Long-range power-law correlations in local daily temperature fluctuations

Eva Koscielny-Bunde[1,2], H. Eduardo Roman[1], Armin Bunde[1], Shlomo Havlin[1,2] and Hans-Joachim Schellnhuber[3]

[1] Institut für Theoretische Physik III, Universität Giessen, D-35392 Giessen, Germany.

[2] Minerva Center and Department of Physics, Bar-Ilan University, 52900 Ramat Gan, Israel

[3] Potsdam Institute for Climate Impact Research, D-14412 Potsdam, Germany

Abstract

We have studied long time daily temperature records (between 56 and 218 years) obtained from 12 meteorological stations in Europe and North-America, from various climatological zones. To analyse the fluctuations of the daily temperatures around their average values, we have calculated directly the autocorrelation function $C(\ell)$ between two days separated by ℓ days, and applied also random-walk fluctuation analysis and wavelets methods, which can systematically overcome nonstationarities in the data. In addition, we have also analysed the distribution of the daily temperature fluctuations, which in most cases is well approximated by a Gaussian. Our analysis suggests that the persistence, characterized by the autocorrelation function $C(\ell)$, is long-ranged and approximately decays with a power law, $C(\ell) \sim \ell^{-\gamma}$, with roughly the same exponent $\gamma \cong 2/3$ for all stations considered. This universal persistence law seems to be valid at least for one decade of years, but we cannot exclude the possibility that the range of the power-law correlations even exceeds the range of the temperature series considered.

1. Introduction

Every day experience as well as systematic meteorological observations teach us that the weather is persistent on short time scales. The prediction that today's weather will be mimicked during the following one or two days is a remarkably succesful one. This evidence remains a serious challenge to "sophisticated" weather forecasting, which tries to beat the "trivial" prognosis with the help of involved fluid-dynamical models and advanced digital machinery (Essenwanger 1986, Newman 1991).

In this paper we attempt to corroborate quantitatively the persistence properties of the weather over the full range of temporal scales. For finding the law that governs the persistence we have analysed long temperature records obtained from several meteorological stations in Asia, Europe and North-America, within different climatological zones. At each day i, we have usually considered the maximum temperature T_i. The total number N of days available for a given weather station ranges typically from 36 000 days (Luling) to 80 000 days (Prague). All temperatures are expressed in degree centigrades.

In the analysis of long temperature records that extend over several years, it is necessary to eliminate the periodic seasonal trends. To this end, we have considered the departures of the T_i, $\Delta T_i = T_i - \bar{T}_i$, from the mean maximum daily temperature $\bar{T}_i$ for each calendary date i, say 1st of April, which has been obtained by averaging over all years in the temperature series. We have used several mathematical techniques (direct calculation of the autocorrelation function, random walk theory and wavelets) to analyse the ΔT_i time series. For the application of some of these methods to heartbeat time series and DNA sequences, we refer to Peng et al. (1994).

Our analysis, part of which has been presented elsewhere (Koscielny-Bunde, Bunde, Havlin, Roman, Goldreich, and Schellnhuber, 1997), suggests that the temperature fluctuations at days i and $i+\ell$ are long-range correlated, i.e., the correlation function behaves like

$$C(\ell) \equiv \langle \Delta T_i \Delta T_{i+\ell} \rangle \sim \ell^{-\gamma} \tag{1}$$

with an apparently universal exponent $\gamma \cong 2/3$ for all weather stations considered. The brackets in (1) denote an average over all pairs of temperature data separated by ℓ days,

$$\langle \Delta T_i \Delta T_{i+\ell} \rangle = \frac{1}{N-\ell} \sum_{i=1}^{N-\ell} \Delta T_i \Delta T_{i+\ell}. \tag{2}$$

Without persistence, the ΔT_i are uncorrelated and $C(\ell)$ is zero. If persistence exists up to a certain number of days ℓ_p, the correlation function will be positive up to ℓ_p and vanish above ℓ_p.

¿From our results we can conclude that, within the pertinent error bars, the correlations range at least over one decade of years. We did not find any evidence for a crossover to uncorrelated behavior at very large time scales, and therefore cannot exclude the possibility that the range of the power-law correlations is larger than the range of the temperature series considered.

The paper is organized as follows: In Section 2 we describe the methods for analysing the temperature series. In Section 3 we present our results for the temperature fluctuations from the meteorological station at the city of Prague, which constitutes the longest time series in this study, and present representative results from other meteorological stations considered. We conclude with a discussion of the results.

