Earth Syst. Dynam. Discuss., 5, C331–C337, 2014
www.earth-syst-dynam-discuss.net/5/C331/2014/

Earth System Dynamics
Discussions

Interactive comment on "Statistical significance of rising and oscillatory trends in global ocean and land temperature in the past 160 years" *by* L. Østvand et al.

Anonymous Referee #3

Received and published: 5 August 2014

1 Paper summary

Østvand *et al.* present a scheme for evaluating various statistical models (including those with long range autocorrelation, or memory) of the global surface temperature record. They apply their scheme to historical records of the global land surface air temperature, and the global sea surface temperature. They find that there is a significant upwards trend in both of these records, and a (barely) significant oscillatory trend for the land surface data. They suggest that the global land record may be better suited for detection of global warming and oscillating trend than the ocean record.

2 Review summary

In my view, this manuscript does not fulfil the criteria for publication in ESD, and should therefore not be considered for publication without at least a major re-write. The paper does not offer sufficiently novel conclusions, or new insight into the Earth system. The techniques used to reach those conclusions seem unnecessarily complex, to the point where it is difficult to clearly see the process that the authors used. This is not helped by a poor structure and many unnecessary digressions, particularly in the description of the methodology. The statements of the conclusions are too confident, given the analysis as presented.

If my reading is correct, the analysis method used by the authors cannot hope to provide the strong and unambiguous conclusions that they draw - in fact, it will simply give them back the information that they put in. A more standard statistical model building process would help clarify if the authors are adding genuinely novel insight.

Further, many of the conclusions that are drawn - that there was a strong upwards trend in global temperatures during the 20th Century; that temperature change in the land surface is easier to detect than in the ocean - are already very well established through other methods, or are established from basic physical principles. The main conclusion that could be described as novel - that there is a significant oscillation of 70 years in the land surface temperature record - seems inevitable given the structure of the testing methods.

As it stands, the core analysis of the paper seems circular. The authors note the appearance of a strong upward trend, and perhaps an oscillatory trend in the global surface temperature. They then use a hypothesis testing structure to see if the data could be modelled with a long range memory (LRM) process. They fit a LRM model that could generate this data, but then reject that, in favour of a model with prescribed upwards and oscillatory trends that they noted previously. In addition, the authors make a classic mistake in assuming that the rejection of a null hypothesis is good evidence for

the alternative hypothesis. In this case the alternative hypothesis is arbitrarily defined.

At the core of this paper is an attempt to distinguish between a long-memory process, and some other kind of trend. I suggest that this information is not recoverable - certainly using the techniques described - from the time series of global surface temperature alone without reference to the physical processes involved. As such, we learn little or nothing about the Earth system itself. A supplementary approach to considering physical understanding might be to use a standard statistical model selection procedure that penalised complexity, instead of simply prescribing a model structure with little or no justification.

Here, the authors have ignored the vast majority of a substantial literature that concerns itself with the detection and attribution of climatic change. New and fresh ideas and approaches are welcome, but they must be well thought through and executed in order to be convincing. I am not convinced by this manuscript.

I suggest that if the authors wish to resubmit that they simplify the text significantly, focusing on the core idea of detecting trends in the presence of long range memory processes. A good way to demonstrate the efficacy of the methods presented here would be a blind test, where the time series might be a draw from a long memory process, or a simulated record from a climate or perhaps even "toy" model. If the authors were able to demonstrate the effectiveness of their method in this sort of test, I suggest that ESD should consider publishing the paper.

3 Review details

Copy reviewed http://www.earth-syst-dynam-discuss.net/5/327/2014/esdd-5-327-2014.pdf

P328 L13 *"The parameters of the null models are estimated from the instrumental*

record, but are also checked to be consistent with a Northern Hemisphere temperature reconstruction prior to 1750 for which an anthropogenic trend is negligible."

I cannot see further reference to this statement (from the abstract) in the rest of the text. It seems an important part of the analysis - was it conducted? If so, how?

P328 L23. *"...the conceptually simplest approach to detection of anthropogenic global warming should be the estimation of trends in global surface temperature throughout the instrumental observation era..."*

In order to detect anthropogenic global warming, we must be able to reliably distinguish between natural and anthropogenic variability. The statement above is (correctly) contradicted in the text almost immediately afterwards.

