Clim. Past Discuss., 6, C541–C544, 2010
www.clim-past-discuss.net/6/C541/2010/

Climate of the Past Discussions

Interactive comment on "Objective identification of climate states from Greenland ice cores for the last glacial period" *by* D. J. Peavoy and C. Franzke

M. Crucifix (Referee)

michel.crucifix@uclouvain.be

Received and published: 29 July 2010

The authors nicely introduce the subject: there is a need for a formal framework to identify and characterise events in Greenland Ice Core records. A discrete hidden state model embedded into a Bayesian inference framework seems a natural and sound option.

The present paper contributes to a current and welcome trend of introducing more formal and better framed statistics in (palaeo-)climatology. With this trend comes the difficulty of traning scientists with mathematical background in the complexity of the climate system, and conversely communicating abstract and complicated mathematical concepts and methods to the bulk of climate scientists. The present paper features a multi-stage Gibbs sampler, which as a general rule requires a fairly mature understand-

ing of Bayesian statistics to be comprehended. I am unsure to be qualified enough to guarantee the correctness of the statistical implementation but I will do my best to formulate constructive comments.

1 the statistical model

1.1 On the use of 'objective', in particular in the title

. As pointed out by another reviewer the present article formulates statistical models of Greenland delta-18O dynamics. Any model implies a number of choices about conditional structure and parameter prior distribution. These are necessarily subjective, although once formulated, the inferrence process follows the objective rules of Bayesian statistics. I therefore share the concern about whether the wording 'objective identification' is appropriate.

1.2 About the structure and content of section 2

. Section 2 is awkwardly structured, with a long introduction and a shorter 2.1, with the warning : 'Readers (not ?) familiar with Bayesian inference can skip to Sect. 3'. A 'not' must be missing because section 2.1 requires some Bayesian statistics training to be understood. I believe that this section should be expanded:

1. The α hyperparameter of the Dirichlet distribution controlling the prior on λ is usually a vector of dimension M. All its components are presumably here equal to some value, not explicitly specified. λ is then presumably sampled within the Gibb's sampling procedure, according to the conjugate distribution $\mathrm{Dir}(\alpha + \beta)$, where β is — if I understand correctly — the histogram of the S_i. This supposes that α and β are sufficient for λ. All of this should be confirmed and discussed.

2. A table with values given to a, b, σ_μ and α with justification is needed.
3. Notations have to be better defined : $P(D|M)$ (?); d is presumably the total number of parameters; θ should be more explicitly defined (I understand this is the concatenation of λ_j, τ_j, μ_j.
4. Incidentally it could be specified that the Laplace approximation is here first order.

The lesson is that section 2.1 as it stands fails to comfortably reassure a climate scientist reasonably introduced to Bayesian statistics, and therefore needs to be partly rewritten and expanded.

1.3 About the Bayesian model itself

The likelihood function p. 1214 (loosely called 'probability of the data', which at least should be termed 'probability distribution function'), is as often the critical bit of the statistical edifice. It seems to suppose here that X_i are true indices for climate, with no error on observations. The authors should be urged to comment on this point. Otherwise, the rationale for the choice of priors arguably needs to be further defended, although they overall seem reasonable.

2 About the practical implementation

There is some oblique clue in the text that the X_i are sampled every 50-years, and that GISP2 underwent a running average. All of this should be much more clearly explained, with a discussion of the consequences of possible data pre-processing. For example, why does '50' appear in the interpretation of λ_j p. 1217?

2.1 Result presentation

Table 1 should also feature experiments with 4 states, in order to confirm that 3 states is optimum. Fig. 5 are too small and little legible.

3 Result discussion

1. Interpretation for μ p. 1218 is confusing. My understanding is that the 'increment model' is selected, in which case μ has unit 'delta 18-O per mil per time slice' (what is the time slice?) and therefore cannot be interpreted as a relaxation state nor compared with the threshold used by Rahmstorf (2003).
2. I concur with previous reviewer's comments that the authors are taking too much risk in the interpretation of their results in terms of climate mechanisms. They natural have the right to issue speculative statements about a possible interpretation of their results, but the tone adopted here is uncomfortably affirmative about the behaviour of the ocean circulation given that the study only considers the Greenland records, without being substantiated by further references. My advise would be to concentrate on the identification method, which is complicated enough, possibly propose further applications, but remain more open about specific climate mechanisms.

3.1 Miscellaneous

Equation numbering would have been helpful.

Interactive comment on Clim. Past Discuss., 6, 1209, 2010.