Statistical analysis of climatic time series: uncertainty and insights

(Former title: Long-term persistence and uncertainty on the long term)

by Demetris Koutsoyiannis and Alberto Montanari

Reply to the referees' comments

First of all we wish to thank the Editor and the referees for the prompt review process. Even if the outcome of the review was not positive, we sincerely appreciate the effort of the referees and respect their opinions. Particularly, we are grateful to Referee #2 for evaluating our paper as "superb".

The remarks of the referees stimulated a deep and long discussion between us which led us to prepare a major revision of the paper. The structure of the work and the presentation of the analysis are significantly changed and possibly improved. The feeling that the major revision of the paper may have resulted in improvement of the study and that several negative comments by Referee #1 may have been replied in this revision led us to consider the possibility to re-submit the revised manuscript (RM). Before listing point by point how each of the remarks provided by the referees was addressed, we would like to summarize here below the main changes that were applied to the original manuscript (OM) and our main objections on the strong review comments of Referee #1.

a) We have the feeling that the critical attitude of Referee #1 was mainly addressed to previous (already published) studies and only in a minimal part to our specific findings. This equivocal situation was certainly induced by lack of clarity of our presentation, but nevertheless we had the feeling to read a review of what has previously been done by others and not addressed to our own manuscript. We quote here two examples from his points that caused this feeling, along with our replies:

   Quote 1 by Referee #1

   "… the claims of existence of LTP in this and related previous studies…"

   Our reply – quote from OM

   "…even the presence of LTP can be disputable on purely statistical grounds…"

   Quote 2 by Referee #1

   "The obvious test that the authors (and other researchers attempting to attribute low-frequency climate variability to "LTP" or related phenomena) must pass is to be able to demonstrate that the LTP analysis procedure works correctly when applied to a realistic synthetic example where the answer is known a priori not to be 'LTP'. In other words, the procedure has to be demonstrated not to yield unacceptably high 'false positives'.

   Our reply – quote from OM

   "Besides we generated with this process [an ARMA process known a priori not to be 'LTP'] *a synthetic series with sample size 2000, and all estimation methods we tried gave incorrect values of H in the order 0.79-0.93. Continuing this experiment, we also found that we need a series with length of about 20 000 to correctly estimate H, viz. to find a value around 0.50* [the correct value, as ARMA is known a priori not to be 'LTP']. *These examples clearly point out even the distinction between the extreme cases H = 0.5 and H → 1 is not statistically decidable with typical sample sizes."*

b) Another critical point made by Referee #1 is his contrast of LTP vs. long-term deterministic variability.

In fact, we do not exclude in any respect the deterministic nature of the temperature signal. On the contrary, we think that such a deterministic nature does not only apply to long-term variability but also short-term variability, in other words to any timescale. The fact that we use statistical or stochastic descriptions should not be confused with the deterministic, causal nature of the signal. Stochastic descriptions do not generate the uncertainty, they simply describe it; what generates the uncertainty is the enormously complex nonlinear dynamics of the system (as is now well known from the impressive results of chaos literature). The requirement for stochastic descriptions is to describe the uncertainty as faithfully as possible, thus providing a useful insight of the system behavior.

We hope the RM, which was extensively rewritten in the abstract, introduction, explanation of the analysis, and conclusions, now better conveys the motivations and goals of our work.

c) Another fundamental point made by Referee #1 is his contrast of “physical” vs. “stochastic” aspects.

The fact that a certain quantity or property is defined on probabilistic, statistical, or stochastic grounds, does not mean that it does not have a physical meaning. Take for instance the standard deviation σ, which is the quantity used almost exclusively in our paper (equation (2) and beyond). To make things simpler, think of the same quantity in a simpler thermodynamic system, for instance a volume of gas (in an unmoving isolated container) described in terms of momenta of its particles. The quantity σ is defined on probabilistic grounds but, simultaneously, it is well known that σ^2 is a measure of the (macroscopic) energy of the system. Not only does σ^2 have a physical meaning but it is also an invariant quantity of the system (due to preservation of energy) and thus it can be alternatively viewed as a deterministic quantity.

