JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 102, NO. D21, PAGES 25,305–25,338, NOVEMBER 20, 1997

Evaluation of modeled sulfate and SO_2 over North America and Europe for four seasonal months in 1986-1987

Carmen M. Benkovitz and Stephen E. Schwartz

Department of Applied Science, Brookhaven National Laboratory, Upton, New York

Abstract. A three-dimensional Eulerian transport and transformation model driven by observation-derived synoptic meteorological data has been applied to calculate mixing ratios (MRs) of sulfate and SO_2 and wet deposition of sulfate over the North Atlantic and adjacent continental regions for 1-month periods in each of four seasons in 1986-1987. Model performance is evaluated by comparison of grid-cell average (1.125°) modeled MRs for sulfate (24-hour average) and SO_2 (6- and 24-hour average) in the lowest model level (surface to ~65 m) to surface MRs observed at monitoring stations in North America and Europe. For sulfate ~8000 model-observation comparisons were made employing ~10,000 individual measurements; for 24-hour SO_2, 21,000 comparisons (54,000 measurements) and for 6-hour SO_2, ~71,000 comparisons (211,000 measurements). Subgrid variation of observed MRs is inferred from the spread of multiple simultaneous measurements within individual grid cells. The median spread of the observed MRs is a factor of 1.5 for 24-hour sulfate and 2.2 for 24-hour SO_2. The median spread between observed and modeled MRs is a factor of 2.3 for sulfate and 2.1 for 24-hour SO_2, comparable to that for the observations themselves. This suggests that much of the departure between modeled and observed MRs can be attributed to subgrid spatial variation and nonrepresentative sampling of model grid cells at the stations used for the comparisons. For SO_2 the median ratio of modeled to observed MRs is 0.97, with little seasonal variation, somewhat lower in North America but considerably higher in Europe; little difference was evidenced in comparisons of 6-hour averages versus 24-hour averages. For sulfate the median ratio is 0.51, with the range for the four simulation periods 0.36 to 0.66, lowest in January-February 1987, and with comparable values for Europe and North America. For all four simulations the time series of 24-hour average modeled MRs at most locations rather closely reproduce the magnitudes and temporal episodicity of the observed sulfate and SO_2 MRs. Analysis of correlations of observed and modeled MRs was carried out for all grid cell locations for which at least 25 days of observations were available in a simulation period; 76% of 203 correlations for 24-hour sulfate and 51% of 526 correlations for 24-hour SO_2 were significant at the 95% confidence level. The superior model performance in this respect for sulfate is attributed to the lower subgrid variation in the mixing ratio of this mainly secondary atmospheric species versus the mainly primary emitted species SO_2. Comparisons of modeled and observed sulfate wet deposition (concentration times precipitation amount) for ~300 daily and ~1100 weekly samples, all in North America, indicate a median spread between modeled and observed deposition of a factor of 2.6 and a median ratio of modeled to observed deposition of 0.82. The major contributor to model underestimation of sulfate MR in air is tentatively attributed to the lack of representation in the model of the aqueous-phase conversion of SO_2 to sulfate in nonprecipitating clouds.

1. Introduction

Aerosols influence the global energy budget, and therefore global climate, through changes in the clear-sky albedo and cloud albedo (the direct and the indirect forcing, respectively) [*Charlson et al.*, 1992]. The global shortwave radiative forcing of climate by anthropogenic enhancement of aerosols is estimated to be comparable, but of opposite sign, to current global forcing by anthropogenic enhancement of greenhouse gases [*Charlson et al.*, 1991, 1992; *Wigley and Raper*, 1992; *Kiehl and Briegleb*, 1993]. Current estimates of climate forcing by aerosols are highly uncertain [*Penner et al.*, 1994; *Intergovernmental Panel on Climate Change*, 1996; *National Research Council*, 1996; *Schwartz and Andreae*, 1996]. A major contributor to this forcing is sulfate, derived mainly from atmospheric oxidation of SO_2 emitted in the combustion of fossil fuels. At present, neither data from in situ measurements [*Andreae*, 1996] nor from remote sensing by satellites [*Hansen et al.*, 1993] are adequate for assessing global aerosol loadings; consequently, estimates of the direct forcing of anthropogenic aerosols [*Kiehl and Briegleb*, 1993; *Taylor and Penner*, 1994; *Cox et al.*, 1995; *Kiehl and Rodhe*, 1995; *Mitchell et al.*, 1995; *Chuang et al.*, 1997] and parameterizations of the indirect forcing [*Leaitch et al.*, 1992; *Jones et al.*, 1994; *Boucher and Lohmann*, 1995; *Erickson et al.*, 1995; *Schwartz and Slingo*, 1996; *Chuang et al.*, 1997] have been based on atmospheric chemistry models that describe the loading and distribution of

Copyright 1997 by the American Geophysical Union.

Paper number 97JD02211.
0148-0227/97/97JD-02211$09.00

anthropogenic aerosols, particularly sulfate. This dependence on chemical modeling to estimate current atmospheric aerosol loadings, and future loadings in response to changes in emissions, makes it imperative to quantify and reduce the large uncertainties associated with these model-based estimates [*Penner et al.*, 1994; *National Research Council*, 1996; *Schwartz and Andreae*, 1996].

Evaluation of the accuracy of atmospheric chemistry models requires comparison of model results with observations of the pertinent trace species. For the most part modeling studies of aerosol sulfate loadings have been carried out either with monthly mean winds [*Langner and Rodhe*, 1991; *Langner et al.*, 1993] or with meteorological fields derived from general circulation models (GCMs) [*Taylor and Penner*, 1994; *Pham et al.*, 1995; *Chin and Jacob*, 1996; *Feichter et al.*, 1996; *Chuang et al.*, 1997; *Kasibhatla et al.*, 1997]. In either case, model accuracy is evaluated by comparison of monthly mean modeled and observed concentrations. However, for a species such as sulfate whose mean atmospheric residence time of a few days is comparable to the timescale of synoptic-scale weather systems, and whose concentration at a given location is therefore highly variable, [e.g., *Husain and Dutkiewicz*, 1990; see also *Schwartz*, 1996], much of the variation in concentrations is lost by such averaging, thereby blunting the precision with which modeled and observed concentrations can be compared. Moreover, such averaging precludes examination of the ability of the model to represent the temporal variability of the loading; this variability is certainly important in accurate representation of indirect radiative forcing (in view of the nonlinear dependence of this forcing on aerosol concentrations and possible correlation with other synoptic variables) and perhaps also the direct forcing.

An alternative approach, employed here, is to drive the chemical model by observation-derived synoptic meteorological data. Such an approach allows the model to capture the temporal variation of the trace species, which is not represented when models are driven by time-averaged meteorological data, thereby permitting numerous and highly detailed comparisons of model results to observations at specific times, as opposed to multiyear monthly averages, as has been conventional. Such comparisons are likely to be much more indicative of model performance in estimating mean loadings, as the variation in concentrations due to the intrinsic meteorological variability can, to a great extent, be accounted for in the model. Additionally, this approach permits representation of nonlinear chemical processes, which cannot be accurately represented using time-averaged concentration fields. This approach has been taken in evaluations of regional-scale models [*Dennis et al.*, 1993; *Macdonald et al.*, 1993; *Hass et al.*, 1993] but has not seen widespread application, especially at larger geographical scales. *Dennis et al.* [1993] compared modeled sulfate surface concentrations with 24-hour measurements made over a 1-month period on spatial scales of ~500 km, finding rather convincing representation of the temporal variation of their modeled sulfate concentrations, but substantial model underestimate that required upward adjustment either of the primary sulfate emissions or of the oxidation rate, the latter compensating for inadequate representation of reactions in nonprecipitating clouds. *Macdonald et al.* [1993] have presented comparisons of model calculations with point measurements by aircraft sampling, finding comparable mixing ratios in some instances, but substantial departure in other instances, which they ascribed to inaccurate representation of precipitation in their model.

Despite the inherent advantages in driving a model with observation-derived synoptic data, several concerns must be addressed when comparing such model calculations with observations. Inaccuracies in the meteorological data governing transport direction and location and time of deposition events inevitably lead to mismatch between the model and observations, ultimately necessitating statistical comparisons, such as correlations between modeled and observed mixing ratios (MRs). Another concern is that the models represent averages over the horizontal and vertical dimensions, whereas most anthropogenic sulfur emissions are released from point sources and observations represent values at individual points. Consequently, it is inevitable that especially in source regions there will be substantial subgrid variation in MRs, which can result in departures between model and point observations. This variability not only causes departures from the model value, but may also give rise to bias if, for example, an observation site is located more or less proximate to sources within the grid cell or if there are persistent nonuniformities resulting from terrain features. Thus even with a model that is driven by observation-derived synoptic data, a one-to-one correspondence between model and observations cannot be expected.

We have previously described an Eulerian transport and transformation model for evaluating the spatial and temporal patterns of MRs of atmospheric sulfate aerosol using observation-derived synoptic meteorological data [*Benkovitz et al.*, 1994, hereinafter referred to as B94]. That paper presented time series of modeled and observed 24-hour average MRs for sulfate at 43 locations in North America and Europe for a 30 day period in October-November 1986. Comparisons of these time series indicated that although in general modeled MRs were somewhat lower than observed MRs, the model accurately represented the geographical distribution and temporal episodicity of atmospheric sulfate loadings, lending confidence in the performance of the model. The model calculations indicated that sulfate loading over the mid North Atlantic, well away from source regions, exhibits variation on short spacescale (~1000 km) and short timescale (~12 hours) that is a consequence of the synoptic-scale variability in the meteorological variables (wind speed and direction, precipitation, etc.) that drive the model. The high temporal and spatial variability in sulfate loading indicated in the model calculations supports the need to drive the model by observation-derived synoptic data to permit meaningful comparison with observations.

We have now carried out three additional 1-month simulations, (comprising, with the original simulation, one simulation in each season) to extend the range of conditions over which the model is exercised and to examine seasonal variation of the sulfate and SO_2 MRs and sulfate wet deposition. In this paper we present a variety of comparisons of model results with observations to evaluate the performance of the model over a range of meteorological conditions. Comparisons include differences and ratios between measured and modeled MRs and time series over the several ~30-day simulation periods. Departures between modeled and observed MRs are compared to the magnitude of subgrid variation inferred from multiple simultaneous observations within individual model grid cells.

2. Description of the Model and the Input Data

The model used in this study, Global Chemistry Model driven by Observation-derived meteorological data (GChM-O), is a

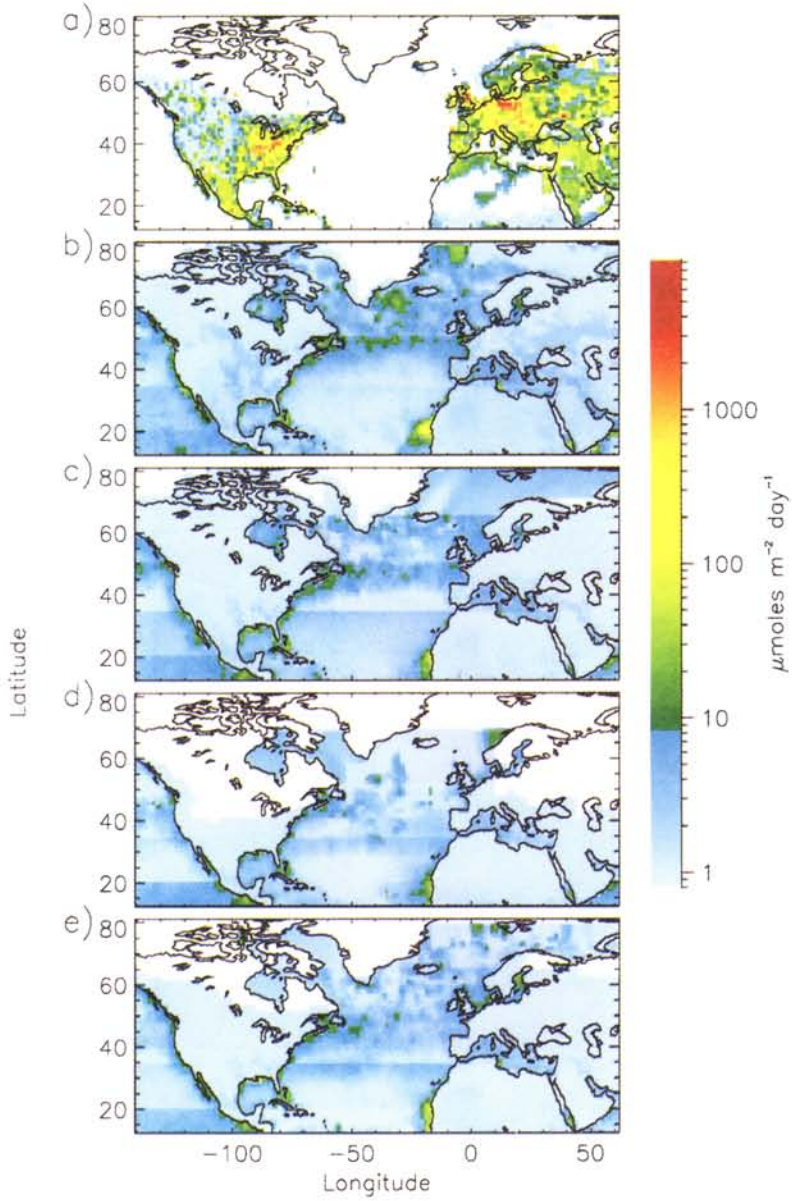

Plate 1. Geographic distribution of sulfur emission fluxes. (a) Anthropogenic emissions of sulfur (SO_2 plus primary sulfate) in the basic emission inventory. European emissions were taken as 33% higher in winter and 33% lower in summer; elsewhere, emissions were taken as seasonally constant. Biogenic sulfur (DMS plus H_2S) emissions. (b) June-July, (c) October-November, (d) January-February, and *e*) March-April. The latitude band structure of the oceanic DMS emissions given by *Bates et al.* [1992] is still apparent after distribution proportional to CZCS data and reallocation t o the model grid, introducing discontinuities in these emissions. Note logarithmic scale; white denotes areas with no emissions.

three-dimensional Eulerian transport and transformation model for sulfate. The physical and chemical mechanisms used in the model have been previously described by B94 and by *Luecken et al.* [1991] and are only briefly described here, with emphasis on features pertinent to the description of seasonal variation. The model represents emissions of anthropogenic sulfur dioxide (SO_2) and sulfate and of biogenic sulfur species, horizontal and vertical transport, gas-phase oxidation of SO_2 and dimethylsulfide (DMS), aqueous-phase oxidation of SO_2 by hydrogen peroxide (H_2O_2) and ozone (O_3) in precipitating clouds, and wet and dry deposition of SO_2, sulfate, and methanesulfonic acid (MSA). The model is driven by observation-derived synoptic data obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF), specifically the 6-hour "first guess" data (uninitialized 6-hour forecast) [*ECMWF*, 1988], to allow comparison with observations at specific times and locations. The horizontal coordinates used are latitude and longitude with 1.125° resolution (approximately 122 × 125 km at 12°N; 62.5 × 125 km at 60°N); the vertical coordinate is the ECMWF "η" (eta) coordinate, a hybrid system of pressure and terrain-following sigma coordinates, with 15 levels between the surface to 100 hPa. The model solves material balances based on the continuity equations with the application of gradient-transport assumptions; numerical approximations at any location are considered mechanistically and dimensionally independent over

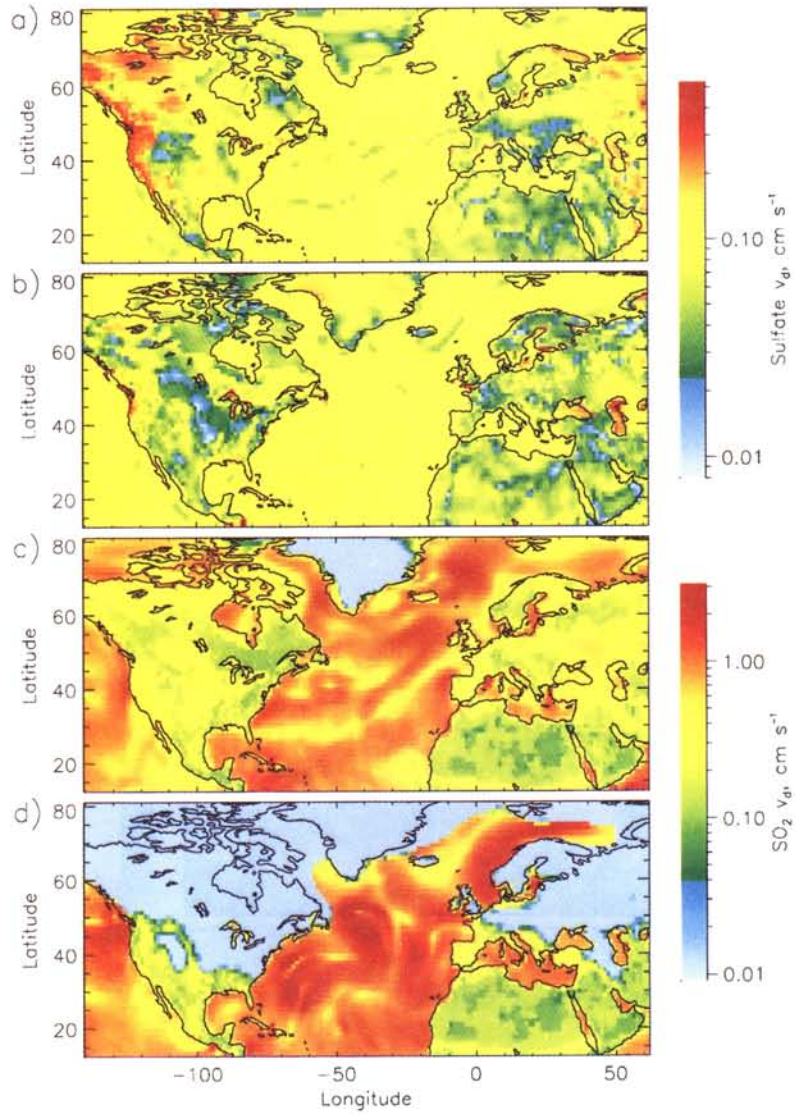

Plate 2. Geographic distribution of deposition velocity. (a) Sulfate, July 14, 1986, 0000 UT. (b) Sulfate, January 28, 1987, 0000 UT. (c) SO_2, July 14, 1986, 0000 UT. (d) SO_2, January 28, 1987, 0000 UT.

short time steps according to the operator splitting technique [*Yanenko*, 1971]. Spatial integrals of the transport are approximated using a modification of the area-preserving flux form developed by *Bott* [1989] as modified by *Easter* [1993]. The vertical eddy diffusivity coefficient was taken as 0.25 $\times 10^5$ cm^2 s^{-1} for the lowest model level (surface to ~65 m), 1 $\times 10^5$ cm^2 s^{-1} [*Pasquill*, 1976] for model levels 2 to 4, and zero for all other model levels. The time step is 0.5 hour for transport and 1 hour for chemical reactions and wet removal. The model output consists of mixing ratios of SO_2 and sulfate at each grid cell location and model level, at 6-hour time intervals. In order to permit interpretation of sources of sulfate, the sulfate variable is distinguished as to origin of material (biogenic or anthropogenic), region of origin (North American, west of 30°W; and European, east of 30°W), and formation mechanism (primary; and secondary by gas-phase or aqueous-phase oxidation).

The model domain extends from ~140°W to 62.5°E and 12.5°N to 81°N, as shown in Plate 1. Because this domain is not global, it is necessary to assume mixing ratios of species that are advected into the model domain. We set these MRs to representative values as presented in Figure 1. As discussed by B94, these "external" contributions to loadings of sulfur species are relatively minor in regions within the model domain that are heavily influenced by anthropogenic emissions but represent an increasingly large contribution to the modeled loading of sulfur species near boundaries of the domain, especially in regions that are usually upwind of anthropogenic sources, that is, toward the western side of the domain.

Model simulations were carried out for the following periods: June 15 to July 31, 1986 (JJ86), October 1 to November 15, 1986 (ON86), January 15 to February 28, 1987 (JF87), and March 15 to April 30, 1987 (MA87). The October-November 1986 simulation, previously reported by B94, is included in the present work to allow comparisons with additional observations identified since B94, to extend the analyses to SO_2, and to permit seasonal comparisons. All simulations were initiated with mixing ratios of sulfur and oxidant species in the model

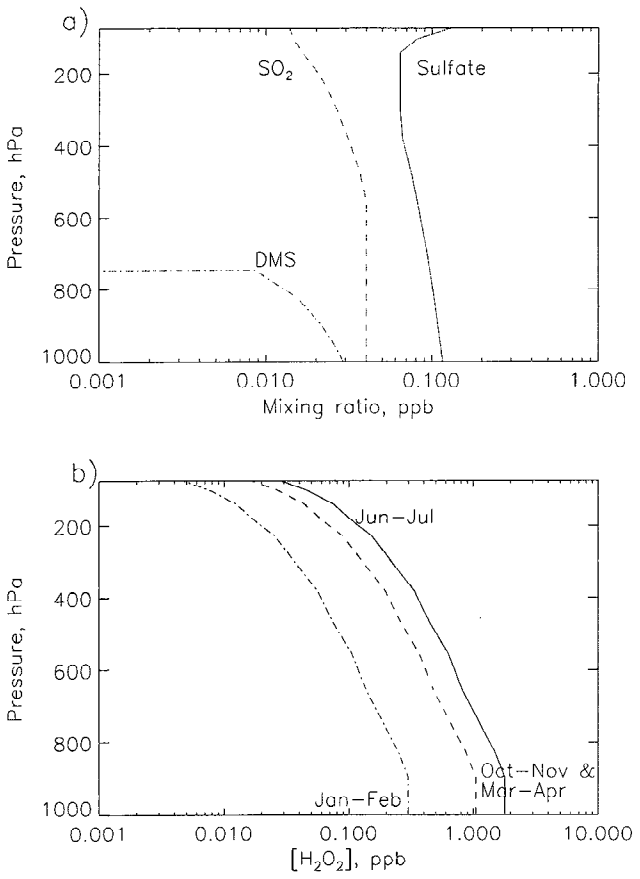

Figure 1. Mixing ratios assumed in the model calculations. (a) Mixing ratios of SO_2, sulfate, and DMS external to the model domain. (b) Maximum mixing ratio of H_2O_2 assumed for the four simulations.

domain set to zero. The first 2 weeks of each simulation are considered model start-up time and are not used in comparisons with observations.

2.1. Emissions

Emissions of sulfur species in the model were taken to be representative of summer, fall, winter, and spring seasons rather than actual emissions for the specific simulation periods. To account for the height of release from tall stacks, emissions from large point sources are introduced at the lowest three model levels (approximate heights above terrain 0 to 65, 65 to 230, and 230 to 540 m). The model immediately distributes emissions uniformly throughout the volume of the grid cell into which the material is emitted. Emissions were based on the inventories compiled for the ON86 simulation, described by B94 and denoted here as the basic GChM-O inventory. For the present simulations the seasonal dependence of emissions was specified as described below. No effort was made to identify departures from these inventories that might arise, for example, from a particular point source being out of operation during the simulation period.

