Bayesian Analysis (2011) **6**, Number 4, pp. 529–532

Comment on Article by Wyse et al.

Paul Fearnhead*

I would like to start by congratulating the authors on a stimulating paper. The focus of the paper is on computationally efficient methods for analysing multiple changepoint models. They build on existing methods (Yao 1984; Barry and Hartigan 1992; Liu and Lawrence 1999; Fearnhead 2006) that allow for iid sampling from the posterior distribution for certain changepoint models. These existing methods require the ability to analytically calculate the marginal likelihood associated with any segment within the data. For analysing a data set with n data points, under the assumption of K changepoints, the CPU cost of these methods is $O(Kn^2)$.

There are two key ideas within this paper. The first is to use ideas from the integrated nested Laplace approximation (INLA) to approximate the segment marginal likelihoods. This enables you to apply these recursive methods to a much wider range of changepoint models, that is to models which include dependence of data within a segment when it is modelled through a Gaussian Markov random field. The second is to implement the recursions on a reduced set of possible changepoint times. This can be thought of as grouping each g consecutive data points into a single observation, and then running the recursions on this reduced set of observations. The motivation for this is purely computational – as it reduces the CPU cost of the recursions by a factor of g^2. In many applications this approximation has a natural interpretation. For example, for the coal-mining disaster data (Section 4 of the paper), different choices of g would relate to analysing data at different levels of aggregation: such as corresponding to data on the number of deaths each day, week, or year. Providing the distance between successive changepoints is larger relative the level of aggregation, we would expect any approximation error to be small.

The key feature of both ideas is to introduce some approximation, but with the gain of being able to analyse a much wider class of models and a much bigger size of data set. I would like to first discuss, via asymptotic arguments, in what sort of situations these approximations are likely to be small; and secondly to look at the idea and approach of summarising the inferences via a MAP estimate of changepoint positions.

*Department of Mathematics and Statistics, Lancaster University, Lancaster, UK, p.fearnhead@lancs.ac.uk

 DOI:10.1214/11-BA620A