Another example of this type, which is a quotation from Karl Popper, we give in the beginning of the introduction of the RM.

Now LTP is a property defined on probabilistic/stochastic grounds, too. This obviously, implies that a stochastic model may or may not have this property. But simultaneously LTP is verified in real world data too. This means that LTP **can result from deterministic dynamics and therefore it is fully compatible with deterministic dynamics**. For example (as also described in the RM), it has been demonstrated that a simple nonlinear model with only two degrees of freedom can produce trajectories exhibiting LTP (Koutsoyiannis, 2006). This model has nothing stochastic or random as it is totally deterministic. Thus, it is possible that a certain deterministic climatic model may also produce trajectories exhibiting LTP.

For this reason, a statistical analysis can be carried out on any kind of data and helps us to derive indications about the nature and behavior of the process. If this analysis reveals the existence of LTP, then this may have some consequences, which we try to explore in the paper. Among these is the simple fact that statistical procedures not admitting LTP are not appropriate for analyzing time series that exhibit LTP. In addition, it may raise some requirements for deterministic models that attempt to reproduce the natural process – but this issue is not in the scope of our paper.

In view of the considerations above, we respectfully disagree with the referee on his point to “clean” the spectrum of temperature series by removing low frequency behaviors accordingly to climatic hypotheses. Of course, this modifies the correlation structure of the process (e.g., it can be fit with an AR(1) model) and removes LTP. But this does not assure us that the hypothesis that was used in order to identify the low frequency behavior is correct. (Note that this reasoning does not apply to the trivial example of the seasonal cycle, whose treatment can be agreed almost unanimously).

In conclusion, **LTP is not an alternative to the presence of a deterministic signal**. LTP is a behavior of the process.

We strongly believe that the analysis of climatic records should be carried out by combining statistical approaches with physical models. Statistical tools are helpful in order to identify macroscopic behaviors of the process; physical models should provide detailed physical explanation for these behaviors. This is the way in which statistical analysis is historically carried out in fields like hydrology, biology and medicine. (For instance, statistical analysis of the incidence of a type of cancer helps in formulating hypotheses about the possible causes for it. These possible causes are subsequently further analyzed in order to identify the connections between them and the cancer itself). Neglecting the role of statistical tools in the analysis of climatic signals would prevent us to benefit from valuable research opportunities in this field.

In this view, it is extremely important to be fully aware of the uncertainty involved in the statistical analysis. This consideration motivated our study.

We clarified these concepts in the introduction of the RM. In particular, we included the following sentences:

"LTP is a behavior defined on statistical grounds (see equation 2 below) and can be easily reproduced by appropriate stochastic models. However, this does not mean that LTP implies necessarily stochastic dynamics. For instance, it has been demonstrated that a simple deterministic nonlinear model (involving no random component) can produce trajectories exhibiting LTP (Koutsoyiannis, 2006). From a practical point of view, LTP indicates that the process is consistent with the presence of fluctuations on a range of timescales, which may reflect the long term variability of several factors such as solar forcing, volcanic activity and so forth. LTP can be also conceptualized as a tendency of clustering in time of similar events (droughts, floods, etc)."

d) As we mentioned above, we believe that the presentation of our work in the OM perhaps led Referee #1 to misunderstand the essence of our paper. Our work aims to prove that a relevant uncertainty is associated with the statistical analysis of climatologic records. We wish to demonstrate that (1) in view of our findings, the conclusions of previous studies could probably be revisited and (2) a better insight into the process is needed in order to correctly quantify uncertainty. We do not aim to prove once again the presence of LTP in temperature records (and we do not focus on the detection/attribution problem, see the introduction of OM and RM). We only wish to demonstrate that if LTP is a plausible hypothesis (as we know from the outcomes of many previous studies for many climatic records), then the hypothesis testing becomes less and less reliable with respect to classical statistics. Up to our knowledge, the concepts and proofs presented in our paper are new and may prove very useful in order to better understand the value of previous and future studies of climate dynamics that use statistical approaches.

