**Climate
of the Past
Discussions**

Interactive comment on "Automated ice-core layer-counting with strong univariate signals" *by* J. J. Wheatley et al.

J. J. Wheatley et al.

j.wheatley@shef.ac.uk

Received and published: 15 August 2012

General comments:

Summary.

Referee:

The CPD paper by Wheatley et al. (2012a) is a nice contribution towards automated ice-core layer-counting. The introduced statistical method yields convincing reculs on the H_2/O_2 series from the Gomez ice core. This good performance is likely owing to (1) the strong annual cycle preserved in that series and (2) the sinusoidal form

of the variations (of the log-transformed H_2/O_2 values). Insofar the adaptation of the method and the convincing results regard to that specific Gomez ice core series, it should be helpful to the reader if that fact could be pointed out more clearly in the revised paper version. In my review I raise three major points; the first deals with a better understanding the performance of the method, the second and third with a wider application of it. The Editors of CPD should check the cited paper Wheatley et al. (2012b) in Environmental and Ecological Statistics for the degree of overlap in content with the present paper. I close with minor points. These are few because, unlike many other papers I had to review (here on CPD and elsewhere), the present paper contains so few errors and is so clearly and well written.

Response:

Firstly we would like to thank the referee for reviewing our discussion paper, and for providing a valuable and thorough proof reading.

The presented method is applicable to other signals with strong annual cycles where symmetry has been induced by transformation. In our response to Anonymous Referee # 1 we include examples on two such signals from the NGRIP ice-core: ammonium and calcium. A third additional analysis on a more challenging signal, Gomez non-sea-salt sulphur, is also presented. These analyses will be included in the revised version. The editor has a copy of our other submitted paper.

Major Point 1: Monte Carlo simulations.

Referee:

The paper by Wheatley et al. (2012a) lacks Monte Carlo simulations. It is good statistical practice to accompany a newly introduced method with Monte Carlo simulations on artificial time series. If the results are satisfactory, and the design of the Monte Carlo

runs is comparable to the problem at hand (here: measured Gomez series), then we can more confidently apply the tested method to the measured series. Roughly: pre-scribe known timescale, generate one artificial series by means of a signal (sinusoid, other) and noise (of red type, additive), estimate timescale using method and compare estimated timescale with prescribed timescale. Repeat the procedure generation estimation many times and calculate error measures (e.g., root of mean integrated squared error). Study the error measure for various data sizes and for various signal/noise ratios. (Data size and signal/noise ratio should be in the same order as the measured series to be analysed.) By imposing non-sinusoidal signals, you can learn about the performance of the method in such mis-specified situations, that is, one can evaluate the robustness of the method. You can compare your method with other methods, for example, that presented by Winstrup et al. (2012). A general reference for Monte Carlo simulations in climate time series analysis is, please let me mention it, my book (Mudelsee, 2010).

Response:

As mentioned above, in response to Anonymous Referee #1 we have carried out analyses of 3 additional ice-core chemistry signals, including various 'thinned' versions to simulate the effect of coarser data. They consist of another signal from the Gomez core, and two different signals from the same stretch of the NGRIP core, allowing cross-checking between different univariate analyses. These analyses of well-understood data-sets thus offer some of the advantages of simulation from a known model or algorithm. In addition, since the focus of our work is on finding regions in a signal where there are apparent departures from an underlying simple model, using these real data-sets has the advantage of ensuring realistic 'complications'. Simulations with that level of realism would be a major undertaking, and we feel that they are not within the intended scope of the current manuscript, given that the extra analyses now give much more information on the performance of our method.

Major Point 2: Error measures of timescale estimate.

Referee:

The method (Wheatley et al. 2012a) does not give timescale error measures. The estimation target of the presented method is the (layer-counted) timescale. Because (1) datasize is less than infinity, (2) noise level is larger than zero and possibly (3) some values may be missing, the estimated timescale cannot be expected to equal the true timescale. Error measures (e.g., standard error) describe the size of the deviation between truth and estimation. It should in my view be straightforward to augment the methodology presented in Section 4 of the paper with the purpose of determining such error measures.

