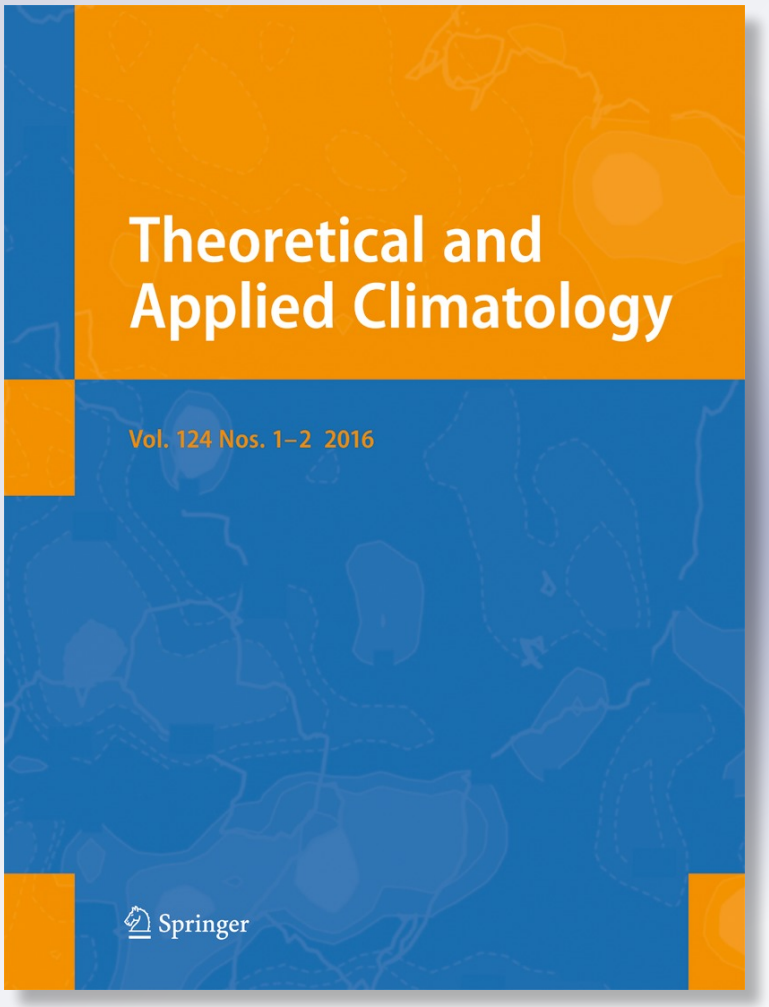
Impact of previous one-step variation in positively long-range correlated processes

Zuntao Fu, Fenghua Xie, Naiming Yuan & Lin Piao

Theoretical and Applied Climatology

ISSN 0177-798X
Volume 124
Combined 1-2

Theor Appl Climatol (2016) 124:339-347
DOI 10.1007/s00704-015-1419-9



Springer

Theor Appl Climatol (2016) 124:339–347
DOI 10.1007/s00704-015-1419-9

ORIGINAL PAPER

Impact of previous one-step variation in positively long-range correlated processes

Zuntao Fu · Fenghua Xie · Naiming Yuan · Lin Piao

Received: 1 April 2014 / Accepted: 23 February 2015 / Published online: 13 March 2015

Abstract In a positively long-range correlated process, variations among consecutive steps are interdependent, especially the influence of previous one-step variation on next steps. How to quantify this kind of impact is of great importance to predict the future variations. In this paper, we demonstrate that this kind of impact depends on the memory strength of underlying processes from two aspects based on the theoretical and observational calculations. More precisely, the conditional calculations and the marginal distribution of the next step variation with given distribution of the previous one-step variation. Both the theoretical and observational calculations demonstrate that the previous one-step variation affect greatly the variation for the next one-step, and the expectation of next step variation will shift to larger value as the increase of memory strength but with a much smaller uncertainty. This is beneficial for our one-step ahead prediction, and will be especially beneficial for multi-step ahead prediction.

1 Introduction

Fluctuations of many atmospheric or climatic variables are considered to be auto-correlated. On small time scales (daily to weekly, which correspond to the general weather scale regimes), it is normally believed that the considered variation can be well approximated by low-order autoregressive (AR) processes. By using a simple linear function.

$$x_i = ax_{i-1} + \epsilon_i, \quad i = 2, 3, 4 \cdots \tag{1}$$

where x_i denotes the past values and ϵ_i stands for a Gaussian white noise process and a stands for the strength of auto-correlation between the past and present states, one could simulate the next step variation adequately. If the considered time series are properly detrended, for many atmospheric or climatic variables, there is even no need to apply a higher-order AR model for the simulation (Bartos and Janosi 2005).

However, besides the short-term correlation on weather scale, on large time scales (monthly, yearly, etc.), it has been recently recognized that many atmospheric or climatic variables are also characterized by long-range correlations (LRC) (Koscielny-Bunde et al. 1998; Yuan et al. 2012, 2014). Long-range correlation means the present variation may be influenced by the variations a long time ago, and normally, it is also named as long-term memory (LTM), long-term persistence (LTP), or the Hurst phenomenon. Different from the weather-scale (short-term) correlation, the so-called LRC is believed to be originated from the interactions between fast-excitation subsystems and slow-response subsystems, such as the atmosphere-ocean interactions. Therefore, understanding this long-range correlation phenomenon is important for our further analysis of the whole climate system. Besides temperature variations, recently, many studies show that many records exhibit long-range correlations, such as DNA sequences (Peng et al. 1994), economic time series (Mantegna and Stanley 1995), heartbeat records (Peng et al. 1995), turbulence, biological and financial data fluctuations, and more generally, in self-organized critical systems (Ivanova and Ausloos 1999). Thanks to the development of nonlinear science, a new