Simultaneously, we are trying to show the limitations of the statistical analysis in this context and therefore we may provide support to the general opinion of Referee #1 (see e.g. our Conclusions).

In order to better explain the essence of our study, we changed the title and completely rephrased the abstract, introduction and parts of the analyses and conclusions.

e) Our work is inserted on the field of the statistical analysis of climatic records. In this respect, we would like to point out some potential features (and possible limitations) of previous studies recently published by GRL. This does not mean that we neglect the value of physically based climatic models. We do not want to consider this side, which is another issue. However, we believe that a work that contributes to point out the features of previous studies could be worth publishing even if (maybe "especially if") one does not agree with the essence of the previous studies themselves.

The statistical methods we consider for the analysis of climatic time series are meant to be an expedient tool (and therefore not an antagonistic strategy) to the development of physically based climatic models, based on the dynamics of long timescale response components of the system (e.g. oceans, cryosphere, land surface processes and biosphere). Our aim is to provide a further insight about the use and usefulness of statistical methods for assessing the features of climatic data. As mentioned before, following the examples provided by researchers in other fields (such as hydrology, biology, medicine and many others) we believe that the knowledge of the statistical behaviors of a process (like variability and dependence structure) may provide a useful support for inferring the dynamics of the underlying physical processes.

In the RM we tried to make clear this point in the introduction. We hope that now the focus of our paper is clearer.

Here below we detail how each of the referees' remarks was addressed.

1. Reply to Referee#1

> I was disappointed by this manuscript. Like many papers attempting to attribute "long-term persistence" ("LTP") to climate time series, the authors naively apply statistical tools to data series without a proper recognition of the physics and dynamics underlying the systems which these series represent. A more circumspect analysis would first analyze the behavior of appropriate synthetic time series whose properties are *a priori* known (e.g. output from a theoretical climate model), and test the results of the analysis procedure in such cases, before attempting to argue that actual time series conform to the implausible statistical model of LTP. As discussed below in my review, the claims of existence of LTP in this and related previous studies is likely simply an artifact of long-term deterministic variability that is incorrectly diagnosed as long-term stochastic behavior by appropriately specified statistical model. The fact that GRL has published other poor papers in this subject area e.g. Cohn and Lins (2005) and Rybski et al (2006) is not a justification for the continued publication of such papers. These papers do not advance the field forward, but simply clutter the peer-reviewed literature in this field with half-baked analyses and false conclusions which, as this present submission makes clear, become self-perpetuating.

Obviously, it is not the purpose of our paper to study the physics and dynamics of the systems but this does not mean that we do not properly recognize them; on the contrary, we believe that a proper stochastic description, consistent with the observed behaviors (as opposed to descriptions that put *ab initio* the postulate of STP) is the most proper recognition of physics and dynamics. We believe that, as the paradigms of statistical thermophysics and quantum physics (which rely on the concepts of probability and statistics and depart from mechanistic physics) have clearly demonstrated, physics and dynamics should not be confused with mechanistic explanations thereof. Specifically, we point out in our manuscript that LTP may be related to the maximum entropy principle; this may be viewed as a non-mechanistic explanation of LTP. Entropy, being a measure of uncertainty, is a probabilistic (rather than a mechanistic) concept, yet a peak research topic today.

We also believe that the application of LTP methods to output of climatic models is not pertinent to the present analysis. We respectfully disagree that the behavior of such series is a priori known. It is likely that climatic models are consistent with LTP as these models must reproduce observed statistical behaviors of climatic data. But if not, this will simply indicate a weakness of models. But in any case, the test the referee proposed, was already contained in the OM (see point a above) and was performed with synthetic data that are known to be non LTP.

