An evaluation of TRACE-P emission inventories from China using a regional model and chemical measurements

Q. Tan,¹ W. L. Chameides,¹ D. Streets,² T. Wang,³ J. Xu,⁴ M. Bergin,¹ and J. Woo⁵

Received 29 May 2004; revised 31 August 2004; accepted 21 September 2004; published 20 November 2004.

[1] We evaluate the TRACE-P emission inventories for gaseous and particulate pollutants from East Asia using chemical measurements made at a rural site in China during the China-MAP Field Intensive in conjunction with a coupled regional climate/chemical transport modeling system. Time-dependent, three-dimensional fields for trace gas and particulate matter concentrations over East Asia are simulated by an updated version of the Regional Acid Deposition Model (RADM) driven by the TRACE-P emission inventories along with meteorology fields calculated by the NCAR Regional Climate Model (RegCM) for the month of November 1999. Model-calculated SO₂ is in good agreement with measurements, while CO and particulate carbon (PC) are significantly smaller, and particulate sulfate (SO₄²⁻) is somewhat smaller. Our calculations suggest that increases in the TRACE-P emission inventory of CO by 110–120% and PC by 60–90% would bring the model-calculated CO, PC, and particulate sulfate concentration into agreement with the China-MAP observations. If these increases were spread uniformly throughout China and the year, it would require that there be additional emissions in China of CO and PC of 60 Tg yr⁻¹ and 2.5–4 Tg yr⁻¹, respectively. Further analysis of high resolution gas species measurements suggests that the missing CO emissions are likely to be associated with SO₂ emissions from coal burning. This in turn suggests that coal-burning facilities in China are operating at significantly lower efficiencies than currently assumed.

INDEX TERMS: 0322 Atmospheric Composition and Structure: Constituent sources and sinks; 0345 Atmospheric Composition and Structure: Pollution—urban and regional (0305); 0365 Atmospheric Composition and Structure: Troposphere—composition and chemistry; 0368 Atmospheric Composition and Structure: Troposphere—constituent transport and chemistry; KEYWORDS: anthropogenic emissions, East Asia, regional modeling


1. Introduction

[2] East Asia is currently one of the major sources of anthropogenic pollutants globally. The rapid industrialization and urbanization of the region, and of China in particular, coupled with its high population density suggests that the region’s emission will grow in the coming decades [Intergovernmental Panel on Climate Change (IPCC), 2000]. The air pollutants from the region have already caused significant adverse environmental effects on local, regional, and perhaps global scales. Based on a World Health Organization report (1998), seven of the world’s ten most polluted cities are in China. Acid-rain, high loading of fine particles, and elevated ozone concentration have been reported over large portions of the region [Galloway et al., 1987; Wang and Wang, 1995; Chameides et al., 1999a; Streets et al., 1999; Bergin et al., 2001; Xu et al., 2002]. Transport of pollutants out of East Asian has also been well-documented and appears to be having significant impacts over the North Pacific Ocean, North America, and beyond [Andreae et al., 1988; Prospero and Savoie, 1989; Parrish et al., 1992; Talbot et al., 1997; Jaffe et al., 1997, 1999, 2003; Jacob et al., 1999; Liu et al., 2003; Jaegle et al., 2003].

[3] In order to assess the possible impacts of anthropogenic pollutants on local, regional, and global scales, an accurate and comprehensive emission inventory is needed. However, several recent studies indicate that significant inaccuracies may exist in the inventory for pollutant emissions from Asia in general and East Asia in particular, and especially for that of carbon monoxide (CO) [Kasibhatla et al., 2002; Pétron et al., 2002; Kiley et al., 2003; Palmer et al., 2003; Carmichael et al., 2003; Arellano et al., 2004;
Allen et al., 2004]. CO pollution arises primarily from the incomplete combustion of fossil fuel and biomass. It is also produced from the oxidation of anthropogenic and natural hydrocarbons. Since CO is the major sink of the hydroxyl radical (OH), the most important oxidant in the atmosphere, the abundance of CO will have large impacts on the lifetime of other air pollutants and greenhouse gases [Seinfeld and Pandis, 1998].

[4] Kasibhatla et al. [2002] (K02) and Pétron et al. [2002] (P02) applied an inverse modeling technique using a global chemical transport model, i.e., GEOS-CHEM (K02) and IMAGES (P02), and NOAA/CMDL background CO measurements. These investigators concluded that CO emissions estimated by the EDGAR/GEIA inventory [Oliver et al., 1996] from combustions of fossil and biofuels (FF+BF) and biomass (BM) in Asia are underestimated. K02 estimated that CO emissions from FF+BF and BM burning in Asia are underestimated by factors of 1.5 and 2 (i.e., ~120 and 60 Tg CO/yr), respectively. P02 estimated that FF+B and BM burning emissions in Asia are underestimated by factor of 2 and 1.25 (i.e., 265 and 18 Tg CO/yr), respectively. More recently, Kiley et al. [2003], Palmer et al. [2003], and Carmichael et al. [2003] concluded, on the basis of a comparison between global/regional model calculations and TRACE-P airborne measurements over the western North Pacific Ocean near the coast of China, that CO emissions from East Asia are underestimated by the TRACE-P inventory prepared by Streets et al. [2003]. (The Streets et al. [2003] inventory is similar to the Asian part of the GEIA inventory.) The posteriori CO emission inventory derived by Palmer et al. [2003] has an additional 60 Tg CO/yr of emissions from fossil and biofuel burning in China over that of the TRACE-P emission inventory; this represents a factor of 1.5 increase in CO emissions from these sectors. On the other hand, Carmichael et al. [2003] concluded that the underestimate in CO emissions arose from domestic sources in central China, and they estimated that these emissions need to be increased by a factor of 3–5 over that in the TRACE-P inventory. Arellano et al. [2004] used space-based CO measurements by MOPITT satellite in their inversion study and concluded that the CO emissions in East Asia are about a factor of 1.8–2.0 higher than current bottom-up scenarios. Allen et al. [2004] combined a modeling study with aircraft (TRACE-P) and satellite (MOPITT) measurements of CO during spring of 2001 and concluded that Asian FF/BF emissions should be increased by factor of 1.6 over that in current CO emission estimates.

[5] Thus, while there appears to be a consensus among previous investigators that CO emission from Asia in general, and particularly in East Asia are underestimated by a factor of about 1.5–2, there are significant differences among these investigators as to the identity of the sources that are responsible for these additional emissions. In this study we use simultaneous measurements of gas and particulate matter pollutants at Linan, a rural site in China, during November 1999 [Wang et al., 2002; Xu et al., 2002] in combination with a regional climate/chemistry model to evaluate the TRACE-P emission inventory for CO, SO$_2$, and particulate carbonaceous matter (PC).

[6] Our approach has a number of aspects that set it apart from the aforementioned studies. Kasibhatla et al. [2002] and Pétron et al. [2002] used a global model and globally-distributed measurements of CO at CMDL background monitoring sites. Palmer et al. [2003] also used a global model, but, instead of globally-distributed measurements, TRACE-P measurements made during a number of flights off the coast of China during February–April, 2001. Carmichael et al. [2003] also used the TRACE-P measurements, but in this case a regional model. The study of Kiley et al. [2003] is based on inter-comparison of four global and three regional models. Arellano et al. [2004] and Allen et al. [2004] used global CTM with the CO measurement by satellite (Allen et al. [2004] also used TRACE-P measurements). In our case, we use a regional model, similar to that of Carmichael et al. [2003], but, in contrast to all the other investigators, make use of more detailed and continuous chemical measurements from a site within one of the most populated and industrialized regions of China. While this might make our results more susceptible to local influences, it makes them less dependent on uncertainties related to the simulation of long-range transport and chemical transformations. It minimizes the confounding influence of emissions from other regions of the globe. Moreover, unlike the investigators who used the TRACE-P measurements, our approach makes use of an entire month of near-continuous observations. This perhaps makes our results less susceptible to the influences of episodic anomalies. Uncertainties related to the simulation of transport and meteorological phenomena are further reduced by the availability of simultaneous data on multiple gaseous and particulate species. This allows us to compare model-calculated and observed ratios of species concentrations in addition to their absolute concentrations, and concentration ratios should be less dependent upon transport and meteorological influences.

