Atmos. Chem. Phys. Discuss., 6, S3755–S3769, 2006
www.atmos-chem-phys-discuss.net/6/S3755/2006/


Atmospheric Chemistry and Physics Discussions



# *Interactive comment on* "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" *by* F. M. San Martini et al.

**F. M. San Martini et al.**

Received and published: 13 October 2006

Comment: *Do you use the same data set, already employed for the determination of the likelihood functions, also for the 'model-observation' intercomparison (presented in Figs. 9-21)? Or did you use an independent control sample?*

The dataset shown in Figs. 9-21 was not used to determine the likelihood functions. Rather, the likelihood functions derived in Section 3.2 are based on reported measurement uncertainties and, in the case of the TILDAS and FTIR ammonia observation,



Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

the belief that there is no *a priori* reason to believe that one measurement technique is more likely than the other. (Note that the data shown in Figs. 9-21 was not used to derive the prior distributions – this would bias the results due to 'double counting'.)

- Comment: *Do you compare the modal values of the a posteriori PDF, i.e., $X_{post}^{mode}$ with $X_{obs}$? Are the latter time-averaged values? In the case of a "perfect prediction", the modal value of $p(\theta|$Data$)$ from the l.h.s of Eq. (2) should be identical with $X_{obs}$ entering $p(Data|\theta)$ at the r.h.s of Eq. (2). The aim of any solver is to ensure validity of the equal gn in Eq. (2) by searching for the solution $X_{post}^{mode}$. Is this correct?*

For the averaging of the observations, see the previous comment. Regarding the second part of the comment, the modal value of the likelihood should, for a 'perfect measurement', indeed be identical to $X_{obs}$. To help differentiate the MCMC method from other methods, however, consider that the objective function, for example, of a non-linear least-squares solver, which searches for the value of $X$that minimizes the difference between $X$ and $X_{obs}$. The MCMC method, however, does not only search for this solution, but searches for the solution that matches all the measurements given the system constraints, our knowledge of thermodynamics, and prior knowledge. The result is not a single value but a probability density function.

- Comment: *p. 5958, line 19-20: "This means that during these periods [...] the TILDAS observations are more consistent with the observations." Actually, we have at least three different values of any variable: the (unknown) true value, an observation, and a model estimation/prediction. Hence, the predicted data can only be compared with observations, but not with the "truth". With respect to $NH_3$ you have two measurement devices: open/long-path FTIR, closed-path TILDAS (=point measurements). The consistency with which observations do you mean? Apart from that, the $NH_3$ posterior PDF given in Fig. 9b is difficult to read.*



ACPD
6, S3755–S3769, 2006





EGU

The reviewer is correct – the prediction can only be compared with observations, but not with the 'truth' (which is unknown). The observations referred to in line 20 are all the observations, i.e., the temperature, relative humidity, AMS, nitric acid and both ammonia observations, as well as the prior probability functions. This has been clarified by substituting the sentence with:

This means that during these periods, given our understanding of aerosol thermodynamics, the TILDAS observations are more consistent with the temperature, relative humidity, AMS, and gas-phase observations.

Figure 9 has been enlarged to make reading Fig. 9b easier.

- Comment: *Section 4.1: The message is, that the prediction of gas phase $HNO_3$ and HCl is sensitive against deliquescence/efflorescence, but the prediction of particle phase concentrations and $NH_3$ is not. Is this correct?*

This is correct. This is because the uncertainty in the ammonia observation is smaller than for nitric and hydrochloric acid.

- Comment: *On p. 5961, line 14: The wet aerosol is predicted to be acidic even at high ammonia concentrations. Can you physicochemically explain this effect or it is a side effect of the uncertainty of the activity model, when the concentration of the solution becomes large (at comparatively low relative humidities considered here)?*

I know of no physiochemical explanation for this prediction. The explanation that the acidity may be an artifact of the activity coefficient model is reasonable to me. The partial dissociation of the bisulfate ion is notoriously difficult to predict, especially at ionic strengths on the order of $\sim$40 mol/kg (and above).



ACPD
6, S3755–S3769, 2006





EGU

- Comment: *On p. 5962, line 4–6: Please insert, which figure you are referring to (Figs. 11, 20). According to Subsection 3.1, Mozurkewich's revised constant has been used in Fig. 11. Even if it overpredicts the afternoon $HNO_3$ concentration on April 27, the revised equilibrium constant enhances the $HNO_3$ prediction compared to the original parameterisation.*

I have added a reference to Figure 20 (which compares the model predictions based on Mozurkewich's and the equilibrium constant used in the original formulation of ISORROPIA).

