

## Response to the comment of Le Mouël, Blanter, Shnirman and Courtillot

We appreciate the comments by Le Mouël et al. on our paper. We thank the editor for enabling us to answer and do however regret that this was not possible for other scientists.

The publications of Le Mouël et al. (2008, 2009) appeared in two different journals (J. Atmos. Solar Terr. Phys. and C. R. Geoscience). Our paper applies to both articles. Clim. Past offers an open discussion in which Le Mouël et al. are invited to participate (which they did). We highly value this review mechanism, which broadens the discussion and we know this would not have been the case with other journals.

We start our reply with a striking statement of Le Mouël et al.: p. 6 (l. 125 to 136), they write “Our commentators may have failed to notice that in none of our papers have we implied a direct causal link between magnetic and temperature variability, in the sense that one would be the direct cause of the other.”

The title of Le Mouël et al. (2009) is: “Evidence for solar forcing in variability of temperatures and pressures in Europe” and the first sentence of the conclusion of Le Mouël et al. (2009) is: “In the present paper, we have provided evidence of significant solar forcing of short-term variations in European temperature lasting up to the present.” [we outlined the three key words.] The term “significant” even implies that statistical tests were performed in the analysis, which they were not (this is not clearly stated in their papers). (High level publication standards like those for *Science* insist on such aspects — see <http://www.sciencemag.org/about/authors/prep/res/style.dtl>.) Le Mouël et al. could consider spending time to revisit their JASTP paper if they believe that their conclusion is not grounded — as admitted in the present comment.

Incidentally, Le Mouël et al. write (p. 10, l. 210-212): “In these papers (Le Mouël et al., 2008; Courtillot et al., 2010), we have seen that some conclusions could be extended for instance to the US: the Sun may not influence only the UK or Europe...” This is also a contradiction with their statement of p. 6 (l. 125 to 136).

In the sequel, we follow Le Mouël et al.’s comments on our paper.

p. 2 (l. 42 to 66). The discussion on data homogeneity was postponed to the manuscript of (Legras et al., 2010). Le Mouël et al. obviously do not admit the key importance of the process of data homogenization (which was done for the geomagnetic data used by (Le Mouël et al., 2009), but they did not mention it). We do not doubt their alleged “long experience on geomagnetic data” but we also emphasize, as experienced members of the climate and meteorological communities, that homogenization of meteorological data is a vital step before starting any interpretation of observations.

We never claimed anywhere that we used the same time series they took. Instead, we explained the criterion we adopted for the choice of time series. Our criterion might be different (we do not claim that it is better) but the remaining stations are rather similar, given the latest, updated ECA&D temperature set. If the general conclusion of their papers is tied to a particular choice of meteorological stations, then they must agree that there is a problem. Our paper, with a (slightly) different choice of stations should be seen as a test of robustness.

p. 3 (l. 66 to 71): Removing the seasonal cycle in meteorological data is done for a simple reason: to obtain an estimate of AR(1) parameters to compare to the L estimates. By doing so, we do acknowledge that there is structure in the anomalies which becomes more visible. This allows a comparison of L estimates with the standard Yule-Walker estimates of “memory”. We already stated in the manuscript that removing the seasonal cycle did not impact estimates of Q, zeta or L. The remarks of Le Mouël et al. on “iron curves” are irrelevant: we do not remove any seasonal cycle in the geomagnetic data. We note that Le Mouël et al. never explain (nor mention) the change of behaviour of the geomagnetic data in the 1940s.

p. 4 (l. 79 to 83): although this is marginally relevant to the message of the manuscript, the statement that the Paris anomalous warming since 1987 is due to an urban heat island effect, without quantification or documentation is speculation; the paper of (Perrier et al., 2005) does not quantify such an effect. The Paris meteorological station has been within the city walls for more than a century, and the urban heat is generally proportional to the logarithm of the population (Oke, 1982). The Paris demography has stalled since 1990, therefore there is no objective reason to attribute the excessive surface warming in Paris to an urban heat effect. We remind that Perrier et al. (2005) do not make such a claim.