[7] As noted above, our analysis focuses on the month of November. We chose this month since this was the time when simultaneous measurements of both gaseous and particulate species were made at the Linan site, and the availability of both gaseous and particulate measurements provides a valuable constraint for evaluating the model results. However, there are other advantages to our November focus. Regional air pollution episodes over eastern China are frequent and severe during this month [Luo et al., 2000; Wang et al., 2002]. During November, the influence of dust storms and boreal fires tends to be at a minimum and the dominant wind flow out of the north-west minimizes the influx of pollutants from biomass burning in Southeast Asia. This allows us to use the data to focus on pollutant emissions from China. Moreover, the climatology over China during month of November tends to be characterized by relatively stable conditions and relatively little precipitation. This reduces model uncertainties related to the simulation of convective transport, cloud processing, and wet deposition.

2. Climate-Chemistry Modeling System
2.1. Model Description
[8] Our simulations make use of a modified version of the regional three-dimensional coupled climate-chemistry modeling system used by Chameides et al. [1999a, 1999b, 2002], Luo et al. [2000], and Giorgi et al. [2002]. The modeling system consists of two components: a climate
model, the NCAR Regional Climate Model (RegCM2) [Giorgi et al., 1993a, 1993b, 1999, 2002] which provides the meteorological fields that drive a chemical transport model, an enhanced version of Regional Acid Deposition Model (RADM) [Luo et al., 2000]. RegCM is built upon the PSU/NCAR MM4 and is equivalent to the hydrostatic version of the current MM5 with additional and modified versions of several parameterizations needed for long-term climate simulations. RADM was originally developed by Chang et al. [1987]. The version we used is based on the one used and evaluated for East Asia by Luo et al. [2000] with a number of updates and modifications as described below. More detailed description of the RegCM and RADM can be found in above references.

Our model domain covers the major part of East Asia, i.e., ~80–160°E in east-west direction and 10–55°N in the south-north direction (Figure 1) with 60 km × 60 km horizontal grid cells on a Lambert conformal projection and 15 vertical levels on σ coordinates that top-off at 80 hpa. Five of the vertical layers are within the planetary boundary layer. The RADM model domain is slightly smaller than that of RegCM to minimize near-lateral-boundary effects from the RegCM. Prior to simulating the gas and aerosol distribution of November, 1999, both RegCM and RADM are run for a two month spin-up period.

In this study we made several modifications to the version of RADM used by Luo et al. [2000]. The photochemical reaction rates have been updated based on Sander et al. [2003]. RADM now uses the same vertical resolution as RegCM, i.e., 14 layers instead of 6 as in Luo et al. [2000]. The dry deposition of species is calculated using a resistance-in-series module that accounts for seasonal variations in land-surface characteristics and the specific physical/chemical properties of the depositing tracer species [Walcek et al., 1986; Wesely, 1989; Walmsley and Wesely, 1996]. Improvements made to the original deposition scheme of Wesely [1989] include calculating the aerodynamic resistance based on Jacobson [1999] and Byun et al. [1999], the surface resistance over tundra based on Jacob et al. [1992], the surface resistance of NO2 based on Ganzeveld and Lelieveld [1995], and the surface resistance of SO2 over snow based on Valdez et al. [1987] and Ganzeveld et al. [1998]. Our model-calculated dry deposition velocities are within the range of observed velocities [Ganzeveld et al., 1998, and references therein] and comparable to other regional and global models results: Environmental Protection Agency (EPA) Model-3 [Hu et al., 2004], GO-CART [Chin et al., 2000], and ECHAM [Ganzeveld et al., 1998].

In the previous applications of the modeling system [e.g., Luo et al., 2000], RADM and RegCM used different land-cover/land-use data sets. In order to provide consistency between the models, the Global Land Cover Characterization (GLCC) 30° data set is used to derive the land surface

Figure 1. Spatial distribution in the annual average emission rates of CO, SO2, and particulate carbon (PC) over the model domain interpolated from the TRACE-P inventory [Streets et al., 2003].

Figure 2. Contributions of various source-types to the total SO2, CO, and Particulate C emissions from China in the TRACE-P emission inventory of Streets et al. [2003].
Figure 3
types for our model domain and used as inputs to both RegCM and RADM. (The two models require different land use categorizations, and so somewhat different aggregations were used.)

The lateral boundaries of our model domain are located in areas that are relatively far away from significant anthropogenic emissions. Moreover, Linan is located ~1700–2600 km away from the model’s lateral boundaries but adjacent to significant sources of anthropogenic emissions. Thus the chemical boundary conditions chosen for our model simulations generally have a negligible impact upon the pollutant concentrations calculated by our model for the grid containing the Linan site. The boundary conditions for air pollutants with relatively short life-times and small natural sources, such as SO₂ and carbonaceous aerosols, are set to negligibly small values. O₃ and CO on the other hand have relatively long life times, and thus can be transported into our model domain from distant sources.

In the case of O₃, the lateral boundary conditions are set to values normally associated with clean tropospheric air [Luo et al., 2000]. In the case of CO, we use a concentration of 100 ppbv based on the average CO concentration measured at a NOAA CMDL site in Qinghai Province, China (36N, 101E), i.e., 108 ppb for November, 1999 http://www.cmdl.noaa.gov/). The Qinghai site is the only CMDL site that close to our model lateral boundaries in the upwind direction. Sensitivity calculations with CO at the model’s lateral boundaries set at 200 ppbv (a CO concentration higher than that measured at all northern hemispheric CMDL background sites in November, 1999, except at Hegyhatsal, Hungary, inside of Europe), resulted in only a few ppbv’s increase in the CO calculated at Linan.

2.2. Emission Inventory

Estimates of anthropogenic emissions of air pollutants that were specifically focused on East Asia began in

---

**Figure 3.** Average temperature, T (K) and humidity, Qv (kg/kg) at the surface and 500 hpa during November 1999 from the NCEP Reanalysis data and the RegCM simulation.

**Figure 4.** Time series of temperature, T(K), humidity, Qv(kg/kg), wind vector, u and v (m/s), and wind speed, w (m/s) per 6 hours during November 1999, at Linan from measurements made at the site (red dashed line, T and Qv only), and from interpolation of the surrounding grids from the NCEP reanalysis data (black solid line) and RegCM simulations (blue dots).
the early 1990’s and have been continuously improved [Fujita et al., 1991; Kato and Akimoto, 1992; Akimoto and Narita, 1994; Arndt et al., 1997; Klimont et al., 2001; Streets and Waldhoff, 2000; Streets et al., 2000, 2001a, 2001b, 2001c]. Most recently, Streets et al. (2003) developed an emission inventory of gaseous and primary aerosols for the year 2000 for the NASA TRACE-P [Jacob et al., 2003] and the NSF/NOAA ACE-Asia experiments [Huebert et al., 2003]. It is one of the most comprehensive and updated analyses of anthropogenic emissions from East Asia to date. The inventory includes all major sources of pollutants for essentially the entire East Asian region on various spatial scales. In our model simulations, we used the TRACE-P inventories with resolutions of $1^\circ \times 1^\circ$ and the $30^\circ \times 30^\circ$, and obtained similar results. The calculations presented here used the $1^\circ \times 1^\circ$ inventory.

