**Atmospheric
Chemistry
and Physics
Discussions**

Implementation of a
MCMC method to
aerosol modeling in
MCMA

F. M. San Martini et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures

◄◄	►►
◄	►
Back	Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Implementation of a Markov Chain Monte Carlo Method to inorganic aerosol modeling of observations from the MCMA-2003 Campaign. Part I: Model description and application to the La Merced Site

F. M. San Martini[1,*], E. J. Dunlea[2], M. Grutter[3], T. B. Onasch[4], J. T. Jayne[4],
M. R. Canagaratna[4], D. R. Worsnop[4], C. E. Kolb[4], J. H. Shorter[4], S. C. Herndon[4],
M. S. Zahniser[4], J. M. Ortega[1,**], G. J. McRae[5], L. T. Molina[1,6], and M. J. Molina[1,***]

[1]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of
Technology, Cambridge, MA, USA
[2]Cooperative Institute for Research in the Environmental Sciences (CIRES), University of
Colorado at Boulder, Boulder, CO, USA
[3]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, Mexico City,
Mexico
[4]Aerodyne Research Inc., Billerica, MA, USA

[5] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

[6]Molina Center on Energy and the Environment, La Jolla, CA, USA

[*] now at: the Board on Chemical Sciences and Technology, National Academies, Washington D.C., USA

[**] now at: Sandia National Laboratory, Livermore, CA, USA

[***]now at: University of California San Diego, La Jolla, CA, USA

Correspondence to: F. M. San Martini (ico@alum.mit.edu)

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◄◄	►►
◄	►
Back	Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

Abstract

The equilibrium inorganic aerosol model ISORROPIA was embedded in a Markov Chain Monte Carlo algorithm to produce a powerful tool to analyze aerosol data and predict gas phase concentrations where these are unavailable. The method directly
5   incorporates measurement uncertainty, prior knowledge, and provides for a formal framework to combine measurements of different quality. The method was applied to aerosol- and gas-phase precursor observations taken at La Merced during the Mexico City Metropolitan Area (MCMA) 2003 Field Campaign and served to discriminate between diverging gas-phase observations of ammonia. The model reproduced obser-
10   vations of particle-phase ammonium, nitrate, and sulfate well. The most likely concentrations of ammonia were found to vary between 4 and 26 ppbv, while the range for nitric acid was 0.1 to 55 ppbv. During periods where the aerosol chloride observations were consistently above the detection limit, the model was able to reproduce the aerosol chloride observations well and predicted the most likely gas-phase hydrochloric acid
15   concentration varied between 0.4 and 5 ppbv. Despite the high ammonia concentrations observed and predicted by the model, when the aerosols were assumed to be in the efflorescence branch they are predicted to be acidic (pH~3).

1   Introduction

Air quality managers continually face decisions with information constraints, where
20   data may have poor time resolution, high uncertainty, or may simply be lacking. Examples range from the Southeastern Aerosol Research and Characterization Project (SEARCH) in the U.S., which did not include the requisite ammonia measurements (Blanchard and Hidy, 2003), to the 1999 Atlanta Supersite Experiment, where a systematic bias of aerosol sulfate concentrations was postulated (Zhang et al., 2003).
25   In particular, while measurements of urban PM are often available, the corresponding gas phase concentrations needed for modeling and devising control strategies are

often lacking. The need for gas phase observations to devise control strategies for inorganic PM was confirmed by Blanchard et al. (2000), who identified two parameters to determine which precursor species limit aerosol nitrate formation: both parameters require particulate and gas phase observations.

5   Mexico City has long been well-known for its poor air quality (Molina and Molina, 2002). The 2003 Mexico City Metropolitan Area (MCMA) field campaign was an intensive 5-week campaign focused on providing a scientific base for devising emissions control strategies to reduce exposure to harmful pollutants in the MCMA, as well as insights to air pollution problems in other megacities. In this paper we use data from
10 the MCMA-2003 campaign to illustrate the application of a Bayesian method to infer missing and uncertain measurements, focusing on the inorganic aerosol system. We focus on the inorganic aerosols because the health risk of air pollution in the MCMA is dominated by the effect of particles (Evans et al., 2002), and, although organic aerosols comprise a majority of the fine particle mass in the MCMA, uncertainties in both model-
15 ing (Ansari and Pandis, 2000b; Clegg et al., 2001, 2003) and measurements (McMurry, 2000) of organic aerosols and gas phase precursors remain large relative to the inorganic aerosols.

The Bayesian method introduced here combines available measurements with knowledge of aerosol thermodynamics to infer missing variables, and allows for the
20 direct incorporation of measurement uncertainty, the use of prior knowledge, and provides a framework for combining measurements of different quality. Here we apply the method to observations taken at the La Merced site. This site was chosen due to the availability of direct observations of gas-phase nitric acid as well as two co-located instruments to measure gas-phase concentrations of ammonia; this dataset provides
25 a unique opportunity to demonstrate how the method can be used to discriminate between diverging observations. The need to better understand ammonia emissions and concentrations was highlighted by San Martini et al. (2005), who showed that reductions in ammonia concentrations are likely to be less effective at reducing $PM_{2.5}$ in Mexico City than expected, while reductions in nitrate and sulfate are expected to be

effective. A companion paper will discuss the application of the Bayesian method to three other fixed sites in the MCMA-2003 campaign, CENICA, Pedregal, and Santa Ana.

2 Experimental

La Merced (19°24′ N, 99°07′ W, 2250 m a.s.l.) is an area in downtown MCMA that includes both commercial and residential buildings and has heavy traffic. A major bus station (known as TAPO) is located ∼500 m northeast of the site, and the Mexico City international airport is ∼2–3 km east of the site (Moya et al., 2004). The routine monitoring network in Mexico City (Red Automática de Monitoreo Atmosférico, RAMA) operates a site at La Merced. The RAMA station measures temperature and relative humidity, as well as other meteorological observations (wind speed and direction, UV, etc.) and criteria pollutant concentrations (e.g., ozone, NO_x, SO_2, PM_{10}[1]), on a per-minute basis. In addition, hourly averaged observations, based on the per-minute data, are available from RAMA if the data logger fails.

During MCMA-2003 gas-phase measurements of ammonia and nitric acid were taken at La Merced using an open-path Fourier Transform Infrared (FTIR) spectrometer. A full description of the experiment and location is presented elsewhere (Grutter et al., 2003). The FTIR instrument used a bistatic telescope system installed on top of two four-story buildings along the 426-m lightpath, approximately 20 m above the surface. The RAMA monitoring station is located ∼30 m away to the north from the west end of the optical trajectories and is only ∼8 m above the surface (Grutter et al., 2005). The IR radiation is modulated with a Nicolet interferometer and captured with a HgCdTe detector at 77 K. Approximately 180 interferograms are co-added during 5 min to produce an infrared transmission spectrum with $0.5\,cm^{-1}$ resolution. The concentrations are retrieved by performing a classical least squares regression using a synthetic back-

---

[1] Monitoring of $PM_{2.5}$ by RAMA started in 2004, after the MCMA-2003 campaign.

ground and references generated from the HITRAN spectroscopic database (Rothman et al., 1998). For the quantitative analysis of NH_3 and HNO_3, the regions 920–1090 and 875–900 cm^{-1} are used, respectively.

The Aerodyne Mobile Laboratory (AML) was parked at the La Merced site from 25 April to 27 April 2003. The AML contains a suite a fast-response instruments capable of measuring trace gas concentrations at sub ppb levels, an aerosol mass spectrometer (AMS) to measure the non-refractory chemical components of fine airborne particles, as well as selected commercial fast response instruments (Herndon et al., 2005; Kolb et al., 2004). Included in the suite of instruments on the AML was a quantum cascade tunable infrared laser differential absorption spectroscopy (TILDAS) instrument capable of measuring NH_3 concentrations with one second time resolution. The TILDAS instrument is a closed path system where the laser output is coupled into a multiple pass absorption cell with a 56 m pathlength. The laser (Alpes Lasers) operated in the 967.35 cm^{-1} region, overlapping a strong ammonia feature. The laser linewidth was 0.014 cm^{-1} (hwhm), and the laser tuning rate was determined from a Germanium etalon. Concentrations were calculated based on the HITRAN database (Rothman et al., 2003) and measured sample pressure and temperature.

The AML did not include an instrument to directly measure gas-phase HNO_3. However, an estimate of the HNO_3 concentration can be derived based on observations of NO, NO_2, and total NO_y. The AML included a commercial total NO_y instrument, which measures both NO_y and NO using the chemiluminescence (CL) technique, but configured differently than a standard CL NO_x monitor so as to exploit the molybdenum converter's ability to detect more gas phase reactive nitrogen species. NO_2 was measured with two instruments on board the AML: a fast-response TILDAS and a commercial NO_x instrument. The operation of the NO_2 TILDAS is described in (Li et al., 2004). (Dunlea et al., 2006)[2] compared observations of NO_2 from the AML and other

---

[2]Dunlea, E. J., Herndon, S. C., Nelson, D. D., Volkamer, R. M., San Martini, F. M., Zahniser, M. S., Shorter, J. H., et al.: Evaluation of Standard Measurement Techniques for Nitrogen Dioxide in a Polluted Urban Environment?, in preparation, 2006.

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◄◄ ►►

◄ ►

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

instruments at three sites during MCMA-2003 and showed that the TILDAS observations are the most reliable. From the total NO_y and NO measurements, along with the TILDAS NO_2 measurement, we calculate the non-NO_x fraction of NO_y, referred to as NO_z:

5   $$NO_z = NO_y - NO - NO_2 \tag{1}$$

NO_z provides an (approximate) upper bound to the HNO_3 concentration since NO_z may comprise HNO_3, RNO_3, PAN, HONO, NO_3, N_2O_5 and particulate NO_3. Section 3.2 discusses the uncertainties in this measure of HNO_3.

The AMS has been described in detail elsewhere (Jayne et al., 2000), and an
10   overview of its application during the MCMA-2003 campaign is provided by Salcedo et al. (2006). The AMS measures non-refractory (NR) species, operationally defined to include all species that evaporate in a few microseconds after a sampled aerosol particle impinges on the AMS heated vaporization surface, in particles smaller than about 1 μm (NR-PM_1) (Salcedo et al., 2006). NR species internally mixed with refrac-
15   tory species can be detected quantitatively (Katrib et al., 2005; Slowik et al., 2004). Therefore, all inorganic aerosol species of interest are observed except for crustal materials and sea salt. The AMS observations used here are 4-min averages. All other observations were averaged to the AMS timestamp, with the exception of the 5-min averaged FTIR observations, which were interpolated to the AMS timestamp.
20   In sum, observations of gas-phase precursors and inorganic aerosol species, as well as temperature and relative humidity, are required in order to model the inorganic aerosol system. Thus, although CENICA was considered the supersite for MCMA-2003, La Merced can be considered the inorganic aerosol system supersite because it was the only location during the MCMA-2003 campaign where co-located NH_3 and
25   HNO_3 observations were both available. The only species relevant to the inorganic aerosol system that were not directly measured at La Merced are crustal species and gas-phase hydrochloric acid. The method used to estimate these species is presented in Sect. 3.3.

Printer-friendly Version

Interactive Discussion

EGU

3 The Bayesian approach

Uncertainty can be divided into two categories: aleatory and epistemic uncertainty (Pate-Cornell, 1996). Aleatory uncertainty (also known as inherent or stochastic uncertainty) represents randomness or variability in nature, and in general cannot be completely eliminated. Epistemic uncertainties represent a lack of knowledge of the system, which may be due to statistical uncertainty (due to lack of sufficient data) and model uncertainty (due to lack of understanding of the physics or chemistry). In principle, epistemic uncertainties can be reduced as knowledge increases and more data becomes available. The inorganic aerosol system modeled here is characterized by uncertain observations, missing variables, and stochastic processes; therefore, we require a tool that treats both types of uncertainty. The statistical theory that allows the measurement and combination of aleatory and epistemic uncertainties is Bayesian statistics (Pate-Cornell, 1996). The tool we will use for the Bayesian analysis is the Markov Chain Monte Carlo (MCMC) method.

Bayes' Theorem describes conditional probability:

$$p\left(\theta|\text{Data}\right) = \frac{p\left(\text{Data}|\theta\right) p\left(\theta\right)}{p\left(\text{Data}\right)} \tag{2}$$

where Data and θ are the observations and unknown variables. In Eq. (2), $p\left(\theta|\text{Data}\right)$ is the posterior, $p\left(\text{Data}|\theta\right)$ is the likelihood function, $p\left(\theta\right)$ the prior, and $p(\text{Data})$ is a normalizing constant (equal to the probability of the observations). The vector of unknown variables θ may be composed of observables that have not yet been observed and parameters, which are inherently unobservable. From a Bayesian perspective, there is no fundamental distinction between observables and parameters: all are considered random quantities. Bayes' Theorem provides a powerful tool to predict observables and infer parameters based on observations.

A variety of methods are available to solve Eq. (2), including conjugate analysis, asymptotic analysis, the use of closed-form approximations, and sampling based approximations. MCMC is an example of the sampling based approach, where the key

idea is that while it would be nice to calculate $p(\theta|\text{Data})$, we are just as happy to simulate a large number of random draws from $p(\theta|\text{Data})$ (Draper, 2006[3]). Thus, rather than calculate the posterior, which in multi-dimensional problems involves computing expensive multi-dimensional integrals, the posterior is estimated by directly drawing random
5 samples from the distribution. These random samples are then used to generate the descriptive characteristics of the posterior distribution.

The question of how to implement a stochastic simulation from which random draws can be obtained and that is described by the posterior distribution $p(\theta|\text{Data})$ was originally answered by Metropolis et al. (1953). Metropolis proposed generating a Markov
10 chain, a stochastic process whose next state depends on the past only through the value of the present state (Bertsekas and Tsitsiklis, 2002), that has the same state space as θ and whose equilibrium distribution is $p(\theta|\text{Data})$. Algorithms to generate the Markov chain include Gibbs sampling, the independence sampler, Metropolis-Hastings, and others. The algorithm used in this work, and probably the most widely
15 used algorithm, is the Metropolis-Hastings algorithm. First, an initial guess θ^0 must be specified. Then, the algorithm is as follows:

– Current position is θ

– Generate proposed new θ^*

– Calculate the acceptance probability α (see below)

20 Compare α to U, where U is a random number generated on the interval [0,1]. If $\alpha > U$, the proposed step is accepted.

The acceptance probability α is given by:

$$\alpha = \min\left\{1, \frac{p(\theta^*|\text{Data})}{p(\theta|\text{Data})} \frac{PD(\theta|\theta^*)}{PD(\theta^*|\theta)}\right\} \tag{3}$$

[3]Draper, D.: Bayesian Hierarchical Modeling, New York, Springer-Verlag, in preparation, 2006.

where *PD(θ)* is termed the probing distribution (see Appendix A).

The two key components of the MCMC method are a model relating the unknown variables to the observations and a probability model describing the likelihood of the observations. These will be discussed in turn. Next, Sect. 3.3 describes how prior information was included in the analysis. Appendix A briefly discusses two implementation issues that are critical to successfully applying the MCMC method: selecting a probing distribution and convergence monitoring strategies.

3.1 Inorganic aerosol model

The assumption that local equilibrium exists for volatile species between the gas and aerosol phases has been frequently invoked, and equilibrium models have been under development for over twenty years (for example see Ansari and Pandis, 1999b; Bassett and Seinfeld, 1983, 1984; Nenes et al., 1998; Pilinis and Seinfeld, 1987; Wexler and Seinfeld, 1991). A variety of researchers have shown generally good agreement between equilibrium predictions and field observations (for example, see Allen et al., 1989; Hildemann et al., 1984; Pilinis and Seinfeld, 1988; Russel et al., 1988), though under certain conditions the equilibrium time scale is too long to justify the equilibrium assumption (Wexler and Seinfeld, 1990, 1992). Factors that favor equilibrium are small particle size, high particle number concentrations, and higher temperatures. Conversely, low aerosol mass concentrations, low temperatures, and large particle sizes will increase the equilibrium time scale. Given the size of the particles sampled by the AMS ($<1\,\mu m$), their high concentrations, and the high temperatures and low relative humidities observed at La Merced, we expect that the equilibrium assumption is reasonable for the period of study.