2.1.1. Anthropogenic emissions. Emissions for North America in the basic GChM-O inventory were taken from the National Acid Precipitation Assessment Program (NAPAP) 1985 version 2 inventory [*Saeger et al.*, 1989]. According to this inventory there is very little seasonal variation (91.3% of the emissions varied less than 1% between seasons, the remaining varied less than 4% between seasons), so in this work, North American emissions were taken as equal for all four simulation periods. Emissions for Europe in the basic GChM-O inventory are the European Modeling and Evaluation Programme (EMEP) annual-average emissions [*Iversen et al.*, 1990]. To account for seasonal variation in European emissions, a sinusoidal variation was imposed on the basic values to calculate emissions in January (× 1.33) and in July (× 0.67) [*Berge*, 1993]. Primary sulfate emissions in the United States and Canada, obtained from the NAPAP inventory, constitute 1.4% by mole of the total sulfur emissions. For Europe, primary emissions were taken as 5% of total sulfur emissions [*Eliassen*, 1978; *Eliassen and Saltbones*, 1983]. Although we consider 5% primary sulfate to be an overestimate [*Dietz and Wieser*, 1983], we have retained this value to preserve the homogeneity across the four simulation periods. Outside of Europe, anthropogenic emissions in the basic GChM-O inventory were taken as constant for the four simulations. The geographic distribution of the basic anthropogenic emissions, used for the ON86 and MA87 simulations, is presented in Plate 1. The distribution is highly nonuniform spatially, with the greatest density of emissions in the industrialized areas of Europe and North America. The emissions exhibit considerable short range variation, indicative of the highly structured pattern of emissions, even when aggregated to the 1.125° × 1.125° grid.

2.1.2. Biogenic emissions. Marine biogenic emissions of DMS in the basic GChM-O inventory were derived from the work of *Bates et al.* [1992] distributed to a 1° × 1° grid based on global monthly gridded pigment concentrations derived from Coastal Zone Color Scanner (CZCS) images. Global monthly composite pigment concentrations derived from CZCS images [*Feldman et al.*, 1989], obtained from the National Aeronautics and Space Administration Goddard Space Flight Center National Space Science and Data Center, were used to distribute the oceanic emissions. The CZCS data, which are provided on a 512 km × 512 km grid, were initially transformed to a 1° × 1° latitude/longitude grid. Missing data were estimated iteratively from the surrounding valid data by a valid-neighbors weighted algorithm. The July, February, and April composites were used to distribute the summer, winter, and spring oceanic DMS emissions to a 1° × 1° grid, and this grid was converted to the GChM-O 1.125° × 1.125° grid as described by B94.

Biogenic emissions of sulfur compounds from terrestrial sources were calculated by month on a 1° × 1° grid using the methodology of Lamb [*Bates et al.*, 1992] described by B94. Emissions were estimated based on land use [*Wilson and Henderson-Sellers*, 1985], average seasonal temperature [*Shea*, 1986], and seasonal biomass conversion factors [*Zimmerman*, 1979]. As DMS emissions are ~88% of the total biogenic emissions (B94), emissions of DMS and H_2S from terrestrial biogenic sources were treated entirely as DMS. The geographic distributions of biogenic emissions in the model domain for the four simulations are presented in Plate 1. The highest emissions are located over the high productivity coastal areas, with emissions from land areas being lower than oceanic emissions. For the winter period there are no emissions over land areas in the midlatitudes. In all seasons these emissions are considerably lower and more uniformly distributed than the anthropogenic emissions.

Emissions from volcanoes were not included in the simulations; for the modeling domain of this study, volcanic emissions are estimated as approximately 11% of biogenic

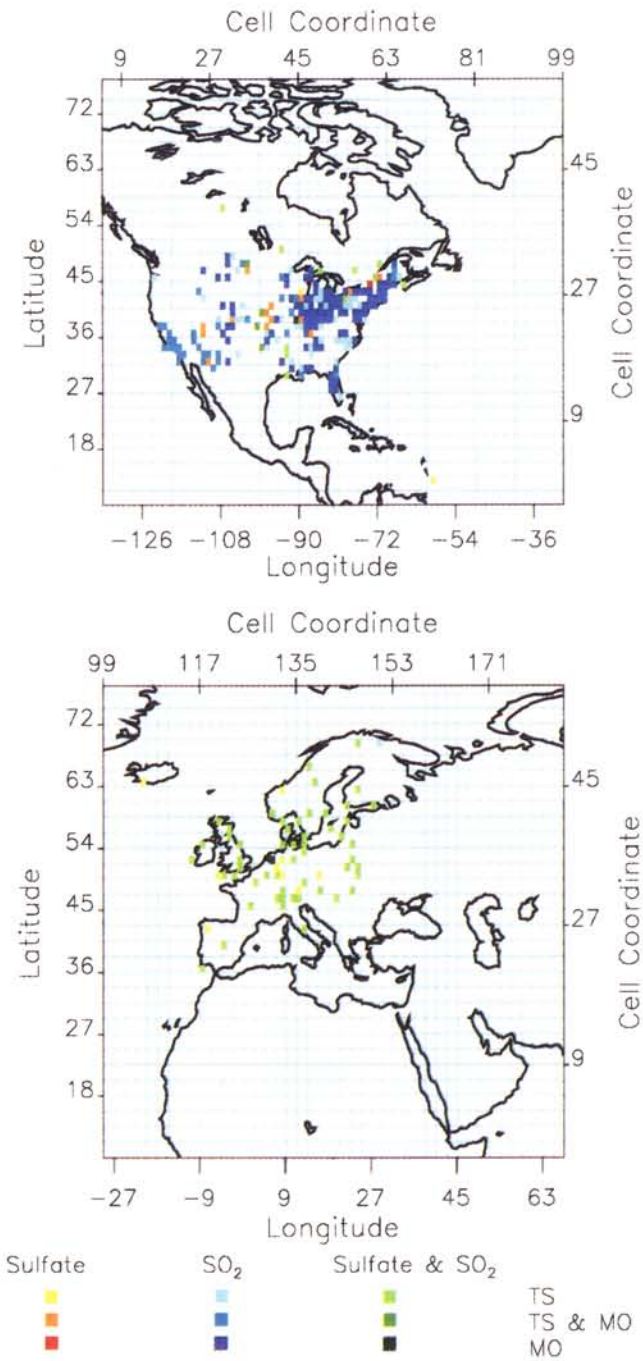

Plate 3. Geographic distribution of locations of comparisons of observed and modeled mixing ratios. Scales at top and right show indices (x, y) used to identify cell locations. The algorithm relating the longitude of the center of a grid cell to grid cell index x is longitude = $-140.625 + 1.125(x+0.5)$; likewise latitude = $12.375 + 1.125(y+0.5)$. The colors indicate whether observations at a given location are for SO_2, sulfate, or both, and whether the observations consist of time series (TS), multiple observations (MO), or both.

emissions and approximately 0.6% of anthropogenic emissions in the basic GChM-O inventory (B94).

2.2. Dry Deposition

As described by B94, time- and location-dependent dry deposition velocities of SO_2 and accumulation mode particulate sulfate were calculated from surface resistances, atmospheric stability, and surface wind speeds as described by *Wesely* [1989]. Plate 2 presents examples of the geographic distribution of SO_2 dry deposition velocity for two specific times, July 14, 1986, at 0000 universal time (UT) and January 28, 1987, at 0000 UT. The differences in the deposition velocities for these two dates are mainly at midlatitudes, where the majority of emissions sources are located. In JF87 the dry deposition velocities over land areas of North America and Europe are very low because of the presence of ice or snow or because surface temperatures are below -4°C, conditions that inhibit deposition of SO_2 resulting from aqueous dissolution and reaction. The structure of the dry deposition velocity fields over the ocean is governed by the structure of the wind fields.

Histograms of the deposition velocities for SO_2 and sulfate weighted by the modeled MRs at the lowest model level (surface to approximately 65 m) are presented in Figure 2, and statistics are presented in Table 1. Rather slight differences are exhibited for the several simulations, except that the SO_2 deposition velocity over land areas is considerably less for the JF87 simulation than for the other three simulations. For this period, surface temperatures in midlatitudes were frequently below -4°C; at these temperatures the rate of SO_2 deposition to land surfaces is specified as 0.01 cm s^{-1}, as reflected in the high peak at this value in Figure 2a. For ocean areas the maximum deposition velocity of 0.1 cm s^{-1} reflects the low diffusion coefficients of accumulation mode aerosol particles. Such low deposition velocities lead to this process being a minor sink for sulfate, relative to wet deposition [*Slinn*, 1983].

2.3. Reactions and Oxidant Concentrations

Gas-phase oxidation of SO_2 and DMS is represented as a second-order reaction using diurnal-average OH concentrations. As described by B94, OH concentrations were obtained from the photochemical model calculations of *Spivakovsky et al.* [1990]. OH concentrations were calculated as the 24-hour average every 5 days for 1 year using meteorological conditions derived from a GCM and climatological averages for the chemical species. Thus the OH concentrations obtained are climatologically representative rather than representative of specific dates and times. Time series plots of the OH concentration at two geographic locations are presented in Figure 3. There is a great decrease in the OH concentrations between JJ and JF. The concentrations are fairly constant in JJ and JF but decrease gradually in ON and likewise increase in MA. These changes are driven mainly by the length of daylight and by the variation of the solar zenith angle. OH concentrations over the high emissions areas of Europe are generally somewhat lower than over the high emissions areas of North America because of the higher latitude.

Aqueous-phase reactions were represented only in precipitating clouds, as the liquid water content of nonprecipitating clouds cannot be derived from the available ECMWF data. The vertical distribution of liquid water content for precipitating clouds was apportioned as described by B94. Oxidation of SO_2 by H_2O_2 was limited by the lesser of either the SO_2 or the H_2O_2 concentration. H_2O_2 was advected in the same way as the sulfur species, depleted by reaction with SO_2, and regenerated in the gas phase at a fixed rate until a seasonally dependent maximum MR (presented in Table 2) was reached. Aqueous-phase oxidation of SO_2 by O_3 was represented by a pseudo-first-order reaction of gaseous SO_2 with a rate constant that takes into account the solubility and dissociation of gaseous SO_2 in water and the cloud liquid water content. The O_3 MRs

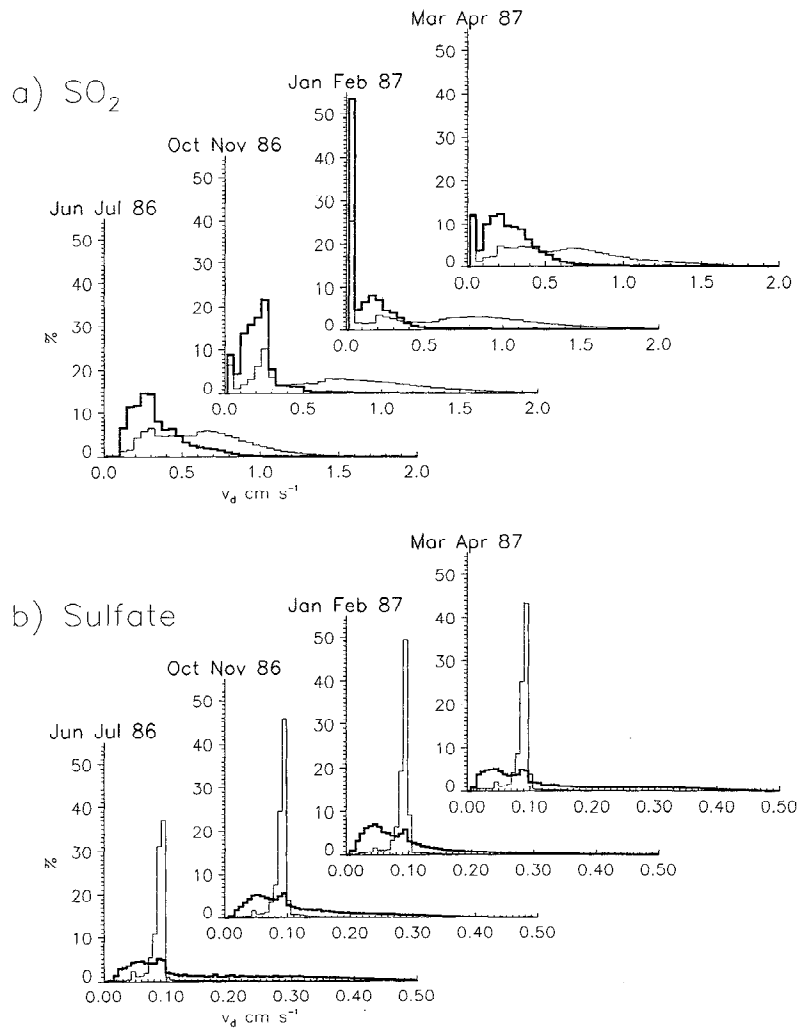

Figure 2. Histograms of deposition velocities for (a) SO_2 and (b) sulfate, weighted by the modeled mixing ratio in the lowest model level (extending to approximately 65 m above terrain) for the four simulation periods, June-July 1986, October-November 1986, January-February 1987, and March-April 1987. Thick lines represent values for land; thin lines represent values for ocean. The high peaks at the low end of the distributions for SO_2 for October-November 1986, January-February 1987, and March-April 1987 reflect low deposition velocities over ice and snow and when the surface temperature is less than -4°C, and the imposition of a minimum deposition velocity for stable conditions, all situations prevalent over the northern areas of the modeling domain and in midlatitudes during winter.

and the H_2O_2 generation rates and maximum MRs are presented in Table 2 and Figure 1b. Cloudwater *p*H, governing the rate of SO_2 oxidation by O_3, was taken as 4.5 in all simulations.

Pertinent to the comparison of modeled and observed MRs is the importance of the aqueous-phase oxidation of SO_2 to sulfate and primary sulfate MRs. To gain a sense of possible underestimate of sulfate production due to lack of representation of reaction in nonprecipitating clouds and its seasonal dependence, we note that the fraction of the area in the modeling domain covered by nonprecipitating clouds is 27-35% for the JJ86 simulation, 28-43% for ON86, 36-48% for JF87, and 34-47% for MA87. Thus the importance of the missing aqueous-phase oxidation of SO_2 to sulfate may be greater for the JF87 simulation because of the greater fraction of the modeling domain covered by nonprecipitating clouds together with lower rate of gas-phase oxidation of SO_2, although, to be sure, H_2O_2 mixing ratios are also lower. The fraction of the area in the modeling domain where primary sulfate contributes 50% or more of the surface MR of sulfate ranges from 6 to 9% for the JJ86 simulation, mainly in small sections of northern and northeast Europe, along the Pacific coast and in small sections of northern North America. For the ON86 and MA87 simulations the fractional area is 12-22% and 13-17%, respectively, in somewhat larger areas of central and northern Europe and northern and western North America. For the JF87 simulation the fractional area is 27 to 45%, covering most of Europe and larger sections of northern and western North America. As noted above, the European values are likely to be overestimates in view of the high primary sulfate emission factor assumed.

3. Observed Surface Mixing Ratios and Wet Deposition Amounts

The measurements used in the comparisons with model results were made by a number of surface monitoring networks having somewhat different measuring protocols, as described in Appendix A. Reported concentrations of sulfate and SO_2 in air were converted from the original units to mixing ratios in units

Table 1. Statistics of the Dry Deposition Velocities Weighted by the Mixing Ratio in the Lowest Model Level (Surface to ~65 m) for Land and Ocean Areas for the Four Simulations

Species	June-July 1986		October-November 1986		January-February 1987		March-April 1987	
Percentile	Land	Ocean	Land	Ocean	Land	Ocean	Land	Ocean
SO_2								
10	0.13	0.23	0.04	0.12	0.01	0.01	0.03	0.03
50	0.27	0.59	0.18	0.62	0.03	0.53	0.21	0.52
90	0.57	0.99	0.35	1.31	0.32	1.27	0.46	1.10
Sulfate								
10	0.04	0.07	0.03	0.08	0.03	0.07	0.03	0.07
50	0.11	0.09	0.09	0.09	0.07	0.09	0.09	0.09
90	0.36	0.09	0.26	0.10	0.21	0.10	0.32	0.09

Dry deposition velocities are in cm s^{-1}.

of parts per billion (ppb) (nmol per mol air); MRs reported at or below the stated limit of detection (LOD) were assigned the LOD value. Some stations did not report a value for the LOD, so multiple instances of identical minimum values were interpreted as the LOD for the station. The observed wet deposition of sulfate was calculated as the product of the sulfate concentration in precipitation and the precipitation amount and converted to mol S m^{-2}. Sampling times, which differed among the several programs, were all converted to universal time to permit comparison with model results. The spatial area used for the comparison of modeled and observed MRs is the model grid cell (1.125° × 1.125°), which we refer to as a location. Locations of the comparisons of sulfate and SO_2 MRs are shown in Plate 3.

As described in the following sections, we compared modeled and observed MRs for sulfate (24-hour observations) and SO_2 (1-hour observations averaged to 6 and 24 hours, or 24-hour observations). Where simultaneous observations were available at multiple stations within a location, these were averaged to obtain the observed MR. Table 3 summarizes the numbers of locations, measurement stations, observations, and comparisons between observations and model according to simulation period, network, and compared species. The difference between the number of observations and the number of comparisons is due to averaging multiple simultaneous observed MRs within a location to yield a single comparison. Roughly equal numbers of comparisons were available for each of the four simulation periods.

For sulfate a total of 9,978 observed MRs were available, yielding 8,057 model-observation comparisons. Because of greater station density and sampling frequency, the majority (63 to 73%) of the observed comparisons for the four simulation periods were from the EMEP measurement program in Europe; 10 to 14% were from Canada, 8.3 to 11% were from National Air Monitoring Stations (NAMS)/State and Local Air Monitoring Stations (SLAMS) (United States), and fewer than 5% were from other networks. The fraction of comparisons by region, simulation period, and observed MR is displayed in Figure 4a. Overall, more than 55% of the observed MRs were

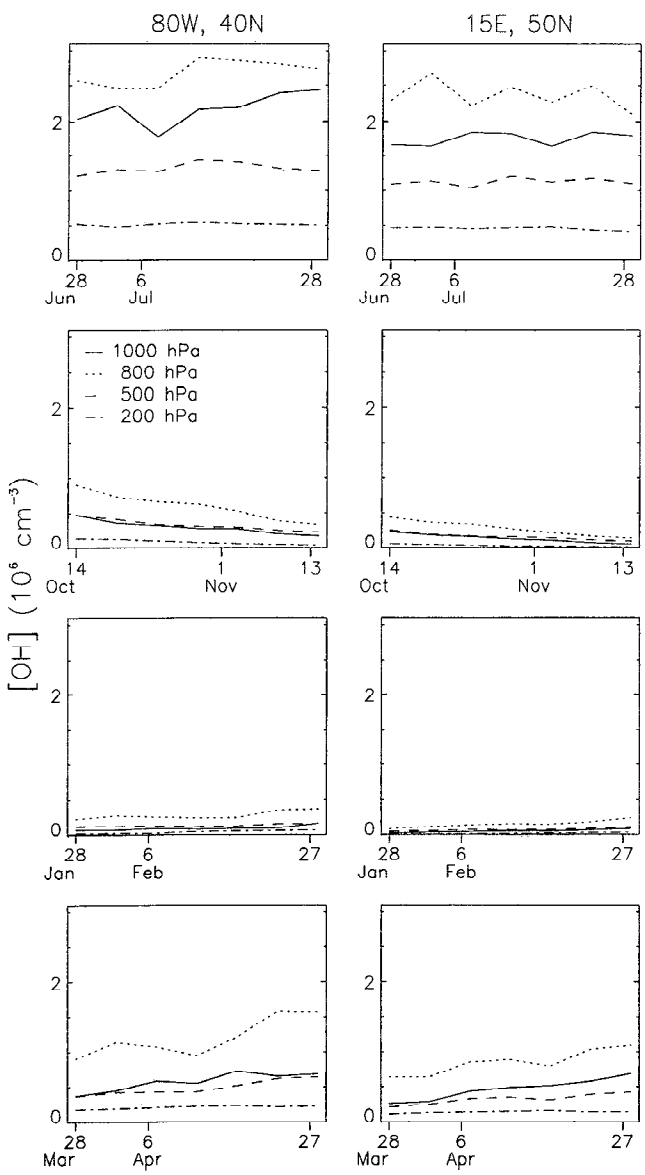

Figure 3. Time series of 24-hour average OH concentration for the four simulations at 40°N, 80°W (southwestern Pennsylvania) and 50°N, 15°E (Czech Republic). Both locations are in areas of high SO_2 emissions.

Table 2. Values of the Parameters Used in the Calculation of Aqueous Phase Oxidation of SO_2 by H_2O_2 and O_3

	H_2O_2 maximum MR, ppb	H_2O_2 generation rate, ppb h^{-1}	O_3 MR, ppb	Rate constant for O_3 oxidation, 10^{-3} s^{-1}
June-July 1986	2.4	0.040	50.0	2.24
October-November 1986	1.4	0.021	37.5	1.68
January-February 1987	0.4	0.002	25.0	1.12
March-April 1987	1.4	0.021	37.5	1.68

less than 1 ppb, and fewer than 4% were greater than 5 ppb. As might be anticipated, the MRs are not uniformly distributed, with stations in Canada and Barbados exhibiting a greater proportion of low MRs.

The SO_2 24-hour data used in the comparisons were 24-hour averages (0000 UT to 2359 UT) of the observed 1-hour MRs at the U.S. NAMS/SLAMS stations and at the Glen Canyon, Arizona, station of SCENES, and the 24-hour observed MRs at stations of the CAPMON and EMEP networks. A total of 54,644 observations were available, providing 21,647 comparisons between modeled and observed MRs. The great majority (almost 90%) of the observations in all the simulation periods were from the NAMS/SLAMS network. Figure 4b presents the distribution of the observed MRs by region, simulation period, and mixing ratio. Note the preponderance of higher MRs from U.S. stations, reflecting their mainly urban and suburban locations; fewer than 1% of the MRs observed at these stations were below 1 ppb, and over 50% of the observed MRs were between 5 and 30 ppb. Stations in Canada are located mainly in areas not usually impacted by proximate emission sources; 51% (JF87) to 82% (JJ86) of the observed MRs were less than 1 ppb. Over 50% of the observed MRs at European stations were less than 5 ppb; observed MRs exceeding 5 ppb occurred mainly in JF87.

The SO_2 6-hour data were evaluated as the 6-hour averages (0000-0559 UT, 0600-1159 UT, 1200-1759 UT and 1800-2359 UT) of the 1-hour observed MRs at the NAMS/SLAM stations and the SCENES Glen Canyon, Arizona, station. A total of 211,471 observations were available, providing 70,941 comparisons between modeled and observed MRs; the high ratio of observations to locations is indicative that many of the observations were made in source regions with a high density of observations. Over 40% of the observed MRs for JJ86 and MA87 were between 1 and 5 ppb, and approximately 40% of the observed MRs for JF87 were between 10 and 30 ppb. For ON86, approximately 33% of the observed MRs were between 1 and 5 ppb, and another 33% were between 10 and 30 ppb. The much lower proportion of the MRs less than 1 ppb compared to the 24-hour values (Figures 4b and 4c) is indicative of preferential location of these stations in regions influenced by proximate sources.