p. 4 (l. 87 to 91): the magnetic field is a three dimensional vector. We note, however, that the geomagnetic data available at Eskdalemuir is available in only two dimensions (<http://www.wdc.bgs.ac.uk/catalog/master.html>). We never objected to the fact that only the vertical dimension (Z) can be used in such a study.

p. 5. As explained in our manuscript, AR(1) processes provide a natural interpretation of L(t) in term of a memory parameter (which we do define). The annex 1 of (Le Mouël et al., 2009) gives elements of interpretation of L(t) for an AR(1). Our paper provides additional elements and tries to set a rigorous statistical framework for the “empiric indicators of long-term behaviour”. Our presentation allows a qualitative discussion of significance, which is absent in the papers of Le Mouël et al. (2008, 2009) or Courtillot et al. (2010). Note that their adjective “empiric” means that the indicators have no justification, i.e., they do not refer to a theoretical model and their interpretation is thus pure speculative. Furthermore, they have no established statistical properties (bias, variance) and their dependence on parameters (such as window length) has not been analysed.

p. 5 (l. 101 to 111). First, the L(t) transform can be applied to any time series that yields a variance. The interpretation of L(t) in terms of “process memory” (which we do define in our paper) has only been shown to apply to simple cases, such as AR(1). Other interpretations can be only speculations.

We stated, that the concept of “lifetime” belongs to the framework of *survival theory* and gave a reference for it (Lawless, 2003). Since Le Mouël et al. decided to ignore this source of confusion, we take the opportunity to explain it (although this is not central to the discussion): If T is an random variable denoting the time of death (or an ecosystem, device...), t is some time, then the survival function S(t) is:  $S(t) = \Pr(T > t)$ . The lifetime function F(t) is  $F(t) = 1 - S(t) = \Pr(T \leq t)$ . The so called *survival theory* studies the properties of F(t) given the distribution of T. We do not see a connection between F(t) and what Le Mouël et al. call “lifetime”, and they never provide a reason for using that terminology.

p. 5 (l. 112 to 117): We never wrote that the Sun activity does not affect temperature. We just showed that the demonstration given by (Le Mouël et al., 2009; Le Mouël et al., 2008) is not convincing because the statistics they use is not sufficient to give evidence for it.

p. 6 (l. 118 to 124): Le Mouël et al. (2009) never make estimates of total solar irradiance (TSI) or its variations, as claimed by Le Mouël et al. None of their figures show variations of TSI. Instead, they show the “lifetime” of geomagnetic activity. They maintain the confusion between a variable and its “lifetime” transform.

p. 6 (l. 129-136): Le Mouël et al. write “[...] when one observe a correlation, it is better to analyze it rather than ignore it”. We point out that the correlation is built from “empiric indicators” whose meaning is not clearly defined and whose significance is low, and that Le Mouël et al. never provide an analysis in terms of statistical assessment or quantitative physics, but give a long list of speculations. The entire climate science community is still waiting for a physical model that describes how interannual temperature variability can be affected by daily variability of photons, cosmic ray deflection, cloud formation, solar wind, charged particles, etc.

p. 6-7: the absolute value of the correlation coefficient is not an indicator of statistical significance, it is only the first step of the analysis. The second step evaluates whether this value is likely to occur for uncorrelated series or not (see every basic text books on statistics (von Storch and Zwiers, 2001)). This crucial step is missing in the papers of Le Mouël et al. (2008, 2009) and we note that they have never taken into account the numbers of degrees of freedom (d-o-f). Le Mouël et al. never mention that  $Q(t)$  or  $L(t)$  divide the number of d-o-f by the size of the window, so that the actual number of d-o-f of their series is between 6 and 10. Getting a correlation of 0.84 with such a low number of d-o-f is relatively easy; we do quantify what is meant by “relatively” with the use of p-values (von Storch and Zwiers, 2001), Le Mouël et al. do not in their papers.

p. 7: we note that Le Mouël et al. still fail to show what  $L(t)$  looks like for Z (or H), and hence fail to explain the behaviour we showed (our Figure 5).

p. 8 Our section on bias explains that  $L(t)$  is not a priori a bad idea to estimate the memory parameter of an AR(1). We never claimed that temperature data should be represented by an AR(1), but, as we stated earlier, without a statistical model, an interpretation of  $L(t)$  in terms of memory is speculation.

