

Nonlinear correlations of daily temperature records over land

I. Bartos and I. M. János

Department of Physics of Complex Systems, Eötvös University, Budapest, Hungary

Received: 21 August 2006 – Revised: 13 October 2006 – Accepted: 14 October 2006 – Published: 17 October 2006

Abstract. We present a near global statistics on the correlation properties of daily temperature records. Data from terrestrial meteorological stations in the Global Daily Climatology Network are analyzed by means of detrended fluctuation analysis. Long-range temporal correlations extending up to several years are detected for each station. In order to reveal nonlinearity, we evaluated the magnitude of daily temperature changes (volatility) by the same method. The results clearly indicate the presence of nonlinearities in temperature time series, furthermore the geographic distribution of correlation exponents exhibits well defined clustering.

1 Introduction

The time evolution of weather or climate is conveniently characterized by its memory. Short-term memory of a dynamical system is associated with a finite integral timescale ensuing from an exponential decay of the autocorrelation function (von Storch and Zwiers, 1999). Long-term memory is characterized by a diverging integral timescale and linked to power-law behavior of the autocorrelation function (Fraedrich, 2003). Irrespective of the functional form, 2-point correlations reveal only one aspect of the temporal complexity. In general, higher order statistics would be needed to fully characterize the statistical properties of a dynamical system (Mendel, 1991).

When higher order (3-point, 4-point, ...) correlations are trivially related to the 2-point correlation function, the process is termed “linear” and “monofractal”. In case of a nontrivial relationship, the process is called “nonlinear” and “multifractal” (Kalisky et al., 2005). Instead of the rather complicated direct methods for higher order statistics, Ashkenazy et al. (2001) suggested a simple measure for non-

linearity of a time series: when the magnitude of increment series (called volatility) obeys nontrivial long-range correlation, the original time series is nonlinear.

Besides biomedical (Ashkenazy et al., 2001; Kantelhardt et al., 2002) or hydrological (Livina et al., 2003) applications, the evaluation of volatility correlations is widely used in econometric time series (Liu et al., 1999; Qiu et al., 2006). It has been recognized quite a time that the magnitudes of price changes exhibit long-range correlations, reflecting the fact that economic markets experience quiet periods with clusters of less pronounced price fluctuations, followed by more volatile periods with pronounced jumps up and down. A similar behavior for temperature records was detected too (Govindan et al., 2003), even on a paleoclimatic time scale (Ashkenazy et al., 2003), or deep in the equatorial Pacific (Kalisky et al., 2005). A fundamental difference between econometric and temperature records is that the prices themselves are uncorrelated having a white noise spectrum, in contrast temperature data exhibit long-range correlations (Pelletier, 1997; Koscielny-Bunde et al., 1998; Talkner and Weber, 2000; Kantelhardt et al., 2001; Weber and Talkner, 2001; Király and János, 2002; Blender and Fraedrich, 2003; Fraedrich, 2003; Fraedrich and Blender, 2003; Fraedrich et al., 2003; Pattantyús-Ábrahám et al., 2004; Király and János, 2005; Király et al., 2006).

In this paper we provide an analysis of correlation properties for surface temperature data on a nearly global scale. We use the method of detrended fluctuation analysis (Peng et al., 1994, 1995), which became widely accepted in the past years mostly because of its capability of treating nonstationary signals. We show that temperature volatility records obey asymptotic correlations for most of the stations. The geographic distribution of scaling exponents exhibits definite spatial correlations.

Correspondence to: I. M. János
(janosi@lecco.elte.hu)