4. Subgrid Variation of Observed Mixing Ratios

The premise of the present study is that model skill can be evaluated by comparison of modeled and observed surface MRs. A concern with such comparisons as a measure of model skill is that because of subgrid temporal and spatial variation, the MRs observed at a given station within a location are not necessarily representative of the grid cell as a whole and are therefore not wholly suitable for evaluation of the model, which is meant to represent the average MRs over the grid cell. Thus departure of modeled and observed MRs may not be attributable entirely to inaccuracy of the model but may be due (in addition to any measurement error) to nonrepresentativeness of the observed MRs arising from subgrid variation, and in particular from the influence of proximate sources. This concern arises especially when using observations taken at sites which were not chosen with the objective of being representative of the grid cell as a whole, but which, as is the case of the NAMS/SLAMS observations used in the present study, were taken mainly at stations chosen for the purpose of evaluating population exposure or compliance with air quality regulations. For this reason we first examine the magnitude of subgrid variation at locations with simultaneous observations at multiple stations, with the objective of inferring the subgrid variation that might be expected in the data set as a whole. The magnitude of the variation at the examined locations qualifies the agreement that may be expected between modeled and observed MRs in the data set as a whole.

Simultaneous observations from multiple stations within individual locations were available for 7 to 12% of the comparisons for sulfate 24-hour MRs, depending on simulation period, and for 40 to 45% of the comparisons for SO_2 24-hour MRs. With the exception of a single location in the Netherlands, which included two stations, the locations for which multiple simultaneous observations were available included only stations from the U.S. Environmental Protection Agency (EPA) NAMS/SLAMS network, and as noted above, are concentrated near source regions.

Examples of data illustrating subgrid spatial and temporal variation in observed MRs are given in Plate 4, which presents time series plots of observed SO_2 1-hour MRs from March 28 to April 30, 1987, for three of nine continuous NAMS/SLAMS monitoring stations in model location (58, 24), in the vicinity of Camden, New Jersey, and adjacent areas of Pennsylvania. The very high peaks are ascribed to the transient impact of plumes from proximate sources. Consecutive 1-hour MRs occasionally differ by an order of magnitude or more, for example, on March 31 (Plate 4a), and the high peaks are not concurrent at all stations (Plate 4b). It is not the intent of this modeling effort to capture this temporal and spatial variation, which can be represented only in transport and transformation models with spatial and temporal resolutions much greater than employed here and which are not feasible in hemispheric- to global-scale models. Rather, it is the intent here only to represent average values; the difficulty is in using point observations for model evaluation. To further illustrate this point, we present time series plots of the 6-hour (red) and the 24-hour (green) averages of the observed MRs from all stations in this location (Plate 4c). These plots show a substantial damping of the high frequency temporal and short range spatial variation of the observed MRs illustrated in Plate 4a.

Table 3. Number of Locations and Comparisons of Modeled and Observed Mixing Ratios by Measurement Network

Species Network	Locations	Comparisons	Stations	Observations
Sulfate 24-hour				
CAN	8	955	8	955
NYH	2	175	2	175
USH	5	317	5	317
USE	42	820	139	2,437
USS	5	160	6	195
BAR	1	121	1	121
EUE	62	5,509	63	5,578
Total	124	8,057	209	9,978
SO_2 24-hour				
CAN	8	736	8	736
EUE	52	5,216	53	5,329
USE	155	15,609	538	48,493
USS	1	86	1	86
Total	214	21,647	598	54,644
SO_2 6-hour				
USE	167	70,560	590	211,090
USS	1	381	1	381
Total	168	70,941	591	211,471

Number of observations is the total number of observations. Measurement networks are as follows: CAN, CAPMON network (Canada); NYH, New York State Department of Health (US); USH, Harvard six-cities study (US); USE, U.S. EPA National Air Monitoring (NAMS) and the State and Local Air Monitoring Stations (SLAMS) networks (US); USS, Subregional Cooperative Electric Utility, Department of Defense, National Park Service, and EPA Study (SCENES, US); BAR, AEROCE station in Barbados; EUE, EMEP network (Europe).

To further illustrate subgrid spatial variation, we show in Plate 5 examples of time series plots of observed SO_2 24-hour average MRs at multistation locations. These plots show the MRs at the individual stations (points, represented by a single color per station) and the average of the several observed MRs (red "cityscape"). (Because of different operating schedules at individual stations, averages at a given location do not always consist of observations from the same number or distribution of stations.) In any panel the vertical spread of different-color points at each date represents the spatial variation of the observed 24-hour MRs; the spread of same-color points across multiple dates represents the day-to-day variation of the observed 24-hour MRs. Note the frequent considerable spread of the observations about the daily mean, indicative of the influence of proximate sources. Plate 5 also presents plots of the observed sulfate 24-hour MRs (points) and the average of the observed MRs (red cityscape) at the single location (north central Netherlands) with multiple stations and with sufficiently frequent observations to give meaningful time series plots. The within-location spatial variation of concurrent observed sulfate MRs was generally less that for the SO_2 MRs at the same location, the maximum spread being about a factor of 2.

As quantitative measures of the spatial variation, of multiple observed mixing ratios, MR_o, within a given location at a given time, we introduce two quantities and evaluate them at locations for which multiple simultaneous observations are available. The first, which we denote as the difference characteristic spread, $D_{o\text{-}o}$, is a measure of the magnitude of the differences among several observed MRs. For the number of observed MRs n equal to 2, $D_{o\text{-}o}$ is equal to the absolute value of their difference,

$$D_{o-o} \equiv \left|MR_1 - MR_2\right| \qquad n = 2$$

For more than two observed MRs this definition is generalized to twice the standard deviation of the observed MRs:

$$D_{o-o} \equiv 2\text{sd.}(MR_o) = 2\left[\overline{\left(MR_o - \overline{MR_o}\right)^2}\right]^{1/2}$$
$$= 2\left[\frac{1}{n}\Sigma\left(MR_o - \overline{MR_o}\right)^2\right]^{1/2} \qquad n \geq 2$$

According to this definition, for $n > 2$, $D_{o\text{-}o}$ is somewhat less than the difference between the maximum and minimum MRs. For example, for the set of numbers {1, 2, 3, 4, 5}, $D_{o\text{-}o}$=2.83; for {11, 12, 13, 14, 15}, $D_{o\text{-}o}$ is also equal to 2.83; for {2, 4, 6, 8, 10}, $D_{o\text{-}o}$ = 5.66, that is, twice $D_{o\text{-}o}$ for the set {1, 2, 3, 4, 5}, and for {1, 2, 4, 8, 16}, $D_{o\text{-}o}$ = 10.9. Because the difference characteristic spread is a measure of the absolute difference of a set of values, it results in the same value irrespective of whether a given spread is about a low or a high value. Note also that $D_{o\text{-}o}$ increases with increasing value of a set of observed MRs

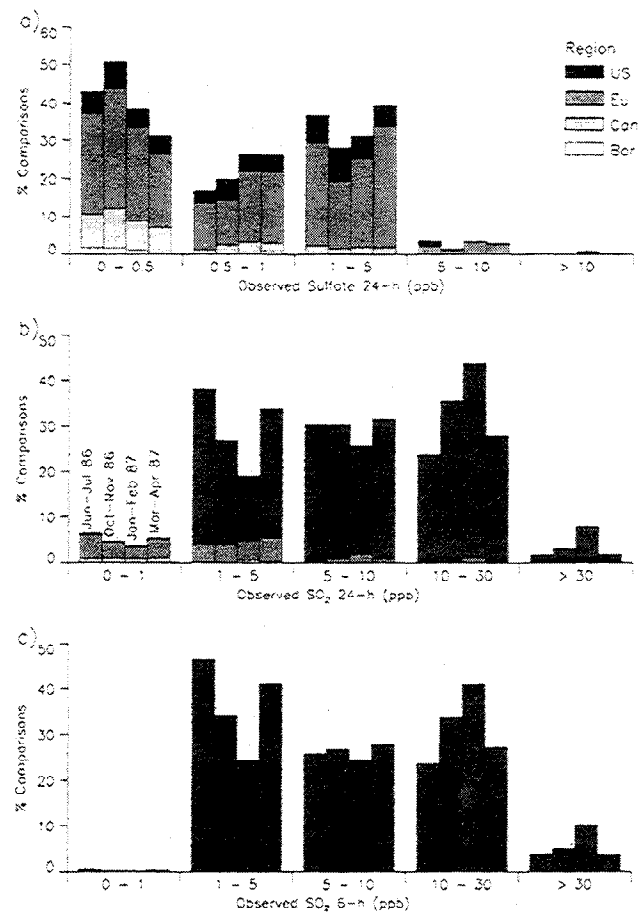

Figure 4. Distribution of the fraction (percent) of comparisons of modeled and observed mixing ratios as a function of location-average observed mixing ratios, by measurement network. In each cluster, each bar represents one simulation period, from left to right June-July 1986, October-November 1986, January-February 1987 and March-April 1987. Normalization is by simulation period, with number of comparisons N as follows: (a) Sulfate 24-hour; $N \approx 2000$. (b) SO_2 24-hour; $N \approx 13{,}000$. (c) SO_2 6-hour; $N \approx 50{,}000$. US, United States (Measurement Programs 3 and 4 described in Appendix A); Eu, Europe (Program 2); Can: Canada (Program 1); Bar, Barbados (Program 6).

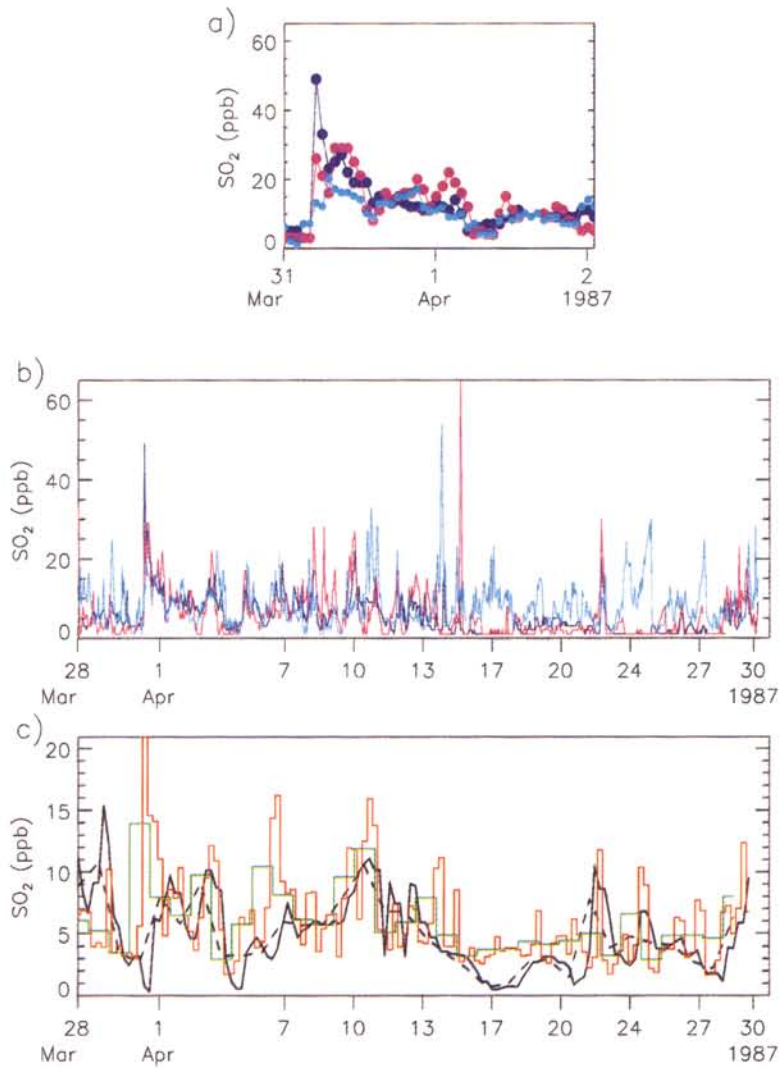

Plate 4. Examples of time series of SO_2 mixing ratios during March-April 1987 at three of nine stations in a location [(58, 24) 75.8°W, 39.9°N, Camden, New Jersey-Pennsylvania] for which there was continuous monitoring at multiple stations. (a) Detail of the observed SO_2 1-hour mixing ratio for March 31 to April 2, 1987 showing individual 1-hour measurements. (b) Time series of the observed 1-hour MR March 28 to April 30, 1987. (c) 6-hour (black solid curves) and 24-hour average (black dashed curves) modeled MR; 6-hour average observed MR (red), and 24-hour average observed MR (green) for the three stations. Peaks in observed MRs in (a and b are attributed to the transient influence of proximate sources.

that differ from each other by a constant factor. For this reason it useful also to consider a measure of the spread of the ratios of a set of values. Thus we define and evaluate a second quantity, which we denote as the ratio characteristic spread, $S_{o/o}$, a measure of the variation of the ratios of the observed MRs. For two observed MRs, $S_{o/o}$ is equal to the ratio of the greater to the lesser of the two values,

$$S_{o/o} \equiv \frac{\text{greater MR}_o}{\text{lesser MR}_o} = \exp\left|\ln\frac{\text{MR}_1}{\text{MR}_2}\right| \qquad n = 2$$

For $n \geq 2$, this definition is generalized analogously to that for $D_{o\text{-}o}$ such that $\ln S_{o/o}$ is twice the standard deviation of the logarithm of the observed MRs, or

$$S_{o/o} \equiv \exp\left\{2\text{s.d.}(\ln \text{MR}_o)\right\}$$
$$= \exp\left\{2\left[\frac{1}{n}\sum\left(\ln \text{MR}_o - \overline{\ln \text{MR}_o}\right)^2\right]^{1/2}\right\} \qquad n \geq 2$$

For $n > 2$, $S_{o/o}$ is somewhat less than the ratio of maximum to the minimum values. For example, for the set of numbers {1, 2, 3, 4, 5}, $S_{o/o} = 3.12$; $S_{o/o} = 3.12$ also for the set {2, 4, 6, 8, 10}; for {11, 12, 13, 14, 15}, $S_{o/o} = 1.25$; and for {1, 2, 4, 8, 16}, $S_{o/o} = 7.1$. Because $S_{o/o}$ involves ratios of MRs, observed MRs at the limit of detection were excluded from the evaluation of this quantity.

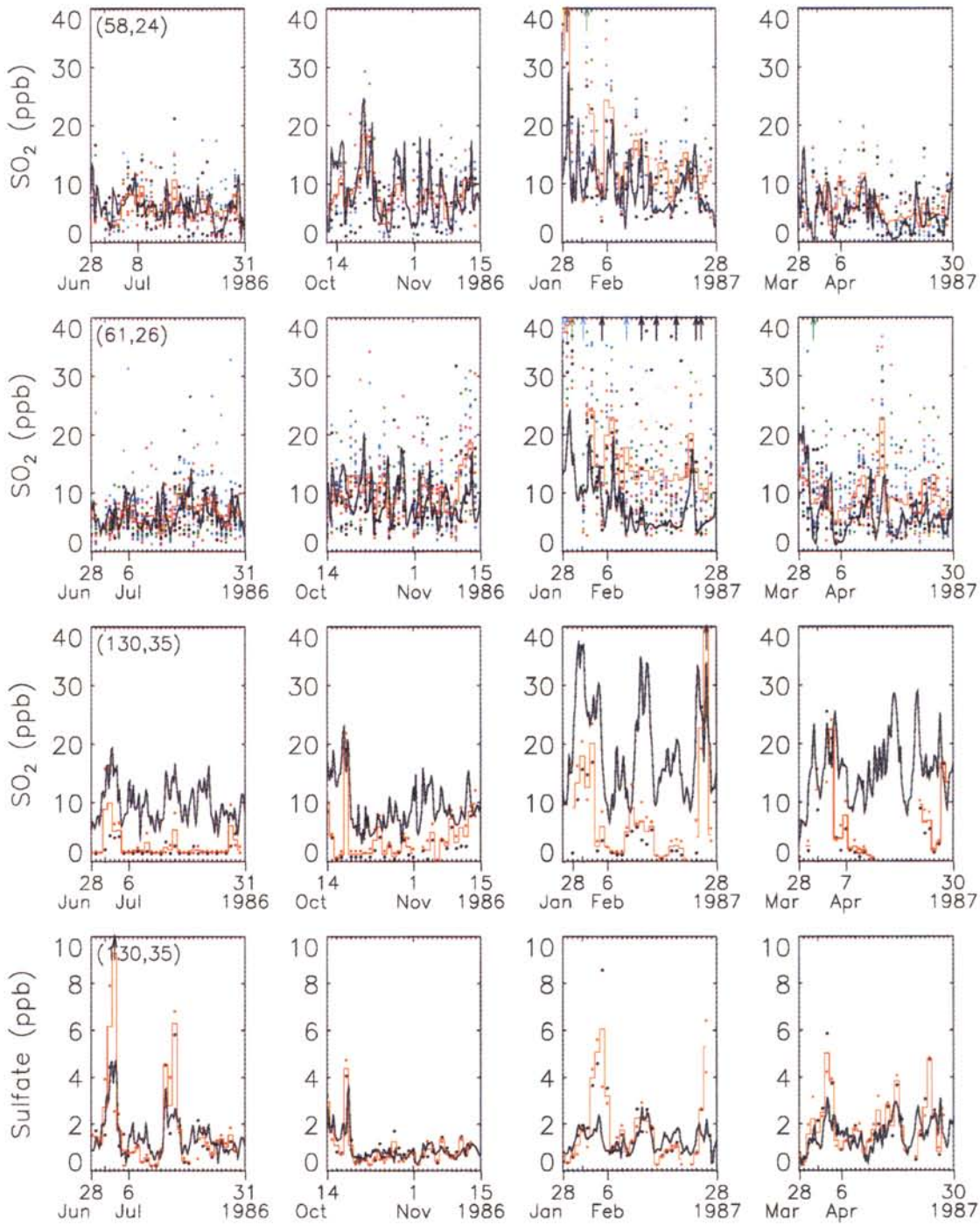

Plate 5. Examples of time series of SO_2 24-hour average mixing ratio and of sulfate 24-hour mixing ratio at locations with more than one station. Points of a given color represent observations at one station. The vertical spread of different-color points on a given date represents the within-location spatial variation of the MR on that date. The arrows indicate one or more observations exceeding the limit of the plot. The location-average observed MR (red cityscape) and the 6-hour modeled MR (black curve) are also displayed. (a) Camden, New Jersey-Pennsylvania; (b) Boston, Massachusetts; (c and d) north central Netherlands.

Distributions of the ratio characteristic spread $S_{o/o}$ of the observed sulfate and SO_2 MRs, (Figures 5a and 6a; Table 4) quantify the relative variation of the mixing ratios of these species within the 1.125° × 1.125° locations. For sulfate 24-hour MR the median ratio characteristic spread is a factor of 1.5; this finding applies to all simulation periods. For SO_2 $S_{o/o}$ is considerably greater than for sulfate, 1.9-2.4 for the 24-hour MRs and 2.3-2.9 for the 6-hour MRs, depending slightly on simulation period. The difference between the distributions for the 6-hour and 24-hour data suggests a decrease in relative within-location spatial variation with increasing averaging time. This inference might be questioned, however, in view of the different data sets comprising the 24-hour data and the 6-hour data and the differences in the distributions of mixing ratios

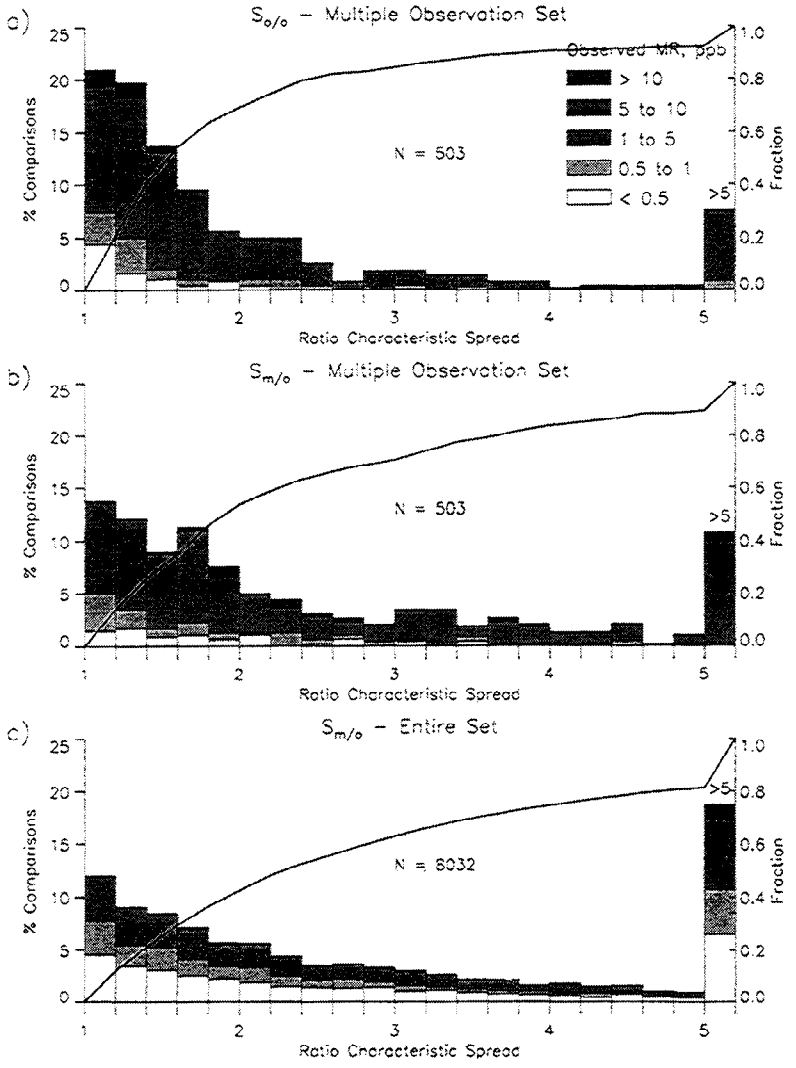

Figure 5. Histograms of the ratio characteristic spread for sulfate 24-hour mixing ratios. (a) Characteristic spread of the observed MRs within individual locations, (b) model-observation comparisons for data set included in Figure 5a, and (c) All available model-observation comparisons. Curves (right hand ordinate) give the cumulative fraction of comparisons exhibiting a spread less than the value denoted by the abscissa. In each panel, N denotes the total number of comparisons.

noted in Figure 4. For this reason we also examined the distribution of $S_{o/o}$ for the subset of the SO_2 24-hour MRs for which 6-hour MRs were also available (SO_2 24-hour*, Table 4); essentially identical distributions were observed, indicating that the decrease in $S_{o/o}$ with increasing averaging times is, in fact, attributable to the increase in averaging time and not an artifact of the difference in the data sets. Note for all quantities the fairly high values of $S_{o/o}$ at the tails of the distributions, indicative of a high degree of within-location spatial variation in a small fraction of the locations. No systematic seasonal differences were exhibited among the four simulation periods for either sulfate or SO_2.