Figure 1 illustrates the spatial distribution of the TRACE-P CO, SO$_2$, and particulate carbonaceous (PC) aerosols emissions interpolated to our model grid and Figure 2 illustrates the total CO, SO$_2$, and PC emissions by source type or sector. The spatial distribution of the CO, SO$_2$, and PC emissions are quite similar, with large sources located near the major urban-industrial centers (e.g., Beijing, Shanghai, Chengdu). However, because significant amounts of CO and PC are emitted by biomass burning, domestic sources, and transport-related sources, their emissions tend to be somewhat more disperse than that of SO$_2$, whose emissions are dominated by point sources. This will turn out to be an important distinction in our later discussion on the identity of the additional sources of CO that we conclude are needed to rationalize model simulations with observations.

2.3. Meteorological Conditions

[15] The meteorological fields needed to drive the RADM simulations are provided by the RegCM. The large-scale boundary and initial conditions used to drive the RegCM simulations are obtained from the NCEP reanalysis data provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/. These data are available 4 times per day on a $2.5^\circ \times 2.5^\circ$ grid at the surface and at different pressure levels.

[16] Figure 3 illustrates the monthly averaged NCEP and RegCM simulated temperature (T) and humidity (Qv) fields at the surface and 500 hpa over our model domain for November 1999. In general, the RegCM simulations reproduced the monthly-averaged spatial distribution of the NCEP reanalysis data over both the surface and mid-troposphere. One noticeable difference is that the simulated temperature and humidity over central to south China are slightly lower than those of NCEP. These differences may
be due, at least in part, to differences in the topographies used for the NCEP and RegCM grids.

Figure 4 illustrates the time series of temperature, humidity, wind components (u and v), and wind speed (w) measured at the sampling site, Linan, during November, 1999, as well as the values appropriately interpolated to the Linan site’s location from NCEP-Reanalysis data and RegCM simulations. The RegCM simulated temperature and Qv generally follow the trend of the two sets of measurements. During early November, however, the RegCM temperature is several degrees lower than both measurements. RegCM winds agree with NCEP data quite well in both magnitude and directions with some fluctuations. One significant difference is on November 15, when the NCEP wind speed is much larger than that of RegCM simulation. We checked the NCEP data and found that there was a strong frontal system, with meridional winds (v) as high as 20 m/s, just off the coast of China, i.e., about 150–200 km away from Linan. Since the original NCEP data have a horizontal resolution of 2.5° (i.e., ~240 km), the high v-value at Linan in the NCEP data may reflect errors.

![Figure 6](image-url)

**Figure 6.** Model-calculated and observed pollutant concentrations as a function of time at Linan during November 1999. Concentrations for gas-phase species (CO and SO₂) are hourly averages, and concentrations of particulate species (SO₄ and Particulate C) are 24-hour averages. Observed concentrations were obtained from Wang et al. [2002] for gas phase species (black asterisks) and from Xu et al. [2002] for the particulate species (black and green asterisks). Model simulations are shown by red diamonds for gas phase species and red and blue diamonds for particulate species. (See Table 1 for detailed description about ‘Meas1’ and ‘Meas2.’)

![Table 1](image-url)

**Table 1.** Arithmetic Mean of Measured and Model-Simulated Gaseous and Particulate Species at Linan in November 1999

<table>
<thead>
<tr>
<th>Species</th>
<th>Measurementb</th>
<th>Modelc</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>763</td>
<td>587</td>
</tr>
<tr>
<td>O₃</td>
<td>33.4</td>
<td>24.7</td>
</tr>
<tr>
<td>SO₂</td>
<td>19.8</td>
<td>14.8</td>
</tr>
<tr>
<td>NOₓ</td>
<td>12.2</td>
<td>17.2</td>
</tr>
<tr>
<td>Meas_1</td>
<td>21.2</td>
<td>13.4</td>
</tr>
<tr>
<td>Meas_2</td>
<td>7.74</td>
<td>3.5</td>
</tr>
<tr>
<td>NH₄</td>
<td>8.56</td>
<td>6.05</td>
</tr>
<tr>
<td>PC</td>
<td>35.0</td>
<td>15.8</td>
</tr>
<tr>
<td>SO₄⁻</td>
<td>146.3</td>
<td>146.3</td>
</tr>
<tr>
<td>NO₃⁻</td>
<td>38.4</td>
<td>38.4</td>
</tr>
<tr>
<td>NH₄⁺</td>
<td>78.5</td>
<td>78.5</td>
</tr>
<tr>
<td>PC</td>
<td>15.8</td>
<td>15.8</td>
</tr>
</tbody>
</table>

*Units: CO, O₃, SO₂, and NOₓ: ppb; SO₄⁻, NO₃⁻, NH₄⁺, and PC: µg/m³.

*Monthly average of hourly measurement by Wang et al. [2002].


*Average of daily measurements by Xu et al. [2002], where ‘Meas_1’ and ‘Meas_2’ are two sets of measurements using different instruments and spanning slightly different time periods: ‘Meas_1’ included 33 24-hour filter samples taken over a period spanning October 29 to December 1, 1999; ‘Meas_2’ included 27 24-hour filter samples taken from November 4 to November 30, 1999. The filters in the two measurements started from a different time each day.

*Average of model simulations over the same time period as aforementioned ‘Meas_1’ and ‘Meas_2.’
produced from the interpolation of the coarsely gridded NCEP data to a specific point.

In summary, while episodic inconsistencies exist between the RegCM simulations and the meteorological observations, it appears that the RegCM provides a reasonable reproduction of average conditions during the month of November, 1999 (see Figure 3). This suggests that the RegCM-RADM modeling system used here will not be able to accurately simulate the day-to-day variations in pollutant concentrations at a specific grid point in the model domain. On the other hand, if other aspects of the modeling system are accurate, it should be able to do a reasonable job on longer weekly to monthly time scales. This was in fact found to be the case when this modeling system’s ability to simulate ozone pollution episodes over East Asia was evaluated by Luo et al. [2000].

2.4. Model Results

Figure 5 illustrates the monthly averaged CO, SO\(_2\), and PC concentrations simulated by RADM during November, 1999. Their spatial distributions generally follow the pattern of their emissions (Figure 1). In Figure 6, we plot the time series of gaseous and particulate species simulated by RADM at Linan, and Table 1 lists the monthly-averaged model simulated concentrations for a number of key gaseous pollutants. Figure 6 and Table 1 also include measured concentrations; these are discussed below.

3. Observations

An intense field campaign to monitor the concentrations of gaseous and aerosol species was carried out at a rural site, Linan, in the Yangtze Delta of China in November 1999. Gaseous species, including SO\(_2\), CO, NO\(_x\), and O\(_3\), were measured by Wang et al. [2002], and aerosol mass and composition were measured by Xu et al. [2002].