Moreover, we would like to point out that there is an extended literature on the detection of LTP (as well as scaling properties and so forth) on climatic data – not only the two recent studies published in GRL. Referee #1 expresses, in our opinion, a general aversion towards these studies. We question the validity of such an aversion as a reason for rejecting our contribution, which is aimed to shed more light on results that were previously published in the literature. We believe that such a general aversion could be the motivation for a comment or a review paper, but not a point for a review of a manuscript, which should be an impartial assessment of its scientific value. It is indeed difficult for us to accept a negative point that applies to recent papers published by GRL, i.e. papers that recently received a positive rating by scientific referees and Editors. We do not question the referee's right to have this aversion for these papers, but we believe that the appropriate vehicle for expressing opinions on published papers is the public dialogue (e.g. journal discussions) rather than a confidential and anonymous review on our submitted manuscript.

In the RM we better explained the focus of our analysis and our opinion about the usefulness of statistical approaches in climate studies.

> There is a history in recent decades of scientists from other fields naively and inappropriately applying certain statistical methods to climatic data. The methods may be appropriate in the context of other problems in the fields these scientists come from, but they are not appropriate in general for the analysis of climate data. Most notable are the various attempts to calculate fractal dimensions, scaling laws, Hurst coefficients, and various measures of long-range or 'long-term' persistence ("LTP") from climate time series. Invariably, there is in such analyses no clear recognition of the extensive body of existing literature on statistical modeling of climate behavior based that is based on principles embracing the underlying physics and dynamics. It is of some concern that the citations of past work the authors here use to motivate their application of LTP to climate data come entirely from the hydrology and biological literature. There does not in fact, to my knowledge, exist a body of literature providing an *ab initio* physical motivation for LTP to the long-term behavior of surface temperature or other climate variables.

Again in this case, the criticism of Referee #1 is not addressed to our paper, but to "…various attempts to calculate fractal dimensions, scaling laws, Hurst coefficient ….". The same consideration we mentioned above applies here, too. We also would like to point out that it is not the scope of our paper to look for a possible physical explanation for LTP in climatic behaviors. This point has been the subject for numerous papers published along a period of over 50 years – and both authors have made some contributions on this. LTP does not apply only to temperature and other climatic variables but, as we point out in the manuscript, it is omnipresent in natural, biological, socio-economical and technological processes. So (as also discussed above), its physical motivation should be traced in a more general physico-mathematical principle, and not particularized in a single phenomenon.

In the RM we mentioned that the maximum entropy principle can be a plausible physical explanation for LTP and made clear that LTP can also originate from deterministic dynamics.

> There are numerous studies going back decades that demonstrate that the natural, stochastic ("noise") variability in surface temperature on interannual and longer timescales is governed primarily short-term autocorrelation (so-called "red noise") associated with the interaction of high-frequency stochastic forcing (i.e. 'weather') with the long timescale response components of the system (e.g. the oceans, cryosphere, land surface processes and biosphere). See e.g. Gilman et al, 1963; Hasselmann, 1976; Wigley and Raper, 1990; Mann and Lees, 1996).
>
> The authors appear to obliquely recognize this. For example, they do discuss Markovian processes and ARMA models, but they fail to recognize firstly that neither MA models nor high-order AR models are faithful to the actual noise processes that impact climate variables, which are best interpreted as a simple short-term autocorrelated "red noise" process which can be represented statistically by an AR(1) model. Such a simple model for climatic noise has the advantages that it is (i) motivated by the actual underlying physics (see e.g. Hasselmann, 1976 and Wigley and Raper, 1990), (ii) parsimonious (characterized by only 1 statistical parameter) and yet (iii) provides a null hypothesis for climatic "noise" that is in fact extremely difficult to in general reject.

