How to estimate the significance of global warming when taking explicitly into account the long-term persistence in temperature?

Armin Bunde, Institut für Theoretische Physik, Universität Giessen, Germany

1. Long-term Persistence in climate and its detection

Long-term persistence (LTP), also called long-term correlations or long-term memory, plays an important role to characterize records in physiology (e.g. heartbeats), computer science (e.g. internet traffic) and in financial markets (volatility). The first hint that LTP is important in climate has been given in the classic papers by Hurst more than 50 years ago when studying the historic levels of the Nile River.

We can distinguish between uncorrelated records ("white noise"), short-term persistent records (STP) and long-term persistent records. In white noise all data points $x_1, x_2, \ldots, x_N$ are independent of each other. In STP records, each data point x_i depends on a short subset of previous points x_{i-1}, x_{i-2},.. x_{i-m}, i.e., the memory has a finite range m. In LTP records, in contrast, x_i depends on all previous points. The simplest model for STP is the "AR1 process" where x_i is proportional to the foregoing data point x_{i-1}, plus a white noise component η_{i-1}, $x_i = a x_{i-1} + \eta_{i-1}$, $0 < a < 1$, $i = 1,2,\ldots,N$.

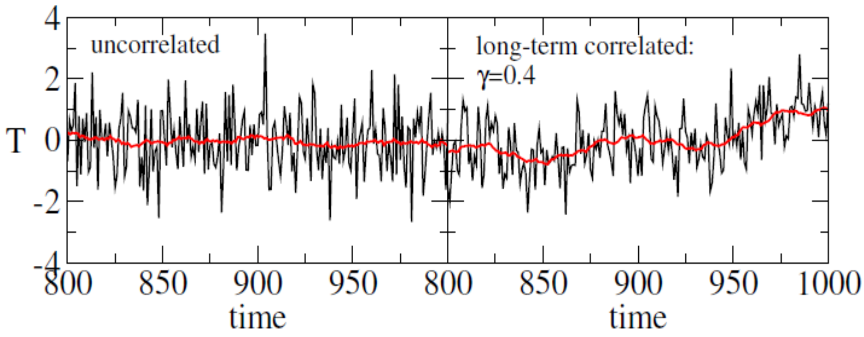
Despite the evidence that temperature anomalies cannot be characterized by the AR1 process, most climate scientists have used the AR1 model when trying to describe temperature fluctuations and estimating the significance of a trend. This usually leads to a considerable overestimation of the external trend and its significance.

There are several methods to quantify the memory in a given sequence. (For a recent review see [1] and references therein). The first one is the autocorrelation function $C(s)$ where s = 1,2,3,… is the lag time between 2 data points. For white noise, there is no memory and C(s) = 0. For the AR1 process, C(s) decays exponentially, $C(s) \sim \exp\{-s/S\}$ where $S = 1/|\ln a|$ is the "persistence length". For infinitely long stationary LTP data C(s) decays algebraically,

$$C(s) \sim (1-\gamma) s^{-\gamma},\ 0 < \gamma \le 1,$$

where γ is called correlation exponent.

The first figure shows parts of an uncorrelated (left) and a synthetic long-term persistent (right) record, with γ =0.4. The full line is the moving average over 30 data points. For the uncorrelated data, the moving average is close to zero, while for the LTP data, the moving average can have large deviations from the mean, forming some kind of mountain-valley structure that looks as if it contained some external deterministic trend. The figure shows that it is not a straightforward task to separate the natural fluctuations from an external trend, and this makes the detection of external trends in LTP records a difficult task. I will return to this later.



One can show analytically [2], that in LTP records with a finite length N, the algebraic dependence of C(s) on s can be seen only for very small time lags s, satisfying the inequality $(s/N)^{\gamma} << 1$. Already for $\gamma = 1/2$ and records of length 600 (which corresponds to 50 years of monthly data), this condition can only be met for very small time lag times s, roughly s < 6. For larger time lags, C(s) decays faster than algebraically. This is an artifact of the method called "Finite Size"-Effect. If one is not aware of this effect, one may be led to the wrong conclusion that there exists no long-term memory in sequences of a finite length.

A similar mistake may happen, when one uses the second traditional method for detecting LTP, the power spectrum (spectral density) S(f). The discrete frequency f is equivalent to an inverse lag time, f=1/s, and a multiple of 1/N. For white noise, S(f) is constant. For STP data, S(f) is constant for f well below m/N (since the data are uncorrelated at time lags s above m), and then decreases monotonously.

For LTP records, S(f) decreases by a power law,

$$S(f) \sim f^{-(1-\gamma)},$$

so one may detect LTP also by considering the power spectrum. However, due to the discreteness of f, the algebraic decay cannot be clearly observed at frequencies below 50/N, which again may lead to the wrong conclusion that there is no long-term memory.

