

1. Introduction

The simulation and prediction of air quality is a complicated problem, involving both meteorological

*Corresponding author. Tel.: +1 303 497 6924; fax: +1 303 497 7262.
E-mail address: Georg.a.Grell@noaa.gov (G.A. Grell).

1352-2310/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2005.04.027

factors (such as wind speed and direction, turbulence, radiation, clouds, and precipitation) and chemical processes (such as deposition, and transformations). In the real atmosphere, the chemical and physical processes are coupled. The chemistry can affect the meteorology, for example, through its effect on the radiation budget, as well as the interaction of aerosols with cloud condensation nuclei (CCN). Likewise, clouds and precipitation strongly influence chemical transformation and removal processes, and localized changes in the wind or turbulence fields continuously affect the chemical transport.

Until recently, the chemical processes in air quality modeling systems were usually treated independently of the meteorological model, (as in the Community Model for Air Quality (CMAQ, (Byun and Ching, 1999); i.e., "*offline*", except that the transport was driven by output from a meteorological model, typically available once or twice per hour. This approach is computationally very attractive since retrospective offline chemical transport simulations only require a single meteorological dataset to produce many chemical transport simulations to examine a scientific research question. However, this separation of meteorology and chemistry can cause a loss of important information about atmospheric processes that quite often have a time scale of much less than the output time of the meteorological model, e.g., wind speed and direction, rainfall, and cloud formation. This may be especially important in air quality prediction systems, in which horizontal grid sizes on the order of 1 km may be required.

Over the past few years, several research institutes have collaborated in the development of a new state-of-the-art Weather Research and Forecasting (WRF) model (http://www.mmm.ucar.edu/wrf/users/document.html). WRF is non-hydrostatic, with several dynamic cores as well as many different choices for physical parameterizations to represent processes that cannot be resolved by the model. This allows the model to be applicable on many different scales. The dynamic cores include a fully mass- and scalar-conserving flux form mass coordinate version. Similar approaches have recently been implemented in the Operational Multiscale Environment Model with Grid Adaptivity (OMEGA, Bacon et al., 2000) as well as the Japanese numerical weather prediction model (Satoh, 2002). This makes the WRF model ideally suited to be the cornerstone for a next generation air quality prediction system.

"The Workshop on Modeling Chemistry in Cloud and Mesoscale Models," a first step toward the implementation of chemistry into WRF, was held at the National Center for Atmospheric Research (NCAR) on 6–8 March 2000. The goal of this workshop was to produce a community assessment of approaches and methodologies used for chemistry modeling in cloud and mesoscale models. Since then, various chemical modules have been implemented into the WRF framework, creating an "online" WRF/Chem model. Transport of species is done using the same vertical and horizontal coordinates (no horizontal or vertical interpolation), and the same physics parameterization with no interpolation in time. The WRF/Chem model presented here maintains the physical and chemical formulations of MM5/Chem (Grell et al., 2000), in particular those related to ozone pollution. The MM5/chem model has previously been rigorously evaluated (Eder et al., 2005; McKeen et al., 2003). Therefore, comparisons of WRF/Chem with observations as well as with MM5/Chem for several key pollutants provide the basis of our model evaluation. The chemical aspects of the model are described in Section 2. The setup for retrospective runs that were used for the model evaluation is explained in Section 3. The results and summary are provided in Sections 4 and 5, respectively.

2. Model description

In general, most air quality modeling systems consider a variety of coupled physical and chemical processes such as transport, deposition, emission, chemical transformation, aerosol interactions, photolysis, and radiation. Details on the modules that describe these processes within WRF/Chem are given below. For details describing the conservative split-explicit time integration method that is used in the mass coordinate version of the WRF model, see http://www.mmm.ucar.edu/individual/skamarock/wrf_equations_eulerian.pdf. The time splitting method is described in Wicker and Skamarock (2002), and an overview of the physics is available at http://www.mmm.ucar.edu/wrf/users/wrf-doc-physics.pdf. Here we only discuss the aspects of the model that directly relate to the chemical part.

2.1. Transport

All transport of chemical species is done online. Although WRF has several choices for dynamic cores, in this model we use the mass coordinate version of the model, called Advanced Research WRF (ARW). The prognostic equations integrated in the ARW model are cast in conservative (flux) form for conserved variables; non-conserved variables such as pressure and temperature are diagnosed from the prognostic conserved variables. In the conserved variable approach, the ARW model integrates a mass conservation equation and a scalar conservation equation of the form

$$\mu_t + \nabla \cdot (\mathbf{V}\mu) = 0,$$
$$(\mu\phi)_t + \nabla \cdot (\mathbf{V}\mu\phi) = 0.$$

In these equations μ is the column mass of dry air, V is the velocity (u, v, ∇), and ϕ is a scalar mixing ratio. These equations are discretized in a finite volume formulation, and as a result the model exactly (to machine roundoff) conserves mass and scalar mass. The discrete model transport is also consistent (the discrete scalar conservation equation collapses to the mass conservation equation when $\phi = 1$) and preserves tracer correlations (c.f. Lin and Rood (1996). The ARW model uses a spatially 5th-order evaluation of the horizontal flux divergence (advection) in the scalar conservation equation and a 3rd-order evaluation of the vertical flux divergence coupled with the 3rd-order Runge–Kutta time integration scheme. The time integration scheme and the advection scheme is described in Wicker and Skamarock (2002). Turbulent transport in the boundary layer is performed using a level 2.5 Mellor–Yamada closure parameterization (Mellor and Yamada, 1982).

For the chemical mechanism used in this version of the model, 39 chemical species are fully prognostic. For the aerosol module (see description below), another 34 variables are added, including the total number of aerosol particles within each mode, as well as all primary and secondary species (organic and inorganic) for both Aitken and accumulation mode, and three species for the coarse mode (anthropogenic, marine, and soil-derived aerosols).

2.2. Dry deposition

The flux of trace gases and particles from the atmosphere to the surface is calculated by multiplying concentrations in the lowest model layer by the spatially and temporally varying deposition velocity, which is proportional to the sum of three characteristic resistances (aerodynamic resistance, sublayer resistance, surface resistance). The surface resistance parameterization developed by Wesley (1989) is used. In this parameterization, the surface resistance is derived from the resistances of the surfaces of the soil and the plants. The properties of the plants are determined using land-use data and the season. The surface resistance also depends on the diffusion coefficient, the reactivity, and water solubility of the reactive trace gas.

The dry deposition of sulfate is described differently. In the case of simulations without calculating aerosols explicitly, sulfate is assumed to be present in the form of aerosol particles, and its deposition is described according to Erisman et al. (1994).

When employing the aerosol parameterization, the deposition velocity, $\widehat{v}_{dk}$, for the kth moment of a polydisperse aerosol is given by

$$\widehat{v}_{dk} = (r_a + \widehat{r}_{dk} + r_a \widehat{r}_{dk} \widehat{v}_{Gk})^{-1} + \widehat{v}_{Gk}, \quad (1)$$

where r_a is the surface resistance, $\widehat{v}_{Gk}$ is the polydisperse settling velocity, and r_{dk} is the Brownian diffusivity (Slinn and Slinn, 1980; Pleim et al., 1984).

2.3. Gas-phase chemistry

This atmospheric chemical mechanism was originally developed by Stockwell et al. (1990) for the Regional Acid Deposition Model, version 2 (RADM2 Chang et al., 1989). The RADM2 mechanism is a compromise between chemical detail, accurate chemical predictions, and available computer resources. It is widely used in atmospheric models to predict concentrations of oxidants and other air pollutants.

The inorganic species included in the RADM2 mechanism are 14 stable species, 4 reactive intermediates, and 3 abundant stable species (oxygen, nitrogen and water). Atmospheric organic chemistry is represented by 26 stable species and 16 peroxy radicals. The RADM2 mechanism represents organic chemistry through a reactivity aggregated molecular approach (Middleton et al., 1990). Similar organic compounds are grouped together in a limited number of model groups through the use of reactivity weighting. The aggregation factors for the most emitted volatile organic compounds (VOCs) are given in Middleton et al. (1990).

A quasi steady-state approximation method with 22 diagnosed and 38 predicted species is used to predict chemical production and loss tendency terms in the numerical solution. The time integrations for 38 predicted species are then solved using these tendencies and a Backward Euler approximation.

Initial and boundary conditions for the prognostic gas-phase variables are based on those of McKeen et al. (2002). These consist of laterally invariant vertical profiles representing clean, oceanic, midlatitude conditions from measurements collected onboard previous NASA-sponsored aircraft missions. No adjustments for boundary conditions based on potential vorticity are applied to the stratospheric or upper-tropospheric model grids.

2.4. Emissions

Anthropogenic emissions are treated similar to that of McKeen et al. (2002) with updates to the April 2002 release of the EPA NET-96 inventory version 3.12 (US EPA, 1998). Hourly temporal allocation, VOC speciation, and spatial partitioning within a specified county are based on the older, yet detailed information within the National Acid Precipitation Assessment Program (NAPAP) emissions database (US EPA, 1998). Canadian emissions are also taken directly from the NAPAP modeler's inventory. In order to adhere to the RADM2 mechanism, reactivity weighting of the various NAPAP-lumped VOC species is used to derive the emissions of

the RADM2-lumped species according to Stockwell et al. (1990, 1997). Sources of VOC, CO, NO_x, and aerosol from forest fires and prescribed burning are omitted from the emissions since these are specific to 1996 within the EPA inventory.

WRF/Chem uses a biogenic emission module based on the description of Guenther et al. (1993, 1994), Simpson et al. (1995), and Schoenemeyer et al. (1997). The module treats the emissions of isoprene, monoterpenes, other biogenic VOC (OVOC), and nitrogen emissions by the soil. For the use in the RADM2 photochemistry module, the emissions of monoterpenes and OVOC are disaggregated into the RADM2 species classes.

The isoprene emissions by the forests depend on both temperature and photosynthetic active radiation. Guenther et al. (1993) have developed a parameterization formula for these emissions, in which the isoprene emission rate is proportional to the isoprene emission rate at a standard temperature and a standard flux of photosynthetic active radiation. A radiation flux correction term and a temperature correction term for forest isoprene emissions are applied. The isoprene emissions of agricultural and grassland areas are considered to be functions of the temperature only (Hahn et al., 1994).

The emissions of monoterpenes, OVOC, and nitrogen are also treated as functions of the temperature only. Little is known about the emission of OVOC; therefore, the same temperature correction is applied for OVOC as for monoterpenes according to Simpson et al. (1995).