2. Temperature series analysis techniques

For analysing the time series, we have studied the autocorrelation function (i) directly by evaluating Eq. (2) and (ii) indirectly by applying random walk techniques. In (ii), we do not study the ΔT_i directly, but, for reducing the level of noise present in the finite temperature series, consider the "temperature profile"

$$Y_n \equiv \sum_{i=1}^{n} \Delta T_i. \tag{3}$$

The fluctuations of the profile, on a given length scale ℓ, are related to the correlation function $C(\ell)$. For the relevant case of long-range power-law correlations, $C(\ell) \sim \ell^{-\gamma}, \quad 0 < \gamma < 1$, the fluctuations $F(\ell)$ increase by a power law (see, e.g. Barabasi and Stanley (1995)),

$$F(\ell) \sim \ell^{\alpha}, \quad \alpha = 1 - \gamma/2. \tag{4}$$

For uncorrelated data (as well as for short-range correlations $\gamma \geq 1$), we have $\alpha = 1/2$.

To find how the fluctuations scale with ℓ, we divide the profile into nonoverlapping segments of size ℓ. We calculate the fluctuations $F_\nu(\ell)$ in each segment ν and obtain $F(\ell)$ by averaging over all segments.

We have used two methods for obtaining $F(\ell)$, (a) the standard fluctuation analysis method (FA) and (b) several discrete wavelet methods that differ in the way the fluctuations are measured and possible nonstationarities are eliminated. We begin with the FA.

2.1 Fluctuation Analysis (FA)

We first divide each series of N successive daily temperatures into $K_\ell = [N/\ell]$ nonoverlapping subsequences of size ℓ starting from the beginning and K_ℓ nonoverlapping subsequences of size ℓ starting from the end of the considered temperature series. For each subsequence ν we calculate the square of the difference of the profile at both ends of the interval ν, $F_\nu^2(\ell)$, and average $F_\nu^2(\ell)$ over the K_ℓ subsequences obtained by starting from one end of the series, and over the K_ℓ subsequences obtained by starting from the other end.

2.2 Wavelets Techniques

The wavelets methods are more advanced methods, which are based on the determination of the mean values $\bar{Y}_\nu(\ell)$ of the profile in each segment ν (of length ℓ), and the calculation of the fluctuations between neighboring segments. First we divide, as above, the temperature series in $2 \times K_\ell$ subsequences. Then we determine in each segment ν the mean values $\bar{Y}_\nu(\ell)$ of the profile. The various techniques we have used here differ in the way the fluctuations between the average profiles are treated and possible nonstationarities are eliminated.

(i) In the first-order wavelets method (WL1), we simply determine the fluctuations from the first derivative

$$F_\nu^2(\ell) = (\bar{Y}_\nu(\ell) - \bar{Y}_{\nu+1}(\ell))^2. \tag{5}$$

This way, trends in the profile of a weather station originating, e.g., from an approximately linear increase of temperature due to urban development around the station, are not eliminated.

(ii) In the second-order wavelets method (WL2), we determine the fluctuations from the second derivative

$$F_\nu^2(\ell) = (\bar{Y}_\nu(\ell) - 2\bar{Y}_{\nu+1}(\ell) + \bar{Y}_{\nu+2}(\ell))^2. \tag{6}$$

So, if the profile consists of a trend term linear in ℓ and a fluctuating term, the trend term is eliminated.

(iii) In the third-order wavelets method (WL3), we determine the fluctuations from the third derivative

$$F_\nu^2(\ell) = (\bar{Y}_\nu(\ell) - 3\bar{Y}_{\nu+1}(\ell) + 3\bar{Y}_{\nu+2}(\ell) - \bar{Y}_{\nu+3}(\ell))^2. \tag{7}$$

By definition, WL3 eliminates linear and parabolic trend terms in the profile.

Finally, we average in each case the quantity $F_\nu(\ell)$ over the $2 \times K_\ell$ subsequences of the temperature series considered.

Methods (i-iii) are called wavelets methods, since they can be interpreted as transforming the profile by discrete wavelets representing first-, second-, and third-order cumulative derivatives of the profile (Koscielny-Bunde et al. 1997). The first-order wavelets are known in the literature as Haar wavelets (see e.g. Kantelhardt, Roman and Greiner (1995)). In principle, one could also use different shapes of the wavelets (e.g. Gaussian wavelets with width ℓ), which have been used by Arneodo et al. (1995) to study long-range correlations in DNA. We believe that discrete wavelets are more suitable to study temperature fluctuations for the following reason: Instead of studing the daily temperature deviations, one could start with the annual temperature deviations from the average temperature value, thus neglecting correlations on scales lower than one year. It can be easily seen that the discrete-wavelets method, applied to the annual temperature departures, is identical to the discrete-wavelets method, applied to the daily temperature, for $\ell = n \times 365$, where n is the number of years.