The reviewer is correct that Mozurkewich's revised constant was used in Fig. 11, and that the revised equilibrium appears to enhance the $HNO_3$ prediction compared to the original parameterization. As was noted in the text, we wish to emphasize, however, that carefully controlled laboratory rather than field conditions are a more appropriate means to determine and validate thermodynamic parameters.

- Comment: *Assessments like "extremely well" (5962/19), "excellent" (5962/19), "excellent job" (5962/21, 5963/24) insinuate, that the method has already reached its final state. What comes beyond "excellent"? Hopefully, there is something left to improve with respect to the method etc. I think, even a "good agreement" is always a very good result. Anyway, the conclusions are found to be sound and conclusive.*

I thank the reviewer for his thoughtful comments. While embedding ISORROPIA in a MCMC framework allows unobserved gas-phase concentrations to be predicted, and reproduces the particle-phase observations very well, enhancements to the method are, indeed, needed. Future work that permits the inclusion of secondary organic aerosol species to the aerosol model will be invaluable (see companion paper). In addition, the rich dataset provided by the MCMA-2003 field campaign will allow future



ACPD
6, S3755–S3769, 2006

work to consider the dependence of the prior distributions on time of day, meteorology, and other factors (see response to other reviewer's comments).

- Comment: *A list of all abbreviations used in the text would be useful.*

The following list of abbreviations was added to the manuscript:

AML = Aerodyne Mobile Lab

AMS = Aerosol mass spectrometer

CENICA = National Center for Environmental Research and Training (Centro Nacional de Investigación y Capacitatión Ambiental)

FTIR = Fourier Transform Infrared

hwhm = half width at half maximum

IMADA-AVER = Investigación sobre Materia Particulada y Deterioro Atmosférico–Aerosol and Visibility Evaluation Research

MCMA = Mexico City Metropolitan Area

MCMC = Markov Chain Monte Carlo

pdf = probability density function

RAMA = Red Automática de Monitoreo Atmosférico

TILDAS = tunable infrared laser differential absorption spectroscopy

- Comment: *p. 5935, line 3: "...to develop (?) a powerful tool"*

The text was modified accordingly.



ACPD

6, S3755–S3769, 2006

- Comment: *p. 5935, line 5: "...provides a basis (?) for a formal framework"*

The text was left unchanged.

- Comment: *p. 5935, line 7: "...to particle- and gas-phase observations of ammonia" (use here and elsewhere "particle phase" instead of aerosol phase; have in mind the definition of aerosol: an aerosol is at least a two-phase system; dispersion of gas and particles) (see also p. 5944, line 22, 24, 25; p. 5953, line 18)*

The text was modified accordingly

- Comment: *p. 5935, line 15: "...varying between 0.4 and 5 ppbv"*

The text was left unchanged.

- Comment: *p. 5937, line 3: explain the abbreviation of CENICA, when it appears the first time (see part II, p. 6003)*

The sentence was modified:

A companion paper will discuss the application of the Bayesian method to three other fixed sites in the MCMA-2003 campaign, the National Center for Environmental Research and Training (Centro Nacional de Investigación y Capacitatión Ambiental, abbreviated as CENICA), Pedregal, and Santa Ana.

- Comment: *p. 5937, line 17: "A full description of the experiment and location is presented in Grutter et al. (2003)." (one can avoid "elsewhere", when it is already known, where "else" is; see also p. 5939, line 9; p. 5954, line 22.)*



ACPD
6, S3755–S3769, 2006





EGU

The text was modified accordingly

- Comment: *p. 5938, line 5: "The AML contains a suite of fast-response instruments ..."*

The text was modified accordingly.

- Comment: *p. 5938, line 15: What does the abbreviation "hwhm" mean?*

hwhm = half width at half maximum; see List of Abbreviations added to the text.

- Comment: *p. 5938, line 25: Please check parenthesises in quotations (many times)! Example: "The operation of the NO2 TILDAS is described in Li et al. (2004)." (see also: p. 5942, line 25; p. 5957, line 19 etc.)*

The text was modified accordingly.

- Comment: *p. 5939, line 7: NO3-*

The text was modified accordingly.

- Comment: *p. 5939, line 22: "Thus, although CENICA was considered to be (OR as) (?) the*

*supersite ..."; Constructions, in which I would intuitively add "to be" or "as" appear several times.*

The text was left unchanged.



