Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2019-16-AC2, 2019

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Joint state-parameter estimation of a nonlinear stochastic energy balance model from sparse noisy data" *by* F. Lu et al.

F. Lu et al.

feilu@math.jhu.edu

Received and published: 28 June 2019

Dear Referee,

We thank you for carefully reading the manuscript and for the valuable feedback. In the following, we respond to your comments and suggestions.

Reviewer's comments: Paleoclimate reconstruction seems to be an interesting and nonstandard question. So it might be worthy to give some details on how the SEBM is formulated, for example, what do each parameter θ stands for. It is also better to give

some references on why the specific parameters are chosen. For example, it seems that data are available with time-interval $\triangle = 0.01$ (what is the time unit here?). But is this practically true?

The SEBM follows the atmospheric model from Fanning and Weaver (1996), which contains parameterizations for incoming shortwave radiation, outgoing longwave radiation, radiative air-surface flux, sensible air-surface heat flux, and the latent heat flux into the atmosphere. The parametrizations contain terms of 0th, 1st, and 4th order, which are aggregated together such that roughly the 0th order terms correspond to incoming shortwave radiation and albedo effects, the 1st order terms correspond to air-sea heat (energy) exchange, and the 4th order terms correspond to longwave radiative transfer expressed as graybody emissivity. The prior distributions of θ aggregate the contributions of the different energy sources and sinks according to their parameterized polynomial order using the parameter and uncertainty estimates from Fanning and Weaver (1996). This first guess is adjusted using estimates of the current earth energy balance from Trenberth et al. (2009) to increase the physical consistency of the estimates. Thereby, at the equilibrium temperature, the contribution of the individual parameterized processes is very close to the estimates from Trenberth et al. (2009).

It should be noted that our parameterizations and the corresponding prior distributions are idealized. For example, spatial variations of the parameters are not yet included. However, we do not anticipate that increasing the realism of the SEBM would reduce the ill-posedness, as it would lead to a more complex parameter dependence structure without increasing the number of available (paleo-)observations.

We have briefly explained the prior range and model structure in the paragraphs in lines 28-30 on page 3 and 7-13 on page 4. We add the sentence "The nonlinear function $g_\theta(u)$ aggregates parametrizations from Fanning and Weaver (1996) for incoming shortwave radiation, outgoing long-wave radiation, radiative air-surface flux, sensible air-surface heat flux, and the latent heat flux into the atmosphere according to their polynomial order." to this description as additional clarification of the model structure.

The time unit is year. We add a sentence ‘and one time unit represents a year” to line 6 in page 4. The time-interval $\triangle = 0.01$ represents about 4 days. In practice, missing or temporally integrated observations should be expected. Missing observations can be incorporated easily in our model, and would increase the reconstruction uncertainty similar to the reduction of the observed nodes (cf. Section 5.1). To include integrated observations, the observation operator can be interpolated (please see a new 'Section 5.4' in the revised manuscript in response to the first referee).

Reviewer's comments: While I agree the Fisher information matrix may be ill-conditioned, but I don't see immediately why the strong regularization approach is the right or natural way to fix it. The numerical results show that the regularized posterior has obvious biases, and sometimes close to being the prior. The strong regularization used here might be the cause of this. An alternative approach might be using the following version instead of (21)

$$p^N(\theta|u,y) \;\propto\; p(\theta)^\alpha [p_\theta(u)]^{1/N} \tag{1}$$

where α is a parameter in [0, 1], and it can be tuned for a better posterior.

We thank the reviewer for the above alternative approach, which is in line with our strongly regularized posterior (with $\alpha = 1$), and it aims to balance the contributions of the prior and the likelihood. As we admitted in the original manuscript, (page 12, line 16-17) the factor $1/N$ might not be optimal. However, it is not clear how to tune the optimal α or the factor $1/N$, and the development of a strategy for optimal regularization factors is beyond the scope of the current study. Therefore, we postpone it to future work. We have added a comment in the manuscript about this alternative approach.
The need of a strongly regularized posterior is due to the degenerate likelihood, which is indicated by the ill-conditioned Fisher information matrix, as demonstrated numerically in Section 3.2 for different data sizes. In this case, the regularization by a prior in the standard Bayesian approach is not sufficient (as indicated by the Bernstein-von

Mises theorem in the asymptotic setting). Therefore, we need a stronger regularization that increases the contribution of the prior.

