Nonlin. Processes Geophys. Discuss., 2, C385–C393, 2015
www.nonlin-processes-geophys-discuss.net/2/C385/2015/

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Efficient Bayesian inference for ARFIMA processes" *by* T. Graves et al.

T. Graves et al.

christian.franzke@uni-hamburg.de

Received and published: 5 September 2015

Response to "Interactive comment on "Efficient Bayesian inference for ARFIMA processes" by T. Graves et al." by Anonymous Referee 1

The reviewers comments are in italics while our responses are in normal font.

Thank you very much for reviewing our manuscript and your valuable comments.

B. Recommendation In my opinion, the method presented in this paper is a novel way of applying Bayesian reasoning thoroughly to the chosen problem of inferring LRD in time series data. The authors have done a thorough job of grounding their approach on a rigorous mathematical foundation and presenting the theory unambiguously and carefully. That being said, the chosen geo-scientific example in the manuscript, viz. the

Nile river minima time series was given only a short, cursory treatment and no effort was made to infer the results obtained by the method in this case. Given the focus of NPG and its target audience of geoscientists, I feel that the authors could move some of the mathematically intensive sections of their method to an appendix and thereafter devote more space and attention to the application of the method to the Nile data and how the method relates to physically understandable features of the Nile river system. I thus recommend publication in NPG only after a few major shifts in presentation and focus of the manuscript, which I hope the authors can address with a major revision. Besides these points, there are certain technical issues with the presentation that I list out below. I hope that these suggestions help the authors to better their manuscript.

Thank you very much for your positive comments. We have improved the presentation and moved some mathematical parts into an appendix. We also enhanced the geophysical examples by including an analysis of the Central England Temperature time series.

C. General Comments 1. A first major point that could be improved is the focus given to the application of the method developed by the authors to real data. In the current version, the manuscript gives much space to the theory and constructed, illustrative examples which, even though necessary and absolutely crucial, should ideally in the end lead to an equally detailed example from the real-world with a discussion on how the method performs in that case. I feel that, given the particular focus of NPG, which is at the juncture of theory and application, it is important to properly discuss the relevant real-world implications of the results obtained when applying the method to the Nile river minima for instance. This is lacking in the current version. One such difficulty for me in understanding the final application was that I was unable to relate the parameters p, d, and q to the physical system that gave rise to the Nile data set.

As stated above we have included a new section were we discuss the Central England Temperature. A physical interpretation of the parameters d, p and q has been elusive as the model was originally introduced in econometrics as phenomenological rather

than physical or structural. The simplest explanation, which is generic rather than application specific, is that d describes the behavior on very long time scales whereas p and q describe fluctuations on shorter time scales. However, very recent progress is being made in statistics and physics on bridging continuous time linear dynamical systems and the discrete time ARFIMA models. We have noted this new work by citing Slezak and Weron [2015] as a representative entry point.

2. Another main issue for me was that there were discrepancies between the title, abstract, and the numerical examples in the text in terms of the main focus. The title indicated that the manuscript is about "Bayesian inference for ARFIMA processes", the abstract suggested that the main focus was LRD inference, and finally the text devoted quite some space to "short memory" as well. The authors should take care to clearly state what are their objectives (which they seem to do even now) and thereafter, relate every new topic/application/idea later in the text to the stated objectives (which is lacking).

We have changed the manuscript title to "Efficient Bayesian inference for natural time series exhibiting both Short- and Long-Memory using a parametric ARFIMA process". We also have changed long-range dependence to long-memory throughout the manuscript.

3. A last major issue that I have with the manuscript is that the Tables and Figures do not have adequate captioning. The table captions do not indicate what are the various parameters represented and for what kind of numerical experiments. The figures do not have a clear "visual" legend and neither do the captions indicate what kind of numerical analysis gave rise to the figure. I feel that this is an extremely crucial part in communicating results.

We agree with the reviewer that captions are very important. We have improved the captions of all tables and figures.

D. Specific / Technical comments 1) TITLE: Perhaps the title is a too vague. Maybe the

authors can consider giving a more informative title?

We have changed the title into "Efficient Bayesian inference for natural time series exhibiting both Short- and Long-Memory using a parametric ARFIMA process".

2) P574, L7: "Rather than Mandelbrot's fractional Gaussian noise ..." There is no clear comparison of the method from this study to Mandelbrot's fractional noise. What is the purpose of mentioning it here?

We have deleted this statement.

3) P574, L15: "We illustrate our new methodology on the Nile water level data..." What are the main results? How do they enhance our understanding of the Nile river data and LRD?

Thanks for this comment. We have augmented the abstract to explain that we treat the Nile data and as a new addition the Central England Temperature time series, noting that comparisons are favorable and that the CET involves an extension to seasonal long memory. Summarizing the results for both analysis seemed inappropriate for the Abstract section. However we have added some brief discussion into our Introduction Section.

4) P574, L26: "A standard definition..." even so, please cite a reference for this definition.

We now have added a few references here.

5) P576, L1: "It would offer the ability to marginalize out aspects of a model apparatus and data, such as short memory..." Perhaps this idea is key to the extensions of short memory later in the text? If so, the authors should address this is more detail here and later as well.

Yes, this is a key idea. We have modified some of the text, in the following two paragraphs actually, to clarify that designing a computationally efficient means of marginal-

izing out nuiscance parameters (seasonal and short memory terms) is a novel and important contribution in the paper. However, rather than additing detail into an already long introduction section, we reiterate later in the paper (just before Section 4.1) the importance of this marginalization.

6) P576, L5: ARFIMA has not been defined before this in the main text (only in the abstract).

We now also define ARFIMA in the text.