3. Results for the daily temperature fluctuations

We begin our analysis with the temperature records for Prague, which is the longest series in this study. Figure 1 shows the maximum daily temperatures T_i and their variations for the year 1905 in Prague. Values of T_i larger than the mean maximum temperature of the calendary date i, $\bar{T}_i$, are indicated in light grey, and values $T_i < \bar{T}_i$ in black. Qualitatively, persistence is represented by relatively large patches of positive and negative ΔT_i. Indeed, when the data are randomly shuffled, the large patches disappear (Koscielny-Bunde et al. 1997).

Figure 2 shows the normalized distribution $\mathcal{H}(\Delta T)$ of the temperature variations ΔT_i for Prague. The distribution represents the number of days with ΔT_i in the interval $(\Delta T, \Delta T + \epsilon)$, with $\epsilon = 1$ ^{o}C, divided by the total number of days. The figure shows, that the distribution is quite accurately described by a Gaussian, with a half width of about 4.3 ^{o}C.

Figure 3 shows the autocorrelation function $C(\ell)$ obtained by direct evaluation of Eq. (2). The plot shows that two days separated by up to 1 year, are approximately correlated by a power law, Eq. (1), with an exponent $\gamma \cong 2/3$. For ℓ above one year, the data start to scatter due to a lack of statistics. Nevertheless, the change of the slope for ℓ well above one year may indicate the effect of trends due to urban warming of the city of Prague, which can not be eliminated within this method. In the figure, we discarded all data points after the first negative data point.

Figure 4 shows the temperature landscape of Prague, and Fig. 5 shows the fluctuation functions $F(\ell)$ obtained from the four methods described above (FA and the wavelet methods WL1 - WL3). In the log-log plots, all curves are approximately straight lines for $\ell > 10$ days, with a slope $\alpha \cong 2/3$. For ℓ of the order of few days, the slope is a little larger. This result suggests, that there exists long-range persistence expressed by the power-law decay of the correlation function, with an exponent $\gamma \cong 2/3$. A closer look on the curves in Fig. 5 indicates that the effects of trends and correlations (seen already in $C(\ell)$ (Fig. 3)) can be, to a certain extent, separated by the methods. At about 10^3 days, the curves of FA and WL1 show a slight crossover towards a larger exponent α (Eq. 4). This behaviour can be interpreted as the effect of the warming of Prague due to urban development. In contrast, WL2 and WL3 yield approximate straight lines until about 10^4 days above which the data start to scatter. The systematic crossover at about 10^3 days does not occur here, since WL2 and WL3 eliminate the (roughly) linear trend of warming.

To see, whether the power law correlations obtained for Prague are a generic feature of the atmospheric variability, we have examined the temperature fluctuations of 11 more weather stations on the globe. Representative results for four stations (Moscow, St. Petersburg, Tucson (Arizona), and Luling (Texas)) are shown in Figs. 6-9. Figure 6 shows, in analogy to Fig. 1b, the temperature variations of these stations, for the year 1905 and Fig. 7 shows the distribution functions of ΔT, corresponding to Fig. 2. It is apparent that the distributions for Moscow and St. Petersburg, belonging to the same climatological zone as Prague according to Köppen's classification scheme, are reasonably well approximated by a Gaussian, while for Tucson and Luling deviations from a Gaussian are observed. But this feature does not have consequences on the long-range behavior of the correlations. This is quite nicely seen in Figs. 8 and 9, where we show, on a log-log scale, the behavior of $C(\ell)$ (Fig. 8) and the fluctuation functions (Fig. 9). We observe quite similar behavior as for Prague, with the same slope -2/3 of the autocorrelation function and the same slope 2/3 for the fluctuation function.

The range of correlations can not be detected from the data, but it is clear that the power-law correlations range at least over one decade. Since we obtain the same behavior for all stations considered, our results point to the possibility that there may be a universal law of persistence, with a universal exponent $\gamma \cong 2/3$.

We do not speculate here on fundamental physical mechanisms that might generate the surprisingly homogeneous correlation structure discovered through our analysis. Let us point out, instead, that the correct description of large-scale and long-term weather variability has become a quite hot topic in climatology in recent years. This development is related to the scientific quest for identifying anthropogenic enhancement of the green-house effect within the sea of natural atmospheric fluctuations (for a recent discussion see Hasselmann (1997)). Our findings provide a useful test bed for any computer simulations of these fluctuations by so-called general correlation models.

