General comments

The ms 'Bayesian analysis of early warning signals using a time-dependent

model' interprets geoscience time series containing DO events through the lens of an AR1 process and assumes a low-order Taylor expansion time-dependent propagator. The propagator's parameters are determined through Bayesian learning from the time series' segments that are quasi-stationary. From the analysis, the authors identify some of the DO events as bifurcations, while others are seen as merely noise-induced.

To the best of my knowledge, utilizing a time-dependent AR1 process to diagnose a system approaching a bifurcation is novel. By definition, data are exploited best by utilizing a statistic that explicitly contains the time-dependence, as against utilizing a moving window in combination with a static AR1 process. Here the ms provides a great service to the geoscience community in demonstrating such an approach can be implemented. I very much would like to see this ms being published with a 'peer-reviewed' status in an EGU journal.

Answer: We thank the referee for the helpful review. The comments will be addressed point-by-point below.

However, the ms should be modified in two main respects. Firstly, it should become clearer what is the domain of applicability of the utilized method. When analyzing time series through the lens of an AR1 process, one lives in a quadratic approximation of the potential as shown in Fig1. This in turn can only be justified in a small noise expansion. However, the ms provides no hint why the small noise expansion might be justified. Quite the contrary, the Conclusions section suggests that a subset of DO events is rather noise- than bifurcation-induced. So, the noise level is seen as large enough to trigger jumping to another equilibrium. This raises doubts whether a small noise expansion is compatible with the time series at hand.

Answer: We will add a comment addressing this in the revised manuscript.

Secondly, I would expect that most readers of EGUsphere are no trained statisticians. Most natural scientists might have heard of Bayes' formula. However, it might be useful to recover the Bayes' principle in a short Appendix. I personally found some of the wording on page 8 inaccessible, such as 'latent field'. I find it necessary that anything transcending elementary Bayesian learning is clearly defined somewhere – either in the main text (preferred) or an Appendix. So here I am asking for a didactical upgrade of the statistical method used in view of a natural science audience.

Answer: We will rewrite this section to include a high-level discussion on Bayesian inference, with more details in the appendix.

Overall, a timely and exciting to read article which is apparently on a very high technical level.

Technical corrections

1. P4: On the history of early warning systems, the following additions might be in order. (1) First mentioning of a noise-induced precursor of a bifurcation: Wiesenfeld (1985), (2) first extension to a complex system, justifying 1D AR1: Held and Kleinen (2004), (3) first utilization on real data (in fact, ice core data): Dakos et al. (2008).

Answer: These additions are appreciated and will be added to the manuscript.

2. L90-100: How does this § relate the other parts of the ms? Later on, an AR1 process is utilized, while this § seems to suggest that it should not be utilized. Furthermore, the Green's function as of Eq9 might be interpreted as a superposition of Green's functions as of Eq7 which could easily occur in multi-dimensional systems. Should one then simply expand the presented formalism to larger dimensions than one? So, I am confused about the logical positioning of that § in the overall ms.

Answer: We agree that the discussion of fractional Gaussian noise and the Hurst exponent is tangential and not really relevant to the main message of the paper which focuses on the AR(1) process. We will remove this in the revised manuscript.

3. I have issues following the rationale of Section 3.1. In the context of Bayesian learning, I would expect that an observation variable is defined (is it x or is x observed through an additive layer of noise – please clarify), and the conditional probability of an observation (in our case a time-correlated time series) on uncertain parameters is presented. In this context, I do not understand the definitions of eta, beta, z, y, for what reason I need to postpone my review of that part to another iteration.

Answer: Section 3.1 addresses how we fit our model into a hierarchical Bayesian modeling framework. Essentially, the data (y) follows a regression model expressed by the predictor (eta) which may depend on some fixed effects (beta_i) of known covariates (z_i) and noise (epsilon), here a time dependent AR(1) process. This section includes details which may be overly technical. Moreover, we used x to denote all random variables included in the predictor (as well as the predictor itself) while in the previous section we used x to denote the time dependent AR(1) process.

We understand the confusion and will update our notation to be more consistent in the revised manuscript. We will also rewrite this section to attempt to make it easier to follow

4. Eq29: Why do we need non-constant time-steps? Is it due to missing data?

Answer: In certain applications, one may not have data sampled at constant time steps. For example, the NGRIP/GICC05 dataset used in this paper is given in constant 5cm steps in ice core depth, but the corresponding time is irregular.

5. Caption of Fig5: What are the intervals showing? Are they 2.5-97.5% quantiles of the posteriors of those variables? Or are they confidence intervals in the frequentist's sense? Would the latter logically consistent with a Bayesian setup?

Answer: The intervals are indeed the 95% credible intervals taken between the 2.5 and 97.5% quantiles of the posterior distribution. We will mention this in the main text.

6. L110: Why are you utilizing a Bayesian approach at all? You emphasize the benefit of having a PDF as output, which has tremendous advantages when later being utilized in economic decision theory. However, you do not mention the drawback of the Bayesian approach: That you need to justify the choice of a prior. What is your justification and what prior was chosen? Is it a 'vague' prior (L185, whatever that means). In the whole ms, I could not find a single reason why a Bayesian approach as against a frequentist's approach was necessary. The utilized likelihoods could likewise have been utilized for a frequentist's approach. Hence, some more motivation of the choice would be helpful.

Answer: The Bayesian framework is very useful in detecting early warning signals since we get uncertainty quantification expressed by the posterior distributions. It also allows for prior knowledge to be incorporated through the prior distribution.

The choice of prior distribution is an essential aspect of Bayesian analysis and should have been given more attention in our manuscript. Since we do not have prior knowledge about the parameters we use vague Gaussian priors, i.e. prior distributions with very large variance on all parameters. In our revised manuscript we use instead a gamma distribution for the precision \kappa=1/\sigma^2, and uniform priors on *a* and *b*. Since *a* depend on *b* we assign a conditional uniform prior on *a*. We will discuss this in more details in the revised paper. Moreover, we will discuss how robust our model is with respect to how informative our priors are.

Specific comments

1. L168: stochastic -> probabilistic? (to distinguish a situation from aleatoric uncertainty?)
2. Define ‘hierarchical Bayesian model’.
3. Define kappa before it is utilized in Eq23.
4. Activate ‘day month year’ (eg L330, to be found more than once in the ms).

Answer: These comments will all be addressed in the revised manuscript

Literature

Dakos, V.; M. Scheffer; E.H. Van Nes; V. Brovkin; V. Petoukhov; and H. Held. 2008. Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences* 105:14308-14312.

Held, H. and T. Kleinen. 2004. Detection of climate system bifurcations by degenerate fingerprinting. *Geophysical Research Letters* 31:L23207.

Wiesenfeld, K. 1985. Virtual Hopf phenomenon: A new precursor of period-doubling bifurcations. *Physical Review A* 32:1744.