The emissions at the standard temperature and the standard photosynthetically active radiation (PAR) flux are given in Grell et al. (2000). They are taken from Guenther et al. (1994) for deciduous, coniferous and mixed forest and from Schoenemeyer et al. (1997) for agricultural and grassland. For the use with RADM2, all nitrogen emissions are treated as NO. This is a maximum estimate, because the emission of N_2O is neglected.

It must be noted that from the land-use categories used in WRF, the nature of biogenic emissions can be estimated only roughly. These categories are based on the USGS 24-class land use/land cover system classification, which in its original WRF implementation does not include any tree species information or fractional coverages.

2.5. Parameterization of aerosols

The aerosol module is based on the Modal Aerosol Dynamics Model for Europe (MADE, (Ackermann et al., 1998) which itself is a modification of the Regional Particulate Model (Binkowski and Shankar, 1995). Secondary organic aerosols (SOA) have been incorporated into MADE by Schell et al. (2001), by means of the Secondary Organic Aerosol Model (SORGAM). Since the different components of the module are well documented in the above-cited references, only a brief summary of the most important features shall be given here.

2.5.1. Size distributions

The size distribution of the submicrometer aerosol is represented by two overlapping intervals, called modes, assuming a log-normal distribution within each mode:

$$n(\ln d_p) = \frac{N}{\sqrt{2\pi}\ln \sigma_g}\exp\left[-\frac{1}{2}\frac{\left(\ln d_p - \ln d_{p_g}\right)^2}{\ln^2 \sigma_g}\right], \quad (2)$$

where N is the number concentration (m^{-3}), d_p the particle diameter, d_{pg} the median diameter, and σ_g the standard deviation of the distribution. The kth moment of the distribution is defined as

$$M_k = \int_{-\infty}^{\infty} d_p^k n(\ln d_p)\, d(\ln d_p) \quad (3)$$

with the solution

$$M_k = N d_{p_g}^k \exp\left[\frac{k^2}{2}\ln^2 \sigma_g\right]. \quad (4)$$

M_0 is the total number of aerosol particles within the mode suspended in a unit volume of air, M_2 is proportional to the total particulate surface area within the mode suspended in a unit volume of air, and M_3 is proportional to the total particulate volume within the mode suspended in a unit volume of air.

2.5.2. Nucleation, condensation, and coagulation

The most important process for the formation of secondary aerosol particles is the homogeneous nucleation in the sulfuric acid–water system. It is calculated by the method given by Kulmala et al. (1998).

Aerosol growth by condensation occurs in two steps: the production of condensable material (vapor) by the reaction of chemical precursors, and the condensation and evaporation of ambient volatile species on aerosols. In MADE the Kelvin effect is neglected, allowing the calculation of the time rate of change of a moment M_k for the continuum and free-molecular regime. The mathematical expressions of the rates and their derivation are given in Binkowski and Shankar (1995).

During the process of coagulation, the distributions remain log-normal. Furthermore, only the effects caused by Brownian motion are considered for the treatment of coagulation. The mathematical formulation for the coagulation process can be found in Whitby et al. (1991) and Binkowski and Shankar (1995).

The change in moments due to coagulation is modified from that described by Whitby et al. (1991). Whereas Whitby et al. (1991) suggest that the collisions of particles within a mode result in the formation of a particle within that mode, MADE allows a particle

resulting from two particles colliding within the Aitken mode to be assigned to the accumulation mode. For this, MADE calculates the diameter, d_{eq}, at which the two modes have equal number concentrations. Colliding particles in the Aitken mode, where at least one exceeds this diameter, are then assigned to the accumulation mode.

2.5.3. Aerosol chemistry

The inorganic chemistry system, based on MARS (Saxena et al., 1986), and its modifications by Binkowski and Shankar (1995), calculates the chemical composition of a sulphate-nitrate-ammonium-water aerosol according to equilibrium thermodynamics. Two regimes are considered depending upon the molar ratio of ammonium and sulphate. For values less than 2, the code solves a cubic polynomial for hydrogen ion molality, and if enough ammonium and liquid water are present, it calculates the dissolved nitrate. For modal ionic strengths greater than 50, nitrate is assumed not to be present. For molar ratios of 2 or greater, all sulphate is assumed to be ammonium sulphate and a calculation is made for the presence of water. The Bromley method is used for the calculation of the activity coefficients.

The organic aerosol chemistry is based on SORGAM (Schell et al., 2001), which assumes that SOA compounds interact and form a quasi-ideal solution. The gas/particle partitioning of SOA compounds is parameterized according to Odum et al. (1996). Due to the lack of information, all activity coefficients are assumed to be unity. SORGAM treats anthropogenic and biogenic precursors separately, and may be used with a chemical mechanism such as RACM (Stockwell et al., 1997) that provides the biogenic precursors. Since we currently use the RADM2 mechanism (Stockwell et al., 1990) in WRF/Chem, the biogenic precursors and their resulting particle concentrations are set to zero.

Similar to the emissions of gas-phase species, aerosol $PM_{2.5}$ primary emissions are taken from the EPA NET-96 inventory version 3.12 (US EPA, 1998). The mobile and non-mobile area sources within this inventory include primary emissions of $PM_{2.5}$ sulfate, ammonium, nitrate, organic carbon, elemental carbon, and an unspeciated category hereafter referred to as "fine". Within the NET-96 inventory all point source $PM_{2.5}$ emissions are specified within the unspeciated "fine" category. Mass concentrations of the six $PM_{2.5}$ subspecies are individually transported, deposited and accounted for within the MADE/SORGAM formalism. Total dry $PM_{2.5}$ is determined by summing the six individual sub-species.

2.6. Photolysis frequencies

[31] Photolysis frequencies for the 21 photochemical reactions of the gas-phase chemistry model are calculated at each grid point according to Madronich (1987). The photolysis frequency of the gas i, J_i, is given by the integral of the product of the actinic flux $I_A(\lambda)$, the absorption cross sections $\sigma(\lambda)$, and the quantum yields $\Phi(\lambda)$ over the wavelength λ:

$$J_i = \int^{\lambda} I_A(\tau, \lambda)\sigma_i(\lambda)\Phi_i(\lambda)\,d\lambda. \quad (5)$$

For the calculation of the actinic flux, a radiative transfer model, which is based on the delta-Eddington technique (Joseph et al., 1976), is used. This radiative transfer model accounts for absorption by O_2 and O_3, Rayleigh scattering, and scattering and absorption by aerosol particles and clouds as described by Chang et al. (1989). The absorption cross-sections and the quantum yields for the calculation of J_{gas} are given by Stockwell et al. (1990). The integral in the above equation is solved for 130 wavelengths between 186 and 730 nm.

The profiles of the actinic flux are computed at each grid point of the model domain. For the determination of the absorption and scattering cross-sections needed by the radiative transfer model, predicted values of temperature, ozone, and cloud liquid water content are used below the upper boundary of WRF. Above the upper boundary of WRF, fixed typical temperature and ozone profiles are used to determine the absorption and scattering cross-sections. These ozone profiles are scaled with total ozone mapping spectrometer (TOMS) satellite observational data for the area and date under consideration.

The radiative transfer model permits the proper treatment of several cloud layers each with height-dependent liquid water contents. The extinction coefficient of cloud water β_c is parameterized as a function of the cloud water computed by the three-dimensional model based on a parameterization given by Slingo (1989). For the present study, the effective radius of the cloud droplets follows Jones et al. (1994). For aerosol particles, a constant extinction profile with an optical depth of 0.2 is applied.

An online computation of the photolysis frequencies is preferred here, since it has advantages over offline techniques and is more versatile. One advantage is that the absorption cross-sections of ozone are temperature dependent. Furthermore, this treatment can be used to account for the humidity dependence of the extinction by aerosol particles. As shown by Ruggaber et al. (1994), aerosol particles have a strong effect on the photolysis frequency of NO_2. Another possible option for the model is the parameterization of cloud droplets as a function of the sulfate content according to Jones et al. (1994).

The photolysis model may be applied at any timestep. However, for numerical efficiency, the photolysis routine is called with time intervals of 30 min.

3. Test-bed setup

Composition and meteorological data collected from several measurement platforms for a 47-day period, 5 July–20 August 2002, are used to evaluate the WRF/Chem model forecasts relative to MM5/Chem forecasts. This period was previously used to evaluate the real-time performance of MM5/Chem with two other air quality forecast models (McKeen et al., 2003). Most of the evaluation is based on observations of O_3 and several other pollutants in the Northeast US that were collected routinely during the NEAQS-2002 (New England Air Quality Study–2002) field campaign. These include observations at four Atmospheric Investigation, Regional Modeling, Analysis and Prediction (AIRMAP) sites operated by the University of New Hampshire, the Harvard Forest field site operated by Harvard University, and ship-based measurements taken aboard the NOAA research vessel *Ronald H. Brown* operating in the Gulf of Maine. Additionally, the EPA's ambient air quality AIRNOW O_3 network is also used in this evaluation to assess the models' ability to forecast surface O_3 over the entire eastern third of the US.

To be able to compare the previous MM5/Chem evaluation, the WRF/Chem model configuration was chosen to be as similar as possible to the original MM5/Chem setup. A series of 36-h simulations, conducted every 12 h (0000 and 1200 UTC) are performed on a roughly 3600-km × 3000-km numerical grid having 27-km horizontal resolution and centered at 86°W longitude and 34.5°N latitude. The domain of both models extends vertically to ~15.5 km with a vertical mesh interval smoothly increasing from 15 m near the surface to approximately 1.5 km at the domain top. However, vertical coordinates are not coincident between the two models due to the inherent difference between the mass coordinate system used in WRF/Chem compared to the reference-state vertical coordinate system used in the non-hydrostatic MM5/Chem model. The number of vertical levels for the WRF/Chem model (34 in this case) was increased relative to the 29 vertical levels within the MM5/Chem model to allow for a smoother transition between 800 and 2500 m, and for increased resolution in the upper troposphere.

Further information about the configuration of the WRF/Chem model is provided in Table 1. Both models obtained meteorological initial conditions from the Rapid Update Cycle (RUC) model analysis fields generated at the Forecast Systems Laboratory (FSL), and lateral boundary conditions from the NCEP Eta-model forecast. Aside from the advection scheme and the inclusion of aerosols, three model physics options within MM5/Chem are different than WRF/Chem in Table 1: the land surface models, the convective parameterizations, and the microphysics parameterizations. These options were part of the MM5/Chem formulation but were not available within the WRF model framework at the time of this work.

Both modeling systems use approximately the same amount of computing time, and were run on 36 Linux PC based processors of a distributed memory parallel computer. Although WRF/Chem has significantly more calculations, it also uses a larger time step for the dynamics. Computationally most expensive in both models are the chemical interactions, followed closely by the transport of the species. The meteorology only accounts for a minor amount of the total computational time.

