Session HS41: Statistical concepts in understanding and modelling hydro-climatic change

Long term persistence and uncertainty on the long term

Demetris Koutsoyiannis[(1)] and Alberto Montanari[(2)]

(1) Department of Water Resources
National Technical University of Athens, Greece
dk@itia.ntua.gr
http://www.itia.ntua.gr/dk/

(2) Faculty of Engineering
University of Bologna, Italy
alberto.montanari@unibo.it
http://www.costruzioni-idrauliche.ing.unibo.it/people/alberto

Work partially supported by the Italian Government
SPIE and CUBIST research projects

Premise 1 / 3

What is long term persistence (LTP)?

LTP has a long history – It dates back to the pioneering work of Hurst (1951) who first detected his presence when analysing flow records of the Nile River.

Definition of “LTP” (Beran, 1994):

$$\rho(k) \approx c|k|^{-\alpha} \quad \text{with} \quad 0 < \alpha < 1$$

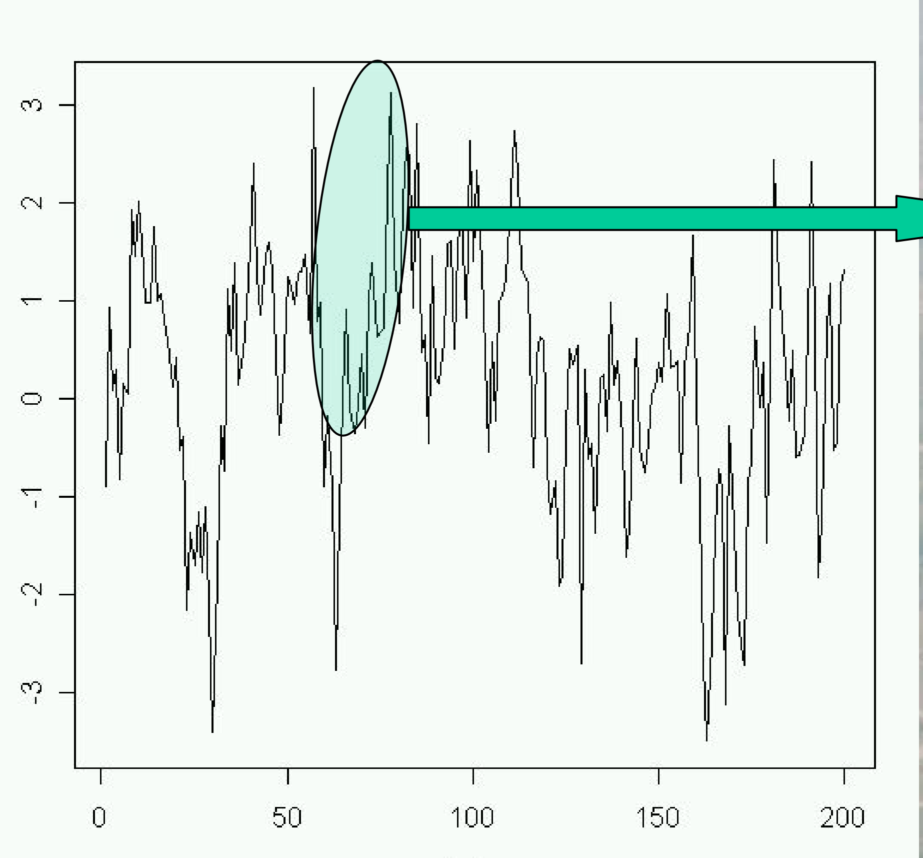
By contrast, short term persistence (STP, the traditional persistence) is characterised by:

