Response to comment:

"On misleading solar-climate relationship

by B. Legras, O. Mestre, E. Bard and P. Yiou,

Clim. Past Discuss., 6, 767-800, 2010".

Vladimir Kossobokov
International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Moscow, Russia

Jean-Louis Le Mouël
Institut de Physique du Globe de Paris, Place Jussieu, Paris, France

Vincent Courtillot
Institut de Physique du Globe de Paris, Place Jussieu, Paris, France

Submitted to CPD June 22, 2010

We respond to the comments formulated by Legras et al (2010, hereafter referred to as LMBY), on two papers published in the *Journal of Atmospheric and Solar Terrestrial Physics* (Le Mouël et al, 2010a, hereafter referred to as LMKC; and Kossobokov et al, 2010, KLMC) on evidence and statistical significance of solar signatures in some of the longest time series of temperature measurements available in Europe. LMBY provide Mathematica codes in their Supplementary material, which can be used as a tool-box by other parties interested in trying to check and reproduce our results and obtain their own independent evaluation of the measurements available, in particular the air temperatures provided, among other data, by the ECA team (Klein Tank et al, 2002; http://eca.knmi.nl). The three conclusions of LMBY as summarized in their abstract are: "*1) that correlation with solar forcing alone is meaningless unless other forcings are properly accounted and that sunspot counting is a poor indicator of solar irradiance, 2) that long series of temperature require homogenization to remove historical artefacts that affect long term variability, 3) that incorrect application of statistical tests leads to interpret as significant a signal which arises from pure random fluctuations*". We believe all of these conclusions can either be rejected, or are valid but do not apply to our work, as discussed below.

LMBY's introduction contains a series of rather general statements which are not specifically addressed to our papers but which nevertheless deserve brief comments. For instance, the lack of multi-decadal trend in either total solar irradiance or the UV part of the solar spectrum cannot be ascertained from the present (mostly satellite based) data which are available only for the past three to four decades. This is well illustrated by ongoing controversies, such as that between Scafetta and Willson (2009) and Krivova et al. (2009). We also strongly disagree with the statement that "*Correlations are a basis of knowledge in areas (...) where the theory is qualitative and does not provide*

mathematical tools for prediction". We believe that science must rely on a balanced mix of observation, theory and numerical modeling, and that this balance may not be sufficiently maintained in certain areas; also, correlations can be, in almost all disciplines and certainly in geophysics, a very powerful guide to the advancement of understanding and formulating theories, and are part of the observational checks that allow to verify them.

Before we start, we would like to quote in full several paragraphs from Press et al (1992) which we try to adhere to as we believe everyone should:

"*Data consist of numbers, of course. But these numbers are fed into the computer, not produced by it. These numbers to be treated with considerable respect, neither to be tampered with, nor subjected to a numerical process whose character you do not completely understand. You are well advised to acquire a reverence for data that is rather different from the "sporty" attitude that is sometimes allowable, or even commendable, in other numerical tasks.*

The analysis of data inevitably involves some trafficking with the field of statistics, that gray area which is not quite a branch of mathematics - and just as surely not quite a branch of science. In the following sections, you will repeatedly encounter the following paradigm:

. apply some formula to the data to compute "a statistic"

. compute where the value of that statistic falls in a probability distribution that is computed on the basis of some "null hypothesis"

. if it falls in a very unlikely spot, way out on a tail of the distribution, conclude that the null hypothesis is false for your data set.