4. Results

4.1. Comparisons with AIRMAP and Harvard Forest data

Fig. 1 shows the location of the surface observing sites that are used for the evaluations discussed below. Details on the location and characteristics of the four

Table 1
WRF-Chem and MM5-Chem model configuration

Options	WRF-Chem	MM5-Chem
Advection scheme	Runge–Kutta 3rd order	Smolarkiewicz 3rd order
Microphysics	NCEP 3-class simple ice	Mixed phase (Reisner)
Long-wave radiation	RRTM	RRTM
Short-wave radiation	Dudhia	Dudhia
Surface layer	Monin–Obukhov (Janjic Eta)	Monin–Obukhov
Land-surface model	OSU	RUC LSM (Smirnova)
Boundary layer scheme	Mellor–Yamada–Janjic TKE	Mellor–Yamada TKE
Cumulus parameterization	Betts–Miller–Janjic	Grell–Devenyi
Photolysis scheme	Madronich (1987)	Madronich (1987)
Chemistry option	RADM2	RADM2
Aerosol option	MADE/SORGAM	None

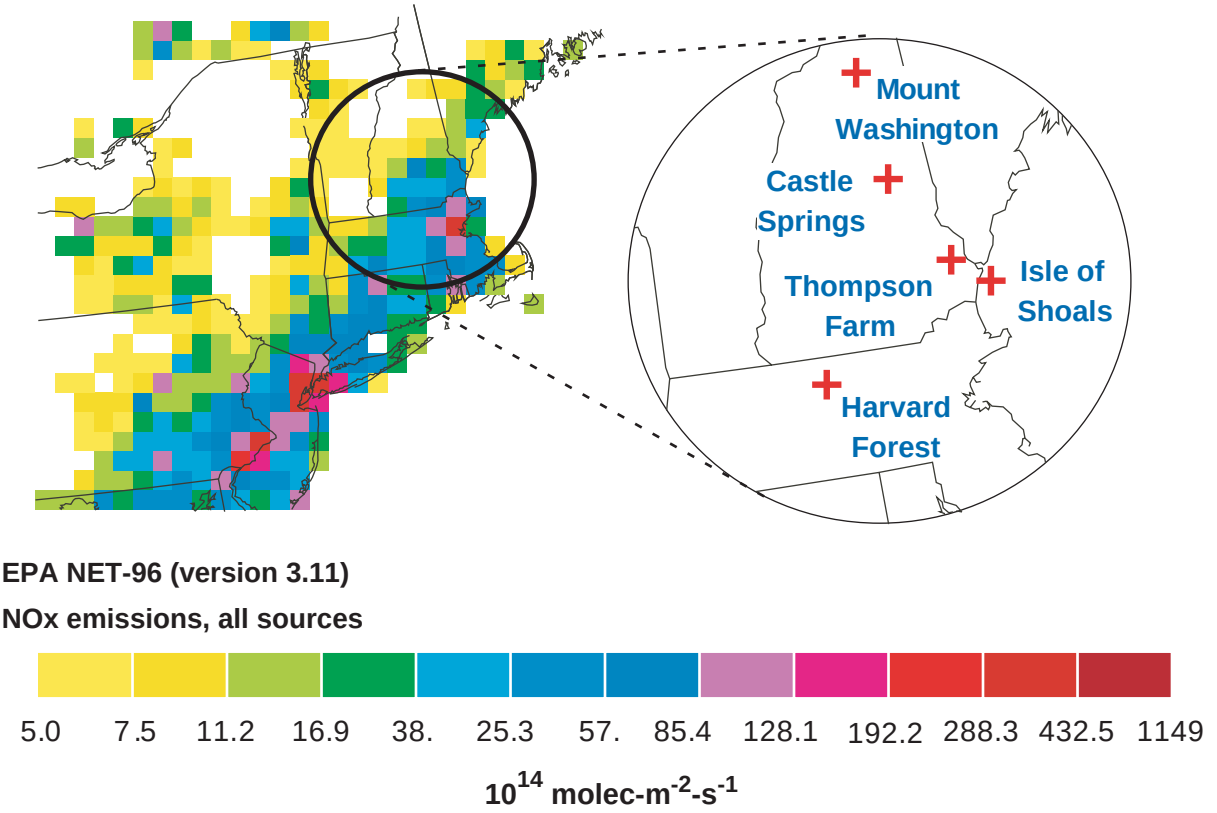

Fig. 1. Diurnally averaged summertime weekday NO_x emissions for the 27-km horizontal grid in the New England region, and the location of the five surface sites used in the statistical evaluation.

surface sites can be found at internet address http://www.airmap.unh.edu/home. The elevations of the Thomson Farm, Castle Springs, and Mount Washington sites are 75, 400, and 1915 m, respectively. Both model results and observations suggest a strong influence from the Boston region on the air quality at Thompson Farm and Isle of Shoals, a mixed source of urban coastal and more regional sources from the Boston–Washington corridor affecting the Castle Springs site, and except for nearby emissions from a cable car and parking lot, only long-range regional sources affecting the Mount Washington site. Gas-phase species directly comparable between the Air Quality Forecast Models (AQFMs) and the individual sites include CO and O_3 at Isle of Shoals; CO, O_3, NO, NO_y and SO_2 at Thompson Farm and Castle Springs; and CO, O_3, NO, and SO_2 at the Mount Washington site. The only aerosol data available for model comparison at the land base sites are at Thompson Farm, which consists of $PM_{2.5}$ aerosol mass data. Data from the AIRMAP sites are archived on a 1-min time base, and hourly averages are calculated for comparisons with hourly snapshots of the model results.

The Harvard Forest site has been collecting air quality data for more than a decade and is well characterized in terms of anthropogenic and natural sources and transport paths (e.g. Goldstein et al., 1995; Munger et al., 1998), as well as O_3 and related photochemistry (e.g. Hirsch et al., 1996). Air quality at this site is most often impacted by southwesterly airflow from the New York City–Washington DC corridor. The Harvard FOREST site archived hourly averaged CO, O_3, NO_y, NO, NO_2 and PAN.

For comparison with the AIRMAP and Harvard FOREST data, the time period of the statistical analysis extends from 0000 UTC 13 July to 0000 UTC 20 August 2002. Each model had complete coverage during this period allowing 38 days of model-measurement overlap. Only data and model results for the 11:00 AM to 1900 EDT (1500 to 2300 UTC) are used in the analysis. These hours usually bracket the maximum diurnal O_3 concentrations at all the sites. These late-morning to afternoon hours are also chosen so the observations are representative of a larger spatial footprint due to the efficiency of boundary layer turbulent mixing. Nighttime and early morning comparisons, particularly for O_3

precursors, tend to be less well correlated, and highly dependent on forecast stability. Because of the short lifetime of NO and its extreme variability, comparisons between model and measured NO are not considered here. Some final caveats in the statistical comparisons should be noted. Because of intermittent outages and problems with data logging at Harvard Forest, only about two-thirds of the total possible hourly averages are available. Because of the direct influence of the parking lot below Mount Washington, all 1-min samples with NO greater than 8 ppbv are removed from the analysis at this station. The influence of forest fires or other biomass burning is not expected to be significant for the Northeast US during the study period considered here. Though high, unexpected levels of CO recorded at the AIRMAP monitoring sites, and visible satellite images of fires originating in central Quebec are noted between 6 and 10 July 2002 (DeBell et al., 2004), there is no corresponding evidence for significant fire influence to the Northeast US after this period.

Fig. 2 shows an example of hourly averaged O_3 at Thompson Farm with the WRF/Chem model results. The 15–23 h forecasts correspond to the 0000 UTC daily forecast, while the 3–11 h forecasts correspond to the daily 1200 UTC run of this particular model and resolution. There are 342 comparison points for each forecast lead time, allowing for high confidence in the statistics derived in these comparisons. Two statistical measures (Fig. 2) are used to compare the various model forecasts: the Pearson's r^2 correlation coefficient as a measure of forecast skill, and the median error (model minus observation) as a measure of model bias. Determination of this latter quantity is illustrated in Fig. 2b for the three separate forecast lead times and the combined dataset. Model errors are sorted, and the error at the midpoint of the sorted distribution is noted, along with the errors at the 1/6 and 5/6 quantiles to describe the error spread within the central two-thirds of the error distribution set.

The r^2 and median error statistics for all of the WRF/Chem and MM5/Chem O_3 predictions are summarized graphically in Fig. 3. For the MM5/Chem model, results from all three-model resolutions are shown for completeness. The r^2 coefficients derived from 8-h averages are also included in these plots, as discussed further. Several important aspects of the model statistics have been discussed in a report that compares the MM5/Chem results with another AQFM (McKeen et al., 2003). The most relevant comparisons for the purposes of this study are between the WRF/Chem results (shown as crosses), and the 27-km horizontal resolution MM5/Chem results. For O_3, the WRF r^2 coefficients (based on hourly averages) are higher than those of MM5/Chem for 12 out of the 15 possible lead-time/site combinations. Biases are generally indistinguishable between all of the model cases. One can conclude that the WRF/Chem

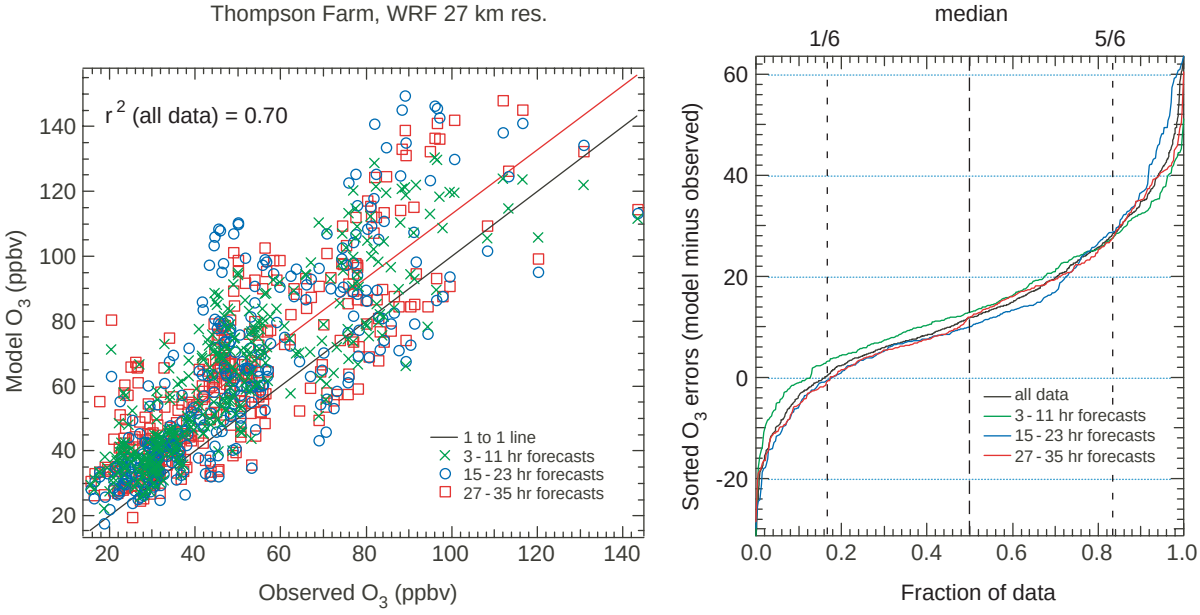

Fig. 2. Scatterplot (a) of WRF/Chem model versus observed O_3 at Thompson Farm, NH, between 13–20 July 2002, windowed between 1500 and 2300 UTC hours. Observations are averaged over hourly intervals (coincident with model results). The red line is the regression line for the linear-least squares fit for all forecasts. Also shown is the distribution of model errors (b), sorted in ascending order from the data shown in Fig. 2a. The dotted lines show the position of the 1/6 and 5/6 quantiles; the dashed line is at the median.

