Clim. Past Discuss., 9, C783–C791, 2013
www.clim-past-discuss.net/9/C783/2013/

Climate of the Past Discussions

CPD

9, C783–C791, 2013

Interactive Comment

Interactive comment on "Bayesian parameter estimation and interpretation for an intermediate model of tree-ring width" *by* S. E. Tolwinski-Ward et al.

S. E. Tolwinski-Ward et al.

tolwinsk@ucar.edu

Received and published: 15 May 2013

1 Response to Comments from Reviewer 1

We appreciate the reviewer's constructive suggestions, which have substantially improved the clarity of the manuscript and the rigor of the evidence presented.

1.1 General comments

We have rewritten section 2, as well as parts of the introduction and abstract, to make the structure of the likelihood clearer as recommended by the reviewer. While climate is indeed a spatio-temporal phenomenon, trees only experience and respond to local environmental conditions. This is reflected in VS-Lite, which models tree-rings locally and deterministically. This is explicitly discussed in a new introductory paragraph to section 2, but we also note the independence of the parameter inferences at each site throughout the paper. The likelihood for the data at a given site is now given in equation 5.

1.2 Section 2.2 comments

In fact our priors are not species dependent as suggested by the reviewer, and we have made this modeling choice clear and described our reasons for it in section 2.2 (pp 5, lines 4-13 and pp 8, lines 12-25 in revised manuscript). We also touch on future work that might proceed in this direction in the Discussion (pp 21, line 25- pp 22 line 1).

We thank the reader for pointing out the errors in our parameter specification for the beta priors. In both the manuscript and the code, we have replaced our original parameterization in terms of the scaling and intercept of the transformation from the unit interval with the standard four-parameter parameterization in terms of the lower and upper limits of the support. We have also added statements of 90% probability intervals to the description of all the priors as suggested.

As suggested by R1, we checked the reproducibility of the point estimates given by the estimation procedure by running it twice at two representative sites for both the PPE with SNR = 1 and for the OMPC. Due to improvements in the suitability of SETW's computing environment for batch runs, the number of MCMC draws was also increased to 1,800 (after subsampling from three chains with 30,000 autocorrelated draws each) to

CPD
9, C783–C791, 2013

improve the reproducibility of reported results. In both the duplicated PPE and OMPC runs, the differences between estimates of the two temperature thresholds T_1 and T_2 from two runs of the algorithm was less than a tenth of a degree Celcius; moisture thresholds M_1 and M_2 were reproduced to within 0.004 v/v; the ratio of posterior to prior variance was reproduced in all cases to within 0.05; and the estimated signal-to-noise ratios were within 0.02 of one another for the two runs. This result is given on page 15 in a new paragraph about general experiment-independent results (lines 8-22), and a table of the estimated values for the reproducibility experiments is included in the supplementary material.

The smoothed density plots drawn in figure 1 are estimates of the density function given by running a normal kernal smoother over the frequency plot of discrete posterior draws; this is now stated explicitly in the figure caption.

Regarding computational expense, running 3 MCMC chains of length 30,300 in series at a given site on a Macbook laptop with a 2.7 GHz processor takes less than three-and-a-half minutes for sites using the white noise model, and less than nine minutes for sites using the AR(1) model, which is reported in the general results section (pp 15, lines 4-7).

1.3 Section 2.3 comments

The intellectual framework in which the PPE and the observed proxy calibration are carried out is now explained at the beginning of section 2. Although we have not restructured the divide between the pseudoproxy and observed proxy analysis, we believe these should be clearer to read due to this last new paragraph, and because we no longer split the intervals for calibration of different quantities in the PPE.

CPD
9, C783–C791, 2013

1.4 Section 3 comments

In addition to revising the contents of the Results section in accordance with the change in study design, we have also taken care to include more interpretations in this part of the manuscript.

1.5 Section 4 comments

That posterior inference is determined by the prior model when the data series are short and/or noisy is a general feature of Bayesian inference, which is now clearly reflected in a comparison of the posterior to prior variance ratios of the PPE performed with SNR = 1 and SNR = 0.25. Much less "learning" occurs in the latter experiment, and so the posterior to prior variance ratios tend to be closer to one than in the former experiment. To the extent that the priors represent scientifically-based understanding of the parameter values and their uncertainty, a lack of learning from noisy data is not a problem in our view; rather this feature of the analysis just underscores the importance of careful prior elicitation. We have revised this part of the discussion to make these views clearer.

Though we have not made a new section of the paper about related future work, we have expanded our description of possible related future projects in the discussion section.

1.6 Minor typographical errors

The manuscript has been checked and revised for typographical and spelling errors.

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

2 Response to Comments from Reviewer 2

We appreciate Referee 2's review of the structural aspects of our study. We believe that re-running our experiments with some changes made as per R2's suggestions has substantially improved the work.

2.1 Major Comment 1

The reviewer states that the pseudoproxy validation experiment would be both more convincing and more rigorously Bayesian if stochastic noise was added to the synthetically-generated data. We are convinced by this argument and thank the reviewer for clearly explaining the reasoning behind it. In two revised versions of our previous pseudoproxy experiment (PPE), we now estimate the model parameters given pseudoproxy data which have been appropriately scaled and have had stochastic noise added to them with two signal-to-noise-ratios: SNR = 1 and SNR = .25. These represent optimistic and pessimistic estimates of the real-world SNR for climate proxies and are values that are both typically tested in climate reconstruction pseudoproxy experiments (see eg. Smerdon 2012, cited in the main research article.)