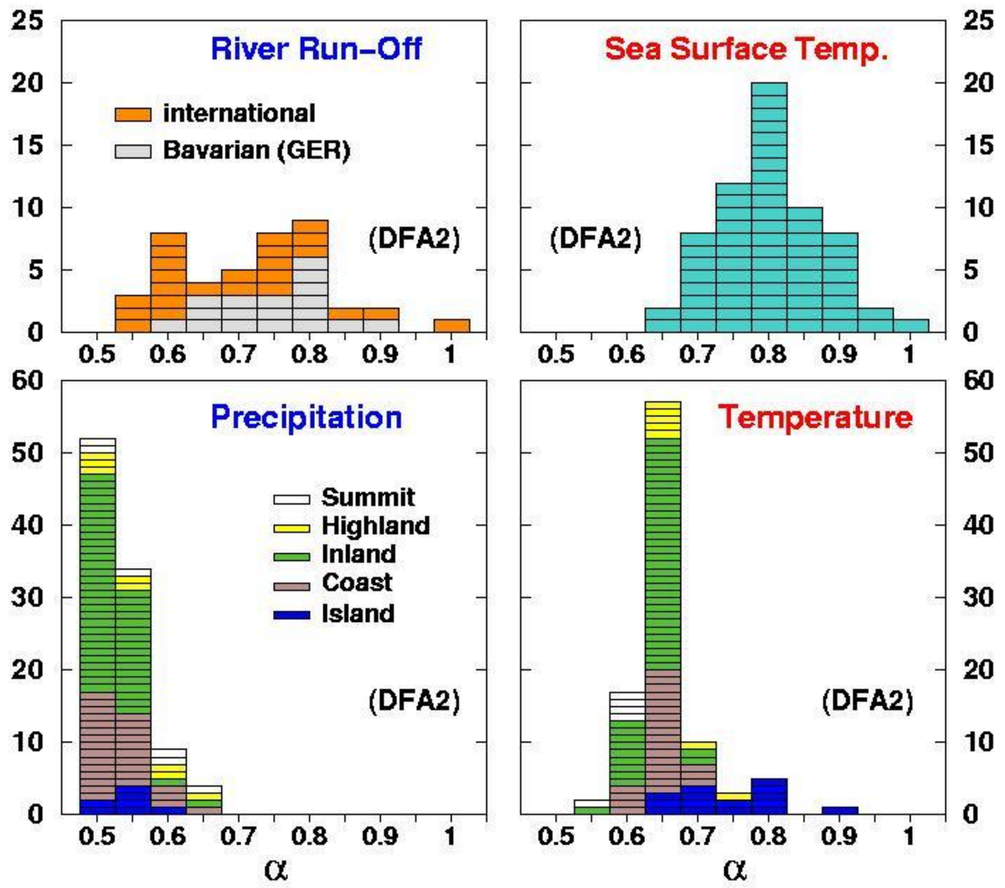
In addition to the remarkable finite size resp. discreteness effects, both methods lead an over-estimation of the LTP in the presence of external deterministic trends.

In recent years, several methods (see, e.g.,[1,3]) have been developed where long-term correlations in the presence of deterministic polynomial trends can be detected. These methods include the detrended fluctuation analysis (DFA2) and Haar-wavelet analysis (WT2), where linear trends are eliminated systematically. DFA2 is quite accurate in the time window $8 \le s < N/4$ while WT2 is accurate for $1 \le s < N/50$. In both methods, one determines a fluctuation function F(s) which measures the fluctuations of the record in time windows of length s around a trend line. For LTP records with correlation exponent γ, F(s) increases as

$$F(s) \propto s^{\alpha}, \ \alpha = 1 - \gamma/2,$$

where α is usually called Hurst exponent. By combining DFA2 and WT2 one can obtain a consistent picture on time scales between s=1 and s=N/4. For a meaningful analysis, the records should consist of more than N = 500 data points. I like to emphasize again that in the case an external deterministic trend cannot be excluded, the evaluation of the LTP and the determination of the Hurst exponent must be done with trend-eliminating methods, e.g., DFA2 and WT2, as described above.

The second figure summarizes the results of our earlier analysis (for references, see [1]) for a large number of atmospheric and sea surface temperatures as well as precipitation and river run-offs. Each histogram shows how many stations have Hurst exponents around 0.5, 0.55, 0.6, 0.65 and so on.



For the daily precipitation records and the continental atmospheric temperatures the distribution of Hurst exponents is quite narrow. For daily precipitation, the exponent is close to 0.5, indicating the absence of persistence (see also [3]), while for the daily continental temperature records, the exponent is close to 0.65, indicating a "universal" persistence law. Both laws can be used very efficiently as test bed for climate models and paleo reconstructions (for references, see [1] and [3]).