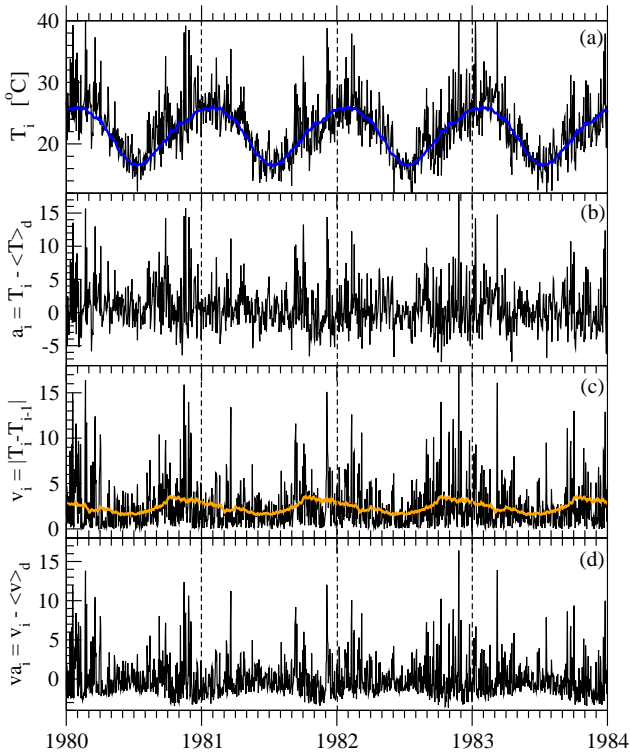


Fig. 1. Example time series evaluated in this work: Sydney station (Australia, 33.87° S, 151.20° E. (a) Four years (out of the 138) of daily maximum temperatures (black) with the empirical annual cycle (blue). (b) Daily maximum temperature-anomaly time series. (c) Volatility series with the empirical annual cycle (orange). (d) Volatility-anomaly series.

2 Data and methods

We evaluated daily minimum and maximum temperature records from the Global Daily Climatology Network (GDCN – <http://www.ncdc.noaa.gov/oa/climate/research/gdcn/gdcn.html>). We concentrated our analysis on the 7320 stations (out of the total 14 737) where the number of recorded data exceeds 8000 (approximately 22 years) with less than 1% missing intervals. Since daily minima and maxima show very similar correlation properties (Király et al., 2006), we present results for maximum values only.

As a first step of the analysis, the annual periodicity is removed from the daily temperature values T_i (Fig. 1a) by the long-time climatological mean $\langle T \rangle_d$ for the given calendar day $d=1 \dots 366$ (when leap days are included), as usual. Note that this procedure, which provides the temperature-anomaly data $a_i = T_i - \langle T \rangle_d$ (Fig. 1b), cannot remove slow trends from the original time series, such as a gradual shift of the annual means. The volatility is defined as the absolute value of temperature-anomaly difference or temperature difference $v_i = |a_i - a_{i-1}| \approx |T_i - T_{i-1}|$. The two definitions give practically the same numerical values, because the deviation of

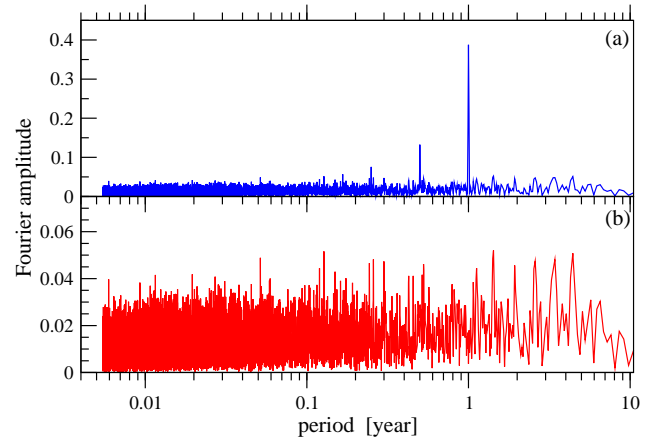


Fig. 2. The high frequency part of the power spectrum as a function of period for (a) the volatility series shown in Fig. 1c, and (b) the volatility-anomaly series shown in Fig. 1d. Note the different vertical scales. The horizontal axes are logarithmic.

climatological averages for consecutive days is typically less than the resolution of temperature data (0.1°C). The volatility series shown in Fig. 1c clearly exhibits a residual annual periodicity as a consequence of seasonally changing variability, which is a well-known attribute of extratropical climate. Figure 1d shows the volatility-anomaly series $va_i = v_i - \langle v \rangle_d$, produced by subtracting the climatological mean values from the signal, as above. At first sight, the removal of annual periodicity might seem to be unsuccessful, however standard Fourier tests (Fig. 2) demonstrate convincingly that the volatility-anomaly series do not have residual periodicities. (Note that the vertical scale in Fig. 2a is six times larger than in Fig. 2b.) Alternatively, the annual component can be removed by normalization with the climatological mean values: $va'_i = v_i / \langle v \rangle_d$. Such series have positive sign everywhere and strongly suppressed magnitudes compared to va_i , nevertheless the correlation properties reflected in the Fourier spectra or DFA scaling (see below) are fully equivalent.