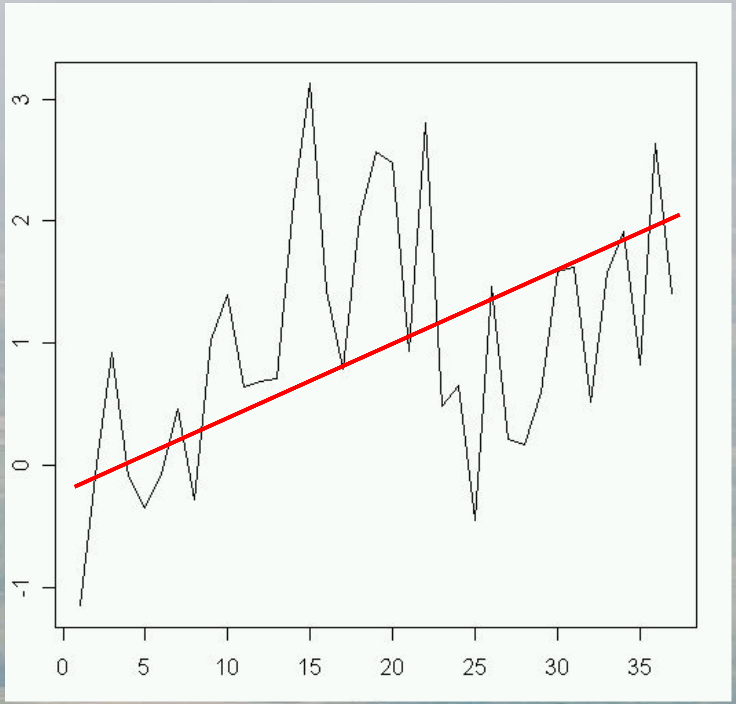
$$\rho(k) \approx c^{-k}$$

In the presence of LTP *“The dependence between events that are far apart in time dimishes very slowly with increasing distance”*.
Overall the time series looks stationary. When one only looks at short time periods, then there seems to be cycles or local trends.

Premise 1 / 3

Example of series generated by a LTP model





LTP has a significant influence on trend estimation. Classical statistics needs to be revisited to account for LTP

Premise 2 / 3

Mandelbrot (1968, 1977, 1983a) detected the possible presence of LTP in several geophysical records. In reference to the famous statement in the Genesis (41, 29-30):

“*Seven years of great abundance are coming throughout the land of Egypt, but seven years of famine will follow them*”

Mandelbrot proposes to name the presence of long term cycles in a river flow record (which can be given by LTP) as “Joseph Effect”. Indeed, the flows of the Nile River were found to be possibly affected by LTP by many authors (Hurst, 1951; Beran, 1994; Montanari et al., 2000; Montanari, 2003).

PROBLEM: to detect LTP is not easy. It is a long term behavior and therefore very long records are needed.

BUT: LTP is very important in the detection of climate change. It induces a significant uncertainty in the long term.

Premise 3 / 3

The intensity of LTP can be measured through the value of the Hurst coefficient H, which varies between 0 and 1.

$0 < H \leq 0.5$ ➔ no LTP

$0.5 < H < 1$ ➔ LTP with increasing intensity.

For positively correlated processes $H \geq 0.5$.
Many methods are available for computing H. We do not go into details here.

Effect of LTP on statistical uncertainty: a classical example

Denote with the symbol μ_n the estimator of the mean of a record $\mathbf{X}_n = (X_1, X_2, \ldots X_n)$ of size n. The standard deviation of the estimator, in the presence of LTP, is:

$$\sigma(\mu_n) = \frac{\sigma(X)}{n^{1-H}}$$

where $\sigma(X)$ is the standard deviation of the process X

If $n = 10^2$, $\sigma = 1$ and $H = 0.5$ we obtain: $\sigma(\mu_n) = 0.1$
If $n = 10^{10}$, $\sigma = 1$ and $H = 0.9$ we obtain: $\sigma(\mu_n) = 0.1$

In practice the situation is even worse as H is estimated from the sample

Goal of the study: analyse the influence of LTP on trend detection in mean annual global temperature series