Reviewer's comments: Page 9, Fisher information matrix: since most NPG readers are likely to be geoscientists, maybe you should explain that in statistics, Fisher information dictates the asymptotic inference difficulty and give references. Also, there should be some explanations on why this matrix is ill-conditioned, not just some simulation plots.

While we agree with the reviewer that the Fisher information matrix dictates the asymptotic inference, we would prefer not to distract the readers to asymptotic inference, particularly because this manuscript focuses on non-asymptotic study.

Following the reviewer's suggestion, we added an intuitive argument on why the matrix is ill-conditioned on page 9 line 17: “As $N \to \infty$, the Fisher information matrix converges, by ergodicity of the system, to its expectation $\left(\Delta t\sigma_f^{-2}\mathbb{E}[(\mathbf{A}u_n)^{\circ k} A_{\mathcal{T}}^T \mathbf{C}^T \mathbf{C} A_{\mathcal{T}} (\mathbf{A}u_n)^{\circ l}]\right)_{k,l=0,1,4}$, where the matrices $\mathbf{A}$, $A_{\mathcal{T}}$ and $\mathbf{C}$, arising in the spatial-temporal discretization, are defined in Section A1. Intuitively, neglecting these matrices and viewing the vector u_n as a scalar, this expectation matrix could be reduced to $(\Delta t\sigma_f^{-2}\mathbb{E}[u_n^k u_n^l])_{k,l=0,1,4}$, which is ill-conditioned because u_n has a distribution concentrated near one with a standard deviation at the scale of 10^{-2} (see Figure 1).”

Reviewer's comments: Figure 2. The figure caption below "Data size $\log_{10} N$" is garbled. This happens to many figures later on as well. This might be a problem with my own computer/printer. But you better check.

We checked the captions, but garbled captions did not appear on our computers.

Reviewer's comments: Page 11, what is u_i? And it is better to define u_c and σ_o with mathematical terms.

In Eq. (18), the subindex i in u_i indices the time. In the revised manuscript, u_i is replaced by u_n since n is used throughout the manuscript. To avoid confusion of notation, we rewrote u_c as $\overline{u}_c$. Since u_c and σ_o denote the mean and standard deviation of the observations, we believe that mathematical formulas for them are not necessary. In particular, they would lead to notational complexity.

Reviewer's comments: Page 13, I think Markov chain might be too abstract a term for NPG. I think you can replace it with MCMC for the same meaning.

We change the title of Sect. 4.1. to “Diagnosis of the Markov chain Monte Carlo algorithm". However, in the following MCMC refers to the method while Markov chain refers to the output of the MCMC algorithm. Thus, replacing Markov chain by MCMC throughout the section would be wrong and we still use Markov chain in the revised manuscript.

Reviewer's comments: Page 18, line 2, there should not be parenthesis for θ_0, θ_1.

We clarify the previous formulation by replacing it with “the correlations between θ_4 and θ_0 as well as θ_1 are weakened".

Reviewer's comments: Page 29, line 7, "see e.g. [". You miss some content here.

Thanks for finding the typo. The correct form is “(see e.g. Doucet and Johansen, 2011)".

References

Fanning, A. F. and Weaver, A. J.: An atmospheric energy-moisture balance model: Climatology, interpentadal climate change, and coupling to an ocean general circulation model, J. Geophys. Res. Atmos., 101, 15 111–15 128, 1996.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J.: Earth's Global Energy Budget, Bull. Amer. Meteor. Soc., 90, 311–324, 2009.

Please also note the supplement to this comment:
https://www.nonlin-processes-geophys-discuss.net/npg-2019-16/npg-2019-16-AC2-supplement.pdf

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2019-16, 2019.