Z. Fu (✉) · F. Xie · N. Yuan · L. Piao
Lab for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China
e-mail: fuzt@pku.edu.cn

N. Yuan
Chinese Academy of Meteorological Science, Beijing 100081, China

method, detrended fluctuation analysis (DFA), has been developed recently (Peng et al. 1994). By calculating the DFA exponent, α (equal to the Hurst exponent H, details can be found in references (Koscielny-Bunde et al. 1998; Peng et al. 1994; Yuan et al. 2012, 2014), one could tell whether the considered time series are characterized by LRC, and how strong the LRC is. When $\alpha = H > 0.5$, the corresponding processes are taken to be long-range positively correlated, when $\alpha = H = 0.5$, no correlation exists, and when $\alpha = H < 0.5$, the corresponding processes are taken to be long-range anti-correlated.

Due to the existence of the two kinds of correlations, variations among consecutive steps are normally interdependent. Therefore, how to quantify their connections is of great importance to predict the future variations based on the historical information. In this study, we will first mainly focus on the relations between two consecutive steps. On small time scales, e.g., the daily scale, it is obvious that yesterday's variation should have strong links with today's state, since they are likely governed by the same weather systems. But on large time scales, e.g., the monthly scale, how the past month's variation influence the present month's state is still unclear and is thus the key issue we aim to address. Therefore, in this work, we will try to reveal the impacts from previous one-step variation on the variation of the present step, when the long-range correlations are considered.

One interesting phenomenon, which may be closely related to the existence of long-range correlation, is the clustering of extreme events (Altmann and Kantz 2005; Bunde et al. 2005). Clustering means an extreme event is more likely to be followed by another extreme event. This kind of clustering behavior has been observed in floods, winter storms, and avalanches in central Europe (Mudelsee et al. 2003). But detailed statistical research on the relations between the clustering behavior and long-range correlation is difficult due to the limited record length (Fuentes et al. 2013). Therefore, our work, which is to study the impacts of previous one-step variation on the variation of the present step, may provide us useful clues to the mechanisms of the clustering phenomenon (Bunde et al. 2005; Yuan et al. 2012).

To study the impacts of previous one-step variation on the variation of the present step in a positively long-range correlated process, we first try to explore their relations theoretically by using the fractional Brownian motion (fBm) (Mandelbrot and van Ness 1968) theory. The increments of fBm (which is fractional Gaussian noise, fGn for short) are long-range correlated, which illustrates that the behavior of a real process after a given time t does not only depend on the situation at t but also of the whole history of the process up to time t (Decreusefond and Üstünel 1999). We find the stronger LRC is, the closer relations between two consecutive steps are. With the information of previous one-step, we can not only calculate the expectation of the present step, but also determine the uncertainty range (standard deviation) quantitatively. By using the concept of marginal distribution, which has been applied to extreme events studies for spatial analysis, we further generalize this new method for the confirmation of the relations between two consecutive steps in a long-range correlated process. Our results are tested by both the artificially generated data and the observational temperature records.

2 Data

In this paper, a modified Fourier filtering technique (Makse et al. 1996) is applied to generate artificial time series with different LRC (the scaling exponents alpha ranges from 0.5 to 0.95). All the generated data are normalized to zero mean and unit standard deviation with a length of 100,000. To reach a reliable numerical statistics, there are 1000 samples in each group with the same scaling exponent.

Since many studies have shown that daily mean surface air temperature variations (Koscielny-Bunde et al. 1998; Yuan et al. 2012, 2014) are positively long-range correlated, one daily mean surface air temperature (SAT) record is chosen in this study. It is observed at Stockholm (59.3° N, 18.1° E), Sweden, from 1756 to 2000, which has been homogenized in the EU-funded IMPROVE project. This time series has been already extensively utilized, see Qian et al. (2009) and references therein. This temperature record is chosen for our research because it has long histories (more than 200 years) with fewer missing points. Before our analysis, we first standardize the data to generate biweekly data by (a) removing the seasonal trend through subtracting the annual cycle, as $T_{\mathrm{di}} - \langle T_{\mathrm{di}} \rangle$ (Koscielny-Bunde et al. 1998), where T_{di} is the daily temperature and $\langle T_{\mathrm{di}} \rangle$ is a long-time climatological average for each calendar day. Note that this procedure can not remove slow background trends such as a gradual shift of annual mean temperature (Bartos and Janosi 2005); and (b) averaging the daily anomaly data over 2-week-long non-overlapping windows to remove the short-term correlations due to general weather regimes; (c) normalizing biweekly temperature anomaly series $\{T_i\}$ by subtracting sequence mean and dividing by sequence standard deviation, $T_i' = (T_i - \langle T_i \rangle)/\sigma(T_i - \langle T_i \rangle)$, and $\{T_i'\}$ are series that we use for analysis. At the same time, 50-day averaged data can be generated by the same way.

In order to find long-range correlation in the original biweekly temperature anomaly variations, surrogate procedures such as shuffling and phase randomize surrogate

(PRS) (Govindan et al. 2007) have been applied to this normalized biweekly temperature anomaly series from Stockholm to generate surrogated data with 1000 samples for each group. These surrogated series will be analyzed and compared with original temperature anomaly series to investigate effect of long-range correlation on the conditional calculations and marginal distribution of one-step ahead variations.