ACPD
6, S3755–S3769, 2006





EGU

- Comment: *p. 5939, line 24: ...where both co-located NH3 and HNO3 observations were available*

The text was left unchanged.

- Comment: *p. 5943, line 26: "For the four models they examined, Ansari and Pandis (1999a) found minor differences in predicted chloride concentrations." This way, one can avoid quasi-double citation in one quotation (several times).*

The text was modified accordingly.

- Comment: *p. 5944, line 5: The sentence "Using the same dataset, San Martini..." is an example for a long sentence, which could be splitted.*

The sentence was modified to:

Using the same dataset, San Martini (2004) compared predictions from ISORROPIA and a new equilibrium model that directly minimizes the Gibbs free energy and includes complex and hydrate species. Only small differences in model predictions were found (San Martini, 2004).

- Comment: *p. 5950, Eq.(25) (several times): leave out brackets for the prefactor 0.3*

The text was modified accordingly.

- Comment: *p. 5952, line 14-15: Laplace's Principle of Insufficient Reason – I would like to see this statement explicitly commented in a footnote or appropriately referenced.*



ACPD
6, S3755–S3769, 2006





EGU

Laplace's Principle of Insufficient Reason says that in the absence of evidence to the contrary, all possible outcomes are equally likely. A reference to (Ferson and Ginzburg, 1996) was added to the text. (Laplace's Principle of Insufficient Reason can be considered the probabilistic analogy to Occam's Razor.)

- Comment: *p. 5952, line 24: Please resolve the abbreviation pdf for "probability density function" in the text.*

This was resolved on p. 5950, line 4. See also List of Abbreviations.

- Comment: *p. 5954, line 13: "Previous observations have found..."*

The text was modified accordingly.

- Comment: *p. 5955, line 12-14: Please examine this sentence. Make clear, to what part the half sentence "found using the method of moments" is referring to. Please add the page numbers in the citation of the textbook of Seinfeld and Pandis (1998).*

This sentence was modified to:

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., see p. 1270 in (Seinfeld and Pandis, 1998)).

- Comment: *p. 5957, line 3: Structurise the sentence by a comma.*



ACPD
6, S3755–S3769, 2006





EGU

This sentence was modified to:

Most emitted organic chloride is expected to be converted to HCl given the high level of photochemical activity generally present in the MCMA.

- Comment: *p. 5959, line 26: $\sim$ O(? ppbv)*

This sentence was modified to:

The HCl (g) concentrations are well constrained in this period, with concentrations generally on the order of ppbv, though higher concentrations are predicted on the $27^{th}$ (on the order of 10 ppbv).

- Comment: *p. 5962, line 7–11: Please completely reformulate this sentence. I suspect the message is, it does not matter, which equilibrium constant is used.*

This sentence was modified to:

Regardless of which value of $K_p(NH_4NO_3)$ is used, the $NH_3$ TILDAS point observations are more consistent with all the available measurements and our knowledge of thermodynamics.

- Comment: *p. 5963, line 9: "did not reveal (?) a comparable difference"*

This sentence was modified accordingly.

- Comment: *p. 5963, line 10: "versus those (?) in the AML"*

This sentence was modified accordingly.



ACPD
6, S3755–S3769, 2006





EGU

- Comment: *p. 5963, line 29: "Finally, ..." (check comma placement, several times)*

This sentence was modified accordingly.

- Comment: *p. 5964, line 17: Please define, what an "overdispersed version of the posterior distribution" is.*

This sentence was modified to:

Choose a probing distribution that approximates an overdispersed (i.e., with a larger variance) version of the posterior distribution that is being sampled from (Gelman and Rubin, 1992)

- Comment: *p. 5964, line 19: Please reformulate the second half sentence of the second item "[...] is to remain at $\theta_t$[...]"*

This sentence was modified to:

Choose a probing distribution whose expected value for each proposed move is to stay put, i.e., $E(\theta^*|\theta_t)$=$\theta_t$, where $\theta^*$ and $\theta_t$, are the proposed and current states.

- Comment: *p. 5964, line 21: "... The second suggestion ensures that there will be an approximate left-right balance, which encourages rapid exploration of the entire solution space ..." In cases, this sentence is important to understand the approach, please make clear, what is meant [I have at least a suggestion]. In cases it is not, please try to generalise the message and add a reference, which is freely accessible.*

Hastings (1970) generalized the algorithm of Metropolis et al. (1953) for the case where a probing distribution that is not symmetric is used. For a review of different probing



ACPD
6, S3755–S3769, 2006

distribution strategies see, for example, p. 330 in (Chib and Greenberg, 1995). For a further discussion of efficient jumping rules for symmetric probing distributions see Gelman et al. (1996).