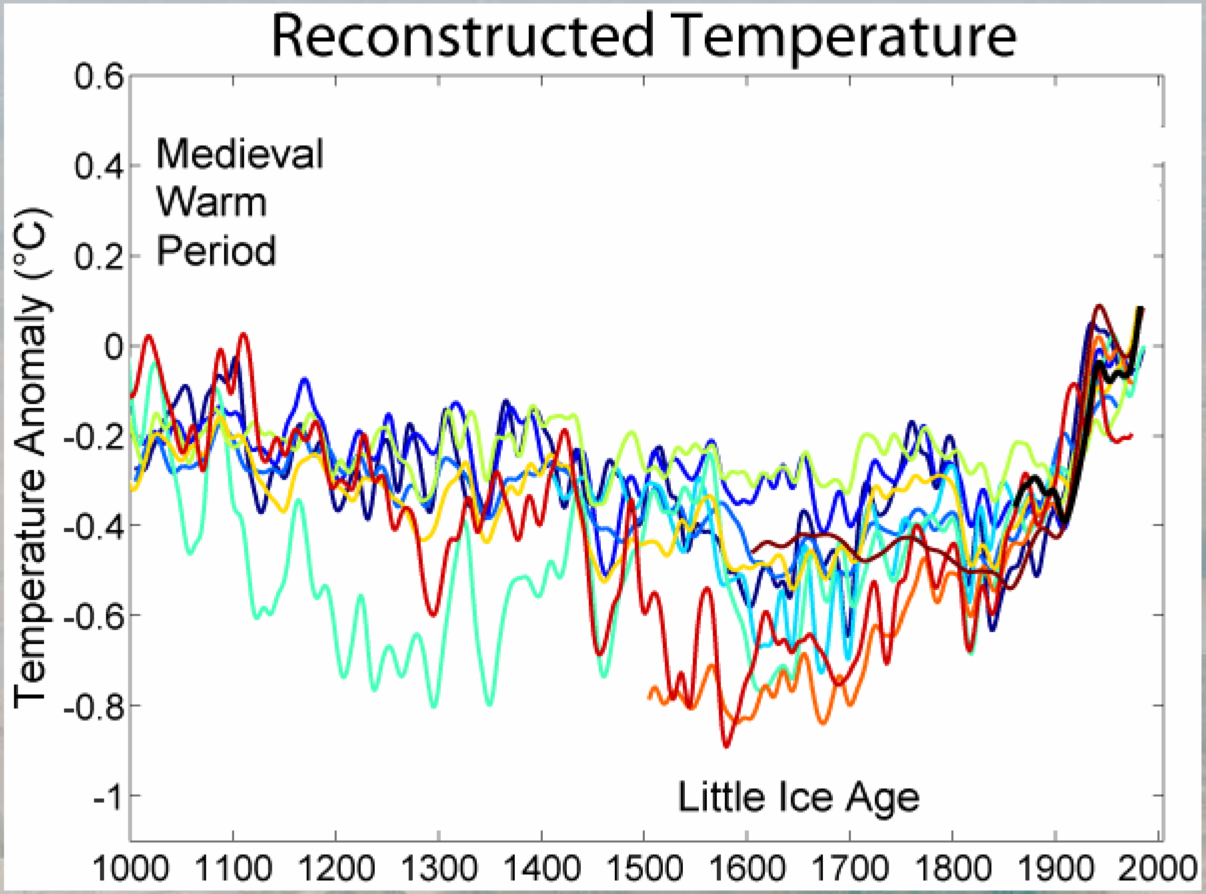
Record for case study

- Climatic Research Unit record (Observed in the period 1880-2005; CRU);

Additional information extracted from:

1. Reconstruction by Jones et al (1998, J98);
2. Reconstruction by Mann et al. (1999, M99);
3. Reconstruction by Briffa (2000, B00);
4. Reconstruction by Esper et al. (2002, E02);
5. Reconstruction by McIntyre and McKitrick (2003, M03);
6. Reconstruction by Moberg et al. (2000, M05).

Goal of the study: analyse the influence of LTP on trend detection in mean annual global temperature series



Is LTP present in the temperature records ?

Data series	CRU	J98	M99	B00	E02	M03	M05
Sample size	150	992	981	994	1162	581	1979
Estimated standard deviation	0.27	0.23	0.13	0.14	0.14	0.17	0.22
H by R/S	1.07	0.90	0.89	0.89	0.93	0.97	0.92
H by ASD	0.93	0.88	0.91	0.91	0.94	0.92	0.94
ρ	0.84	0.53	0.65	0.64	0.81	0.66	0.91

Information contained in reconstructed series

- All reconstructed series indicate strong long term persistence (H = 0.88-0.94), which complies with the instrumental series (H = 0.93)
- Redoing all analyses for the pre-instrumental period (1400-1855) the statistical characteristics, including H, are almost the same

However:

- The different series show different evolutions of temperature even though all are supposed to represent the same physical quantity
- Their uncertainty increases progressively as we move toward the past

Conclusion:

- These series are good to obtain a rough picture of the temperature evolution and a general statistical behaviour - but not good for statistical testing

Trend test statistic

Ribsky et al. (2006) proposed to use the following statistic for detecting the presence of trend:

$$D^i_{l,n} = \mu^i_n - \mu^{i-l}_n$$

- Compute the mean of a sub-sample of size *n* starting from time *i*.
- At time *i-l*, repeat the same computation.
- Compute the difference between the two computed means.
- Compare the computed statistic with a confidence interval of the zero value.

Standard deviation of the test statistic

$$\sigma\left(D^i_{l,n}\right) = \sqrt{2}\sigma\left(\mu_n\right)\sqrt{1-\rho^n_{l/n}}$$

where:

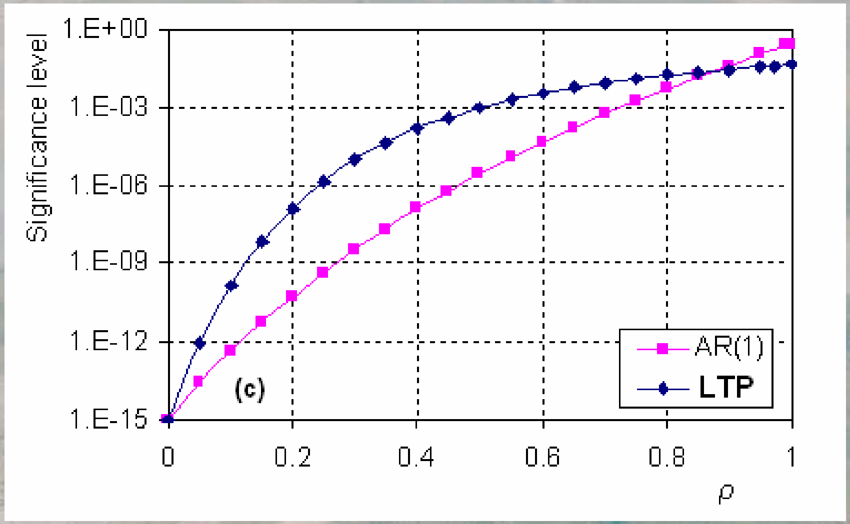
$\rho^n_{l/n}$ is the correlation coefficient of μ_n, i.e. the process X averaged at scale n, at lag l/n, which can be theoretically estimated from the autocorrelation function at scale 1 (annual)

The null hypothesis (no trend) is not rejected if $D^i_{l,n} < 2.58\ \sigma(D^i_{l,n})$ for 1% significance level

Test statistic: an example

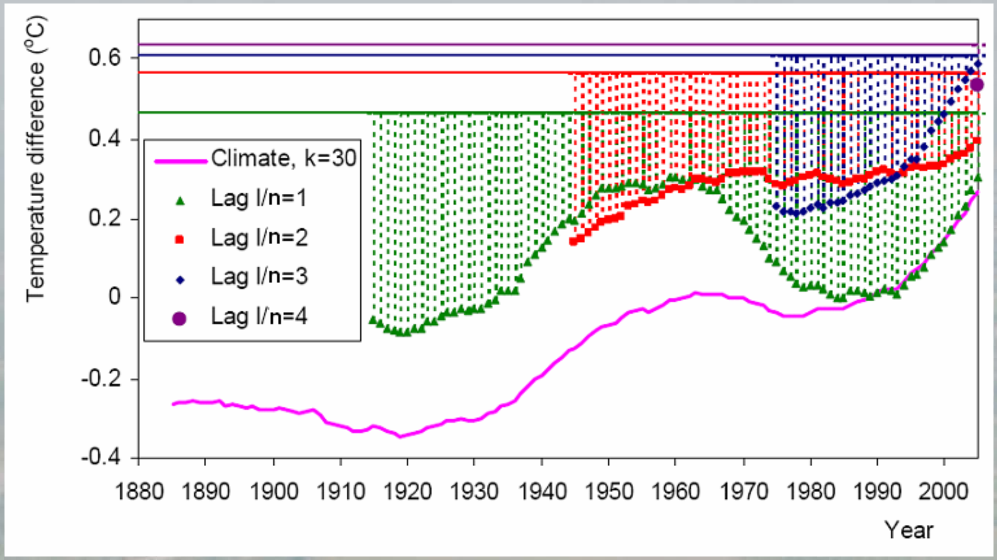
Let's assume that a time series of sample size equal to 150 is available and that, under the assumption of no correlation, the test was computed by assuming $n = 30$ and $l/n = 3$. Let's also assume that the null hypothesis (no trend) has been rejected with a very low risk (10^{-15}). (This in practice means that a trend is likely to be present).

Let's investigate the effect of the presence of correlation in the original time series. The picture below shows the risk associated to the rejection of the null hypothesis as a function of the lag 1 autocorrelation (ρ) of the original data.



Trend detection on the CRU series

- **The pseudo-test did not reject the null hypothesis (no change, no trend).**



- As noted before, a real test would be even less likely to reject the null hypothesis.
- This result agrees with Cohn and Lins (2005).
- Rybski et al. (2006) arrived at an opposite conclusion. However, they may have underestimated some uncertainty factors.

Summary of the tests

- Statistical uncertainty is dramatically increased in the presence of dependence, especially if this dependence is LTP.

- Before conclusions can be drawn, a rigorous methodological framework, based on both physical and statistical arguments, should be built.

- **Quoting from Cohn and Lins (2005): “From a practical standpoint … it may be preferable to acknowledge that the concept of statistical significance is meaningless when discussing poorly understood systems.”**

http://www.costruzioni-idrauliche.ing.unibo.it/people/alberto

Koutsoyiannis, D. & Montanari, A.: Statistical Analysis of Hydroclimatic Time Series: Uncertainty and Insights.
Rejected twice by GRL ☹ – Now in press on WRR ☺