3 Theoretical results

3.1 Conditional calculations

Fractional Brownian motion (fBm) offers a convenient modeling for non-stationary processes with long-term correlations (Mandelbrot and van Ness 1968). The increments of fBm, which are fGn, are stationary and long-range correlated with parameter H. Considering an fBm sequence $B_H(t) = \sum_{i=1}^{t} y(i)$, where $y(i+1) = B_H(i+1) - B_H(i)$ is a long-range-correlated fGn with Gaussian distribution, the non-stationary character of fBm is proved by its covariance structure:

$$E[B_H(t)B_H(\tau+t)] = \frac{\sigma^2}{2}\left\{t^{2H} + (\tau+t)^{2H} - \tau^{2H}\right\} \quad (2)$$

where $E[]$ stands for expectation operator. Without loss of generality, we set $B_H(0) = 0$, then

$$\frac{E[B_H(t+\tau)|B_H(t)]}{B_H(t)} = \frac{E[B_H(t+\tau)B_H(t)]}{E[B_H^2(t)]} = \frac{t^{2H} + (\tau+t)^{2H} - \tau^{2H}}{2t^{2H}} \quad (3)$$

Considering fGn is the increment of fBm, we have

$$E\left[\sum_{j=t+1}^{t+\tau} y(j)|\sum_{i=1}^{t} y(i) = B_H(t)\right] = \frac{(t+\tau)^{2H} - t^{2H} - \tau^{2H}}{2t^{2H}}\left(\sum_{i=1}^{t} y(i)\right) \quad (4)$$

with memory kernel defined by $K(t,\tau,H) = \frac{(t+\tau)^{2H} - t^{2H} - \tau^{2H}}{2t^{2H}}$, and memory kernel is a factor to quantify the impact of historical information.

Considering the simplest case, when $t = 1$ and $\tau = 1$, from Eq. 4, we have

$$E[y(2)|y(1)] = (2^{2H-1} - 1)y(1) \quad (5)$$

which means the conditional mean value $E[y(2)|y(1)]$ is totally determined by memory kernel $K(1,1,H)$ if $y(1)$ is given. While the memory kernel $K(1,1,H)$ is determined totally by only one parameter, Hurst exponent H. When $H = 1/2$, which corresponds to white noise process and no memory exists, $K(1,1,1/2) = 0$, which means that historical information has no no effect on determining the expectation value of next-step variation. When $H > 1/2$, $K(1,1,H) > 0$, which means when $y(1)$ takes a positive value, then $y(2)$ will has a great chance to take a positive value. This fact also explains why an extreme event in a positively long-range correlated process usually follows by another extreme event, which leads to the clusterings of extreme events.

This theoretical result can also be confirmed by the numerical calculations, see Fig. 1a, b. For each H, $E[y(2)]/y(1)$ is all constant and does not vary with $y(1)$. The value of $E[y(2)]/y(1)$ is totally determined by Hurst exponent H, with increasing H, the value of $E[y(2)]/y(1)$

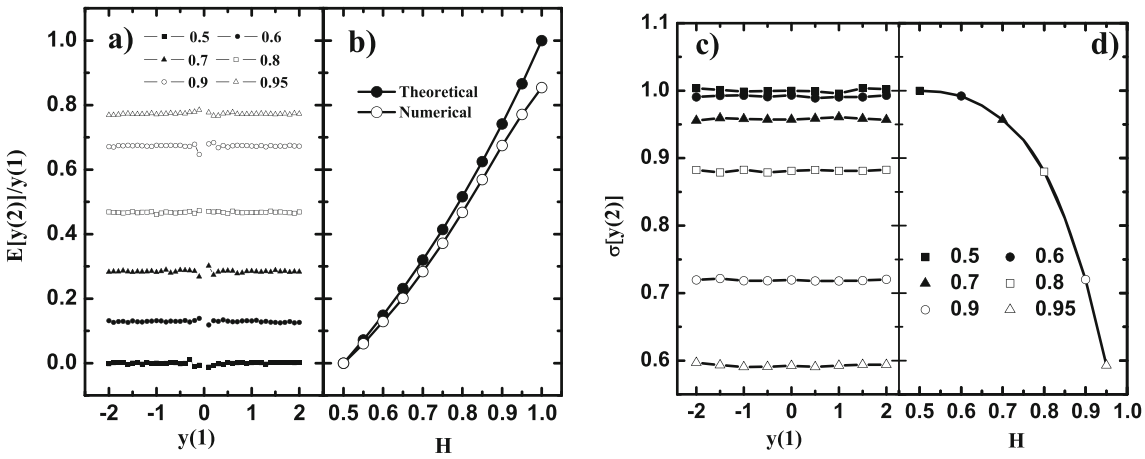


Fig. 1 Theoretical results for conditional calculations. **a** The variation $E[y(2)]/y(1)$ with $y(1)$. **b** The comparison between theoretical and numerical variation $E[y(2)]/y(1)$ with different H. **c** Numerical results for conditional standard deviation $\sigma[y(2)]$ with $y(1)$ from fGn processes. **d** The variation $\sigma[y(2)]$ with different H

Table 1 Standard deviation, theoretical and numerical memory kernel from different H

H	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
$\sigma[y(2)]$	1.0	0.998	0.992	0.978	0.957	0.927	0.880	0.811	0.720	0.593
$K_{\text{Theor.}}$	0	0.072	0.15	0.23	0.32	0.41	0.52	0.62	0.74	0.87
$K_{\text{Num.}}$	0	0.060	0.13	0.20	0.28	0.37	0.47	0.57	0.67	0.77

increases. From Fig. 1b, we can also find that with increasing value of H, deviation of numerical results from the theoretical results increases, meanwhile the total deviation is not large, see Table 1. The reason for this deviation is that the interdependence range will enlarge with memory intensity increasing, more than one step variations will affect the future variation in a positively long-range correlated process.

