Available online at www.sciencedirect.com
SCIENCE DIRECT
Atmospheric Environment 40 (2006) 3737–3749
ATMOSPHERIC ENVIRONMENT
www.elsevier.com/locate/atmosenv

Evaluation of the GEM-AQ air quality model during the Québec smoke event of 2002: Analysis of extensive and intensive optical disparities

N.T. O'Neill[a,*], M. Campanelli[b], A. Lupu[c], S. Thulasiraman[a], J.S. Reid[d], M. Aubé[a], L. Neary[c], J.W. Kaminski[c], J.C. McConnell[c]

[a]*CARTEL, Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Qué., Canada J1K 2R1*
[b]*Institute of Atmospheric Sciences and Climate, Area della Ricerca di Torvergata, Via Fosso del Cavaliere 100, 00133 Rome, Italy*
[c]*CRESS, Department of Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto, Ont., Canada M3J 1P3*
[d]*Naval Research Laboratory, 7 Grace Hopper Avenue, Stop 2, Monterey 93943-5502, CA, USA*

Received 31 October 2003; received in revised form 28 February 2006; accepted 4 March 2006

Abstract

The root-mean-square (rms) differences between the Canadian air quality model GEM-AQ and measurements for intensive and extensive optical variables (aerosol optical depth or AOD and Ångström exponent or α) were investigated using data from the July 2002 Québec smoke event. In order to quantify regional differences between model and measurements we employed a three component analysis of rms differences. The behaviour of the two absolute amplitude rms components of AOD (difference of the means and the difference of the standard deviations) enabled us to infer emission properties which would otherwise have been masked by the larger 'anti-correlation' component. We found the inferred emission fluxes to be significantly higher than the original geostationary, satellite-derived FLAMBÉ (fire locating and modelling of burning emissions) emissions flux estimates employed as inputs to the simulations. The model captured the regional decrease of the intensive α exponent (increase of particle size with trajectory time), while the agreement with the extensive AOD parameter was marginal but clearly dependent on the nature of the spatio-temporal statistical tools employed to characterize model performance. In establishing the α versus trajectory time trend, the modelled AOD data was filtered in the same way as the measured data (very large AODs are eliminated). This processing of modelled results was deemed necessary in order to render the α results comparable with the measurements; in the latter case it was difficult, if not impossible, to discriminate between measured α trends due to instrumental artifacts (non-linearities at low signal strength) versus trends due to coagulative effects.

Keywords: Forest fire; Aerosol optical depth; Ångström exponent

1. Introduction

Model evaluation is a multi-disciplinary research activity which ranges from qualitative 'eyeball' comparisons to sophisticated spatio-temporal

*Corresponding author.
E-mail address: Norman.T.ONeill@USherbrooke.ca (N.T. O'Neill).

1352-2310/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2006.03.006

statistical analyses. It is a somewhat subjective endeavour given the different criteria and objectives of the different communities with direct or indirect interests in model evaluation. Difficulties include the gap between model output parameters and available measurement data, the characterization of model uncertainty, the introduction of measurement bias due to data reduction techniques, and a general culture of poor interaction and communication between the separate disciplines involved in the evaluation process (Berk et al., 2002).

All-encompassing and popular statistical parameters such as the root mean square (rms) difference between model and measurement, while arguably an advance on eyeball comparisons, can be so phenomenological or top-down as to be of limited value (Berk et al., 2001). The correlation coefficient between ensembles of modelled and measured data (across time and/or space) can be instructive but is often given in an incomplete context in the sense of how its value impacts on the error budget for a given model parameter. A graphical tool such as a Taylor diagram (Taylor, 2001) divides the rms error into its general component contributions related to correlative and amplitude differences but does not show all contributions at once, nor the sign and absolute magnitude of these contributions. A spatio-temporal evaluation of differences between model and measurements is certainly the most fundamental approach to model evaluation. However, aerosol optical data is arguably the most temporally and spatially sparse atmospheric data available and typically does not lend itself well to data intensive processes such as spatial or temporal auto-correlation techniques. A further complication is related to retrieval as well as statistical biases; high AODs in particular are very sparse.

The literature on model evaluation in terms of aerosol optical parameters is largely concentrated on generic single-wavelength total aerosol optical depth (AOD) comparisons, while the interpretation of these comparisons is largely of a qualitative or semi-qualitative nature. There is, moreover, minimal literature on the evaluation of model performance in terms of both extensive optical properties, such as AOD, and intensive, size and/or type dependent optical properties, such as the Ångström exponent (α). Lesins and Lohmann (2005) did, for example, present a study on the use of AERONET and MODIS fine mode fractions as prescribed inputs to a GCM; this process provided an indirect evaluation of an intensive variable (fine mode fraction) in terms of how its variability influenced the modelled versus measured agreement of (extensive) AOD. Finally, we would note that there has been a tendency to use the measured optical parameters as is with little analysis of possible environmental or instrumental biases such as the presence of thin homogeneous cloud or radiometric artifacts.

The principal motivation for the current work was to evaluate the performance of the GEM-AQ air quality model (the Canadian air quality model described below) within a context of both extensive and intensive variability. Given the array of choices and difficulties associated with model evaluation approaches, we chose to focus our evaluation efforts on strong-signal events whose aerosol optical variations were well understood. A natural first choice was to investigate the numerous Boreal forest fire smoke events which affect nearly all parts of Canada each year. This process-level approach allows one to investigate the behaviour of model evaluation criteria across significant variations in both AOD and the Ångström exponent, while ensuring that cloud contamination and instrumental artifacts are avoided or at least minimized.

Boreal forest fires emit large amounts of aerosol particles and trace gases that can be transported as smoke to distant locations. These smoke events can have a substantial impact on air quality, visibility and regional radiative balance. Smoke emissions from Canadian wildfires have been known to influence optical measurements as far away as Europe (Wandinger et al., 2002), or to alter the photochemical production of ozone in the south-eastern United States (Wotawa and Trainer, 2000). The extremely warm and dry summer of 2002 created conditions for widespread wildfire activity across central and eastern Canada. On July 2 and 3, thunderstorms sparked more than 85 fires in central and western Québec. Cloudiness and weak tropospheric winds prevented these fires from intensifying until July 5 and 6. The smoke from the fanned fires was transported rapidly and with little dispersion into southern Québec, central and eastern Ontario, and north-eastern United States (Colarco et al., 2004; Taubman et al., 2004; O'Neill et al., 2005; Pahlow et al., 2005; Sapkota et al., 2005).