Response:

Our method calculates a discrete probability distribution as a measure of uncertainty on the cycle count. The shape of this distribution is not necessarily close to Gaussian, so standard error may be a misleading measure of uncertainty in this case. Figure 8 in the manuscript shows examples of such plots that relate to the whole of the Gomez core, and hence to the date at the bottom of the core, since this is the single point most likely to be of interest. But such distributions could equally be calculated at any required depth; we will amend the manuscript to make this clear. In addition, our method generates several types of 'layer markings': change points, apexes, nadirs. If these correspond to a 'certain' run they have probability 1, if they are from an issue they have a probability attached. So distributions based on these features could be obtained similarly, again at any required depth.

Major Point 3: Simulated timescales.

Referee:

The method (Wheatley et al. 2012a) does not produce simulated timescales. By using the error measures from Major Point 2, it should be straightforward to generate simulated timescales (one has to take into account the serial dependence between depth points, however). Why generate simulated timescales? Please let me cite own work: "Construction of age-depth curves for climate archives on the basis of dating points, constraints (e.g. strictly monotonically increasing curves) and the physics relevant for describing the archive's growth is a challenging task. It is currently being tackled by means of Bayesian and other simulation-based approaches. Construction of these curves is, however, not a means in itself. Age-depth modelling must also provide simulated curves, which can then be fed into modern resampling methods of climate time series analysis, resulting in realistic measures of uncertainty in our knowledge about the climate" (Mudelsee et al., 2012, p. 1991, my italics here in this review). That paper from which I quote, for example, takes simulated timescales as input for a bootstrap algorithm in nonparametric regression and studies the effects of dating errors on the widths of bootstrap confidence bands.

Response:

Whilst our method does not in itself provide a way of simulating signals, there is a way to simulate timescales via the regression model fitted to the log 'certain' run lengths. This model describes their mean trend as a linear function of depth, and their local variation (the residual error) with run classification as a factor covariate. A minor adjustment to the model would be needed, taking the run starting depth as a covariate instead of the run central depth. Then the model could be applied iteratively. Firstly a starting run classification is chosen (from P, D, T, or A), this could be done at random. To generate the first run length: get the fitted value for a log run length of this classification starting at depth 1; add on some noise from the $N(0, \sigma^2)$ distribution (where σ is the residual

error); then take its exponent. We now know the starting depth and classification of the second run (as they must follow the $P\ D\ T\ A$ pattern), and so can generate its length. One could continue to generate cycles in this way until a required number or depth is reached. To incorporate the serial dependence between depth points into this scheme it should be possible to measure the correlation between log run lengths of all possible pairs of types, and take this into account when adding noise. We will add a discussion of this process in to the revised version of the paper.

Minor Point 1. p. 2477, title.

Referee:

Specify that this is a kind of pilot study on the Gomez series (e.g., via a subtitle).

Response:

Although we present a dating of the Gomez core from its H_2O_2 signal, this paper is intended to also introduce a general framework for providing an ice-core chronology with a measure of uncertainty. We have added several other analyses—see response to the anonymous referee. We have clarified which comments are specific to a given data-set.

Minor Point 2. p. 2478, l. 14.

Referee:

Write "In some cases the chemical or isotopic signals ..."

Response:

Agreed, thanks.

Minor Point 3. p. 2478, l. 17.

Referee:

Since the meaning of "robustness," which is a term coined by statistician George Box, is often misunderstood by climatologists, I would highly appreciate one extra sentence on the definition of that term.

Response:

It is perhaps not appropriate to include such technical detail here, right at the start of the paper. Perhaps we should simply change "no robust method" to "no formal, systematic method". Consideration of robustness in the statistical sense comes later in the paper.

Minor Point 4. p. 2479, l. 22.

Referee:

"... requires no prior knowledge" is a bit too strong wording. Later (e.g., p. 2484, l. 13–17) you explain nicely how expert knowledge may set in at a later stage of the analysis.

Response:

We will replace with: "...requires no prior manual assessment of the chronology, ...".

Minor Point 5. p. 2479, l. 24.

Referee:

"Depth i": I have problems here since depth in general is not an integer. It is an integer when you count it in units of bags of the ice core, but other ice-core measurements (continuous-flow analysis) exist and also other archives exist. Please re-write.

Response:

As we state in the paper (p.2485, l. 25.), we assume—primarily for ease of presentation—that the data points are equally spaced in depth, and depth can therefore be represented as an integer. This is the case for Gomez (2cm bags) and NGRIP (1mm CFA). Relaxing this assumption requires a slight extension of the notation: the ith depth rather than depth i; difference in depth rather than number of points. We should mention this earlier in the discussion paper.