Two excellent reviews of available inorganic aerosol models are provided by (Zhang et al., 2000) and (Ansari and Pandis, 1999a). Zhang et al. (2000) compared predictions of MARS-A, SEQUILIB, SCAPE2, EQUISOLV II, and AIM2 under a variety of conditions and found that PM compositions are generally comparable for most ambient gas-phase compositions. These findings were confirmed by Ansari and Pandis, who

F. M. San Martini et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◄◄	►►
◄	►
Back	Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

compared the predictions of GFEMN, ISORROPIA, SCAPE2 and SEQUILIB both for a series of theoretical cases and against observations taken during the Southern California Air Quality Study (SCAQS) (Ansari and Pandis, 1999a). Applying the models to SCAQS, Ansari and Pandis found minor discrepancies in predictions between the models and general agreement with the SCAQS observations, though nitrate is underpredicted (Ansari and Pandis, 1999a). Observations of crustal species were not available, however, and likely contributed to this underprediction.

Overall, Ansari and Pandis found small discrepancies in the overall prediction of aerosol behavior of the four models, where GFEMN was used as a reference aerosol. For ammonia rich environments, the mean predictions of the four models of aerosol nitrate and total dry inorganic PM agreed within 3%; ISORROPIA's aerosol nitrate predictions showed better agreement with GFEMN than SEQUILIB and SCAPE2 (Ansari and Pandis, 1999a). Although ISORROPIA tended to predict lower aerosol water than GFEMN, SCAPE2, and SEQUILIB, relative to the predictions of GFEMN, the mean normalized bias and error of ISORROPIA's aerosol water were approximately an order of magnitude smaller than for SCAPE2 and SEQUILIB. Previous observations indicate that Mexico City is an ammonia rich environment (Chow et al., 2002b; Moya et al., 2004), suggesting that ISORROPIA is a particularly suitable choice of model.

The model treatment of chloride species is worth particular mention (see Sect. 4). GFEMN, SCAPE2 and ISORROPIA predict that sodium will preferentially bind with available HNO_3 to form sodium nitrate ($NaNO_3$). If sufficient HNO_3 is not available, the excess sodium is predicted to bind with available HCl (forming NaCl). EQUISOLV II and SEQUILIB, however, assume that the partitioning of sodium between $NaNO_3$ and NaCl (in the presence of HNO_3 and HCl) is governed by:

$$NaCl(s) + HNO_3(g) \leftrightarrow NaNO_3(s) + HCl(g) \tag{4}$$
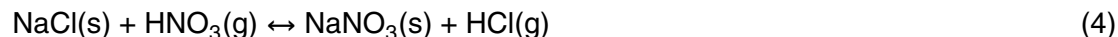
For the four models they examined, Ansari and Pandis found minor differences in predicted chloride concentrations (Ansari and Pandis, 1999a). However, Zhang et al. found that the particulate chloride exhibited good agreement between AIM2 and

EGU

SCAPE2, but they found significant differences between the predictions of SEQUILIB and EQUISOLV II.

Moya et al. (2001) confirmed the finding of minor discrepancies between the four models examined by Ansari and Pandis by applying the models to data from the 1997 IMADA-AVER field campaign in Mexico City. Using the same dataset, San Martini compared predictions from ISORROPIA and a new equilibrium model that includes complex and hydrate species and directly minimizes the Gibbs free energy, and similarly found only small differences in model predictions (San Martini, 2004).

Nenes et al. (1999) incorporated ISORROPIA into the three-dimensional airshed model UAM-AERO and compared it with predictions of UAM-AERO with SEQUILIB; they found good agreement between model predictions as well as experimental results. These researchers also point out that ISORROPIA is significantly faster than other inorganic aerosol models; this characteristic makes ISORROPIA an attractive model for use in a sampling based technique.

Based on its agreement with other models, computational speed and the high ammonia concentrations previously observed in the MCMA, the model selected to relate the unknown variables to the observations is a modified version of the inorganic equilibrium model ISORROPIA. The major reactive inorganic atmospheric aerosol components are ammonia, sulfate, nitrate, sodium, and chloride; water is the most important solvent for constituents of atmospheric particles and drops (Ansari and Pandis, 1999b). ISORROPIA predicts the equilibrium partitioning of inorganic species between the gas and aerosol phase given inputs of temperature, relative humidity, and total pollutant concentrations (Nenes et al., 1998). For sulfate and sodium, the total concentration is the aerosol phase concentration, while for ammonia, nitrate, and chloride the total concentration is the sum of the gas and aerosol phase:

$$NH_3^t = NH_4^+(aerosol) + NH_3(g) \tag{5}$$

$$NO_3^t = NO_3^-(aerosol) + HNO_3(g) \tag{6}$$

$$Cl^t = Cl^-(aerosol) + HCl(g) \tag{7}$$

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

One modification made to ISORROPIA was the value of the equilibrium constant K_p (ppb^2) for the dissociation of solid ammonium nitrate:

$$NH_4NO_3(s) \leftrightarrow NH_3(g) + HNO_3(g) \tag{8}$$

The value of K_p at a temperature T is evaluated according to:

$$K_P(T) = K(T^o) \exp\left\{ a\left(\frac{T^o}{T} - 1\right) + b\left(1 + \ln\left(\frac{T^o}{T}\right) - \frac{T^o}{T}\right) \right\} \tag{9}$$

ISORROPIA, like SCAPE2 (Kim et al., 1993) and EQUISOLV II (Jacobson, 1999; Jacobson et al., 1996), used the NBS Thermodynamic Tables (Wagman et al., 1982) to determine Kp for Eq. (8). Mozurkewich has conducted the most comprehensive review to date of available thermodynamic data to determine the equilibrium constant of ammonium nitrate (Mozurkewich, 1993); therefore, the thermodynamic parameters of Mozurkewich were substituted for those used in the original formulation of ISOR-ROPIA. Table 1 shows the values of $K(T^o)$, a, and b used by different models. Note that neither AIM2 nor GFEMN use equilibrium constants, rather they directly minimize the Gibbs free energy to determine equilibrium. The thermodynamic parameters used by these two models can be used to calculate a value of $K(T^o)$ of 43.6 ppb^2 for AIM2 and 42.5 ppb^2 for GFEMN. These values compare well with the value of 41.99 ($\pm$12%) ppb^2 suggested by Mozurkewich.

3.2   Likelihood of the observations

Uncertainty associated with a measurement can be partitioned into two components: random noise and measurement bias (or systematic error) (Ferson and Ginzburg, 1996). The observed and "true" variables can be related by:

$$X_{\text{true}} = X_{\text{obs}} + \varepsilon \tag{10}$$

where ε is the difference between the "true" and observed values of X. For the case where the variation is due to a combination of many small errors, with each of the

errors being equally likely of being positive or negative, ε will be described by a normal distribution. Indeed, the central limit theorem tells us that even if some of the error sources have non-Gaussian distributions, ε will still be normally distributed as long as the number of error sources is large. The Gaussian distribution has been found to describe more real cases of experimental and instrument variability than any other distribution (Coleman and Steele, 1999). For an unbiased observation ε is described by a normal distribution with zero mean and variance σ^2:

$$\varepsilon \sim N(0, \sigma) \tag{11}$$

X_{true} now is a normally distributed random variable with mean X_{obs} and variance σ^2. For the case where the measurement uncertainty is proportional to the observation

$$\sigma \propto X_{\text{obs}} \tag{12}$$

$$\sigma = s \times X_{\text{obs}} \tag{13}$$

Combining Eqs. (11) and (13) gives the likelihood model for an unbiased observation whose uncertainty is proportional to the observation:

$$X \sim N(X_{\text{obs}}, s \times X_{\text{obs}}) \tag{14}$$

Equation (14) describes the likelihood function of an unbiased measurement with Gaussian error whose uncertainty is proportional to the measurement, i.e.,

$$p\left(X_{\text{obs}}|x\right) = \frac{1}{\sqrt{2\pi}sX_{\text{obs}}} \exp\left(-\frac{1}{2}\left(\frac{x - X_{\text{obs}}}{sX_{\text{obs}}}\right)^2\right) \tag{15}$$

3.2.1 Likelihood function for ammonia observations

Figure 1 shows the FTIR and TILDAS NH_3 observations taken at La Merced. Between approximately midnight (CDT) and 11:00 a.m. these two observations diverge

markedly. In theory, the error associated with the FTIR NH_3 concentrations is between 15% and 20% (Moya et al., 2004). The estimated uncertainty for the TILDAS ammonia concentrations is 20%. However, given the discrepancies in observed concentrations, a more conservative error estimate of ±29% (at the 95% confidence level) was used for both NH_3 instruments (see Table 2). Combining the likelihood expression given by Eq. (15) with the parameters in Table 2 yields the likelihood for the FTIR or TILDAS measurement:

$$p\left(\text{FTIR}|NH_3\right) = \frac{1}{\sqrt{2\pi}\left(0.15\left(NH_3\right)_{\text{obs}}^{\text{FTIR}}\right)} \exp\left(-\frac{1}{2}\left(\frac{NH_3-\left(NH_3\right)_{\text{obs}}^{\text{FTIR}}}{0.15\left(NH_3\right)_{\text{obs}}^{\text{FTIR}}}\right)^2\right) \tag{16}$$

$$p\left(\text{TILDAS}|NH_3\right) = \frac{1}{\sqrt{2\pi}\left(0.15\left(NH_3\right)_{\text{obs}}^{\text{TILDAS}}\right)} \exp\left(-\frac{1}{2}\left(\frac{NH_3-\left(NH_3\right)_{\text{obs}}^{\text{TILDAS}}}{0.15\left(NH_3\right)_{\text{obs}}^{\text{TILDAS}}}\right)^2\right) \tag{17}$$

The goal of this work, however, is not to determine likelihood for the FTIR or TILDAS measurement. Rather, given the "true" ammonia concentration, we wish to determine the likelihood of both the FTIR and TILDAS observations, i.e., we wish to determine p(Data $|NH_3$), where Data=(FTIR, TILDAS).

Given the evident discrepancies in the two NH_3 observations, great care was taken to ensure quality assurance and control of these observations. We therefore have a high degree of confidence in both observations. We define an augmented model space $\theta \equiv [NH_3, M]$, where $M \equiv [M_{\text{FTIR}}, M_{\text{TILDAS}}]$, to incorporate both observations into the likelihood function. M is a binary variable, where $p(M_{\text{TILDAS}})$ and $p(M_{\text{FTIR}})$ are the probabilities that the TILDAS and FTIR instrument reflect the true state of nature, and we assume that $p(M_{\text{TILDAS}})+p(M_{\text{FTIR}})=1$. We have a high degree of confidence in both observations and hence no a priori reason to believe one observation is more likely than the other, i.e.,

$$p\left(M_{\text{FTIR}}\right) = p\left(M_{\text{TILDAS}}\right) = 0.5 \tag{18}$$

EGU

The augmented likelihood function is thus given by:

$$p\left(\text{Data}|\text{NH}_3, M\right) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\left(0.15(\text{NH}_3)_{\text{obs}}^{\text{FTIR}}\right)} \exp\left(-\dfrac{1}{2}\left(\dfrac{\text{NH}_3-(\text{NH}_3)_{\text{obs}}^{\text{FTIR}}}{0.15(\text{NH}_3)_{\text{obs}}^{\text{FTIR}}}\right)^2\right) & \text{if } M = M_{\text{FTIR}} \\ \dfrac{1}{\sqrt{2\pi}\left(0.15(\text{NH}_3)_{\text{obs}}^{\text{TILDAS}}\right)} \exp\left(-\dfrac{1}{2}\left(\dfrac{\text{NH}_3-(\text{NH}_3)_{\text{obs}}^{\text{TILDAS}}}{0.15(\text{NH}_3)_{\text{obs}}^{\text{TILDAS}}}\right)^2\right) & \text{if } M = M_{\text{TILDAS}} \end{cases} \quad (19)$$

Note that with an augmented model space $\theta \equiv [\text{NH}_3, M]$, the expression for the acceptance probability now is:

$$\alpha = \min\left\{1, \frac{p(\text{Data}|\text{NH}_3^*, M^*)p\left(\text{NH}_3^*|M^*\right)p\left(M^*\right)}{p(\text{Data}|\text{NH}_3, M)p\left(\text{NH}_3|M\right)p\left(M\right)} \frac{PD(\text{NH}_3|\text{NH}_3^*, M^*)PD\left(M\right)}{PD(\text{NH}_3^*|\text{NH}_3, M)PD\left(M^*\right)}\right\} \quad (20)$$

The ammonia prior is obtained from observations taken before the experiment is begun; it is therefore independent of M:

$$p\left(\text{NH}_3|M\right) = p\left(\text{NH}_3\right) \quad (21)$$

Similarly, the ammonia probing distribution is independent of M:

$$PD\left(\text{NH}_3^*|M, \text{NH}_3\right) = PD\left(\text{NH}_3^*|\text{NH}_3\right) \quad (22)$$

Combining Eqs. (20–22):

$$\alpha = \min\left\{1, \frac{p(\text{Data}|\text{NH}_3^*, M^*)p\left(\text{NH}_3^*\right)p\left(M^*\right)}{p(\text{Data}|\text{NH}_3, M)p\left(\text{NH}_3\right)p\left(M\right)} \frac{PD(\text{NH}_3|\text{NH}_3^*)PD\left(M\right)}{PD(\text{NH}_3^*|\text{NH}_3)PD\left(M^*\right)}\right\} \quad (23)$$

3.2.2 Likelihood function for nitric acid observations

The nominal error associated with the FTIR HNO_3 concentrations is approximately 40% (Moya et al., 2004). The higher uncertainty in the HNO_3 determination relative to that for NH_3 is due to the small HNO_3 infrared fingerprint and the strong water interference

EGU

in the spectral window. A more conservative estimate of ±49% is used here (see Table 2).

The uncertainty in the estimate of HNO_3 based on the NO_z observations is large. In theory, NO_z provides an upper bound to the HNO_3 concentration since NO_z may comprise HNO_3, RNO_3, PAN, HONO, NO_3 and N_2O_5. However, the uncertainties in the NO_z concentrations are large because the concentrations were derived from three measurements (see Eq. 1). The measurement errors are therefore additive. The average concentration of NO_y and $(NO+NO_2)$ measured by the AML at La Merced were 100 and 90 ppb, respectively. Typical urban concentrations of HNO_3 range from sub-ppb to 10's of ppb; the average NO_z observation at La Merced was 4 ppb, and the mode of the distribution was 3 ppb. For reference, a 10% uncertainty in the NO, NO_2, and NO_y observations results in an uncertainty that is approximately a factor of five larger than the most frequently reported value of NO_z. Figure 2 compares the FTIR HNO_3 observations with the determined NO_z values. Despite the large uncertainties in the NO_z values, the diurnal variations of HNO_3 evident in the FTIR observations can be seen in the NO_z time series.

Due to the high uncertainty associated with the NO_z values, the likelihood function for HNO_3 is based only on the FTIR HNO_3 observation when it is available, i.e., we neglect the NO_z observation if the FTIR HNO_3 observation is available. Out of a total of 612 data points analyzed, there are three points except either the FTIR HNO_3 observation is missing or negative. For these three points only the NO_z observation is used in the likelihood function, where the uncertainty is assumed to be ±49% (see Table 2).