Smoke events generally provide strong and unambiguous fine mode aerosol optical signals. The Québec smoke event was characterized by extraordinarily large AODs which yielded clear

spatio-temporal patterns of aerosol abundance and which provided an optimal environment for short term coagulative growth effects. Smoke radiances are often easy to discriminate from cloud contaminated radiances in ground-based and satellite data, and source emissions are relatively strong and easily identified. The evaluation of an air quality model in the presence of an event such as the Québec smoke event is thus, in numerous ways, an excellent proving ground for model evaluation strategies as well as for testing model performance on important and robust indicators of intensive and extensive aerosol behaviour.

2. Model and measurements employed

In order to simulate the 2002 Québec smoke event, we used the Global Environmental Multiscale Air Quality model (GEM-AQ) which is based on the global variable-resolution multiscale model developed by the Meteorological Service of Canada for operational weather prediction (Côté et al., 1998), and includes a size-resolved multi-component aerosol module (Gong et al., 2003) and an online gas-phase chemistry module based on the condensed mechanism of Lurmann et al. (1986). The host meteorological model can be configured to simulate the atmosphere over a broad range of scales, from the global scale down to the meso-γ scale. GEM-AQ was run for the period 6–11 July on a 191×174 resolution horizontal grid. The grid has a uniform core of 100×100 cells where the size resolution is 0.22°. Moving away from the core, the grid size increases by a constant factor of 1.1. The grid is centered on 45°N, 75°W, and has 28 hybrid vertical levels up to 10 hPa. High resolution (0.22°) objective analysis data interpolated onto the model grid were used to update the meteorological fields every 12 h. The model was run with a time step of 15 min.

The emission fluxes of smoke aerosol particles were generated from the real-time Fire Locating and Modeling of Burning Emissions (FLAMBÉ) system (Reid et al., 2004, www.nrlmry.navy.mil/flambe/) which uses the GOES Wildfire Automated Biomass Burning Algorithm (WF_ABBA) fire products (Prins et al., 2001). The WF_ABBA is a real-time automated dynamic thresholding algorithm which uses visible, shortwave-infrared (3.9 μm) and longwave-infrared (10.7 μm) data to locate fire pixels and characterize sub-pixel burning. The net emission for an individual fire burning in a particular biome is determined by factoring in the land area burned, fuel loading, carbon fraction of the fuel, combustion fraction and emission factor (Reid et al., 2005). This general methodology is all that is available for real-time air quality applications.

The smoke particles were assumed to be 6% black carbon (BC) and 94% organic matter (OM) by mass (as per Reid et al., 1999; we argue below that the optical impact of this nominal choice of internal mass fractions is not large). The particle mass was distributed across 12 logarithmically spaced radius size bins between 0.01 and 1 μm. Initial size distributions were taken as unimodal and lognormal. The particles were allowed to grow hygroscopically and by coagulation (Gong et al., 2003). In this simulation, the chemical module was turned off since detailed organic chemistry is not accounted for in the model. By this we mean that organic chemistry processes likely dominate smoke particle evolution and thus that the application of the existing chemistry processing package would arguably be irrelevant because the most important chemical processes would be neglected. This decision was mitigated by the fact that the optical outputs of the model are largely dependent on the bulk properties of the smoke aerosols (number density and size) and only weakly dependent on aerosol type and mixtures of aerosol type (cf. the discussion in Section 4.1). Condensation of inorganic species onto smoke particles was also neglected, since mass or optical properties are not likely to be altered in a significant way.

The aerosol optical depths were calculated for seven wavelengths using a Mie theory code to compute matrices of size-bin and relative humidity (RH) dependent extinction cross sections which were applied off-line to each bin (each size-discriminated tracer) for all grid cells at all model time steps. Each bin was optically processed as an internally and homogeneously mixed size distribution of OM and BC in volume proportions derived from the mass fractions given above. The optical implications of the assumed BC/OM mass composition as well as the assumption of internal homogeneous optical mixing are discussed in the model uncertainty section below.

The output of the model was compared with AOD and Ångström exponent (α) values measured at 10 AERONET/AEROCAN stations across Eastern Canada (Egbert, Howland and Halifax) and North-Eastern United States (MDSC, CCNY, COVE, Norfolk State, SERC, Wallops and GSFC) (O'Neill et al., 2005). Both parameters were

computed at a reference wavelength of 500 nm from a second order polynomial applied to the 380, 440, 500, 670, 870, 1020 nm AERONET channels (ibid). Level 1 (non-cloud-screened) data were used in order to ensure that legitimate smoke data was not removed by the automated cloud screening process normally applied to AERONET data (ibid). In lieu of the cloud screening algorithm, the AOD data set was carefully monitored for obvious cloud contamination using AVHRR animations and MODIS imagery, temporal excursions in AOD and α, and minimum thresholds on α and AOD. A minimum-signal (voltage) suppression filter was applied in the Ångström exponent computational chain in order to reject low count (non-linear signal response) measurements related to excessively large AOD values at the smaller wavelengths; such conditions induce non-linear spectral flattening and obvious artifacts in α (ibid). During the 6 July–11 July period, when all stations were determined to be affected by smoke for at least a few hours, smoke-event time windows were selected as detailed in O'Neill et al. (2005).

3. Model evaluation methodology

3.1. Choice of comparative runs

Several runs over a variety of initial conditions were performed in order to study the sensitivity of the GEM-AQ model to the input parameters (Table 1). Run *B* was chosen as a reference run for assessing the influence of perturbations from a base set of run conditions. It corresponded roughly to what one might characterize as the best run (after a large number of model evaluation iterations). However we did not pursue the idea of a best run any more than was necessary given (a) the interpretative flexibility associated with what exactly one means by a best run, and (b) given the uncertainty in performance indicators (associated with model uncertainty for example).

The range of initial size distributions was constrained by limitations found in the literature (Reid et al., 2005). The choice of injection height ranges was influenced by the back-trajectory analyses of (Colarco et al., 2004). While these former parameters could be tuned to improve the model performance indicators to a certain extent, we could not produce a reasonable level of agreement between model and measurements given the emission flux estimates from the FLAMBÉ system. Use of the raw FLAMBÉ emissions resulted in AODs which were significantly smaller than those observed. We note that the location of the fires and the diurnal variation of the emissions are parametric inputs which are fairly well constrained by the satellite input data to the FLAMBÉ system and thus that there is little variational liberty in these aspects of the source modelling. Gradual increases of FLAMBÉ emissions by single-digit factors led to better agreement with measurements. This trial and error approach to emissions tuning was the source of the three emissions factor choices described immediately below.

