

Rev #1

The paper makes an important contribution in specifying how to estimate statistics in the presence of long-range dependence, where classical estimates of many statistics are erroneous. I am surprised that this has not been done before; I have not been following the hydrological or climatological statistical literature over the last 10 years, so I cannot say for certain it has not. If it truly hasn't, then this paper is long overdue and is a very important contribution to the body of hydrological knowledge. It is eminently suitable for Water Resources Research, and overall is very good.

#000679

The paper is technically sound and while I have not verified the derivation of the statistical results they appear to be sensible. The methods are described in sufficient detail. The empirical functions on page 17 are a little clumsy, and alternatives are suggested in the comments.

The paper is well organised, easy to read and grammatical. The figures are generally well executed (exceptions noted in the comments) and support the arguments of the paper. There is no particular reason to shorten the paper. The abstract accurately reflects the contents of the paper.

My only serious criticism of the paper is the over-emphasis on the distinction between stochastic and deterministic processes (mainly in the Introduction, pages 10 and 11) and a repeated tendency to overstate the implications of the analysis – the paper shows the data are consistent with scaling processes but not (as the author claims) that the processes behind the data must therefore be a scaling process rather than a deterministic one with varying means. As such, some of the statements in the conclusions and the abstract should be modified along the lines of the measurements being consistent with a scaling model, rather than being conclusively proved to be due to a scaling process.

Specific comments to the author

I was surprised to discover that this has not been done before. Is there really no previous literature describing statistics for simple scaling processes?

Page 10. I find the discussion here concerning determinism rather unproductive, and does not provide a sufficiently strong introduction to the good work that follows. I do agree with your statement at the end of that page that the separation of chaotic signals such as climatic data into deterministic trends and random fluctuations may not be the best approach. Mandelbrot talked about the lack of distinction between signal and noise in fractal processes (can't remember the exact reference).

However, the argument you present does not lead to this conclusion.

Your argument appears to be:

- the diagnostic character of determinism is that it is predictable
- the trends were identified *a posteriori*, not predicted
- therefore they are not deterministic

For this logic to hold, you need to know not just that the trends were not predicted but that the trends *could not have been* predicted. The "could not have been" condition might be made dependent on the available knowledge (the pragmatic view of determinism), in which case the apparent trends may have been unpredictable at the beginning of the century but predictable using current knowledge, making them deterministic from our current viewpoint. More generously, you could allow that we might yet gain the knowledge to predict these trends, so that they should be considered non-deterministic now but in principle deterministic. But in any case, what does this argument gain?

To support the contention that the "trend plus noise" view of climatic data is misleading, you might be better off demonstrating that there are trends at multiple scales, so trends at one scale seem to be part of the random fluctuations at a broader scale. Or maybe show that "trends" come and go without any apparent cause. From a more theoretical base, you could use the behaviour of a dynamical system that appears to have trends and noise that both arise from a single process.

Page 11, line 1: where you say "the large-scale trends in the time series are closely related to the well-known Hurst phenomenon", mightn't it be better to say they "are a manifestation of"? Isn't an increasing variability with increasing duration exactly what the Hurst phenomenon is?

Page 11, second paragraph: Again I think you are making too much of the distinction between deterministic and stochastic processes. Isn't it true that your results hold in any

case wherever the data behave as an SSS? Putting such emphasis on the philosophical underpinnings is, I believe, more likely to result in semantic disputes rather than ready acceptance of your statistical results.

Page 15, after equation 13. You state without proof that equation 13 is an unbiased estimator regardless of the type of the process. This seems fairly obvious to me, but not absolutely transparent. Can this be demonstrated, or supported by a reference? Or is this so fundamental (e.g., the definition of the mean) that it doesn't need any support?

Page 17, after equation 21. Why is S an "approximately unbiased estimator"? Where does the approximate status come from, if S^2 is unbiased?

Page 17, equations 24 and 25. The behaviour of $\kappa(H)$ and $\lambda(H)$ might be better expressed as single functions rather than as piecewise functions. $\kappa(H)$ has a small discontinuity at $H = 0.6$ ($\kappa(0.6+\epsilon)$ is not the same as $\kappa(0.6-\epsilon)$) and $\lambda(H)$ has an abrupt change in slope. Using your functions, I was able to come up with:

$$\kappa(H) = \frac{1.4 - H}{(1 - H)^{1.3}}$$

$$\lambda(H) = \frac{1.02(1 - H)}{(1.02 - H)(1 + H^6)}$$

which are both continuous functions that are reasonably close to your functions. If you chose to adopt this type of function you would need to adjust the parameters to retain the identity with the classical formula for $H = 0.5$.

Page 23, first paragraph: Figure 7 indeed shows that the true probability distribution lies within the 95% confidence limits of the SSS estimate, but there is a substantial and consistent bias across the distribution. Is this simply because of the effect of using only one sample, and the estimates being uncertain due to the nature of the scaling process, or is there something systematic about it? In other words, if you did the analysis for a large number of synthetic samples would the average estimated distribution converge to the theoretical one or does it remain biased?

Page 26, end of second paragraph: Claiming that "the empirical autocorrelation function agrees *perfectly* with the model" is a bit strong; something like "fits well" would be more suitable. In fact, I wonder if the irregularities in the autocorrelation plots of Figure 11 (down) actually indicate significant departure from simple scaling. There is an unexplained peak at a lag of about 30, autocorrelation is lower than the expected at around 100, and then increases from about 0.3 to 0.4 from lag 100 to 200 where the model shows a decrease from 0.38 to 0.33. Is that a significant departure? You don't need to answer that, but you should be more critical of your own results.