Minor Point 6. p. 2480, l. 1.

Referee:

Detrending: give further details (regression model type, estimation method).

Response:

We refer to subtracting μ, as discussed in section 2.1, we will make sure the wording in the revised version is more clear.

Minor Point 7. p. 2480, l. 16–17.

Referee:

Perhaps "... moving standard deviation of x" is easier to comprehend.

Response:

We do not use a moving standard deviation of x, although that might achieve something similar. We use a moving standard deviation of $(x - \mu)$, where μ is the estimate of the mean trend in the signal and x is 'de-trended' by subtracting μ.

Minor Point 8. p. 2481, l. 6–7.

Referee:

Give more details on that algorithm used to segment the series into β subsections.

Response:

We will include more detail on this algorithm in the revised version of our paper. It works like this: firstly the signal is split into β sections using $(\beta - 1)$ boundaries that are equally spaced in depth. Then iterate:

1. calculate the expected average cycle length, l_j, for each section from their ACF, $j \in (1, 2, \ldots, \beta)$

2. redistribute the boundaries so that section j contains $[n l_j / \sum_j l_j]$ points

3. recalculate l_j for each section

4. if expected number of cycles in all sections are equal then stop, otherwise go to 2.

Minor Point 9. p. 2481, l. 11.

Referee:

Write "Fourier analysis (McGwire et al., 2011) and ice flow modelling (Shimohara et al., 2003)."

Response:

Agreed, thanks.

Minor Point 10. p. 2481, l. 12.

Referee:

Mention your core: "... and fifth (stars) section of the Gomez ice core ..."

Response:

Agreed, thanks.

Minor Point 11. p. 2481, l. 15–16.

Referee:

$6 \times 24.2 = 145.2 \neq 142$

Response:

Well spotted, thanks—this is now less of an underestimate!

Minor Point 12. p. 2482, l. 2–6.

Referee:

You should say that your method and the derived probabilities are not necessarily related to the "3/4 consensus ratio".

Response:

We deem a number of cycles as being 'certain' if they are well defined in the standardised signal with respect to our repeatable algorithm. However there is some probability that we have got this wrong, though hopefully less than $1/4$. We don't attempt to quantify the probability *at this stage*; as the referee suggests, it is not an attempt to match the '$3/4$ consensus ratio' and we should emphasize that.

Minor Point 13. p. 2482, l. 12-15.

Referee:

Omit the spaces before and after the slashes.

Response:

Agreed, thanks.

Minor Points 14-16

Referee:

. . . replace "(top)" by "Fig. 3a" and so on. . .

Response:

Agreed, thanks.

Minor Points 17-19

Referee:

. . . relates specifically to the Gomez core data. Do mention this core here. . .

Response:

Agreed, thanks.

Minor Point 20. p. 2487, l. 3.

Referee:

Expand in one short paragraph on the "standard linear modelling assumptions" as: fit method, goodness-of-fit measures, determination of predictive intervals, etc.

Response:

The model is linear regression, with independent Gaussian errors with constant variance, fitted using ordinary least-squares estimation within the `lm` function in R (citation to be added). The standard regression diagnostics and residual plots in `plot.lm` in R were all satisfactory.

Minor Point 21. p. 2487, l. 24.

Referee:

As in Minor Point 14.

Response:

Agreed, thanks.

Minor Point 22. p. 2488, l. 19–25.

Referee:

Strictly speaking, you have not done anything on $\nu > 0.8$: re-write mathematically correctly.

Response:

We have applied the method for $\nu > 0.8$ and $\nu < 0.3$: it failed for the reasons listed. We will re-word to make this clear.

Minor Point 23. p. 2489, l. 3.

Referee:

As in Minor Point 14.

Response:

Agreed, thanks.

Minor Point 24. p. 2489, l. 4.

Referee:

Write "For $\beta < 5$ and $\nu = 1/\sqrt{2}$"

Response:

Actually this is true for $\beta < 5$ generally. But the paragraph does need to be clarified. Thanks.

Minor Points 25-26

Referee:

Do not capitalize title of [references].

Response:

Agreed, thanks.

Minor Points 27-28

Referee:

Give ν and β values [in captions, Fig. 7 and Fig. 9]

Response:

In both cases, $\nu = 0.5$, $\beta = 6$.