If a statistic falls in a reasonable part of the distribution, you must not make the mistake of concluding that the null hypothesis is "verified" or "proved". That is the curse of statistics, that it can never prove things, only disprove them! At best, you can substantiate a hypothesis by ruling out, statistically, a whole long list of competing hypotheses, every one that has ever been proposed. After a while your adversaries and competitors will give up trying to think of alternative hypotheses, or else they will grow old and die, and then your hypothesis will become accepted. Sounds crazy, we know, but that's how science works! »

"1) correlation with solar forcing alone is meaningless unless other forcings are properly accounted and that sunspot counting is a poor indicator of solar irradiance". In our papers we have attempted to test the hypothesis that a solar signature can be recognized in long records of temperatures (or in transforms of these records) from European stations. We have been looking for possible evidence of forcings with, for instance, 11-yr quasi periodicity, which of course means we have not been mixing this with other, non periodical forcing factors. And we have in no way implied that the mechanism was fully understood (existence versus unicity of a solution) but we have made suggestions that could lead to identifying such mechanisms. This certainly does not in itself require full modeling of all forcings, in particular if one keeps in mind that some forcings may yet have to be discovered and properly modeled, and that several depend on key parameters (sensitivities, feedbacks) that are highly uncertain (e.g. Lindzen, 1999). The present view of the majority of scientists involved in climate change is that forcings include solar effects, but with parameters that lead them to conclude that the contribution of such solar forcing is minor. Regarding the use of sunspots as an indicator of solar activity, we are well aware of the fact that the former does not reflect the latter in a one to one correspondence. But, as we have shown for instance in Le Mouël et al (2009, Figure 4), the long term (multi-decadal) changes in sunspots, the aa-index and a number of geomagnetic proxies all behave in the same important general pattern with a succession of rising, decreasing, rising and then decreasing again segments over the 20th century (what we had called the "*overall magnetic tendency*" in Le Mouël et al (2005) and which should better be termed the "*secular variation of solar activity*") (see also Blanter et al, 2006; Le Mouël et al, 2007; Shnirman et al, 2009). In that particular case, should formal

104 correlation coefficients be used with so few degrees of freedom, most tests would likely
105 fail (indicating that rigorous statistics is not always the best guide for physical insight).
106 Regarding for instance the effect of volcanic eruptions (which we have some experience
107 with, e.g. Chenet et al, 2005), we note that their effects tend to last only a few years and
108 significantly less than a solar cycle, and also that, whereas some eruptions apparently
109 coincided with periods of low sunspot numbers, three major eruptions occurred in the
110 second half of the 20th century which LMBY note is a period of high (multi-decadal)
111 solar activity.

112 The international sunspot numbers are the only and, therefore, the best available
113 proxy of solar activity in the last 250 years. Better proxies of solar activity, such as solar
114 irradiance, cover at most the past three or four solar cycles and, regretfully, cannot be
115 used in the study we have undertaken (LMKC, KLMC).

116 *"2) long series of temperature require homogenization to remove historical*
117 *artefacts that affect long term variability"*. The question of homogenization of the data
118 has already been raised by Yiou et al (2010) and we have responded to it (Le Mouël et al,
119 2010b). We repeat here that we have a profound disagreement with these authors' view of
120 what they call data homogenization, based among other reasons, on decades of
121 experience with magnetic observatory data that raised the same questions (baseline
122 change, change in instrument or measurement location or observer's practice, influence
123 of local magnetic anomalies, numerical errors in transferring the data to a database,…;
124 e.g. Le Mouël et al, 2004; Chulliat et al, 2005). We warn against automated correction
125 and homogenization of these data: it is highly unlikely that one can do better (except of
126 course for easily detected very large errors) than the original observers, particularly for