Page 27, lines 3-5: the temperature anomalies quoted here (e.g. 99% quantile of annual temperature anomaly is about 0.6°C) don't seem to agree with the values in Figure 12.

review#1
#000679

There appears to be a bias of about -0.3°C between your numbers and those of Figure 12. For that matter, shouldn't the temperature anomaly at probability of 0.5 be 0? Is this bias the same problem seen in Figure 7? Is it just a plotting problem? If not, can it be fixed?

Pages 29-31, Conclusions. I agree with your overall conclusions and specifically with the conclusion that statistical analysis of time series must take account of long range variability associated with the Hurst phenomenon. However, I am not sure that you have proved that the time series you explore are results of SSS processes; they are certainly consistent with this hypothesis but that's not the same as identifying the cause of the variability. The same arguments have been made in soil science and in the study of land surfaces: the data show scaling properties, so soils (or landscapes) are fractal, and are the result of a stochastic scaling process. But this does not follow: the same results can be obtained by a variety of non-scaling processes operating at different scales.

Some of the strong statements in the conclusions should therefore be modified. In particular:

- Page 30, line 11. The statement that trends or jumps should not be removed "as the shifts are *in fact* stochastic rather than deterministic" (my emphasis) is not defensible (see earlier comments on the focus on determinism). You have shown (as have others) that the SSS hypothesis is plausible for these time series, but not that they are the result of a scaling process. It could be that there are *in fact* shifts and trends due to specific processes that come and go.
- Page 31, line 10. Change "agree perfectly with" to "are consistent with".
- Page 31, lines 11-14. The claim that the trends or shifts "are nothing more than regular behaviour" should be amended to something like "are consistent with the scaling hypothesis".

The same issues are evident in the Abstract: the statement that "changes of the climate on all scales ... is nothing more than a simple scaling behaviour" is overly strong. Those changes can be described by a scaling model, but that doesn't mean the processes underlying the behaviour are scaling processes.

In summary, what you have shown is that there is a viable alternative explanation for the shifts or trends in the observed data that does not require an explanation of changing trends. In some ways, the hypothesis of a scaling process is simpler than that of shifts or trends, but in some ways it explains nothing. What lies behind these Hurst phenomena? A statistical model is not an explanation but a description. But the point that the statistical models should include the effect of long-term dependence is well made: the statistical description is better when these effects are accounted for.

#000679
review#1

Minor comments

Abstract, line 3: Change “leaded” to “led”

Page 9, line 13: local *overyear* average – should that be *multi-year*?

Page 12, line 2: references should be in italics as elsewhere.

Page 18, 3rd last line: change “0.067 is 0.043” to “0.067 and 0.043”.

References: The Hirsch et al 1993 and Salas 1993 references should include a chapter number within the book.

Paper: WRR 000679

Climate change, Hurst phenomenon, and hydrologic statistics

By Koutsoyiannis, D.

review^{#12}

#00679

I do not find this paper acceptable in its present form. The paper may be acceptable after major revisions.

I would be willing to review the manuscript again if resubmitted.

I do not agree to be acknowledged by name.

Very shortly, the papers contains some contributions related to statistical estimation of simple scaling stochastic processes. I find the motivation for the application of this kind of processes to hydrologic or climatic series very confusing, lacking rigor and ignoring important contributions in other fields. Moreover, in the paper there is not an open minded approach to the problem of what can one learn and understand about climate change and climate variability from the analysis of recorded time series. The paper is so weak in this respect that I recommend to take a more simple and pragmatic approach to motivate the presentation of the paper's contributions. It is enough to quote references that have used simple scaling in climate and hydrology, state the definition and to go directly to the present section 3 and the case studies.

From the physical side, we know that some driving processes of climate and hydrology have known periods or time scales (astronomical variations in solar radiation due to earth orbit parameters, solar variability). We understand some of the mechanisms of climate variability: ice-albedo feedback, CO₂ cycles and green house effects in general, ocean deep water circulation, ocean-atmosphere interactions, land-atmosphere interactions, etc. The dynamics of each one points clearly to dominant time scales. Some of those processes are not independent and interact, for instance ENSO (2 to 6 years time scale) is coupled with the annual cycle. Statistical analysis of climatic time series can ignore this (partial) knowledge? For instance, Thompson, 1994, shows how in the temperature time series one can observe clearly earth orbit precession and anthropogenic effects.

The paper ignores alternative explanations for the observed statistical behavior in the time series: Large but finite correlation length or scale of fluctuation (Mesa and Poveda, 1993); Composite random processes with components with significantly different scales of fluctuation (Vanmarcke, 1988, pag. 225); Power law trends (Bathacharya et al, 1983); and probably many more. Even in the case the paper will be concentrated in the self similar model, it should mention other alternative models. Most urgently, there is a need for tools that could discriminate among the competing alternative models. The paper does not provide any contribution in this regard.

Summarizing, the paper should be reduced to a technical showing only the statistical techniques, and briefly one application.

References

review^{tz}
#000679

Vanmarcke, E. Random Fields: Analysis and Synthesis. The Mit Press, Cambridge, 1988.

Mesa, O. J. and G. Poveda. The Hurst Effect: The Scale of Fluctuation Approach. Water Resources Research. Vol 29, NO 12:3995-4002, December, 1993.

Bathacharya, R. N., Gupta, V. J. and Waymire, E. The Hurst effect under trends. Journal of Applied Probability. 20, pp. 649-662, 1983

Rest of the evaluation follows by fax and regular mail