Distributions of the difference characteristic spread $D_{o\text{-}o}$ of the observed sulfate and SO_2 MRs are presented in Figures 7a and 8a and Table 4. As might be anticipated, smaller values of $D_{o\text{-}o}$ are exhibited predominantly at small average observed mixing ratio $\overline{MR_o}$, and larger values are exhibited predominantly at large $\overline{MR_o}$, consistent with an expected scaling of the magnitude of $D_{o\text{-}o}$ with the magnitude of $\overline{MR_o}$. However, a considerable fraction of comparisons exhibit values of $D_{o\text{-}o}$ well less than $\overline{MR_o}$, indicative of fairly uniform spatial distribution within the location. For the four simulation periods, values of $D_{o\text{-}o}$ for sulfate were similar, with no evident systematic seasonal differences. For SO_2 the distributions of $D_{o\text{-}o}$ for the subset of the 24-hour MRs corresponding to the 6-hour MRs (SO_2 24-hour*, Table 4) are systematically narrower than for the 6-hour MRs, in contrast to the finding for the ratio characteristic spread $S_{o/o}$. This is attributed to a decrease in the spread of the distributions of the MRs themselves associated with the longer averaging period, together with the scaling of $D_{o\text{-}o}$ with $\overline{MR_o}$. Also in contrast to $S_{o/o}$, a systematic difference is exhibited between the 24-hour data as a whole and the subset corresponding to the 6-hour averages, with greater values of $D_{o\text{-}o}$ for the subset, again indicative of the scaling of $D_{o\text{-}o}$ with $\overline{MR_o}$; compare the greater MRs for the 6-hour data than for the 24-hour data indicated in Figure 4. Somewhat higher values of $D_{o\text{-}o}$ for SO_2 were exhibited in JF87, and lower values were exhibited in JJ86, probably as a consequence of SO_2 MRs being somewhat lower in JJ86 than in the JF87 (Figure 4). Because the ratio characteristic spread $S_{o/o}$ is not subject to this

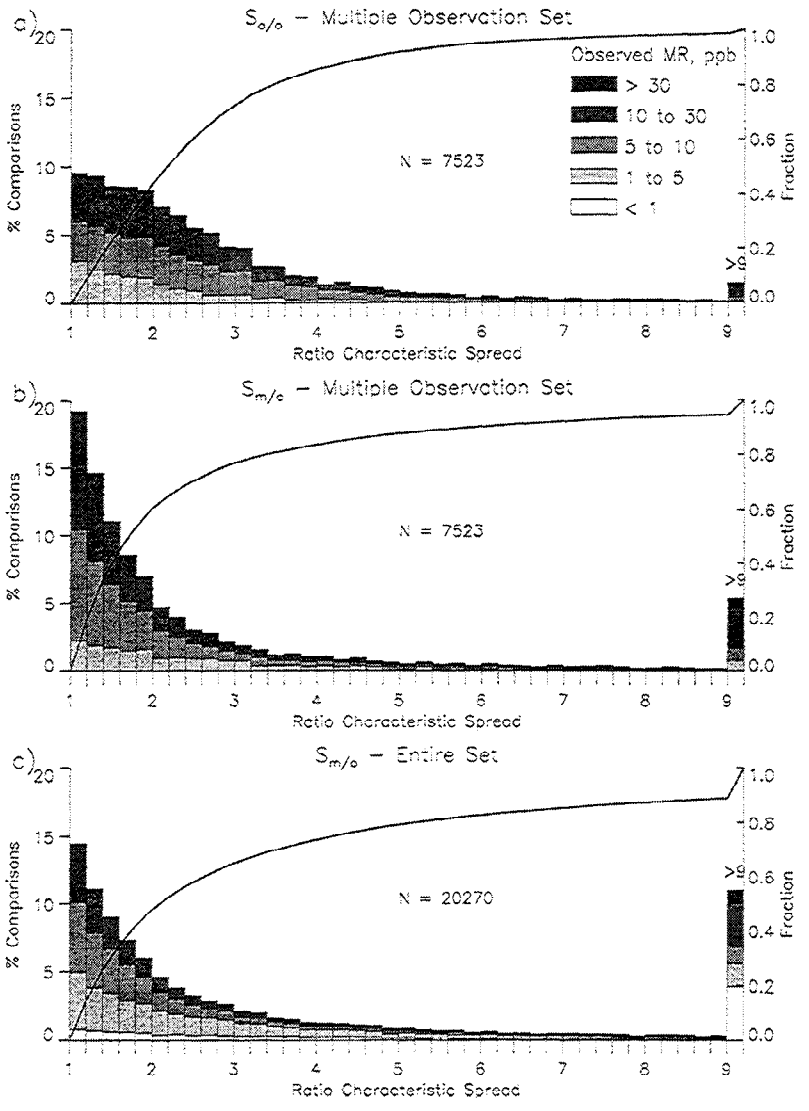

Figure 6. Histograms of the ratio characteristic spread for SO_2 24-hour mixing ratios, as in Figure 5.

scaling, we consider this quantity to be a more intrinsic measure of the characteristic spread of a data set than $D_{o\text{-}o}$.

The magnitude of the within-location variation of the observed MRs of both sulfate and SO_2 substantiates the concern over the direct comparison of modeled and observed MRs and qualifies the level of agreement that can be expected in these comparisons, especially for the primary emitted species SO_2 in the near-source regions. Although the stations used here are limited mainly to urban and suburban areas of the United States, it may be inferred that the within-location variation of the sulfate and SO_2 MRs is similar in other regions of the modeling domain where observed MRs may be influenced by emissions from proximate sources, such as much of central and eastern Europe. We surmise that the within-location variation of observed MRs would be somewhat less in regions not directly influenced by proximate sources.

5. Comparison of Model Results and Observations

To assess the performance of the model, we carried out a variety of comparisons of modeled and observed MRs and sulfate wet deposition. The objective was to examine the accuracy of the model and to identify and determine the extent of any model bias, either for the data set as a whole or for a particular simulation period or geographical region, or as a function of observed or modeled mixing ratio. Surface observations were compared to modeled MRs for the lowest model level, which extended from the surface to ~65 m, for averaging times of 24 hours for sulfate and 24 and 6 hours for SO_2. Multiple simultaneous observed MRs within a location were averaged; in comparisons involving ratios, observed MRs at or below the LOD were excluded. For evaluation of modeled sulfate wet deposition, we compared daily or weekly cumulative modeled sulfate wet deposition with the corresponding daily or weekly cumulative observed sulfate wet deposition.

Previous comparisons of modeled and observed MRs have been restricted typically to comparisons of monthly means. Comparisons of time series of daily or subdaily MRs at specific times present a much more detailed picture of the temporal pattern of model performance relative to observations at a given location. Additionally, we present several statistical characterizations of model performance. Correlation analysis provides a measure of the fraction of the variance in the

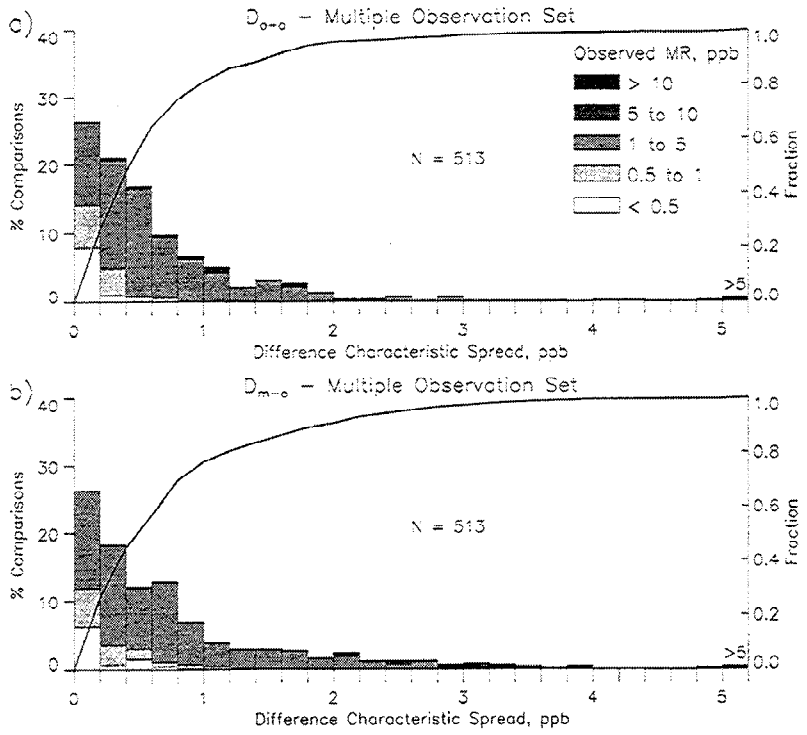

Figure 7. Histograms of the difference characteristic spread for sulfate 24-hour mixing ratios, as in Figure 5. Data for all available model-observation comparisons are not shown because the difference in the distribution of MRs themselves precludes meaningful comparisons of $D_{m\text{-}o}$.

observations that is explained by the model. Model accuracy is examined as the magnitude of departure between model and observation; because of availability of observed MRs at two or more sites within a given model grid cell, it is possible to compare departure between observed and modeled MRs with the spread of observed MRs within individual locations as evaluated above. Model bias is determined by examination of the ratio of modeled to observed MR, $MR_m/\overline{MR_o}$, and the difference between modeled and observed MR, $MR_m - \overline{MR_o}$.

A key component of our evaluation of model performance consists of examination of the characteristic spread between modeled and observed MRs, $D_{m\text{-}o}$ and $S_{m/o}$, defined analogously

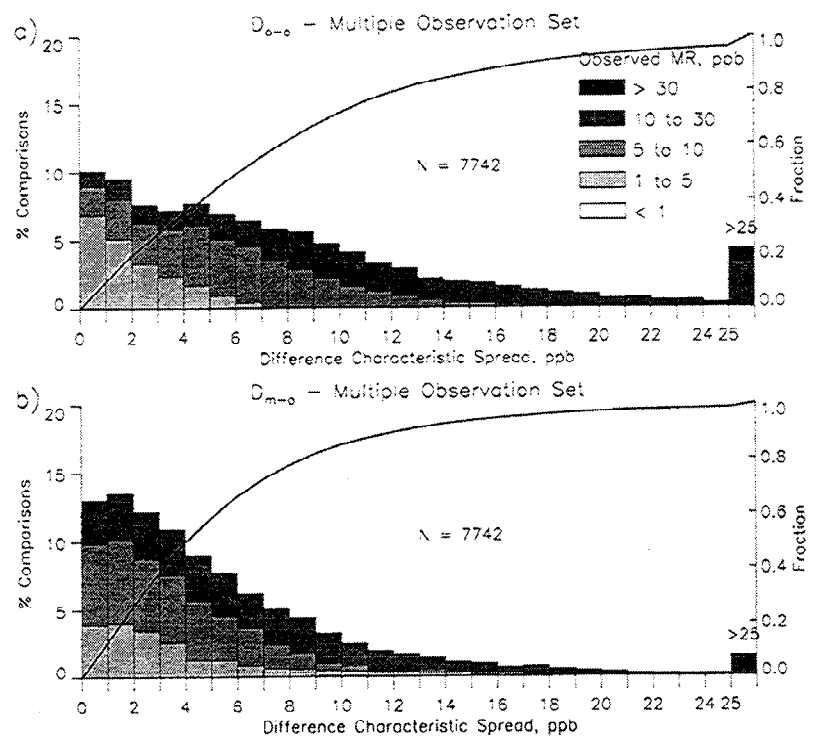

Figure 8. Histograms of the difference characteristic spread for SO_2 24-hour mixing ratios, as in Figure 5. Data for all available model-observation comparisons are not shown because the difference in the distribution of MRs themselves precludes meaningful comparisons of $D_{m\text{-}o}$.

Table 4. Distributions of Ratio and Difference Characteristic Spread and Ratios and Differences Between Observed and Modeled Mixing Ratios, by Percentile, for the Four Simulations Taken Together

Species	Quantity, ppb							
Percentile	$S_{o/o}$	$S_{m/o}$ (a)	$S_{m/o}$ (b)	$MR_m/\overline{MR_o}$	D_{o-o}	D_{m-o} (a)	D_{m-o} (b)	$MR_m - \overline{MR_o}$
Sulfate 24-hour	*503*		*7,907*		*513*		*7,918*	
10	1.1	1.1	1.2	0.15	0.06	0.07	0.05	-1.8
25	1.2	1.4	1.5	0.27	0.18	0.19	0.14	-0.76
50	1.5	1.9	2.3	0.51	0.42	0.49	0.37	-0.26
75	2.3	3.2	4.0	0.93	0.84	0.96	0.89	-0.02
90	3.8	5.4	7.1	1.7	1.5	1.9	1.9	0.28
SO_2 24-hour	*7,523*		*20,007*		*7,742*		*20,270*	
10	1.2	1.1	1.1	0.17	1.0	0.76	0.4	-9.0
25	1.5	1.3	1.4	0.45	2.7	1.9	1.3	-3.7
50	2.2	1.7	2.1	0.97	6.1	4.0	3.2	-0.09
75	3.1	2.8	4.2	2.0	10.7	7.4	6.7	2.8
90	4.6	5.8	9.7	4.4	17.5	12.1	12.1	7.0
SO_2 6-hour	*30,136*		*65,260*		*33,307*		*70,941*	
10	1.3	1.1	1.1	0.11	0.83	0.79	0.66	-12.1
25	1.7	1.3	1.4	0.32	2.6	2.0	1.7	-5.8
50	2.5	1.9	2.1	0.77	6.4	4.5	4.1	-1.2
75	4.0	3.2	4.1	1.5	12.5	8.8	8.5	2.6
90	6.2	6.6	9.6	2.8	22.7	14.4	14.4	7.1
SO_2 24-hour*	*7,269*		*14,144*		*7,336*		*14,407*	
10	1.2	1.1	1.1	0.13	1.1	0.76	0.64	-10.8
25	1.6	1.3	1.3	0.35	3.0	1.9	1.7	-5.4
50	2.2	1.7	1.9	0.75	6.4	4.0	3.7	-1.5
75	3.1	2.8	3.4	1.3	11.0	7.4	7.2	1.9
90	4.6	5.7	7.8	2.2	17.7	12.1	12.4	5.7

$S_{o/o}$ is ratio characteristic spread of simultaneous observed mixing ratios at multiple stations within a location. $S_{m/o}$ is ratio characteristic spread between modeled and observed mixing ratios. $S_{m/o}$(a) includes only cases with multiple observations; $S_{m/o}$(b) includes all comparisons. $MR_m/\overline{MR_o}$ is ratio of modeled to observed mixing ratios. D_{o-o} is difference characteristic spread of simultaneous observed mixing ratio at multiple stations within a location. D_{m-o} is difference characteristic spread between modeled and observed mixing ratios. $MR_m - \overline{MR_o}$ is difference between modeled and observed mixing ratios. For $MR_m/\overline{MR_o}$ and $MR_m - \overline{MR_o}$ underscore demarcates model underestimation (above) and overestimation (below). SO_2 24-h* denotes subset of SO_2 24-hour data included in the SO_2 6-hour data set. Number of comparisons N is given in italics.

to D_{o-o} and $S_{o/o}$ as

$$D_{m-o} \equiv \left|MR_m - \overline{MR_o}\right|$$

$$S_{m/o} \equiv \exp\left|\ln \frac{MR_m}{\overline{MR_o}}\right|$$

According to this definition, $S_{o/o} = 2$ for $MR_m/\overline{MR_o}$ equal to either 2 or 0.5. These definitions permit examination of the magnitude of the departure between modeled and observed MRs, a measure of model accuracy, and comparison to the within-location variation of the observations themselves measured by D_{o-o} and $S_{o/o}$. These characteristic spreads cannot, however, be compared to $MR_m - \overline{MR_o}$ and $MR_m/\overline{MR_o}$, which, because they indicate the sign of the departure as well as the magnitude, assume values greater and less than zero (or greater than and less than unity, respectively); that is, they are measures of model bias rather than accuracy.

5.1. Sulfate Mixing Ratios

We examine the model performance for the sulfate 24-hour MRs first by means of $S_{m/o}$ and D_{o-o}. We calculated these quantities both for locations for which multiple simultaneously observed MRs were available and which are represented in the calculation of $S_{o/o}$ and D_{o-o} (which we denote the multiple-observation set) and for the entire set of model-observation comparisons (the full set). Values of characteristic spread calculated for the multiple-observation set are expected to be directly comparable to the characteristic spread of the observed MRs themselves; however, because locations for which there are multiple simultaneous observations are preferentially in the vicinity of sources, both the magnitude and spread of the observed MRs evaluated for the multiple-observation set might be expected to be greater than for the full set of model-observation comparisons. These quantities are presented in Table 4.

Comparison of the distributions of the departure between modeled and observed MRs with those of the within-location variation of the observations themselves for the same subset of the data (Figures 5a and 5b, and 7a and 7b) shows the distributions to be similar but slightly broader for the model-observation comparisons. The median ratio characteristic spread between model and observations $S_{m/o}$ is a factor of 2.3, versus 1.5 for the spread of the observations $S_{o/o}$. Examined seasonally, the distributions of both the difference and ratio characteristic spread of the comparisons were likewise similar to those of the observations, except for the JF87 simulation, for

Table 5. Distributions of Ratio Characteristic Spread and Ratios of Observed and Modeled Mixing Ratios, by Region and Percentile, for the Four Simulations Taken Together

Species	Quantity			
	North America		Europe	
Percentile	$S_{m/o}$	$MR_m/\overline{MR_o}$	$S_{m/o}$	$MR_m/\overline{MR_o}$
Sulfate 24-hour	*2,400*		*5,507*	
10	1.2	0.16	1.2	0.14
25	1.5	0.27	1.5	0.28
50	2.3	0.47	2.2	0.54
75	3.8	<u>0.78</u>	4.1	<u>1.00</u>
90	6.6	1.4	7.4	1.9
SO_2 24-hour	*14,791*		*5,216*	
10	1.1	0.14	1.2	<u>0.45</u>
25	1.3	0.36	1.6	1.04
50	1.9	<u>0.79</u>	2.9	2.3
75	3.6	1.4	6.0	5.1
90	8.1	2.4	13.9	12.4

$S_{m/o}$ is ratio characteristic spread of modeled and observed mixing ratios. $MR_m/\overline{MR_o}$ is ratio of modeled to observed mixing ratio; underscore demarcates model underestimation (above) and overestimation (below). Comparisons are for the entire data set. Number of comparisons N is given in italics.

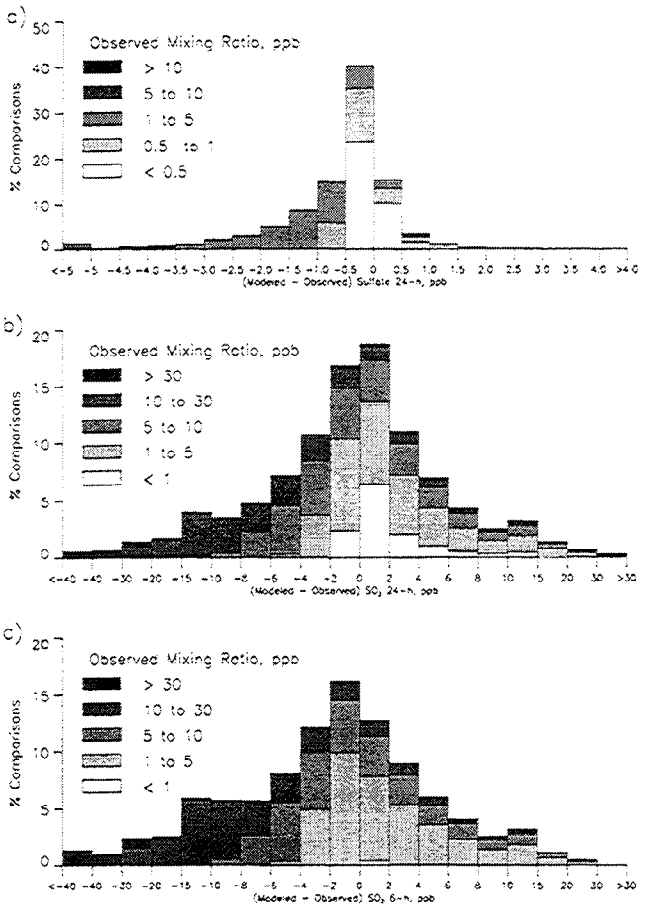

Figure 9. Histograms of the difference between modeled and observed mixing ratios $MR_m - \overline{MR_o}$ for (a) sulfate 24-hour observations, (b) SO_2 24-hour averages, and (c) SO_2 6-hour averages. Comparisons are classified by the magnitude of the observed mixing ratio. Note the increase in bin width at large values of the difference.

which the distribution of $S_{m/o}$ is considerably broader than that for $S_{o/o}$. The median values of $S_{m/o}$ are comparable for European and North American locations (Table 5), again except for the JF87 simulation, for which the distribution for North America is somewhat broader. With these exceptions the comparisons suggest that by this measure the model represents the observed mixing ratios with an accuracy comparable to the within-location variation of the observed mixing ratios.

To assess model bias, we examine the differences between modeled and observed MRs (Figure 9a). The distribution of the difference between modeled and observed MRs, $MR_m - \overline{MR_o}$ peaks at a low value of this difference, a measure of the overall lack of bias of the model. (Distributions for the individual simulation periods were similar.) Importantly, the peak in the distribution of $MR_m - \overline{MR_o}$ includes a fairly wide range of observed MRs, not just very low MRs, showing that the accuracy of the model indicated by this peak is not a consequence simply of small differences at low MR. It is evident from the figure that there are more instances of model underestimation than overestimation. For the data set as a whole, the model underestimates the observed MR for 79% of the comparisons. This bias toward underestimation is manifested also in Table 4. Here attention is called to the underscores in the columns for $MR_m/\overline{MR_o}$ and $MR_m - \overline{MR_o}$, which demarcate the regions of model underestimation and overestimation; a lack of model bias would be indicated for this demarcation point (corresponding to $MR_m/\overline{MR_o} = 1$ or $MR_m - \overline{MR_o} = 0$) at the median of the distribution. Model underestimation for sulfate was indicated in more than 75% of the comparisons for all simulations except JJ86.