3.2.3 Likelihood function for AMS observations

The mass concentrations measured with the Aerodyne AMS during the MCMA-2003 campaign have a range of uncertainty of approximately −30% and +10% (Salcedo et al., 2006). This asymmetric likelihood distribution is due to the uncertainty in particle collection efficiency. The mode of the likelihood function will therefore be the observation (i.e., the most likely value is the observation itself), with the probability density

decreasing from the observation.

The probability model selected to represent this likelihood function is a mixed Gaussian model, where most of the probability density comes from a normal probability density function (pdf) centered at the observation, and the remaining probability density is provided by a second more diffuse Gaussian centered at 85% of the observation:

$$p\left(X_{\text{obs}}|X\right) = m \times N_1\left(\mu_1, \sigma_1\right) + (1-m) \times N_2\left(\mu_2, \sigma_2\right) \quad (24)$$

where:

m=mixing proportion=0.7

μ_1 = the AMS observation = X_{obs}

$\sigma_1 = 0.061 \times X_{\text{obs}}$

$\mu_2 = 0.85 \times X_{\text{obs}}$

$\sigma_2 = 0.15 \times \mu_2 = 0.1275 X_{\text{obs}}$

Substituting the parameters into the likelihood function yields

$$p\left(X_{\text{obs}}|X\right) = 0.7 \times N\left(X_{\text{obs}}, 0.061X_{\text{obs}}\right) + (0.3) \times N\left(0.85X_{\text{obs}}, 0.1275X_{\text{obs}}\right) \quad (25)$$

Figure 3 shows the likelihood pdf for an observation of unity (arbitrary units). The most likely concentration is the observation, and the probability that the "true" concentration is greater than 1.1 is 4.3% and less than 0.7 is 3.6% (i.e., there is an 8% chance that the "true" concentration is either below 0.7 or above 1.1).

The AMS observations were measured in 4-min intervals with a 50% duty cycle. The first two minutes of each interval were used to characterize the internal particle beam shape and are not included in the averaged data presented here. The detection limits for ammonium, nitrate, and sulfate were 0.37, 0.05, 0.11 μg/m^3, respectively. Although the nominal detection limit for chloride is 0.05 μg/m^3, a higher detection limit was used. The higher detection limit for chloride was selected because on average the chloride observations are between one and two orders of magnitude smaller (on a molar basis) than the other inorganic aerosol species, and due to the negative observations reported (see Fig. 12d and Part II). The detection limit used for chloride observations

EGU

was $0.15\,\mu g/m^3$. Moreover, given the uncertainty of the small chloride mass concentrations evidenced by the negative observations, the standard deviation for the chloride likelihood was doubled for observations between one and two times the detection limit, i.e., for chloride observations between 0.15 and $0.30\,\mu g/m^3$ the likelihood function is:

$$p\left(X_{obs}|X\right) = 0.7 \times N\left(X_{obs}, 0.122X_{obs}\right) + (0.3) \times N\left(0.85X_{obs}, 0.255X_{obs}\right) \tag{26}$$

The chloride observations and predictions are discussed further in Results section and in Part II.

For all species, if an observation was below the detection limit, the likelihood function given by Eq. (25) was not used. Rather, for ammonium, nitrate, and sulfate, the uncertainty in the observation is assumed to be constant and equal to ±49% of the detection limit (i.e., ±0.18, 0.025, and $0.054\,\mu g/m^3$ for ammonium, nitrate and sulfate, respectively, at the 95% confidence level). Chloride observations below the detection level were assumed to be negligible. The AMS time series at La Merced comprises a total of 612 AMS observations. Of these 612 observations, seventeen ammonium observations were below the detection limit, while the nitrate and sulfate observations were always above the detection limit. A little less than half (38%) of the chloride observations were above the detection limit. Of the 232 chloride observations above the detection limit, 115 were above $0.30\,\mu g/m^3$ and 117 were between 0.15 and $0.30\,\mu g/m^3$.

3.2.4   Likelihood function for temperature and relative humidity

Figure 4 shows the observed temperature and relative humidity at La Merced. Both the per-minute and per-hour RAMA observations are shown. As can be seen from the gaps in the per-minute data, the data logger failed for extended periods of time. For the periods where the per-minute data is not available, the hourly averaged data was interpolated. The uncertainty in the averaged temperature and relative humidity measurements was assumed to be ±0.6°C and ±1.4%, respectively, at the 95% confidence

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
◄◄	►►
◄	►
Back	Close
Full Screen / Esc	

Printer-friendly Version

Interactive Discussion

EGU

level. The likelihood functions are:

$$p\left(T_{\text{obs}}|T\right) \sim N(T_{\text{obs}}, 0.3) \tag{27}$$

$$p\left(RH_{\text{obs}}|RH\right) \sim N(RH_{\text{obs}}, 0.7) \tag{28}$$

3.3 Selecting the prior

The prior $p(\theta)$ represents the uncertainty of θ before the data arrives: the prior thus contains all the information available about the unknown variables before the experiment begins. Two desirable characteristics of a prior are that it be well-centered near the actual value of the unknown variables and the uncertainty bands should correspond well to the realized discrepancies between actual and predicted values (Draper, 2006)[3].

Information for the prior may come from previous experiments, the scientific literature, expert opinion, constraints provided by knowledge of the physics and chemistry of the system, and so on. For example, if all that is known about a parameter is that it must be greater than zero and below an upper bound, (by Laplace's Principle of Insufficient Reason) an appropriate prior would be a uniform probability density function:

$$X \sim U[\text{lower bound, upper bound}] \tag{29}$$

In general, some information is almost always available, and one of the advantages of the Bayesian approach is that it provides a formal and intuitive mechanism to utilize this information.

A natural selection for random variables that must be positive is the lognormal distribution. Kahn provides an elegant explanation for the applicability of the lognormal distribution to air pollution concentrations (Kahn, 1973). More recently, Ott proposed the theory of successive random dilutions as a physical explanation for the lognormality of pollutant concentrations (Ott, 1990). The use of lognormal pdf's to describe pollutant concentrations has been used by a wide variety of researchers (e.g., Beier,

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

1999; Georgiadis et al., 1998; Hadley and Toumi, 2003; Kan and Chen, 2004; Kao and Friedlander, 1995; Lorenzini et al., 1994; Lu, 2002; Murphy, 1998; Tripathi, 1994).

The lognormal pdf of X is given by

$$p(X) = \frac{1}{\sqrt{2\pi}\sigma X} \exp\left(-\frac{1}{2}\left(\frac{\ln(X) - \mu}{\sigma}\right)^2\right) \qquad (30)$$

Care must be exercised in the notation used: $\overline{X}$ is the mean of the random variable X while μ is the mean of the natural logarithm of X (i.e., $\mu = \overline{\ln(X)}$). These two size parameters are related by

$$\mu = \ln(\overline{X}) - \frac{1}{2}\sigma^2 \qquad (31)$$

The standard deviation of the natural logarithm of X is σ. The standard deviation of X, denoted by σ_X, is given by

$$\sigma_X = \sqrt{\left(e^{2\mu+\sigma^2}\right)\left(e^{\sigma^2} - 1\right)} \qquad (32)$$

Finally, the mode ($\tilde{X}$) of the lognormal distribution, which is the most likely value of the distribution, is given by

$$\tilde{X} = e^{\mu - \sigma^2} \qquad (33)$$

Since we are often interested in the most likely value of a variable, if X is lognormally distributed, this will be denoted as

$$X \sim \log N(\tilde{X}, \sigma) \qquad (34)$$

3.3.1 Prior for aerosol phase species

A uniform prior was used for the inorganic AMS species. The maximum concentration observed for ammonium, nitrate, sulfate, and chloride was 9, 21, 12, and 3 μg/m^3,

EGU

respectively. None of these maxima is above what one would expect in a polluted atmosphere like Mexico City.

The only inorganic aerosol species for which observations are not available are crustal species (Koloutsou-Vakakis and Rood, 1994). Numerous researchers have highlighted the importance of including crustal species in predicting aerosol behavior (e.g., Ansari and Pandis, 1999a; Jacobson, 1999; Koloutsou-Vakakis and Rood, 1994; Moya et al., 2001). Common crustal elements include Al, Si, Fe, Ca, Mg, K, and Na. Since Al, Si, and Fe are present in the form of stable oxides, they do not participate in reactions and do not significantly affect the partitioning of species (Moya et al., 2001). Ca, Mg, K, and Na compounds generally exist as oxides and/or carbonates and can be transformed to water-soluble species, and can affect the distribution of species (Kim and Seinfeld, 1995).

Previous observations has found that geologic material comprises a significant fraction of $PM_{2.5}$ in Mexico City (Chow et al., 2002a), and size resolved observations of aerosols at La Merced found appreciable concentrations of Na, K, and Ca in aerosols with aerodynamic diameter smaller than $1 \mu m$ (Moya et al., 2004). The dry salt lake of Texcoco is ~15 km northeast of the site, covers ~12 km^2 (Moya et al., 2004) and is a likely source of crustal species. We therefore need to allow for the presence of crustal material in our calculations.

While crustal material was not measured at La Merced during MCMA-2003, impactor aerosol collection followed by PIXE analysis was conducted at the CENICA site. The experimental setup and results are described elsewhere (Johnson, 2006). Here we use the 6-h averaged concentrations of elemental Na, K, Mg, and Ca from the two smallest stages (0.07–0.34 μm and 0.34–1.15 μm) to estimate a lognormal prior for equivalent Na, defined as

$$Na_{equiv} = Na + K + 2Mg + 2Ca \qquad (35)$$

where all the concentrations are in molar units (μmol/m^3). Moya et al. (2001) found good agreement between the predictions of SCAPE2, which explicitly includes K, Mg,

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀◀ | ▶▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

and Ca, with those of ISORROPIA, where crustal species are included as equivalent Na, when the models were applied to data from the 1997 IMADA-AVER campaign.

The PIXE observations provide an upper limit to the equivalent sodium concentration for two reasons. First, a relatively small critical orifice was selected for the AMS onboard the AML to better sample smaller particles from fresh vehicle exhaust during chase experiments. Thus, the AML AMS size-dependent collection efficiency was shifted to smaller sizes with a 50% cutoff for large particles of $\sim$0.8 μm. On average, approximately half of the equivalent Na is from the smallest stage (0.07–0.34 μm) of the impactor. Second, the PIXE observations are elemental concentrations, while we are only interested in the concentration of crustal species that can interact with the other inorganic aerosol species.

Since the PIXE observations provide an upper limit for Na_{equiv} for our system, we halved the mode and doubled the standard deviation of the lognormal fit to the PIXE measurements found using the method of moments (e.g., Seinfeld and Pandis, 1998). Figure 5 shows the frequency distribution of the equivalent sodium concentration and the fitted lognormal prior (Observations $\sim$logN(12×10^{-3}, 0.33), as well as the lognormal prior selected for Na_{equiv} ($Na_{equiv}\sim$ logN(6×10^{-3}, 0.65)). Finally, we additionally imposed an upper-limit cut-off of $Na_{equiv}=80\times10^{-3}\mu$mol/m^3.

3.3.2 Lognormal prior for NH_3 and HNO_3

The lognormal prior distributions for NH_3 and HNO_3 used in this work are described in detail by San Martini (2004). Briefly, two sources are used to determine the prior distribution for ammonia: the 1997 IMADA-AVER campaign (Edgerton et al., 1999) and an exploratory campaign undertaken at La Merced during February 2002 (Grutter, 2002). The IMADA-AVER campaign provides 6-h averaged measurements at La Merced (only), and 24-h averaged measurements at 25 different sites throughout the MCMA (Chow et al., 2002a), while the 2002 exploratory campaign yields 6-min NH_3 concentrations measured using the same FTIR system used here. Figure 6 shows

the frequency distribution of the 6-h and 24-h averaged IMADA-AVER observations, the 6-minute averaged observations from February 2002, and the fitted lognormal prior ($NH_3 \sim \log N(0.5, 0.9)$).

During the IMADA-AVER campaign nitric acid measurements were taken as 6-h averages at the La Merced site only. No observations of nitric acid are available from the 2002 exploratory campaign. Given the scarcity of the HNO_3 observations a very diffuse prior was selected (see Fig. 7).

3.3.3 Prior for HCl(g)

To the authors' knowledge no direct observations of HCl(g) are available for the MCMA. Previous work, however, indicates that concentrations of HCl(g) may be appreciable (San Martini et al., 2005). Typical sources of HCl(g) include volatilization of chloride from sea salt particles and other primary particulate matter emissions from natural (e.g., soil dust) and anthropogenic sources. While Mexico City is hundreds of kilometers from the ocean, the dry salt-lake in the northeast of the city is a source of salt particles. Previous work has shown a clear gradient in $PM_{2.5}$ and PM_{10} chloride concentrations, with decreasing concentrations further from the dry lakebed (San Martini, 2004; San Martini et al., 2005).

In addition to the dry lakebed, other sources of (direct or indirect) HCl may be present in the MCMA. The combustion of chlorine- or chloride-containing fossil fuels and the incinerations of chlorine- or chloride-containing refuse are the two major anthropogenic sources of HCl (Saxena et al., 1993). In the U.S., most of the HCl emissions are believed to be due to (bituminous) coal combustion (Saxena et al., 1993); this will not be the case for Mexico City as there is negligible coal combustion in the MCMA. In general, large anthropogenic sources of molecular chlorine (Cl_2) include chemical production facilities, water treatment plants, smelters, and paper production operations (Tanaka et al., 2000). Other anthropogenic sources that commonly contribute to the chlorine budget include dry cleaning operations and solvent use.

The emissions inventory for chlorine sources in the MCMA is sparse. The 2000

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

emissions inventory for the MCMA reports usage of chlorine in the production of aluminum, as well as evaporative emissions of methyl chloroform and perchloroethylene (Secretaría del Medio Ambiente, 2000). Given the high level of photochemical activity generally present in the MCMA efficient conversion of most emitted organic chloride to HCl is expected. Despite the sparseness of the chlorine emissions inventory, the presence of the dry lakebed and the variety of industry in the MCMA suggests appreciable atmospheric emissions of chlorine- and chloride-containing compounds in the MCMA.

In addition to anthropogenic sources, an additional source of chlorine relevant to the MCMA may be the volcano Popocatépetl located southeast of the MCMA. Volcanoes are a major source of HCl to the atmosphere. Allen et al. (2002) measured emissions from the Masaya Volcano, Nicaragua, and observed concentrations of HCl up to $1300\,\mu g/m^3$. Given the predominant winds, the relatively low volcanic activity, and the distance of the volcano from the city (approximately $\sim$60 km from the center of Mexico City), it is unlikely that emissions from the volcano will significantly impact concentrations in the MCMA. However, there may be episodes of higher than normal volcanic activity that coincide with winds from the southeast that contradict this assumption. Raga et al. found that aerosol composition in Mexico City is affected by emissions from Popocatépetl (Raga et al., 1999). They suggest that recirculating flows as observed by (Fast and Zhong, 1998) would provide the mechanism to transport pollutants from aloft into the city. In addition, Moya et al. examined size-differentiated aerosol particles during December 2000–October 2001 and found significantly higher sulfate concentrations during April and June (Moya et al., 2003). The authors attribute this observation to an increase in volcanic activity and predominantly easterly winds (i.e., from the volcano to the city) during this period, as well as ambient conditions that favor sulfate production (high humidity). Thus, given appropriate meteorological conditions and volcanic activity, Popocatépetl may contribute to HCl concentrations in the MCMA.