p. 9 (l. 197): We did give a definition of what we call “memory” (Le Mouël et al. do not, nor do (Courtilot et al., 2010; Le Mouël et al., 2009; Le Mouël et al., 2008)). Referring to that definition, we claim that if the process under study is not AR(1),  $L(t)$  does not give any hint about the memory. Unless Le Mouël et al. give us a proper mathematical demonstration, their claims are speculations and play on words.

p. 9 (l. 199): Le Mouël et al. write “we expect that second order moments contain information about the common solar signal that can be extracted through the evolution of lifetime.” This is precisely where the problem lies. The heuristic tests provided in the Appendix of Le Mouël et al. (2009) do not provide any hint about how common second order moment information can be retrieved from  $L(t)$  when the signals are not piecewise AR(1). Our paper does make the test in a simple case, and the result is less than impressive. Moreover, Le Mouël et al. compare  $Q(t)$  for geomagnetic activity and  $L(t)$  for temperature, and never provide a justification for this particular choice.

The reference to (Lockwood et al., 2010) is not relevant here since the sentence (l. 199) is a specific statistical issue (variations of second order moments) that the paper of Lockwood et al. (2010) does not address.

p. 10 (l. 221). We do not think we made an error. The R code is available for checking. In addition, in their Figure 2, Le Mouél et al. show that two curves resemble each other, but the number of d-o-f of the plotted series is less than 6! How do they explain that the curves do not resemble at all before 1940?

p. 10 (l. 230 to 239): The tests of stability mentioned by Le Mouél et al. do not prove anything about statistical stability, robustness, bias, etc. They test their results by interchanging two similar time series (H and Z for geomagnetic activity). This kind of test always works. They never test that the resulting curves and their variations are due to the method itself. We suggest that Le Mouél et al. check the meaning of stability and robustness in a statistics textbook (von Storch and Zwiers, 2001; Wilks, 1995), and make the appropriate tests in a rigorous way (i.e. following a procedure that is accepted by statisticians).

p. 11 (l. 245 to 256). Le Mouél et al. claim that the correlation of zeta in one variable and L in another variable is interesting without giving any reason for it. They state that “the correlation between temperature lifetime and squared daily differences of the geomagnetic field Z component is natural and expected, because both are influenced by solar forcing.” Again, this is an unjustified claim. Le Mouél et al. assume that proposition P (“temperature is influenced by solar forcing”) is right then claim that they have found evidence for P. This is pointless. The second point is that if the same data transform is applied to temperature and geomagnetic activity, then the correlation is weak and can even change sign (our table 1). Choosing the type of data transform (L or zeta) in order to maximize a correlation is not very useful to draw any scientific conclusion.

p. 12 (l. 257 to 268). Le Mouél et al. again fail to mention the actual number of d-o-f in the series for which they compute a correlation. This is crucial for a test of statistical significance.

p. 12 (l. 274 to 276): Our sentence [“the correlation between L transforms does not allow for an inference of a mutual relation between the original time series, unless a specific model of covariation is provided”] is true: A correlation (even statistically significant) between two variables does not prove a mutual relation (or interaction or causality), because a third variable could influence the two correlated variables without any mutual effect. Moreover, what Le Mouél et al. write does not contradict our statement.

p. 12 (l. 276 to 281). There are standard ways to define a stable correlation: e.g. when time translations do not alter it in a significant way. Reliability has no meaning here. The correlations found by Le Mouél are not stable through time, neither when the windows are changed. The use of data filtering is done in electrical engineering when the signal itself is known and its statistical properties are well constrained. This is not the case for L(t) (or Q(t)) of geophysical time series.

p. 13 (l. 282). The Pearson correlation coefficient (used Le Mouél et al.) is typically a good measure for Gaussian or nearly Gaussian distributions. Other measures of correlation, as the rank correlation should be used for non Gaussian variables. These do, however, give smaller correlations here.

p. 13 (l. 296 to 300). We computed correlations for De Bilt (NL) as Le Mouël et al. We do not see a problem for doing the same for Paris for illustration purpose.