Next we consider the anomaly series as increments of a random walk process. The “trajectory” or “profile” of the signal x_i is given by simple summation as $y_j = \sum_{i=1}^j x_i$, here x_i denotes either a_i or va_i . We divide the profile into nonoverlapping segments of equal length n . In each segment, the local trend is fitted by a polynomial of order p and the profile is detrended by subtracting this local fit. The usual measure of fluctuations is the standard deviation of the detrended segment averaged over all the segments $\langle F_p(n) \rangle$. A power-law relationship between $\langle F_p(n) \rangle$ and n indicates scaling with an exponent δ (DFA p exponent):

$$\langle F_p(n) \rangle \sim n^\delta. \quad (1)$$

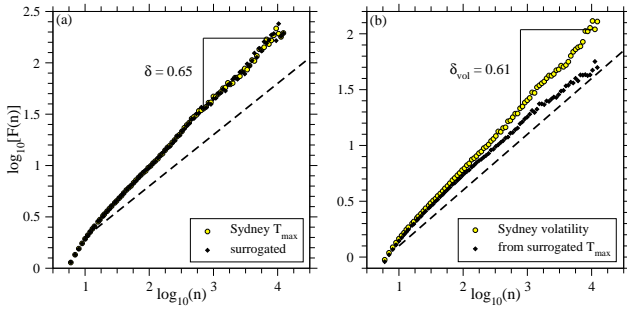


Fig. 3. DFA2 curves for the signals shown in Fig. 1. **(a)** Daily maximum temperature-anomalies (yellow circles) and its surrogate (black diamonds). The latter is produced by an iterative inverse Fourier algorithm matching the autocorrelation functions and probability distributions (see text). **(b)** Volatility-anomalies (yellow circles) and the volatility-anomalies obtained from the surrogate temperature record (black diamonds). Dashed lines illustrate a slope of $1/2$ (uncorrelated processes), exponent values are indicated.

Notice that such a process has power-law autocorrelation function and power spectrum

$$A(\tau) = \langle a_i a_{i+\tau} \rangle \sim \tau^{-\alpha} , \quad S(f) \sim f^{-\beta} , \quad (2)$$

where stationarity requires $0 < \alpha < 1$ and $0 < \beta < 1$. The relationships between the correlation exponents are (Koscielny-Bunde et al., 1998; Talkner and Weber, 2000)

$$\alpha = 2(1 - \delta) , \quad \beta = 2\delta - 1 . \quad (3)$$

Processes of long-term memory are characterized by DFA exponents $\delta > 1/2$, uncorrelated time series (e.g., pure random walk) obey $\delta = 1/2$. Signals with $\delta < 1/2$ are called antipersistent, expressing the tendency that an increasing trend in the past implies a decreasing trend in the future, and vice versa.

Local polynomial fits of order p eliminate polynomial background trends of order $p - 1$. In practice, long-term correlation is inferred when the exponent δ does not depend on p .

Figure 3 illustrates results for the Sydney station. Since we observed slope differences between DFA1 and DFA2 curves for a few stations but not between DFA2, DFA3 and DFA4 within the numerical accuracy of the fitting procedure, we performed the overall evaluation by second order local detrending (DFA2). Exponent values were obtained by linear fits in the interval $[30, 2000]$ days ($\log_{10}(n) \in [1.47, 3.30]$).