Apart from the expectation of conditional mean value $E[y(2)|y(1)]$ can be derived, the expectation of conditional standard deviation value $\sigma[y(2)|y(1)]$ can also be determined numerically, see Fig. 1c, d. It shows that the value of $\sigma[y(2)]$ is determined totally by H. With increasing H, the value of $\sigma[y(2)]$ will decrease greatly.

Combining the results from conditional mean and standard deviation of $y(2)$, we find the expectation of $y(2)$ will shift to larger value as the increase of memory strength but with a much narrower band, see Figs. 1b, d. This will be beneficial for our prediction, especially for one-step-ahead prediction.

3.2 Marginal distribution

Marginal distribution has been applied to extreme event studies, such as spatial dependencies of heavy rainfall (Bernard et al. 2013) and spatial modeling of extreme temperature (Fuentes et al. 2013). In these studies, extreme value theory (EVT) has been indirectly used to model spatial structure of climate extremes (Coles and Tawn 1996; Naveau et al. 2005), where it's supposed that the marginal distribution of $M(x)$ is known with $M(x)$ to be maximum value of measured variable recorded at the location x in a stationary field. Without loss of generality, the marginal distribution can be assumed as

$$P(\mu) = P(M(x) < \mu) = \exp(-1/\mu) \tag{6}$$

with μ is a given threshold value. And in bivariate case, it can be rewritten as

$$P(M(x) < \mu, M(x+h) < \mu) = \exp(-\theta(h)/\mu) = P(u)^{\theta(h)} \tag{7}$$

where $\theta(h)$ is extremal coefficient, and h is a parameter used to measure spatial distance between observational stations. It should be noted that, the derivation of Eqs. 6 and 7 have been taken some assumptions; details about these derivations and assumptions can be found in the references, such as Bernard et al. (2013), Coles and Tawn (1996), Fuentes et al. (2013), and Naveau et al. (2005).

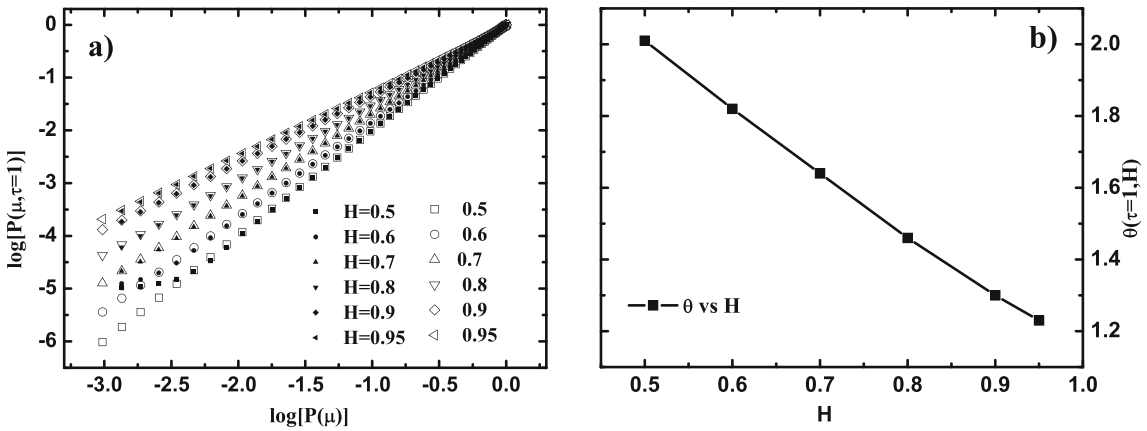


Fig. 2 Theoretical studies of marginal distribution. **a** Marginal distribution $\log[P(\mu, \tau = 1)]$ as distribution function $\log[P(\mu)]$. The *hollow symbols* are theoretical results for different H and the *solid symbols* are numerical results from fGn processes with different H. **b** The variation of $\theta(\tau = 1, H)$ with different H