We are aware that the short-term autocorrelation is easily understandable and has a simple explanation. This happens also in hydrological processes. However, this does not mean that nature should follow the short-term dependence behavior that we have devised for her. On the contrary, there are numerous studies that prove the consistency of hydroclimatic records with the LTP hypothesis (e.g. Markovic and Koch, GRL, 2005; Bloomfield, 1992). The AR(1) model is not simpler and more parsimonious than the simple scaling model; the latter involves one parameter as well and, as can be verified from a simple look at the two sets of equations in our manuscript, it yields simpler expressions than the AR(1) model; see our text in lines 163-165:

> *"Note that both AR(1) and SSS involve a single parameter each and that the equations (2) and (3) of SSS are simpler than (4) and (5) of AR(1), even though the former has been regarded by many as very complicated."*

This point (LTP versus STP in climatic records) is thoroughly discussed in the RM. The implausibility of AR(1) (according to our opinion) is discussed in the introduction, lines 59-67:

> "*Although many have considered the Markovian behavior physically more plausible for the climate system (e.g. Mann and Lees, 1996), its two aforementioned features (non influence of the past, exclusiveness of a single scale of fluctuation) and other features discussed below might make it implausible, in our opinion. Moreover, climatic records do not verify a hypothesis of Markovian behavior. Thus, its adoption has been usually combined with a decomposition of a climatic series into components, one of which is Markovian (equation (6) in Mann and Lees, 1996); this decomposition is made on stochastic grounds (by spectral methods) and its physical fundament may be disputable, in our opinion.*"

Additional support is provided later. For instance, the reviewer may have not noticed that an AR(1) process at the annual scale is no longer an AR(1) process at a climatic scale, when the climatic process is formed (as usually done) by time averaging; see our text in lines 160-162:

> "*only at scale k = 1 (annual) is the process Markovian (i.e., $\rho_j = \rho^j$); at all other scales the autocorrelation structure in (5) (i.e. $\rho_j^{(k)} = \rho_1^{(k)} (\rho^k)^{j-1}$) is identical to that of an autoregressive moving average (ARMA) process of order (1, 1), another classical example of STP.*"

Furthermore, if the statistical analysis is done correctly, even the AR(1) dependence may result in practically equivalent reduction of significance; as we point out in lines 317-319:

> "*For higher values of ρ both the SSS and the AR(1) processes yield significance levels that are very close to each other; this may be interesting to those who do not trust the LTP hypothesis and prefer to assume an STP behavior.*"

> However, there is an even deeper problem with the present analysis. The analysis presented by the authors assumnes that all variability is stochastic in nature. Yet the stochastic ("noise") component is only part of the behavior of the system. Much of the variability in surface temperatures and other climate variables is not stochastic at all, but deterministic, driven by long-term changes in forcing, e.g. changes in radiative balance. The simplest and most obvious example is the seasonal cycle in surface temperatures, which represents the thermal response to changes in solar insolation over the course of the year, and is characterized by a nearly sinusoidal emperature signal which dominates the variance in monthly temperature series. The seasonal cycle is the most obvious example of a completely deterministic signal whose presence would obviously compromise any attempt to estimate the parameters of a stochastic model based on direct application to actual monthly temperature time series. Indeed, this is why the first thing that climatologists typically do before analyzing temperature time series is to remove the seasonal cycle, yieldings residuals that are known as temperature "anomalies". But this, of course, does not remove all of the "signal". There is much else in the data that is almost certainly not stochastic. Much of the decadal and lower-frequency variability is known to represent a response to long-term changes in radiative forcing due both to natural causes (e.g. volcanic and solar forcing) and in the most recent centuries, anthropogenic impacts (anthropogenic increases in greenhouse gas concentrations and anthropogenic tropospheric aerosols). Stochastic models for surface temperature simply represent the "null hypotheses" for detecting these signals. Indeed, the Wigley and Raper (1990) article represented an attempt to use a simple model for the "red noise" stochastic temperature variations as a null hypothesis for detecting a forced (largely anthropogenic) 20th century trend in surface temperatures.
>
> The current paper shows a complete lack of recognition of the distinction between the stochastic and deterministic components of climate variability, both of which contribute substantially to observed records, and both of which will influence the estimation of any statistical parameters from the raw temperature data. Obviously, the presence of a substantial forced "signal" will compromise any analysis which attempts to describe the record entirely in terms of a stochastic process ("noise"). It is almost certain that any apparent "long-term persistence" in the authors analyses is simply an artifact of the presence of deterministic forced long-term variability, which compromises the naive estimation of stochastic parameters directly from the data.