In order to demonstrate that long-range correlation in the volatility-anomaly series is indeed a consequence of higher order (nonlinear) correlations, we implemented the method of data surrogation by Schreiber and Schmitz (2000). (For technical details see: Hegger et al., 1999). This procedure does not affect 2-point correlations, whatever is their nature (short-term or long-term), but effectively destroys

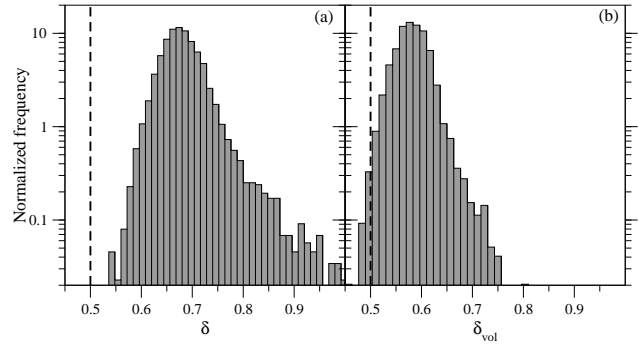


Fig. 4. Normalized histogram of asymptotic DFA2 exponents for **(a)** daily maximum temperature-anomalies, and **(b)** volatility-anomalies. Statistics for 7320 stations is shown.

higher order correlations in the series. The DFA2 curve of volatility-anomaly series obtained from the surrogate temperature record (Fig. 3b, black diamonds) clearly indicates the lack of long-range correlations, while the surrogate series itself has the same 2-point correlation behavior (Fig. 3a).

3 Results

Figure 4 shows the overall statistics of DFA2 exponents for the analyzed stations. Since exponent values could be obtained with a numerical accuracy of ± 0.05 (see also: Király et al., 2006), we conclude that daily temperature anomalies obey significant long-term correlations everywhere. The histogram of volatility-anomaly exponents δ_{vol} (Fig. 4b) is narrower, significant scaling behavior ($\delta_{vol} > 0.55$) can be declared for 6526 stations (89%) out of the total 7320.

The geographic distribution of exponent values is shown in Fig. 5. Unfortunately South-America and Africa are strongly underrepresented in the GDCN data base, furthermore the spatial coverage is very uneven (Király et al., 2006). Nevertheless it seems that relatively low δ values are characteristic at the south-eastern continental boundaries (Fig. 5a). As for the volatility exponents δ_{vol} , the geographic pattern is quite different (Fig. 5b). Probably the only solid conclusion which can be drawn is the tendency of clustering: there are relatively large areas where the exponent values for neighboring stations are similar. This suggests that the results are not numerical artefacts, but they can have a climatological origin.

The comparison of maps in Figs. 5a and b already suggests that there is no strong correlation between the exponent values δ and δ_{vol} for individual stations. The scatter plot shown in Fig. 6 confirms completely this suspicion.

The original observations on a possible link between non-linearity and long-range volatility correlations (Ashkenazy et al., 2001) were fully empirical. Later, Kalisky et al. (2005) and Jun et al. (2006) have developed analytical tools to establish relations between the scaling exponents of a time series

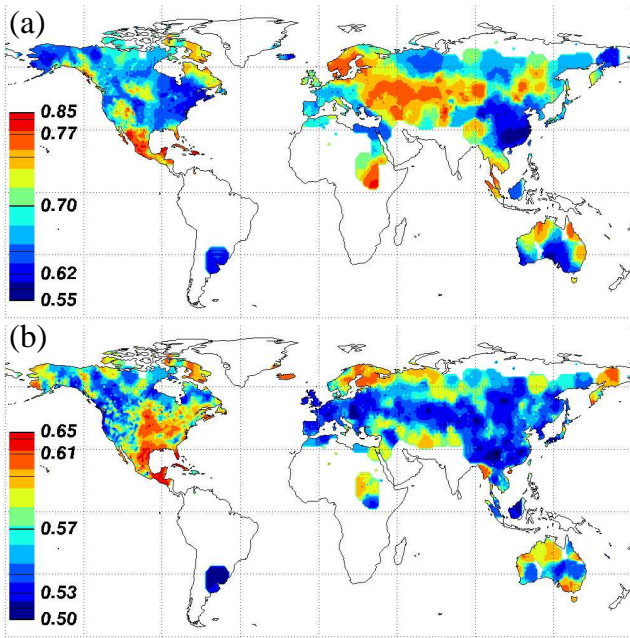


Fig. 5. Geographic distribution of asymptotic DFA2 exponent values for 7320 stations. (a) δ for daily maximum temperature anomalies. (b) δ_{vol} for volatility anomalies. (The color scale is nonlinear and codes different intervals for the two cases.)