When classified by the magnitude of the observed MR, the distributions exhibit considerably more bias. For observed MR 0.5 to 1.0 ppb the fraction of the modeled results within 0.5 ppb of the corresponding observation for the four simulation periods ranged from 57 to 75%; for observed MR 1.0 to 5.0 ppb the fraction of model results within 1.0 ppb of the corresponding observation ranged from 44 to 53%. As indicated in Table 6, for

Table 6. Fraction (%) of the Comparisons Between Modeled (MR_m) and Observed ($\overline{MR_o}$) Mixing Ratios for which $MR_m > \overline{MR_o}$, Grouped According to the Value of the Observed Mixing Ratio, for the Four Simulations Taken Together, for the Individual Simulations, and Separately for North America and Europe

Species Observed MR, ppb	All	June-July 1986	October-November 1986	January-February 1987	March-April 1987	North America	Europe
Sulfate 24-hour							
All	20.9	32.3	22.3	17.0	20.2	12.6	24.7
<0.5	34.2	49.6	31.1	25.0	32.0	25.2	38.7
0.5 to 1	22.5	36.4	21.1	18.2	21.9	9.1	27.7
1 to 5	9.1	13.6	8.3	8.2	11.4	3.1	12.1
5 to 10	0	0	0	0	0	0	0
>10	0	0	0	0	0	0	0
SO_2 24-hour							
All	48.8	44.8	*52.7*	46.2	51.4	39.5	*75.9*
<1	*82.3*	*74.5*	*83.2*	*86.6*	*86.6*	*93.1*	*79.7*
1 to 5	*64.2*	*56.7*	*69.4*	*68.9*	*63.7*	*57.1*	*78.4*
5 to 10	42.3	31.7	*50.4*	46.2	40.2	39.8	*65.9*
10 to 30	20.4	14.5	24.5	21.4	18.7	18.8	49.4
>30	2.5	0.0	1.0	4.7	0.0	1.3	18.8

Absence of model bias is indicated by entries in the table equal to 50%. A lower value indicates model underestimation; a higher value (in italics) indicates model overestimation. For locations with multiple simultaneous observations, average mixing ratios are employed.

observed MR less than 0.5 ppb the modeled MR is less than the observed MR in 66% of the comparisons (range 50 to 75% for the four simulations) increasing to essentially 100% for observed MR greater than 5 ppb. This bias is more pronounced in the comparisons by region, with appreciably greater fraction of model underestimates in North America than in Europe, albeit with rather similar distributions of $MR_m/\overline{MR_o}$ (Table 5). Possible causes of model underestimation include errors in oxidant concentrations employed and lack of representation of aqueous-phase conversion in nonprecipitating clouds; the more frequent underestimates at higher observed MR may be also due to the influence of proximate sources on the observations. The fraction of underestimates was least for the JJ86 simulation, perhaps reflecting higher oxidant concentrations in summer and/or a greater amount of in-cloud reaction being represented, because of a greater fraction of precipitating clouds.

Correlation coefficients were calculated for the data as a whole, by region, and by simulation, as summarized in Table 7. The data set as a whole and the several subsets exhibited lognormal distributions; therefore we used the Pearson product-moment correlation coefficient R and Student t test to evaluate the significance of the coefficients. Correlation coefficients were all positive and significant at the 99% confidence level (i.e., null hypothesis probability ≤ 0.01) for the data set as a whole and for the several subsets. However, the statistic R^2, which is an estimate of the fraction of the variance in the data explained by the correlation, was generally rather low, 0.44 for the data set as a whole and ranging from 0.32 (JF87) to 0.50 (ON86).

To further characterize model skill, we examined correlations at individual locations, denoted "time series locations," for which observed MRs were available for at least 25 days in a 30-

Table 7. Statistics on the Pearson Product-Moment Correlation Coefficient R Between Observed and Modeled 24-hour Sulfate and 24-hour SO_2 Mixing Ratios, for the Four Simulations Taken Together, for the Individual Simulations, and Separately for North America and Europe

	Sulfate 24-hour			SO_2 24-hour		
	N	R	R^2	N	R	R^2
Entire domain						
All	*7,895*	0.66	0.44	*20,015*	0.43	0.18
June-July 1986	*1,818*	0.70	0.49	*4,362*	0.47	0.22
October-November 1986	*1,705*	0.71	0.50	*4,701*	0.44	0.19
January-February 1987	*1,940*	0.57	0.32	*5,230*	0.40	0.16
March-April 1987	*2,432*	0.65	0.42	*5,452*	0.34	0.12
NA (west of 30°W)	*2,388*	0.68	0.46	*14,799*	0.34	0.12
Eu (east of 30°W)	*5,507*	0.64	0.41	*5,216*	0.56	0.31

N indicates the number of comparisons in each set. R^2 indicates the fraction of variance explained by the correlation. NA, North America; Eu, Europe.

Table 8. Statistics on Spearman Rank-Order Correlation Coefficients of Observed and Modeled Mixing Ratios Applied to Individual Locations for which Observations were Available for at Least 25 Days in an Individual Simulation

	Individual Simulations					Multiple Simulations		
	June-July 1986	October-November 1986	January-February 1987	March-April 1987	All	North America	Europe	All
Sulfate								
Number of correlations	*46*	*44*	*50*	*63*	*203*			*53*
percent significant at 95% c.l.	85	75	64	81	76			91
SO_2								
Number of correlations	*110*	*122*	*147*	*147*	*526*	*125*	*50*	*175*
Percent significant at 95% c.l.	37	50	63	49	51	70	84	74
Percent significant with negative correlation coefficient	2.5	1.6	3.2	0	1.9	1.1	0	0.8

Results are presented for the four individual simulations, for the four simulations taken together, and, for all locations for which the 25-day criterion was met for more than one simulation period, for the grouped data sets (multiple simulations). SO_2 correlations are distinguished also by region; sulfate correlations are not distinguished by region because of the small number of locations in North America for which correlations could be determined. For sulfate all of the significant correlations exhibited positive correlation coefficients. The abbreviation c.l. denotes confidence level.

to 34-day simulation period. The correlations by individual location were calculated by simulation and, for locations where the 25-day criterion was satisfied for more than one simulation, for the several simulations together (combined simulations). As the majority of the cases did not exhibit lognormal distributions, it was necessary to apply a nonparametric analysis. Therefore we calculated the Spearman rank-correlation coefficient [*Sokal and Rohlf*, 1981], Table 8. For the combined simulations, 91% of the coefficients were significant at the 95% confidence level, with all significant coefficients positive. For the analysis by individual simulations the fraction of correlations significant at the 95% confidence level was 76%, ranging from 64% (JF87) to 85% (JJ86), again all with positive coefficients.

Scatterplots of observed versus modeled MRs at individual locations are shown in Plate 6 for all locations for which observed MRs were available for at least 25 days in at least one simulation period. The color code in the figure indicates the simulation period to which an individual point belongs. In each plot the one-to-one line is indicated by the dashed line, which is shown in bold for correlations that are significant at the 95% confidence level. The sense of the figures is that a point above the one-to-one line denotes model underestimation, and below the line it denotes overestimation. There is no evident consistent indication of segregation by simulation period in modeled or observed MR that would indicate a systematic seasonal bias in the quality of the correlations, although in some instances there is a tendency in this regard, for example locations (72, 0) Barbados, (121, 24) Toledo, Spain, and (135, 30) Arabba, Italy.

Also shown in each panel of Plate 6 (magenta curves) are the cumulative distributions of the ratio characteristic spread $S_{m/o}$ for the individual scatterplots. The value of the ordinate (range 0-1, as indicated by the scale panel (last panel) denotes the fraction of the individual observed mixing ratios whose departure from the one-to-one line is within a factor denoted by the value on the abscissa scale (range 1-3). Thus, for location (29, 39) Cree Lake, Saskatchewan, 47% of the observed MRs are within a factor of 2 of the one-to-one line, and 64% are within a factor of 3. These curves provide a convenient visual means of evaluating the accuracy of the model in each of these locations. It is seen that in several locations the model exhibits a rather high accuracy, with well more than half the values of $S_{m/o}$ within a factor of 2, for example, locations (44, 23) St. Louis, Missouri-Illinois, (124, 37) High Muffles, England, (130, 35) north central Netherlands, and (136, 36) Neuglobsow, Germany. Occasionally, such accuracy is indicated even when the correlation is not significant, for example, locations (132, 30) Jungfraujoch, Switzerland (station elevation 3600 m). Several locations exhibit substantial model underestimation, for example, location (72, 0) Barbados, (118, 26) Bragança, Portugal, and (124, 34) Barcombe Mills, England; much more rarely is there systematic overestimation, most notably location (132, 30) Jungfraujoch, Switzerland. There are also numerous instances of substantial scatter about the one-to-one line that also result in low values of the cumulative distribution of $S_{m/o}$, for example, locations (130, 41) Skreådalen, Norway and (144, 34) Jarczew, Poland. Overall, however, the picture is one of a substantial majority of the comparisons within a factor of 2 to 3 of the one-to-one line.

As a further measure of model performance, we directly compared the time series of modeled and observed MRs. Figure 10 presents time series plots of the modeled and observed sulfate MRs for JJ86; the plots are similar for the other simulation periods (not shown, but see B94 for ON86). These plots show first that in general the model accurately represents the spatial distribution of the observed sulfate MRs; that is, for the most part at locations where the observed MRs were low, the modeled MRs were also low, for example, locations (29, 39) Cree Lake, Saskatchewan, (72, 0) Barbados, (106, 45) Irafoss, Iceland, and (137, 47) Tustervatn, Norway, and where observed MRs were high, the modeled MRs were also high, for example, (130, 35) north central Netherlands, (134, 35) Langenbrugge, Germany, and (145, 37) Suwalki, Poland. There were several instances of remarkably good quantitative agreement, for example at locations (49, 30) Algoma, Ontario, (136, 36) Neuglobsow, Germany, (136, 38) Vavihill, Sweden, (139, 31) Illmitz, Austria, and (144, 42) Utö, Finland. Even in instances

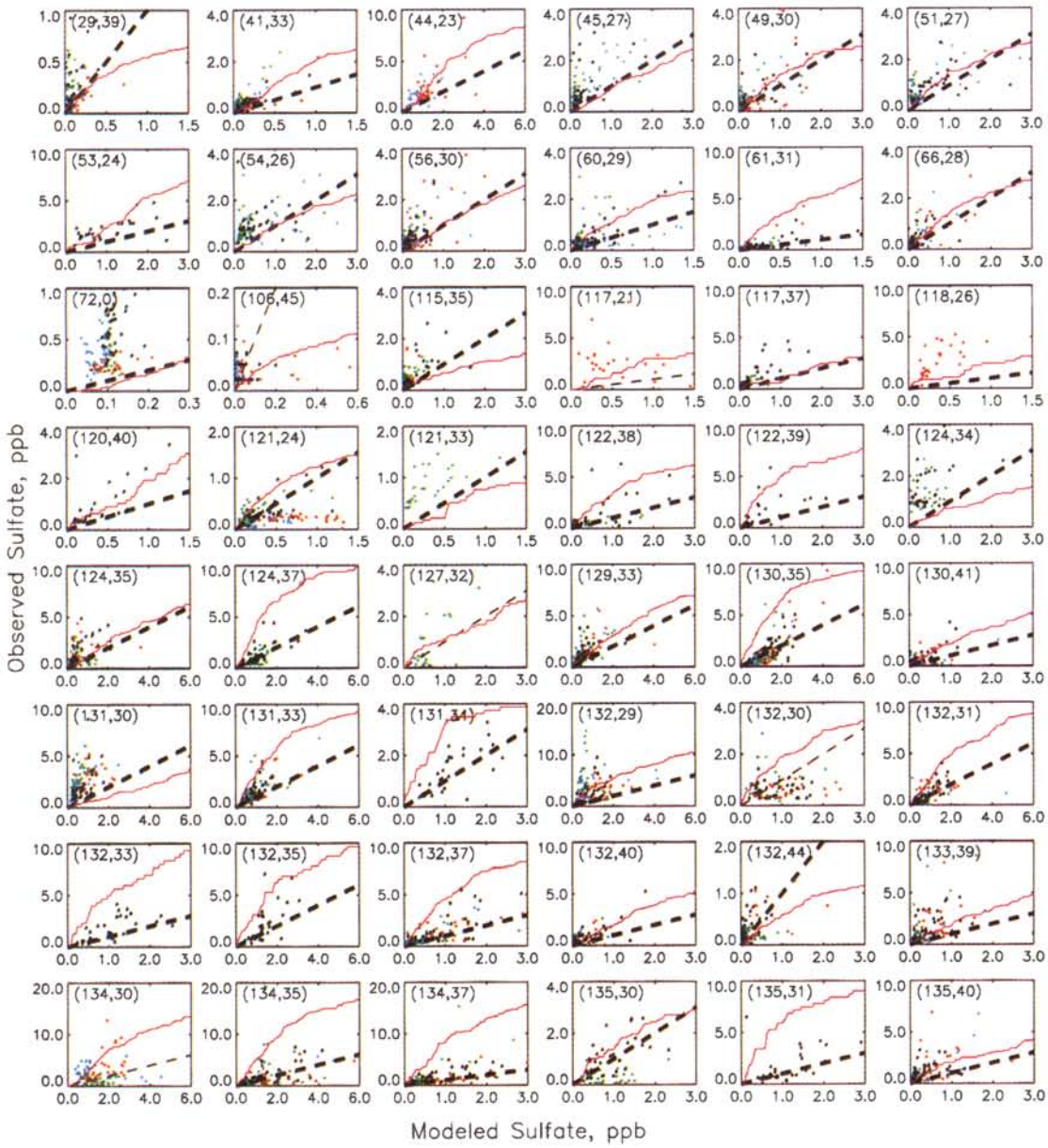

Plate 6. Scatter plots of the modeled and observed sulfate 24-hour mixing ratios (ppb, nmol/mol air) at all locations for which MRs were available for at least 25 days in one or more of the four simulation periods; the individual simulations are distinguished by color. The abscissa and ordinate of each graph are scaled by the data, as shown. The dashed line is the 1:1 line; thick lines indicate correlations that are significant at the 95% confidence level (null hypothesis probability ≤ 0.05). Correlations are for the data from the one or more simulation periods indicated by the colors of the points in the several graphs. In each graph the thin magenta curve denotes the cumulative probability distribution of the ratio characteristic spread between modeled and observed MRs, $S_{m/o}$, plotted for an ordinate range 0-1 and abscissa range 1-3 as given at the bottom of the plate. Indices (x, y) denote locations shown in Plate 3; longitude = -140.625 + 1.125 $(x+0.5)$; latitud e = 12.375 + 1.125 $(y+0.5)$.

where quantitative agreement was less close, the episodicity (pattern of short-term temporal variability) of the observations was well represented by the model; that is, peaks and valleys of observed and modeled MRs exhibit similar temporal fluctuations, with coincident or nearly coincident peaks and valleys, for example at (134, 30) Stelvio, South Tyrol, Italy; (130, 35) north central Netherlands (see also Plate 4); (134, 35) Langenbrugge, Germany; and (135, 40) Rorvik, Sweden. Observed MRs at (117, 21) Faro, Portugal, (118, 26) Bragança, Portugal, and (136, 26) Montelibretti, Italy indicate considerable apparent episodic influence of proximate sources that is not captured by the model.

5.2. SO_2 Mixing Ratios

5.2.1. SO_2 24-hour mixing ratios. We examine model performance for SO_2 24-hour MRs by the same comparisons as for sulfate. We note first that the distributions of the ratio and

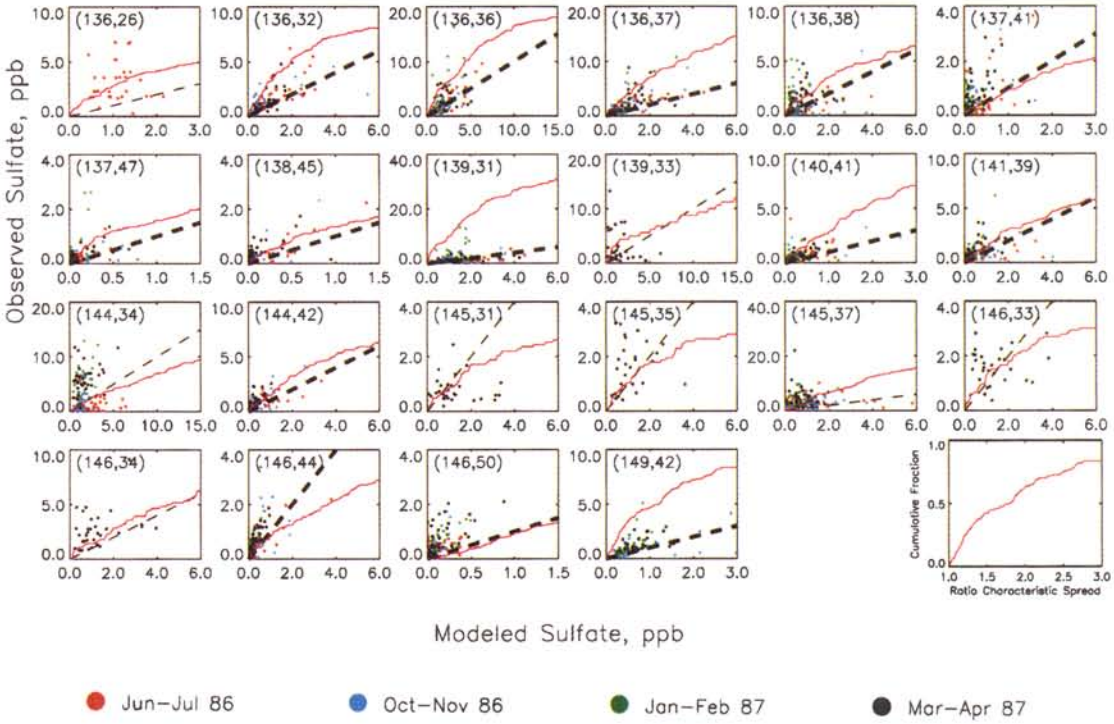

Plate 6. (continued)

difference characteristic spreads between the modeled and observed MRs for the multiple-observation set ($S_{m/o}$ and $D_{m\text{-}o}$, Figures 6b and 8b) are appreciably more strongly peaked at low values than the distributions of the within-location variation of the observations themselves ($S_{o/o}$ and $D_{o\text{-}o}$, respectively); that is, that the modeled MRs agree with the observed MRs to an accuracy that is within the subgrid variation of the observed MRs themselves. For the ratio characteristic spread this finding holds also not just for the multiple-observation set but also for the entire set of model-observation comparisons (Figure 6c versus 6a). Median values of $S_{m/o}$ (Table 4) are a factor of 1.7 for the multiple-observation set and 2.1 for the full set, compared to 2.2 for $S_{o/o}$. The accuracy of the model in the multiple-observation set is somewhat surprising in view of the rather large within-location variation in mixing ratio arising from the fact that the multiple observations are primarily in source regions. Only at the higher end of the distribution (90th percentile for the multiple-observation set and 75th and 90th percentiles for the full set) does the ratio characteristic spread between model and observations exceed the spread of the observations themselves; note the peak at $S_{m/o}$ >9 in Figures 6b and 6c. The model accuracy is somewhat greater in North America than Europe (Table 5), with half of the modeled MRs within a factor of 1.9 of the observed, versus a factor of 2.9 for Europe.

Differences between modeled and observed MRs are presented in Table 4 and Figure 9b. As was found for sulfate, the distribution of the difference between modeled and observed MRs, $MR_m - \overline{MR_o}$ peaks at a low value of this difference, indicative of the overall accuracy of the model; the fractions of model underestimates and overestimates are approximately equal (49% overestimations, Table 6). Again, the peak in the distribution of $MR_m - \overline{MR_o}$ at low $MR_m - \overline{MR_o}$ includes a wide range of observed MRs, not just very low MRs, with most of the large differences at the tails of the distribution arising from higher observed MRs. Note also that the demarcation points between model underestimation and overestimation in Tables 4 and 5 lie fairly close to the median. As with sulfate, greater model bias is indicated when the distribution of $MR_m - \overline{MR_o}$ is examined as a function of observed MR, Table 6. For observed MR less than 5 ppb, modeled MRs predominantly exceed observed MRs. In contrast, for observed MR between 5 and 10 ppb, fewer than half of the modeled MRs exceed the observed MRs, and for observed MR greater than 10 ppb, fewer than 25% of the modeled MRs exceed the observed MRs. The systematic underestimation for observed MRs greater than 10 ppb presumably reflects the inability of the model to represent the influence of proximate sources, as discussed above. A systematic bias is noted between North America and Europe (Table 5), with a tendency toward underestimation in North America (median $MR_m/\overline{MR_o}$ 0.79) and overestimation in Europe (median $MR_m/\overline{MR_o}$ 2.3). Model underestimation at high observed MR is more pronounced in North America than Europe (Table 6). This may reflect the difference in siting strategies for monitoring sites in the two regions and is reflected in differences in distributions of observed MRs evident in Figure 4; sites in North America are intended mainly for determining population exposure and compliance with air quality standards. The distributions of $MR_m - \overline{MR_o}$ are similar for the several simulation periods, except that the JF87 simulation exhibits a greater fraction of observed MRs exceeding 5 ppb and thus a broader distribution, with more values in the tails.

As with sulfate, we examined model predictive skill by examining the correlation coefficients of observed and modeled MRs. Here also, the distributions of the data set as a whole, by region, and by simulation were found to be lognormal, but the distributions for individual locations were not lognormal. All correlations by region and simulation are significant at the 99% confidence level (null hypothesis probability ≤ 0.01), Table 7.

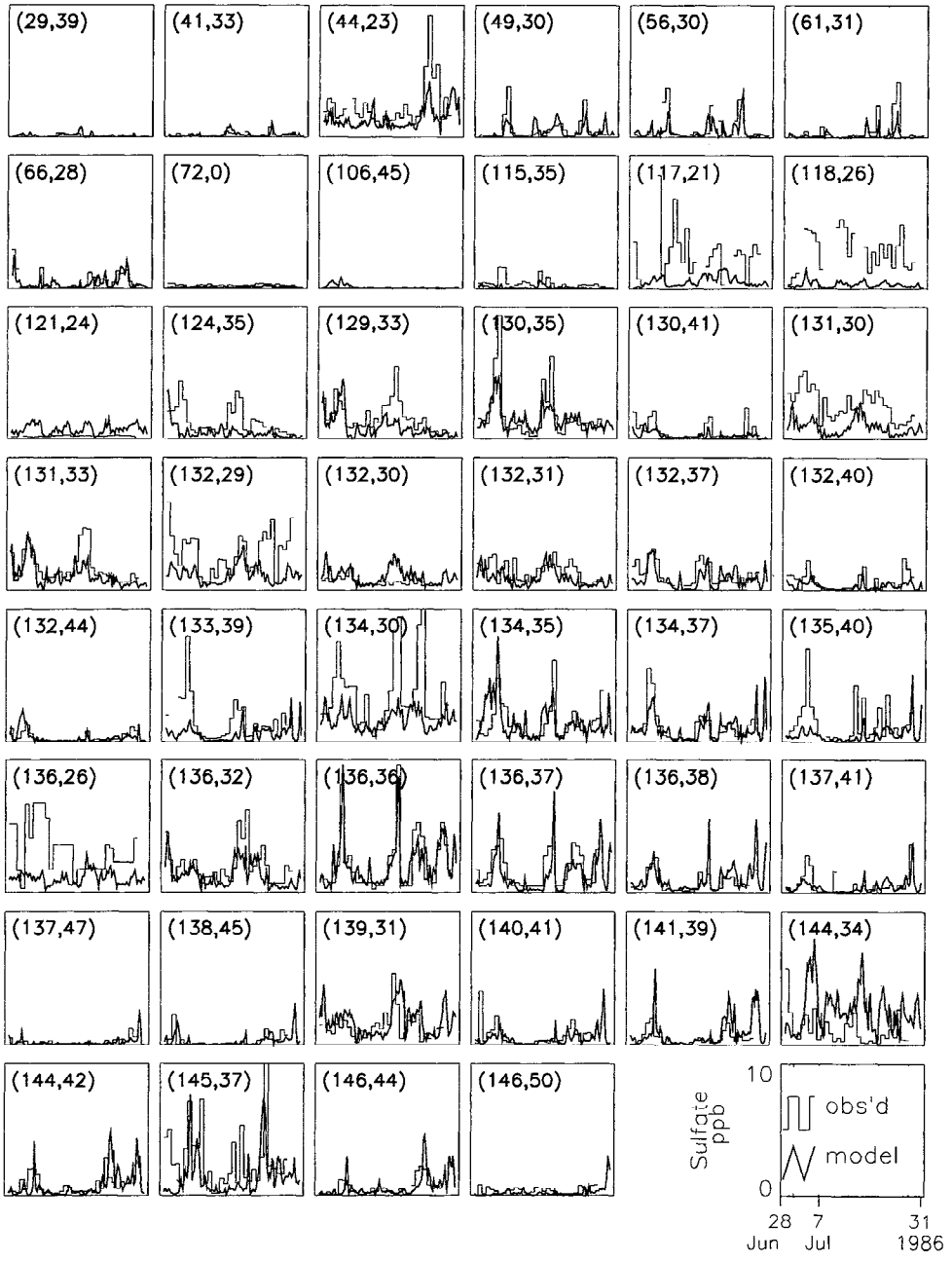

Figure 10. Time series of the modeled (thick continuous curve) and observed (thin line cityscape) sulfate 24-hour mixing ratios at all locations for which MRs were available for at least 25 days of the June-July 1986 simulation period. Ordinate scale, 0-10 ppb (nmol/mol air), and date scale, June 28 to July 28, shown at the bottom of the figure are the same for all graphs. Indices (x, y) denote locations shown in Plate 3; longitude = $-140.625 + 1.125\,(x+0.5)$; latitude = $12.375 + 1.12\,(y+0.5)$.