In this work we use the measurements of CO, SO\(_2\), and particulate SO\(_4^{2-}\) and carbon to compare to our model simulations in order to assess the accuracy of the TRACE-P emissions inventory for CO, SO\(_2\), and biomass aerosols. Because of uncertainties in the measurements of
concentrations of CO, SO2, NO, and O3 in Linan are considered a rural site, it often experiences air pollution episodes that cause significant spikes in the concentrations of CO, SO2, NOx, and particulate matter [Chameides et al., 1999a; Luo et al., 2000; Wang et al., 2002; Xu et al., 2002]. This was clearly the case in November, 1999. As illustrated in Figure 6, the period was marked by a number of occasions when CO exceeded 1 ppmv and SO2 exceeded 50 ppbv.

[21] The monthly average concentrations of pollutants at Linan were also quite high (see Table 1). The average CO at Linan was ~760 ppb for November 1999. By contrast the average CO at 36 northern hemisphere NOAA/CMDL background sites is only 111 ppb for the same month (http://www.cmdl.noaa.gov/). The average SO2 of 19.8 ppb at Linan is higher than averaged annual mean SO2 concentration measured at urban sites in United States (http://www.epa.gov/airtrends/index.html). At the same time, the PM2.5 and PM10 mean concentration measured at Linan of 90 μg/m3 and 98 μg/m3 are 5 and 1 times higher than the U.S. EPA NAAQS (National Ambient Air Quality Standards) promulgated to protect human health, i.e., 15 μg/m3 for PM2.5 and 50 μg/m3 for PM10 respectively (http://www.epa.gov/air/criteria.html). Thus the measurements from Linan provide a database that has been significantly impacted by anthropogenic emissions from China and, as a result, well-suited for analyzing the accuracy of pollutant emission inventories for the region.

4. Model-Measurement Comparison

[24] Inspection of Figure 6 reveals a general qualitative similarity between the model-calculated and observed temporal variations in the concentrations of the gaseous and particulate species of interest during November, 1999. On the other hand, while both observations and model-calculated species exhibit significant day-to-day variability, the timing of the observed concentration spikes does not always correspond to the timing of the model-calculated spikes. Moreover, the model-calculated concentrations generally vary more smoothly with time

Figure 8. Distribution of percentiles (10%, 25%, 50%, 75%, and 90%) of measured [Wang et al., 2002] and model-calculated CO and SO2 concentrations in November 1999 at Linan, where ‘Meas_all’ and ‘Model_all’ are derived from measurements and model simulations for the whole month, ‘Meas_photo’ and ‘Model_photo’ are restricted to measurements made in photochemically aged air (defined in the text) and the model simulations for the local afternoon.

Table 2. Measured and Model-Simulated Geometric Mean of Gaseous Species Concentration (ppb) and CO to SO2 Ratios at Linan, China, in November 1999

<table>
<thead>
<tr>
<th>Species</th>
<th>Measurement</th>
<th>Model</th>
<th>Diff</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>683</td>
<td>479</td>
<td>-30</td>
<td></td>
</tr>
<tr>
<td>O3</td>
<td>44</td>
<td>38</td>
<td>-14</td>
<td></td>
</tr>
<tr>
<td>SO2</td>
<td>11.5</td>
<td>11.5</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>NOx</td>
<td>10.1</td>
<td>13.2</td>
<td>30.7</td>
<td></td>
</tr>
<tr>
<td>CO: SO2</td>
<td>59</td>
<td>42</td>
<td>-28.8</td>
<td></td>
</tr>
<tr>
<td>ΔCO: SO2</td>
<td>50</td>
<td>33</td>
<td>-34</td>
<td></td>
</tr>
</tbody>
</table>

*Selected photochemically aged air from measurements by Wang et al. [2002] (see text for details).
*Model simulations for the local time 12:00–18:00.
*Diff = (Model − Measurement)/Measurement *100.
*Monthly geometric mean of measured ozone (12:00–18:00).
*Because of uncertainties in the measurements of reactive nitrogen as discussed in Wang et al. [2003; see also NARSTO, 2000], we do not consider the NOx measurements made at Linan and do not address issues related to the emission inventory for NOx.
*ΔCO: CO increment with respect to the background concentration. Measured CO background level is 111 ppb based on monthly average of CO measured at 36 North Hemisphere NOAA/CMDL background sites in November 1999; model uses 100 ppb as background CO level. ΔCO is included as a way of demonstrating that our conclusion that CO is underestimated by the model is robust and not affected by our choice of background concentration.
than do the observed concentrations. As discussed in detail by Luo et al. [2000], this is to be expected when comparing day-to-day variations in observations at a specific locale with concentrations calculated using a regional modeling system of the sort used here. In this work, therefore, we focus on a more appropriate comparison: namely between monthly-averaged observed and model-calculated concentrations.

[23] Inspection of Table 1 indicates significant differences in the simulated and observed monthly-averaged concentrations at Linan, with the model results well below the observations in all but one case (NOX). A comparison of the percentile distributions in the modeled and measured concentrations (Figures 8a and 8b) indicates that much of the model-calculated underestimate arises from the fact that there were a significant number of observations with especially high concentrations that are not reflected in the model simulations. It is quite possible that these discrepancies, especially for the higher percentile concentrations, are due, at least in part, to the incommensurability between the model-calculated and measured concentrations [NARSTO, 2000]. While, in our case, the model-calculated concentrations represent averages over a 60 km × 60 km grid, the observations are at a single point in space and are thus subject to short-term perturbations caused by local emissions and sub-grid scale meteorological processes. To limit these influences, we made two adjustments to the method used to calculate the observed averaged concentrations: 1. we use geometric means instead of arithmetic means; and 2. we limit the comparison of gaseous species to the afternoon hours when the air at the sampling site is ‘photochemically aged’ (i.e., when O3 >35 ppbv) [Trainer et al., 1993; Parrish et al., 1993]. (Since Linan is a rural site, there is little afternoon titration of O3 by local NO sources, and as a result ~80% of the afternoon data collected at Linan qualify as being photochemically aged.) The use of geometric means (measure of central tendency) tends to limit the influence of short-term concentration spikes on the resulting averages, and the requirement for photochemically-aged air during the afternoon increases the probability that the air being sampled was well-mixed and had not been recently impacted by direct pollutant emissions and is thus more representative of the area surrounding the sampling site [Trainer et al., 1993; Parrish et al., 1993]. (Note that in making the comparison between model-calculated and measured particulate matter concentrations we could not limit the comparison to the afternoon hours or to photochemically aged air, since the measurements are based on filter samples collected over 24-hour intervals.)

[26] Finally to limit the influence of model uncertainties related to the simulation of the boundary layer and transport, we compare model-calculated and observed ratios of species’ concentrations as well as the absolute species concentrations. Since chemical species are all impacted by boundary-layer mixing and transport processes equally, the use of ratios should tend to normalize out model errors arising from the inaccurate simulation of these processes and thereby allow us to focus on inaccuracies related to the emission inventories [see, e.g., Parrish et al., 1991; Chameides et al., 1992].

[27] A comparison of the resulting monthly geometric means of measured and model-calculated concentrations and concentration ratios of gaseous and particulate species are presented in Table 2 and Table 3, respectively. In the case of the comparison of particulate concentrations in Table 3, comparisons with two sets of measurements carried out using somewhat different instrumentation and spanning somewhat different time periods: Meas_1 involved analyses of 33 24-hour filter samples taken over a period spanning Oct. 29–Dec. 1, 1999 while Meas_2 involved analyses of 27 24-hour filter samples taken from Nov. 4–Nov. 30, 1999.

[28] Inspection of the tables reveals that while the model-simulated monthly average SO2 was quite close to the observations, the model-simulated averages for CO, particulate SO4−2 and particulate C are all significantly underestimated. Inspection of Figure 8 reveals that modeled CO concentrations at each percentile range (10, 25, 50, 75, and 90) are smaller than that of the measurements while the percentiles of modeled SO2 all fall within the range of measured values. When we compare the ratio of CO-to-SO2, we find that the model-calculated ratio is ~20% smaller than that of the measurements.

