Two comments on “How Red are my Proxies?” by David Ritson on *Real Climate*

Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, dk@itia.ntua.gr - http://www.itia.ntua.gr/dk

In his commentary article and his treatise “Deriving AR1 Autocorrelation Coefficients from Tree-Ring Data” (http://www.realclimate.org/supp/nred.pdf), David Ritson explains that in his tree ring analysis he decomposed the data series into a Markovian noise and a deterministic fluctuating signal component with comparatively large excursions over multi-decadal periods. In these two comments it is maintained that this methodology is fundamentally flawed. Also, more consistent stochastic methodologies are discussed.

Posted in http://www.realclimate.org/index.php/archives/2006/05/how-red-are-my-proxies/, May 2006

Comment 1*

Dear Professor Ritson,

With all respect to your work and with the risk of some misunderstanding because of my different scientific origin (engineering hydrology), I would like to make a few comments on your treatise “Deriving AR1 Autocorrelation Coefficients from Tree-Ring Data” that you link in your above post and seems to be the background document for the post.

1. In my opinion it is useful that the author of a scientific text underlines the hypotheses which he/she uses to derive the results -- and not leave the reader to guess them.

2. Apparently you use the hypothesis of stationarity and ergodicity – the latter is obvious from your notation, which is in terms of time averages rather than ensemble averages. Of course these are not strong hypotheses and everybody uses them; however, one must have always in mind that ergodicity has an asymptotic character (e.g. a stationary process is mean-ergodic if

* Comment #34 in http://www.realclimate.org/index.php/archives/2006/05/how-red-are-my-proxies/

its time average tends to the ensemble average as time tends to infinity; Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, 1991, p. 428).

3. You also use the hypothesis that the process $X(j)$ (representing the growth-amplitudes) is an AR(1) process (please note that I have dropped your subscript i to simplify notation). Even though you put this hypothesis for a component of $X(j)$ that you call “noise amplitude”, it becomes also the case for the initial (decomposed) process $X(j)$, given that you assume equality in time for what you call “slow component”. Such a hypothesis is a strong yet unjustified one; in my opinion there is no reason that nature’s signals should be AR(1).

4. Loosely speaking, the AR(1) hypothesis is equivalent with an hypothesis that a single time scale (e.g. the annual) dominates in nature. But I am glad to see in your paper the recognition of a “signal component with comparatively large excursions over multi-decadal periods”. So I agree with you that, in addition to fluctuations on the annual scale, there exist fluctuations on over-annual scales. In the case that we follow a multi-scale thinking, a maximum entropy consideration will result in a non Markovian (non AR(1)) dependence, and most probably in a process with long-range dependence (LRD) or long term persistence (LTP). This I tried to show in Koutsoyiannis (2005).

5. Even with a simpler thinking, just with the superposition of fluctuations on three time scales, e.g. annual, decadal and centennial, one arrives at a process that is virtually equivalent (meaning for lags as high as 1000 years) to a process with LRD. This I demonstrated in Koutsoyiannis (2002).

6. From a more philosophical – if you allow me to say – standpoint, viewing complex natural phenomena as AR(1) processes, which means Markovian processes, may be too simplified. Recall from the theory of stochastic processes that a Markovian process is by definition "a stochastic process whose past has no influence on the future if its present is specified" (Papoulis, ibid., p. 635). Thus, for me it is very difficult to imagine that only the present state of a complex natural system matters for its future and that we can drop our knowledge of its past. On the other hand, compared to a time independent (like head/tail outcomes in coin tossing) view of natural processes, in which even the present does not matter for the future, certainly a Markovian view is a progress.

7. I was able to verify your main result in your treatise that $\alpha = 1 + 2\,\rho$, where α is the lag one autocorrelation of the process $X(j)$ and ρ is the lag one autocorrelation of the process $Y(j)$ as you define it. (Here I have used the notational convenience ρ for your fraction in your penultimate equation – I hope that my understanding is correct that this is lag one autocorrelation). In my opinion there is no need to do – as you did – a decomposition of the

process $X(t)$ into a slow component and a noise amplitude. I think that such a decomposition is fuzzy, subjective, and not necessary because you can obtain your result without any decomposition (and without your accompanying assumption $s(j) = s(j + n)$, which may not be justified). If one simply defines $Y(j) := X(j) - X(j + 1)$ (i.e. in terms of the actual process rather than the decomposed one) and also assumes an AR(1) autocorrelation function, one directly obtains $\alpha = 1 + 2\rho$.