However the R^2 values are quite low, much lower than for sulfate, 0.18 for the data set as a whole. R^2 for European locations, 0.31, is greater than for North American locations, 0.12. Thus, despite highly significant correlations, the model explained a fairly small fraction of the variance of the observed SO_2 MRs when measured in this point by point comparison of the data set as a whole, by region, or by simulation. We also calculated correlation coefficients at individual locations (shown in Plate 3) for which observed MRs were available for at least 25 days in a 30- to 34-day simulation period. These correlations were examined individually by simulation and for all simulations together (combined simulations); again, nonparametric analysis was employed, Table 8. As measured by these statistics, the model exhibits a somewhat lower skill for SO_2 than for sulfate. For the combined simulations, 81% of 152 correlations are significant at the 95% confidence level, compared to 91% for sulfate. A negative correlation coefficient was exhibited in a single regression; this is attributed to a nonsignificant correlation falsely indicated as significant within the statistical confidence level. Of a total of 526 correlations for the individual simulation periods, 51% are significant at the 95% confidence level, compared to 76% for sulfate, including some

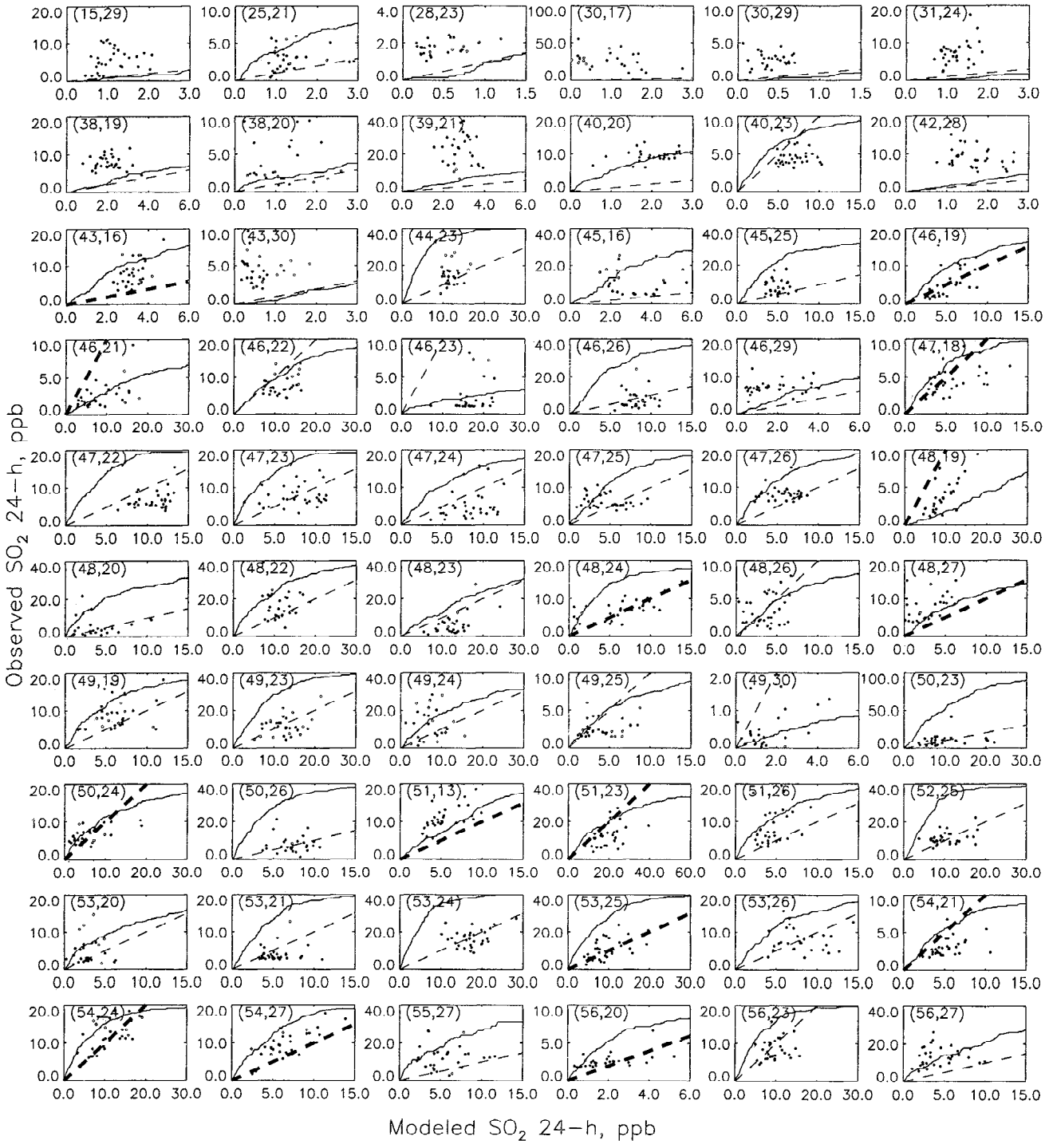

Figure 11. Scatterplots of the modeled and observed SO_2 24-hour mixing ratios (ppb, nmol/mol air) for June-July 1986 at all locations for which MRs were available for at least 25 days of the simulation period. The abscissa and ordinate of each graph are scaled by the data, as shown. The dashed line is the 1:1 line; thick lines indicate correlations that are significant at the 95% confidence level (null hypothesis probability ≤ 0.05). In each graph the thin curve denotes the cumulative probability distribution of the ratio characteristic spread between modeled and observed MRs, $S_{m/o}$, plotted for an ordinate range 0-1 and abscissa range 1-3 as given at the bottom of Figure 11 (continued). Indices (x, y) denote locations shown in Plate 3; longitude = -140.625 + 1.125 (x+0.5); latitude = 12.375 + 1.125 (y+0.5).

3% with negative correlation coefficients. A somewhat better performance was indicated in the European comparisons than in the North American comparisons. The poorer performance of the model in North America relative to Europe (Tables 7 and 8) may be attributable to the greater influence of local sources and the inability of the model to represent subgrid variability that influences the observed MRs.

Scatterplots of observed versus modeled MR are shown in Figure 11 for the JJ86 simulation; the plots for the other simulations are similar. As in Plate 6 for sulfate, a bold one-to-

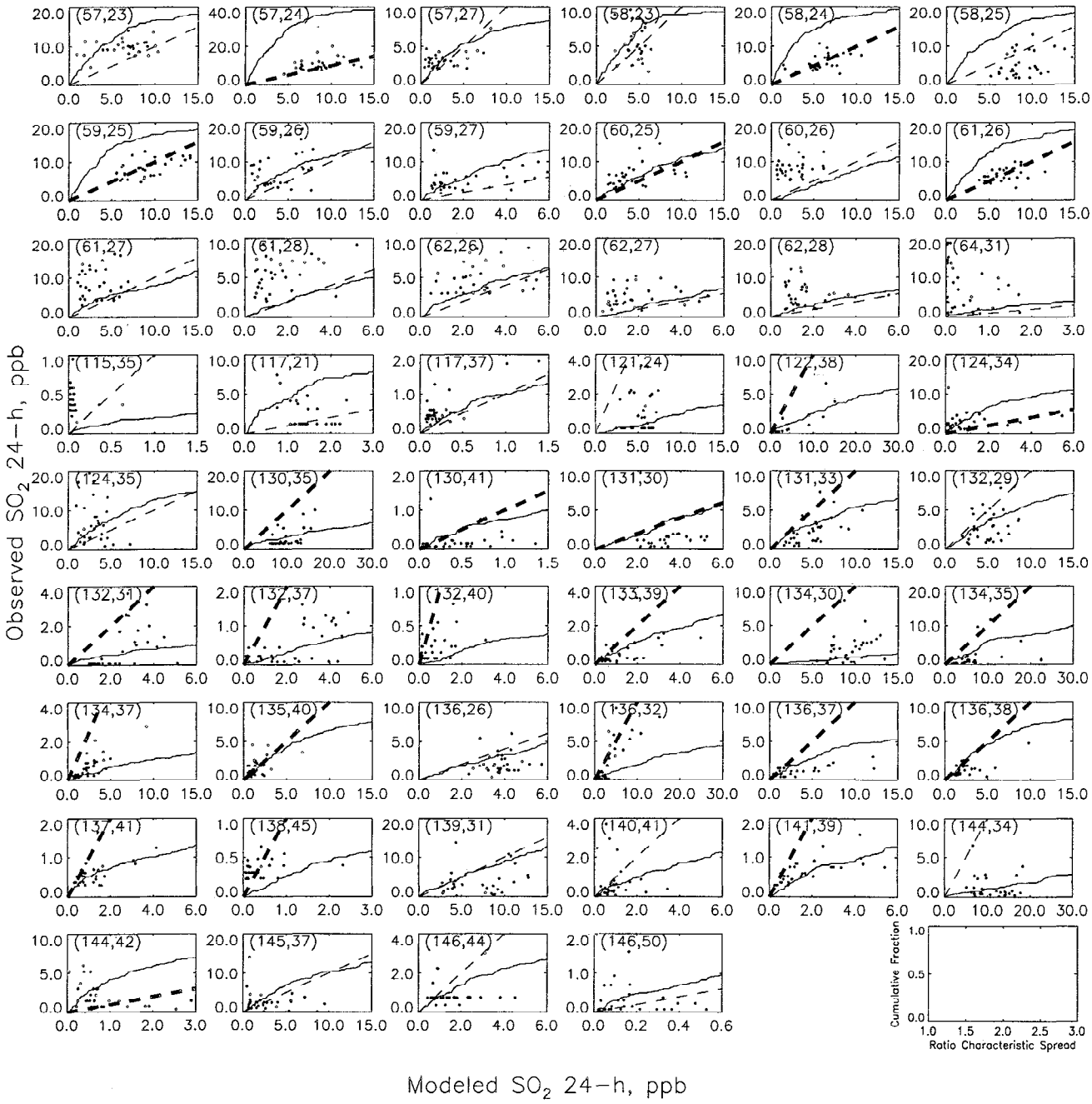

Figure 11. (continued)

one line indicates that the correlation is significant at the 95% confidence level. The figure also indicates the distribution of the cumulative fraction of values of the ratio characteristic spread between model and observations $S_{m/o}$ over the range 1 to 3 . Figure 11 shows some fairly convincing examples of model skill, as evidenced by significant correlation and by a high fraction of the values of $S_{m/o}$ that are close to unity, for example (47, 18) Birmingham, Alabama, (54, 24) Johnstown, Pennsylvania-Maryland, and (58, 24) Camden, New Jersey-Pennsylvania. However, there are numerous instances for which the modeled points lie fairly close to the one-to-one line that may be taken as a measure of model skill, even where the correlation is not significant, for example, locations (46, 22) Evansville, Indiana-Illinois-Kentucky, (50, 26) Toledo, Ohio-Michigan, and (52, 25) Cleveland, Ohio. These examples also indicate substantial skill as measured by the cumulative plots of $S_{m/o}$. For a fair number of the significant correlations the points lie close to the one-to-one line, but with considerable scatter that leads to a rather low fraction of the values of $S_{m/o}$ near unity; for example, at locations (60, 25) New Haven, Connecticut and (136, 32) Brotjacklriegel, Germany. On the other hand, numerous other significant correlations are at considerable variance with the one-to-one line, for example, (46, 21) Paducah, Kentucky, (130, 35) north central Netherlands, and (138, 45) Bredkälen, Sweden; for these locations the fraction of values of $S_{m/o}$ near unity is fairly low, despite the correlation being significant. Finally, there are numerous locations for which little model skill is indicated by either standard.

We also examined time series of modeled and observed MRs (see Figure 12 for JJ86; the plots for the other simulation periods are similar). These plots show first (as with the plots for sulfate MRs) that in general the model accurately represents the spatial distribution of the observed SO_2 MRs, with locations exhibiting low observed MRs also exhibiting low modeled MRs, for example, (28,23) Grand Junction, Colorado, (49,30) Algoma, Ontario, (115,35) Valentia, Ireland, (130,41) Skreådalen, Norway, and (146,50) Jergul, Norway, and vice versa, for example, (44,23) St. Louis, Missouri-Illinois, (49,23) Cincinnati, Ohio-Indiana-Kentucky, and (53,24) Pittsburgh, Pennsylvania-Ohio. Again, there are quite a few instances of remarkably good quantitative agreement between modeled and observed MRs, for example, (52,25) Cleveland, Ohio, (53,25) Youngstown, Ohio-Pennsylvania, (58,24) Camden, New Jersey-Pennsylvania (see also Plate 4), (61,26) Boston, Massachusetts (also Plate 4), (135,40) Rorvik, Sweden, and (136,32) Brotjacklriegel, Germany. In some instances where the model did not match the magnitudes of the observed MRs the episodicity (pattern of short-term temporal variability) of the observations was nonetheless well represented by the model, for example, (50,24) Dayton, Ohio, (122,38) Eskdalemuir, Scotland, and (132,37) Westerland, Germany. Instances of poor quantitative agreement and episodicity, for example, (30,17) El Paso, Texas-New Mexico, (39,21) Tulsa, Oklahoma, (45,16) Gulfport, Mississippi, are often at or within one grid cell of large point source emissions. Observed MRs at locations (46,23) Mt. Vernon, Illinois and (144,34) Jarczew, Poland are also close to large point sources and show indication of occasional influence of plumes from large nearby sources; these plumes are not represented in the model, which instantly averages the emissions over the entire volume of the grid cell into which the material is emitted. In all four simulations the model overestimates the SO_2 MRs at these locations. Possible causes for model overestimation include errors in the emissions information for these locations, or perhaps that sources were not operating at full capacity during the simulation periods. Location (134,30) Stelvio, South Tyrol, Italy, is in mountainous terrain (station elevation 1400 m), where observed MRs are influenced by subgrid terrain features which may be responsible for ducting flows.

Comparison of the measures of model skill indicated by correlation analysis and cumulative plots of $S_{m/o}$ (Figure 11) and time series plots (Figure 12) indicates that good model skill by all of these measures is exhibited at several locations, for example, locations (46,19) Florence, Alabama, and (53,25) Youngstown, Ohio-Pennsylvania. However, locations such as (44,23) St. Louis, Missouri-Illinois and (62,26) Beverly, Massachusetts, despite showing rather convincing model skill in the time series comparisons, nonetheless exhibit correlations that are not significant. The slight displacements of the peaks and valleys of the time series at these locations prevent good agreement when MRs are compared in a scatterplot. Because of comparable episodic behavior of the modeled and observed MRs, locations (46,21) Paducah, Kentucky, and (134,30) Arabba, Italy (station elevation 2000 m), exhibit statistically significant correlations, despite substantial model overestimations. Additionally, despite a higher fraction of significant correlations in Europe than in North America, the cumulative plots of $S_{m/o}$ in Figure 11 show considerably lower skill in Europe than in North America. These considerations suggest that correlation analysis of scatterplots does not by itself yield a wholly meaningful evaluation of model performance and that this evaluation can be substantially enhanced by examination of the distribution of $S_{m/o}$ and by comparison of time series of modeled and observed MRs.

5.2.2. SO_2 6-hour mixing ratios. In view of the high variability of the SO_2 1-hour observations illustrated in Plate 4, we compared the 6-hour modeled MRs (the time resolution of the meteorological data driving the model) to the corresponding 6-hour averages of the observed 1-hour MRs to ascertain whether comparisons on such shorter timescales yielded any improvement in the representation of the observed MRs over that indicated in the 24-hour comparisons. As noted above (Table 4), the 6-hour observations exhibit a considerable systematically greater within-location ratio characteristic spread $S_{o/o}$ than do the 24-hour observations, indicative of the spatial smoothing that results from temporal averaging. As with the 24-hour averages, the median values for the model-observation ratio characteristic spread $S_{m/o}$ and difference spread D_{m-o} for the several simulation periods are invariably less than the corresponding within-location spread of the observations themselves (Tables 4 and 5). The median values of the ratio characteristic spread between model and observations $S_{m/o}$ of the 6-hour data are essentially identical to those in the corresponding subset of the 24-hour data, indicative of comparable model accuracy on this averaging period and hence of a slight relative improvement in model performance at 6-hour averaging in comparison to magnitude of the within-location spread measured by $S_{o/o}$. (The difference characteristic spread is greater for the 6-hour data than for the 24-hour data, again because it scales with the values of the MRs.) Comparison of the distributions of $MR_m/\overline{MR_o}$ and $MR_m - \overline{MR_o}$ for the 6-hour versus the corresponding subset of the 24-hour data indicates no discernible difference in the bias between the 6-hour and 24-hour averaging periods. In sum, comparing modeled and observed MRs for 6-hour averaging periods indicates little improvement relative to the 24-hour comparisons.

5.3. Sulfate Wet Deposition

As noted in B94, comparison of modeled and observed wet deposition presents additional problems and caveats beyond those for comparison of atmospheric MRs. In particular such comparisons are subject to error due to inaccuracy in the representation of location and amount of precipitation in the meteorological forecast model that drives the chemical transport model, in addition to any error in the atmospheric loading of the substance and in the representation of wet removal processes. Evaluation of the precipitation data for ON86 by B94 concluded that the meteorological data from the ECMWF 6-hour forecast model used to drive GChM-O underestimates precipitation for all precipitation amounts and that the geographic precipitation patterns showed some displacement from the observed patterns. The absence of precipitation at a model grid cell at a specific time could be due to the meteorological forecast model not representing the event at all or to the event being displaced from the point of observation, in either space (to a nearby grid cell) or in time (in the same model grid cell but at a different time) or both. As discussed in Appendix B, the numbers of modeled and observed precipitation events are similar, but a substantial fraction of these events are displaced; that is, modeled events are not seen in observations and vice versa. Because only instances of coincident modeled and observed precipitation events are compared, comparisons of modeled versus observed sulfate wet deposition amounts are probably weighted more heavily by

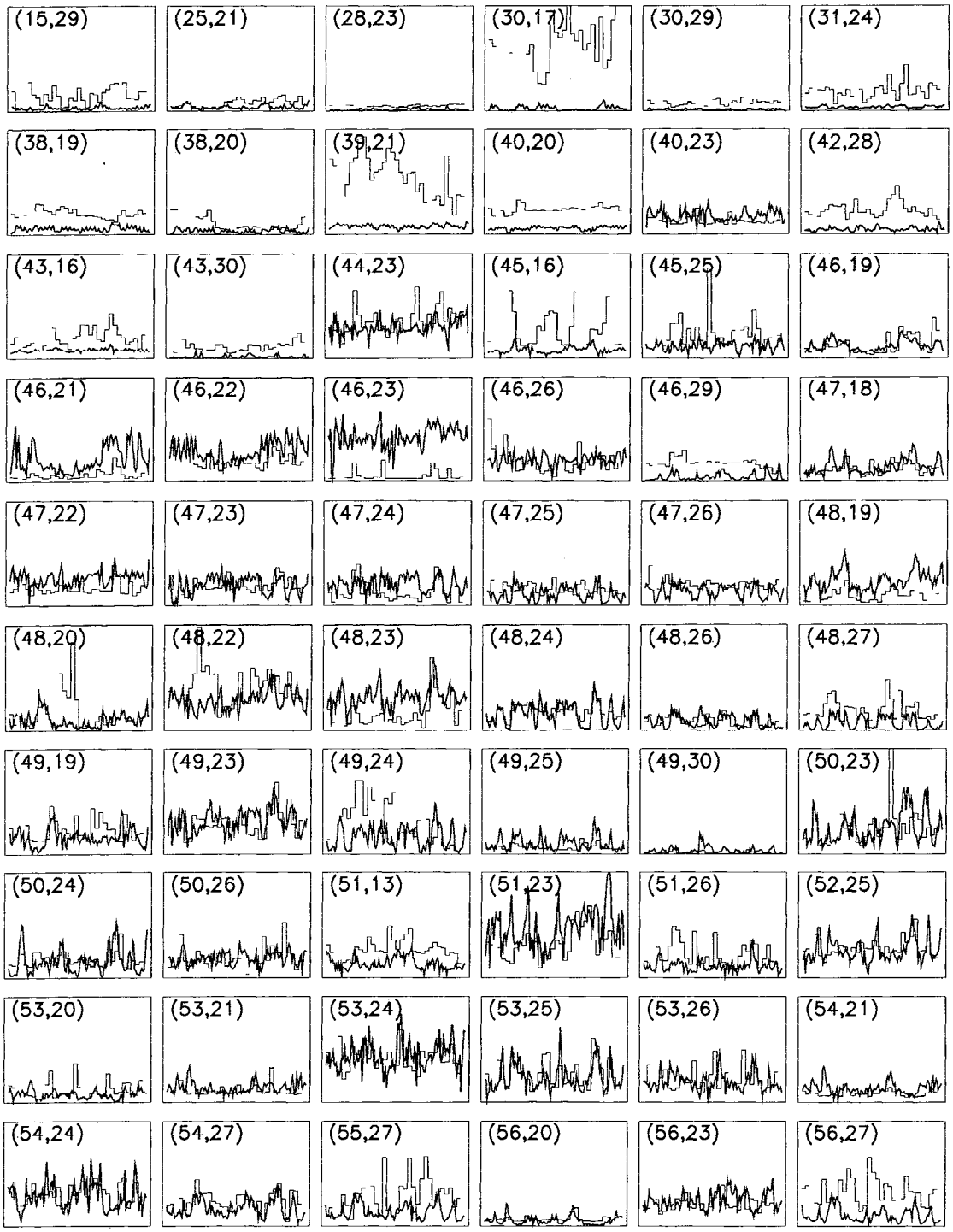

Figure 12. Time series of the modeled (thick continuous curve) and observed (thin line cityscape) SO_2 24-hour mixing ratios at all locations for which MRs were available for at least 25 days of the June-July 1986 simulation period. Ordinate scale, 0-40 ppb (nmol/mol air), and date scale, June 28 to July 28, shown at bottom of Figure 12 (continued) are the same for all graphs. Indices (x, y) denote locations shown in Plate 3; longitude = -140.625 + 1.125 (x+0.5); latitud e = 12.37 5 + 1.125 (y+0.5).

widespread frontal rain events than by local convective precipitation and may thus be nonrepresentative. Especially in weekly samples, the temporal variation in sulfate loadings over time period of aggregation inevitably contributes to error in comparisons between observed and modeled sulfate deposition. A further problem with the comparisons conducted in this way is that reported precipitation events do not start or end on the 6-hour time resolution of the meteorological data; the modeled wet sulfate deposition was aggregated to the precipitation sampling period, so discrepancies necessarily arise from this aggregation as well.