Finally, a characteristic of the air pollution in the MCMA that is of particular relevance in the question of HCl concentrations is the high concentrations of alkanes (Blake and Rowland, 1995). Chlorine radicals react rapidly with alkanes via hydrogen abstraction

to form HCl(g). Therefore, an urban atmosphere with high concentrations of alkanes and a source of chlorine radicals is likely to have appreciable concentrations of HCl(g).

With no measurements of HCl(g) available for the MCMA, we turn to observations of HCl in other locations to estimate the likely range of HCl(g) concentrations. San
[5] Martini reviewed ambient HCl concentrations in urban locations worldwide, including locations close and far from the coast (San Martini, 2004). Figure 8 summarizes this review, where for each literature source the minimum and maximum observed concentration are shown; the units of the ordinate are arbitrary. Also shown is the assumed prior (HCl~logN(0.02, 1.4)), which allows for HCl concentrations an order of magnitude
[10] greater and smaller than the largest and smallest observation.

4 Results

Figure 9a shows the observed and predicted ammonia concentrations for the period of study. Shown are both the long-path and point observations, as well the mode (black) and 95% confidence interval (dashed black) of the NH_3 posterior distribution. The NH_3
[15] posterior probability density surface is shown in Fig. 9b. The model is able to reproduce the observations well when the two ammonia time series agree. The most significant discrepancies between the two ammonia time series are evident at night and in the morning hours; during these times, the NH_3 posterior probability density is centered on the TILDAS observations. This means that during these periods, given our under-
[20] standing of aerosol thermodynamics, the TILDAS observations are more consistent with the observations. As will be discussed later, the open-path instrument apparently detected a source of ammonia that was not seen in the point measurement. Note that the predicted 95% confidence interval encompasses the long-path FTIR observations, indicating that the FTIR observations are plausible, and that the Markov Chain has
[25] explored the entire solution space.

During the afternoon of 26 April 2003, the NH_3 posterior probability density is more closely centered on the FTIR than TILDAS observations. In this period the FTIR ob-

servations are ~3 ppbv larger than the TILDAS observations; this discrepancy is within the uncertainty of the two time series. Figures 10a and b compare the mode of the NH_3 posterior distribution with the FTIR and TILDAS observations, where the points are colored by the time of day.

Figure 11 shows the observed and predicted nitric acid concentrations. The observed nitric acid concentrations are low (~5 ppbv) at night and in the early morning. At approximately 11:00 a.m. the concentrations of nitric acid start to increase: this increase occurs despite the rise in the boundary layer, clearly pointing to photochemical production of HNO_3 from rush-hour NO_x emissions. The maximum nitric acid concentration is at ~03:00 p.m. The model captures this diurnal profile well on 26 April 2003. On 27 April 2003 the model appears to over-predict afternoon nitric acid concentrations. Moreover, nitric acid concentrations at night and during the morning (before ~11:00 a.m.) are predicted to be significantly below the observations. It is during these periods when the concentrations of HNO_3 are lowest and closest or below the minimum detection limit.

Figures 12 and 13 compare the observed and predicted aerosol concentrations. The model does an excellent job at reproducing the aerosol observations. In particular, note that while inorganic equilibrium models have traditionally performed poorly in predicting aerosol nitrate concentrations, the model has no such difficulties here. The chloride concentrations deserve particular mention. The top panel of Fig. 14 shows the chloride observations, where only those observations above the detection limit $(0.15 \, \mu g/m^3)$ are shown, as well as the mode of the predicted chloride posterior distribution (black squares). The model accurately predicts the chloride observations when the chloride observations are consistently above the detection limit (~01:00 a.m. to ~11:00 a.m.). The HCl(g) concentrations are well constrained in this period, with concentrations generally ~O(ppbv), though higher concentrations are predicted on the 27th (~O(10 ppbv)). In particular, the concentration of HCl(g) is predicted to go from sub-ppbv in the early morning hours to ~1 ppbv at 09:30 a.m. (see Fig. 14c). The concentration of HCl (g) is predicted to increase to ~5 ppbv until ~10:30 a.m., at which

Title Page

Abstract   Introduction

Conclusions   References

Tables   Figures

◄   ►►

◄   ►

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

point the predicted aerosol chloride continues to match the observations well. After this the HCl(g) concentrations increase to ~10 ppbv and higher; however, despite these high gas-phase concentrations, the chloride is predicted to partition mostly to the gas-phase, and the aerosol chloride predictions and observations no longer match well.

5 This behavior is also seen in the afternoon periods, where occasionally the AMS chloride observation is above the detection limit. During these periods the Markov Chain will search in extremely high HCl(g) concentration solution space (~100 ppbv), and still the most likely aerosol phase chloride concentration is negligible. The FTIR setup has not been optimized for detection of HCl(g); however, it is expected that it would de-

10 tect concentrations above 5 ppbv, and certainly concentrations of ~ 100 ppbv. No such signal was detected.

In sum, when the aerosol chloride concentration is consistently above the 0.15 μg/m^3 detection limit, the model is able to accurately reproduce the aerosol chloride concentrations. During these periods predicted HCl(g) concentrations are well constrained

15 and on the order of a couple ppbv. Conversely, when the aerosol chloride signal only occasionally goes above the detection limit, the model either fails to match the aerosol phase concentrations or predicts gas phase concentrations that are unreasonably high.

4.1 Deliquescence versus efflorescence

The aerosols are predicted to be dry during the period of study. Figure 4 shows the RH

20 profile at La Merced: the maximum and minimum relative humidities during the period of study were 56% and 24%. Ansari and Pandis (2000a) have highlighted the importance of considering (metastable) efflorescence behavior. In addition, previous work has indicated that the presence of aerosol chloride may be indicative of metastable behavior (San Martini, 2004). We therefore investigated the effect of assuming the

25 aerosols were metastable (wet) for the period (04:42 a.m.–11:22 a.m.) on 26 April 2003. During this period the RH started at 49%, reached a maximum of 56%, and then decreased to 34%. The chloride concentration was consistently above the detection limit between 04:42 a.m. and 10:06 a.m.. Subsequent to this, the chloride signal var-

ied between being above and below the detection limit until ~11:00 a.m., after which the signal remained below the detection limit. We examined the effect of assuming the aerosols were in the metastable branch for this period only because the activity coefficient model used by ISORROPIA breaks down at high ionic strengths.

5   Differences between the predicted aerosol phase concentrations for the stable versus metastable case were found to be negligible (see Fig. 15). Similarly, the predicted ammonia concentrations for the two cases are very similar (see Fig. 16). Conversely, differences between the predicted gas phase nitric and hydrochloric acid are evident (see Figs. 17 and 18), where in both cases the acid concentrations are slightly higher

10 for the metastable case. The available data are insufficient to discriminate between stable and metastable behavior in this case given the excellent agreement between the ammonia and aerosol phase predictions for the two cases, as well as the uncertainties in the nitric acid observations.

    Finally, it is interesting to note that despite the high concentrations of ammonia ob-

15 served at La Merced, when the aerosols are assumed to be metastable they are predicted to be acidic. The mode of the pH posterior distribution varies between 2.5 and 4.0 pH units during this period. Note, however, that after ~10:00 a.m. the ionic strength of the aerosols is predicted to be higher than 60 mol/kg. The errors associated with the activity coefficient model used by ISORROPIA are significant at these high ionic

20 strengths. Excluding these points yields an average ionic strength of 41 mol/kg.

4.2   Equilibrium constant $K_P(NH_4NO_3)$

As discussed in Sect. 3.1, the value of the equilibrium constant used by ISORROPIA for the dissociation of solid ammonium nitrate was changed based on the work of (Mozurkewich, 1993). Figures 19 and 20 compare the observations with the model

25 predictions of NH_3 and HNO_3, and Fig. 21 compares the mode of the NH_3, HNO_3, and HCl distribution, respectively, predicted using the original $K_p(NH_4NO_3)$ based on the thermodynamic tables of (Wagman et al., 1982) and the modified $K_p(NH_4NO_3)$ of (Mozurkewich, 1993). Differences between the NH_3 and HCl concentration for the

two cases are small; for HNO_3 the higher value of $K_p(NH_4NO_3)$ used by the original formulation of ISORROPIA results in HNO_3 concentrations that are ~20% higher. At night and during the early morning hours the model predictions of HNO_3 are below the observations regardless of which equilibrium constant is used. The over-prediction of HNO_3 for the afternoon of the 27th discussed previously is exacerbated by the use of $K_p(NH_4NO_3)$ based on Wagman's data. We wish to emphasize, however, that carefully controlled laboratory conditions rather than field conditions are a more appropriate means to determine and validate thermodynamic parameters. The finding that the NH_3 TILDAS point observations are more consistent with all the available measurements and our knowledge of thermodynamics is robust using both the value of $K_p(NH_4NO_3)$ suggested by Mozurkewich and that used in the original formulation of ISORROPIA.

5 Conclusions

ISORROPIA was embedded in a Markov Chain Monte Carlo algorithm to produce a powerful tool to analyze concentrations of inorganic aerosol and gas-phase precursors. The method allows for the direct incorporation of measurement uncertainty, provides a formal framework for including prior knowledge and datasets of different quality. The method was successfully applied to data taken at La Merced during the MCMA-2003 field campaign. The model was able to reproduce observed aerosol concentrations extremely well, as well as provide an excellent constraint for gas-phase concentrations.
   Observations of ammonia from two different instruments, a long-path FTIR and a TILDAS point sampler, were provided to the model. The model did an excellent job at reproducing the gas-phase ammonia concentrations during periods where the observed ammonia concentrations agreed. During periods when the two time series diverged the model predicts that the observations from the TILDAS, which was co-located to the aerosol measurements, are more likely to be consistent with all the available observations and our knowledge of aerosol thermodynamics.
   The FTIR and TILDAS NH_3 observations are not expected to be identical. Long-

path observations tend to smooth out gradients from very local sources and pick up plumes that are not observed with the point samplers. Moreover, the different sampling heights (~20 m versus ~2 m) of the two instruments may result in the sampling of different air masses. In particular, the largest discrepancies are seen at night: a plausible explanation would be a highly stratified nighttime atmosphere.

While both vertical and horizontal inhomogeneities are plausible explanations to the observed discrepancies, one would expect that these inhomogeneities would also be reflected in other pollutants. However, an intercomparison of observations of SO_2, formaldehyde, CO, CO_2, toluene and benzene did not find a comparable difference in the time series from instruments measuring along a ~0.5 km open-path versus in the AML. It is therefore unlikely that vertical inhomogeneities cause the difference in NH_3 observations. Rather, it is hypothesized that during the day, when there is good mixing in the boundary layer, mobile sources dominate NH_3 emissions. During these times the two time series agree well. At night and early morning, another (local) ammonia source becomes dominant presumably closer to the east-end of the light path and thus is not observed by the TILDAS instrument, which was located at the west-end of the FTIR beam trajectory. The low mixing height at night emphasizes the importance of the local source; the rising boundary layer during the day, combined with increased NH_3 emissions and mixing, allows the measurements to converge during the day.

The gas phase concentration and diurnal profile of HNO_3 measured with the FTIR along the open path is reasonably well predicted by the model. This indicates that this gas is more homogenously distributed around the measurement site and can be accurately measured with the spectroscopic technique. The model predicts the rise in photochemically produced nitric acid in the early afternoon, and does an excellent job at predicting the maximum concentration on the 26 April, while on the 27 April the maximum concentration is slightly overpredicted. Nighttime and morning concentrations appear to be significantly below the measured concentrations, with sub-ppb concentrations predicted between ~03:00 a.m. and ~09:00 a.m. both days.

Finally the model is able to predict HCl concentrations despite the lack of measured

data. When aerosol chloride concentrations are consistently above the $0.15\,\mu\text{g/m}^3$ detection limit, the model is able to accurately reproduce the aerosol chloride concentrations and the most likely concentration of HCl(g) is predicted to vary between 0.4 and 5 ppbv.

Appendix A

Practical implementation issues

A1   Probing distribution

The question of what is the "best" probing distribution for a particular problem is a question that has bedeviled MCMC practitioners from the inception of the method to this day. In part, this is because essentially any probing distribution will (eventually) work: the stationary distribution for just about any probing distribution is the desired $p(\theta|\text{Data})$ (Gilks et al., 1996). To date, no one has established a general method of choosing a probing distribution that always leads to a well-mixed chain. Given this caveat, two suggested characteristics of a successful probing distribution are (Draper, 2006)[3]:

1. Choose a probing distribution that approximates an overdispersed version of the posterior distribution that is being sampled from;

2. Choose a probing distribution whose expected value for each proposed move (θ^*), given the current state θ_t, is to remain at θ_t (i.e., $E(\theta^*|\theta_t)=\theta_t$).

The second suggestion ensures that there will be an approximate left-right balance, which encourages rapid exploration of the entire solution space (Draper, 2006)[3]. One way of ensuring this is to select a symmetric probing distribution, i.e., $PD(\theta^*|\theta)=PD(\theta|\theta^*)$.

EGU

θ is a ten-dimensional vector comprising 9 continuous (T, RH, NH_3, HNO_3, HCl, NH_4, Na, NO_4, SO_4, Cl, H_2O) and one binary (M) variable. For the continuous variables we use a random-walk Metropolis algorithm, which is easy to program and has good MCMC convergence properties (Draper, 2006)[3]. The algorithm uses a multi-variate

5 normal distribution whose mean is the current position and variance plays the role of a tuning parameter that can be varied to ensure good mixing. However, since generating a multi-variate distribution is computationally inconvenient, we use the approach outlined in (Rao, 1992) to generate the correlated normal draws. A modified Choleski decomposition is applied to an estimate of the probing distribution covariance matrix.

10 We sample from a uni-variate normal distribution 9 times and multiply this vector by the diagonal matrix calculated in the modified Choleski decomposition to generate a 9-element vector of independently distributed normal numbers. The proposed Markov step is generated by multiplying this vector by the lower triangular matrix calculated from the modified Choleski decomposition of the estimated probing distribution covari-

15 ance matrix (San Martini, 2004). The probing distribution for the binary variable, which is independent of the other components of θ, is derived from a uniform distribution:

$$M = \begin{cases} M_{FTIR} & \text{if} \quad 0 \le u \le 0.5 \\ M_{TILDAS} & \text{if} \quad 0.5 < u \le 1.0 \end{cases} \tag{A1}$$

A well-chosen probing distribution will favor convergence. We want a Markov chain that mixes well, or, in the words of Draper, "that moves around freely, happily jumping

20 all over the place" (Draper, 2006)[3]. A MCMC simulation with either too high or low an acceptance probability is suspicious: a high acceptance probability indicates that the Markov steps are too small so that the simulation moves very slowly through the target distribution, while a small acceptance probability may lead the Markov chain to stand still most of the time. Adaptive Metropolis sampling (Gelman et al., 1996) was used in

25 this work to ensure an optimal acceptance probability ($\sim$20%).

A2 Convergence diagnostics

Convergence diagnostics attempt to answer the question of when it is reasonable to believe that the samples generated by the MCMC simulation are representative of the underlying equilibrium distribution. This is a more general notion of convergence than for usual iterative procedures because the MCMC simulation provides neither a single value nor a distribution, but rather a sample from a distribution. Moreover, the nature of the MCMC algorithm is that the samples will generally be correlated (Cowles and Carlin, 1996). While ensuring convergence is crucial when applying MCMC methods in general, it is particularly important here due to the bimodal nature of the gas phase posterior distributions.

