Correlation and Return Interval Analysis of Tree Rings Based Precipitation Reconstructions

A. Bunde[1], J. Ludescher[1], J. Luterbacher[1], U. Büntgen[2] and H. von Storch[3]

[1]*University of Giessen, Germany,* [2]*University of Bern, Switzerland*
[3]*Institute of Coastal Research, Geesthacht, Germany*

We analyze recent tree rings based precipitation reconstructions from Central Europe, the southern Colorado Plateau and the high mountains of North Pakistan, by (i) detrended fluctuation analysis (DFA2), (ii) the Haar wavelet technique (WT2) and (iii) conditional averages. We also study (iv) the PDFs of the return intervals for return periods of 5y, 10y, 20y, and 40y. We find a strong persistence that can be modeled by Hurst exponents α between 0.8 and 0.9. This result, however, is not in agreement with neither observational data of the past two centuries nor millennium simulations with contemporary climate models, which both suggest α close to 0.5 for the precipitation data. This strong contrast in precipitation (highly correlated for the reconstructed data, white noise for the observational and model data) rises concerns on tree rings based climate reconstructions, which will have to be taken into account in future investigations.

Data

We consider tree ring based precipitation reconstructions from Central Europe[1], the southern Colorado Plateau[2] and the high mountains of North Pakistan[3]. For comparison we use recent model data (ECHAM6) from the MPI in Hamburg. In the model, we selected only those time spans which have been used in the reconstructions(April-June for Central Europe, October-July for Colorado, the whole year for North Pakistan). We also compare with observational data (Potsdam, Germany). In order to eliminate the short-term persistence due to weather fronts ("Grosswetterlagen"), we averaged the observational data over two weeks. For eliminating the annual seasonal trend, we subtracted the seasonal mean and divided by the seasonal standard deviation.

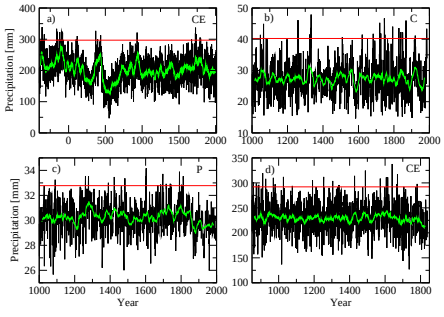

Fig. 1: Reconstructed precipitation in a) Central Europe (398 BC - 2000 AD, April-June), b) Colorado (1000-1988, Oct-July) and c) North Pakistan (1000-1998, all year) compared with d) model simulation from Central Europe (850-1850, April-June). The red line denotes a threshold which is passed on average every 40 years. The green line represents the 30y running average.

Correlation analysis

To reveal the memory in the data, we use 4 techniques.

[A] Detrended Fluctuation Analysis (DFA2) and Haar Wavelet analyses (WT2)

In both DFA2 and WT2 one considers a record $\{x_i\}, i = 1, 2, \ldots, N$, with zero mean, and divides the record into non-overlapping windows ν of lengths s.

In **WT2** one determines in each segment ν, the mean value $\overline{x}_\nu(s)$ of the data and considers the linear combination $\Delta_\nu^{(2)}(s) = \overline{x}_\nu(s) - 2\overline{x}_{\nu+1}(s) + \overline{x}_{\nu+2}(s)$, averages $[\Delta_\nu^{(2)}(s)]^2$ over all segments ν and takes the square root to arrive at the desired fluctuation function $F(s)$. For uncorrelated records ("white noise"), $F(s) \sim s^{-1/2}$. For long-term persistent records we have

$$F(s) \sim s^{\alpha-1}. \quad (1)$$

The exponent α can be associated with the Hurst exponent, and is related to the correlation exponent γ and the spectral exponent β by $\alpha = (1+\beta)/2$ and $\alpha = 1 - \gamma/2$.

In **DFA2** one focuses, in each segment ν, on the cumulated sum Y_i of the data and determines the variance $G_\nu^2(s)$ of the Y_i around the best polynomial fit of order 2. After averaging $G_\nu^2(s)$ over all segments ν and taking the square root, we arrive at the desired fluctuation function $G(s)$. One can show that $G(s) \sim s^{1/2}$ for uncorrelated records, while for long-term persistent records

$$G(s) \sim s^{\alpha}. \quad (2)$$

[B] Conditional averages and persistence length

We distinguish between "rainy" and "dry" years (where the precipitation is either above or below the median). We study the average precipitation $\bar{p}_n$ after n successive rainy (or dry) years and determine the distribution $H(l)$ of the length of "rainy" or "dry" periods. For uncorrelated records, $\bar{p}_n$ does not depend on the history and $H(l)$ is a simple exponential.

[C] Return interval approach

We consider events above a certain threshold Q (as in Fig.1) and are interested in the distribution $P_Q(r)$ of the lengths of the intervals between them. For uncorrelated data, $P_Q(r) = \frac{1}{R_Q} e^{-r/R_Q}$ where R_Q is the mean return interval at fixed threshold Q. For long-term correlated data characterized by a correlation exponent γ, $0 \le \gamma \le 1$, $P_Q(r)$ decays, for $r/R_Q \gg 1$, by a stretched exponential, $\ln(R_Q P_Q(r)) \sim -(r/R_Q)^\gamma$, while for $r/R_Q \ll 1$, $P_Q(r)$ decays by a power law, $R_Q P_Q(r) \sim (r/R_Q)^{\gamma-1}$ [4].