and its volatility. In Fig. 6 we show a comparison with the predictions of Kalisky et al. (2005). The black dashed line indicates the expected behavior for a linear process, where the 2-point correlations obey a power law with DFA exponent δ . Up to the value $\delta \approx 0.75$, the volatility exponent is practically constant $\delta_{vol} = 0.5$, and then changes to an approximately linear increase. Unfortunately the number of stations with the strong correlation $\delta > 0.75$ is quite limited, c.f. Fig. 4a, thus the possible crossover in the empirical data is not clear enough. Nevertheless we can probably reject the hypothesis that the temperature records can be fully represented by linear processes with 2-point (power-law) correlations only.

The blue line in Fig. 6 shows the predictions for multifractal processes with a parameter $\sigma = 0.05$, which characterizes the width of the multifractal spectra (Kalisky et al., 2005). The red cloud of the scatter plot does not follow this prediction neither, but at least the blue line crosses its center. One possible explanation of the unstructured scatter plot in Fig. 6 can be that the numerical differentiation used to produce volatility series is known to strongly enhance inherent noise level in the data, therefore the weak dependence of δ_{vol} on δ remains hidden.

In the original proposal, Ashkenazy et al. (2001) decomposed the derivatives of the time series into a volatility and a sign series. The latter is formed by assigning +1 when the one-day difference $T_i - T_{i-1}$ is positive, and -1 when it is negative. We evaluated the sign series for daily maximum

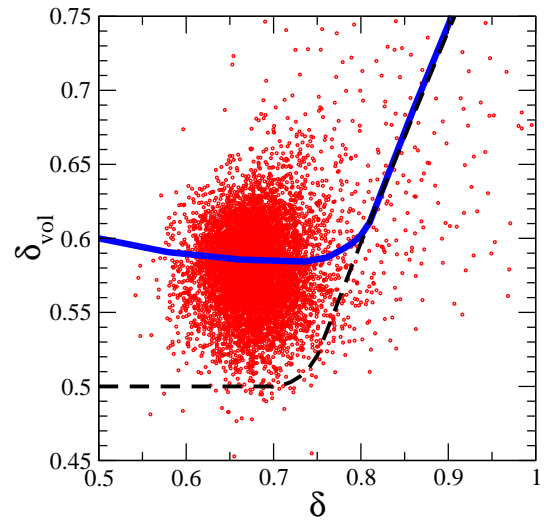


Fig. 6. Correlation plot for the temperature anomaly exponent (δ) and volatility anomaly exponent (δ_{vol}) for each station. Black dashed line indicates the theoretical prediction by Kalisky et al. (2005) for long-range correlated linear processes, the blue line is the same for multifractal series with the parameter $\sigma = 0.05$.

temperatures as well, but we did not find nontrivial results. Figure 7 reveals that the sign of temperature changes can have a weak anticorrelation for time lags up to 4 days, and a clear annual periodicity reflecting the seasonal changes. Otherwise the sign series are uncorrelated. The DFA method is not really suitable to analyze such series, because the typical magnitude of fluctuations around the local trends is extremely small.

4 Conclusions

Although the application of detrended fluctuation analysis has its own pitfalls (Metzler, 2003; Maraun et al., 2004), we can conclude with low risk that terrestrial temperature records obey long-range nonlinear correlations. The term “long-range correlation” obviously does not mean an infinite memory, it reflects simply the fact that we did not observe any breakdown of empirical DFA curves in the available time intervals for temperature anomaly and volatility signals without annual periodicities (c.f. Fig. 3). Indeed, Blender and Fraedrich (2003) and Fraedrich and Blender (2003) demonstrated by globally coupled models that the scaling regime might be restricted, and correlations fade away over ~ 150 years for terrestrial locations far from the oceans. Huybers and Curry (2006) suggest also a change in correlation scaling at around the same timescale, however in the opposite direction: they predict larger exponent values for periods longer than centuries. A clear empirical discrimination between such predictions would require instrumental or proxy