[29] In the case of the comparison of particulate concentrations, we find better agreement between the model and Meas_2 than between the model and Meas_1. However, in

---

**Table 3. Comparison of Model-Calculated Geometric Mean of Particle Concentration (µg/m³) With Measurements Made at Linan, China, in November 1999**

<table>
<thead>
<tr>
<th></th>
<th>Meas 1*</th>
<th>Model</th>
<th>Diff, %</th>
<th>Meas 2*</th>
<th>Model</th>
<th>Diff, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO4²⁻</td>
<td>17.7</td>
<td>10.1</td>
<td>−43</td>
<td>15.0</td>
<td>12.1</td>
<td>−19</td>
</tr>
<tr>
<td>NO³⁻</td>
<td>5.4</td>
<td>3.3</td>
<td>−40</td>
<td>4.8</td>
<td>3.8</td>
<td>−22</td>
</tr>
<tr>
<td>NH4⁺</td>
<td>7.1</td>
<td>4.9</td>
<td>−31</td>
<td>6.8</td>
<td>5.7</td>
<td>−16</td>
</tr>
<tr>
<td>Particulate C</td>
<td>29.2</td>
<td>14.3</td>
<td>−51</td>
<td>24.6</td>
<td>15.6</td>
<td>−37</td>
</tr>
<tr>
<td>SO2⁻² : PC</td>
<td>0.61</td>
<td>0.71</td>
<td>16</td>
<td>0.61</td>
<td>0.78</td>
<td>28</td>
</tr>
</tbody>
</table>

*From Xu et al. [2002]. 'Meas_1' and 'Meas_2' are described in Table 1.

**Table 4. Description of Various Test Runs**

<table>
<thead>
<tr>
<th>Test Run</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>STND MOD</td>
<td>simulations driven by TRACE-P emissions inventory [Streets et al., 2003]</td>
</tr>
<tr>
<td>2 CO MOD</td>
<td>simulations driven by doubled CO emissions</td>
</tr>
<tr>
<td>2 PC MOD</td>
<td>simulations driven by doubled PC emissions</td>
</tr>
<tr>
<td>3 PC MOD</td>
<td>simulations driven by tripled PC emissions</td>
</tr>
</tbody>
</table>
both cases, the underestimate in PC is larger than that of particulate SO$_4^{2-}$, and the model-calculated ratio of particulate SO$_4^{2-}$ to-PC ratio is ~15–30% higher than that of the measurements.

### 4.1. Test Run With Increased Emissions

A modeling system such as the one used here contains a variety of uncertain inputs and parameters that could in principle give rise to the inconsistencies between model-calculated and observed concentrations and concentration ratios discussed above. Examples, include the description of landuse and land cover, the parameterizations for dry and wet deposition, the photochemical mechanism, and boundary layer mixing, as well the pollutant emission rates. We have attempted to examine the sensitivity of our model results to many of these uncertainties, and, with the exception of the emission rates, were unable to find a reasonable adjustment in model inputs and parameterizations that was able to uniformly improve model performance with respect to the observations at Linan. For example, while downward adjustments in OH concentrations can lead to model-calculated CO concentrations that are closer to the observations, they make the simulated SO$_2$ and SO$_4^{2-}$ concentrations further from the observations. Similarly adjustments to the boundary layer mixing heights in the model in order to improve CO, particulate SO$_4^{2-}$, and PC would have little effect on the CO:SO$_2$ and particulate SO$_4^{2-}$:PC ratios and would worsen the simulated SO$_2$. Increases in dry deposition velocities improve the simulated CO:SO$_2$ ratio, but do not improve the simulated CO concentration and worsen the simulated concentrations for SO$_2$, particulate SO$_4^{2-}$ and PC.

Of all the sensitivity tests we carried out, the only adjustment that uniformly improved the model-predicted concentrations of all relevant species and species ratios were the tests involving an increase in the emissions of CO and PC. To illustrate this, we carried out three additional model runs: ‘2 CO MOD’ with CO emissions increased uniformly by factor 2 and ‘2 PC MOD’ and ‘3 PC MOD’ with PC emissions increased by factors of 2 and 3, respectively (Table 4). The results are summarized in Table 5. Not surprisingly, since CO and PC are primary pollutants an increase in the emissions of these species leads to an approximate proportionate increase in their model-simulated concentrations, bringing them more in line with the observations. An interesting and somewhat unexpected result is the fact that increasing the PC emissions also leads to an increase in the model-simulated particulate SO$_4^{2-}$ concentrations and brings it closer to the measured value as well. This occurs because the model includes a heterogeneous pathway for converting SO$_2$ to particulate SO$_4^{2-}$ [see Chameides et al., 2002].

### Table 5. Comparison of Simulated Geometric Mean of the Gas and Aerosol Concentration Together With Their Ratios From Different Test Runs With Geometric Mean of Measured Photochemically Aged Air

<table>
<thead>
<tr>
<th></th>
<th>Meas STND MOD</th>
<th>2 CO MOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO, ppb</td>
<td>684</td>
<td>848</td>
</tr>
<tr>
<td>SO$_2$, ppb</td>
<td>11.5</td>
<td>10.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Meas STND MOD</th>
<th>2 PC MOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO$_2$, ppb</td>
<td>11.5</td>
<td>10.8</td>
</tr>
<tr>
<td>Sulfate, µg/m$^3$</td>
<td>17.7</td>
<td>29.2</td>
</tr>
<tr>
<td>PC, µgC/m$^3$</td>
<td>10.2</td>
<td>14.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Meas STND MOD</th>
<th>2 PC MOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfate, µg/m$^3$</td>
<td>15.0</td>
<td>24.6</td>
</tr>
<tr>
<td>PC, µgC/m$^3$</td>
<td>12.2</td>
<td>15.6</td>
</tr>
</tbody>
</table>

### Table 6. Percentage of Increase in Model-Calculated CO and Particulate C (PC) Concentration at Linan for Different Model Test Runs

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Run</td>
<td>%$^a$</td>
<td>%$^a$</td>
</tr>
<tr>
<td>2 CO MOD$^b$</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2 CO in Subregion A$^c$</td>
<td>75</td>
<td>2 PC in Subregion A$^c$ 76</td>
</tr>
<tr>
<td>2 CO in Subregion B</td>
<td>89</td>
<td>2 PC in Subregion B 91</td>
</tr>
<tr>
<td>2 CO in Subregion C</td>
<td>96</td>
<td>2 PC in Subregion C 96</td>
</tr>
<tr>
<td>2 CO in Subregion D</td>
<td>91</td>
<td>2 PC in Subregion D 92</td>
</tr>
</tbody>
</table>

$^a$Relative to increment obtained from ‘2 CO MOD’ or ‘2 PC MOD’ test runs.
$^b$Test runs that are defined in Table 4.
$^c$Test runs with emissions of CO or PC doubled just in subregions ‘A,’ ‘B,’ ‘C,’ and ‘D,’ which are defined in Figure 7.
and, thus an increase in the concentration of PC increases the rate of production of particulate $SO_4^{2-}$ and, in turn, its concentration.