We note furthermore that our analysis can be systematically refined by studying large clusters of meteorological time series from distinct climatic zones separately. Most interesting results can be expected, in particular, by correlating temperature and precipitational data. Different climates may account for small, but clear-cut and characteristic corrections to the average global persistence law reported in this paper.

LIST OF REFERENCES:

Arneodo, A., et al, 1995, Phys. Rev. Lett. **74**, 3293.

Barabasi, A.-L., and Stanley, H.E., 1995, *Fractal Concepts in Surface Growth*, (Cambridge: Cambridge University Press).

Essenwanger, O.M., 1986, *General Climatology: Elements of Statistical Analysis*, Vol.-1B, Landsberg, H.E. (Ed.), World Survey of Climatology (Amsterdam: Elsevier).

Hasselmann, K., 1997, Science **276**, 914.

Ivanov, P.Ch., et al, 1996, Nature **383**, 323.

Kantelhardt, J.W., Roman, H.E., and Greiner, M., 1995, Physica A **220**, 219.

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H.E., Goldreich, Y., and Schellnhuber, H.J., 1997 (preprint).

Newman, S.B., 1991, Weather and Forecasting **6**, 111.

Peng, C.K., et al., 1994, Phys. Rev. E**49**, 1685 (1994).

Schönwiese, C.D., and Rapp, J., 1997, *Climate Trend Atlas of Europe Based on Observations 1891-1990*, (Dordrecht: Kluwer).

Fig. 1: (a) Maximum daily temperatures T_i for the year 1905 in the city of Prague. Values of T_i larger than the mean maximum temperature of the calendary date i, $\bar{T}_i$, are indicated in light grey, and values $T_i < \bar{T}_i$ in black. Here, $\bar{T}_i$ has been obtained by averaging T_i over the period (1775-1992), consisting i.e. of 218 years. We have excluded the 29th February from the bissextile years. (b) Daily temperature variations $\Delta T_i = T_i - \bar{T}_i$ for the same data shown in (a).

Fig. 2: The distribution $\mathcal{H}(\Delta T)$ of the temperature variations ΔT [^{o}C] for Prague over 218 years. The line is a Gaussian fit, $P(\Delta T) = (2\pi\sigma^2)^{-1/2}\ \exp[-(\Delta T)^2/(2\sigma^2)]$, with $\sigma = 4.27$ ^{o}C.

Fig. 3: The autocorrelation function $C(\ell)$ for two days separated by ℓ days, for the city of Prague. The straight line has slope $\gamma = -2/3$ and is drawn as a guide to the eye.

Fig. 4: The temperature landscape $Y_n = \sum_{i=1}^{n} \Delta T_i$ versus number of days n for the city of Prague. Here, $1 \leq n \leq 218 \times 365 = 79570$.

Fig. 5: Fluctuation analysis for the city of Prague. (a) FA (circles), (b) WL1 (triangles), WL2 (diamonds) and WL3 (stars). The straight lines have slopes 2/3 and are drawn as a guide to the eye.

Fig. 6: Daily temperature variations $\Delta T_i = T_i - \bar{T}_i$ during the year 1905 for the weather stations in: (a) Moscow, (b) St. Petersburg, (c) Tucson (Arizona) and (d) Luling (Texas).

Fig. 7: The distribution of temperature variations ΔT_i for: (a) Moscow (1880-1994, 115 years), (b) St. Petersburg (1884-1994, 111 years), (c) Tucson (Arizona) (1895-1991, 97 years) and (d) Luling (Texas) (1902-1991, 90 years). The lines are Gaussian fits with: (a) $\sigma = 5.05$ ^{o}C, (b) $\sigma = 4.62$ ^{o}C, (c) $\sigma = 3.99$ ^{o}C and (d) $\sigma = 4.72$ ^{o}C.

Fig. 8: The autocorrelation function $C(\ell)$ versus ℓ for: (a) Moscow (115 years), (b) St. Petersburg (111 years), (c) Tucson (Arizona) (97 years) and (d) Luling (Texas) (90 years). The straight lines have slope $\gamma = -2/3$ and are shown as a guide to the eye.

Fig. 9: Fluctuation analysis for: (a) Moscow, (b) St. Petersburg, (c) Tucson (Arizona) and (d) Luling (Texas). The different methods are denoted by different symbols: FA (circles), WL1 (triangles), WL2 (diamonds) and WL3(stars). The straight lines have slopes 2/3 and are drawn as a guide to the eye.