Bayesian Analysis (2008) **3**, Number 1, pp. 45–56

Comment on article by Sansó et al.

Jonathan Rougier*

1 Introduction

This paper represents a very welcome combination of Statistics and Climate Science. I am sure that no-one who has studied the paper is in any doubt about how demanding this type of collaboration is: it is splendid that statisticians and climate scientists are working together to understand better uncertainty in future climate.

As a statistician developing methods for computer experiments, I like climate science precisely because it is so challenging. In particular, the models are still quite poor on the scales for which we would like to use them (transient and regional behaviour). That is to say they have large *structural errors*: errors that cannot be removed simply by tuning the model parameters. They are also some of the most expensive models in the world to evaluate. Typical performance is about three model-years of output per day at the main research centres. Tony O'Hagan (2006) has termed the consequence of this paucity of evaluations 'code uncertainty'. In some applications we also have to contend with the scale of the model outputs: the state vector can easily have millions of components.

The MIT2DCM of Sansó *et al.* is of relatively low resolution, and in this case-study the focus is on just three uncertain model-inputs, so code uncertainty is not going to be a problem. As a consequence of the low resolution, though, and the small number of uncertain inputs, structural error is going to be crucial. For the last forty years, the trend in climate science has been towards higher and higher resolution models, and this will continue because so much of the important physics is missing even at current high resolutions (where a grid-cell in the solver is typically about 250 km a side). There are important questions concerning how much we can learn from low resolution models, and one of the projects I am working on addresses exactly that, by trying to understand the structural links between models along a spectrum of modelling refinements.

Sansó *et al.* are interested in calibrating MIT2DCM, i.e. using observations on climate to learn about the correct setting of this model's parameters. Probabilistic learning requires a statistical model that links (i) evaluations of MIT2DCM, (ii) the model's parameters, and (iii) observations on climate. A crucial component of this statistical model is the treatment of structural error. One thing I particularly like about this paper is that Sansó *et al.* have explicitly included a term for MIT2DCM's structural error (which they term ξ). Actually, in their treatment this term combines structural error, representation error (incommensurability of the grid-averaged model outputs with the point observations) and observation error, but the first of these is likely to dominate. Currently the predominant practice in climate science is to invoke the caveat

*Department of Mathematics, University Walk, Bristol, UK, mailto:j.c.rougier@bristol.ac.uk

 ba0009