Because the CO and PC concentrations vary in an approximate linear fashion with their emissions, we can use the results summarized in Table 5 to estimate the percentage increase in CO and PC emissions that would be needed to bring the model-simulations into agreement with the observations. For CO, the estimated increase is about 52% (i.e., a factor of 1.52 over the TRACE-P emissions). For PC, the increase is 100% for Meas_1 and 60% for Meas_2. If the increase occurred uniformly throughout China and throughout the year, it would correspond to 61 Tg of CO yr$^{-1}$ and 2.5–4.5 Tg of PC yr$^{-1}$. The underestimates quantified based on this study is within the range of uncertainties suggested by Streets et al. [2003] for CO ($\pm$185%), BC ($\pm$360%), and OC ($\pm$450%).

![Figure 9](image1.png)

**Figure 9.** Scatterplot of hourly averaged observed and model-calculated CO and SO$_2$ concentrations in photochemically aged air during the afternoon hours in November 1999. Model results are illustrated for the standard model run (STND MOD) and test run with CO emissions doubled (2CO MOD). Also indicated are linear regressions of CO-SO$_2$ for each data set. The 2-day back trajectory for each of the observed data points highlighted with a blue cycles is illustrated in Figure 10.

![Figure 10](image2.png)

**Figure 10.** Two-day back-trajectories for selected air masses (highlighted by blue cycles in Figure 9) passing over Linan. The back trajectories were calculated using the NOAA ARL HYSPLIT model http://www.arl.noaa.gov/ready/hysplit4.html and meteorological data from NCEP/Reanalysis data. Each asterisk along a trajectory denotes 12 hours. Also indicated are locations of power plants in China (circles, with the size of the circles being proportional to the capacity of the plant), and the metropolitan areas of Shanghai and Hangzhou (triangles).
[33] Our estimate, 176 Tg/yr of total CO emissions from China, is similar to the posterior scenario suggested by Palmer et al. [2003], i.e., 180 Tg/yr. The CO emission from biomass burning in China is estimated to be about 12–15 Tg/yr [cf. Streets et al., 2003; Palmer et al., 2003], and thus we would estimate a total CO emission rate from FF+BF burning in China of ~161–164 Tg/yr. When compared with previous studies summarized by Allen et al. [2004], in which the posterior scenarios of whole Asia are scaled to China by a factor of 0.48 [Palmer et al., 2003], our estimate falls at the lower end of these previous estimates: 262 Tg/yr CO [Petron et al., 2002], 208 Tg/yr [Arellano et al., 2004]; 198 Tg/yr [Carmichael et al., 2003], 175 Tg/yr [Kasibhatla et al., 2002], 168 Tg/yr [Palmer et al., 2003], and 145 Tg/yr [Allen et al., 2004].

4.2. Estimate of the Spatial Extent of the Required Emissions Increase

[34] Based on the above analysis, we conclude that the CO and PC emissions in the Streets et al. [2003] inventory are underestimated by about 50 and 60–100% respectively. We then estimated the magnitude of the emissions increase that would be needed if that increase were spread over the whole domain. However, our analysis is based on a comparison with measurements at one site, and it may not be appropriate to conclude that the underestimation we inferred from these measurements is applicable to the whole domain. In this section we present model simulations in which we increased CO and PC emissions over four subregions of the model domain surrounding Linan to determine the minimum spatial extent required for the emissions increase. The subregions are in Figure 7: Subregion A comprises a 600 km × 600 km box with Linan at the center; Subregion B a 1200 km × 1200 km box with Linan at the center; Subregion C a 600 km × 1800 km rectangle that includes emissions from Beijing, a major source of pollutants; and subregion D a 2000 km × 600 km rectangle that includes emissions from the Sichuan Basin but not Beijing.

[35] Table 6 summarizes the results of a series of model runs in which CO emissions and PC were doubled in either the whole model domain or in one of the subregions. We find that increased emissions in Subregion A are responsible for only about 75% of the model-calculated increases in CO and PC concentrations in Linan obtained from proportionate increases in emissions throughout the whole domain. Further significant increases in the model-calculated concentrations at Linan are obtained by including portions of the model domain to the north of Linan, including Beijing (i.e., Subregions B and C). Inclusion of emission to the west had a negligible effect (compare results for Subregion D and Subregion B). These results suggest that either: 1. The required emission increases cover a fairly large portion of China (i.e., virtually the entire eastern half of China) although they do not necessarily have to cover all of China and may not include the Sichuan basin; or 2. The emissions increases are constrained to a smaller area around Linan. For example if the increases are constrained to the 360,000 km² area around Linan (Subregion A in Figure 7), then they would need to be 30–40% larger than inferred from the extrapolation of our model calculations to all of China.

5. Characteristics of the Missing CO Sources

[36] In this section we use the high resolution, continuous measurements of CO, SO₂, and NOₓ at Linan to gain insights into the characteristics of the missing CO source or sources that we concluded in the previous section are not accounted for in the TRACE-P emissions inventory. (Because particulate matter measurements at Linan were not continuous (i.e., 24-hour filter samples collected on selected days), a similar analysis is not possible for PC emissions.)

5.1. Analyses Based on 1-Hour Averaged Data

[37] We begin by considering a scatter plot of the measured and model-calculated hourly CO and SO₂ concentrations at Linan (Figure 9). Inspection of the regression lines also included in the figure reveals that the CO-SO₂ regression line for the standard model (STND MOD) generally underestimates CO for a given SO₂ concentration. The regression line for the 2 CO Test run, on the other hand, overestimates CO. This is consistent with our earlier conclusion that CO emissions need to be increased by a factor of ~2.5 to bring the model-results into agreement with the observations.

[38] It is also interesting to note that the observation data points tend to fall into two groups: one group that tends to fall close to the regression line from the Standard Model run, and the other that tends to fall closer to the regression line from the 2 CO Test run. This suggests, perhaps, that (1) there were times when the sampling site was impacted by a set of CO and SO₂ sources whose magnitudes were consistent with the TRACE-P emissions inventory used in the standard model; but (2) there were other times when some additional source or sources of CO not accounted for in the TRACE-P emissions inventory were impacting the sampling site. It is for this second set of points that an enhanced CO emissions inventory is needed in the model.

[39] We next consider a subset of the observed data points plotted in Figure 9: those data points, 17 in all, are highlighted by the blue circles in the figure. These data points all have CO concentrations greater than 1000 ppbv and they all fall close to the 2CO test run regression line in Figure 9. Thus the blue-highlighted data points all appear to have occurred during periods when the Linan site was

Figure 12. The probability distribution function (PDF) of the running correlation coefficients of randomly mixed correlated and noncorrelated synthetic data sets. Here ‘n’ is the window size by which the correlation coefficient is calculated.
impacted by the hypothesized missing CO source(s). In Figure 10 we plot the 2-day back-trajectories for these air masses that passed over the Linan site when these data points measured. Inspection of the figure reveals that 12 of the 17 trajectories passed over the urban-industrial areas of Shanghai or Hangzhou; most of them also passed near one or more of the coal-fired power plants in the region. This suggests that these high CO events that our model was not able to reproduce were caused by emissions associated with urban/industrial activities opposed to biomass burning.

[40] To further investigate the identity of the hypothesized missing CO source, we next plot the modeled error in CO (i.e., observed CO - calculated CO) as a function of the observed SO2 concentration (Figure 11). Once again we find that the points tend to fall into two distributions: one in which the modeled CO error is small and essentially independent of SO2 and the other in which the modeled CO error is positively correlated with observed SO2. Most significantly we find that the largest underestimation in model-calculated CO occurred for data points with high SO2 concentrations. This suggests that the hypothesized CO source(s) is not only associated with urban/industrial activities, but, more specifically, may be spatially and temporally linked to sources of SO2.