8. However as I wrote above, the AR(1) hypothesis is a strong one and it would be better to avoid it. In this case, one can easily obtain that your relation $\alpha = 1 + 2\rho$ (equivalently $\rho = -(1 - \alpha)/2$) becomes $\rho = -(1 - 2\alpha + 2\alpha_2)/(2 - 2\alpha)$, where α_2 is the lag two autocorrelation of the process $X(j)$. Your formula is a special case of the general one, obtained by substituting $\alpha_2 = \alpha^2$ (i.e. assuming a Markovian process). Given that α_2 is unknown in an approach such as yours, we cannot estimate α from ρ. But we can estimate its upper and lower bounds. Assuming stationarity, we can put the restriction that the size 3 autocorrelation matrix of the process $X(j)$ is positive definite. In this case, a positive determinant results in the constraint $-1 + 2\alpha^2 \leq \alpha_2 \leq 1$. From this constraint, using simple algebra, we can find an interval for α given the value of ρ.

9. You may say that this interval of α is too wide and thus not helpful in an accurate point estimation of α. Well, this is the optimistic view. The interval for α would be that wide if we knew precisely the value of the lag one autocorrelation ρ. But we only have a sample estimate of ρ – thus the range of alpha is even wider. More specifically, in Koutsoyiannis (2003), I have demonstrated that the classic estimator of autocorrelation (that you use) implies high bias if the process exhibits LRD. You may also find there citations pointing that bias exists also in the AR(1) process. In addition to bias, there also exists significant variability and thus uncertainty in estimates. Therefore one should be very careful in such statistical calculations, because they entail bias and uncertainty -- in contrast to typical arithmetic calculations.

10. Having some experience with statistical uncertainties and particularly with complex interactions of uncertainties (and the magnification of the total uncertainty) when ones combines two or more random variables in a single expression, personally I would avoid calculating statistics of a process X(j) based on the differenced process Y(j) = X(j) - X(j +1) (or, much worse, on a process involving differences of some subjectively defined components of X(j) as you did). You can check the magnification of uncertainty even with arithmetic calculations, assuming for instance a pair of values X(j) and X(j +1) close to each other and attributing a certain percentage of uncertainty in each of the two. In this respect, I would prefer to base my estimations on the process X(j) per se and in addition to be as aware and careful as possible of the uncertainty and bias in statistical estimations, especially for

processes which might exhibit LRD -- a case not well covered so far in classical statistical texts.

Comment 2†

Dear Professor Ritson,

Thanks for your kind reception of my comments and your responses. Based on your responses, I have the feeling that we can converge, at least partially. Therefore, I will put my emphasis not to some different views that I may have for some of your points, but to points that I feel we can converge.

I am happy that you "assume along with everybody else that the climate signal contains regions of warmth Medieval warm period and Little ice-age for example." This could be a good point for convergence. I also agree with your statement that "These may result deterministically from externals such as solar forcing." But I wish to discuss it further and first your term "deterministically". I hope you could agree with me that a specific storm that causes severe damages results deterministically from some atmospheric dynamics. This dynamics is in fact the basis of the meteorological prediction of the storm, cast some days earlier. At the same time, nobody would accuse meteorologists for not having predicted the storm a year or a century earlier. Because of the complexity and chaotic behaviour, we all recognize that it may be impossible to accomplish such a long-term prediction. Therefore, in engineering, given that we have to design works that will last say a century, we use a probabilistic or stochastic approach to describe storms and to construct what we call "design storm", a hypothetical severe storm that has some pre-specified probability of occurrence.

We could expand this logic to other simpler phenomena, e.g. the movement of a die. There is some deterministic dynamics in this movement; however we all say that the outcome of the die is random (cf. Einstein's apothegm "God does not play dice").