With these caveats we nonetheless present comparisons of modeled and observed amounts of wet-deposited sulfate A_m and A_o, Table 9; comparisons are restricted to the Canadian

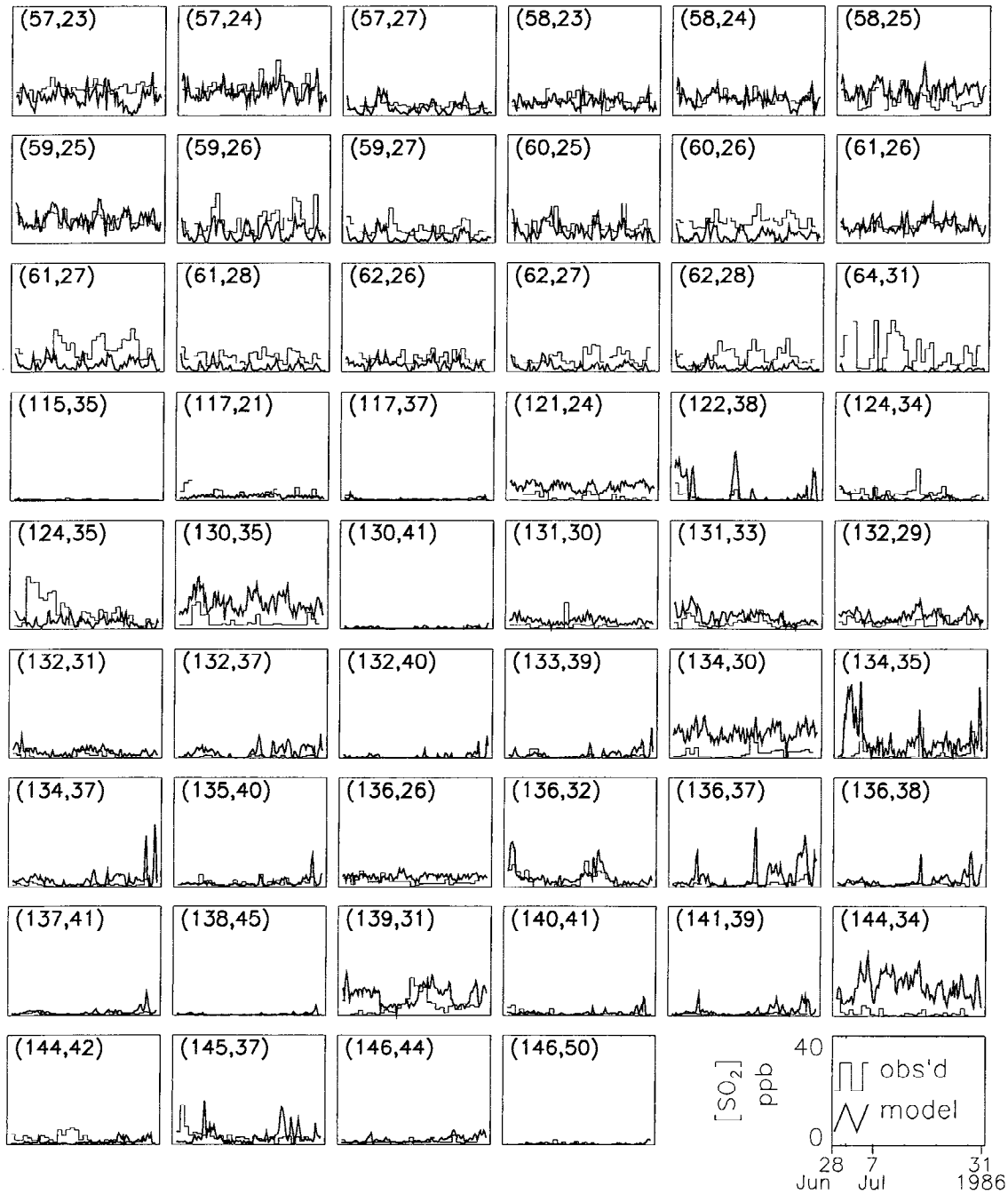

Figure 12. (continued)

CAPMON network (daily samples) and the U.S. National Trends Network (NTN) (weekly samples). The comparisons were made using the same statistical methods as for mixing ratios. Model accuracy is reflected in the distributions of the ratio characteristic spread $S_{m/o}$. Median values of this quantity ranged from 1.9 to 3.3 for the two networks and four simulation periods with median 2.6 for the entire data set; that is, half the modeled mixing ratios were within this factor of the observed values. However, there are substantial tails to the distribution, with the 90th percentile values in the JF87 simulation being a factor of 37 and 200 for the CAPMON and NTN data sets, respectively. Examination of the distributions for the ratio A_m/A_o (Table 9) and difference A_m-A_o (not shown), which quantities indicate the sense of the departure as well as the magnitude, shows that the large values of $S_{m/o}$ are due mainly to model underestimates. Still, in all but one instance (JF87, CAPMON data) the demarcation point between underestimation and overestimation lies between the 25th and 75th percentiles of the distributions, with overall median value of A_m/A_o of 0.82, indicative of a fairly symmetric distribution. There is little systematic difference in the accuracy or bias evident in the comparisons for the two networks, despite one (CAPMON) comparing daily samples and the other (NTN) comparing weekly samples.

Table 9. Distributions of Ratio Characteristic Spread $S_{m/o}$ and Ratio A_m/A_o of Modeled and Observed Sulfate Wet Deposition Amount, by Observing Network, for the Four Simulations Taken Together

	Network					
	CAN		NTN		Both	
Percentile	$S_{m/o}$	A_m/A_o	$S_{m/o}$	A_m/A_o	$S_{m/o}$	A_m/A_o
	342		*1100*		*1442*	
10	1.1	0.23	1.2	0.08	1.2	0.09
25	1.6	<u>0.58</u>	1.5	0.26	1.5	0.30
50	3.0	1.5	2.5	<u>0.67</u>	2.6	<u>0.82</u>
75	6.0	4.0	5.2	1.5	5.4	1.9
90	13.9	8.8	16.1	3.1	15.0	4.4

CAN denotes the CAPMON network (Canada), daily samples; NTN is the National Trends Network (US), weekly samples. Underscore demarcates model underestimation (above) and overestimation (below). Number of comparisons N is given in italics.

6. Discussion

It is now recognized that aerosols exert a shortwave radiative influence on climate that is of comparable magnitude on regional scales (and potentially on hemispheric to global scales) to longwave forcing by anthropogenic greenhouse gases, but opposite in sign. It is thus increasingly becoming appreciated that accurate representation of anthropogenic influences in climate models requires representation of the influences of aerosols in addition to those of greenhouse gases. A major difference between aerosols and the greenhouse gases is that aerosols, because of their short residence time, highly nonuniform distribution of sources, and sporadic removal processes, exhibit geographical distributions that are highly nonuniform in space and time. Because aerosol influences on climate can be anticipated to be highly nonlinear, especially the indirect effect, this situation seems to imply, in contrast to the situation for greenhouse gases, that accurate representation of aerosol influences in climate models requires modeling the distribution of aerosols "on-line" in these models rather than "off-line," with subsequent incorporation of a representative aerosol field in a climate model.

The present study represents an initial step in formulating a model capable of representing the sources and removal processes governing sulfate aerosol in a manner that might be suitable for incorporation in a climate model and in evaluating the performance of the sulfate model by comparison with observations. However, even this step presents challenges. Evaluation of model performance by comparison with observations requires driving the model by observationally derived synoptic meteorological data. Thus, in addition to any inaccuracy in the chemical module, the model will also reflect inaccuracies in the meteorological driver. Additionally, this study requires accurate representation of sources and a representative set of measurements against which to compare model output.

Both the meteorological and chemical components of the model represent compromises. First is the requirement for a self-consistent meteorological field driven by observationally derived synoptic data. To meet this requirement, we employ the available 6-hour forecast fields of ECMWF. As noted above and also in B94, this data set exhibits limitations in representing clouds and precipitation, most importantly lack of liquid water content of nonprecipitating clouds and the resultant underestimation of aqueous-phase reactions in the model. Likewise, as already noted, model inaccuracy in representation of the location and amount of precipitation limits the comparability of wet deposition; this inaccuracy results in a complementary inaccuracy in the modeled aerosol sulfate mixing ratio.

With respect to the chemical module, a key requirement for suitability for incorporation in climate models is that the set of variables to be actively modeled be kept as small as possible consistent with maintaining a representation that is based in known chemical reactions, rather than, for example, a constant percent per hour conversion rate. To this end, we have employed a reaction-based but simplified treatment of oxidation reactions making use of climatological fields of Ohio and O_3 concentrations and of H_2O_2 production rate generated "off-line" rather than modeling these species themselves. Consequently, any inaccuracy in concentrations of these oxidants, including lack of variability and any lack of correlation with the sulfur species, will be reflected in inaccuracy in the model output. Finally, in this initial effort we have restricted ourselves to a single aerosol component "sulfate" without any representation of the size distribution and the influence of size on removal processes, and likewise neglecting any influence of other aerosol species on the size distribution of sulfate.

With respect to sources of sulfur species, we have employed the best available representation of anthropogenic emissions, taking into account major point sources and their effective height of emission, and representing smaller sources as distributed areal emissions. However, it should be stressed that we have employed emissions that are generically representative of the mid-1980s as opposed to emissions for the specific time periods modeled. Further, we have not adjusted the emissions to account for specific operating patterns of individual facilities. Finally, we have used only a fairly simple representation of the seasonal pattern of emissions rather than any detailed model for this pattern. We have also used biogenic emissions that are viewed as representative of the typical June, October, etc., rather than emissions that are specific for June 1986, October 1986, and so forth.

A crucial requirement for evaluating the performance of an Eulerian model is the availability of a large, geographically distributed set of observations with which the spatial distribution and temporal variation of the modeled quantities may be compared. To meet this requirement in this study, we have made use of available data from monitoring networks. In our comparisons we have focused mainly on 24-hour mixing ratios of sulfate and SO_2 at the surface, and to lesser extent 6-hour average SO_2 MR and daily and weekly wet deposition of sulfate. As noted above, a major potential limitation of these data sources is nonrepresentativeness resulting from siting strategies intended to meet other requirements. To assess the representativeness of the observational data, we examined the within-location spatial variation of simultaneously observed mixing ratios, where such data were available. For 24-hour average sulfate mixing ratio the median value of the ratio characteristic spread of the observed MRs, $S_{o/o}$, was a factor of 1.5; for 24-hour SO_2, it was 2.2; and for 6-hour SO_2, it was 2.5. We attribute the decrease in within-location variation between the 6-hour and 24-hour SO_2 MRs to the longer averaging period. We attribute the decrease between 24-hour SO_2 and 24-hour

sulfate to a smoother concentration field for the longer-lived secondary species sulfate than for the shorter-lived primary emitted species SO_2. To the extent that these values are representative of the data set as a whole, they indicate substantial within-location spread of the observed MRs to which the modeled MRs are to be compared and thus a potential limitation on the ability to evaluate the model using these observations. The median values of the ratio characteristic spread between modeled and observed MRs $S_{m/o}$ were in fact comparable to the values for the observations themselves. For sulfate 24-hour MRs, $S_{m/o}$ was slightly greater than $S_{o/o}$, 1.9 for the subset for which multiple observations were available and 2.3 for the entire data set. For SO_2 the within-location spread of the observed MRs actually exceeded the spread between modeled and observed MRs: for the 24-hour average MRs, 1.7 and 2.1 for the multiple-observation set and for the entire set, respectively; and for the 6-hour averages, 1.9 and 2.1, respectively. By this rather broad standard it would thus seem that the model can be viewed as agreeing with the measurements to near or within the subgrid variation of the observations themselves.

Despite the apparent success of the model by this standard, in our view, this comparison presents a rather limited assessment of model performance, as it masks other information that is available in the comparison of modeled and observed MRs. In particular, we note that the several comparisons of modeled and observed MR show evidence of distinct biases. In the case of sulfate MR the model systematically underestimates sulfate in all simulation months, and in both Europe and North America. For the several simulation periods the median ratio of modeled to observed MR ranges from a high of 0.66, in JJ86, to a low of 0.36, in JF87 (0.27 in North America in JF87). This systematic departure from unity would appear to be indicative of substantial model underestimation despite the near agreement of the magnitude of the characteristic spreads, and suggests that much of the spread between model and observation may be due to this systematic underestimation. As suggested above, a possible reason for this may be the lack of representation of oxidation in nonprecipitating clouds. Modeling studies by others [*Langner and Rodhe*, 1991; *Pham et al.*, 1995; *Chin and Jacob*, 1996; *Feichter et al.*, 1996] indicate that 67 to 90% of sulfate derives from aqueous-phase oxidation; the present model, for ON86, gives 74% (B94). The fact that the underestimate is greatest in winter and least in summer would be consistent with insufficient oxidation in the model, due to greater fraction of this oxidation taking place in nonprecipitating clouds in winter than in the other seasons. The systematic increase in fraction of underestimates with increasing $\overline{MR_o}$, on the other hand, may be indicative of sample bias in the observed MRs that is greater in source regions. For 24-hour SO_2 the median ratio of modeled to observed MR is much closer to unity, 0.97 for the data set as a whole and ranging from 0.88 to 1.07 for the four simulations. This, together with the fact that the model-observation spread is less than the spread of the observations themselves, may suggest that the spread between modeled and observed MR is much more attributable to subgrid variation than is the case with sulfate. With SO_2 as with sulfate the fraction of model underestimates increases systematically with increasing $\overline{MR_o}$, suggestive of sample bias in the observed MRs in near-source regions. The systematic model underestimate of SO_2 in North America may likewise be attributable to nonrepresentative sampling.

In addition to examining the performance of the model over the domain as a whole and by large regional areas, we also examined, by correlation analyses and by comparisons of the time series, the ability of the model to represent the temporal variability at individual locations. We have already noted the skill of the model in representing the geographical distribution of mixing ratios, as measured by the coincidence of high and low MRs as a function of location in Figures 10 and 12. This skill would be seem to be a measure mainly of accuracy in the distribution of sources. In contrast, the variation of modeled MR at an individual location is influenced mainly by variation in the meteorological variables controlling the transport and removal of material, as is also the case in the real world. Hence the correlation between measured and modeled MRs, as measured by the correlation coefficient R and the fraction of the variance explained R^2, is a measure more of the skill of the meteorological driver of the model than of the accuracy of the representation of the chemistry and removal processes (under the assumption also that the observed MRs are representative). The same can be said for the accuracy of the temporal coincidence of the features in the time series comparisons. It would seem, however, if the meteorological variability is well represented by the model, as indicated by skill in the correlation or in the episodicity of the time series plots, then the further agreement or lack of agreement in magnitude of modeled and observed MR in the time series is a measure of accuracy in the representation of the chemistry and removal processes. In this respect we have already called attention to the numerous instances of fairly close quantitative match between modeled and observed MRs for sulfate indicated in Figure 10 in the great majority of the instances of significant correlations indicated in Plate 6. We consider the latter especially to be a strong indicator of model skill. These locations tend also to be characterized by a high fraction of the values of the ratio characteristic spread between modeled and observed MRs, $S_{m/o}$, within a rather small factor of unity. For SO_2 the situation of consistent quantitative match in the scatterplots and time series comparisons is much less frequent than for sulfate. We attribute this to the fact that SO_2, as a primary emitted species, is much more likely than sulfate to be influenced by local emissions, which cannot be well represented by the model.

A final concern that requires comment is the issue of primary sulfate. We have noted that for European emissions we took primary sulfate emissions to be 5%, consistent with European modelers, although this would seem to be an overestimate. We note that in the JF87 simulation, primary sulfate contributed over half the modeled sulfate in much of the model domain. This situation forces us to note that model skill in the European locations, especially in the JF87 simulation, may be due in part to a compensation of too low a rate of atmospheric conversion by too great a rate of primary sulfate emissions. In this regard we note that the distribution of the ratio of modeled to observed MR for sulfate $MR_m/\overline{MR_o}$, while indicative of model underestimation in all seasons and both regions, shows less underestimation in Europe (Table 6) than North America, especially so in JF87. These considerations suggest that the issue of primary sulfate, as well as that of representing SO_2 oxidation in nonprecipitating clouds, needs to be examined further in future work.

One further qualification of the present comparisons is that they are restricted to surface measurements; comparisons of the vertical structure of modeled and observed MRs are thus needed

for further evaluation of model performance. We anticipate using observational data sets taken during recent and future intensive field campaigns to expand the evaluation of the model performance. Comparison of modeled column burden with instantaneous aerosol optical depth determined by satellite observation [*Wagener et al.*, 1997] may also serve as a valuable means of model evaluation.

7. Summary and Conclusions

A three-dimensional Eulerian transport and transformation model has been applied to calculate SO_2 and sulfate mixing ratios (MRs) and sulfate wet deposition over the North Atlantic and adjacent continental regions for actual times and locations for specific 1-month periods in each of the four seasons. Modeled MRs of sulfate and SO_2 for the lowest model level (surface to ~65 m) for June 28 to July 31, 1986, October 14 to November 15, 1986, January 28 to February 28, 1987, and March 28 to April 30, 1987 were compared to observed MRs at the surface from monitoring networks in the United States, Canada, and Europe, and modeled amounts of wet-deposited sulfate (concentration times precipitation amount) were compared with deposition observed by networks in the United States and Canada.

A principal concern in these comparisons is the degree of representativeness in observational data sets available from monitoring networks arising from the subgrid temporal and spatial variation of the quantities. To assess this, we examined the short term temporal and spatial variation of the observed MRs. Within-location variation of the observations was measured by means of quantities which we denote as the ratio and difference characteristic spreads. The ratio characteristic spread emerges as a useful measure of this within-location variation whose magnitude is independent of the magnitude of the MRs themselves. Median values for the ratio characteristic spread of simultaneous observed MRs within a 1.125° × 1.125° location are approximately a factor of 1.5 for the sulfate 24-hour MRs, 2.2 for the SO_2 24-hour average MRs, and 2.5 for the SO_2 6-hour average MRs, with no major differences among the four simulations. The sense of the progression of these median values is attributed to the longer time interval averaged and to the greater spatial uniformity of a secondary versus a primary emitted species. This within-location variability qualifies the level of agreement that can be expected for the comparisons of modeled and observed MRs, especially for a primary emitted species such as SO_2 at locations influenced by proximate sources.

A variety of comparisons were carried out between model and observations to ascertain the performance of the model. The characteristic spreads between modeled and observed MRs are comparable to those of the observed MRs themselves, indicating that agreement of modeled and observed MRs is near or within the subgrid variation of the observed MRs, with the departure of the model from the observations substantially greater than the within-location spread of the observations mainly at the extreme values of these quantities. Agreement between modeled and observed MRs for sulfate is greater than for SO_2 probably because observed MRs for SO_2 are more influenced by proximate sources and because subgrid temporal and spatial variation is not well represented by the model; sulfate being a longer-lived and mainly secondary species is less influenced by proximate sources and subgrid variability. No improvement was indicated in the comparisons of the modeled and observed SO_2 6-hour MRs over those for the 24-hour MRs.

In all four simulations the distribution of the difference between modeled and observed MRs, a measure of overall model bias, peaks at low values, and these peaks contain substantial contributions from a wide range of values of observed MRs. However, the model generally underestimates sulfate MR, with the median ratio of observed to modeled MR ranging from 0.36 to 0.66 for the four simulation periods. The fraction of model underestimates increases with increasing observed MR. For SO_2 the median is much closer to unity (0.97), but the model overestimates at low observed MR and underestimates at high observed MR. These trends imply that the model produces MR fields that are smoother than indicated by the observations.

Correlations of observed versus modeled MRs are highly significant for the entire set or by simulation period or by large region (Europe, North America), but exhibit rather low values of R^2, 0.44 for 24-hour sulfate, 0.18 for 24-hour SO_2. At individual locations, for the individual simulations or for the composite of multiple simulations, considerable model skill was indicated in many instances. For sulfate, 76% of 203 correlations were significant at the 95% confidence level, and for SO_2, 51% of 526 correlations were significant. However, detailed examination of scatterplots and of time series of modeled and observed MRs suggests that correlation analysis is not a wholly satisfactory measure of model performance, as data may be at considerable variance from the one-to-one line and yet still yield a significant correlation, or alternatively, rather convincing model performance is evidenced in the time series comparisons, and yet the correlation is not significant. Examination of the distribution of the ratio characteristic spread between modeled and observed sulfate and SO_2 24-hour MRs at individual locations indicated that a high proportion of modeled MRs were within a factor of 2 of the observed MRs at many locations. Substantial departure at a number of locations is attributed to local source or terrain influences which are not accounted for by the model and/or which make the observed MRs nonrepresentative.

Comparisons of modeled and observed wet deposition amount reflect inaccuracies both in the location and amount of precipitation in the meteorological data used to drive the model and in representation of chemical and wet removal processes in the model. Because of inaccuracies in the forecast of precipitation events, comparison of the 24-hour cumulative modeled sulfate wet deposition with 24-hour observational data was possible in only 43% of the observed precipitation events; comparison of weekly cumulative modeled wet deposition with weekly composite observational data was possible in 66% of the observed events. Approximately one third of the modeled amounts of wet-deposited sulfate were within a factor of 2 of the observed amount, and more than half were within a factor of 3. The median ratio of modeled to observed wet deposition amount for the four simulations and two networks ranged from 0.47 to 2.2, with an overall median of 0.82, indicative of fairly low model bias by this measure.

In conclusion, we have demonstrated that by using observation-derived synoptic meteorological data we are able to directly and quantitatively compare model results with measured surface mixing ratios of sulfate and SO_2 and wet deposition of sulfate in daily averages, not just in monthly averages, as has been the common practice with large-scale models. Comparisons between modeled and observed MRs yield a variety of measures of the accuracy with which the model

reflects the magnitude, spatial distribution, and temporal episodicity of the observed sulfate and SO_2 MRs. Perhaps key among our findings is the fact that the modeled MRs agree with the observed MRs within or close to the subgrid variation of the observations themselves. Further, the near zero peak of the distributions of the difference between observed and modeled MR for both SO_2 and sulfate and the agreement of modeled and observed sulfate wet deposition amount lend additional support to the accuracy of the model. We take these findings to be encouraging especially with respect to the utility of incorporating this modeling approach in climate models to evaluate the radiative influence of atmospheric sulfate.

Appendix A: Measurement Program and Protocols

Data were obtained from several measurement programs as briefly described here; see also B94.

