Nonlin. Processes Geophys. Discuss., 1, C140–C142, 2014
www.nonlin-processes-geophys-discuss.net/1/C140/2014/

Nonlinear Processes in Geophysics
Discussions
Open Access

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 #1

Received and published: 16 May 2014

General Comments:

The manuscript reports a Bayesian mixture of sparse regression models for clustering using the Dirichlet process mixture model for feature selection. The purpose is to be able to find interpretable clusters that can allow for climate model downscaling based on key features. This is a potentially useful approach to the discovery of unique features in highly complex datasets for use in statistical downscaling.

The authors thoroughly describe the model setup and theory, but they spend less time on the intended purposes of using this approach to better understand and discover real-world covariates to be used in statistical downscaling. The purpose of sparse re-

gression for enhancing the interpretability is clearly stated by the authors. However, the results from the experiments suggest that many challenges with simultaneous clustering and discovery exist, including temporal averaging impact (annual vs. monthly) and regional spatial and temporal impacts on final feature selections and clusters. These challenges should be discussed in greater detail. After reading through the model theory and description, I had high hopes that this mixture of sparse regression models would lead to great insights into the variability of precipitation along the Western U.S. However, this was not quite the case, as the author's interpretation of the results were general and did not show features that could be clearly defined and understood.

Greater insight needs to be expressed in the manuscript by detailing how these cluster results could be applied in downscaling, which would assist the reader in understanding the broader point of the paper. It is not clear how downscaling will benefit from this approach. The manuscript is of publication quality, but needs a major revision starting with answering the specific questions below and providing additional descriptions and interpretation of the results as they apply to regional characteristics and the potential use in downscaling.

Specific Comments:

Section 4.1.2, Page 617, Line 14-20 This argument needs to be justified more clearly. See Bader et al., 2008

Section 4.1.2 With the degrading NMI value for such a small value of K, given ideal case (i.e. simulated synthetic data), what does this say about how this model perform with weather and climate data, which is likely to have varying values of K depending on season, location, year, co-dependence of climate indices?

Section 4.1.2 By averaging precipitation to an annual value, you have reduced your ability to interpret results looking at both atmospheric and climate features. Why was annual averaging chosen? The temporal averaging (i.e. annual averages) may also limit the new discoveries and insight that can be drawn from the results by smoothing

out any variability that is likely to be teleconnected to atmospheric and climate related phenomena.

Section 4.2, Page 638, Line 6 This is unclear.

Section 4.2, Page 638, Line 10 Annual/Seasonal average? I thought it was only the annual average for each variable.

Section 4.2.1, Page 637, Line 15 A description of the relaxation procedure is necessary. Quantification of this sensitivity may also be informative of how sensitive your model is to spatial and temporal dependence. This would be an interesting aspect of the evaluation of your model.

Section 4.2.1, Page 637, Line 19-22 So does this mean that the model did not do its job of finding interpretable results for feature selection and downscaling?

Section 4.2.1, Page 638, Line 1-10 It is not clear how the authors came to this conclusion. Neither the figures nor text clearly state the basis of their interpretation. It is unclear how the model expressed “dependence” on particular features. Is this based on inference from the selected sparse model? This needs to be made clearer.

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