

This manuscript presents an overview of open source spatio-temporal modelling software aimed at the paleo-environmental community. The modelling approaches are formulated in a hierarchical Bayesian context with different implementations sampling algorithms of varying complexity (and calculation speed) and these are demonstrated with a variety of case studies (limited here to sea-level models). Overall, the idea of putting these tools out there is to be welcomed, and I feel this will be a useful contribution in terms of advertising the availability of the software. In the end, success will be determined by the uptake by non-expert users. To this end, at least from my experience, key aspects are good documentation, relative ease of use and some tutorial examples that allow a user to repeat, and develop on, previous studies. Given the nature of the methods implemented, there is a degree of technical know how required to understand the statistical details but as far as I can see, the examples are all available and should allow a new user to at least try things out without perhaps necessarily being on top of the details. So, the manuscript needs a bit of work (minor changes), but I think this is worth publishing in the sense of put it out there and see how it goes - hopefully well and ideally leading to a wider range of applications than presented here.

#### Some points

I felt there are perhaps some technical terms/acronyms that could be defined for the less experience potential users. Too many undefined terms and/or jargon will tend to put off the non-expert and/or those not in the sea-level community – I have highlighted some below. In addition to the definition of terms table... perhaps a table defining the symbols, etc used may be use.

There is no mention of transdimensional MCMC methods, nor model choice (in terms of model complexity for example)...the latter is not always simple but is there any sense of this in the current version (or planned for future versions ?)

L13 out-of-sample data ??? Not sure what this means.

L14 (and L45) for me data are plural and should use are not is (data are), but this may be a slightly old fashioned view.

L17 do we mean reconstructing paleo-environment or paleo-environmental signals ?

L18 – perhaps define hierarchical here...

L19 clearer (than what ?)

L20 – It has been...you mean hierarchical statistical models from the previous sentence...They have been ?

L23 GMSL – Global Mean Sea Level...

L26 I would start a new paragraph after 2013..

L26 such techniques ? Remind us in a word or two.

L45 Marginal likelihood, or evidence...(could define it  $p(y) = \int p(y|\theta)p(\theta) d\theta$ )

Perhaps the phrase static observations is a little ambiguous (confusion with temporal data)...would it be better to say a given data set (which could include temporal data).

Table 1

Hyperparameter – parameter of a prior distribution (to be inferred) ?  
MCMC – random sampling (need a space)  
Residuals – definition implies this should not be plural.

L53 RSL – Relative Sea level ?

L55 inverting conditional probabilities – not sure we formally inverting conditional probabilities, but use Bayes' theorem (historically this was often considered as the inverse probability problem).

L58 Eqn 3 – not too important, but the last term could be written as the product of the two probabilities as  $\theta_s$  and  $\theta_d$  are being treated as independent ?

L66 Conditional model parameters ? Also could the unobserved physical parameters being treated as nuisance parameters (that we need to know, but are not that interested in...but perhaps here we can not integrate them out).

L68 linear rate in a linear sea-level ..constant rate (too many linears ?)

L94 Paszke et al. 2017...perhaps this the best citation for PyTorch ?

L104 L3 should be L4 ?

Eqn 4 perhaps just redefine  $y$  here as observed data (or prediction ...as uncertainty is defined as difference in prediction and true value).

L140 – for me the covariance matrix contains the errors covariance, not data covariance and the common assumption of  $\text{cov} = 0$  is that the errors are not correlated...this is not always the case in time series (as implied in the text). This seems to be stated later..L238 – uncorrelated uncertainty ?

See also Sambridge, (2016), Reconstructing time series and their uncertainty from observations with universal noise, *J. Geophys. Res. Solid Earth*, 121, 4990–5012, doi:10.1002/2016JB012901 and/or Bodin et al. (2012), Transdimensional inversion of receiver functions and surface wave dispersion, *J. Geophys. Res.*, 117, B02301, doi:10.1029/2011JB008560.

L145 Why Whereas ? ...Perhaps in practice ?

L153 paleo-env. signal is modeled ?

L158 for me a linear model is not just a straight line, it is where the forward problem is linear, such that model parameters can be separated from a kernel matrix (so a polynomial of order 20 is still linear)....perhaps just say straight line model

L165 – what happens if we do not know  $m$ , the number of changepoints ? any comments on transdimensional change point models (or indeed transdimensional models in general)..e.g. Gallagher et al. (2011), Inference of abrupt changes in noisy geochemical records using Bayesian Transdimensional changepoint models, *Earth Planet. Sci. Lett.*, 311, 182–194.

L173 – not clear to me...the product of  $\beta_{k-1}$  and the difference of  $\gamma_k$  and  $\gamma_{k-1}$ ....what does this mean in practice ?

L181 – covariance matrix - data or data errors ?

L196 why Yet ?

L205 fundamental (not foundational)..

L213 – will allow...should allow ?

Table 2 caption  $t$  denotes age of the data...a little clumsy ...could be the age of the analysis...a few months. Age of the sample ?

L253 – again a little clumsy...to take a range of probable values....to take on a range of values,

potentially with different probabilities

L267 – why do you say relative (explain this...).

L282 -283 – not clear for me...explicit prior for the parameters (distinguish from hyperparameters, if relevant, with an example ?)

L284 optimize hyperparameters using their maximum likelihood estimates..

L318 – to demonstrate (rather than enhance)...

L330 – we can look at residuals to assess if our model produces normal distributed uncertainties ?

L345 – noticeable, but significant ? Does not really look too important in terms of the uncertainty ranges on the slope.

Fig 3 – can we explain why the uniform likelihood gives multimodal distributions on the rate and intercept ?

How do the correlations of these two parameters differ for the two likelihoods... ?

L355 – could we let the number of changepoints vary - and the same for fig 4 – if we had 10 changepoints...perhaps we capture the GP model form better ?

L336 – language a little clumsy – through a variational Bayesian manner...rephrase.

Fig 4 – are the likelihoods much different for the preferred model and for the MCMC samplers can we see the likelihood as a function of iteration ?

L366 – environmental statistics problems ? How do we know the model is too smooth ?

L378 their underlying uncertainty.

L383 bias of the perhaps overly simple approach ?

L390 no need for The...

Fig 5 the bottom right panel is a little hard to read with the different colours overlapping...I could not figure out where the orange region is ...you could trace the bounding curves to show that

L396 define GIA ?

L398 ensemble

L409 their prior distributions

L415 mean expectation of what ?

L416 – mean = expected ?

L472 weighted according ?

L482 – something missing between hierarchical model and any statistical model

L499 – comma rather than and before Piecuch

L549 In contrast to what ?