Table 1
List of GEM-AQ runs. Run *B* is the reference run. The variable input parameters relative to run *B* are shown in bold

Run	Initial r_M (μm)	Initial σ_M	Injection layer (km)	Emissions
B	0.14	1.7	1–5	$6 \times \Phi_{FLAMBÉ}$
*S*1	0.14	**2.0**	1–5	$6 \times \Phi_{FLAMBÉ}$
*S*2	0.14	**1.5**	1–5	$6 \times \Phi_{FLAMBÉ}$
*R*1	**0.12**	1.7	1–5	$6 \times \Phi_{FLAMBÉ}$
*R*2	**0.16**	1.7	1–5	$6 \times \Phi_{FLAMBÉ}$
*H*1	0.14	1.7	**0.5–5**	$6 \times \Phi_{FLAMBÉ}$
*H*2	0.14	1.7	**0.5–3.5**	$6 \times \Phi_{FLAMBÉ}$
*H*3	0.14	1.7	**2–6.5**	$6 \times \Phi_{FLAMBÉ}$
*E*1	0.14	1.7	1–5	$\mathbf{5 \times \Phi_{FLAMBÉ}}$
*E*2	0.14	1.7	1–5	$\mathbf{7 \times \Phi_{FLAMBÉ}}$
P[a]	0.14	1.7	1–5	$6 \times \Phi_{FLAMBÉ}$

[a]Meteorology updated every 24 h.

Because the initial (source) particle distribution in the model was assumed to be single-mode lognormal, evaluation tests were performed for variable geometric mass radii (r_M) and geometric standard deviations (σ_M), ranging respectively from 0.12 to 0.16 μm (runs *B*, *R*1, *R*2) and from 1.5 to 2 (runs *B*, *S*1, *S*2). Four runs were performed for smoke being uniformly injected into grid columns between altitudes of 1–5, 0.5–3.5, 0.5–5, and 2–6.5 km, respectively (runs *B*, *H*1, *H*2, *H*3) with source fluxes being updated every 30 min. For three runs (*B*, *E*1, *E*2) the emission fluxes were taken as 5, 6 and 7 times the emission flux estimated from FLAMBÉ. One run (*P*) was performed by updating the meteorological field every 24 h rather than the standard 12 h updating frequency.

3.2. Performance indicators

The starting point for quantifying differences in the observed (x_i) and modelled (y_i) fields is the root

mean square difference. For an ensemble of model runs and an ensemble of measurements, it can be expressed as a quadrature sum of one positive and two sign-dependent contributions:

$$\delta^2_{\mathrm{RMS}} = \delta^2_{\sigma} + \delta^2_{R\star} + \delta^2_{\mathrm{OFF}} = (\sigma_y - \sigma_x)^2 + \left(\sqrt{2\sigma_y\sigma_x(1-R)}\right)^2 + \delta^2_{\mathrm{OFF}}, \quad (1)$$

where σ_x and σ_y are the standard deviation for the fields x_i and y_i, respectively, R is the correlation coefficient, and $\delta_{\mathrm{OFF}} = \langle y\rangle - \langle x\rangle$ is the difference between the means of the two fields. The first term (δ_σ) expresses the variance difference, while the $\delta_{R\star}$ term represents an anti-correlation difference (0 for perfect correlation and maximum for perfect anti-correlation when $R = -1$).

The rms difference is a top-down type of performance indicator which does not allow one to appreciate the more basic variational mechanisms which induce differences between model and measurements. We found that the rms difference expressed in terms of the sign-dependent components defined above permits one to better judge and quantify model performance.

4. Model evaluation results

Fig. 1 shows temporal graphs of the modelled (run *B*) and measured smoke AODs (500 nm) for all 10 AERONET stations listed above (for measurements acquired during the smoke-event windows). Some of the largest AODs ever recorded by North American AERONET instruments were observed on 6 and 7 July in southeast Canada and the northeast USA (AOD spikes with values greater than 4 were observed at a number of sites). Lower AOD values persisted over the Maryland/Virginia area on 8 July while a large smoke mass was advected out to the Atlantic as part of anti-cyclonic movement about a low pressure zone over Maine and New Brunswick. Most of this early smoke was traceable to sources in Northern Québec on 5 July (Colarco et al., 2004). A separate major plume, traceable to sources on 7 July and 8 July, invaded the Maine/Nova Scotia region from 8 July to 10 July (O'Neill et al., 2005). These general events were captured by the model results of Fig. 1 with what can be qualitatively described as varying degrees of success; in this section we attempt to quantify the model performance in terms of the indicators discussed above.

An rms component analysis was applied to the modelled and measured data across the 6 days of the smoke event in order to analyze the performance of each model run (where each point of the rms data ensemble corresponded to a smoke-window average at a given station).

The quantities R, σ_x, σ_y, δ_{OFF} and δ_{RMS} were calculated for α and AOD over the ensemble of station data as a function of all the runs listed in Table 1. As shown for the AOD in Fig. 2 (bottom left), R assumes values generally around 0.4 for all the runs except run $H3$, where R decreases to about 0.2. The value of R in the case of α is about 0.7 for all the runs except P, where it decreases down to 0.45. Given that mean AOD and α values during the smoke events were about 1.8 and 1.3, respectively (500 nm), it is evident that the δ_{RMS} discrepancy is relatively larger in the case of AOD (for which δ_{RMS} is dominated by $\delta_{R\star}$). The variance and absolute amplitude terms play a more significant role in the case of α since the $\delta_{R\star}$ term is relatively much less important.

The δ_σ, $\delta_{R\star}$ and δ_{OFF} components of the AOD rms error were primarily sensitive to the height of the injection layer and to the emission strengths. For an injection height range of 0.5–5 km (runs $H1$ and $H2$), there is a recognizable decrease in δ_σ and δ_{OFF}. This altitude range is somewhat lower than the 2–6 km range reported by Colarco et al. (2004) (however their backtrajectories were performed only for the GSFC station: injection heights for east coast sites were somewhat lower). The δ_{OFF} and δ_σ components for runs B, $E1$ and $E2$ show a clear emissions-dependent negative to positive bias, while the anti-correlation term is only moderately influenced by emissions strength.