5.2. Statistical Analyses of High-Resolution Measurements

[41] The arguments presented above suggest that the missing CO source may be associated with SO2 emissions. To explore this possibility further we turn to a statistical analysis of the high-resolution analysis (i.e., 1-minute averaged as opposed to hourly-averaged) measurements from Linan. By using the 1-minute averaged data we are able to isolate individual plumes that impact the Linan site and investigate how CO and SO2 co-vary within these plumes.

[42] Table 7 summarizes the results of a multivariate analysis of the 1-minute average CO, SO2, and NOx measurements from November, 1999. About 96% of the variability in these three variables can be explained by two factors: one primarily associated with variations in CO and NOx, and one primarily associated with variations in SO2 and NOx. Nevertheless there is some mixing of CO and SO2 loadings, with a small but significant loading of CO in the factor associated with SO2 and vice versa. Thus the multivariate analysis confirms that there is a small subset of the high-resolution data when CO and SO2 are correlated.

[43] However, the multivariate analysis does not reveal whether there is a consistent distinction between the CO concentrations observed at Linan when CO is correlated with SO2 and when CO is not correlated with SO2, and, if so, the nature of that distinction. To examine this possibility, we used a statistical approach referred to here as the ‘running correlation method’. In this method, we successively calculated the correlation coefficient between the 1-minute CO and SO2 measurements obtained over the time period from t – n/2 to t + n/2, where n is the number of minutes of the time interval considered and t is varied to cover the entire month of the data set (Figure 12). The probability distribution function (PDF) of the correlation coefficients is then plotted together with the average concentrations of CO and SO2 obtained for that corresponding correlation coefficient bin (Figure 13). Time intervals range from n = 5 minutes to n = 240 minutes (i.e., 4 hours).

5.2.1. Illustration of Running Correlation Using Synthetic Data Set

[44] Before discussing the results of the running correlation analysis of the CO and SO2 data it is useful to illustrate the results of this analysis using a synthetic data set with known properties. First we generated three random data series, namely ‘A’, ‘B’, and ‘C’, where r(A, B) = r(A, C) = r(B, C) = 0 and r(D, D’) is the correlation coefficient between the D and D’ data series. Then we randomly mixed ‘A’ and ‘B’ to form a new data series ‘X’, and randomly mixed ‘A’ and ‘C’ to form data series ‘Y’. Thus ‘X’ and ‘Y’ are mixtures of a correlated data series (A vs. A) and a non-correlated data series (B vs. C). We adopted a number of different methodologies for mixing the data series to obtain ‘X’ and ‘Y’. We mixed the data series using fixed blocks or intervals of data from ‘A’ and then from ‘B’ and ‘C’, and we also used randomly selected intervals or blocks of data. The analysis shown here was obtained using a fixed interval of 51 with one-third of the intervals containing data from the ‘A’ series and two-thirds from ‘B’ and ‘C’. Analyses using different intervals (or randomly selected intervals) and different mixing fractions yielded similar results.

[45] After forming the ‘X’ and ‘Y’ data series, we then calculated the running correlation between ‘X’ and ‘Y’ as a function of n, the window size (i.e., the number of data points included for calculating each correlation coefficient). The resulting PDF’s for various values of n are illustrated in Figure 12. Note that the PDF of the correlation coefficients between ‘X’ and ‘Y’ has two peaks, one around 0 and the other at 1. These two peaks correspond the non-correlated and correlated components of the synthetic data series ‘X’ and ‘Y’. Thus we see that the running correlation method is able to separate a mixed data series into its correlated and non-correlated components.

5.2.2. Application of Running Correlation to Linan Data Set

[46] Ideally, we would use the results obtained using a small value of n, since this would allow us to focus on CO and SO2 variations in individual plumes that presumably arose from specific source types. However, the correlation coefficients calculated with a very small value of n are not useful because they have no statistical significance. (The statistical significance of a correlation coefficient increases with the square root of n-2, degree of freedom.) Conversely, choosing a large value of n will yield statistically significant correlation coefficients, but ones that are not very informative with regard to CO and SO2 emissions since they average together the variabilities for many plumes and begin to be effected by longer-term processes not related to emissions such as expansion and contraction of the boundary layer.

[47] To determine the optimum value of n for our analysis, we calculated the average confidence level for all

Figure 13. The probability distribution function (PDF) of CO-SO2 correlation coefficients, and average CO and SO2 concentrations (in ppbv) for each correlation coefficient bin as a function of the time interval, n, varying from 5 to 241 min.
correlation coefficients obtained from the month-long data set as a function of n using the student’s t-test (Figure 14). We find that the average confidence level is in fact quite low for n < 10 minutes and increases rapidly as n is increased. When n ~ 50 minutes, the average confidence level reaches 90% and then flattens out with relatively little increase with further increases in n. It would therefore appear that the optimum n is in the range of about 50–70 minutes. Examination of the PDF distribution in Figure 13 for this range of n-values reveals that there are two major components to the data, much like that of the synthetic data series discussed above. The one component, which comprise most of the time intervals, has little or no correlation between CO and SO\textsubscript{2}. The other component, which comprise ~13 of the time intervals, has a significant positive correlation (r > 0.7) between CO and SO\textsubscript{2}.

The unique aspect of the running-correlation analysis is that it yields values for the average CO concentration measured during those time intervals when CO is correlated and not correlated with SO\textsubscript{2}. Interestingly we find that the time intervals when CO is highly correlated with SO\textsubscript{2} are generally those time intervals when the CO concentrations are highest (i.e., >700 ppbv); these in fact are the high concentrations that caused the observed, average CO at Linan to be larger than the model-calculated average CO. When CO is not correlated with SO\textsubscript{2}, the CO concentration averages about 500–510 ppbv. This latter CO concentration is in fact quite close to the average CO concentration of ~480 ppbv predicted by our Standard Model simulations using the TRACE-P inventory. In other words, if we eliminate those portions of the Linan data set when CO and SO\textsubscript{2} are correlated, our model-calculated CO would agree with the average of the remaining CO observations. Our model-calculated SO\textsubscript{2}, however, would be greater than average of the remaining SO\textsubscript{2} observations. One interpretation of these results is that the hypothesized missing CO source(s) that caused the Standard Model to underpredict the CO concentration at Linan is (are) in large part also sources of SO\textsubscript{2} that are already accounted for in the TRACE-P inventory.

5.2.3. Application of Running Correlation to a Synthetic Data Set Similar to Measurements

In order to validate the analysis of the real measurements in the previous section, we applied the running correlation method to an additional set of synthetic data that more closely resembles the actual database from Linan. The pseudo-data X and Y are generated by randomly mixing a correlated data set (with r = 0.7) ranging in values from 950–1050 (average: 1000) with an uncorrelated data set (with r = 0) ranging in values from 450–550 (average: 500). The databases were mixed in consecutive fixed intervals of 100 datapoints. For the appropriate window size range (35–71), the running correlation analysis is able to filter out two peaks in the probability density function, one centered around the correlation coefficient bin of r = 0.7, with an average value of ~1000, and the other around r = 0, with an average value of ~500 (Figure 15). This analysis further demonstrates the effectiveness of running correlation method.

5.3. Scenarios That Would Provide the Missing CO Emissions

Suppose, as suggested in Section 4, the estimated 50% increase in CO emissions required to bring the Standard Model calculations into agreement with the Linan measurements were uniformly spread throughout China and throughout the year; this would require an additional source of 60 Tg CO yr\textsuperscript{−1}. Moreover suppose as suggested by the analysis presented in Section 5.2, that these additional CO emissions all occurred from sources that also emitted SO\textsubscript{2}. What kinds of modification would be required in the emission factors used to develop the TRACE-P emission inventory for CO?