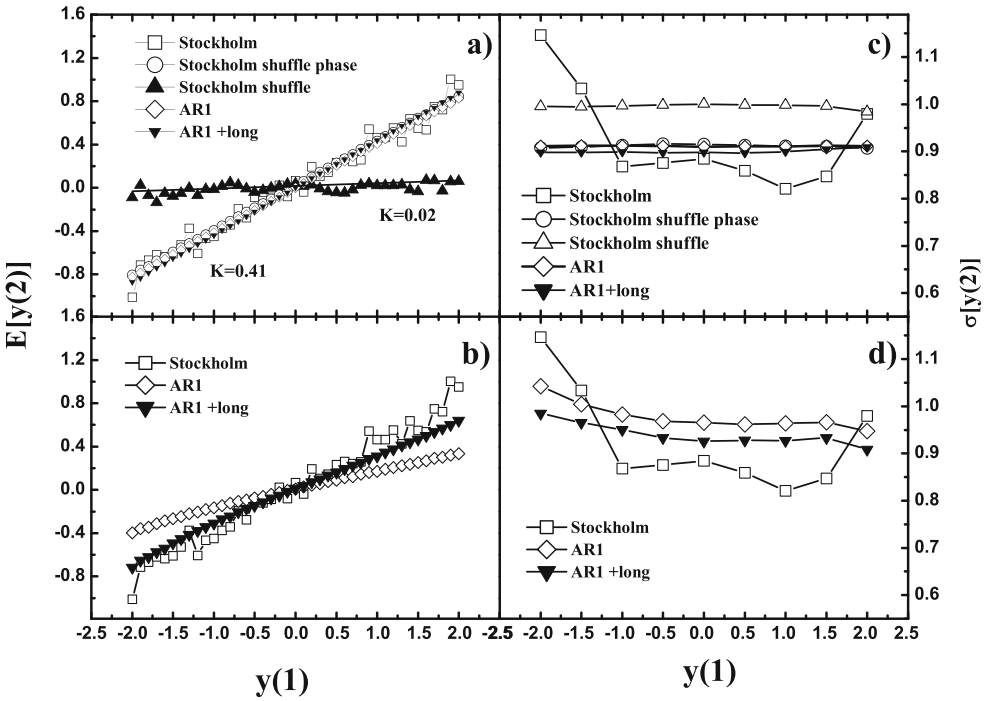


Fig. 3 Conditional calculations for observational temperature anomaly series for Stockholm. **a**, **b** Conditional expectations, i.e., the variation $E[y(2)|y(1)]$ with $y(1)$. **c** and **d** Conditional standard deviation of $\sigma[y(2)|y(1)]$ varies with $y(1)$. **a** and **c** Original observations and its surrogates and fitted by AR(1) process and AR(1) plus long-range correlated noise process. **b** and **d** Original observations and its one-step ahead prediction based on AR(1) process and AR(1) plus long-range correlated noise process. Where *hollow squares* are for original series, *hollow circles* are for PRS surrogated series, *solid up-triangles* are for shuffling series, rhombuses for AR(1) process and *solid down-triangles* for AR(1) plus long-range correlated noise process

Equation 7 can be borrowed to build the relationship between marginal distribution of $y(\tau + 1)$ with distribution of given $y(1)$, i.e.,

$$P(\mu, \tau) = P(\{y(\tau + 1)\} \leqslant \mu | \{y(1)\} \leqslant \mu) \tag{8}$$

Equation 8 estimates probability of $y(\tau + 1) \leqslant \mu$ when $y(1) \leqslant \mu$ is given, and it can be assumed to take the following formula

$$P(\mu, \tau) = P(\mu)^{\theta(H,\tau)} \tag{9}$$

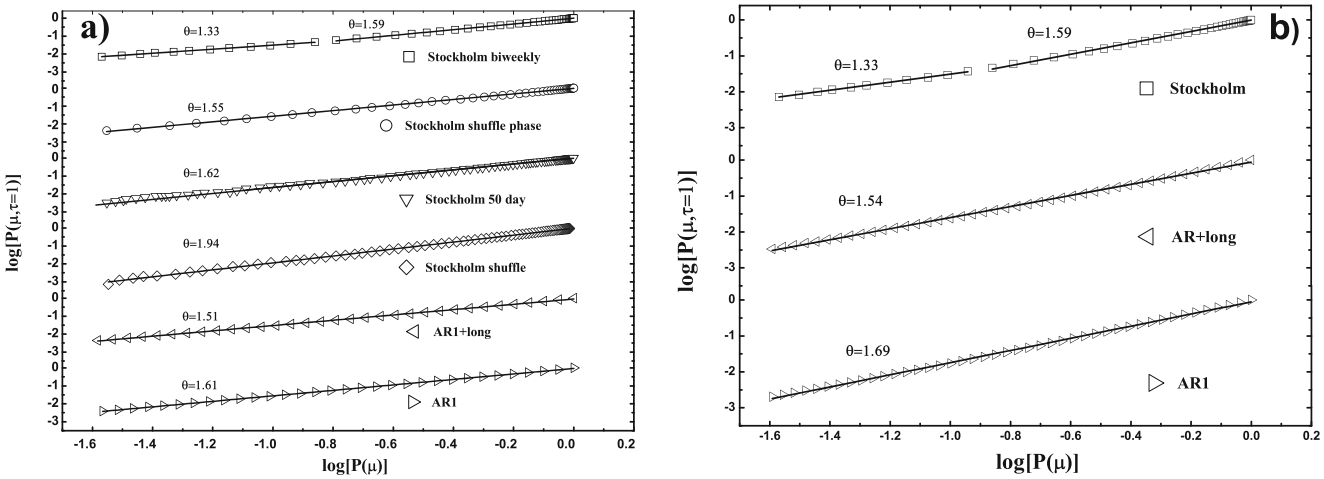


Fig. 4 Marginal distribution for observational temperature anomaly series for Stockholm. **a** is for the original observations and its surrogates and fitted by AR(1) process and AR(1) plus long-range correlated noise process. **b** is for the original observations and its one-step-ahead prediction based on AR(1) process and AR(1) plus long-range correlated noise process. The *squares* denote the results from biweekly averaged original series, circles for PRS surrogated series from biweekly averaged original series, *down-triangles* for the results from 50-day averaged original series, rhombuses for shuffling series from biweekly averaged original series, *left-triangles* for AR(1) plus long-range correlated process and right-triangles for AR(1) process, respectively. *Solid lines* with slope value of $\theta(1, 1, H)$ are fitted over certain ranges

where $P(\mu) = P(\{y(1)\} \leqslant \mu) = \int_{-\infty}^{\mu} f(y)dy$ is cumulative distribution function (CDF) for distribution function $f(y)$ of y, and $\theta(H, \tau)$ is marginal distribution coefficient. If $y(1)$ at time $t = 1$ is independent to $y(\tau + 1)$ at time $t = \tau + 1$, then conditional distribution coefficient $\theta(\tau, H)$ is 2.