p. 14 (l. 301 to 307): It is striking that the word “regularity” (which does have a meaning in statistics, but does not seem to be employed as such) never appears in the paper of Le Mouël et al. (2009). We also note that Le Mouël et al. admit that they manipulate the data (spatial averages, lifetime transforms, etc.) to seek [Le Mouël et al.’s words, l. 303] a correlation (unless “signature” means something else than “correlation” here). A skilled statistician can transform any random unrelated data sets to yield a positive correlation. This is *data snooping* (White, 2000). Unless Le Mouël et al. formulate clearly a “forcing” (their term) model before they do any mathematical transform of data, their discussion has no statistical validity.

p. 14 (l. 308 to 316): indeed, we fail to be impressed by a list of results whose statistical significance is low, and whose outcome can be due to the method of analysis itself. This failure lead us to redo the analyses and write this paper.

p. 14 (l. 317 to 326). Le Mouël et al. write “ “The increase of temperature (or temperature anomalies) after 1940 is still unexplained by the variations of the geomagnetic field anomalies.” [our words] It should be clear that nowhere in the papers referred to by the commentators do we propose such an explanation, which might make erroneously believe.” We are glad of this clarification, but Le Mouël et al. should re-read what they wrote p. 10 (l. 211)!

p. 19 (figure 2). Those figures are fascinating indeed. How do Le Mouël et al. explain that the red and blue lines are so different before 1940? And how do they explain that even after 1960, the phasing between the red and blue lines can change signs, with temperature “lifetime” leading “geomagnetic” daily variation? How does that fit with their photon, solar wind, cosmic rays, etc. scenario? To experienced statisticians, these figures are not very surprising. It is very likely, that two arbitrary and unrelated but sufficiently smoothly varying series show a finite interval of covariation (or coherence). If this is to be taken as evidence for a relation of the two underlying processes, one has a) to check whether this interval is longer than expected for a plausible null hypothesis and b) find plausible reasons for the disappearance of this visual covariation outside the interval.

We hope that this reply will close the review process. Further work can be done on many of the issues addressed by our paper and the rich discussion and comments on it. We thank all the participants for their contribution.

## References

- Courtillot, V., Le Mouël, J.L., Blanter, E. and Shnirman, M., 2010. Evolution of seasonal temperature disturbances and solar forcing in the US North Pacific. *Journal of Atmospheric and Solar-Terrestrial Physics*, 72(1): 83-89.
- Lawless, J.F., 2003. *Statistical models and methods for lifetime data*. Wiley series in probability and statistics. Wiley-Interscience, Hoboken, N.J., xx, 630 p. pp.
- Le Mouël, J.L., Blanter, E., Shnirman, M. and Courtillot, V., 2009. Evidence for solar forcing in variability of temperatures and pressures in Europe. *Journal of Atmospheric and Solar-Terrestrial Physics*, 71(12): 1309-1321.

- Le Mouél, J.L., Courtillot, V., Blanter, E. and Shnirman, M., 2008. Evidence for a solar signature in 20th-century temperature data from the USA and Europe. *Comptes Rendus Geoscience*, 340(7): 421-430.
- Legras, B., Mestre, O., Bard, E. and Yiou, P., 2010. On misleading solar-climate relationship from inaccurate data and inadequate statistics. *Clim. Past Discuss.*, 6: doi:10.5194/cpd-6-767-2010.
- Lockwood, M., Harrison, R.G., Woollings, T. and Solanki, S.K., 2010. Are cold winters in Europe associated with low solar activity? *Environ. Res. Lett.*, 5(024001): doi:10.1088/1748-9326/5/2/024001.
- Oke, T.R., 1982. The Energetic Basis of the Urban Heat-Island. *Quarterly Journal of the Royal Meteorological Society*, 108(455): 1-24.
- Perrier, F., Le Mouél, J.L., Poirier, J.P. and Shnirman, M.G., 2005. Long-term climate change and surface versus underground temperature measurements in Paris. *International Journal of Climatology*, 25(12): 1619-1631.
- von Storch, H. and Zwiers, F.W., 2001. *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.
- White, H., 2000. A reality check for data snooping. *Econometrica*, 68(5): 1097-1126.
- Wilks, D., 1995. *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, San Diego.