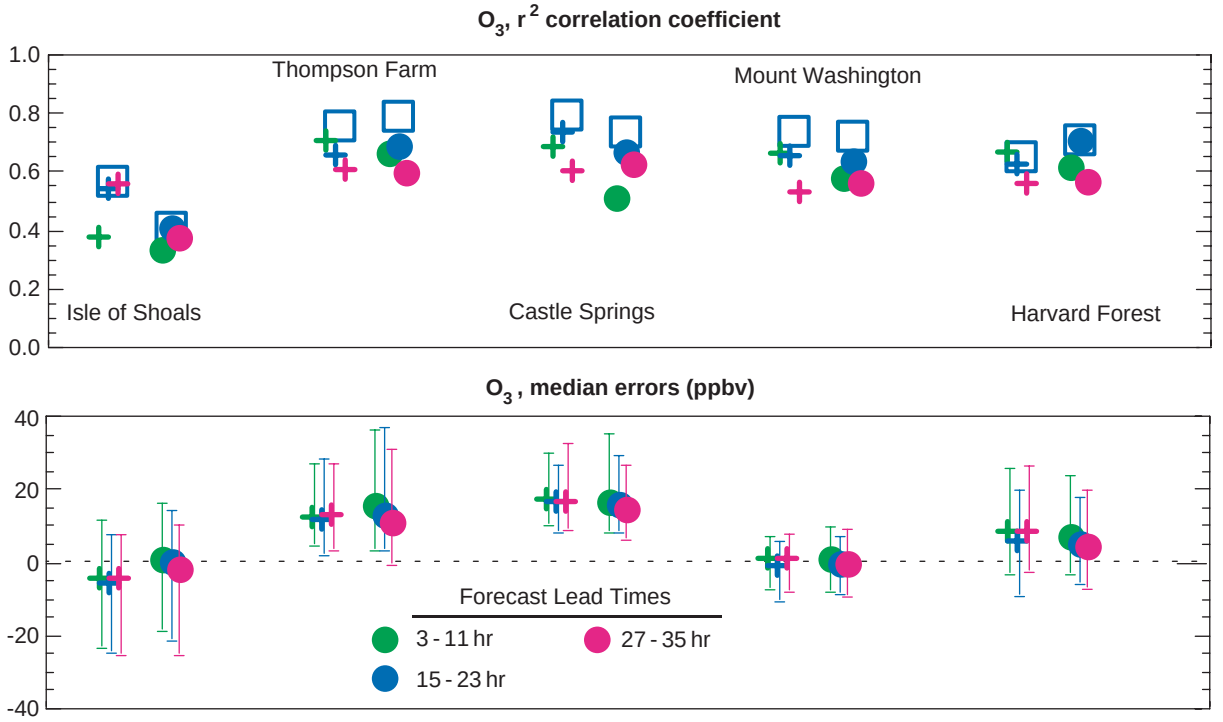

Fig. 3. Summary statistics (r^2 correlation coefficients, and median errors with bars showing 1/6 and 5/6 quantiles) for O_3 at the four AIRMAP sites and Harvard Forest. Data have been windowed for comparisons between 1500 and 2300 UTC from 13 July and 20 August 2002. The abscissa shows the five stations with some small scatter to distinguish the various forecast lead times (green—3–11 h, blue—15– 23 h, purple—27– 35 h, red—39– 47 h). MM5/Chem results are solid filled circles (hourly averages used in comparisons), crosses correspond to the WRF/Chem. The squares are statistics for the 8-h (1500 and 2300 UTC) averages of the 15- to 23-h lead-time forecast.

model exhibits improved model skill relative to MM5/Chem for O_3. Although less confidence is associated with the r^2 values derived from 8-h averages (only 38 points in the linear regressions), they are always as large or larger than the r^2 values derived from 1-h averages. This may imply that model/observation correlations at each site are driven by the models' ability to simulate large scale, day-to-day variability in O_3, as opposed to variability forced by processes acting over timescales from one to several hours.

Unlike O_3, NO_y has negligible photochemical sources, and provides a more direct link between anthropogenic source regions and transport to the various sites. Fig. 4 shows the statistical measures for NO_y for those surface sites with NO_y measurements. In the case of NO_y, sorted distributions of observations and model results generally conform to a log-normal distribution rather than just a normal distribution. For this reason, Pearson r^2 values of the log-transformed mixing ratios are used as a measure of forecast skill, and median values of sorted distributions of the model/observation ratio are used as the measure of model bias. The patterns for the NO_y statistical measures for the hourly averages show that 8 out of 9 lead-time/site combinations show improved r^2

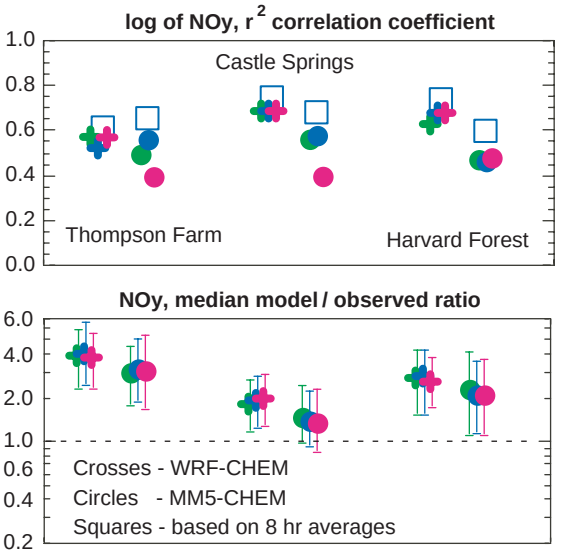

Fig. 4. As in Fig. 3, except for NO_y. Because of log-normal distributions in NO_y concentrations and model errors, Pearson r^2 correlation coefficients are from logarithms of mixing ratios, and biases are represented by median model/observed ratios.

values with WRF/Chem compared to MM5/Chem. All model cases overpredict NO_y by a factor of two or more at the Thompson Farm and Harvard Forest, which could be due to coarse spatial partitioning in the emissions inventory, inefficient vertical mixing and dispersion, or the partitioning of NO_y into forms of odd-nitrogen other than HNO_3 (which is efficiently removed by surface deposition). However, at all sites the WRF/Chem model is biased higher than the MM5/Chem model. The most likely cause of this persistent model difference is related to the parameterizations of the planetary boundary layer (PBL) physics used in the two formulations, specifically with respect to the parameterization of the surface fluxes and the way that they are coupled to the boundary layer. The fact that O_3 (photochemically produced well above the surface) does not show a difference in model bias, but NO_y (primary sources from surface emissions) does show a difference (CO showed the same behavior as NO_y, but is not shown here), suggests that upward transport (by turbulence or vertical advection) out of the bottom few model layers is sufficiently different between the models to affect the statistics.

Fig. 5 shows a scatterplot of hourly averaged $PM_{2.5}$ from the model and the observations using the continuous ambient mass monitoring system, environmental monitoring EMS Andersen instrument (CAMMS) at Thompson Farm. There is clearly a correlation between the model and observations, particularly at the high end. However, the median model $PM_{2.5}$ under-prediction is 55%. The major source of model $PM_{2.5}$ mass is from unspeciated primary emissions, rather than the condensation of gas-phase inorganic and organic species. The shallow slope of the linear regression (0.26) shown in Fig. 5 suggests that either $PM_{2.5}$ emissions from the EPA-NET 96 inventory

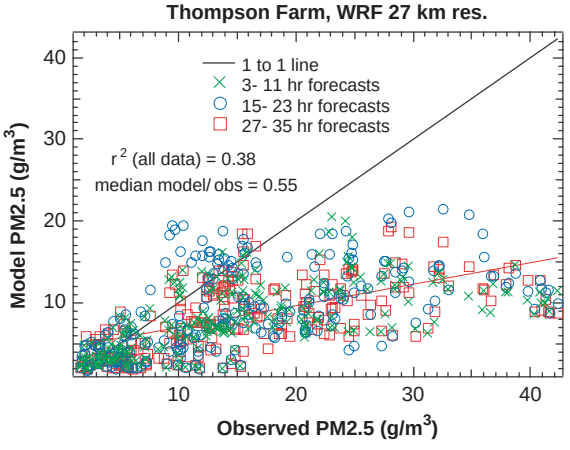

Fig. 5. As in Fig. 2, except for $PM_{2.5}$.

are too low, or that the model is not adequately treating the exchange of mass from the gas to aerosol phase during high pollution transport to this site. This point is discussed further below in regard to $PM_{2.5}$ comparisons from measurements on board the *Ron Brown*.

4.2. Comparisons with the NOAA R/V Ron Brown data

Details concerning some of the instrumentation, science objectives, and cruise tracks of the *Ron Brown* research vessel within the NEAQS-2002 field campaign are outlined in de Gouw et al. (2003) and Warneke et al. (2004). Although a much larger set of gas-phase, particulate, and radiative measurements were made aboard the *Ron Brown*, we restrict the data comparisons to those gas-phase species that were also measured at the land-based sites cited above, and the $PM_{2.5}$ aerosol mass and speciation determined by the aerosol mass spectrometer (AMS, Aerodyne Corp.) instrument. Comparisons are also restricted to dates when the *Ron Brown* was north of 39° latitude, or between 14 July and 7 August 2002. During this time, the cruise track of the ship was centered in the region of Isle of Shoals and the Massachusetts coastline except for a single run up the Maine coastline under very clean conditions between 24 and 25 July 2002. Data for 6 August 2002 are also excluded, since the ship's heading and wind direction were nearly identical for most of this day, and measurements were strongly influenced by the ship's exhaust.