127 old data. Proper homogenization of data would imply extended stays in the original
128 observatories. Yiou et al (2010) had already noted the existence of homogeneity problems
129 in the data base and pointed out that "more than 94% of stations are flagged as "*doubtful*"
130 or "*suspect*"". The definition of data quality and suspect stations should be specified: the
131 ECA&D data base has three quality control checks for data values: "*Flag=0*" or "*valid*",
132 "*Flag=1*" or "*suspect*" and "*Flag=9*" or "*missing*". Stations are put in three classes
133 according to four homogeneity tests: "*Class 1*" or "*useful*" when no more than 1 test
134 rejects the null hypothesis at the 1% level, "*Class 2*" or "*doubtful*" when 2 tests reject the
135 null hypothesis at the 1% level and "*Class 3*" or "*suspect*" when 3 or 4 of the four tests
136 reject the null hypothesis at the 1% level. Note that due to the definition used by the
137 database editors, the best data in their data base are designated as "*suspect*"! LMBY share
138 these views and recommend performing a data massage such as "*homogenization*" in
139 order to smooth out the roughness of unusual details. We are against this practice. Table 1
140 in LMBY provide a special tabloid "*excerpt*" on homogeneity checking results for the
141 1901–2007 period pointing to the five "*suspect*" stations, but not reminding readers that
142 only 7 out of 126 stations in Europe are found to be "*useful*", whereas 118 are "*suspect*"
143 and 1 "*doubtful*"! How could many of the results we and others have found based on
144 these European observations have been obtained with suspect data containing only noise
145 and no useful information? And how many useful stations are there in the world? It is
146 clear that the homogeneity checking, which rejects 95% of all European data and stations,
147 as recognized by LMBY, is not a useful or reasonable test. LMBY also point out a
148 suspect potential artefact in Bologna, which we had also noted and discussed. LMBY
149 note that there was a change of thermometer in 1867 and a relocation "*to a different*

place" (not mentioned by the ECA team) in 1881. These "*meta data*" do not fit the "*artefact*", which begins in 1865 and ends before 1880: the dramatic changes in (maximum) TX, i.e. rise and fall, were recorded using the same thermometer and at the same place, and thermometers did not show any suspect behavior in variation of (minimum) TN. Finally, we repeat that "*TN and TX values are all of the highest quality code in ECA at each of the three locations*", i.e., "*Flag=0*" that is "*valid*". As a conclusion, we reject LMBY's assertion that "*detection and correction of these heterogeneities are absolutely necessary before any climate study can be based on the instrumental series*". Of course it is better to correct obviously erroneous data points, but we reject blind systematic corrections. What our studies show is that using the raw data does produce interesting information that cannot be due to noise or chance.

"3) incorrect application of statistical tests leads to interpret as significant a signal which arises from pure random fluctuations". This general "lesson" is of course right, but cannot be implied as a criticism of our studies. On the contrary, we have identified a major error in LMBY which invalidates most of their remaining criticism on statistical validity of our results.

Some of the main issues of statistical significance are left for the second part of KLMC, which LMBY have left for future analysis elsewhere. Without waiting for this new piece of comment, we wish to simply refute here our commentators' claim that we have been using the *standard error* in an erroneous way in order to ascertain the separation of the averages. LMBY have been splitting our averages into "averages of averages" and have thereby created confusion in quantifying the statistical significance of potential solar signals. In fact, the sizes of samples $\{T_{H\ ik}\}$ and $\{T_{L\ ik}\}$ are $21 \times N_H$ and

$21 \times N_L$, and the standard errors on these involve the square of these quantities in the denominator. LMBY have erroneously reduced these numbers to N_H and N_L, i.e. the numbers of years in the periods *H* and *L,* which leads them to use $21 \times N_H^2$ and $21 \times N_L^2$ in estimating the standard error of the mean of the mean, $E(E(T_{ik}))$. LMBY therefore overestimate confidence intervals by a factor of about square root of 21.

In addition, as can be seen in their supplementary material, when trying to account for dependencies in a 21-day interval (which we select), LMBY use 90- and 150-day intervals that naturally are affected by the seasonal variability of temperatures (plot and output on page 21, SM to LMBY). Figure 1 actually shows that autocorrelations of the daily temperatures in 21-day intervals fall below 0.2 in less than 3 days, while autocorrelation for the daily range of temperatures ΔT (which LMBY fail to consider) falls below 0.2 on the second day. This is why we used a 21-day interval and not a longer one biased by seasonal effects (from LMKC: "*A 21-day (i.e. 3-week) centered moving average is applied: indeed, this is both long enough to stabilize the still noisy averaged calendar values and yet short enough that features with monthly and longer time constants are well preserved.*"). The effective number of degrees of freedom in our estimates is therefore not reduced as thought by LMBY. As a consequence of their error, the attempts of LMBY to use the T-test technique are biased by the LMBY 9-day dependencies attributed to weekly variability of air temperatures (LMBY page 780 and supplementary material). Should this 9-day correlation apply to our Planet Earth, in our Solar System, weather forecasts would be greatly improved... The sentence "*the number of effective degree of freedom is about 9 times smaller than estimated by LKMC and consequently the estimated variance of the ensemble average is about three times larger*"

196 is therefore false.