In the above subsection, it has been found that historical information can affect not only the expectation of next step variation, but also, its variance for fGn process. Using these results, Eq. 9 can be rewritten as

$$P(\mu, \tau = 1) = \int_{-\infty}^{\mu} f(y)dy \int_{-\infty}^{\mu} f\left(\frac{z - Ky}{\sigma}\right)dz \qquad (10)$$

where K is memory kernel and σ is standard deviation.

If $f(y)$ is a Gaussian distribution function, then Eq. 10 can be simplified as

$$P(\mu, \tau = 1) = \int_{-\infty}^{\mu} \Phi(\frac{\mu - Ky}{\sigma})d\Phi(y) \qquad (11)$$

with $\Phi(y)$ being a cumulative normal distribution function for y. Eq. 11 can be solved numerically to derive the theoretical results for Gaussian processes, see Table 1 for the standard deviation σ and memory kernel K.

To verify the above generalization, we use artificially generated sequences which has been introduced in Section 2 to test the simplest case, i.e., $\tau = 1$, the results are shown in Fig. 2. We can see that both theoretical and numerical results indicate a power law relation between $P(\mu)$ and $P(\mu, \tau = 1)$ (Fig. 2a). For each H, the numerical results nearly collapse to the theoretical ones over wide ranges. The conditional distribution coefficient are 2.01 ($H = 0.5$), 1.82 ($H = 0.6$), 1.64 ($H = 0.7$), 1.48 ($H = 0.8$), 1.30 ($H = 0.9$), and 1.23 ($H = 0.95$) for (Fig. 2b). These results demonstrate the generalization and assumption (8) is suitable to bridge the connection between $y(2)$ and $y(1)$ in a positively long-range correlated process. At the same time, the theoretical results of Gaussian distribution function indicate $\{y(2)\}$ admits a Gaussian distribution, the expectation $E[y(2)]$ and standard deviation $\sigma(y(2))$ are fully enough to describe the behaviors of $y(2)$.

4 Confirmation in observational analysis

In the abovementioned section, we discuss conditional calculations and marginal distribution of $\{y(2)\}$ by employing theoretical derivation and numerical calculations of idealized artificially generated sequences when $y(1)$ is given. In this section, we will check whether these results can be obtained from the observational record.

First of all, we have analyzed the conditional calculations $\{y(2)\}$ with given $y(1)$ in the recorded temperature variations, and this can be found in Fig. 3a. For the original sequences, $E[y(2)]$ will change linearly with $y(1)$, its slope, i.e., $K(1, 1, H)$, is 0.41. For shuffling case where long-range correlations will be lost when shuffling procedure is carried on the original series, the expectation of $y(2)$ doesn't depend on the variation of $y(1)$. All these results are in line with what has been found in above theoretical results. As for the standard deviation (see Fig. 3c), just like the theoretical results, $\sigma[y(2)]$ keeps invariant over wide ranges as $y(1)$ changes for both original and surrogated series, and much large deviation from the constant behavior found for larger $y(1)$ is caused by finite-size sample effect, and similar but smaller fluctuations emerged for original series result from finite-size sample effect, either. The similar calculations for the fitted AR(1) processes to the observational series show that both expectations $y(2)$ and the standard deviation $\sigma[y(2)]$ from the fitted AR(1) processes nearly collapse to these from the observational series (see Fig. 3a, c). Secondly, we check the marginal distribution of $\{y(2)\}$ with given $y(1)$ in the recorded temperature variations, this can be found in Fig. 4a. There are two power law-scaling ranges in the original series, which are different from what have been found in theoretical results. The reason for this cross scaling behavior is the averaging window is too narrow to smooth off the small scale motions. When the window size is enlarged, this kind of behavior will disappear, see Fig. 4a for results of 50-day window averaged series, the power law behavior holds over the whole ranges, and similar behaviors can be found for surrogated series in Stockholm. When no memory exists, which corresponds to the shuffling case in Fig. 4a , θ approaches the limit 2 as expected, a little bit deviation from the limit 2 for Stockholm maybe results from the deviation from Gaussian distribution of temperature anomalies found in Fig. 5.

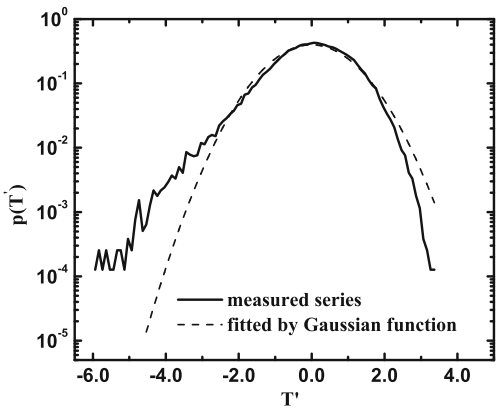


Fig. 5 Probability distribution of temperature anomaly variations over Stockholm. The *solid lines* are from observational series and *dash lines* are fitted by Gaussian distribution function (Dumouchel 1973)

The fitted AR(1) process leads to nearly the same results as the from the original biweekly measured series on the larger scales.