First note that almost all of the SO\textsubscript{2} emitted in China comes from coal-burning: 49% from coal-burning power plants, 36% from industrial coal consumption, and 10% from domestic coal use [Streets et al., 2003]. Thus the requirement that the additional CO emissions come from sources that also emit SO\textsubscript{2} essentially requires that the additional CO emissions also come from coal-burning.

However, while China is the world’s largest consumer of coal, coal-burning in China represents a relatively small source of CO in the current TRACE-P inventory (i.e., about 18 Tg yr\textsuperscript{−1} or 16% of all CO emissions from China); about 9% of the total comes from industry, 7% from domestic sources, and power plants are negligibly small. If the additional emissions are to come from coal-burning, the emission factors used in the TRACE-P inventory for one or more of these sources would need to be adjusted upwards.

In general, CO emission factors from coal combus-
tors are highly variable. If a unit is operated improperly or not well maintained, the resulting CO emissions may

Figure 14. The average confidence level of the CO-SO\textsubscript{2} running correlation coefficients as a function of the time interval. Confidence levels calculated using the student’s t-test.

Figure 15. The probability distribution function (PDF) of the running correlation coefficients of randomly mixed correlated (r = 0.7, average = 1000) and noncorrelated (r = 0, average = 500) synthetic data sets. Also plotted are the average value of each corresponding coefficient bin. Here ‘n’ is the window size by which the correlation coefficient is calculated.
increase by several orders of magnitude \cite{U.S. EPA, 1998}. Thus an increase in CO emissions assigned to coal-burning activities in China would require that the coal-burning facilities in China are being operated at significantly lower efficiency than was assumed when the emission inventory was developed.

5.3.1. Coal-Fired Power Plants

\cite{54} When estimating CO emission rates, it is generally assumed that power plants operate with high efficiency. For example, the emission factor (EF) for most power plants in the United States is 0.25 g/kg, i.e., this corresponds to an operation efficiency in excess of 99.99\% \cite{U.S. EPA, 1998}. The CO emission factor (EF) for power plants in China is estimated to be 4 g/kg, which corresponds to an efficiency \textasciitilde99.8\%. Even though power plants are responsible for more than half of the total coal consumption in China (i.e., \textasciitilde500 Tg yr\(^\text{-1}\)) \cite{Sinton and Fridley, 2000}, the resulting CO emissions from power plant is negligibly small (i.e., only 2 Tg yr\(^\text{-1}\)). If power plants were responsible for all of the additional 60 Tg yr\(^\text{-1}\) of CO emissions hypothesized here, then the CO emission factor from coal-fired power plants in China would have to be increased from 4 g/kg to 120 g/kg. While this represents a rather large increase in power-plant operating efficiency from the current estimate of 99.8\% to 95.6\%. Moreover, recent aircraft measurements of coal-fired power plant plumes in United States (U.S.), revealed significantly elevated CO concentrations in some plumes. These observations suggested that the CO emission factors from these power plants is about 30 times higher than that normally assigned to U.S. plants under normal operating conditions \cite{Nicks et al., 2003}. Thus it would appear that there are significant uncertainties in CO emission rates from power-plants. Even larger uncertainties apply to power plants operating in China. Field measurements in China, similar to those of Nicks et al. \cite{2003}, would help to constrain these uncertainties.

5.3.2. Industrial and Domestic

\cite{55} Another possible category of unaccounted for emissions CO is from coal combustors used in small industrial and domestic facilities, such as boilers and cook stoves. Smaller combustors typically emit more CO per unit of fuel because they burn at lower temperatures than larger ones \cite{U.S. EPA, 1998}. The CO emission factors used in the TRACE-P inventory for the industrial and domestic sectors are 38 and 74 g/kg respectively; these are already 2–3 times higher than comparable sources in the United States \cite{Streets et al., 2003}. These emission factors yield 10 and 8 Tg CO yr\(^{-1}\) from industrial and domestic coal burning, respectively. In order to make up the hypothesized 60 Tg yr\(^{-1}\) in missing CO emissions, the emission factors from these two sectors would have to be increased by a factor of \textasciitilde3, i.e., \textasciitilde110 g/kg for industrial boilers and 200 g/kg for cooking stoves. These factors fall near the high-end in emission factors observed for these types of facilities. For example, the \textit{U.S. EPA} \cite{1998} suggests an emission factor for industrial hand-fired coal combustors (which typically operate at low efficiency) of 138 g/kg. Zhang et al. \cite{1999} observed CO emission factors from cooking stoves in China ranging from 18–170 g/kg.

\cite{56} Of course the above estimated emission factors assumed that the additional CO emissions were coming either entirely from power plants or entirely from industrial and domestic sources. If the emissions were equally divided between these sectors, it would require CO emission factors from power plants, industrial burners, and domestic stoves of \textasciitilde60, 55, and 100 g/kg. These are large, but perhaps not out of the realm of the possible.

6. Conclusion

\cite{57} A comparison between model simulations and field measurements of gas and particle species at a rural site in one of the more industrialized regions of China, the Yangtze Delta, during November, 1999 suggests that while the TRACE-P emission inventory for sulfur oxides is accurate, the emissions inventory for CO and PC are underestimated by \textasciitilde50 and 60–90\%, respectively. Further analysis of the high resolution measurements of CO and SO\(_2\) suggest that the additional source(s) of CO needed to bring the model into agreement with the measurements is most likely associated with SO\(_2\) emissions from coal burning. Significant increases in the CO emission factors typically assigned to power-generating, industrial, and domestic coal-burning processes would be needed to account for the missing CO emissions from these sources. It is not at all clear that such large emission-factor increases are appropriate. A more comprehensive effort to monitor CO and related pollutants in smoke-stacks, pollutant plumes, and ambient air are needed to better constrain the CO emission rate from China and East Asia and identify the contributions from various sources. A similar effort is needed in the case of PC emissions as well.

\cite{58} Acknowledgments. This work was supported in part by the National Science Foundation under grant ATM-0129495. T. Wang acknowledges support for his effort by the Research Grants Council of Hong Kong (project PolyU 5063/01E).

References

Akimoto, H., and H. Narita (1994), Distribution of SO\(_2\), NO\(_x\), and CO emissions from fuel combustion and industrial activities in Asia with 1° × 1° resolution, \textit{Atmos. Environ.}, 28, 213–225.


Chameides, W. L., and M. Bergin (2002), Enhanced: Soot takes center
Chameides, W. L., et al. (1999a), Is ozone pollution affecting crop yields in
Chameides, W. L., et al. (1992), Ozone precursor relationships in the am-
Ganzeveld, L., J. Lelieveld, and G. J. Roelofs (1998), A dry deposition
Parrish, D. D., et al. (1993), The total reactive oxidized nitrogen levels and Convective processes and assimilation of lateral boundary conditions,
Parrish, D. D., et al. (1993), The total reactive oxidized nitrogen levels and Convective processes and assimilation of lateral boundary conditions,


M. Bergin, W. L. Chameides, and Q. Tan, School of Earth and Atmospheric Science, Georgia Institute of Technology, Atlanta, GA 30332, USA. (tan@eas.gatech.edu)

D. Streets, Argonne National Laboratory, Argonne, IL 60439, USA.

T. Wang, Hong Kong Polytechnic University, Hong Kong, China.

J. Woo, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA 02114, USA.

J. Xu, Desert Research Institute, Las Vegas, NV 89119, USA.