197 The checks on our two papers in response to LMBY lead us to conclude that their

198 criticisms are either irrelevant or erroneous and to reaffirm all of our conclusions.

199

200

201

202 Acknowledgements: IPGP Contribution NS xxxx.

203

References

Blanter, E.M., Le Mouël, J.L., Perrier, F., et Shnirman, M.G., 2006. Short term corrélation of solar activity and sunspot : evidence of lifetime increase, *Solar Physics,* 237, 329–350.

Chenet, A.L., F. Fluteau, and V. Courtillot, Modelling massive sulphate aerosol pollution, following the large 1783 Laki basaltic eruption, *Earth and Planetary Science Letters*, *236*, 721-731, 2005.

Chulliat, A., Blanter, E., Le Mouël, J.L., and Shnirman, M., , J. Geophys. Res., 110, A05301. On the seasonal asymmetry of the diurnal and semidiurnal geomagnetic variations, *J. Geophys. Res.*, **110**, A05301, doi:10.1029/2004JA010551, 2005.

Kossobokov, V., Le Mouël, J.-L., and Courtillot, V., A statistically significant signature of multi-decadal solar activity changes in atmospheric temperatures at three European stations, *J. Atmos. Solar-Terr. Phys.*, **72**, 595–606, doi:10.1016/j.jastp.2010.02.016, 2010.

Krivova, N. A., Solanki, S. K., and Wenzler, T., ACRIM-gap and total solar irradiance revisited: Is there a secular trend between 1986 and 1996? *Geophys. Res. Lett.*, **36**, L20101, doi: 10.1029/2009GL040707, 2009.

Le Mouël, J.L., Kossobokov, V., and Courtillot, V., On long-term variations of simple geomagnetic indices and slow changes in magnetospheric currents, The emergence of anthropogenic global warming after 1990?, *Earth Planet*. *Sci*. *Lett*., **232**, 273-286, 2005.

Le Mouël, J.L., Blanter, E., Chulliat, A., and Shnirman, M., On the semiannual and annual variations of geomagnetic activity and components, *Ann*. *Geophys*., **22**, 3583-3588, 2004.

Le Mouël, J.L., Shnirman, M., et Blanter, E., 2007. The 27-Day Signal in Sunspot Number Series and the Solar Dynamo, *Solar Phys.*, 246, 295–307.

Le Mouël, J.-L., Blanter, E., Shnirman, M., and Courtillot, V., Evidence for solar forcing in variability of temperatures and pressures in Europe, *J. Atmos. Sol.-Terr. Phys.*, **71**, 1309–1321, doi:10.1016/j.jastp.2009.05.006, 2009.

Le Mouël, J.-L., Kossobokov, V., and Courtillot, V., A solar pattern in the longest temperature series from three stations in Europe, *J. Atmos. Solar-Terr. Phys.,* **72**, 62–76, doi:10.1016/j.jastp.2009.10.009, 2010a.

Le Mouël, J.-L., Blanter, E., Shnirman, M., and Courtillot, V., Response to comment: "Statistical issues about solar-climate relations by P. Yiou and 7 co-authors, Clim. Past Discuss., 6, 461-487, 2010", submitted to *J. Atmos. Sol.-Terr. Phys.*, 2010b.

Legras, B., O. Mestre, E. Bard and P. Yiou, On misleading solar-climate Relationship, *Clim. Past Discuss.*, **6**, 767-800, 2010.