Table 2 summarizes the forecast skill and model bias for SO_2, O_3, NO_y, and $PM_{2.5}$ aerosol mass from the two models and for NOAA's *Ron Brown* from the 0000 UTC forecasts. The statistics for O_3 are very similar to those obtained at the AIRMAP Isle of Shoals site, which is not surprising since the *Ron Brown* spent a significant fraction of its time in or near the model grid cell that includes Isle of Shoals. The bias is near zero for both model cases, however, the WRF model shows significant improvement in forecast skill. For NO_y and SO_2 the WRF model shows no improvement in skill, and shows an increased positive bias. The statistics for NO_y for both models are very similar to the results for the AIRMAP Thompson Farm site, except for a much reduced skill for both models in forecasting the hourly averages. Since the comparisons with hourly averages of the ship data involve both spatial and temporal variations, and the 27-km grid resolution is too coarse to resolve many of the plumes advected off the US coastline, the reduced skill in the hourly averaged NO_y versus the 8-h averages is to be expected. The statistics for the SO_2 comparisons are likewise similar to that at Thompson Farm, except the positive bias for MM5/Chem is smaller compared to the WRF/Chem bias. Observations from the ship transects yield a Boston urban plume width ranging from 16 to 20 km, while the

Table 2
Summary statistics for comparisons between WRF and MM5 chemistry models with observations collected on the *Ron Brown* between 14 July and 7 August 2002, and excluding 6 August 2002

	1-h averages			8-h averages		
	r^2	Model-obs median	1/6 and 5/6 quantiles	r^2	Model-obs median	1/6 and 5/6 quantiles
O_3						
WRF/Chem	0.57	−0.2	−14/12	0.60	−1.1	−11/10
MM5/Chem	0.36	2.7	−17/19	0.41	1.3	−16/14
	r^2 of logs	Model/obs median ratio	1/6 and 5/6 quantiles	r^2 of logs	Model/obs median ratio	1/6 and 5/6 quantiles
No_y						
WRF/Chem	0.32	3.0	1.4/9.0	0.58	2.7	1.5/5.2
MM5/Chem	0.33	2.3	1.0/6.4	0.59	1.9	1.1/4.6
SO_2						
WRF/Chem	0.22	4.1	0.7/11.6	0.51	2.9	0.8/6.6
MM5/Chem	0.29	1.6	0.3/7.4	0.52	1.9	0.6/4.6

Model/observation comparisons are only done for the 1500– 2300 UTC time period. Only statistics for the 0000 UTC forecast (15–23-h forecast lead time) are shown.

model's plume is smeared over 3 or 4 of the 27 km grid cells.

Fig. 6 compares $PM_{2.5}$ aerosol mass measured on the *Ron Brown* with WRF/Chem predictions for the time period between 22 and 25 July 2002. On 22 and 23 July 2002, the ship made transects through pollution plumes moving offshore from Massachusetts, in contrast to the clean conditions sampled on 24 and 25 July off the coast of Maine. The model captures the relative differences between polluted and clean air, the diurnal behavior, and the approximate magnitude of observed $PM_{2.5}$ mass (Fig. 6, upper panel). The model accurately simulates the observed large drop in $PM_{2.5}$ at about 1800 UTC on the 23 July, when the *Ron Brown* briefly left the plume. On the other hand, WRF/Chem misses much of the observed fine structure of $PM_{2.5}$ from nearby point source plumes. This is expected because of the coarse resolution of the model grid (27-km).

Fig. 6 also shows the mass of the emitted unspeciated aerosol category ("$PM_{2.5}$ Fine") and the speciated aerosol components. Even when unspeciated $PM_{2.5}$ mass is included in the total, the model underpredicts the observations during peak pollution periods. The directly emitted "Fine" category represents roughly 70% of the modeled $PM_{2.5}$ mass during selected polluted (1400–2000 UTC 23 July 2002) and clean conditions (1400–2000 UTC 24 July 2002). The modeled speciated $PM_{2.5}$ mass is a comparatively small fraction of the total and is predominantly secondary, i.e., produced after emission by either uptake of gas- phase species into the aerosol phase or through chemical reactions within aerosols.

Some portion of the model's unspeciated "fine" aerosol might be composed of sulfate (SO_4) and organic mass (OM), which could partly explain the model underestimates of these species during selected polluted and clean periods. Model SO_4 mass is about a third of that observed during the polluted period but is comparable to the observations during the clean period. The model suggests that much of the SO_2 observed on the *Ron Brown* on 23 July 2002 was transported from power plants along the Ohio River, for which the SO_2 emission inventories are relatively well characterized. The model underestimate of aerosol SO_4 may therefore indicate WRF/Chem did not convert SO_2 to aerosol SO_4 rapidly enough. One explanation for the discrepancy could be in-cloud oxidation of SO_2, which was not included in the NEAQS-2002 WRF/Chem runs. WRF/Chem also underpredicts aerosol organic mass by a factor of 10 during the polluted period and by a factor of 4 during the clean period. As noted previously, the RADM chemical mechanism used in WRF/Chem during NEAQS-2002 does not include terpene chemistry, so the model did not account completely for biogenic secondary organic aerosol formation.

4.3. Comparisons with the AIRNOW surface O_3 data

The US EPA Air Quality System (AQS) maintains an archive of O_3 collected from various national, state, tribal, and local agencies through its Aerometric Information Retrieval system currently known as AIRNOW. Quality assured hourly averaged ozone data for 832 monitors in the eastern half of the US were taken from the AQS archive facility (http://www.epa.gov/ttn/airs/airsaqs/archived%20data/downloadaqsdata.htm). Eder et al. (2005) present a detailed statistical comparison of three air quality forecast models (including MM5/Chem) for the 5–29 August 2002 time period based on an earlier release of the same AIRNOW data for about 470 monitors. We have chosen a 2-week period from 21 July through 4 August 2002 as a period

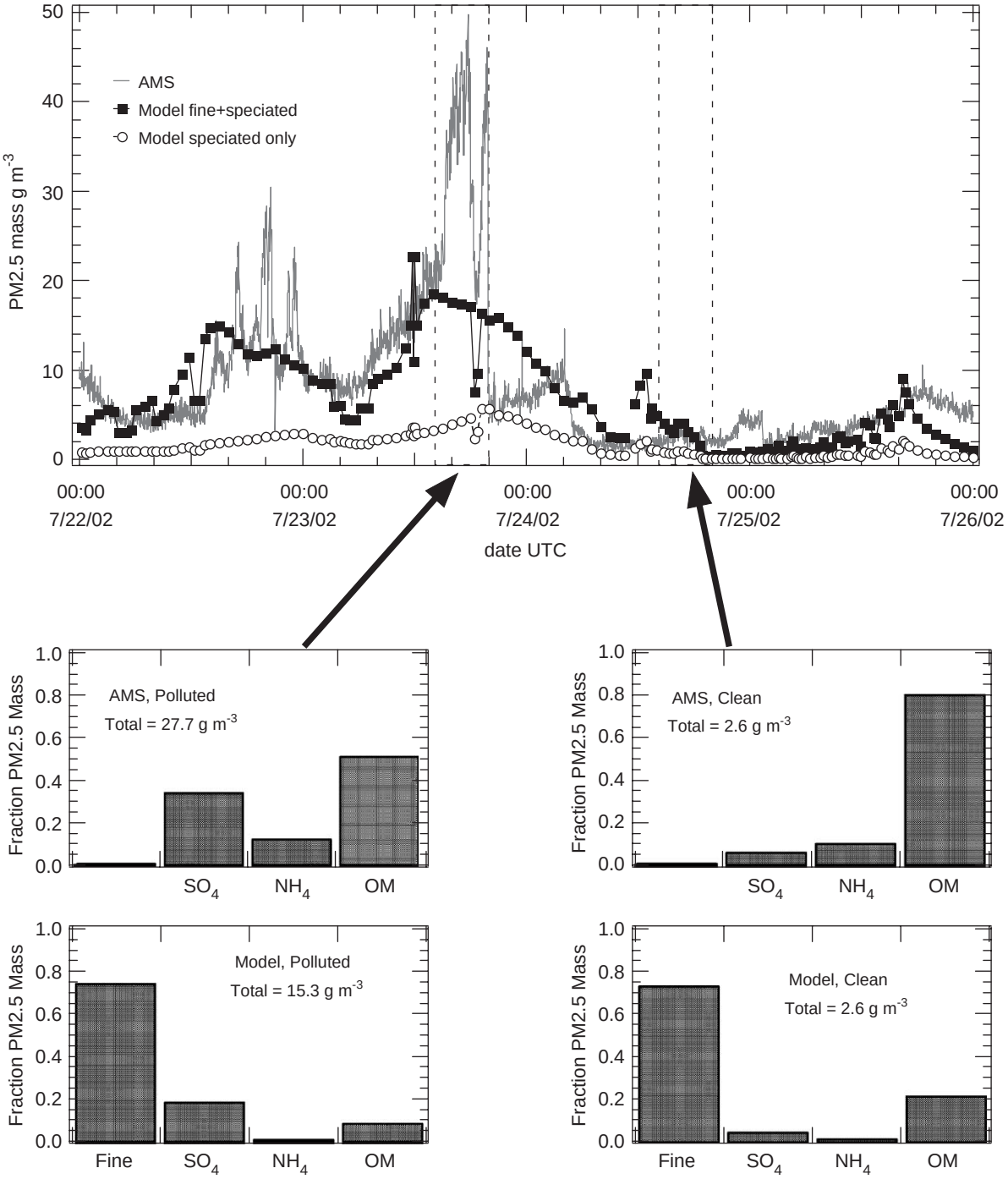

Fig. 6. Time series (upper panel) of observed dry aerosol mass from the AMS instrument compared to model predicted dry $PM_{2.5}$ mass. The model mass is shown for the total (upper panel) and for the sum of species common with the AMS measurement. The observed (middle panels) and model predicted (bottom panels) PM composition averaged over two 5-hour periods (1300–1800 UTC) for the Boston plume and clean conditions east of Maine are also shown. "Fine" refers to the unspeciated, primary $PM_{2.5}$ emissions within the NET-96 inventory.

for comparison, noting that MM5/Chem statistics obtained for this sampling period are quantitatively very similar to those reported in Eder et al. (2005) despite no overlap in the sampling periods. Because of interest in the regulatory aspects of surface O_3, statistical measures based on 1-h maximum, and 8-h maximum O_3 mixing ratios within a 24-h period are of primary interest in the forecast comparisons. Three sets of statistical measures as defined by Eder et al. (2005) are evaluated based on 1-h daily O_3 maximums: a set of discrete statistics consisting of correlation coefficents (r), mean bias (MB) and root mean square errors (RMSE), a set of categorical statistics consisting of 5 measures that characterize a model's ability to predict a 1-h O_3 exceedance (125 ppbv), and a temporal skill score that compares a model's forecast skill with persistence forecast from the previous day's 1-h maximum. For consistency with the Eder et al. (2005) study only the most recent forecast for a given hour is used in the calculations.