This revision to the structure of the PPE necessitated a structural change to the data-level model, or likelihood used in the Bayesian model for the parameters. This change is now clearly described in both sections 2.1 and 2.2. Because both the observed data series and VS-Lite output are standardized to have zero mean and unit variance, a coefficient related to the variance of the additive model error now appears in front of the estimate from VS-Lite (equation 3). This change in the modeling also allows for straightforward computation of the posterior SNR in terms of the posterior estimates of the error model parameters. We now also include frequency plots of the posterior SNR to convince the reader that the Bayesian model can recover the value of the additive noise, as suggested by R2.

2.2 Major Comment 2

The reviewer raises the following two modeling choices related to the context of the parameter estimation:

2.2.1 Spatial independence of parameters

The formal structure of VS-Lite represents the gross mechanisms by which climate influences growth across all trees. As discussed in response to R1's comments, the inference for the model parameters is performed independently at each site, without explicit spatial dependence built into our models. This reflects the fact that trees experience and respond to local conditions, though similarities may emerge across space in the inference given that the driving climate inputs, of course, are spatially correlated. Still, there is no a priori reason to expect that trees close to one another in space should necessarily share similar parameters, as we view the parameters as accounting for differences in tree response to local environmental conditions, and chronologies quite close to one another can exhibit distinct climate controls (eg. sites might be close in space but differ inaspect, species, or elevation, where upper treeline sites are tempreature-limited and lower treeline sites are moisture sensitive). This approach is described clearly in subsection 2.1.

2.2.2 Error modeling

The form of the likelihood for the data has been changed since the original submission to make the error more interpretable. We now consider both white and AR(1) error models for the data. We first run the parameter calibration at each site assuming independent gaussian errors. We then compare the fit of a white noise model and and AR(1) noise model to the residuals of simulations run using the posterior median

CPD
9, C783–C791, 2013

of the estimated growth response parameters. If the AR(1) model has a lesser value of Schwarz's Bayesian Criterion, then the parameter calibration is re-run for that site under the assumption of AR(1) model errors. This procedure results in 106 out of 277 sites with a white noise error model, and 171 sites with an AR(1) model, and is described in the methods.

2.3 Major Comment 3

The reviewer requests improved evidence regarding three points:

2.3.1 MCMC convergence

As suggested, trace plots of the three MCMC chains run for the Sipsey Wilderness site PPE are now provided in the supplementary material. Note that since our original manuscript submission, we also increased the length of the chains run for the analysis.

The reviewer suggests improved transparency in the documentation for the algorithm, but in fact the code for the MCMC was provided with the online supplementary material at Climate of the Past along with the original Discussion Paper. The authors view the availability and usability of the code as an important component of this work's contribution. We have provided the revised code along with the revised manuscript, and feel that the commenting and help file for the code are quite thorough. We plan to archive the code at the NCDC Paleoclimate Software Library as stated in the manuscript as soon as the code and manuscript complete the review process.

We have also included information about the reproducibility of estimates drawn from the posterior in the results section (page 15, lines 8-22) and included a table in the supplementary documentation, which also help quantify the degree to which convergence occurs at two representative sites for the number of samples drawn in this study.

CPD
9, C783–C791, 2013

2.3.2 Qualitative fit of VS-Lite to observed data

The qualitative fit of VS-Lite to observed data at each site is addressed in Figure 7, which compares the correlation of simulated data with observed data in the validation interval for calibrated simulations and for simulations computed using prior median parameters, as well as the RE statistic for calibrationed and uncalibrated simulations. The figure is discussed in the Results and Discussion.

2.3.3 Representativeness of the vector of marginal medians as posterior summary

The posterior distributions of parameters generally appear to be unimodal and without complex shape. We believe the vector of marginal medians provides a representative summary of the center of their distributions, and we present scatterplots of the pairwise posterior samples and the marginal medians in the supplementary documentation for one representative site. While we continued to use the vector of marginal medians in our analysis, we have added the option for the code to return the vector of sampled parameters that maximize the data likelihood as point estimates, rather than the marginal posterior medians.

2.4 Major Comment 4

The reviewer suggests that the setup could be clarified by removing the division between calibration and validation intervals in our experiments. In the pseudoproxy experiment, we now use the entire data interval to estimate the growth response parameters of VS-Lite and the parameters of the error model since both sets of parameters can be validated against known targets. In inferring the growth response and error model parameters using observed data, we maintain a sepatation between intervals ysed for these respective tasks as is standard practice in dendroclimatology and also

CPD
9, C783–C791, 2013

Interactive Comment

provides a check against overfitting (see first paragraph of section 2.3.2).

Regarding R2's comments about the efficiency of our algorithm, in describing the approach as "efficient", we merely meant to suggest that the code to perform the inference, which is publicly available, is quick enough to be used in typical laptop computing environments in times taking on the order of minutes, and is thus efficient enough to be practical for users of VS-Lite who would like an objective means for deciding on a set of model parameters. We have made this clear in both the Introduction (pp 4, line 20) and Results (pp 15, lines 4-7). We do not compare estimates of the parameters made globally versus locally in this manuscript, as the reviewer seems to have misunderstood. The re-writing of secion 2 should help clarify this for the general audience.

Interactive comment on Clim. Past Discuss., 9, 615, 2013.