For an excellent comparative review of MCMC convergence diagnostics see (Cowles and Carlin, 1996). While methods such as MCMC are used in many application areas by a wide population of experimenters, these experimenters often have little knowledge or interest in the fine details of convergence diagnostics (Heidelberger and Welch, 1983). Fortunately, the need for a tool that provides a variety of automated convergence diagnostics has been fulfilled. Researchers at the Medical Research Council Biostatistics Unit in Cambridge, UK have developed and kindly made freely available by anonymous ftp://ftp.mrc-bsu.cam.ac.uk) Convergence Diagnosis and Output Analysis (CODA) software, a set of S+ functions that calculates a variety of MCMC convergence diagnostics (Best et al., 1995; Best et al., 1997). LeSage has implemented some of the CODA diagnostics in Matlab and kindly made these freely available (http://www.spatial-econometrics.com/). The Matlab code of LeSage was used in this work, and the contributions of Best, Cowles, Vines, and LeSage are gratefully acknowledged.

Since it is not possible to guarantee that a finite sample from an MCMC algorithm is representative of the underlying stationary distribution, (Cowles and Carlin, 1996) recommend using a variety of diagnostic tools rather than any single plot or statistic, as well as visually inspecting the chain trace plots. The latter suggestion is impractical

given the number of data points examined here. Rather, we use a two-part convergence test. If the Markov chain fails either of the tests, a trace plot is generated, inspected and, either the resulting distribution is omitted from the final results due to a lack of convergence or the MCMC simulation is repeated using a longer burn in,
more thinning, or a modified probing distribution covariance matrix so that the tests are passed. The two convergence tests utilized here are due to Raferty-Lewis (Raferty and Lewis, 1992) and Geweke (Geweke, 1992).

The Raferty-Lewis test is based on two-state Markov chain theory and standard sample size formulas (Raferty and Lewis, 1992). The convergence diagnostic requires that
the user specify the quantile q to be estimated (here the 2.5th percentile), the desired accuracy r (here ±0.01), and the required probability s of attaining the specified accuracy (here 0.95). The diagnostic then calculates the number of iterations N needed to estimate the specified quantile to the desired precision, as well as N_{min}, the minimum number of iterations needed if the samples were independent, identically distributed
draws. The ratio of these two quantities defines the independence factor $I \equiv N/N_{min}$; thus I is a measure of within-chain correlation, where large values suggest probable convergence failure. Per the recommendation of Raferty and Lewis, the diagnostic indicates convergence failure if $I > 5.0$. The diagnostic is also deemed to fail if N is greater than the number of draws in the Markov chain (7000).

The Geweke diagnostic tests whether an equilibrium has been attained by comparing the means of the first 20% versus the last 50% of the sample (Geweke, 1992). The mean of these two splits is not equal if the Markov chain has not reached an equilibrium state. The output of the test is the two-sided p-value of the tail probability associated with the computed Z statistic. If the null hypothesis is rejected, this indicates that
the chain has not converged. Here, a p-value of 0.05 was considered strong enough evidence to reject stationarity.

References

Allen, A. G., Harrison, R. M., and Erisman, J.-W.: Field Measurements of the Dissociation of Ammonium Nitrate and Ammonium Chloride Aerosols, Atmos. Environ., 23(7), 1591–1599, 1989.

Allen, A. G., Oppenheimer, C., Ferm, M., Baxter, P. J., Horrocks, L. A., Galle, B., McGonigle, A. J. S., and Duffel, H. J.: Primary Sulfate Aerosol and Associated Emissions from Masaya Volcano, Nicaragua, J. Geophys. Res., 107(D23), 4682, 2002.

Ansari, A. S. and Pandis, S. N.: An Analysis of Four Models Predicting the Partitioning of Semivolatile Inorganic Aerosol Components, Aerosol Sci. Technol., 31(2–3), 129–153, 1999a.

Ansari, A. S. and Pandis, S. N.: Prediction of Multicomponent Inorganic Atmospheric Aerosol Behavior, Atmos. Environ., 33, 745–757, 1999b.

Ansari, A. S. and Pandis, S. N.: The Effect of Metastable Equilibrium States on the Partitioning of Nitrate between the Gas And Aerosol Phases, Atmos. Environ., 34, 157–168, 2000a.

Ansari, A. S. and Pandis, S. N.: Water Absorption by Secondary Organic Aerosol and Its Effect on Inorganic Aerosol Behavior, Environ. Sci. Technol., 34(1), 71–77, 2000b.

Bassett, M. E. and Seinfeld, J. H.: Atmospheric Equilibrium Model of Sulfate and Nitrate Aerosol, Atmos. Environ., 17, 2237–2252, 1983.

Bassett, M. E. and Seinfeld, J. H.: Atmospheric Equilibrium Model of Sulfate and Nitrate Aerosols II. Particle Size Analysis, Atmos. Environ., 18, 1163–1170, 1984.

Beier, R.: Calculation of expected air quality parameters based on monitoring results and modelling of planned additional sources, Gefahrstoffe Reinhaltung Der Luft, 59(4), 143–149, 1999.

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◄◄   ►►

◄   ►

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

Bertsekas, D. P. and Tsitsiklis, J. N.: Introduction to Probability. Athena Scientific, Belmont, MA, 430 pp., 2002.

Best, N., Cowles, M. K., and Vines, K.: CODA Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output, Version 0.30, MRC Biostatistics Unit, Cambridge, UK, 1995.

Best, N., Cowles, M. K., and Vines, K.: CODA Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output, Version 4.0 (Addendum to Manual), MRC Biostatistics Unit, Cambridge, UK, 1997.

Blake, D. R. and Rowland, F. S.: Urban Leakage of Liquefied Petroleum Gas and Its Impact on Mexico City Air Quality, Science, 269(5226), 953–956, 1995.

Blanchard, C. L. and Hidy, G. M.: Effects of Changes in Sulfate, Ammonia, and Nitric Acid on Particulate Nitrate Concentrations in the Southeastern United States, J. Air Waste Manage. Assoc., 53, 283–290, 2003.

Blanchard, C. L., Roth, P. M., Tanenbaum, S. J., Ziman, S. D., and Seinfeld, J. H.: The Use of Ambient Measurements to Identify which Precursor Species Limit Aerosol Nitrate Formation, J. Air Waste Manage.Assoc., 50, 2073–2082, 2000.

Chow, J. C., Watson, J. G., Edgerton, S. A., and Vega, E.: Chemical Composition of $PM_{2.5}$ and PM_{10} in Mexico City during Winter 1997, The Science of the Total Environment, 287, 177–201, 2002a.

Chow, J. C., Watson, J. G., Edgerton, S. A., Vega, E., and Ortiz, E.: Spatial Differences in Outdoor PM_{10} Mass and Aerosol Composition in Mexico City, J. Air Waste Manage. Assoc., 52, 423–434, 2002b.

Clegg, S. L., Seinfeld, J. H., and Brimblecombe, P.: Thermodynamic modelling of aqueous aerosols containing electrolytes and dissolved organic compounds, J. Aerosol Sci., 32(6), 713–738, 2001.

Clegg, S. L., Seinfeld, J. H., and Brimblecombe, P.: Thermodynamic Modelling of Aqueous Aerosols Containing Electrolytes and Dissolved Organic Compounds, Aerosol Sci., 32, 713–778, 2003.

Coleman, H. W. and Steele, W. G.: Experimentation and Uncertainty Analysis for Engineers, John Wiley & Sons, New York, 1999.

Cowles, M. K. and Carlin, B. P.: Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review, J. Am. Stat. Assoc., 91(434), 883–904, 1996.

Edgerton, S. A., Bian, X., Doran, J. C., Fast, J. D., Hubbe, J. M., Malone, E. L., Shaw, W. J.,

Arriaga, J. L., Ortiz, E., Ruiz, M., Sosa, G., Vega, E., Limon, T., Guzman, F., Archuleta, J., Bossert, J. E., Elliot, S. M., Lee, J. T., McNair, L. A., Chow, J. C., Watson, J. G., Coulter, R. L., Doskey, P. V., Gaffney, J. S., Marley, N. A., Neff, W., and Petty, R.: Particulate Air Pollution in Mexico City: A Collaborative Research Project, J. Air Waste Manage. Assoc., 49, 1221–1229, 1999.

Evans, J., Levy, J., Hammitt, J., Burgoa, C. S., Casteillejos, M., Ramirez, M. C., Avila, M. H., Rodriguez, H. R., Bracho, L. R., Trespalacios, P. S., Spengler, J. D., and Suh, H.: Health Benefits of Air Pollution Control, in: Air Quality in the Mexico Megacity: An Integrated Assessment, edited by: Molina, L. T. and Molina, M. J., Alliance for Global Sustainability Bookseries, Kluwer Academic Publishers, Dordrecht, Netherlands, 105–136, 2002.

Fast, J. D. and Zhong, S.: Meteorological Factors Associated with Inhomogeneous Ozone Concentrations within the Mexico City Basin, J. Geophys. Res., 103(D15), 18 927–18 946, 1998.

Ferson, S. and Ginzburg, L. R.: Different Methods are Needed to Propagate Ignorance and Variability, Reliability Engineering and System Safety, 54, 133–144, 1996.

Gelman, A., Roberts, G. O., and Gilks, W.: Efficient Metropolis Jumping Rules, in: Bayesian Statistics 5, edited by: Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., Oxford University Press, New York, NY, 1996.

Georgiadis, T., Fortezza, F., Alberti, L., Strocchi, V., Marani, A., and Dal Bo, G.: Probability density functions of photochemicals over a coastal area of Northern Italy, Nuovo Cimento Della Societa Italiana Di Fisica C-Geophysics and Space Physics, 21(1), 75–84, 1998.

Geweke, J.: Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior Moments, in: Bayesian Statistics 4, edited by: Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., Oxford University Press, Oxford, UK, 169–193, 1992.

Gilks, W., Richardson, S., and Spiegelhalter, D.: Introducing Markov Chain Monte Carlo, in: Markov Chain Monte Carlo in Practice, edited by: Gilks, W., Richardson, S., and Spiegelhalter, D., Chapman & Hall, London, UK, 1–21, 1996.

Grutter, M.: Open-path FTIR measurements near downtown Mexico City, Mexico City Air Quality Project Workshop, Mexico City, Mexico, 2002.

Grutter, M., Flores, E., Andraca-Ayala, G., and Báez, A.: Formaldehyde levels in downtown Mexico City during 2003, Atmos, Environ,, 39(6), 1027–1034, 2005.

Grutter, M., Flores, E., Basaldud , R., and Ruiz-Suarez, L. G.: Open-path FTIR spectroscopic studies of the trace gases over Mexico City, Atmos. Oceanic Optics, 16(3), 232–236, 2003.

Hadley, A. and Toumi, R.: Assessing changes to the probability distribution of sulphur dioxide in the UK using a lognormal model, Atmos. Environ., 37(11), 1461–1474, 2003.

Harrison, R. M. and Allen, A. G.: Measurements of Atmospheric HNO_3, HCl and Associated Species in a Small Network in Eastern England, Atmos. Environ., 24A(2), 369–376, 1990.

Heidelberger, P. and Welch, P. D.: Simulation Run Length Control in the Presence of an Initial Transient, Operations Res., 31(6), 1109–1144, 1983.

Herndon, S. C., Jayne, J. T., Zahniser, M. S., Worsnop, D. R., Knighton, W. B., Alwine, E., Lamb, B., Zavala, M., Nelson, D. D., McManus, B., Shorter, J. H., Canagaratna, M. R., Onasch, T. B., and Kolb, C. E.: Characterization of urban pollutant emission fluxes and ambient concentration distributions using a mobile laboratory with rapid response instrumentation, Faraday Discuss., 130, 327–339, 2005.

Hildemann, L. M., Russel, A. G., and Cass, G. R.: Ammonia and Nitric Acid Concentrations in Equilibrium with Atmospheric Aerosols: Experiment vs. Theory, Atmos. Environ., 18(9), 1737–1750, 1984.

Jacobson, M. Z.: Studying the Effects of Calcium and Magnesium on Size-Distributed Nitrate and Ammonium with EQUISOLV II, Atmos. Environ., 33(2), 3635–3649, 1999.

Jacobson, M. Z., Tabazadeh, A., and Turco, R. P.: Simulating equilibrium within aerosols and nonequilibrium between gases and aerosols, J. Geophys. Res., 101(D4), 9079–9091, 1996.

Jayne, J. T., Leard, D. C., Zang, X., Davidovits, P., Smith, K. A., Kolb, C. E., and Worsnop, D. R.: Development of an aerosol mass spectrometer for size and composition analysis of submicron particles, Aerosol Sci. Technol., 33(1–2), 49–70, 2000.

Johnson, K. S. , de Foy, B., Zuberi, B., Molina, L. T., Molina, M. J., Xie, Y., Laskin, A., and Shutthanandan, V.: Aerosol composition and source apportionment in the Mexico City Metropolitan Area with PIXE/PESA/STIM and multivariate analysis, Atmos. Chem. Phys. Discuss., 6, 3997–4022, 2006.

Kahn, H. D.: Note on the distribution of air pollutants, J. Air Pollut. Control Assoc., 23, 973, 1973.

Kan, H. D. and Chen, B. H.: Statistical distributions of ambient air pollutants in Shanghai, China, Biomedical Environ. Sci., 17(3), 366–372, 2004.

Kao, A. S. and Friedlander, S. K.: Frequency-Distributions of Pm(10) Chemical-Components and Their Sources, Environ. Sci. Technol., 29(1), 19–28, 1995.

Katrib, Y., Martin, S. T., Rudich, Y., Davidovits, P., Jayne, J. T., and Worsnop, D. R.: Density changes of aerosol particles as a result of chemical reaction, Atmos. Chem. Phys., 5, 275–

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

291, 2005.

Kim, Y. P. and Seinfeld, J. H.: Atmospheric Gas-Aerosol Equilibrium: III. Thermodynamics of Crustal Elements, Ca^{2+}, K^+, and Mg^{2+}. Aerosol Sci. Technol., 22, 93–110, 1995.

Kim, Y. P., Seinfeld, J. H., and Saxena, P.: Atmospheric Gas-Aerosol Equilibrium I. Thermodynamic Model, Aerosol Sci. Technol., 19, 157–181, 1993.

Kolb, C. E., Herndon, S. C., McManus, B., Shorter, J. H., Zahniser, M. S., Nelson, D. D., Jayne, J. T., Canagaratna, M. R., and Worsnop, D. R.: Mobile laboratory with rapid response instruments for real-time measurements of urban and regional trace gas and particulate distributions and emission source characteristics, Environ. Sci. Technol., 38(21), 5694–5703, 2004.

Koloutsou-Vakakis, S. and Rood, M. J.: The $(NH_4)_2SO_4$-Na_2SO_4-H_2O System: Comparison of Deliquescence Humidities Measured in the Field and Estimated from Laboratory Measurements and Thermodynamic Modeling, Tellus, 46B, 1–15, 1994.

Li, Y. Q., Demerjian, K. L., Zahniser, M. S., Nelson, D. D., McManus, J. B., and Herndon, S. C.: Measurement of formaldehyde, nitrogen dioxide, and sulfur dioxide at Whiteface Mountain using a dual tunable diode laser system, J. Geophys. Res.-Atmos., 109, D16S08, doi:10.1029/2003JD004091, 2004.

Lorenzini, G., Nali, C., and Panicucci, A.: Surface Ozone in Pisa (Italy) – a 6-Year Study, Atmos. Environ., 28(19), 3155–3164, 1994.

Lu, H. C.: The statistical characters of PM10 concentration in Taiwan area, Atmos. Environ., 36(3), 491–502, 2002.

McMurry, P. H.: A Review of Atmospheric Aerosol Measurements, Atmos. Environ., 34, 1959–1999, 2000.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E.: Equations of state calculations by fast computing machines, J. Chem. Phys., 21, 1087–1092, 1953.

Molina, L. T. and Molina, M. J.: Air Quality in the Mexico Megacity: An Integrated Assessment, 2. Kluwer Academic Publishers, Dordrecht, Netherlands, 408 pp., 2002.