5 Conclusion and discussion

In this paper, fBm and fGn processes have been applied to theoretically investigate conditional mean and standard deviation of one-step-ahead variations $\{y(2)\}$ when the previous one-step variation $\{y(1)\}$ has been given. It is found that $E[y(2)]$ changes linearly with $y(1)$ and its slope, i.e., memory kernel $K(1, 1, H)$, which is totally determined by the memory strength H. At the same time, numerical calculations from artificially generated series with different memory strength H confirm the results above. The standard deviation $\sigma[y(2)]$ doesn't depend on $\{y(1)\}$, but its value is determined by the memory strength H, with the memory strength H increasing, the numerical value of standard deviation $\sigma[y(2)]$ will drop down dramatically. The conditional connections between $\{y(2)\}$ and $\{y(1)\}$ can also be obtained in the marginal distribution of $\{y(2)\}$ for given $\{y(1)\}$, a power law relationship between $P(\mu, \tau = 1)$ and $P(\mu)$ being revealed. The theoretical results of fGn processes collapse very well in numerical analysis from artificially generated series with different memory strength H over wide ranges, and the conditional distribution coefficient $\theta(\tau = 1, H)$ decreases with the increase of memory strength H.

These theoretical results are well confirmed in the one longest observational temperature anomaly series from Stockholm. Comparison between theoretical results and observational analysis is demonstrated in Table 2. Since the memory strength of temperature anomaly variation in Stockholm is $H = 0.74$ (see Fig. 6), very close to $H = 0.75$, then we have theoretical results of $H = 0.75$ to compare with those from Stockholm. It is obvious that there are nearly perfect concordances in results between theoretical and observational calculations over Stockholm, where $K_{\text{Theor.}} = 0.41$ and $K_{\text{Stockholm}} = 0.41$, $\sigma_{\text{Theor.}} = 0.93$ and $\sigma_{\text{Stockholm}} = 0.91$, $\theta_{\text{Theor.}} = 1.54$ and $\theta_{\text{Stockholm}} = 1.55$. And it should be noted that in order to reach reliable statistics, both $\sigma_{\text{Stockholm.}} = 0.91$ and $\theta_{\text{Stockholm.}} = 1.55$ are calculated from PRS surrogated series. All results are calculated from the normalized biweekly series, original or surrogated. Actually, it can be found in Fig. 4a that different scaling behaviors emerge when biweekly series are analyzed, which may affect the conditional calculations and marginal distribution of $\{y(2)\}$ when $\{y(1)\}$ has been given. Further investigations show that larger averaged windows will remove cross scaling behavior, there are no cross scaling behaviors when the size of averaged window is 50 days for Stockholm, see Fig. 4a.

Table 2 Comparison of H, K, σ, and θ between theoretical and observational calculations over Stockholm

	H	K	σ	θ
Theor.	0.75	0.41	0.93	1.54
Stockholm	0.74	0.41	0.91	1.55

From the above-mentioned analysis, we can see the fitted AR(1) process seems to be able to reach the same results as those from the observational record over Stockholm. However, the fitted AR(1) process can not capture the long-range correlation found in the measured series. The long-range correlation of temperature anomaly variations can be quantified by detrended fluctuation analysis (DFA for short) method, which has been originally proposed by Peng et al. (1994). DFA results (with DFA 2 Kantelhardt et al. (2001)) have been shown in Fig. 6. We can see that the memory strength found in the fitted AR(1) process approaches to the no memory case $H = 0.5$, but the memory strength of temperature anomaly variation in Stockholm is $H = 0.74$. Here, we can see that the fitted AR(1) process may not capture the exact information hidden in the measured series. In order to find these information, AR(1) model (1) should be revised as

$$x_i = bx_{i-1} + \eta_i, \quad i = 2, 3, 4 \cdots \tag{12}$$

where $\{\eta_i\}$ is long-term correlated noise with zeros mean and unity variance, this model was first proposed by Király and his coauthor Király and Jánosi (2002).

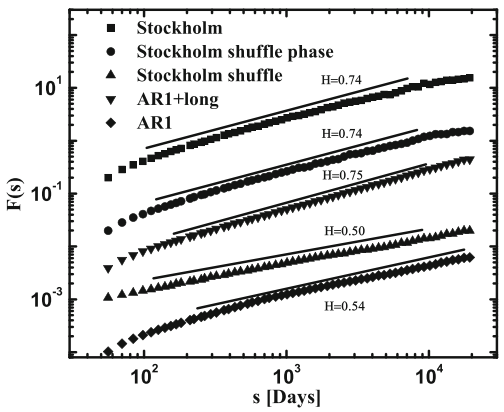


Fig. 6 DFA results from observational temperature anomaly series for Stockholm by DFA 2. The *squares* denote the results from original series, *circles* from PRS surrogated series, *up-triangles* from shuffling series, *down-triangles* and rhombuses for AR(1) plus long-range correlated process and AR(1) process fitted to biweekly averaged original series, respectively. *Solid lines* with slope value of H are fitted over certain ranges

It should be noted that here, b in Eq. 12 is related to a in Eq. 1. If we fit the model (1) to biweekly mean temperature series over Stockholm, we can derive that $a = 0.412$. However, if we take long-term correlated noise into account, i.e., we choose the model (12) to fit biweekly mean temperature series over Stockholm, we can find the short-term correlation decreases sharply, $b = 0.07$ for Stockholm. The benefit of adopting the model (12) is firstly given by its capturing the long-range correlation over the larger scale ranges, see Fig. 6; here, DFA shows the same results for the model series and the original ones over the larger scale ranges. At the same time, the expectation and standard deviation analysis for one-step ahead statistics show that the model (12) works a little bit better than the model (1), see Figs. 3a, c and 4a. Especially, we can see standard deviation of one-step ahead variations $\{y(2)\}$ when the previous one-step variation $\{y(1)\}$ has been given is much closer to the original observational results, see Fig. 3c.