In the case of α, no strong sensitivity was found for $\delta_{R\star}$ as a function of initial conditions. In contrast, δ_σ decreases moderately in (negative) amplitude, while δ_{OFF} decreases from positive to negative values with increasing σ_{M} and r_{M}; this means that for larger initial distribution sizes the model evolution tends to underestimate the downstream Angstrom exponent (overestimates particle size). It is noted that the small σ_{M} value of run $S2$ is more consistent with the few available values of σ_{M} (actually, σ_{Volume}) available from the Dubovik inversion after elimination of the small-mode anisotropy discussed in O'Neill et al. (2005); in terms of successive increases in σ_{M} (from run $S2$ to run B to run $S1$) one notes that the combined errors of δ_σ and δ_{OFF} go through a minimum somewhere

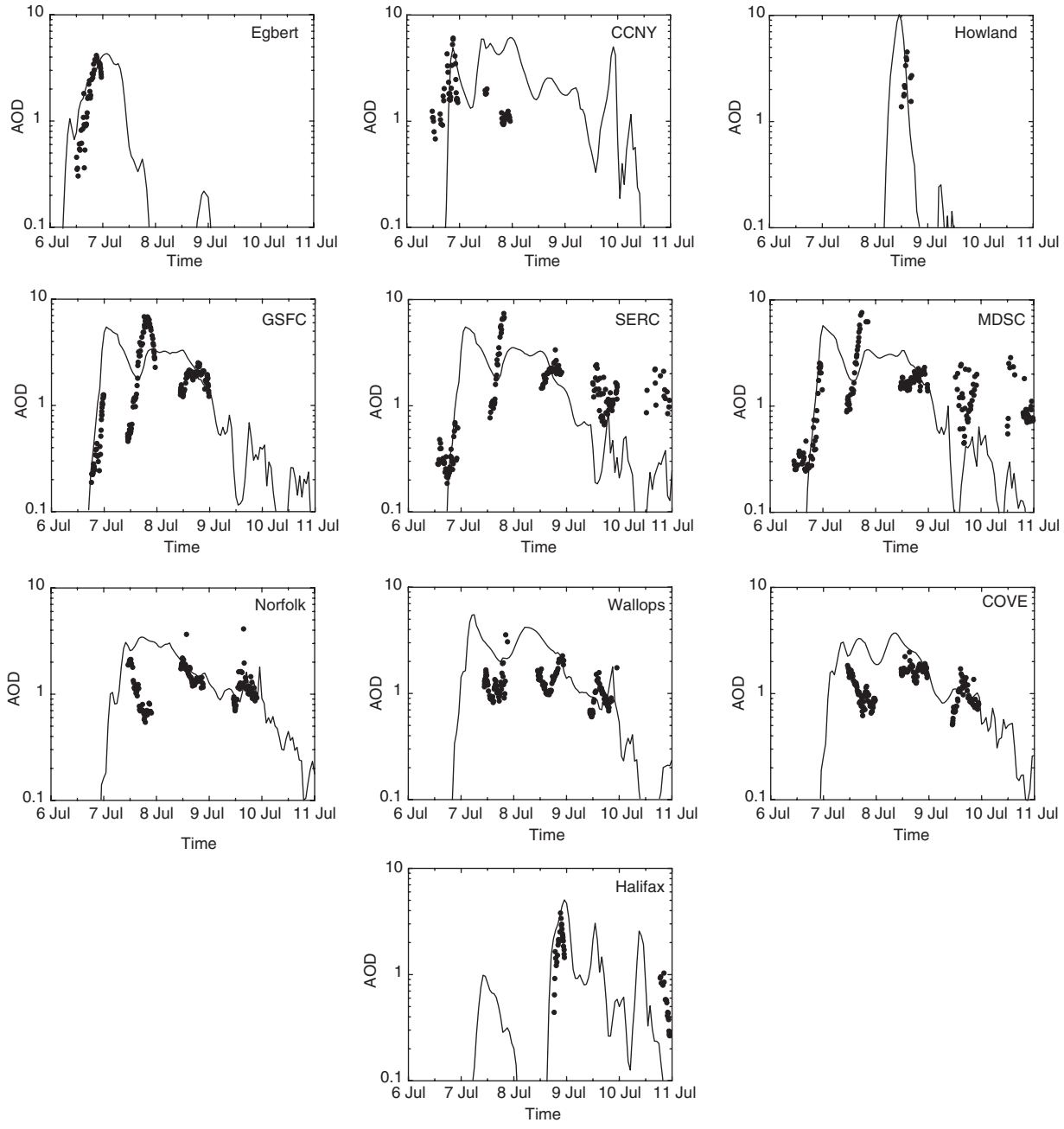

Fig. 1. Time series (solid line) of AOD (500 nm) for all stations during the Québec smoke event of July 2002. The solid line represents the GEM-AQ AOD output while the solid circles represent AERONET data.

between run $S2$ and run B. Some δ_σ and δ_{OFF} sensitivity to emission strength (likely to be attributable to coagulation influences) is observable between the $E1$ and $E2$ cases for α. The most readily apparent α sensitivity to the 24 h meteorology updates is R, which deteriorates from about 0.7 to 0.45 (from run B to P).

4.1. Meteorological and optical model uncertainty

In this subsection we attempt to provide indicators of meteorological and optical model uncertainties whose impact was not explicitly investigated in the rms difference study of Fig. 2. While hardly representing a comprehensive analysis of model

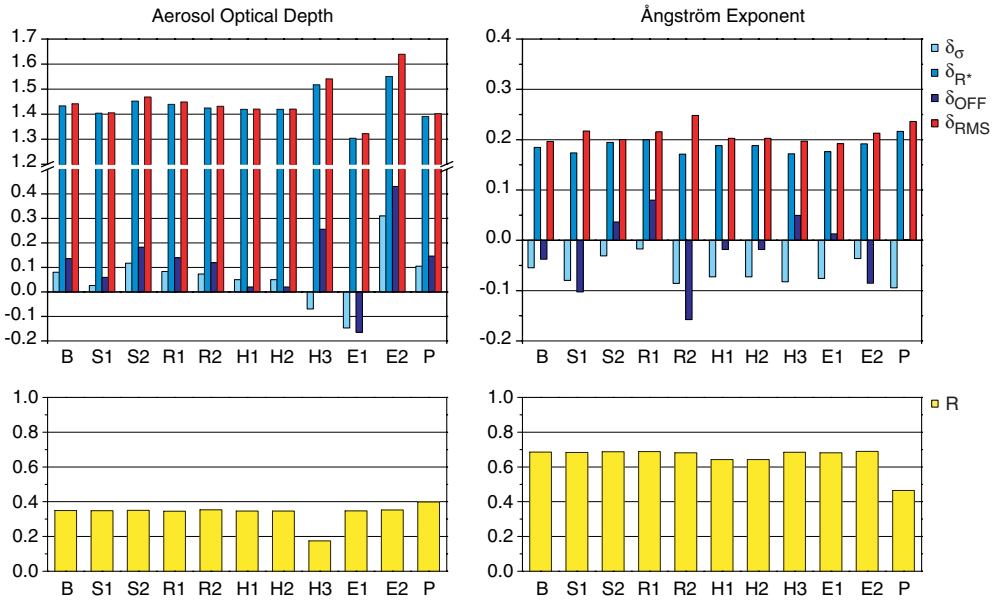

Fig. 2. Comparison of model versus measurements in terms of the RMS error (in red) and its components (different shades of blue), and correlation coefficient (yellow). The letters on the horizontal axis represent the model runs defined in Table 1. AOD results are shown in the left hand column, while Ångström exponent results are shown in the right hand column.

uncertainty, these numbers at least give a context to the significance of the rms differences in that figure.