Our paper does not assume that all the variability of the temperature data is stochastic in nature. There is no hypothesis of this type in the OM. We do not imply any dichotomy of the type "signal" vs. "noise" or "long-term" vs. "short-term" variability. When one calls a part of the time series "noise", one may imply that there do not exist causative dynamics in this, which is not correct. When one says that the long-term variability is deterministic, one may imply that there is no determinism in short-scale variability, which is incorrect.

We analyse the variability of the data, depending on their time aggregation scale, in relationship with the LTP analysis, but it is out of the scope of this manuscript to investigate the reasons for this variability. We only would like to show the high uncertainty that affects the analysis when the variability scales with aggregation timescale according to LTP. We believe that an analysis of the statistical behaviors of climatic records may help to formulate appropriate physical models that are able to provide a physical explanation for the statistical results. In addition, we believe that it is not correct to eliminate a component of the time series according to a preliminary arbitrary hypothesis (especially in the case one is not sure about its suitability as we discussed above).

The only artifact that we can see is the dichotomy "signal" vs. "noise" – but we have avoided it. It is an artifact because it is made by the modeler, not by nature. Even though we recognize that in some modeling attempts it may be a useful practice, it may have catastrophic consequences if it is regarded as a natural behavior.

All these points are briefly discussed in introduction of the RM.

> The obvious test that the authors (and other researchers attempting to attribute low-frequency climate variability to "LTP" or related phenomena) must pass is to be able to demonstrate that the LTP analysis procedure works correctly when applied to a realistic synthetic example where the answer is known *a priori* not to be "LTP". In other words, the procedure has to be demonstrated not to yield unacceptably high "false positives". The obvious test would be to use a climate model simulation which contains both the "red noise" stochastic variability of the sort described by Hasselmann (1976), Wigley and Raper (1990), and many others, but also the deterministic forced variability due to long-term, including anthropogenic, radiative forcing changes (e.g. Crowley, 2000). Jones and Mann (2004) describe several such simulations which have been performed, and in several instances the data have been made publicly available. In this case, , we know the physics present in the models and it is not "LTP". Rather, it is described by a combination of AR(1) "red noise" stochastic variability combined with deterministic long-term forced variability. What would the present authors' analysis methods yield when applied to these simulation results? If the analysis indicates the presence of LTP in this case, it reinforces the interpretation that any apparent "LTP" in these series is simply an artifact of contamination of the estimated parameters of what is wrongly assumed a purely stochastic procedure by the authors' process, by deterministic long-term variability. Until the authors can demonstrate that their claims pass this sort of test, it is difficult not to conclude that the claimed LTP is simply an artifact of long-term deterministic variability that is incorrectly diagnosed as long-term stochastic correlations by their procedure.

As we have replied in the beginning of this report, if the referee had addressed his comments to our manuscript and not to other papers he would not write this comment. The test he suggests here exists both in the OM and the RM.

> 1. The manuscript is unfocused and appears to lack any central hypothesis or conclusions. Much of the paper is devoted to detailed discussion of the concept of LTP. But certain parts, such as the section on "Observation Uncertainty" describe inappropriate specific applications of the concept. The attempt to characterize and explain various different proxy reconstructions of long-term hemispheric temperature variations using the concept of LTP, in particular, is completely nonsensical. First of all, the use of the concept of LTP to characterize hemispheric surface temperature is unjustifiable on physical grounds, as discussed above. It therefore follows that the concept is equally inappropriate for characterizing imperfect reconstructions (from proxy data) of that variable. Equally problematic, the use of LTP to attempt to classify climate reconstructions into different groups simply obscures, rather than enlightens, the true reasons for differences among different reconstructions. The true reasons likely relate to differences in the types of proxy data that have been used in the various reconstructions, the seasonal and spatial representativeness of the various reconstructions, and the differing statistical calibration approaches. See for example the review by Jones and Mann (2004). It is difficult to see how the concept of LTP can provide any possible insight into these issues. There is an entire body of literature which the authors seem to be entirely unware of (see e.g. the Jones and Mann review described above) that provides far more insight into these issues.