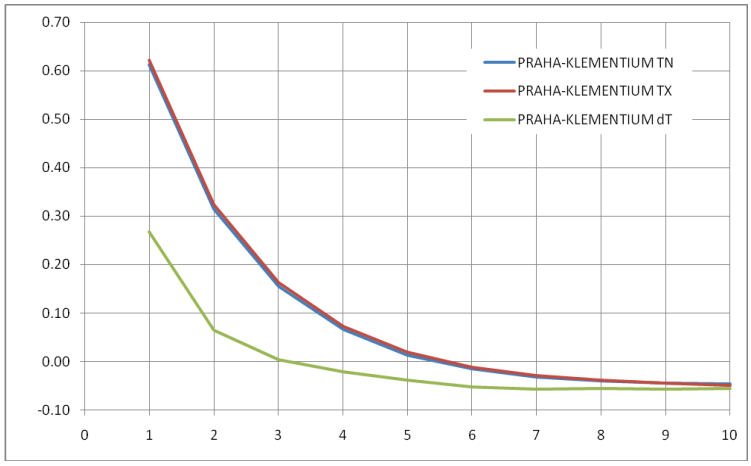
Lindzen, R.S., The Greenhouse Effect and its problems, Chapter 8 in *Climate Policy After Kyoto* (TR Gerholm, editor), Multi-Science Publishing Co., Brentwood, UK, 170pp, 1999.

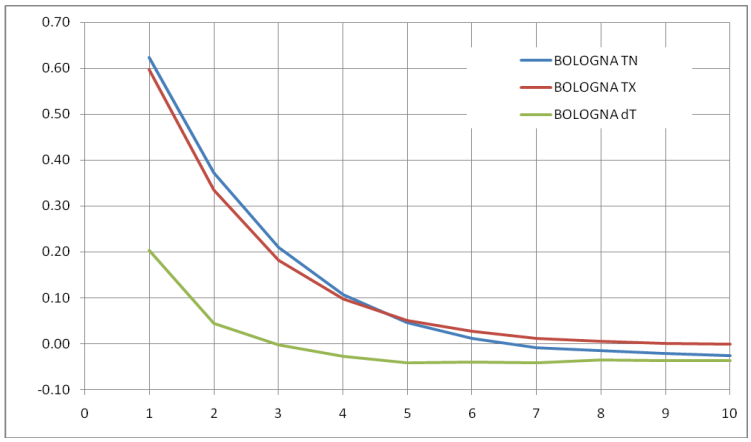
Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). Numerical recipes in C: the art of scientific computing 2nd ed.. New York: Cambridge University Press, 994 p.

Scafetta, N., and R. C. Willson (2009), ACRIM-gap and TSI trend issue resolved using a surface magnetic flux TSI proxy model, Geophys. Res. Lett., 36, L05,701, doi: 10.1029/2008GL036307

Shnirman, M., Le Mouël, J.L., et Blanter, E., 2009. The 27-Day and 22-Year Cycles in Solar and Geomagnetic Activity, *Solar Phys.*, 258, 167–179.

.





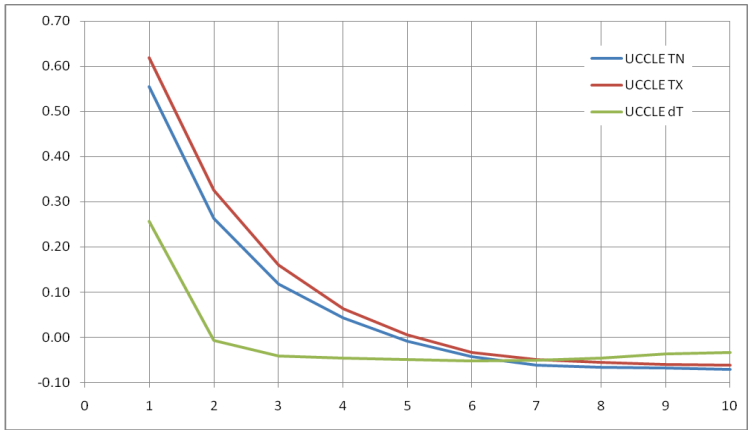


Figure 1.
Autocorrelation functions of TN, TX and ΔT=TX-TN at Prague, Bologna, and Uccle.