1. CAPMON, the Canadian monitoring network (R. Vet, Atmospheric Environment Services, Environment Canada, Toronto, Canada, personal communication, 1992), stations are located in areas with no proximate sources, as described by *Vet et al.* [1988]. Observations for SO_2 24-hour MR and sulfate 24-hour MR were made with filter packs. Twenty-four-hour sulfate wet deposition was determined from wet only collectors and rain gauge measurement of precipitation amount [*Vet et al.*, 1989, 1988; *Sirois and Friche*, 1992].

2. In the SCENES, Subregional Cooperative Electric Utility, Department of Defense, National Park Service, and U.S. Environmental Protection Agency Study, network [*Vasconcelos et al.*, 1994; *Mueller et al.*, 1986] in the southwestern United States (L. Vasconcelos, Washington University, St. Louis, Missouri, personal communication, 1995), 24-hour filter samples were taken approximately every third day in two size ranges, diameter < 2.5 and < 15 μm. In the present comparisons we used data from the < 2.5 μm fraction, appropriate for the accumulation mode sulfate addressed in this work. A single station at Glen Canyon, Arizona, reported hourly SO_2, measured by pulsed fluorescence.

3. At the New York State Department of Health stations at Mayville, New York and Whiteface Mountain, New York [*Husain and Dutkiewicz*, 1990] (L. Husain, New York State Department of Health, Albany, New York, personal communication, 1993), sulfate was measured by filter sampling on an approximately 24-hour basis.

4. In the Harvard six-cities study [*Ferris et al.*, 1986] (G. Allen, Harvard University, Cambridge, Massachusetts, personal communication, 1995), particulate matter was sampled in two size ranges: bdiameter < 2.5 and 2.5 to 15 μm. The analysis reported total sulfur by X ray fluorescence. Only fine particle measurements were used in the model comparisons because the great majority (over 85%) of sulfur was present in this range [*Spengler and Thurston*, 1983]; virtually all sulfur in this size range is present as sulfate [*Forrest and Newman*, 1977].

5. The Aerometric Information Retrieval System of the U.S. Environmental Protection Agency (EPA) [*U.S. Environmental Protection Agency*, 1988] includes monitoring data from the National Air Monitoring Stations (NAMS) and the State and Local Air Monitoring Stations (SLAMS) in the United States. Because the primary objective of this network is to assess population exposure, the majority of the stations are located in urban and suburban settings and thus tend to be influenced by proximate sources. Twenty-four-hour sulfate MRs were measured once every 6 days using glass fiber filter samples. Because sulfate MRs from high-volume sampling using glass fiber filters exhibit an offset of approximately 0.89 ppb relative to comparable measurements using Teflon filters, attributed to filter artifacts [*Lipfert*, 1994], we have subtracted this bias from the data for these networks. SO_2 MRs were measured continuously by a variety of instrumental methods such as pulsed fluorescence, UV-stimulated fluorescence, and to a lesser extent flame photometry, and coulometry and were reported hourly. The limit of detection for the first two methods has been determined to be between 0.5 to 1.0 ppb [*Boatman et al.*, 1988].

6. The AEROCE, Atmospheric Ocean Chemistry Experiment, station located in Barbados, data obtained from J. Prospero (University of Miami, Miami, Florida, personal communication, 1995), measured sulfate on an approximately 24-hour basis with samplers active only for oceanic wind directions; data were included in the comparisons when the sampling period spanned at least 18 and no more than 28 hours.

7. The European Modeling and Evaluation Programme (EMEP) [*Schaug et al.*, 1988; 1989] covers western and central Europe; data were obtained from the Norwegian Institute for Air Research and from B. Arends (Netherlands Energy Research Foundation, Petten, The Netherlands, personal communication, 1995). Although some EMEP stations are located in industrialized areas, for example, Suwalki and Jarczew in Poland, the general intent is that these stations be regionally representative and hence that they be situated to avoid direct impact of proximate sources [*Barrett and Berge*, 1996]. Twenty-four-hour sulfate MR was measured by filter packs and SO_2 24-hour MRs by absorbing solution or impregnated filters.

8. The United States National Atmospheric Deposition Program (NADP)/National Trends Network (NTN) [*National Atmospheric Deposition Program*, 1987, 1988] determined cumulative weekly sulfate wet deposition from wet only collectors and rain gauge measurement of precipitation amount.

Appendix B: Coincidence of Observed and Modeled Precipitation Events

The statistics of observed and modeled precipitation events are summarized in Table B1. The numbers of modeled and observed precipitation events were similar in all simulations, but there are considerable differences in the statistics for the two networks, the Canadian CAPMON network (24-hour samples) and the U.S. NTN network (weekly samples), with data for the CAPMON network indicating a greater number of observed events and the data for the NTN network indicating a greater number of modeled events. Overall, about 60% of the observed events had corresponding modeled events, and likewise about 60% of the modeled events showed corresponding observed events. The fact that these numbers are not 100% indicates displacement between observed and modeled precipitation. Examination of these numbers by season and network indicates considerable variability; note especially the low fraction of observed events captured by the model for the CAPMON network in JF87. Overall, the model exhibited apparent greater forecast skill for the NTN data, but this may be an artifact resulting from the longer period of comparison and the greater likelihood of chance coincidence of modeled and observed events.

Table B1. Number of Wet Deposition Events for which Deposition Amounts were Reported in Model Locations, and Fraction for which the Model Indicated Occurrence of Precipitation During Corresponding Periods; Likewise the Number of Instances for Which the Model Indicated Precipitation and the Fraction for Which Precipitation was Reported During Corresponding Periods

	All Simulations	June-July 1986	October-November 1986	January-February 1987	March-April 1987
Number of observed precipitation events					
Total	*2462*	*730*	*639*	*528*	*565*
CAN (24-hour)	802	233	224	178	167
NTN (7-day)	1660	497	415	350	398
Fraction with modeled precipitation, %					
Both	58.6	47.1	75.4	51.5	60.9
CAN	42.6	45.9	58.9	14.0	46.7
NTN	66.3	47.7	84.3	70.6	66.8
Number of modeled precipitation events					
Total	*2401*	*667*	*643*	*513*	*578*
CAN (24-hour)	443	159	146	38	100
NTN (7-day)	1958	508	497	475	478
Fraction with observed precipitation, %					
Both	60.0	51.6	75.0	53.0	59.2
CAN	77.2	67.3	90.4	65.8	78.0
NTN	56.2	46.7	70.4	52.0	55.7

Data are presented by observing network for the individual simulations and for the four simulations taken together. Occurrence of precipitation in model is defined as precipitation rate exceeding 1.5 mm in a 6-hour period (0.25 mm h^{-1} precipitation rate) within the 24-hour or 7-day sampling period for the Canadian CAPMON network (CAN, 19 locations) and U.S. National Trends Network (NTN, 173 locations), respectively.

Acknowledgments. We thank George Allen for supplying the Harvard Six-Cities Study sulfate data, Trond Iversen for supplying the SO_2 and sulfate data from the EMEP network, Brian Lamb for supplying the program to generate seasonal gridded values of biogenic land emissions, J. David Mobley for supplying the NAPAP emissions inventory for North America, Joseph Prospero for supplying data for Miami and Barbados, David Simpson for supplying the EMEP emissions inventory for Europe, Luis Vasconcelos for supplying the SO_2 and sulfate from the SCENES program, Robert Vet for supplying the SO_2, sulfate, and wet deposition data from the CAPMON network. Meteorological data used to drive the model were obtained from the European Centre for Medium-Range Weather Forecasts. We thank Gunnar Senum for discussions on statistical analysis. This research was supported by the Environmental Sciences Division of the U.S. Department of Energy as part of the Atmospheric Radiation Measurement Program and was performed under contract DE-AC02-76CH00016.

References

Andreae, M.O., Raising dust in the greenhouse, *Nature*, *380*, 389-390, 1996.

Barrett, K., and E. Berge, Transboundary air pollution in Europe MSC-W status report 1996 part one: Estimated dispersion of acidifying agents and of near surface ozone, *EMEP/MSC-W Rept 1/96*, Norw. Meteorol. Inst., Oslo, Norway, 1996.

Bates, T. S., B. K. Lamb, A. Guenther, J. Dignon, and R. E. Stoiber, Sulfur emissions to the atmosphere from natural sources, *J. Atmos. Chem.*, *14*, 315-337, 1992.

Benkovitz, C. M., C. M. Berkowitz, R. C. Easter, S. Nemesure, R. Wagener, and S. E. Schwartz, Sulfate over the North Atlantic and adjacent continental regions: Evaluation for October and November, 1986 using a three-dimensional model driven by observation-derived meteorology, *J. Geophys. Res.*, *99*, 20,725-20,756, 1994.

Berge, E., Preliminary estimates of sulphur transport and deposition in Europe with a regional scale multilayer Eulerian model, *EMEP/MSC-W Note 1/93*, Norw. Meteorol. Inst., Oslo, Norway, 1993.

Boatman, J. F., M. Luria, C. C. Van Valin, and D. L. Wellman, Continuous atmospheric sulfur gas measurements aboard an aircraft: A comparison between flame photometric and fluorescence methods, *Atmos. Environ.*, *22*, 1949-1955, 1988.

Bott, A., A positive definite advection scheme obtained by nonlinear renormalization of the advective fluxes, *Mon. Weather Rev.*, *117*, 1006-1015, 1989.

Boucher, O., and U. Lohmann, The sulfate-CCN-cloud albedo effect: A Sensitivity study with two general circulation models, *Tellus*, *Ser. B*, *47*, 281-399, 1995.

Charlson, R. J., J. Langner, H. Rodhe, C. B. Leovy, and S. G. Warren, Perturbation of the northern hemisphere radiative balance by backscattering from anthropogenic aerosols, *Tellus*, *Ser. AB*, *43*, 152-163, 1991.

Charlson, R., S. E. Schwartz, J. M. Hales, R. D. Cess, Jr. J. A. Coakley, J. E. Hansen, and D. J. Hoffman, Climate forcing by anthropogenic aerosols, *Science*, *255*, 423-430, 1992.

Chin, M., and D. J. Jacob, Anthropogenic and natural contributions to atmospheric sulfate: A global model analysis, *J. Geophys. Res.*, *101*, 18,691-18,699, 1996.

Chuang, C. C., J. E. Penner, K. E. Taylor, A. S. Grossman, and J. J. Walton, An assessment of the radiative effects of anthropogenic sulfate, *J. Geophys. Res.*, *102*, 3761-3778, 1997.

Cox, S. J., W.-C. Wang, and S. E. Schwartz, Climate response to radiative forcings by sulfate aerosols and greenhouse gases, *Geophys. Res. Lett.*, *22*, 2509-2512, 1995.

Dennis, R. L., J. N. McHenry, W. R. Barchet, F. S. Binkowski, and D. W. Byun, Correcting RADM's sulfate underprediction: Discovery and

correction of model errors and testing the corrections through comparisons against field data, *Atmos. Environ.*, *Part A*, *27*, 975-997, 1993.

Dietz, R. N., and R. F. Wieser, Sulfate formation in oil-fired power plant plumes, vol. 1, Parameters affecting primary sulfate emissions and a model for predicting emissions and plume opacity, *Rep. EPRI EA-3231*, Electr. Power Res. Inst., Palo Alto, Calif., 1983.

Easter, R. C., Two modified versions of Bott's positive definite numerical advection scheme, *Mon. Weather Rev.*, *121*, 297-304, 1993.

Eliassen, A., The OECD study of long range transport of air pollutants: Long range transport modelling, *Atmos. Environ.*, *12*, 479-487, 1978.

Eliassen, A., and J. Saltbones, Modelling of long-range transport of sulphur over Europe: A two year model run and some model experiments, *Atmos. Environ.*, *17*, 1457-1473., 1983.

Erickson, D. J. III, R. J. Oglesby, and S. Marshall, Climate response to indirect anthropogenic sulfate forcing, *Geophys. Res. Lett.*, *22*, 2017-2020, 1995.

European Centre for Medium-Range Weather Forecasts (ECMWF), *User Guide to ECMWF Products, Version 1.1*, Reading, England, 1988.

Feichter, J., E. Kjellström, H. Rodhe, F. Dentener, J. Lelieveld, and G.-J. Roelofs, Simulation of the tropospheric sulfur cycle in a global climate model, *Atmos. Environ.*, *30*, 1693-1707, 1996.

Feldman, G., et al., Ocean color: The availability of the global data set, *Eos Trans. AGU*, *70*, 634-641, 1989.

Ferris, B. G., Jr., J. H. Ware, J. D. Spengler, D. W. Dockery, and F. E. Speizer, The Harvard six-cities study, in *Aerosols: Research, Risk Assessment and Control Strategies*, edited by S. D. Lee, et al., pp. 721-730, Lewis, Chelsea, Mich., 1986.

Forrest, J., and L. Newman, Silver-110 microgram sulfate analysis for the short time resolution of ambient levels of sulfur aerosol, *Anal. Chem.*, *49*, 1579-1584, 1977.

Hansen, J., W. Rossow, and I. Fung, Long-term monitoring of global climate forcings and feedbacks, *Proc. Workshop NASA Goddard Institute for Space Studies*, NASA Conference Publication 3234, 1993. 91 pp.

Hass, H., A. Ebel, H. Feldmann, H. J. Jakobs, and M. Memmesheimer, Evaluation studies with a regional chemical transport model (EURAD) using air quality data from the EMEP monitoring network, *Atmos. Environ.*, *Part A*, *27*, 867-887, 1993.

Husain, L., and V. A. Dutkiewicz, A long-term (1975-1988) study of atmospheric SO_4^{2-}: Regional contributions and concentration trends, *Atmos. Environ.*, *Part A*, *24*, 1175-1187, 1990.

Intergovernmental Panel on Climate Change (IPCC), *Climate Change 1995 - The Science of Climate Change, Contribution of WGI to the Second Assessment Report of the Intergovernmental Panel on Climate Change*. Editors: J. T. Houghton, et al. (Cambridge University Press, New York, 1996). 572 pp.

Iversen, T., N. E. Halvorsen, J. Saltbones, and H. Sandnes, Calculated budgets for airborne sulphur and nitrogen in Europe, *Rep. EMEP/MSC-W 1/91*, Meteorol. Syn. Cent.-West, Norw. Meteorol. Inst., Oslo, Norway, 1990.

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

Kasibhatla, P., W. L. Chameides, and J. St. John, A three-dimensional global model investigation of seasonal variations in the atmospheric burden of anthropogenic sulfate aerosols, *J. Geophys. Res.*, *102*, 3737-3759, 1997.

Kiehl, J. T., and B. P. Briegleb, The relative roles of sulfate aerosols and greenhouse gases in climate forcing, *Science*, *260*, 311-314, 1993.

Kiehl, J., and H. Rodhe, Modeling geographical and seasonal forcing due to aerosols, in *Aerosol Forcing of Climate*, edited by R. J. Charlson, and J. Heintzenberg, pp. 281-296, John Wiley, New York, 1995.

Langner, J., and H. Rodhe, A global three-dimensional model of the tropospheric sulfur cycle, *J. Atmos. Chem.*, *13*, 225-263, 1991.

Langner, J., et al., The global atmospheric sulfur cycle: An evaluation of model predictions and observations, *Rep. CM-81*, Dep. Meteorol., Stockholm Univ., Int. Meteorol. Inst., Stockholm, 1993.

Leaitch, W. R., G. A. Isaac, J. W. Strapp, C. M. Banic, and H. A. Wiebe, The relationship between cloud droplet number concentrations and anthropogenic pollution: Observations and climatic implications, *J. Geophys. Res.*, *97*, 2463-2474, 1992.

Lipfert, F. W., Filter artifacts associated with particulate measurements: New evidence and effects on statistical relationships, *Atmos. Environ.*, *28*, 3233-3249, 1994.

Luecken, D. J., C. M. Berkowitz, and R. C. Easter, Use of a three-dimensional cloud-chemistry model to study the transatlantic transport of soluble sulfur species, *J. Geophys. Res.*, *96*, 22,477-22,490, 1991.

Macdonald, A. M., C. M. Banic, W. R. Leaitch, and K. J. Puckett, Evaluation of the Eulerian Acid Deposition and Oxidant Model (ADOM) with summer 1988 aircraft data, *Atmos. Environ.*, *Part A*, *27*, 1019-1034, 1993.

Mitchell, J. F. B., T. C. Johns, J. M. Gregory, and S. F. B. Tett, Climate response to increasing greenhouse gases and sulphate aerosols, *Nature*, *376*, 501-504, 1995.

Mueller, P. K., D. A. Hansen, and J. G. Watson Jr., The Subregional Cooperative Electric Utility, Department of Defense, National Park Service and EPA Study (SCENES) on visibility: An overview, *EA-4664-SR*, Electr. Power Res. Inst., Palo Alto, Calif., 1986.

National Atmospheric Deposition Program, NADP/NTN annual data summary, Precipitation chemistry in the United States, 1986, Natl. Atmos. Deposition Program/Natl. Trends Network, Coord. Off., Fort Collins, Colo., 1987.

National Atmospheric Deposition Program, NADP/NTN annual data summary, Precipitation chemistry in the United States, 1987, Natl. Atmos. Deposition Program/Natl. Trends Network, Coord. Off., Fort Collins, Colo., 1988.

National Research Council, Panel on Aerosol Radiative Forcing and Climate Change (J. Seinfeld, Chair), *A Plan for a Research Program on Aerosol Radiative Forcing and Climate Change*, Natl. Acad. Press, Washington, D. C., 1996.

Pasquill, F., The dispersion of material in the atmospheric boundary layer - The basis for generalization, in *Lectures on Air Pollution and Environmental Impact Analyses*, edited by D. A. Haugen, pp. 1-34, Am. Meteorol. Soc., Boston, Mass., 1976.

Penner, J., et al., Quantifying and minimizing uncertainty of climate forcing by anthropogenic aerosols, *Bull. Am. Meteorol. Soc.*, *75*, 375-400, 1994.

Pham, M., J.-F. Müller, G. P. Brasseur, C. Granier, and G. Mégie, A three-dimensional study of the tropospheric sulfur cycle, *J. Geophys. Res.*, *100*, 26,061-26,092, 1995.

Saeger, M., et al., The 1985 NAPAP emissions inventory (version 2): Development of the annual data and modelers' tapes, *Rep. EPA-600/7-89-012a*, U.S. Environ. Prot. Agency, Research Triangle Park, N. C., 1989.

Schaug, J., J. E. Skjelmoen, S. E. Walker, A. Harstad, K. Nodop, and J. Pacyna, Data report 1986, part 1: Annual summaries, *EMEP-CCC-Rep. 6/88, Ref. O-7727*, Norw. Inst. for Air Res. (NILU), Lillestrøm, Norway, 1988.

Schaug, J., J. E. Skjelmoen, S. E. Walker, U. Pedersen, and A. Harstad, Data report 1987, part 1: Annual summaries, *EMEP/CCC Rep. 1/89, Ref. O-7727*, Norw. Inst. for Air Res. (NILU), Lillestrøm, Norway, 1989.

Schwartz, S. E., The Whitehouse Effect--Shortwave radiative forcing of climate by anthropogenic aerosols: An overview, *J. Aerosol Sci.*, *27*, 359-382, 1996.

Schwartz, S. E., and M. O. Andreae, Uncertainty in climate change caused by aerosols, *Science*, *272*, 1121-1122, 1996.

Schwartz, S. E., and A. Slingo, Enhanced shortwave cloud radiative forcing due to anthropogenic aerosols, in *Clouds, Chemistry and Climate*, *NATO ASI Ser.*, edited by P. J. Crutzen and V. Ramanathan, pp. 191-236, Springer-Verlag, New York, 1996.

Shea, D. J., Climatological atlas: 1950-1979, Natl. Cent. for Atmos. Res., Boulder, Colo., 1986.

Sirois, A., and W. Fricke, Regionally representative daily air concentrations of acid related substances in Canada: 1983-1987, *Atmos. Environ.*, *Part A*, *26*, 593-607, 1992.

Slinn, W. G. H., Air-to-sea transfer of particles, in *Air-to-Sea Exchange of Gases and Particles*, edited by P. S. Liss, and W. G. N. Slinn, pp. 299-405, D. Reidel, Norwell, Mass., 1983.

Sokal, R. R., and F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research*, W. H. Freeman, New York, 1981.

Spengler, J. D., and G. D. Thurston, Mass and elemental composition of fine and coarse particles in six U.S. cities, *J. Air Pollut. Control Assoc.*, *33*, 1162-1171, 1983.

Spivakovsky, C. M., R. Yevich, J. A. Logan, S. C. Wofsy, M. B. McElroy, and M. J. Prather, Tropospheric OH in a three-dimensional

chemical tracer model: An assessment based on observations of CH_3CCl_3, *J. Geophys. Res.*, *95*, 18,441-18,471, 1990.

Taylor, K. E., and J. E. Penner, Response of the climate system to atmospheric aerosols and greenhouse gases, *Nature*, *369*, 734-737, 1994.

U.S. Environmental Protection Agency, *AIRS User's Guide*, Off. of Air Qual. Plann. and Stand., Research Triangle Park, N. C., 1988.

Vasconcelos, L. A., E. S. Macias, and W. H. White, Aerosol composition as a function of haze and humidity levels in the Southwestern U.S., *Atmos. Environ.*, *28* (22), 3679-3691, 1994.

Vet, R. J., W. B. Sukloff, M. E. Still, J. B. Martin, W. F. Kobelka, and A. J. Gaudenzi, Canadian Air and Precipitation Monitoring Network (CAPMON) precipitation chemistry data summary 1986, *Rep. AQRB-88-02*, Atmos. Environ. Service, Downsview, Ont., Canada, 1988.

Vet, R. J., W. B. Sukloff, M. E. Still, J. B. Martin, W. F. Kobelka, and A. J. Gaudenzi, Canadian Air and Precipitation Monitoring Network (CAPMON) precipitation chemistry data summary 1987, *Rep. AQRB-89-1*, Atmos. Environ. Serv., Downsview, Ont., Canada, 1989.

Wagener, R., S. Nemesure, and S. E. Schwartz, Aerosol optical depth over oceans: High space and time resolution retrieval and error budget from satellite radiometry, *J. Atmos. Oceanic Technol.*, *14*, 577-590, 1997.

Wesely, M., Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, *Atmos. Environ.*, *23*, 1293-1304, 1989.

Wigley, T. M. L., and S. C. B. Raper, Implications for climate and sea level of revised IPCC emissions scenarios, *Nature*, *357*, 293-300, 1992.

Wilson, M. F., and A. Henderson-Sellers, A global archive of land cover and soils data for use in general circulation models, *J. Clim.*, *5*, 119-143, 1985.

Yanenko, N. N., *The Method of Fractional Steps*, Springer-Verlag, New York, 1971.

Zimmerman, R. R., Testing for hydrocarbon emissions from vegetation leaf litter and aquatic surfaces, and development of a methodology for compiling biogenic emissions inventories, U.S. Environ. Prot. Agency, Research Triangle Park, N. C., 1979.

C. M. Benkovitz and S. E. Schwartz, Department of Applied Science, Brookhaven National Laboratory, Building 815E, Upton, NY 11973. (e-mail: cmb@bnl.gov; ses@bnl.gov)

(Received May 6, 1997; revised July 24, 1997; accepted July 31, 1997.)