Spatial distributions of the discrete statistics from the 1-h maximum O_3 are shown in Fig. 7 for the MM5/Chem and WRF/Chem models. The WRF/Chem r correlations show general improvement over MM5/Chem. Table 3 also shows a small but significant increase in the median r coefficients for comparisons based on both the 1- and 8-h maximum. Likewise, Fig. 7 and Table 3 show reduced RMSEs for WRF/Chem. The MBs for WRF/Chem show negligible improvement for the 1-h maximum comparisons, but show slightly higher median MB for the 8-h maximum O_3 comparisons (Table 3). In its current configuration WRF/Chem is therefore slightly better than MM5/Chem in terms of capturing observed variance, and RMSE, with essentially no improvement in MBs. The comparisons in Fig. 7 also point out the regional dependence of WRF/Chem statistical improvement with the Northeast US coastline west to Ohio showing general improvement in all three statistical measures. From Lake Michigan southwestward to Oklahoma MBs tend to be negative for WRF/Chem but not for MM5/Chem, and this region generally has lower r coefficients and higher RMSE as well. The source of this regional dependence has not been identified with certainty, however as previously mentioned, the land-use parameterizations of the two models were inherently different, and specification of surface water temperature over the Great Lakes appears colder in the WRF/Chem, which may contribute to the large, negative biases around Lake Superior and Lake Michigan.

Comparisons for the categorical statistics are based upon the location of points within a model versus observed scatterplot relative to exceedance limits imposed on the quantity of interest. Fig. 8 shows the scatterplot for 1 h O_3 maximums relative to observed values for the same set of points used in the discrete analysis relative to a 125 ppbv exceedance limit. The number of points in the 4 quadrants a, b, c, and d in Fig. 8 are used to evaluate accuracy, probability of detection, false alarm rate, critical success index and bias defined in the last column of Table 3. Similar to the discrete comparisons, the results summarized in Table 3 show that WRF/Chem has marginal improvement over MM5/Chem for all the categorical measures except bias, and the critical success index of the 1-h maximum O_3. The values derived here for the discrete and categorical statistics can be compared to a similar set of model/observed surface O_3 analysis by McHenry et al. (2004). Although that analysis was for 8-h maximum O_3 for only the Northeast US and a 10-day O_3 episode in 2001, the numbers reported here for the categorical statistics (other than false alarm rate), MB, and RMSE are consistent and competitive with the two forecast models and predictions from forecast agencies throughout the Northeast US.

Last, the temporal skill score is compared in Table 3, defined as

$$\text{Skill score} = 100\% \times \left[\frac{\text{RMSE}_{\text{prev}} - \text{RMSE}_{\text{modl}}}{\text{RMSE}_{\text{prev}}}\right], \quad (6)$$

where "prev" stands for either the 1- or 8-h maximum observed O_3 of the previous day. A negative value between a possible range from -100 to 100 means that the model forecast for tomorrow is worse than a prediction of O_3 maximum based solely on today's observations, positive means better. The median for all points of this quantity is directly related to the number of points that beat the temporal persistence forecast, and the summary in Table 3 shows that WRF/Chem is the same or slightly worse than MM5/Chem in terms of this statistical measure. The median RMSEs derived from the previous day's forecast are 17.6 and 15.4 ppbv for the 1- and 8-h maximum O_3, respectively. Since both the WRF/Chem and MM5/Chem models have equivalent mean RMSEs, one would expect about half the points to beat persistence based on RMSE comparisons. However, if r coefficients are used as the measure of skill, the models' performance would appear less favorable. In this case, the median r coefficients calculated using previous day's observations are 0.92 for both the 1- and 8-h maximum, which is much higher than the models' median r values. Only ~17% of the model points beat persistence if the r coefficients are used as the measure of skill.

The statistical comparisons outlined above suggest a small but perceptible improvement in surface O_3 statistics with WRF/Chem, but the differences between MM5/Chem and WRF/Chem can be put in the context of other model uncertainties, and how these other model uncertainties contribute to surface O_3 statistical measures. One inherent model uncertainty is with the biogenic emissions of VOC (e.g., isoprene) that

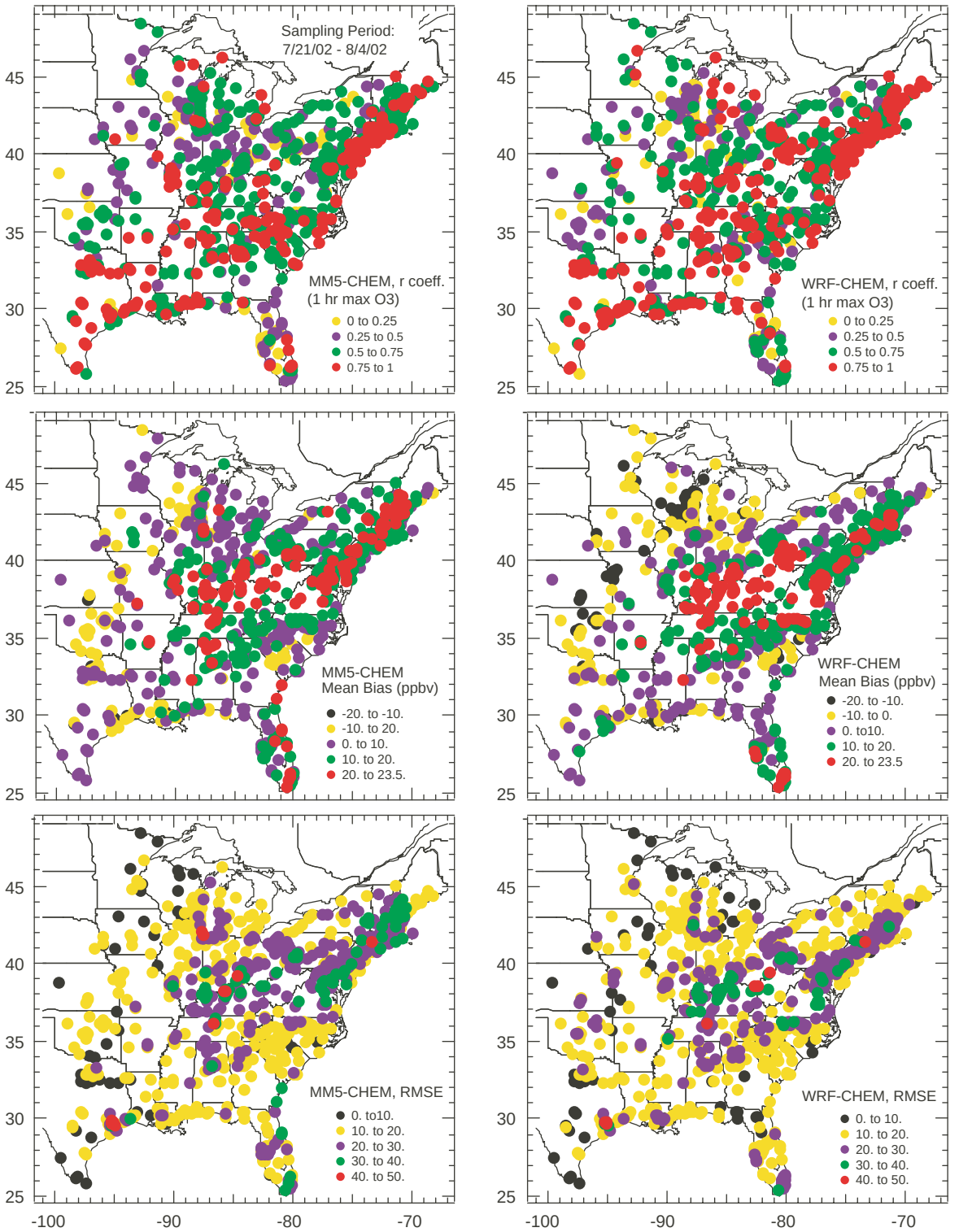

Fig. 7. Pearson *r* correlation coefficients, MB, and RMSE for the MM5/Chem and WRF/Chem models when comparing daily 1-h maximum O_3 with the EPA AIRNOW surface O_3 network between 21 July and 4 August 2002.

Table 3
Discrete and categorical surface O_3 statistics, and temporal skill scores

	1-h max O_3			8-h max O_3			
	MM5	WRF	WRF*	MM5	WRF	WRF*	Units
r correlation coefficient	0.60	0.65	0.63	0.62	0.68	0.67	Unitless
Mean bias	8.92	7.61	0.78	8.37	9.36	3.56	ppbv
Root mean square error	19.15	17.86	15.40	17.30	16.14	13.74	ppbv
Accuracy	97.7	97.7	99.4	83.2	84.7	90.5	$100\% \times (b+c)/(a+b+c+d)$
Probability of detection	17.3	19.0	3.8	63.5	71.8	47.1	$100\% \times b/(b+d)$
False alarm rate	96.0	96.4	91.7	82.5	81.7	74.5	$100\% \times a/(a+b)$
Critical success index	3.3	3.2	2.7	15.9	17.1	19.8	$100\% \times b/(a+b+d)$
Bias	4.35	5.24	0.46	3.64	3.92	1.85	$(a+b)/(b+d)$
Skill score (RMSE)	−1.9	−4.2	3.7	−9.4	−10.7	−1.2	Percent
% > Persistence (RMSE)	46.6	46.1	52.6	41.1	41.0	49.0	Percent
% > Persistence (*r*)	17.3	16.5	24.5	17.4	16.9	24.0	Percent

All values are medians from 833 AIRNOW surface stations. Statistics based on 1-h maximum O_3 and 8-h maximum O_3 values for MM5/Chem (MM5 in column header), WRF/Chem (WRF), and the WRF/Chem model with the leaf temperature approximated by the lowest model level air temperature in the isoprene emissions calculation (WRF*). The *a*, *b*, *c*, and *d* values in the units column refer to the number of points within the corresponding quadrants in Fig. 7. Categorical statistics based on 1-h max O_3 use 125 ppbv limits to define quadrants, while those based on 8-h max O_3 statistics use 85 ppbv limits. The last two rows give the percentage of points beating persistence when either RMSE or *r* coefficients are used for comparison.

