

## ***Interactive comment on “Non-parametric Bayesian mixture of sparse regressions with application towards feature selection for statistical downscaling” by D. Das et al.***

### **Anonymous Referee #2**

Received and published: 3 June 2014

#### General comments:

This paper is about a Bayesian sparse regression technique and an application to statistical downscaling. The work seems to be interesting, and potentially novel.

While it is quite valuable to have interdisciplinary submissions in this field, a significant caveat is that the main contribution needs to be identified in the submission. The downscaling application does not seem to be emphasized in the submission, e.g. it appears to be just an evaluation of their method. (cf. "We have evaluated our method both on synthetic and climate datasets.")

If instead the contribution is on the technical (methodological/algorithmic/modeling)

C210

side, it needs to be more clearly stated in the exposition. In the Conclusion, the authors state "Our major contribution is to develop an efficient and scalable variational inference algorithm for inference in the fully Bayesian model." In order to facilitate evaluation of this claim, the authors would need to delineate and distinguish their contributions from past work, especially in the technical sections. Currently, sections 2 and 3 do not clearly distinguish between prior work in the technical area, including by the authors, and the claimed technical contributions in the submission. This makes it difficult to evaluate the novelty of the contribution.

Finally, if the primary contribution is indeed on the methodological/algorithmic/modeling side, then it would seem much more appropriate to submit the manuscript to a data mining, statistics, AI, or machine learning publication, where it could be refereed by reviewers with the relevant expertise. This reviewer is not aware whether or not similar/parallel submissions have been made by the authors to such venues, but this should be clearly stated by the authors.

#### Specific comments:

It would be helpful to expand the related work discussion to a variety of Bayesian techniques for downscaling by Andrew Robertson (IRI, Columbia LDEO) and collaborators.

Clarifications are needed in the technical section, e.g.:

- "stop when the probabilities  $E[z]$  stop changing any more"
- $E[z]$  is an expectation, not a probability.
- How do you quantify that they "stop changing"? Is there a threshold? How do you set the threshold?
- How is 5k chosen for the non-zero components?
- It would be good to show experiments for various K values.
- How does your method compare to other clustering techniques, including non-

C211

generative techniques such as k-means++?

- Many figures are too small the read the axes.

Technical corrections:

The submission needs extensive copy editing to fix numerous grammatical errors. E.g. there are many instances in which verbs do not agree with nouns (e.g. plurality). Due to these errors (along with typos, missing/repeated words), the submission is very tedious to read in its current form, in particular the introductory and non-technical sections.

A few corrections are listed below; however the above should be done in a concerted way, e.g. by a copy-editor.

GCM is defined twice with two definitions. "Global Climate Model" → "General Circulation Model."

"facing the mankind" → delete "the"

"Variational Bayes inferences" → "Variational Bayesian inference"

"for each partition  $X_k$ " → "for each part  $X_k$ "

---

Interactive comment on Nonlin. Processes Geophys. Discuss., 1, 615, 2014.