



Regional contrasting DTR's predictability over China

Shu Fu ^{a,1}, Yu Huang ^{b,1}, Tao Feng ^{a,1}, Da Nian ^b, Zuntao Fu ^{b,*}

^a Department of Atmospheric Science, Yunnan University, Kunming, 650091, China

^b Lab for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, 100871, China

HIGHLIGHTS

- Both intrinsic and realizable predictability are studied on DTR fluctuations.
- Well defined regime-dependent relation between two predictabilities exists.
- Overall South–North asymmetric predictability is revealed in DTR over China.

ARTICLE INFO

Article history:

Received 6 November 2018

Received in revised form 16 January 2019

Available online 22 January 2019

Keywords:

Diurnal temperature range

Intrinsic predictability

Realizable predictability

Permutation entropy

Forecasting error

ABSTRACT

Predictability is an important topic in weather and climate studies. Model-free and model-based quantification of diurnal temperature range (DTR) time series predictability is provided in this paper. Both intrinsic predictability quantified by the permutation entropy and realizable predictability defined by the model's forecasting error are studied on DTR fluctuations. The intrinsic predictability tells us whether there are more predictive structures in DTR fluctuations and to what extent the maximal predictability can be reached. The realizable predictability measures the degree of prediction to which adopted model reaches. Results show that there is a well defined regime-dependent pattern between the intrinsic predictability and the realizable predictability in DTR fluctuations, and both the intrinsic predictability and the realizable predictability in DTR fluctuations over China are overall South–North (with the dividing line around the latitude of 37°N and almost along the Yellow River, and south and north to this line are called southern China and northern China in this study) asymmetric with higher predictability in southern China and weaker predictability in northern China. This South–North contrasting predictability behavior is closely related to the DTR's South–North asymmetric multi-fractal behavior.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

It has been widely recognized that the study of predictability is multifaced and related to diverse fields [1]. There are different definitions for the predictability in different fields, including mathematics, human–computer interaction, human sentence processing, biology, popular culture, macroeconomics, climate and so on. From Wikipedia, the definition of the predictability is the degree to which a correct prediction or forecast of a system's state can be made either qualitatively or quantitatively. Therefore, predictability is firstly intrinsic to a system, and it can be called the intrinsic predictability [1–3]. This intrinsic predictability is distinct from those from any model, which is taken as the realizable predictability [2,3]. Both

* Corresponding author.

E-mail address: fuzt@pku.edu.cn (Z. Fu).

¹ These authors contributed equally.

intrinsic predictability and realizable predictability can be quantified by the measures estimated from the outputs of any given system [2–21]. It is suggested that the intrinsic predictability [2,3] can be approximated by data-driven measures of time series structured information, such as mean prediction time [4], fractal dimension [5], Lyapunov exponents and improved Lyapunov exponents [11,12,14,17–20], memory or persistence [13,15,21] and entropy [6,10] or permutation entropy (PE) [22–24]. While realizable predictability is often quantified by means of the correlation coefficient between observations and predictions or forecasting error (FE) measures [7,16]. Previous studies also show that there is a well defined correlation between PE and FE [2,3]. If there is any structured information, deterministic or stochastic, in the time series, and then the estimated PE from any given time series will decrease and corresponding estimated FE may also decrease. Well defined structures hidden in any given time series can enhance both intrinsic predictability and realizable predictability [9].

However, both the conjecture that PE can estimate the intrinsic predictability and its well defined relation to FE are only validated in a limited range of outputs of idealized models or empirical observations. As one of the most important predictability study fields, time series from atmospheric system has not been carefully checked. Will the predictive structure in atmospheric time series be effectively quantified by PE? Could the well defined relationship between PE and FE be recovered in atmospheric time series? Previous studies show that there are different ordinal patterns in daily mean temperature (Tmean) variations [25–28], such as rapid cooling and gradual warming at the mid-latitudes [28,29]. Apart from Tmean, there are other temperature variables, such as maximum temperature (Tmax), minimum temperature (Tmin) and diurnal temperature range (DTR, $DTR = Tmax - Tmin$) [30–35]. These temperature variables may have different variation's features from those of Tmean. At the same time, studies show that PE–FE relation suffers from the nonlinearity of underlying dynamics [2,3]. Among above-mentioned different air temperature variables, DTR has dominant regional nonlinearity disparities and it was found that there are distinguished mono-fractal and multi-fractal behaviors for DTR over northern and southern China with a dividing line around 30°N and almost along Yangtze River [36]. Will DTR's contrasting fractal behaviors over northern and southern China affect the DTR's predictability over these regions of China? Therefore, in this paper, DTR will be chosen to answer above raised questions and we also try to find whether there exist the DTR's regional predictability disparities over China.