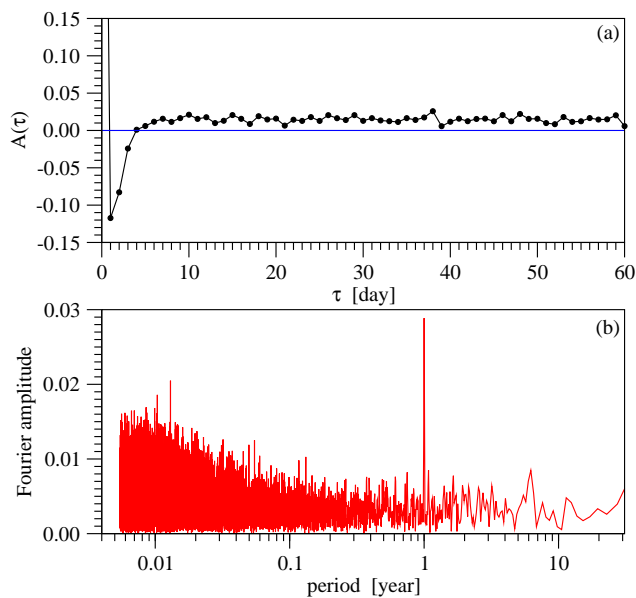


Fig. 7. Autocorrelation function (a) and power spectrum (b) for the sign time series of the derivative of the Sydney record.

records of length ~ 600 years at least, which are not available with the required temporal resolution (≤ 1 month) yet.

As for the detected nonlinearity, the opposite result would be more surprising, since we have no doubt that the atmosphere is inherently nonlinear.

Acknowledgements. This work was supported by the Hungarian Science Foundation (OTKA) under Grant Nos. T047233 and TS044839. IMJ thanks for a János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Edited by: H. A. Dijkstra