Moya, M., Ansari, A., and Pandis, S.: Partitioning of Nitrate and Ammonium between the Gas and Particulate Phases during the 1997 IMADA-AVER Study in Mexico City, Atmos. Environ., 35, 1791–1804, 2001.

Moya, M., Castro, T., Zepeda, M., and Baez, A.: Characterization of Size-Differentiated Inorganic Composition of Aerosols in Mexico City, Atmos. Environ., 37, 3581–3591, 2003.

Moya, M., Grutter, M., and Báez, A.: Diurnal variability of size-differentiated inorganic aerosols

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU

and their gas-phase precursors during January and February of 2003 near downtown Mexico City, Atmos. Environ., 38(33), 5651–5661, 2004.

Mozurkewich, M.: The Dissociation Constant of Ammonium Nitrate and its Dependence on Temperature, Relative Humidity and Particle Size, Atmos. Environ., 27A(2), 261–270, 1993.

Murphy, B. L.: Dealing with uncertainty in risk assessment, Human and Ecological Risk Assessment, 4(3), 685–699, 1998.

Nenes, A., Pandis, S., and Pilinis, C.: ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols, Aquatic Geochem., 4, 123–152, 1998.

Nenes, A., Pandis, S. N., and Pilinis, C.: Continued Development and Testing of a New Thermodynamic Aerosol Module for Urban and Regional Air Quality Models, Atmos. Environ., 33(10), 1553–1560, 1999.

Ott, W. R.: A Physical Explanation of the Lognormality of Pollutant Concentrations, J. Air Waste Manage. Assoc., 40(10), 1378–1383, 1990.

Pate-Cornell, M .E.: Uncertainties in Risk Analysis: Six Levels of Treatment, Reliability Engineering and System Safety, 54, 95–111, 1996.

Pilinis, C. and Seinfeld, J. H.: Continued Development of a General Equilibrium Model for Inorganic Multicomponent Atmospheric Aerosols, Atmos. Environ., 21(11), 2543–2466, 1987.

Pilinis, C. and Seinfeld, J. H.: Development and Evaluation of a Eulerian Photochemical Gas-Aerosol Model, Atmos. Environ., 22(11), 1985–2001, 1988.

Raferty, A. E. and Lewis, S. M.: How Many Iterations in the Gibbs Sampler?, in: Bayesian Statistics 4, edited by: Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M., Oxford University Press, Oxford, UK, 763–774, 1992.

Raga, G. B., Lok, G. L., Baumgardner, D., Baez, A., and Rosas, I.: Evidence for Volcanic Influence on Mexico City Aerosols, Geophys. Res. Lett., 26(8), 1149–1152, 1999.

Rao, S. S.: Reliability-Based Design, McGraw-Hill, Inc., New York, NY, 1992.

Rothman, L. S., Barbe, A., Benner, D. C., Brown, L. R., Camy-Peyret, C., Carleer, M. R., Chance, K. V., Clerbaux, C., Dana, V., Devi, V. M., Fayt, A., Flaud, J.-M., Gamache, R. R., Goldman, A., Jacquemart, D., Jucks, K. W., Lafferty, W. J., Mandin, J.-Y., Massie, S. T., Nemtchinov, V., Newnham, D. A., Perrin, A., Rinsland, C. P., Schroeder, J., Smith, K. M., Smith, M. A. H., Tang, K., Toth, R. A., Van der Auwera, J., Varanasi, P., and Yoshino, K.: The HITRAN molecular spectroscopic database: edition of 2000 including updates through 2001, J. Quant. Spectrosc. Radiat. Transfer, 82, 5–44, 2003.

Rothman, L. S., Rinsland, C. P., Goldman, A., Massie, S. T., Edwards, D. P., Flaud, J. M.,

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Perrin, A., Camy-Peyret, C., Dana, V., Mandin, J. Y., Schroeder, J., McCann, A., Gamache, R. R., Wattson, R. B., Yoshino, K., Chance, K. V., Jucks, K. W., Brown, L. R., Nemtchinov, V., and Varanasi, P.: The HITRAN molecular spectroscopic database and HAWKS (HITRAN Atmospheric Workstation): 1996 edition, J. Quant. Spectrosc. Radiat. Transfer, 60(5), 665–710, 1998.

Russel, A. G., McCue, K. F., and Cass, G. R.: Mathematical Modeling of the Formation of Nitrogen-Containing Air Pollutants. 1. Evaluation of an Eularian Photochemical Model, Environ. Sci. Technol., 22(3), 263–271, 1988.

Salcedo, D., Onasch, T. B., Dzepina, K., Canagaratna, M. R., Zhang, Q., Huffman, J. A., De Carlo, P. F., Jayne, J. T., Mortimer, P., Worsnop, D. R., Kolb, C. E., Johnson, K. J., Zuberi, B., Marr, L. C., Volkamer, R., Molina, L. T., Molina, M. J., Cardenas, B., Bernabé, R. M., Marquez, C., Gaffney, J. S., Marley, N. A., Laskin, A., Shutthanadan, V., Xie, Y., Brune, W. H., Lesher, R., Shirley, T. R., and Jiminez, J. L.: Characterization of Ambient Aerosols in Mexico City during the MCMA-2003 Campaign with Aerosol Mass Spectrometry: results from the CENICA Supersite, Atmos. Chem. Phys., 6, 925–946, 2006.

San Martini, F. M.: Decision Support Tools for Urban Air Quality Management. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 500 pp., 2004.

San Martini, F. M., West, J. J., Sosa, G., Molina, L. T., Molina, M. J., and McRae, G. J.: Modeling Inorganic Aerosols and their Response to Changes in Precursor Concentration in Mexico City, J. Air Waste Manage. Assoc., 55(6), 803–815, 2005.

San Martini, Dunlea, E. J., Volkamer, R., et al.:Implementation of a Markov Chain Monte Carlo Method to Inorganic Aerosol Modeling of Observations from the MCMA-2003 Campaign. Part II: Model Application to the CENICA, Pedregal and Santa Ana Sites, Atmos. Chem. Phys. Discuss., 6, 5999–6040, 2006.

Saxena, P., Mueller, P. K., and Hildemann, L. M.: Sources and Chemistry of Chlorides in the Troposphere: A Review, in: Managing Hazardous Air Pollutants, edited by: Chow, W. and Connor, K. K. ,CRC Press, Inc., Boca Raton, Florida, 173–190, 993.

Secretaría del Medio Ambiente, G. D. F.: Inventario de Emisiones a la Atmosfera, Secretaría del Medio Ambiente, Gobierno del Distrito Federal, Cuidad del Esperanza, Mexico, 2000.

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

Slowik, J., Stainken, K., Davidovits, P., Williams, P. I., Jayne, J. T., Kolb, C. E., Worsnop, D. R., Rudich, D. R., De Carlo, P. F., and Jiminez, J. L.: Particle Morphology and Density Character-

Implementation of a MCMC method to aerosol modeling in MCMA

F. M. San Martini et al.

ization by Combined Mobility and Aerodynamic Diameter Measurements. Part 2: Application to Combustion-Generated Soot Aerosols as a Function of Fuel Equivalence Ratio, Aerosol Sci. Technol., 38(12), 1206–1222, 2004.

Solomon, P. A., Fall, T., Salmon, L., Lynn, P., Vasquez, F., and Cass, G. R.: Acquisition of Acid Vapor And Aerosol Concentration Data for Use in Dry Deposition Studies in the South Coast Air Basin. ARB Research Contract No. A4-144-32, Environmental Quality Laboratory, California Institute of Technology, Pasadena, California, 1988.

Sturges, W. T. and Harrison, R. M.: The Use of Nylon Filters to Collect HCl: Efficiencies, Interferences and Ambient Concentrations, Atmos. Environ., 23(9), 1987–1996, 1989.

Tanaka, P. L., Oldfield, S., Mullins, C. B., and Allen, D. T.: Anthropogenic Sources of Chlorine in Urban Atmospheres. Environ. Sci. Technol., 34, 4470–4473, 2000.

Tripathi, A.: Airborne Lead Pollution in the City of Varanasi, India, Atmos. Environ., 28(14), 2317–2323, 1994.

Wagman, D .D., Evans, W. H., Parker, V. B., Schumm, R. H., Harlow, I., Baily, S. M., Churney, K. L., and Nuttall, R. L.: The NBS Tables of Chemical Thermodynamic Properties, J. Phys. Chem. Ref. Data, 11, Supplement 2, 1982.

Wexler, A. and Seinfeld, J.: Second-Generation Inorganic Aerosol Model, Atmos. Environ., 25A(12), 2731–2784, 1991.

Wexler, A. and Seinfeld, J. H.: The distribution of ammonia salts among a size and composition dispersed aerosol, Atmos. Environ., 24A, 1231–1246, 1990.

Wexler, A. S. and Seinfeld, J. H.: Analysis of Aerosol Ammonium Nitrate: Departures from Equilibrium During SCAQS, Atmos. Environ., 26A(4), 579–591, 1992.

Zhang, J., Chameides, R., Weber, R., Cass, G., Orsini, D., Edgerton, E., Jongejan, P., and Slanina, J.: An Evaluation of the Thermodynamic Equilibrium Assumption for Fine Particulate Composition: Nitrate and Ammonium during the 1999 Atlanta Supersite Experiment, J. Geophys. Res., 108(D7), SOS 2, 1–11, 2003.

Zhang, Y., Seigneur, C., Seinfeld, J. H., Jacobson, M., Clegg, S. L., and Binkowski, F. S.: A comparative review of inorganic aerosol thermodynamic equilibrium modules: similarities, differences, and their likely causes, Atmos. Environ., 34, 117–137, 2000.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Thermodynamic parameters for Eq. (9) for the dissociation of ammonium nitrate used by different models (T^o=298.15 K).

Model	$K(T^o)$ (ppb^2)	a	b
MIT-IAM, MARS-R	41.99	−74.7351	6.025
ISORROPIA[*], EQUISOLV II, SCAPE2	57.46	−74.38	6.12
SEQUILIB	29.86	−75.108	13.456

[*] Earlier versions of ISORROPIA used the same parameters as SEQUILIB. ISORROPIA was modified for use in this work by changing the values of K(To), a, and b for the dissociation of solid ammonium nitrate to the values suggested by Mozurkewich (1993), also used by MIT-IAM and MARS-A.

EGU

Table 2. Measurement uncertainties for NH_3, HNO_3 and NO_z.

Species	Instrument	s	±%
NH_3	FTIR and TILDAS	0.15	29
HNO_3	FTIR	0.25	49
NO_z	NO_y, NO_2 TILDAS	0.25	49

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU
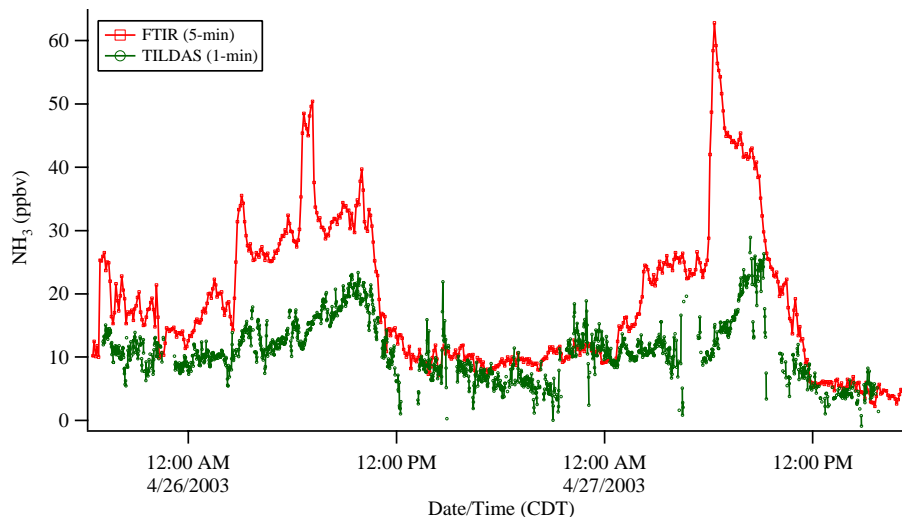
Fig. 1. Ammonia observations at La Merced between 25 April 2003 and 27 April 2003 taken with a TILDAS instrument onboard the AML and with a long-path FTIR instrument located on a building rooftop. The uncertainty in both time series is ±9%.
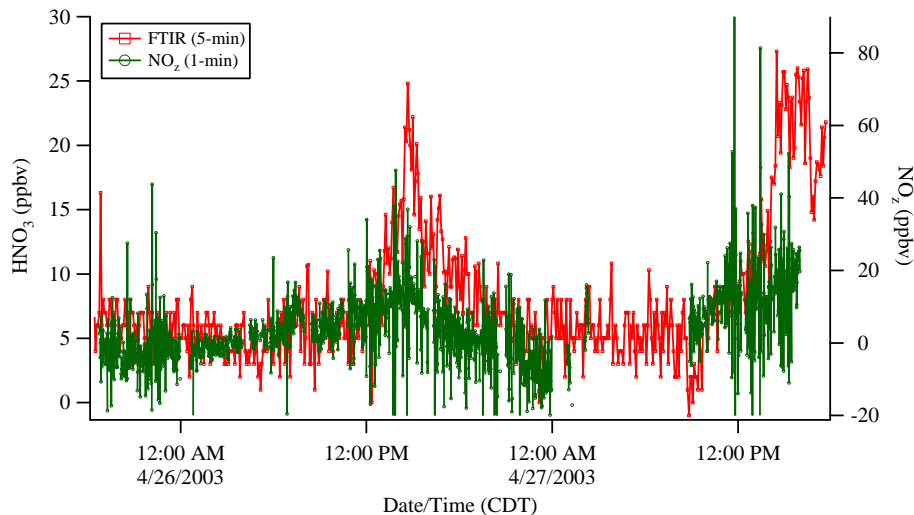
Fig. 2. Nitric acid and NO_z observations at La Merced between 25 April 2003 and 27 April 2003 taken with a NO_y and TILDAS instruments onboard the AML and with a long-path FTIR instrument located on a building rooftop. The uncertainty in both time series is assumed to be ±9%.