P329 L5. *"At the core of the global change debate is how to distinguish anthropogenically forced warming from natural variability"*

This is no longer at the core of the global change debate. An entire literature of detection and attribution is outlined in the IPCC AR5 chapter 10, and gains only a cursory mention in the conclusions section of this paper.

P329 L28. *"In this case we will first construct a parametrized model for the trend based on the appearance of the climate record at hand and our physical insight about the forcing and the nature of the dynamics."*

I think that this is a good idea, and in fact necessary, but can find little evidence that physical insight about the forcing and the nature of the dynamics of the Earth system were used in this paper.

P330 L12. *"rejection of the null hypothesis will be taken as an acceptance of the hypothesis that the estimated trend is significant, and will strengthen our confidence that these trends represent identifiable dynamical features of the climate system."*

It may strengthen our confidence, but this paper appears to make the classic "error

of the transposed conditional" (e.g. Ambaum 2010), assuming that rejection of the null hypothesis is strong evidence for acceptance of the alternative hypothesis. This leads to the paper making much stronger statements about the Earth system than are justified by the analysis.

Ambaum, Maarten H. P., 2010: Significance Tests in Climate Science. J. Climate, 23, 5927–5932. doi: http://dx.doi.org/10.1175/2010JCLI3746.1

P332 L3. *"The construction of forcing time series relies heavily on uncertain observations and modeling, hence there is an obvious case for complementary approaches to trend estimation that do not rely on this kind of information. This is the approach that will be explored in the present paper."*

Modelling and observations are both uncertain, but good approaches take this uncertainty into account. Ignoring the information from these approaches (along with more basic physical understanding) does not seem justified.

P332 L13. *"Emanating from these studies is the recognition that ocean temperature is more persistent than land temperature and that the 20th century rising trend is stronger for land than for ocean"*

This recognition does not emanate from the studies listed, but is well established from a basic physical understanding of the system.

P334 L14. States HADCRUT3 used. This data set has been superseded by HADCRUT4.

Morice, C.P., Kennedy, J.J., Rayner, N.A. and Jones, P.D., 2012: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 dataset. Journal of Geophysical Research, 117, D08101, doi:10.1029/2011JD017187

In addition, this highlights the fact that the authors have not addressed observational uncertainty in their analyses, despite highlighting it as important earlier in the

manuscript.

P334 L22. At this point, the paper goes on a long digression on long range memory models before it has stated the most basic aspects of the methodology approach. This level of discussion may be better suited to supplementary material, as it breaks the flow of the description.

P337 L27. The manuscript never states whether the Bayesian or Frequentist approach is taken.

P337 L9. It would clarify matters considerably if it were pointed out that the trend estimate model is simply an extension of the stochastic model. In this case, the hypothesis test takes on the structure "are the coefficients (A) different significantly from zero".

P339 L5. It would surely be simpler to label δ as A_0?

P339 L12. *"We rather treat f as a fixed quantity which is an inherent part of the alternative hypothesis. In practice we select f from a least-square fit of the trend model to the observed record"*

It is this that leads to the circular argument. With no reference to the possible physical processes involved, this means that the hypothesis test simply gets out what was put in. The fact that an oscillation was observed lead to the situation where an oscillation was being tested for.

P341 L2 *"Another is that the land record analysis establishes beyond doubt that there is a significant global linear trend throughout the last century, and that the reality of an oscillatory trend is probable, but not beyond the 95% confidence limit"*

This kind of statement is far too strong, given the lack of reference to physical mechanism through the manuscript. A rejected null does not mean that the alternative is true.

P341 L17. *"In a Bayesian spirit. . ."*

I struggle to see how assuming something as an established fact and working out the consequences is in a Bayesian spirit.

P346 L27 *"The validity of the method depends, of course, on the assumption that the climate model correctly describes the relevant aspects of the pattern of natural variability, e.g. the long-range correlation structure in space and time."*

This is not quite correct. The validity of the methods depends on the fact that there is some information given by the climate models about the true system. The models do not have to be "correct" (no model ever is), merely useful. Ignoring information given by these models - along with underlying physical principles - is not a viable alternative.

Interactive comment on Earth Syst. Dynam. Discuss., 5, 327, 2014.