In order to clarify these concepts, lines 21-24 were modified to:

A symmetric probing distribution, as originally suggested by Metropolis et al. (1953), fulfills the second characteristic. A symmetric probing distribution (i.e., $PD(\theta^*|\theta) = PD(\theta|\theta^*)$) facilitates the exploration of the entire solution space by assigning equal probability to left and right moves from the current position.

- Comment: *p. 5943, line 27: Add year to Zhang et al.*

The year was added (2000).

- Comment: *p. 5957, line 20: Add year to Moya et al.*

The year was added (2003).

- Comment: *...see also p. 5958, line 4/5 etc.*

The year was added (2004).

- Comment: *Allen et al. (2002): Add page numbers or doi*

The page numbers were added (1591-1599)

- Comment: *Beier (1999): Check syntax of the German title*



ACPD
6, S3755–S3769, 2006





EGU

The reference was checked. However, I found no error.

- Comment: *Li et al. (2004): The JGR appendix "Atmos." can be omitted.*

Atmos. was omitted.

- Comment: *Zhang et al. (2003): Check reference. Please add doi-number.*

The reference was checked and doi number added (10.1029/2001JD001592)

- Comment: *Figs. 3, 7, 8: Check the range of the probability (>1).*

The fact that a probability density function may have values greater than unity is a common source of confusion. The normalization of a probability density function requires that the integral of the pdf is unity, i.e,

$$\int p(x)dx \; = \; 1 \tag{30}$$

where $p(x)$ is the pdf of $x$ and the integral is over all of $x$. This does not mean that $p(x)$ must be smaller than unity.

Consider, for example, a random variable on the interval [0, 1] with a uniform probability density. Its probability density will be unity between 0 and 1, and 0 everywhere else. Contrast this with a random variable on the interval [0, 0.01] described by a uniform probability density. In this case, the probability density will be 100 between 0 and 0.01 and 0 everywhere else.

- Comment: *Fig. 4: Graph colours are difficult to separate.*



ACPD
6, S3755–S3769, 2006





EGU

The figure was modified to make it easier to distinguish the colors.

- Comment: *Figs. 9b, 11b: The probability density surface is difficult to read. The corresponding figures should be rescaled (e.g., there is only 1 cm ordinate length scale for 0-60 ppbv NH3 range), Fig 11b: Check the range of the probability (>1)*

Both Figures have been rescaled so that the probability density surface is easier to read. For the probability range see response above.

- Comment: *Please add the units for the ordinate axis.*

Frequency here refers to the frequency (i.e., number of) observations.

- Comment: *Correct ordinate: "Mode HNO3 (ppbv), Modified Kp(NH4NO3)"*

This has been corrected.

**References**

Barker, A.A., 1965. Monte Carlo Calculations of the Radial Distribution Functions for a Proton-Electron Plasma. Australian Journal of Physics, 18: 119-133.

Chib, S. and Greenberg, E., 1995. Understanding the Metropolis-Hastings Algorithm. The American Statistician, 494(4): 327-335.

Draper, D., Forthcoming. Bayesian Hierarchical Modeling. New York: Springer-Verlag.

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

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



ACPD

6, S3755–S3769, 2006

Gilks, W., Richardson, S. and Spiegelhalter, D., 1996. Introducing Markov Chain Monte Carlo. In: W. Gilks, S. Richardson and D. Spiegelhalter (Editors), Markov Chain Monte Carlo in Practice. Chapman & Hall, London, UK, pp. 1-21.

Gilks, W.R., Roberts, G.O. and Sahu, S.K., 1998. Adaptive Markov Chain Monte Carlo Through Regeneration. Journal of the American Statistical Association, 93(443): 1045-1054.

Hastings, W.K., 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. Biometrika, 57(1): 97-109.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. and Teller, E., 1953. Equations of state calculations by fast computing machines. Journal of Chemical Physics, 21: 1087-1092.

Raferty, A.E. and Lewis, S.M., 1992. How Many Iterations in the Gibbs Sampler? In: J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith (Editors), Bayesian Statistics 4. Oxford University Press, Oxford, U.K., pp. 763-774.

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

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

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

---

Interactive comment on Atmos. Chem. Phys. Discuss., 6, 5933, 2006.



ACPD

6, S3755–S3769, 2006

Interactive Comment



EGU