EGU
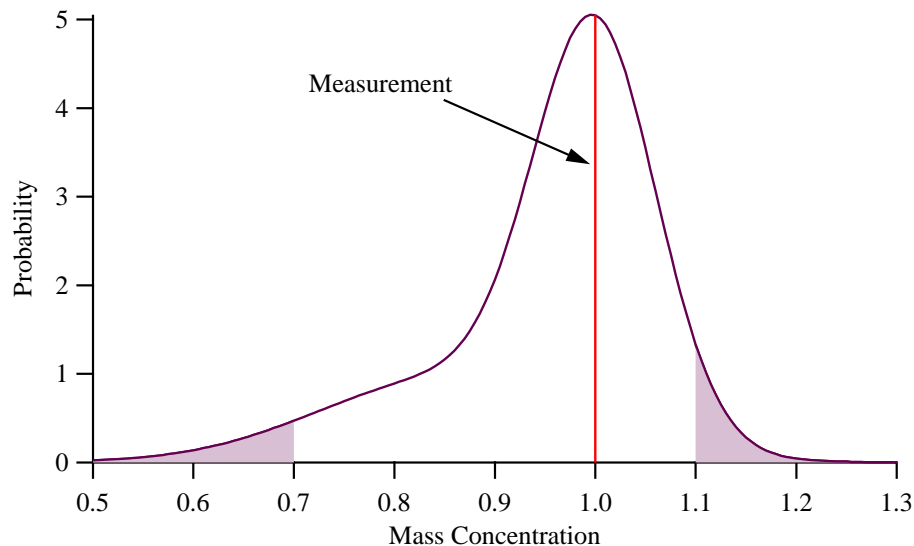
Fig. 3. Asymmetric likelihood function for a hypothetical AMS observation (see Eq. 25) above the detection limit. The observation is 1 (arbitrary units) and has an uncertainty of −30% and +10%. The area of the shaded region is 8%.
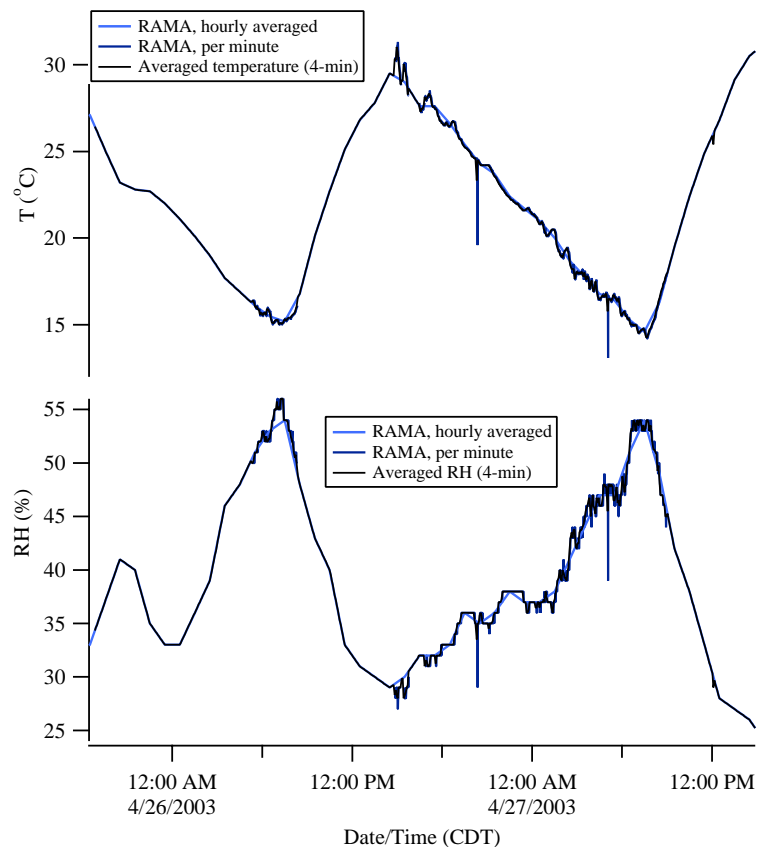
Fig. 4. Temperature and relative humidity profiles from RAMA for La Merced. Where available, we use the per minute data to calculate 4-min averages. For periods when the data logger failed, only hourly averaged data was available. For these periods we interpolated the hourly averaged data to determine the 4-min averages.

EGU
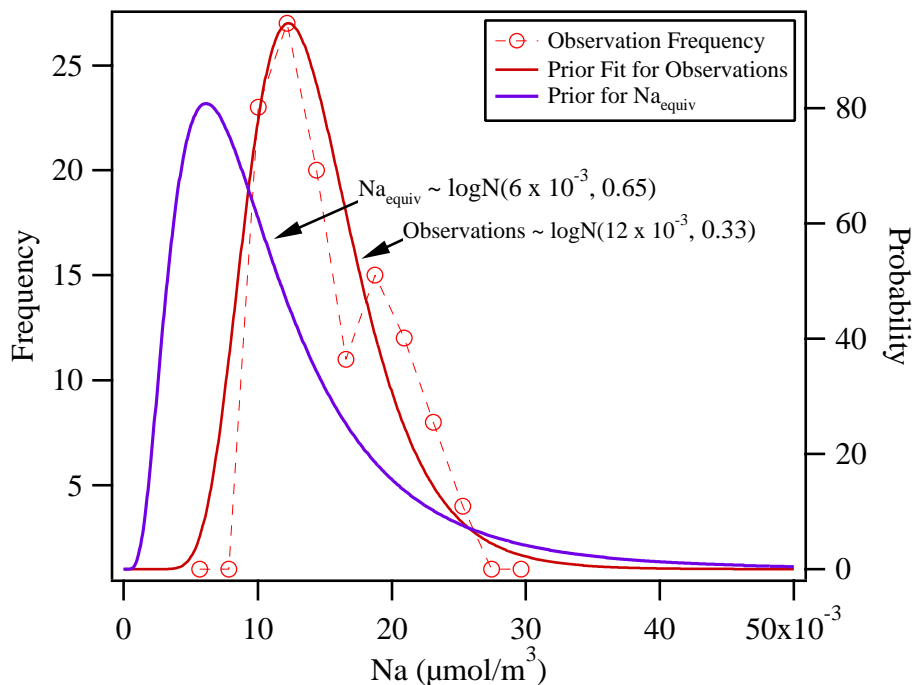
Fig. 5. Distribution of equivalent sodium concentrations (in μmol/m^3) based on PIXE observations (red circles) taken at the CENICA site during MCMA-2003. A lognormal fit (red line) of these observations was fit using the method of moments (Observations $\sim$logN(12×10^{-3}, 0.33)). The lognormal prior for Na$_{equiv}$ used here (heavy purple line) halved the mode of the fitted distribution and doubled the standard deviation to account for the fact that the PIXE observations provide an upper limit to Na$_{equiv}$ (Na$_{equiv}\sim$logN(6×10^{-3}, 0.65)).

Title Page

Abstract   Introduction

Conclusions   References

Tables   Figures

◄◄   ►►

◄   ►

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

EGU
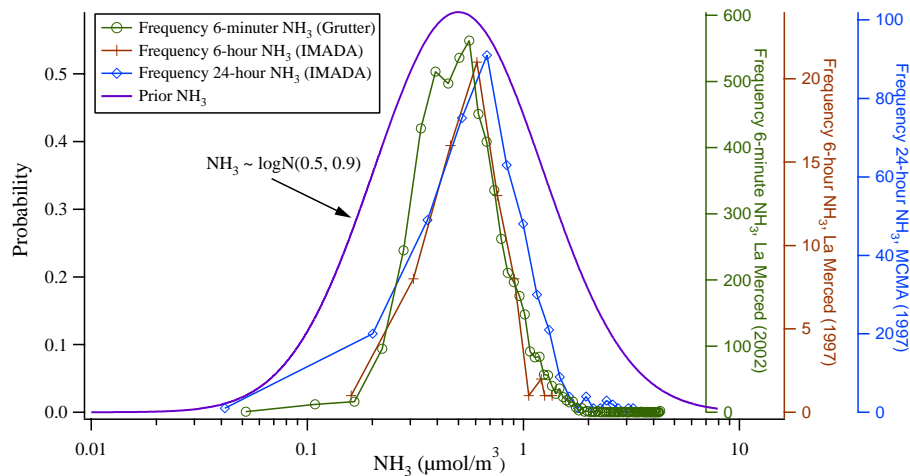
Fig. 6. Distribution of ammonia observations (in $\mu mol/m^3$) from the 1997 IMADA-AVER campaign and an exploratory campaign held at La Merced during February 2002. The lognormal prior is $NH_3 \sim logN(0.5, 0.9)$.

EGU
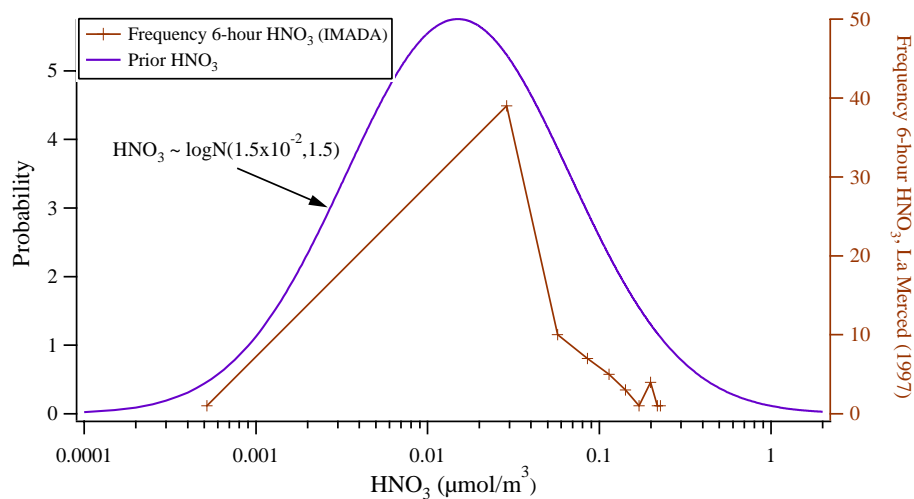
Fig. 7. Distribution of nitric acid observations (in μmol/m^3) from the 1997 IMADA-AVER campaign. The lognormal prior is $HNO_3 \sim logN(1.5 \times 10^{-2}, 1.5)$.
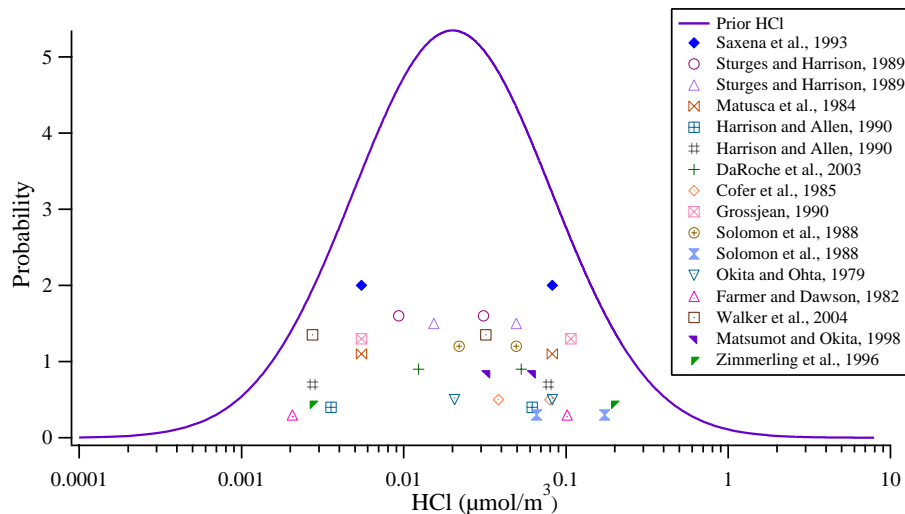
Fig. 8. Observed urban HCl concentrations (in μmol/m^3) in the literature and the proposed lognormal prior distribution, HCl~logN(0.02, 1.4). The units of the literature data for the ordinate are arbitrary. For each literature source the minimum and maximum reported concentration are shown. The Sturges and Harrison (1989) data are 7-day and 24-h samples; the Harrison and Allen (1990) data are 24-h and 3-h samples; the Solomon et al. (1988) data are annual and 24-h maximum samples. San Martini (2004) provides a review of the literature sources.
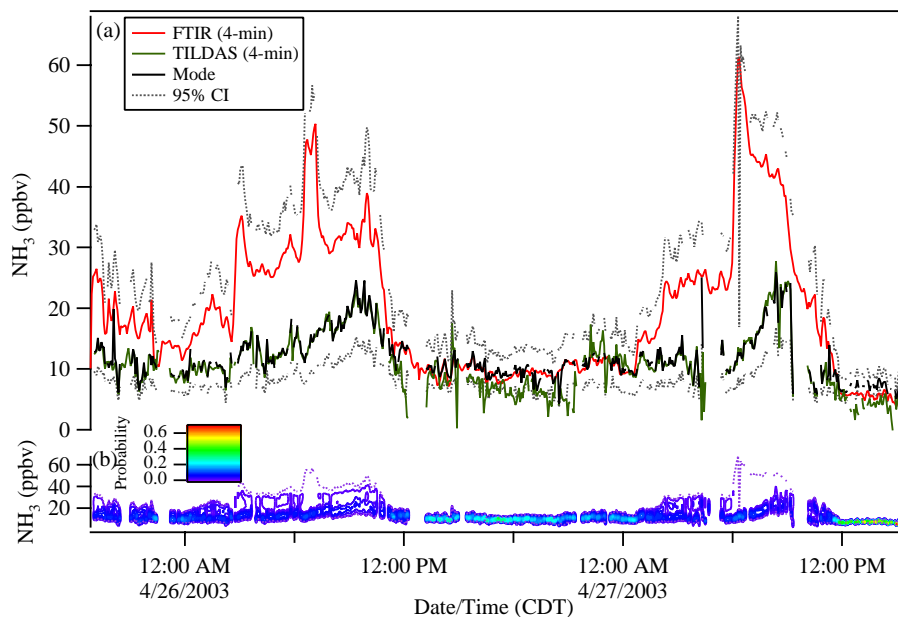
Fig. 9. (a) 4-min averaged ammonia observations at La Merced between 25 April 2003 and 27 April 2003 taken with the FTIR (red) and TILDAS (green) instruments, and the predicted mode (black) and 95% confidence intervals (dashed black) of the ammonia posterior distribution. **(b)** Predicted ammonia posterior probability density surface. The model was not run if both ammonia observations were not available; model runs where the Markov Chain did not converge are also not shown (see text).
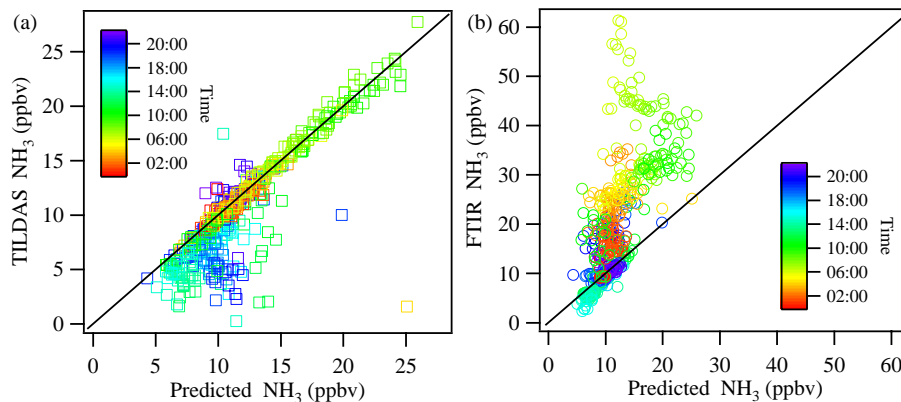
Fig. 10. Comparison between the predicted NH$_3$ concentration and that observed using the TILDAS **(a)** and FTIR **(b)**. Only the mode of the posterior NH$_3$ distribution is shown. The points are shaded are shaded by time of day.
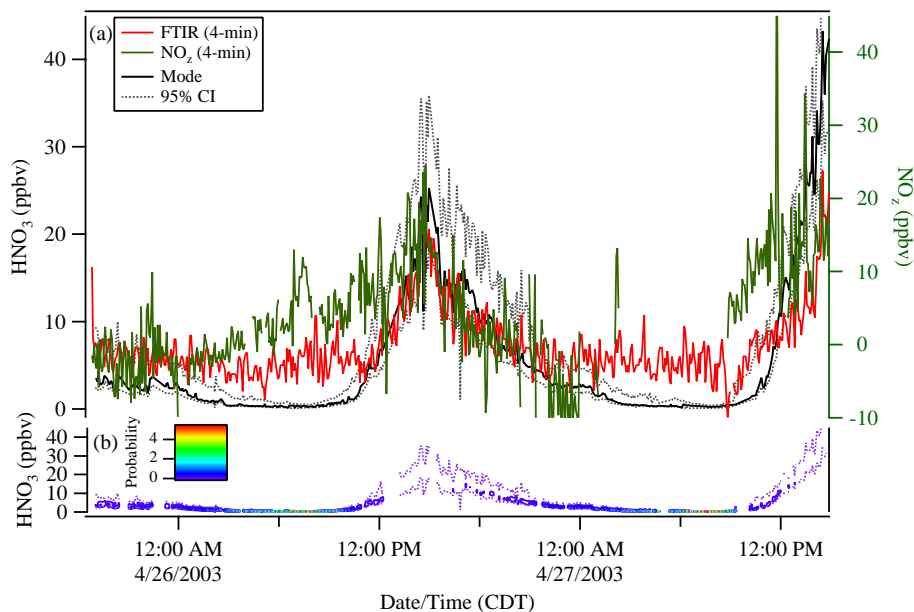
Fig. 11. (a) 4-min averaged nitric acid (red) and NO$_z$ (green) observations at La Merced between 25 April 2003 and 27 April 2003, and mode (black) and 95% confidence intervals (dashed black) of the nitric acid posterior distribution. **(b)** Predicted nitric acid posterior probability density surface.