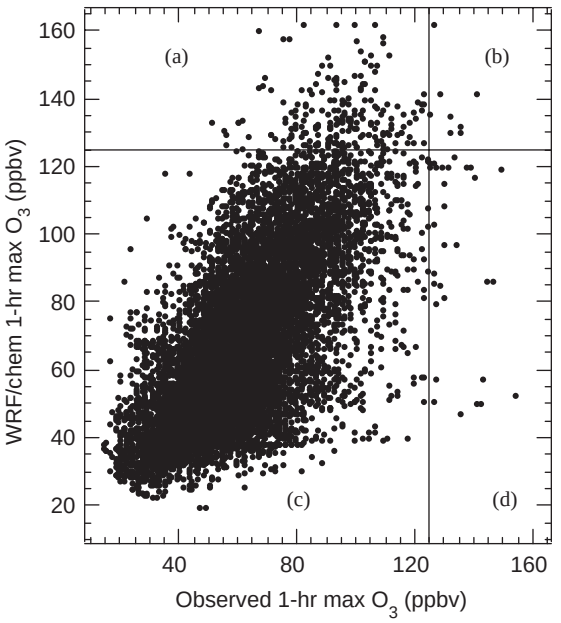

Fig. 8. Scatterplot of WRF/Chem model versus observed daily 1-h maximum O_3 for all EPA AIRNOW surface stations in the model domain between 21 July and 4 August 2002. The quadrants defined by the 125 ppbv exceedance limit are lettered, and used to define the categorical statistics in Table 3.

contribute to O_3 formation. The parameterization of isoprene emissions is highly dependent on the leaf temperature assigned to a particular model grid (Pierce et al., 1998) with about a factor of 2 difference in emissions per 5 °C temperature difference. By design the leaf temperature assigned within the original MM5/Chem (Grell et al., 2000) was taken as the ground surface temperature. The WRF/Chem model also used this assignment for consistency with model comparisons, but to test the influence of the leaf temperature assignment assumption on O_3 statistics, an additional O_3 simulation with WRF/Chem was done with leaf temperature assigned to ambient air temperature, as suggested by Pierce et al. (1998). The statistics for this case, labeled WRF*, are also given in Table 3. This modified WRF/Chem case shows improved statistical comparisons relative to both the original WRF/Chem and MM5/Chem for all measures except the *r* correlation coefficient, the probability of detection, and the critical success index for 1-h O_3 maximums. Fig. 9 illustrates the distribution of *r*-correlation coefficients and mean O_3 biases for the three model cases. The distribution of *r*-coefficients is clearly weighted towards higher values for both of the WRF/Chem simulations compared to that of MM5/Chem. The bias histogram for the base WRF/Chem simulation is somewhat broader than for MM5/Chem, but with little impact on the median, while the adjusted WRF/Chem simulation shows a clear shift in the histogram distribution to lower model bias. Fig. 9 and Table 3 show that the differences between the modified and original WRF/Chem mean O_3 biases are much more significant than between WRF/Chem and MM5/Chem. Biogenic emissions uncertainty represents only one component of the uncertainty associated with the total model system. Several other model components, including

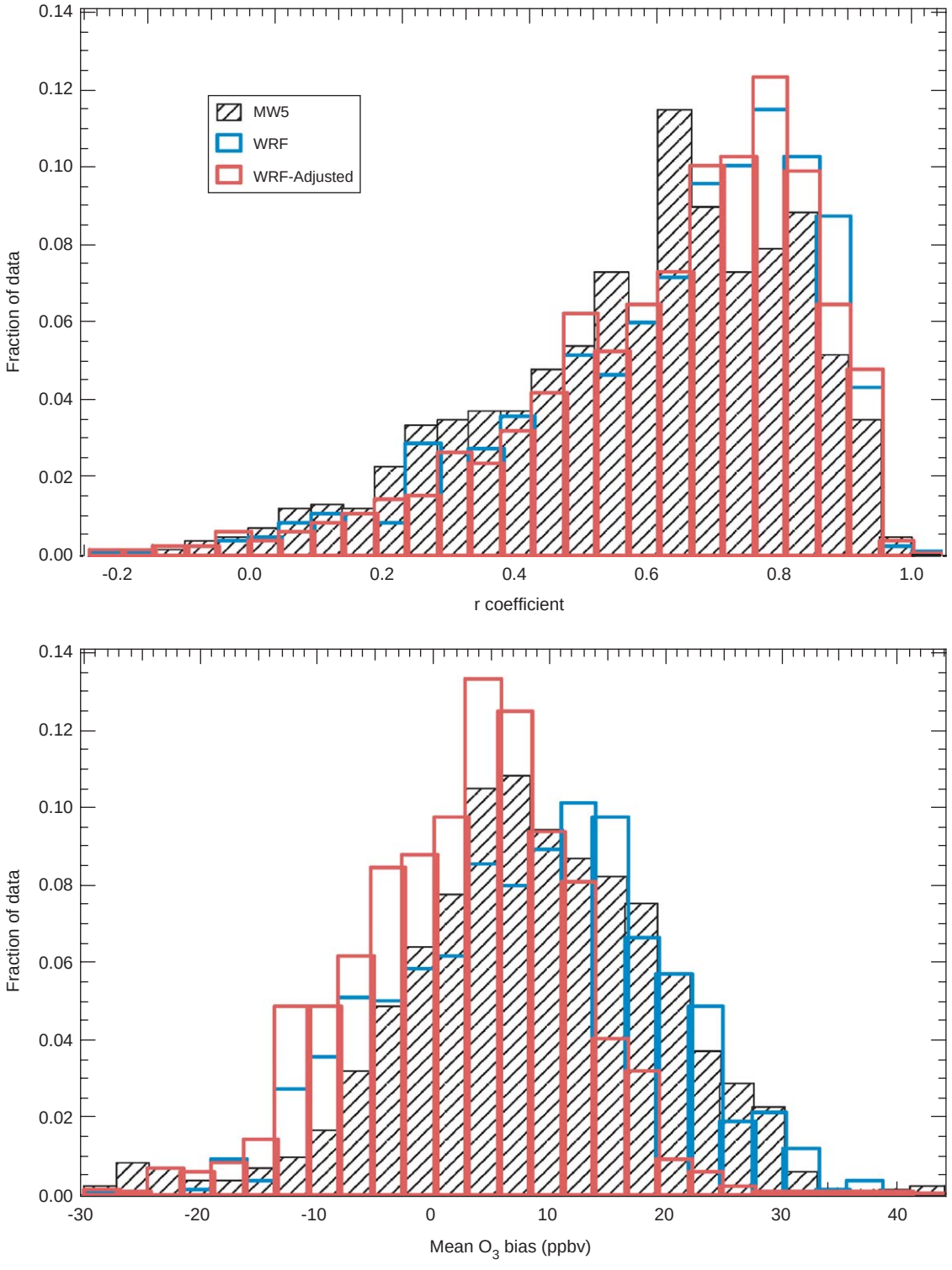

Fig. 9. Histograms of the *r*-correlation coefficient (top panel) and MB (bottom panel) for the 833 AIRNOW O_3 monitor comparisons, and for the MM5/Chem, the base WRF/Chem, and the isoprene-adjusted WRF/Chem model cases. The ordinate is the fraction of the 833 comparison data points that lies within the 28 intervals on the abscissa.

anthropogenic emissions of several key species, cloud, boundary layer, and surface layer exchange parameterizations, and boundary condition assumptions, also contribute to model uncertainty. Thus, even though WRF/Chem shows statistical improvement over MM5/Chem in terms of surface O_3 statistics, these improvements are probably minor compared to the impact that other known model uncertainties would have on the same statistics.

5. Summary

Fully coupled online chemistry has been implemented into the WRF model. The resulting WRF/Chem model was then statistically evaluated in comparison to MM5/Chem and to detailed photochemical data collected during summer 2002 NEAQS field study. The results presented are a summary of statistical comparisons of observed versus model predicted (MM5/Chem and WRF/Chem) atmospheric composition. The photochemistry and emissions are identical between the two models, allowing an examination of the effects of differences between the MM5 and WRF formulations on O_3 photochemical forecasts. Analysis of variance and bias for five surface sites and ship-based measurements of O_3 and its precursors allow some important qualitative generalizations to be made. First, the WRF/Chem model statistically shows better skill in forecasting O_3 than MM5/chem with no appreciable differences between models in terms of bias with the observations. Second, the WRF/Chem model also consistently exhibits better skill at forecasting the O_3 precursors CO and NO_y at all of the surface sites. However, the WRF/Chem model biases of these precursors and other gas-phase species are persistently higher than for MM5/Chem, and are most often biased high compared to observations. The reason behind the higher WRF/Chem biases is probably related to differences in vertical transport between the two models, particularly with the treatment of the bottom few layers within the different PBL physics parameterizations. This points to the importance of vertical transport algorithms and transport rates within the air quality forecasts, and the need for verification of these transport algorithms with appropriate information regarding vertical structure and gradients. Last, when statistical analysis is applied to the 1100–1900 EDT averages of the model and measured data, forecast skill for O_3 and its precursors is always better than the same statistics based on hourly data for the same time periods. The improvement in the forecast skill of WRF/Chem, though not always very large, may be related to improved predictions of larger scale dynamics and physical meteorology within the WRF formalism. The WRF/Chem model also shows improvement in the forecast skill of surface O_3 from the 833 AIRNOW network monitors within the model domain. However, this improvement in forecast skill is minor compared to the impact of other basic model assumptions on the same statistics, as demonstrated by one example of the sensitivity of the statistical measures with respect to the treatment of biogenic VOC emissions.

Acknowledgements

The authors gratefully acknowledge the measurements of the AIRMAP network (University of New Hampshire), Harvard Forest (Bill Munger) and the NOAA RV Ron Brown (Eric Williams, Christoph Senff, Ann Middlebrook, Joost deGouw, Carsten Warneke) that contributed to this work. We are also grateful to Nita Fullerton for editorial assistance, and Brent Shaw for reviewing this manuscript. In addition we thank members from the EURAD group at the Institute for Geophysics and Meteorology, University of Cologne, for providing a version of the aerosol module. Finally, two anonymous reviewers helped significantly improve this manuscript.

References

Ackermann, I.J., Hass, H., Memmesheimer, M., Ebel, A., Binkowski, F.S., Shankar, U., 1998. Modal aerosol dynamics model for Europe: development and first applications. Atmospheric Environment 32 (17), 2981–2999.

Bacon, D.P., Ahmad, N., Boybeyi, Z., Dunn, T.J., Hall, M.S., Lee, P.C.S., Sarma, R.A., Turner, M.D., Waight, K.T., Young, S.H., Zack, J.W., 2000. A dynamically adapting weather and dispersion model: the operational multiscale environment model with grid adaptivity (OMEGA). Monthly Weather Review 128, 2044–2076.