In order to provide a rough estimate of GEM-AQ uncertainty due to wind errors, an ensemble of 12 runs was produced by independently perturbing the two horizontal components of the analyzed wind fields at each grid point with spatially-uncorrelated Gaussian noise having zero mean and $2.5\,m\,s^{-1}$ standard deviation (while keeping all other parameters as in run *B*). The choice of the standard deviation value was based on the error statistics of horizontal winds from the Eta Data Assimilation System, as determined by Lin and Gerbig (2005). The (AOD and α) δ_{RMS} differences between each of the 12 runs and run *B*, for the ensemble of all station data, were not larger than ~0.05 for AOD, and not larger than ~0.01 for α. These two values thus give a meteorological uncertainty context to comparisons such as those presented in Fig. 2.

In order to assess the optical uncertainty due to the assumed smoke-aerosol mass composition of 6% BC to 94% OM, we ran a series of Mie calculations for a variety of related aerosol types. The AOD and α variations at 500 nm were investigated for significant differences in aerosol type (pure sulphates and pure organics as well as sulphates rather than OM as the 94% component), different types of internal mixtures of BC and OM, BC and sulphates and OM and sulphates (both homogeneously mixed and in a shell configuration) and different types of external mixtures with the same three components. The refractive indices for these calculations were those of Hess et al. (1998), while the particle size was taken as the initial size parameters of the reference run *B*. These computations, which yielded variations $<30\%$ in extinction coefficients and $<10\%$ in α, effectively add an additional optically-based component to the stochastic, meteorologically-based estimate of model uncertainty discussed above.

5. Physical interpretation of results

5.1. Aerosol optical depth

Although the gross features of the smoke plume are generally captured by the model, the AOD correlation coefficient across the complete six day period was found to be quite low (R~0.4). This is more likely a statement about the accuracy and spatio-temporal resolution of the analyzed meteorological fields since the major aerosol dynamics phenomenon is clearly advection. A detailed analysis of the model versus the ground-based and satellite results indicated that the model always tracked the AOD measurements in a qualitatively

consistent fashion but that it could lag or lead measurements and that the temporal scale of a given modelled event (for example the duration of a smoke plume) could differ by hours from the measurements. These temporal differences appeared to increase with time; the better comparisons were found for short trajectory times. The predicted AOD amplitude could as well differ from the real amplitude by ~100% or less. Fig. 3 shows an example for 6 July; the comparison between the modelled AOD image and the GOES (smoke enhanced) image clearly demonstrates a significant, if somewhat distorted, spatio-temporal correlation. This correlation was even more apparent when

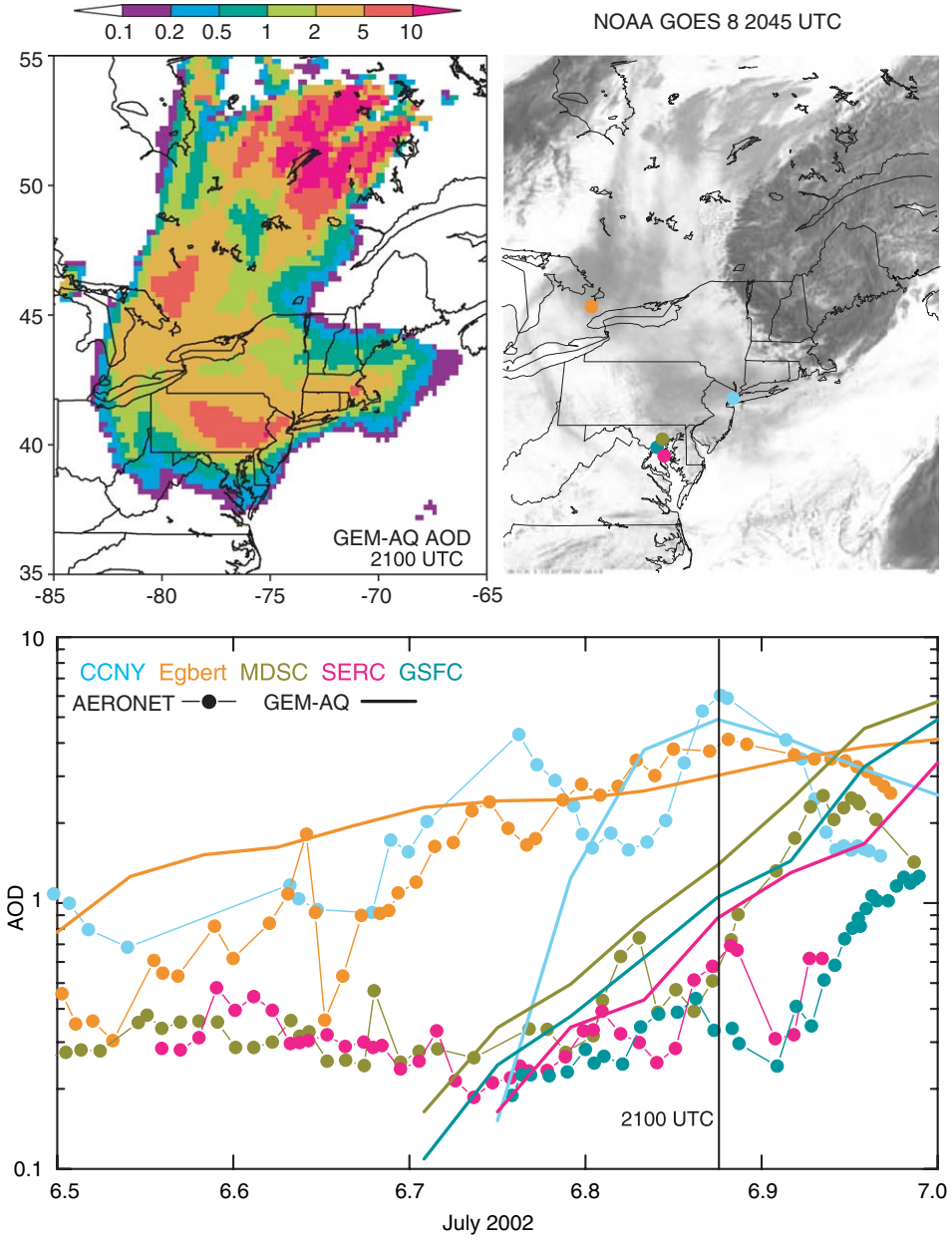

Fig. 3. AOD pattern at or near 2100 GMT for July 6, 2002: (left) output of GEM-AQ, (right) smoke enhanced results for GOES-8 (www.class.noaa.gov), and (bottom) GEM-AQ and AERONET AODs as a function of time (decimal day). The darker gray, textured pattern in the GOES-8 image represents cloud while the lighter gray, more smoothly varying patterns represent thick smoke.

comparing GOES animations of a half-hour frame rate with modelled animations over the daylight hours of 6 July.