EGU
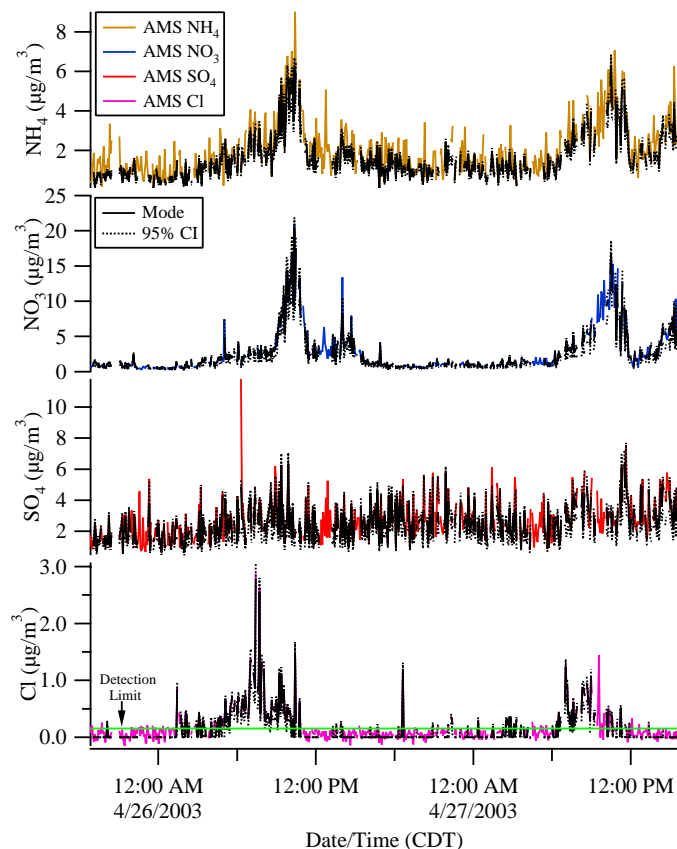
Fig. 12. Predicted (black) and observed (colored) concentrations of ammonium, nitrate, sulfate and chloride at La Merced. The black dashed lines are the predicted 95% confidence intervals; the measurement uncertainty is +10%, −30%. The model was not run if both ammonia observations were not available; model runs where the Markov Chain did not converge are not shown (see text).
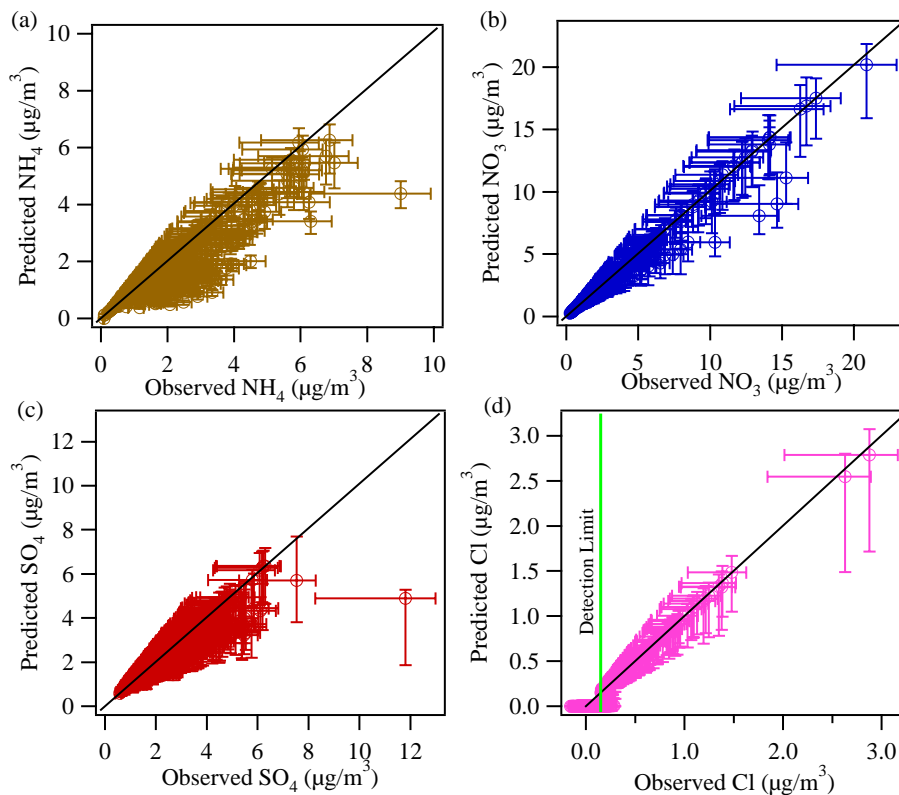
Fig. 13. Correlation plots for **(a)** ammonium, **(b)** nitrate, **(c)** sulphate, and **(d)** chloride for La Merced. The error bars for the predictions represent the 95% confidence interval; the measurement uncertainty is +10%, −30%. The detection limit for chloride is $0.15\,\mu g/m^3$.

EGU
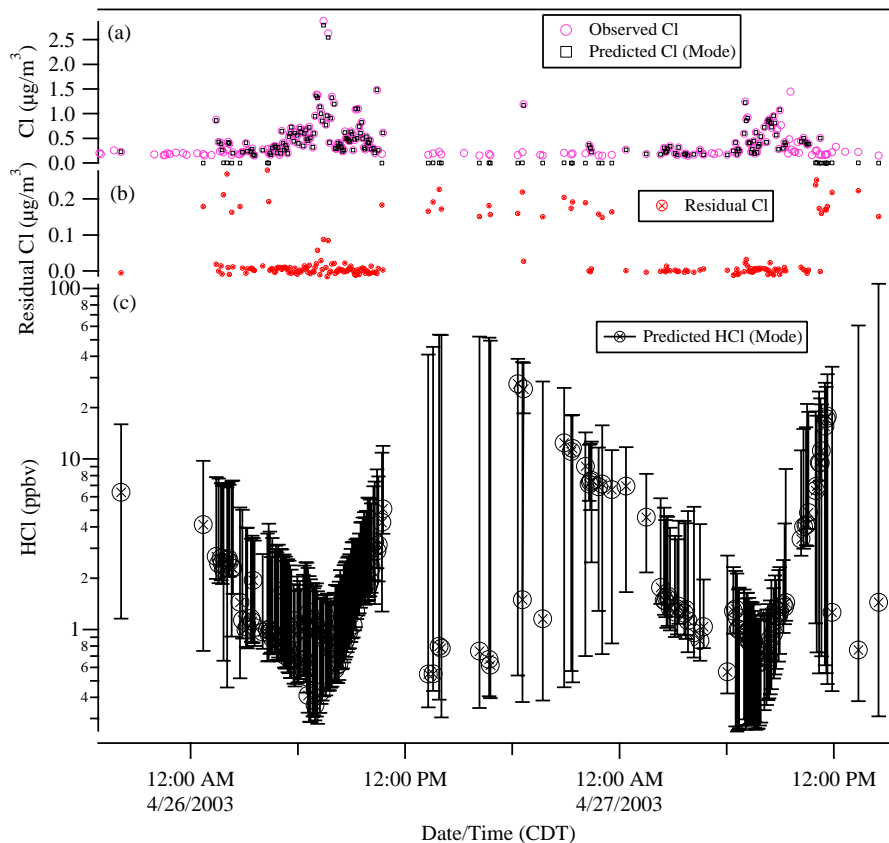
Fig. 14. **(a)** Predicted (mode) and observed aerosol chloride concentrations above the detection limit, **(b)** residual (observed – predicted) aerosol chloride concentrations. **(c)** Posterior distribution of HCl(g) concentrations, where the points represent the mode of the probability density function and the error bars are the 95% confidence intervals.
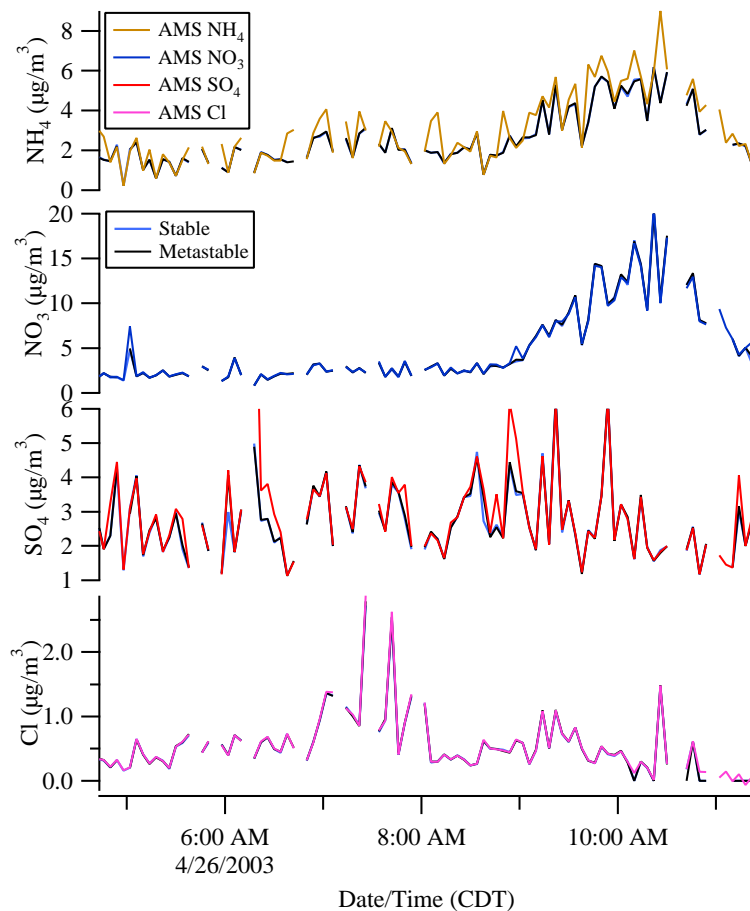
Fig. 15. Observed and predicted aerosol ammonium, nitrate, sulfate, and chloride concentrations. The predicted concentrations are assuming the aerosols are in the deliquescence (black) and efflorescence (light blue) branch. Only the mode of the aerosol posterior probability function is shown.

Fig. 16. Predicted ammonia concentrations for the stable (black) and metastable (light blue) case. The solid lines are the mode of the distribution and the dashed lines represent the 95% confidence interval.

Fig. 17. Predicted nitric acid concentrations for the stable (black) and metastable (light blue) case. The solid lines are the mode of the distribution and the dashed lines represent the 95% confidence interval.

Fig. 18. Predicted hydrochloric acid concentrations for the stable (black) and metastable (light blue) case. The solid lines are the mode of the distribution and the dashed lines represent the 95% confidence interval.
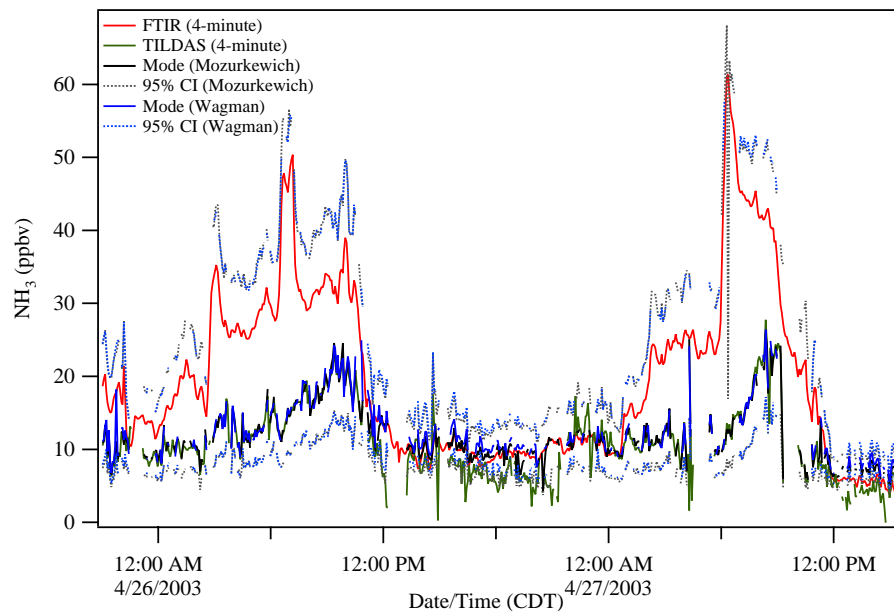
Fig. 19. 4-min averaged observations of ammonia and predictions using the equilibrium constant given by Mozurkewich (black) and based on the thermodynamic data given by Wagman (blue). The original formulation of ISORROPIA uses the value suggested by Wagman. The 95% confidence interval for both cases is shown (dashed line).
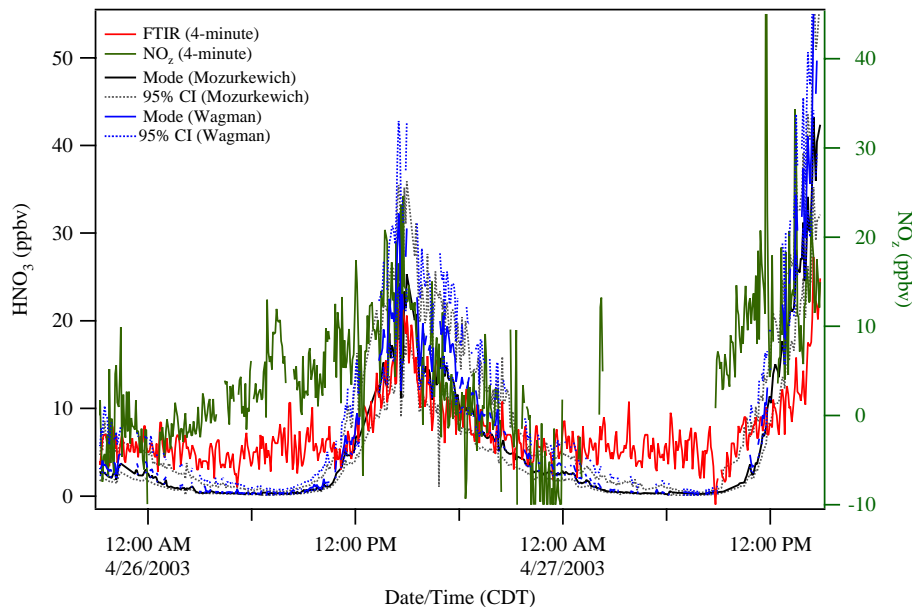
Fig. 20. 4-min averaged observations of nitric acid and NO_z, and predictions using the equilibrium constant given by Mozurkewich (black) and based on the thermodynamic data given by Wagman (blue). The original formulation of ISORROPIA uses the value suggested by Wagman. The 95% confidence interval for both cases is shown (dashed line).

Fig. 21. Comparison of the mode of the posterior distribution for **(a)** ammonia, **(b)** nitric acid, and **(c)** hydrochloric acid predicted using the equilibrium constant suggested by Mozurkewich (modified K_p) and Wagman (original K_p). The points are shaded by the time of day.

EGU