Binkowski, F.S., Shankar, U., 1995. The regional particulate matter model, 1. Mode description and preliminary results. Journal of Geophysical Research 100, 26191–26209.

Byun, D.W., Ching, S. (Eds.), 1999. Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system, EPA/600/R-99/030, US EPA, Research Triangle Park, NC.

Chang, J.S., Binkowski, F.S., Seaman, N.L., McHenry, J.N., Samson, P.J., Stockwell, W.R., Walcek, C.J., Madronich, S., Middleton, P.B., Pleim, J.E., Lansford, H.H., 1989. The regional acid deposition model and engineering model. State-of-Science/Technology, Report 4, National Acid Precipitation Assessment Program, Washington, DC.

DeBell, L.J., Talbot, R.W., Dibb, J.E., Munger, J.W., Fischer, E.V., Frolking, S.E., 2004. A major regional air pollution event in the northeastern United States caused by extensive forest fires in Quebec, Canada. Journal of Geophysical Research 109 (D19), D19305.

de Gouw, J.A., Goldan, P.D., Warneke, C., Kuster, W.C., Roberts, J.M., Marchewka, M., Bertman, S.B., Pszenny,

A.A.P., Keene, W.C., 2003. Validation of proton-transfer-reaction mass spectrometry (PTR-MS) measurements of gas-phase organic compounds in the atmosphere during the New England Air Quality Study (NEAQS) in 2002. Journal of Geophysical Research 108 (D21), 4682.

Eder, B.K., Kang, D., Stein, A.F., McHenry, J., Grell, G.A., Peckham, S.E., 2005. The New England Air Quality Forecasting Pilot Program: Development of an evaluation protocol and performance benchmark. Journal of Air and Waste Management Association 55, 20–27.

Erisman, J.W., van Pul, A., Wyers, P., 1994. Parameterization of surface resistance for the quantification of atmospheric deposition of acidifying pollutants and ozone. Atmospheric Environment 28, 2595–2607.

Goldstein, A.H., Wofsy, S.C., Spivakovsky, C.M., 1995. Seasonal variations of nonmethane hydrocarbons in rural New England: constraints on OH concentrations in northern midlatitiudes. Journal of Geophysical Research 100, 21023–21033.

Grell, G.A., Emeis, S., Stockwell, W.R., Schoenemeyer, T., Forkel, R., Michalakes, J., Knoche, R., Seidl, W., 2000. Application of a multiscale, coupled MM5/chemistry model to the complex terrain of the VOTALP valley campaign. Atmospheric Environment 34, 1435–1453.

Guenther, A.B., Zimmerman, P.R., Harley, P.C., Monson, R.K., Fall, R., 1993. Isoprene and monoterpene emission rate variability: model evaluations and sensitivity analyses. Journal of Geophysical Research 98D, 12609–12617.

Guenther, A., Zimmerman, P., Wildermuth, M., 1994. Natural volatile organic compound emission rate estimates for US woodland landscapes. Atmospheric Environment 28, 1197–1210.

Hahn, J., Steinbrecher, J., Steinbrecher, R., 1994. Studie F: Emission von Nicht-Methan-Kohlenwasserstoffen aus der Landwirtschaft. In: Enquete-Kommission 'Schutz der Erdatmosphaere' des Deutschen Bundestages (Hrsg.), Studienprogramm Band 1 'Landwirtschaft', Teilband 1. Economica Verlag, Bonn.

Hirsch, A.I., Munger, J.W., Jacob, D.J., Horowitz, L.W., Goldstein, A.H., 1996. Seasonal variation of the ozone production efficiency per unit NO_x at Harvard Forest, MA. Journal of Geophysical Research 101, 12659–12666.

Jones, A., Roberts, D.L., Slingo, A., 1994. A climate model study of indirect radiative forcing by anthropogenic sulphate aerosols. Nature 370, 450–453.

Joseph, J.H., Wiscombe, W.J., Weinmann, J.A., 1976. The delta-Eddington approximation for radiative flux transfer. Journal of Atmospheric Science 33, 2452–2458.

Lin, S.-J., Rood, R.B., 1996. Multidimensional flux-form semi-Lagrangian transport schemes. Monthly Weather Review 124, 2046–2070.

Kulmala, Laaksonen, Pirjola, 1998. Parameterization for sulphuric acid/water nucleation rates. Journal of Geophysical Research 103, 8301–8307.

Madronich, S., 1987. Photodissociation in the atmosphere, 1, actinic flux and the effects of ground reflections and clouds. Journal of Geophysical Research 92, 9740–9752.

McHenry, J.N., Ryan, W.F., Seaman, N.L., Coats Jr., C.J., Pudykiewicz, J., Arunachalam, S., Vukovich, J.M., 2004. A real-time Eulerian photochemical model forecast system: overview and initial ozone forecast performance in the Northeast US corridor. Bulletin of the American Meteorological Society 85, 525–548.

McKeen, S.A., Wotawa, G., Parrish, D.D., Holloway, J.S., Buhr, M.P., Hübler, G., Fehesenfeld, F.C., Meagher, J.F., 2002. Ozone production from Canadian wildfires during June and July of 1995. Journal of Geophysical Research 107 (D14) 0.1029/2001JD000697.

McKeen, S.A., Eder, B., Grell, G.A., McHenry, J., Stein, A., Angevine, W.M., 2003. Evaluation of prototype air quality forecast models—chemistry, NOAA/OAR report.

Mellor, G.L., Yamada, T., 1982. Development of a turbulent closure model for geophysical fluid problems. Reviews of Geophysics and Spacephysics 20, 851–875.

Middleton, P., Stockwell, W.R., Carter, W.P.L., 1990. Aggregation and analysis of volatile organic compound emissions for regional modeling. Atmospheric Environment 24A, 1107–1133.

Munger, J.W., Fann, S.-M., Bakwin, P.S., Goldstein, M.L., Goldstein, A.H., Colman, A.S., Wofsy, S.C., 1998. Regional budgets for nitrogen oxides from continental sources: variations of rates for oxidation and deposition with season and distance from source regions. Journal of Geophysical Research 103, 8355–8368.

Odum, J.R., Hoffmann, T., Bowman, F., Collins, D., Flagan, R.C., Seinfeld, J.H., 1996. Gas/particle partitioning and secondary organic aerosol yields. Environmental Science Technology 30, 2580–2585.

Pierce, T., Geron, C., Bender, L., Dennis, R., Tonneson, G., Guenther, A., 1998. Influence of increased isoprene emissions on regional ozone modelling. Journal of Geophysical Research 103, 25611–25629.

Pleim, J.E., Venkatram, A., Yamartino, R., 1984. ADOM/TADAP Model Development Program, The Dry Deposition Module, vol, 4. Ont. Min. of the Environment, Canada.

Ruggaber, A., Dlugi, R., Nakajima, T., 1994. Modeling of radiation quantities and photolysis frequencies in the troposphere. Journal of Atmospheric Chemistry 18, 171–210.

Satoh, M., 2002. Conservative scheme for the compressible nonhydrostatic models with the horizontally explicit and vertically implicit time integration scheme. Monthly Weather Review 130, 1227–1245.

Saxena, P., Hudischewskyj, A.B., Seigneur, C., Seinfeld, J.H., 1986. A comparative study of equilibrium approaches to the chemical characterization of secondary aerosols. Atmospheric Environment 20, 1471–1483.

Schell, B., Ackermann, I.J., Hass, H., Binkowski, F.S., Ebel, A., 2001. Modeling the formation of secondary organic aerosol within a comprehensive air quality model system. Journal of Geophysical Research 106, 28275–28293.

Schoenemeyer, T., Richter, K., Smiatek, G., 1997. Vorstudie uber ein raumlich und zeitlich aufgelostes Kataster anthropogener und biogener Emissionen fuer Bayern mit Entwicklung eines Prototyps und Anwendung fur Immissionsprognosen. Abschluss bericht an das Bayerische Landesamt fur Umweltschutz. Fraunhofer-Institut fuer Atmosphaerische Umweltforschung, Garmisch-Partenkirchen.

Simpson, D., Guenther, A., Hewitt, C.N., Steinbrecher, R., 1995. Biogenic emissions in Europe. 1. Estimates and uncertainties. Journal of Geophysical Research 100D, 22875–22890.

Slingo, A., 1989. A GCM parameterization for the shortwave radiative properties of water clouds. Journal of the Atmospheric Sciences 46, 1419–1427.

Slinn, S.A., Slinn, W.G.N., 1980. Prediction for particle deposition on natural waters. Atmopsheric Environment 14, 1013–1016.

Stockwell, W.R., Middleton, P., Chang, J.S., Tang, X., 1990. The second-generation regional acid deposition model chemical mechanism for regional air quality modeling. Journal of Geophysical Research 95, 16343–16367.

Stockwell, W.R., Kirchner, F., Kuhn, M., Seefeld, S., 1997. A new mechanism for regional atmospheric chemistry modeling. Journal of Geophysical Research 102, 15847–25879.

US Environmental Protection Agency (EPA), 1998. National air pollutant emission trends, procedures document, 1900–1996, Rep. EPA-454/R-98-008, Office of Air Quality Planning and Standards, Research Triangle Park, NC, 148pp.

Warneke, C., de Gouw, J.A., Goldan, P.D., Kuster, W.C., Williams, E.J., Lerner, B.M., Jakoubek, R., Brown, S.S., Stark, H., Aldener, M., Ravishankara, A.R., Roberts, J.M., Marchewka, M., Bertman, S., Sueper, D.T., McKeen, S.A., Meagher, J.F., Fehsenfeld, F.C., 2004. Comparison of daytime and nighttime oxidation of biogenic and anthropogenic VOCs along the New England coast in summer during New England Air Quality Study 2002, Journal of Geophysical Research 109 (D10309), doi:10.1029/2003JD004424, 2004.

Wesley, M.L., 1989. Parameterization of surface resistance to gaseous dry deposition in regional numerical models. Atmospheric Environment 16, 1293–1304.

Whitby, E.R., McMurry, P.H., Shankar, U., Binkowski, F.S., 1991. Modal aerosol dynamics modeling, Rep. 600/3-91/020, Atmospheric Research and Exposure Assessment Laboratory, US Environmental Protection Agency, Research Triangle Park, NC, 1991 (Available as NTIS PB91-1617291AS from National Technical Information Service, Springfield, VA).

Wicker, L.J., Skamarock, W.C., 2002. Time splitting methods for elastic models using forward time schemes. Monthly Weather Review 130, 2088–2097.