Reviewed by: D. Vyushin and two other referees

References

- Ashkenazy, Y., Ivanov, P. C., Havlin, S., Peng, C. K., Goldberger, A. L., and Stanley, H. E.: Magnitude and sign correlations in heartbeat fluctuations, *Phys. Rev. Lett.*, 86, 1900–1903, 2001.
- Ashkenazy, Y., Baker, D. R., Gildor, H., and Havlin, S.: Nonlinearity and multifractality of climate change in the past 420 000 years, *Geophys. Res. Lett.*, 30, 2146, doi:10.1029/2003GL018099, 2003.
- Blender, R. and Fraedrich, K.: Long time memory in global warming simulations, *Geophys. Res. Lett.*, 30, 1769, doi:10.1029/2003GL017666, 2003.
- Fraedrich, K.: Predictability: short- and long-term memory of the atmosphere, in: *Chaos in Geophysical Flows*, edited by: Boffetta, G., Larcotta, S., Visconti, G., and Vulpiani, A., Otto Editore, Torino, 63–104, 2003.
- Fraedrich, K. and Blender, R.: Scaling of atmosphere and ocean temperature correlations in observations and climate models, *Phys. Rev. Lett.*, 90, 108501, doi:10.1103/PhysRevLett.90.108501, 2003.
- Fraedrich, K., Luksch, U., and Blender, R. A.: 1/f-model for long time memory of the ocean surface temperature, *Phys. Rev. E*, 70, 037301, doi:10.1103/PhysRevE.70.037301, 2003.
- Govindan, R. B., Bunde, A., and Havlin, S.: Volatility in atmospheric temperature variability, *Physica A*, 318, 529–536, 2003.
- Hegger, R., Kantz, H., and Schreiber, T.: Practical implementation of nonlinear time series methods: The TISEAN package, *Chaos*, 9, 413–435, 1999, The software package is publicly available at <http://www.mpipks-dresden.mpg.de/~tisean>.
- Huybers, P. and Curry, W.: Links between the annual, Milankovitch, and continuum of temperature variability, *Nature*, 441, 329–332, 2006.
- Jun, W. C., Oh, G., and Kim, S.: Understanding volatility correlation behavior with a magnitude cross-correlation function, *Phys. Rev. E*, 73, 066128, doi:10.1103/PhysRevE.73.066128, 2006.
- Kalisky, T., Ashkenazy, Y., and Havlin, S.: Volatility of linear and nonlinear time series, *Phys. Rev. E*, 72, 011913, doi:10.1103/PhysRevE.72.011913, 2005.
- Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S., and Bunde, A.: Detecting long-range correlations with detrended fluctuation analysis, *Physica A*, 295, 441–454, 2001.
- Kantelhardt, J. W., Ashkenazy, Y., Ivanov, P. Ch., Bunde, A., Havlin, S., Penzel, T., Peter, J. H., and Stanley, H. E.: Characterization of sleep stages by correlations in the magnitude and sign of heartbeat increments, *Phys. Rev. E*, 65, 051908, doi:10.1103/PhysRevE.65.051908, 2002.
- Király, A., Bartos, I., and János, I. M.: Correlation properties of daily temperature anomalies over land, *Tellus A*, 58, 593–600, 2006.
- Király, A. and János, I. M.: Stochastic modeling of daily temperature fluctuations, *Phys. Rev. E*, 65, 051102, doi:10.1103/PhysRevE.65.051102, 2002.
- Király, A. and János, I. M.: Detrended fluctuation analysis of daily temperature records: Geographic dependence over Australia, *Met. Atmos. Phys.*, 88, 119–128, 2005.
- Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H. E., Goldreich, Y., and Schellnhuber, H. J.: Indication of a universal persistence law governing atmospheric variability, *Phys. Rev. Lett.*, 81, 729–732, 1998.
- Liu, Y., Gopikrishnan, P., Cizeau, P., Meyer, M., Peng, C. K., and Stanley, H. E.: Statistical properties of the volatility of price fluctuations, *Phys. Rev. E*, 60, 1390–1400, 1999.
- Livina, V. N., Ashkenazy, Y., Braun, P., Monetti, R., Bunde, A., and Havlin, S.: Nonlinear volatility of river flux fluctuations, *Phys. Rev. E*, 67, 042101, doi:10.1103/PhysRevE.67.042101, 2003.
- Maraun, D., Rust, H. W., and Timmer, J.: Tempting long-memory – on the interpretation of DFA results, *Nonlin. Processes Geophys.*, 11, 495–503, 2004, <http://www.nonlin-processes-geophys.net/11/495/2004/>.
- Mendel, J. M.: Tutorial on Higher Order Statistics (Spectra) in signal processing and system theory: theoretical results and some applications, *Proc. IEEE*, 79, 278–305, 1991.
- Metzler, R.: Comment on “Power-law correlations in the southern-oscillation-index fluctuations characterizing El Niño”, *Phys. Rev. E*, 67, 018201, doi:10.1103/PhysRevE.67.018201, 2003.
- Pattantyús-Ábrahám, M., Király, A., and János, I. M.: Nonuniversal atmospheric persistence: different scaling of daily min-

- imum and maximum temperatures, *Phys. Rev. E*, 69, 021110, doi:10.1103/PhysRevE.69.021110, 2004.
- Pelletier, J. D.: Analysis and modeling of the natural variability of climate, *J. Climate*, 10, 1331–1342, 1997.
- Peng, C. K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L.: Mosaic organization of DNA nucleotides, *Phys. Rev. E*, 49, 1685–1689, 1994.
- Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L.: Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series, *Chaos*, 5, 82–87, 1995.
- Qiu, T., Zheng, B., Ren, F., and Trimper, S.: Return-volatility correlation in financial dynamics, *Phys. Rev. E*, 73, 065103, doi:10.1103/PhysRevE.73.065103, 2006.
- Schreiber, T. and Schmitz, A.: Surrogate time series, *Physica D*, 142, 346–382, 2000.
- Talkner, P. and Weber, R.O.: Power spectrum and detrended fluctuation analysis: Application to daily temperatures, *Phys. Rev. E*, 62, 150–160, 2000.
- von Storch, H. and Zwiers, F.: *Statistical Analysis in Climate Research*, Cambridge University Press, Cambridge, 1999.
- Weber, R. O. and Talkner, P.: Spectra and correlations of climate data from days to decades, *J. Geophys. Res.*, 106, 20 131–20 144, 2001.