

Interactive comment on “Aerosol model selection and uncertainty modelling by adaptive MCMC technique” by M. Laine and J. Tamminen

Anonymous Referee #1

Received and published: 17 July 2008

This paper introduces a Bayesian inference method that incorporates model uncertainty and applies the method to the GOMOS aerosol model selection problem. The method will be of interest to readers of ACP and useful in a variety of applications, and provides sufficient background and references that interested readers should be able to implement the method. I recommend the paper be published with (relatively) minor revisions.

Major comments

I concur with the two major comments made by the second reviewer. Specifically, Figure 5 shows that different models can be used at different heights. By comparing the predictions here with those from, for example, existing operational GOMOS retrieval,

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



the authors can demonstrate the value of the approach. The authors should also discuss why different cross-section models may be applicable at different altitudes (and locations). For example, could variations in aerosol composition and shape with altitude be used to explain the use of different models at different heights? This will also allow the authors to show that this is not a case of introducing more degrees of freedom (in this case, four models rather than one) to enhance the model fit with observations.

Although I have provided some editorial suggestions below, the paper would benefit from further editing to enhance the clarity of the paper.

Minor comments

I concur with second reviewer's suggestion that there is a lack of clarity in the abstract. For example: "We apply Bayesian model selection techniques on the statistical inversion problem of the GOMOS instrument." This may be a minor point, but given that it is the first sentence of your abstract, it would be helpful to be more precise; Bayesian techniques are not applied to the GOMOS instrument. Rather, they are applied to data from the GOMOS instrument to infer model parameters and help select an appropriate model. Similarly, instead of "[t]he algorithm is easy to implement and can readily be employed for model selection as well as for model averaging, to properly take into account the uncertainty of the modeling", consider "[t]he algorithm is easy to implement and can readily be used for model selection and averaging, and to directly incorporate model uncertainty."

In this work no convergence criteria were used beyond visual inspection of the chains. In my experience, inference problems with dimensionality like the problem considered in this work and with a relatively high degree of correlation (see Figure 4) require actually testing for convergence, i.e., visual inspection is (in my experience) insufficient. I urge the authors to use any of the standard convergence tests that are readily available and can be freely downloaded. In addition to the issue of convergence, the manuscript would benefit from a brief discussion of the sensitivity of the results to the initial value

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

and any burn-in and thinning that was used in generating the MCMC chains.

Technical corrections

2/10792: “The motif is to study which type...” Motif? I suggest “The goal is to study. . .”

9/10792: I suggest changing “employed” to “used”

11/10792: Advances in computer resources and algorithms have made the use of increasingly complicated models possible.

22/10792: The approach allows

23/10792: Consider “Prior information from different sources can be directly incorporated and the correlation structure of the unknowns can be fully explored.”

11/10793: Omit “just”

12/10793: Several alternate formulations. . .

13/10793: . . . depending on the types of aerosol at a given location.

13/10793: Data is agnostic. “Consequently, it is advisable to allow for different types of models and to use the data to decide which model to use.”

1-10/10794: Reviewer 3 suggested that these introductory sentences seem overly simple. I concur.

13/10794: “Then it is reasonable to model also the uncertainty in the model, for example by introducing several alternative models and let the data decide which of them to use.” Awkward! How about: In such cases one can include model uncertainty, for example by introducing several alternate models and using the data to decide which model to use.

14/10794: “A problem, similar to a point estimation (or maximum a posteriori estimation) in parameter estimation, would then be the selection of the best model.” Awkward! Maybe replace this sentence simply with “This is the problem of model determination.”

S4969

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



26/10794: ... the constituent line densities. . .

7/10795: "giving the distribution of the observations"? I am not sure what this phrase means. Given the distribution of observations? If it is intended that the likelihood function gives the distribution of the observations, then I do not think this is correct. The likelihood function describes the likelihood of the observations given the unknowns (which may include model(s)).

9/10795: The joint posterior distribution of the unknowns ... conditional to the observed data y is given by Bayes' theorem and can be written as the product of the likelihood and the priors:

11/10796: where the first term on the right. . .

2-10/10797: This paragraph could be shortened considerably or perhaps simply replaced with "MCMC methods overcome the problems posed by high-dimensional integrals by using high-dimensional random walks."

15/10798: Inferences about the unknowns are made using statistics calculated from the chain of values.

19/10799: "... and they thus provide a common scale to perform the between model transformations of the parameters." ? Do you mean: and they thus provide a common scale to perform the transformations of the model parameters?

22/10799: Replace "stand for" with "be"

16/10801: Its success depends on. . .

Throughout the manuscript, sometimes MH is spelled out and sometimes the abbreviation is used. Please fix.

1/10803: ... the so-called. . .

4/10803: First,

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper



5/10803: Second,

6/10803: non-Gaussian... non-linear...

9/10803: I suggest deleting this sentence. This had been already stated numerous times.

10/10803: Delete "We want to note that"

13/10803: Replace employed with used

14/10803: high-dimensional

2/10805: ... after the line densities for all the heights that have...

5/10805: The symbol alpha was used previously for the MH acceptance probability. Use a different symbol.

17/10806: First,

23/10806: Second,

14/10807: Model averaging is useful when the best model cannot be determined

29/10807: This is mainly due to the similarity...

2/10808: Is unidentifiability a word?

1/10809: Delete "as well"

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 10791, 2008.

ACPD

8, S4967–S4971, 2008

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