We agree with the referee that some parts of the OM were unfocused and not clear. Accordingly, we reformulated and shortened the section titled "Observation Uncertainty". However, we still believe that an analysis of the LTP properties of the considered proxy series might be useful. As we mentioned before, we do not believe that the use of LTP is not justified on physical grounds. We mentioned in the introduction of the RM that in practice LTP is related to the presence of long term cycles. We believe that this outcome is physically plausible in climatic series and therefore LTP is fully justified. In this respect, we find it useful to show that the different proxy series (even though all see to exhibit LTP) have different statistical characteristic, which confirms that they are affected by uncertainty. We think that the Referee is right about the possible explanations for such uncertainty and we included in the RM his suggestions in this respect (including the reference).

> 2. The paper is overly pedantic, with much of the manuscript used for the detailed development of mathematical formalisms which are surely already widely available in the publish literature cited. This is especially irritating since even cursory review of the existing literature discussing the underlying physical considerations for the problem at hand (understanding surface temperature variations on interannual and longer timescales), renders all of the assumptions made inappropriate anyway.

We agree with the referee that part of the OM was overly pedantic – but perhaps this very discussion justifies up to a certain degree some persistence in fundamental concepts. However, clearly the mathematical development is new and not presented in the previous literature, with the exception of the equations we used to define LTP and autocorrelation structures. If these were widely available in the published literature, i.e. trivial, then perhaps Referee #2, would not make the following comment:

"Table 1 contains amazing -- almost unbelievable -- results related to the equivalent sample size of the various data sets if we assume that long-term persistence is present. In particular, the 150-year instrumental northern hemisphere temperature record, which exhibits a Hurst coefficient of about H=0.93, contains the equivalent of only about 2 years of equivalent "white noise" information. That seemed incredible, so I checked: The computations are correct. That fact alone should wake people up; the world of long-term persistence -- which both papers concede we're living in -- is not easy to fathom".

As far as the appropriateness of the assumptions is concerned, please see our comments above.

> 3. Five of the citations are to web pages!! This is thoroughly inappropriate, and a disturbing trend if it is indeed the case that other scientists are now doing this. The content of a webpage is even less reliable than the typical 'gray literature' (conference proceedings, technical reports, etc.). At least the latter are typically written by scientists with some expertise in the area. A webpage can be created by anybody, without any expertise whatsoever in the topic under discussion. Neither GRL nor any other journal should permit the citation of webpages as supporting material for scientific arguments made in the peer-reviewed literature.

We believe that the internet for most of us is a global and interactive representation of human knowledge and scientific communication (cf. the Berlin Declaration of 2003). We recognize that our citations to web pages were not in accordance with the AGU reference style (we should have rather put them in the text and not in the reference list, as we have done now in RM). But simultaneously we would like to point out that we have always followed the practice to cite even personal communications, a practice that it is typical in the scientific community aiming to give appropriate credit where credit is due. We cannot imagine that public communications on the internet should be excluded from this practice. However, in the RM we had to exclude most of the citations to web pages (and many others) to make the manuscript shorter – because our initial resubmission trials were not acceptable due to excess length.

> 4. The written English here is quite poor, and at times nearly unintelligible. The manuscript would have benefited greatly from a critical reading by a colleague with better English writing skills.

We confess that we are disabled in terms of this; English is not our native language. We are sorry if the referee had problems in reading and understanding our wordings.