On a purely temporal scale the modelled results (lower graph of Fig. 3), lead or lag the sunphotometer results by a difference which is significantly dependent on how one assigns an oftentimes subjective correspondence between the much more high frequency sunphotometer data and the low frequency model data. A similar remark on the subjective element of the comparisons applies to the AOD magnitude differences which are of the order of 100% or less (except for the case where the model simply failed to capture the broad stable plume preceding the double peaked plume at CCNY). Attempts to perform (variable lag) temporal cross correlations between the model and the sunphotometry results were of limited utility, as one can readily appreciate given examples such as the CCNY and SERC comparisons of Fig. 3.

The sign dependent δ_{OFF} and δ_{σ} components of δ_{RMS} clearly implied that the assumed emissions fluxes in the model were ~5–6 times too small (total emissions of 0.95–1.13 Tg of smoke inferred by eliminating the rms error discrepancy between modelled and measured AOD compared with the value of 0.18 Tg estimated from FLAMBÉ). We also note that our inferred emission flux across the 6-day event was of the same order as the model based total emission estimate of Colarco et al. (2004) (1.5 Tg). A review of the inputs to the FLAMBÉ emission flux estimates indicated that the ground-based estimates of surface burn area (from SOPFEU; Société de protection des forêts contre le feu) were a factor of approximately 2.3 larger than the real-time satellite-based estimates employed in the former, although Stocks (personal communication, 2005) also pointed out a tendency for ground teams to overestimate area burned by including areas not burned within their perimeter estimates. Conversely the FLAMBÉ system does not account for cloud cover, which was at times substantial during the simulation period. Uncertainties in fuel load, fuel consumed, emission factors, etc. for individual fires could well account for a factor of 2 or more (Reid et al., 2005). However both Lavoue (personal communication, 2005) and Stocks (personal communication, 2005) indicated that the carbon consumed figure of approximately 20 t ha^{-1} employed in the FLAMBÉ calculations, i.e., the product of the first three parameters in Eq. (2) of Reid et al. (2005), was reasonable for Canadian Boreal forests. We note, on the other hand, that emission factors for such an extreme event can be many times the more accepted values (Reid et al., 2005). Potential overestimation issues which are not included in Reid's Eq. (2) would be secondary particle production (Reid et al., 2005) and non-smoke contributions to the optical depth measurements. In the latter case, the presence of a fairly thick low-level polluting layer over the southern stations (Taubman et al., 2004) would have contributed to a certain amount of overestimation in our inferred emission fluxes (the pollution AODs being, for example, ~25% of the total AODs on 8 July). If our inferred emission fluxes are approximately correct, then other factors, such as the smoke (PM2.5) emission factor of 40 g kg^{-1} C employed in theFLAMBÉ calculations would arguably be more likely candidates for underestimation.

5.2. Ångström exponent

The movement of air masses affects the value of the AOD by virtue of the number and size of the transported particles. In our case the low correlation coefficients obtained for the AOD combined with the higher correlation coefficients of α (~0.7) are consistent with the idea that the disagreement between the observed and modelled optical thickness is mainly due to variations in the number density of transported particles, rather than variations in their size. The significant correlation of the modelled versus measured α can be explained given that the spatio-temporal variations of α included an important low frequency term which is less dependent on advective meteorology than the variations of optical thickness.

Fig. 4 shows the variation of α versus trajectory time for both model and measurements. A single α point in the comparative data ensemble consisted of an average across a smoke-event time window at a given station (one station might be represented by two or three smoke-event windows). The tendency of decreasing α (increasing particle size) with increasing trajectory time was used by O'Neill et al. (2005) to support arguments for particle growth due to agglomerative aging effects such as coagulation. This tendency, in terms of the modelled results, is generally evident in the series of α frames shown in Fig. 5. Coarsely speaking, smoke from the 5 July sources ages up to three days and this is the cause of the (low α) blue regions (Fig. 5) found over the

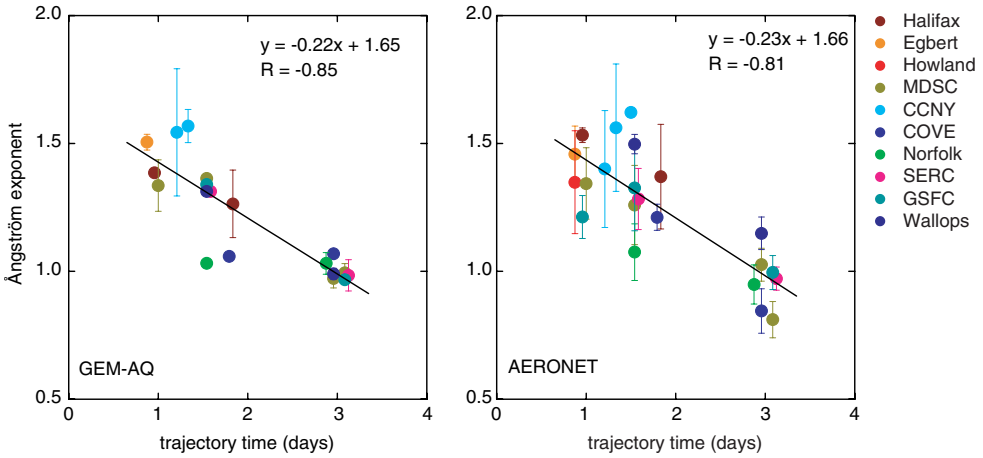

Fig. 4. Ångström exponent from GEM-AQ (left) and measurements (right) versus trajectory time (cf. O'Neill et al. (2005) for details on the computation of trajectory time). The measurements were filtered for small signals (signal less than 50 were excluded for all channels), and the modelled data (run *B* of Table 1) have been similarly filtered in terms of an analogous slant path optical depth filter.

Virginia-area states up until 9 July (i.e. with respect to the modelled data; as pointed out by O'Neill et al., 2005, the measured data on 9 July could not be unambiguously classified as smoke due to the presence of interspersed cloud). Fresher smoke from sources on 7 July is the cause of day-old smoke on 8 July at Howland and Halifax, and smoke on 8 July is the cause of 2-day old smoke on 10 July at Halifax (represented by high α or reddish/orange regions). While the statistics of Fig. 4 (for both the modelled and measured data) are clearly influenced by the clustering of points for the southern stations seen on trajectory day 3 (there are, effectively two distinct populations of points), we would argue that the essential point of a decrease in α is a significant and robust trend.