Results

[A] DFA2 and WT2

Figure 2 shows that, in contrast to the model data and the observational data, which can be approximately considered as white noise with a Hurst exponent $\alpha \simeq 0.5$, the reconstructed data show a strong persistence, over long periods of time, which may be compared with the persistence of long-term correlated records with $\alpha \simeq 0.8$. The result for the observational data is in line with [5], where 100 observational records distributed worldwide were analysed.

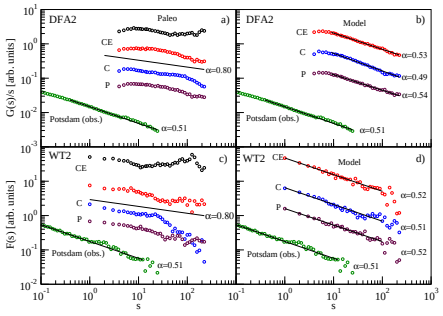

Fig. 2: a) DFA2 fluctuation functions $G(s)/s$ for the reconstructions from (top to down) (i) Central Europe (398 BC - 602 AD), (ii) Central Europe (1000 - 2000 AD), (iii) Colorado (1000-1988 AD), (iv) North Pakistan (1000-1998 AD) and the (v) observational data for Potsdam (1890-2000). b) Same as a) but for the corresponding model data (850-1850). c), d) same as a), b) but for the WT2 fluctuation function $F(s)$.

[B] Conditional averages and persistence length

Figure 3 shows the distribution of the persistence lengths l for the same records as in Fig.2 as well as the conditional averages (inset). The figure shows that reconstructed (a) and model data (c) behave very different. The model data do not show signs of persistence (compare with (d)), while the reconstructed data show a strong persistence, reminiscent of long-term correlated data with α between 0.8 and 0.9 (compare with (b)).

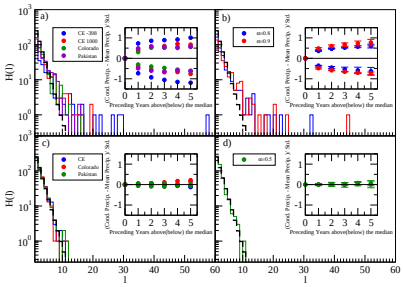

Fig. 3: Conditional average and length of "rainy" and "dry" periods. a) Distribution $H(l)$ of the length of the "rainy" and "dry" periods for the same reconstructed data sets as in Fig. 2(a). Inset: For the same records the conditional mean precipitation above or below the mean precipitation in units of standard deviations. b) Same as a) but for generated long term persistent data with $\alpha = 0.8$ and $\alpha = 0.9$. c) same as a) but for the geographically corresponding model data as in a). Same as b) but for generated long term persistent data with $\alpha = 0.5$.

[C] Return interval approach

We concentrate on the 2400y reconstruction from Central Europe. Figure 4 compares the PDF of the return intervals, in scaled form ($P(r)R_Q$ versus r/R_Q), with the PDF of the model data, for 4 thresholds characterized by the mean return intervals R_Q =5y, 10y, 20y and 40y. We compare also with synthetic long-term persistent data of the same length with $\alpha = 0.9$ obtained by the standard Fourier filtering technique. To reduce the fluctuations, we averaged over 5 records each.

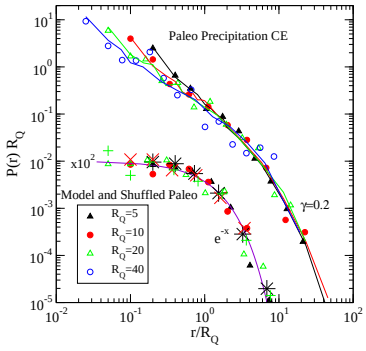

Fig. 4: PDF of the return intervals for (i) reconstructed precipitation for central Europe (398 BC - 2008 AD) (symbols in the upper part) for $R_Q = 5, 10, 20, 40$, (ii) the average over 5 long term persistent records with $\alpha = 0.9$ in corresponding colors, (iii) the shuffled datasets from (i) (same symbols as in (i)) and (iv) model data from Central Europe (remaining symbols in the lower part). The lower straight line shows a simple exponential.

The figure shows that for each data set, the data collapse approximately. For the model precipitation data, the form of $P_Q(r)$ is well described by the simple exponential, which is expected for uncorrelated records. The reconstructed data, in contrast, show a strong persistent behavior at all time scales. The figure shows that this behavior can be quantified by long-term correlations with the comperatively large Hurst exponent $\alpha = 0.9$.

Summary

We have observed a strong contrast in the persistency of tree-ring based precipitation reconstructions on one side and model/observational data on the other side. The Hurst exponent for reconstructed precipitation data is well above the accepted Hurst exponent ($\alpha \simeq 0.65 \pm 0.05$) [6] for continental temperatures. This rises questions on the tree-ring based climate reconstructions.

References

[1] Büntgen U. et al., Science **331**, 578 (2011).
[2] Salzer M.W. et al., Climatic Change **70**, 465 (2005).
[3] Treydte K.S. et al., Nature **440**, 1179 (2006).
[4] Eichner J.F. et al., Physical Review E **75**, 011128, (2007).
[5] Kantelhardt J.W. et al., Journal of Geophysical Research A **111**, D01106 (2006).
[6] Koscielny-Bunde E. et al., Physical Review Letters **81**, 729 (1998).