There are also more intuitive measures of LTP, and one of them is the distribution of the persistence lengths *l* in a record (see, e. g. [3]). In temperature data, *l* describes the lengths of warm resp cold periods where the temperature anomalies (deviation of the daily or monthly temperature from their seasonal mean) are above resp. below zero. It is easy to show that the distribution P(*l*) of the persistence length decays exponentially for uncorrelated data, i.e., ln P(*l*) ~ - *l*. For LTP data, P(*l*) decays via a stretched exponential, $\ln P(l) \sim -l^{\gamma}$, where γ is the correlation exponent (see [1]). Accordingly, in LTP records large persistence lengths are more frequent, which is intuitively clear.

2. Detection of external trends in LTP data

For detection and estimation of external trends ("detection problem") one needs a statistical model. For monthly (and annual) temperature records the best statistical model is long-term persistence, as we have seen in the foregoing section. The main features of a long-term persistent record of length N are determined by the Hurst exponent α. Synthetic LTP records characterized by these two parameters can be easily generated by a Fourier-transformation (see, e.g., [1]) with the help of random number generators.

When using LTP as statistical model we assume that there are no additional short term correlations, generated by "Großwetterlagen" (blocking situations). Since the persistence length of these short term correlations is below 14 days, they are not present in monthly data sets.

For the detection problem, one then needs to know the probability W(x) that in a long-term correlated record of length L and Hurst exponent α, the *relative trend* exceeds x. For temperature data, the relative trend is the ratio between the temperature change (determined by a simple regression analysis) and the standard deviation σ around the trend line. For Gaussian LTP data, an analytical expression for W(x), for given α and N, has been derived in [4], which is easy to implement and can serve also as a very good approximation for Non-Gaussian data. In order to decide if a measured relative trend x_m may be natural or not, one has to determine the exceedance probability at x_m. If W(x_m) is below 0.025, the trend usually is called significant (within the 95 percent confidence interval), if it is below 0.005, the trend is called highly significant (within the 99 percent confidence interval).

From the condition W(y) = 0.025 one may derive error bars y (within the 95 percent confidence interval) for the expected external trend, $x_{external} = x_m \pm y$. If x_m is slightly below y, then the minimum value of the external trend is negative and thus the trend is not significant. But the maximum value of the external trend can be large, and thus an external trend cannot be excluded, even though the trend is not significant. Accordingly, if a trend is *not significant* since W(x_m) is above 0.025, this does *not* mean that one can *exclude* the possibility of an external deterministic trend. It only means that one is *not forced* to assume an external trend in order to describe the variability of the record properly. For example, if we observe a small insignificant positive trend, then this trend may either arise from the superposition of a strong positive natural fluctuation (as in Fig.1b) and a small *negative* external trend or from a strong negative fluctuation (as in Fig. 1b, but downwards) and a *large positive* external trend.

These conclusions are independent of the used model and hold also for the STP model. In previous significance analyses, climate scientists usually used the STP model, where the model parameter *a* has been determined from measuring C(1) of the data, see Sect. 1. The significance of a trend (see below) is clearly underestimated by this model.

3. Detection of climate change within the LTP model

Using our terminology of "significant" and "highly significant" we have obtained a mixed picture of the significance of temperature records, partly reviewed in [1].

(i) The global sea surface temperature increased, in the past 100y, by about 0.6 degree, which is not significant. The reason for this is the large persistence of the oceans, reflected by a large Hurst exponent.

(ii) The global land air temperature, in the past 100y, increased by about 0.8 degrees. We find this increase even highly significant. The reason for this is the comparatively low persistence of the land air temperature, which makes large natural increases unlikely.

(iii) Local temperatures: In local temperature records it is more difficult to detect external trends due to their large variability. We have studied a large number of local stations around the globe. For stations at high elevation like Sonnblick in Austria or in Siberia, we found highly significant trends. For about half of the other stations, we could not find a significant trend. However, when averaging the records in a certain area, this picture changed. Due to the averaging, the fluctuations around the trend line decrease and the temperature increases become more significant.

Our estimations are basically in line with earlier, less rigorous trend estimations in LTP data by Rybski et al [5] and in line with the conclusions of Zorita et al when estimating the probability that 11 of the warmest years in a 162 year long record all lie in the last 12 years.

My conclusion is that the AR1 process falsely used by climate scientists to describe temperature variability leads to a strong overestimation of the significance of external trends. When using the proper LTP model the significance is considerably lower. But also the LTP model does not reject the hypothesis of anthropogenic climate change.

Literature:

[1] A. Bunde and S. Lennartz, Acta Geophysica 60, 562 (2012) and references therein
[2] S. Lennartz and A. Bunde, Phys. Rev. E (2009)
[3] A. Bunde, U. Büntgen, J. Ludescher, J. Luterbacher, and H. von Storch, Nature Climate Change, 3, 174 (2013), see also: online supplementary information
[4] S. Lennartz and A. Bunde, Phys. Rev. E 84, 021129 (2011)
[5] D. Rybski, A. Bunde, S. Havlin, and H. von Storch, Geophys. Res. Lett. 35, L06718 (2006)
[6] E. Zorita, T.F. Stocker, and H. von Storch, Geophys. Res. Lett. 35, L24706 (2008)