It is noted that the modelled data were filtered for large AOD in a manner which was similar to the large-AOD low-signal filtering applied to the measurements. Closer inspection showed that the modelled α temporal trend at each station consisted of high frequency diurnal variations inversely correlated with large diurnal AOD variations superimposed on the low frequency trend evident in Figs. 4 and 5 (variations which clearly result from the coagulation physics). A high frequency α modulation which varied inversely with large AOD was also seen in the raw measured AOD spectra; however this is precisely the data which required filtering due to low signal artifacts. While the modelled α values clearly do not suffer from low-signal artifacts, the measured data is susceptible to these artifacts, while at the same time being subject to coagulative effects. However, it is difficult if not impossible to discriminate between the decrease of α with low-signal non-linearities and the decrease due to large AOD coagulative growth effects. Accordingly, we judged that the only reasonable basis of comparison for the production of Fig. 4 was to filter the modelled data for large AODs just as the measured data was filtered (the result of not filtering is a degraded trajectory time trend in the modelled data which is clearly attributable to diurnal excursions in α).

We note that if one increases the reference wavelength at which α is computed (from 500 nm to a wavelength in the near infrared), fewer large AOD situations are encountered (less low-signal filtering is required). However, in this case α becomes increasingly less dependent on particle size (Reid et al., 1999; O'Neill et al., 2005) and the advantages inherent in an increase in the number of unfiltered points is offset by the decrease in particle size sensitivity. Finally, in order to ascertain that the modelled particle growth effects were due to coagulation and not hygroscopic growth, we re-computed the 12 bin Mie integrations for 0% relative humidity. The Ångström exponent versus trajectory-time graph essentially retained the same form as the modelled results in Fig. 4 (the dry run slope was -0.20 with an R^2 value of 0.62). This is not a surprising result, since the average relative humidities were small at altitudes in the neighbourhood of the smoke layer during the smoke window periods (at 770 mb the average station relative humidities were generally below 30%, with only

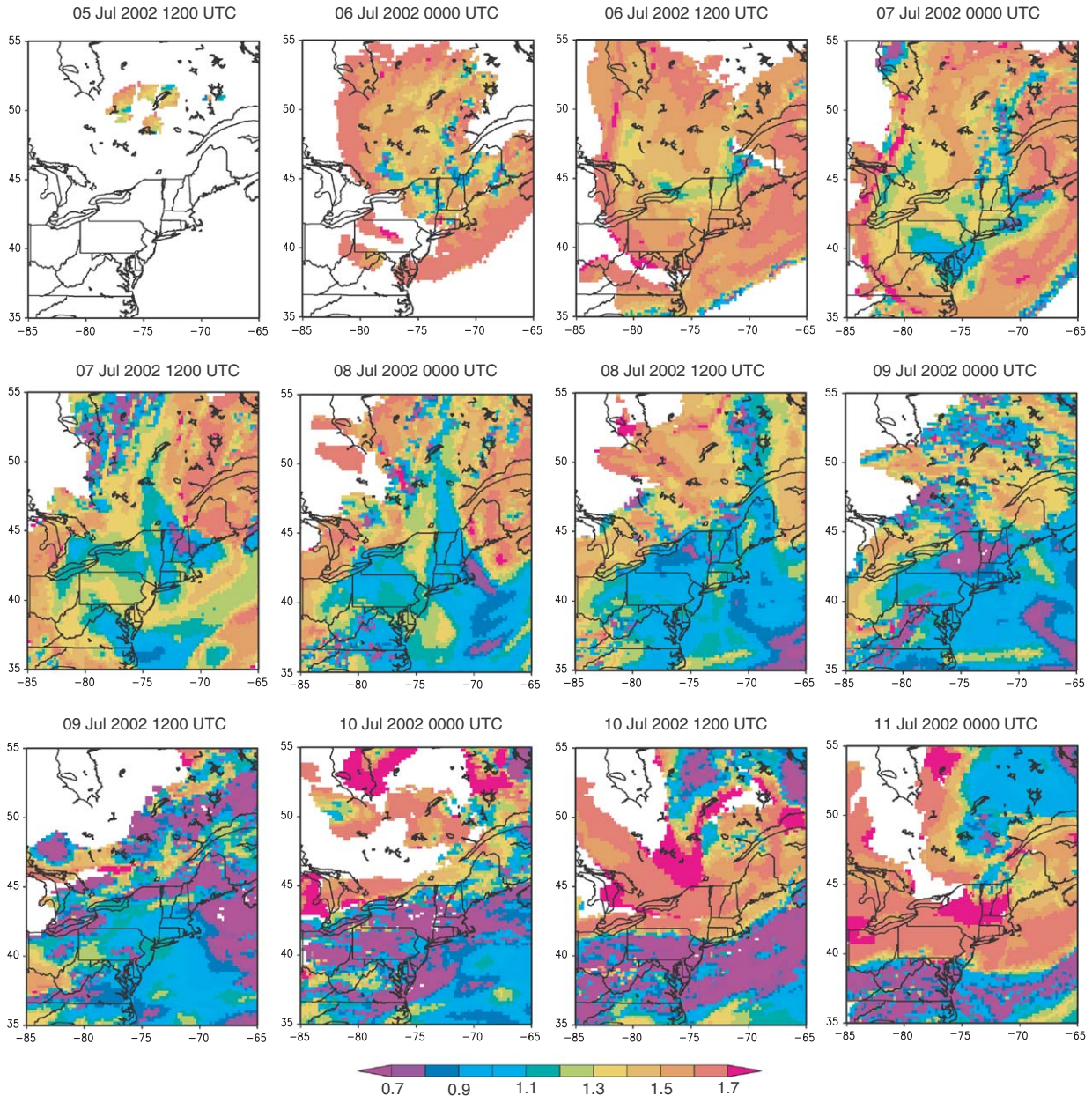

Fig. 5. GEM-AQ animation frames of α taken every 12 h.

Halifax showing values between 60% and 90%, and MDSC, CCNY and GSFC showing values greater than 70% only on 6 July). Relative humidities below 50% only induce small changes in optical extinction, and only moderate changes are induced for relative humidities below 70%; the largest humidity-induced change in α was a decrease of about 0.2 for Halifax on 8 July.