After this, I hope you will agree with me that the Medieval warm period and the Little ice-age are not MORE deterministic than the evolution of mean daily temperature or the mean annual temperature. So, if I have the right to use a stochastic description for the annual temperature, as you did with your proxies, I feel that I have the right to use a stochastic description for over-annual fluctuations or excursions such as the Medieval warm period and the Little ice-age. Of course you may disagree with me. You may say that these excursions should be

---

†Comment #36 in http://www.realclimate.org/index.php/archives/2006/05/how-red-are-my-proxies/

modelled not stochastically but only deterministically. In this case I will ask you: Could you give me your deterministic dynamics for the variations of solar activity and their impacts to the atmosphere and particularly the global average temperature? Could you apply your deterministic dynamics for the past and hindcast the climate over the last 2000 years? Could you apply your deterministic dynamics for the future and forecast the climate over the next 2000 years? In these questions I deliberately used long periods because we need long periods to observe such long-term fluctuations.

As you see, by profession, I do not have any problem to use stochastic descriptions of natural phenomena. In fact I am very satisfied with the answers I am getting from my stochastic descriptions for engineering designs and for supporting water management decisions. But in fact, in hydrology we follow the paradigm of physics. In my knowledge and view, in the late 19th century, physicists abandoned the mechanistic paradigm and were thus able to develop disciplines such as statistical thermophysics (including the entropy concept, first put on probabilistic grounds by Boltzmann) and quantum physics. In both these disciplines probability has a major role and replaces mechanistic concepts, explanations and analogues (e.g. Lavoisier's subtle caloric fluid).

If we accept that one is allowed to use stochastic descriptions the question is: Which stochastic description can be appropriate for hydroclimatic processes, i.e. reproduce the Medieval warm period and Little ice-age, and the persistent droughts and floods of Nile? (I mentioned Nile because we have a lot of information covering many centuries – obviously such behaviours have been observed in other rivers, as well). A Markovian (AR(1)) description? I would say no. I have played a lot with several stochastic models and I think the simplest is a scaling model (also known as fractional Gaussian noise and with many other names – see my post in http://landshape.org/enm/?p=25).

Please allow me to say that a simple scaling stochastic model is not a complex description, as you characterize it in your first response above. It is a very simple description, in some aspects simpler than Markovian. And it has a very simple interpretation: combine fluctuations or excursions on several times scales, and you get a scaling process. Amazingly, the resultant scaling process, by combining different initial components, is simpler than the components. But this may not be a surprise or a unique phenomenon: Combine several weird distribution functions by taking the sum of the different random variables. You get the extremely simple normal distribution – the central limit theorem (notably the normal distribution results also from the maximum entropy principle, regardless of the central limit theorem).

Having said these, it’s a marvel to me that climatologists have been so strongly reluctant to adopt the scaling description for climatic processes. I marvel to read statements that long-term persistence is not “...a proper recognition of the physics and dynamics underlying the [climatic] systems...” and that “… a simple model [an AR(1) model] for climatic noise has the advantages that it is (i) motivated by the actual underlying physics (see e.g. Hasselmann, 1976…" (the quotations are from a review – apparently by a climatologist – that I received recently). A description that cannot reproduce important phenomena, such as the Medieval warm period, the Little ice-age and the persistent multiyear droughts and floods, has been regarded as consistent with the “actual underlying physics”!. At the same time, a description that can reproduce them, lacks “a proper recognition of the physics and dynamics”!

I apologize if I have been verbose. I must stop here saying that my thoughts on these issues are presented in more detail in (just published) Koutsoyiannis (2006).

References

Koutsoyiannis, D., The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal*, 47(4), 573-595, 2002.

Koutsoyiannis, D., Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48(1), 3-24, 2003

Koutsoyiannis, D., Uncertainty, entropy, scaling and hydrological stochastics, 2, Time dependence of hydrological processes and time scaling, *Hydrological Sciences Journal*, 50(3), 405-426, 2005.

Koutsoyiannis, D., Nonstationarity versus scaling in hydrology, *Journal of Hydrology*, 324, 239-254, 2006.