Secondly, the theoretical and observational results on conditional calculations and marginal distribution of $\{y(2)\}$ with given $\{y(1)\}$ are of great importance to prediction, especially to one-step-ahead prediction, since there are extremely close relationships between $\{y(2)\}$ and $\{y(1)\}$. Of course, interdependence among variations is not limited to two successive steps in positively long-range correlated processes. So, if we use the model Eqs. 1 and 12 to make one-step-ahead prediction, i.e., we take each x_{i-1} in Eqs. 1 and 12 from the observational records and apply the model Eqs. 1 and 12 to calculate each one-step-ahead predicted value of x_i. Here, all the model parameters take the same values as the fitted ones to the observational biweekly temperature anomalies over Stockholm. We take the same calculations as given in the above section, the results are shown in Figs. 3b, d and 4b. It is obvious that both the conditional mean and standard deviation of one-step ahead variations $\{y(2)\}$ and marginal distribution for one-step ahead variations $\{y(2)\}$ for the given previous one-step variation $\{y(1)\}$ are sharply distinguished between the prediction based on AR(1) model (1) and those from the modified model (12). Compared to the results from AR(1) model (1), the results from the modified model (12) are much closer to those from the observational records. Although the fitted models, AR(1) model (1) or the modified model (12) can capture nearly the same conditional statistics as the observational records (see Figs. 3a, c and 4a.), the predicted results are sharply distinguished between the prediction based on AR(1) model (1) and those from the modified model (12). Causing these differences is from the basic difference that AR(1) model (1) can not capture the long-term memory found in the observational records, but the modified model (12) can, see Fig. 6. The long-term memory will make the variabilities much predictable, the stronger LTM is, the larger proportion the predictable signals will account for in the whole variations (Yuan et al. 2013; 2014).

Acknowledgments Many thanks are due to valuable suggestions from editors and anonymous reviewers and supports from National Natural Science Foundation of China (no. 41175141).

References

Altmann EG, Kantz H (2005) Recurrence time analysis, long term correlations, and extreme events. Phys Rev E 71:056106

Bartos I, Janosi IM (2005) Atmospheric response function over land: strong asymmetries in daily temperature fluctuations. Geophys Res Lett 32:L23820

Bernard E, Naveau P, Vrac M (2013) Clustering of maxima: spatial dependencies among heavy rainfall in France. J Clim 26: 7929

Bunde A, Eichner JF, Kantelhardt JW, Havlin S (2005) Long-term memory: a natural mechanism for the clustering of extreme events and anomalous residual times in climate records. Phys Rev Lett 94:048701

Coles SG, Tawn JA (1996) A Bayesian analysis of extreme rainfall data. Appl Statist 45:463–478

Decreusefond L, Üstünel AS (1999) Stochastic analysis of the fractional Brownian motion. Potential Anal 10:177–214

Dumouchel WH (1973) On the asymptotic normality of the maximum likelihood estimate when sampling from a stable distribution. Ann Stat 1:945–567

Fuentes M, Henry J, Reich B (2013) Nonparametric spatial models for extremes: application to extreme temperature data. Extremes 16:75–101

Govindan RB, Wilson JD, Preiβl H, Eswaran H, Campbell JQ, Lowery, CL (2007) Detrended fluctuation analysis of short datasets: an application to fetal cardiac data. Physica D 226:23

Ivanova K, Ausloos M (1999) Application of the detrended fluctuation analysis (DFA) method for describing cloud breaking. Physica A 274:349–354

Kantelhardt JW, Koscielny-Bunde E, Rego HHA, Havlin S (2001) Detecting long-range correlations with detrended fluctuation analysis. Physica A 41:295

Király A, Jánosi IM (2002) Stochastic modeling of daily temperature fluctuations. Phys Rev E 65:051102

Koscielny-Bunde E, Bunde A, Havlin S, Roman HE (1998) Indication of a universal persistence law governing atmospheric variability. Phys Rev Lett 81:729

Makse HA, Havlin S, Schwartz M, Stanley HE (1996) Method for generating long-range correlations for large systems. Phys Rev E 53:5445

Mandelbrot BB, van Ness JW (1968) Fractional Brownian motions fractional noises and applications. SLAM Rev 10:422

Mantegna RN, Stanley HE (1995) Scaling behaviour in the dynamics of an economic index. Nature 46:376

Mudelsee M, Börngen M, Tetzlaff G, Grünwald U (2003) No upward trends in the occurrence of extreme floods in central Europe. Nature (London) 425:166

Naveau P, Nogaj M, Ammann C, Yiou P, Cooley D, Jomelli V (2005) Statistical methods for analysis of climate extremes. C R Geoscience 337:1013–1022

Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994) Mosaic organization of DNA nucleotides. Phys Rev E 49:1685

Peng CK, Havlin S, Stanley HE, Goldberger AL (1995) Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 5:82

Qian C, Fu CB, Wu ZH, Yan ZW (2009) On the secular change of spring onset at Stockholm. Geophys Res Lett 36: L12706

Yuan NM, Fu ZT, Li HQ, Mao JY (2012) Effect of extreme value loss on long-term correlated time series. Theor Appl Climot 109:133

Yuan NM, Fu ZT, Shida Liu (2013) Long-term memory in climate variability: a new look based on fractional integral techniques. J Geophys Res 118(12):12962–12969

Yuan NM, Fu ZT, Shida Liu (2014) Extracting climate memory using fractional integrated statistical model: a new prospective on climate prediction. Scientific Rep 4:6577