6. Conclusions

In this paper we attempted to quantify and characterize the rms differences between model and measurements for intensive and extensive optical variables (aerosol optical depth and Ångström exponent). In order to investigate regional differences between model and measurements, we

employed a three component analysis of rms differences. The behaviour of the two absolute amplitude components of the rms differences (δ_{OFF} and δ_{σ}) enabled us to infer emission properties which would otherwise have been masked by the larger anti-correlation rms difference component ($\delta_{R\star}$). The inferred emission fluxes were significantly higher than those estimated by the original FLAMBÉ emissions flux model; it was suggested that this apparent underestimation might be due to underestimated emission factors in very intense fires, or other factors not explicitly accounted for in the emissions model.

The model captured the regional decrease of the intensive Ångström exponent (increase of particle size with trajectory time), while the agreement between the model and measurements for the extensive variable of aerosol optical depth was marginal, but clearly dependent on the nature of the spatio-temporal statistical tools which were employed to characterize model performance. In establishing the modelled Ångström exponent trend, the modelled data were filtered in the same way that the measured data were filtered for non-linear instrumentation artifacts. This processing of modelled results was deemed necessary in order to render these Ångström exponent results comparable with the measurements; in the latter case it was difficult if not impossible to discriminate between measured Ångström exponent trends due to instrumental artifacts versus trends due to coagulative effects. Thus, the only choice left was effectively to process the modelled results in order to better simulate the measurements.

Acknowledgments

This work was partially supported by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS), the Meteorological Services of Canada (MSC), the National Sciences and Engineering Research Council (NSERC), and the AERONET group of NASA. Dr. Reid's participation was funded through the NASA Interdisciplinary Sciences program.

References

Berk, R.A., Fovell, R.G., Schoenberg, F., Weiss, R.E., 2001. The use of statistical tools for evaluating computer simulations—an editorial essay. Climate Change 51, 119–130.

Berk, R.A., Bickel, P., Campbell, K., Fovell, R., Keller-McNulty, S., Kelly, E., Linn, R., Park, B., Perelson, A., Rouphail, N., Sacks, J., Schoenberg, F., 2002. Workshop on Statistical Approaches for the Evaluation of Complex Computer Models. Statistical Science, vol. 17, pp. 173–192.

Colarco, P.R., Schoeberl, M.R., Doddridge, B.G., Marufu, L.T., Torres, O., Welton, E.J., 2004. Transport of smoke from Canadian forest fires to the surface near Washington, DC: Entrainment and optical properties. Journal of Geophysical Research 109, D06203, doi:10.1029/2003JD004224.

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M., Staniforth, A., 1998. The operational CMC-MRB global environmental multiscale (GEM) model. Part I: design considerations and formulation. Monthly Weather Review 126, 1373–1395.

Gong, S.L., Barrie, L.A., Blanchet, J.P., von Salzen, K., Lohmann, U., Lesins, G., Spacek, L., Zhang, L.M., Girard, E., Lin, H., Leaitch, R., Leighton, H., Chylek, P., Huang, P., 2003. Canadian aerosol module: a size-segregated simulation of atmospheric aerosol processes for climate and air quality models, 1, module development. Journal of Geophysical Research 108, 4007.

Hess, M., Koepke, P., Schult, I., 1998. Optical properties of aerosols and clouds: the software package OPAC. Bulletin of the American Meteorological Society 79, 831–844.

Lesins, G., Lohmann, U., 2005. Using MODIS and AERONET to determine GCM aerosol size. Journal of the Atmospheric Sciences, in press.

Lin, J.C., Gerbig, C., 2005. Accounting for the effect of transport errors on tracer inversions. Geophysical Research Letters 32, L01802.

Lurmann, F.W., Lloyd, A.C., Atkinson, R., 1986. A chemical mechanism for use in long-range transport/acid deposition computer modeling. Journal of Geophysical Research 91, 10,905–10,936.

O'Neill, N.T., Thulasiraman, S., Eck, T.F., Reid, J.S., 2005. Robust optical features of fine mode size distributions; application to the Québec smoke event of 2002. Journal of Geophysical Research 110, D11207.

Pahlow, M., Kleissl, J., Parlange, M.B., Ondov, J.M., Harrison, D., 2005. Atmospheric boundary-layer structure observed during a haze event due to forest-fire smoke. Boundary-Layer Meteorology 114, 53–70.

Prins, E., Schmetz, J., Flynn, L., Hillger, D., Feltz, J., 2001. Overview of current and future diurnal active fire monitoring using a suite of international geostationary satellites. In: Ahern, F.J., Goldammer, J.G., Justice, C.O. (Eds.), Global and Regional Wildfire Monitoring: Current Status and Future Plans. SPB Academic Publishing, The Hague, Netherlands, pp. 145–170.

Reid, J.S., Eck, T.F., Chrisopher, S.A., Hobbs, P.V., Holben, B.N., 1999. Use of the Ångström exponent to estimate the variability of optical and physical properties of aging smoke particles in Brazil. Journal of Geophysical Research 22, 27,473–27,489.

Reid, J.S., Prins, E.M., Westphal, D.L., Schmidt, C.C., Richardson, K.A., Christopher, S.A., Eck, T.F., Reid, E.A., Curtis, C.A., Hoffman, J.P., 2004. Real-time monitoring of South American smoke particle emissions and transport using a coupled remote sensing/box-model approach. Geophysical Research Letters 31, L06107.

Reid, J.S., Koppmann, R., Eck, T.F., Eleuterio, D.P., 2005. A review of biomass burning emissions part II: intensive physical properties of biomass burning particles. Atmospheric Chemistry and Physics 5, 799–825.

Sapkota, A., Symons, J.M., Kleissl, J., Wang, L., Parlange, M.B., Ondov, J., Breysse, P.N., Diette, G.B., Eggleston, P.A., Buckley, T.J., 2005. Impact of the 2002 Canadian forest fires on particulate matter air quality in Baltimore City. Environmental Science and Technology 39, 24–32.

Taubman, B.F., Marufu, L.T., Vant-Hull, B.L., Piety, C.A., Doddridge, B.G., Dickerson, R.R., Li, Z.Q., 2004. Smoke over haze: aircraft observations of chemical and optical properties and the effects on heating rates and stability. Journal of Geophysical Research 109, D02206.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research 106, 7183–7192.

Wandinger, U., Müller, D., Böckmann, C., Althausen, D., Matthias, V., Bösenberg, J., Weiß, V., Fiebig, M., Wendisch, M., Stohl, A., Ansmann, A., 2002. Optical and microphysical characterization of biomass-burning and industrial-pollution aerosols from multiwavelength lidar and aircraft measurements. Journal of Geophysical Research 107, 8125.

Wotawa, G., Trainer, M., 2000. The influence of Canadian forest fires on pollutant concentrations in the United States. Science 288, 324–328.