7) P576, L15: Why is it necessary to include/Cite statements about something being "too hard to work with"?

We deleted this statement.

8) P576, L18: "many of the above drawbacks..." It is not clear to me how many drawbacks were mentioned before this line.

We have changed this into "that the above drawbacks ..."

9) P577, L18: "2 Time series definitions and the ARFIMA model" I personally feel that barring the definition of ARFIMA processes, the rest can moved to an appendix.

We have moved most of this section into appendix A.

10) P577, L22: You defined covariance function ngamma (k) as $Cov(X_t, X_{t+k})$ *but what is Cov. This definition is ambiguous and unclear.*

Cov is the standard covariance function in statistics, which can be found on Wikipedia under the heading "Covariance".

11) P577, L25: "the "backshift" operator" - Here and later, the use of quotes for terms and definitions are a bit distracting. Maybe the authors can find some other way of emphasis?

We have deleted most quotes for terms and definitions.

12) P579, L11: I understand that this is the LRD parameter "d" being referred to here, and which was defined in the Introduction. But I find it hard to relate the |d| < 1/2 statement to the preceding equation, i.e., Eq. 5.

Thanks for this comment. This statement, and the surrounding discussion, has been moved into the appendix. We have adjusted the text here to clarify that we are referring back to Equation (1), which defines the ARFIMA process, keeping in mind the immediately previous discussion on ACV/ACFs.

13) P580, L15: "Choosing p = q = 0 recovers FI(d) ARFIMA(0, d , 0)". Is this the definition of ARFIMA? If so, maybe it is possible to add a line stating this more clearly?

As noted above, ARFIMA is now defined in Equation (1), and the rest of this discussion occurs in the Appendix, i.e., much later.

14) P582, L1: 3 Likelihood evaluation for Bayesian inference. I feel it might be better if the authors added a few words here about the interpretation of the parameters p, d, and q of the ARFIMA process here (even at the cost of repetition) before starting with the model inference part. Also, why is the starting point of the inference is a ARFIMA(0, d,0) process and not some other value of p and q? What kind of a process is this?

Thanks. We now clarify that the phrase "having no short-ranged components" is what we mean by p=q=0 and that this corresponds to a fractionally integrated process. For details we refer the reader to our new Appendix. This restriction is made for simplicitiy: if p or q were non-zero the process would be more complicated. This is, or course, generalized later.

15) P582, L9: "causal" in what sense? Maybe briefly mention here.

"Causal" is a time series term. We povide a definition in our appendix, although it can be found in any standard time series text, e.g., Brockwell Davis.

16) P582, L24: It might be ambiguous to infer AR(P) dependence from a time series of length n = P.

If by ambiguous you mean that there are too many degrees of freedom for accurate inference (of the variance for example), then this is of course correct under least squares or maximum likelihood inference. However, under a proper prior for the AR coefficients, or indeed one that encourages sparsity, there is of course no problem technically with Bayesian inference. It is possible, say if seasonal effects are present in the data but not explicitly accounted for in the model, then identifiability might be a concern. But it is not more of a concern than it would be in a P < n setting.

17) P588, L5: with chains moving between and within models... It is unclear what this means for someone unfamiliar with numerical techniques of Bayesian likelihood estimation and MCMC methods.

Thanks. We agree that it would help to be more explicit by linking "between“ moves to choices of p and q and "within“ moves to inference for phi and theta given p and q. The text has been adjusted accordingly.

18) P588, L7: What is RJ?

RJ denotes Reversible Jump. This is now defined in the text.

19) P588, L9: What is FEXP?

FEXP stands for Fractional Exponential Process.

20) P589, L20: Does MVN mean Multivariate Normal?

Yes. This is clarified where it is first used in Section 4.1.

21) P591, L11: What is a transdimensional move?

The only mention of transdimensional that we could find is in Section 1 where the text reads: transdimensional MCMC, in which the model order (the p and q parameters in the ARFIMA model varies and, thus, the dimension of the problem). In other words this is the same as a between model move, in point 17. To better connect the two passages we have added text to the passage referenced above to link transdimentional

to between model moves.

22) P593, L3: The authors should maybe add one more line on why they choose the Poisson distribution here, it is not clear from the parsimony argument they mention.

In the previous sentence we clarify that the more "complicated“ models are "larger“ ones (i.e., bigger p and/or q). Therefore, a prior that prefers parsimony is one that puts more weight on smaller models. We then adjust the text to read that "As a simple representative of potential priors that give greater weight to smaller models we prefer a truncated joint Poisson ...“ In other words, the particular form of the Poisson isn't important.

23) P595, L22: Is it not possible to show the results in a graphical way, such as a histogram instead of a table?

We think that our table provides a very concise presentation of our test results and would prefer to keep it. However, if the editor feels otherwise we are happy the exchange the table with a histogram.

24) P596, L9: Figure 1 legend is unclear. If I understand correctly, for each value of d_l, there are 1024 estimates and thus 1024 "x" markers on the vertical axis - is this correct? I suggest to use better captions to remove such ambiguities.

We have enhanced the caption of this figure to better describe the plot axis labels and the x-markers. Thanks.

25) P600, L8: For the sake of reproducibility, I suggest that the authors state the source of the data and from where it was obtained. Also, maybe they should devote a few lines on the nature of the data, and preprocessing of the data such as removal of missing values, outliers and the like.

The Nile data is part of the R package 'longmemo' and the CET time series can be downloaded from http://www.metoffice.gov.uk/hadobs/hadcet/ We now state this in the manuscript.

26) P601, L11: Reference missing to the dissertation cited here.

We have added this reference.

Interactive comment on Nonlin. Processes Geophys. Discuss., 2, 573, 2015.