2. Reply to Referee#2

> This manuscript addresses a very old topic that has recently resurfaced and taken on enormous importance in two separate letters in GRL (Cohn and Lins ("CL05") and Rybski et al. ("R06")): How to detect trends in climate data in the context of long-term persistent errors. The underlying issue is whether the observed trend in global atmospheric temperature should be attributed to some causal factor (e.g., increasing carbon dioxide) or might be explained by natural variability. CL05 and R06 agree that long-term persistence is present in temperature (and other hydroclimatic) time series. However, they draw quite different conclusions about the impact of long-term persistence with respect to the attribution question above. The purpose of this manuscript is to sort out and reconcile the differences between CL05 and R06, and, to a great extent, the present manuscript achieves this goal. It provides a rigorous foundation for considering the differences, as well as presenting a much-needed review of the statistical problems that arise when long-term persistence is present.

We agree with the referee about the much needed review (and, we would add, remedy) of the statistical problems that arise when LTP is present. It is true that our study was triggered by CL05 and R06 but our purpose was more general than to reconcile these two.

> The manuscript also includes some new material that is very intriguing. Table 1 contains amazing -- almost unbelievable -- results related to the equivalent sample size of the various data sets if we assume that long-term persistence is present. In particular, the 150-year instrumental northern hemisphere temperature record, which exhibits a Hurst coefficient of about H=0.93, contains the equivalent of only about 2 years of equivalent "white noise" information. That seemed incredible, so I checked: The computations are correct. That fact alone should wake people up; the world of long-term persistence -- which both papers concede we're living in -- is not easy to fathom.

We are happy for this comment which makes us feel more confident about the originality and importance of our ideas.

> One minor issue: Although R06 refers to an MM03 "reconstruction" (column 7 in Table 1), Steve McIntyre has been very clear that he never produced such a series [see http://www.climateaudit.org/?p=577] but, rather, modified the Mann proxy series to illustrate the lack of robustness in results. It would likely be best to re-label column 7 in Table 1 from “MM03” to something like "RBHS/MM03" -- and clarify the confusion in the text (i.e., lines 211-213).

We made clear in the introduction of the RM that the MM03 data set was derived by modifying the MBH99 record.

> In short, this is a superb manuscript and I strongly recommend that it be published in GRL. In my opinion, other than some editorial minutiae and cleaning up the wording regarding “reconstruction” as noted above, it could be published as is.

We are grateful for this comment. We do not hear every day that a manuscript of ours is superb.

<u>Minor Edits:</u>

l. 1 "…on the long term" ==> "…in the long term"
l. 10 "...arrive to disagreeing conclusions." ==> "...arrive at different conclusions."
l. 12 "…uncertainty on the long term…" ==> "…uncertainty in the long term…"
l. 21 "...arrived to similar conclusions." ==> "...arrived at similar conclusions."
l. 45 "…in the climatic…" ==> "…in climatic…"
l. 48 "In this respect, with this Letter we wish contribute our thoughts, ..." ==> "In this Letter we offer several thoughts, …"

l. 49 "…also put emphasis to another closely…" ==> "…also highlight a closely…"
l. 50 "…uncertainty on the long…" ==> "…uncertainty in the long…"
l. 52 "...independent identically distributed..." ==> "...independent and identically distributed..."
l. 54 "...mislead us so as fail to..." ==> "...mislead us such that we fail to..."
l. 87 "...it should be reminded..." ==> "...it should be recalled..."
l. 185 "This point has been already done in some studies." ==> "This point has already been made in some studies."
l. 188 "...cannot not be..." ==> "...cannot be..."
l. 346 "It may have some interest..." ==> "It may be of some interest..."
l. 347 "...double sided..." ==> "...double-sided..."
l. 354 "...continue to be an attracting one in..." ==> "...continue attracting attention in..."
l. 355 "…as newer data will be accumulated." ==> "…as newer data accumulate."
l. 403 Date of Koutsoyiannis reference needs to be changed from "(2002)" to "(2003)" Note, however, that the corresponding citations in text appear to be correct.

We are also grateful for the detailed corrections suggested by the referee, which we have done in the RM.