The rest of our paper is arranged as follows. In Section 2, the data source used in this paper will be briefly described. Model to reach the predictions, measures for realizable predictability-FE and intrinsic predictability-PE are outlined in Section 3. Detailed results about descriptive and quantitative features for FE and PE will be presented in Section 4. At last, conclusions given in this paper with some discussions are shown in Section 5.

2. Data

In this paper, homogenized direct observations [37] of daily maximum and minimum temperatures (Tmax and Tmin) were downloaded from the China meteorological data sharing service system (<http://data.cma.cn/en>). And records over 194 meteorological stations taking part in international exchanges were chosen for further Analysis. Since there are consecutive missing points in records over some stations and the beginning times over some stations are later than those over most of stations, missing points and different length sizes will result in biased results. For the consistent derived results and reliable statistics, only records over 185 stations with duration from 1970 to 2013 without missing points were chosen as the major analysis. At the same time, records over 157 stations with longer duration from 1960 to 2015 without missing points taken as validated check were also analyzed in this paper. For records over each station, diurnal temperature range (DTR) can be calculated in each day from Tmax and Tmin by means of $DTR = Tmax - Tmin$.

Suffering from the sun radiation, there are dominated annual cycle or slow changing variations in DTR variations. Actually, these slowly varying periodic variations can give distorted calculating for the statistical measures of fast fluctuations [38]. Since the calculations and analysis in this paper are more involved in local variations [29], these long time-scale variations (such as seasonal trend) were eliminated by the following operations, $DTR'_i = DTR_i - \langle DTR_i \rangle$ [39], where DTR_i is any given daily DTR, $\langle DTR_i \rangle$ is its long-time climatological average for each calendar day, and DTR'_i is called DTR fluctuations. In the following parts, all calculations are based on DTR fluctuations.

3. Methods

3.1. ARFIMA(p, d, q) model and parameters' determination

Predictability studies are involved in model-based and model-free (data driven) ways. It is also generally accepted in the literature that fluctuations of many atmospheric variables can be well modeled by linear processes or models, such as autoregressive (AR) processes [40–42]. In this paper, autoregressive fractionally integrated moving average (ARFIMA) model [43] is chosen to model the DTR fluctuations. And this model is often written as ARFIMA(p, d, q) with p the order of the autoregressive model, q the order of the moving-average model, and d related to the degree of memory (it can take the fractional values). d can be determined by $d = \alpha - 0.5$, where α is long-termed memory (LTM) exponent and it can be estimated by detrended fluctuation analysis (DFA) [39,44–47] method from a given time series. For a given time series x_i , $i = 1, 2, 3, \dots, N$ with data length of N , steps of DFA algorithm can be briefly outlined as follows: (1) calculate its profile: $y_j = \sum_{i=1}^{i=j} x_i - \bar{x}_i$, with the mean $\bar{x}_i = \frac{1}{N} \sum_{i=1}^{i=N} x_i$; (2) for each changing time scale s , divide N into M intervals ($M = \lfloor N/s \rfloor$, $\lfloor \rfloor$ is rounding down calculation) and fit local trend y_t within each interval to reach the detrended profile: $y'_j = y_j - y_t$; (3)

calculate fluctuation function: $F^2(s) = \frac{1}{M} \sum_{k=1}^{k=M} (\frac{1}{s} \sum_{j=1+(k-1)s}^{j=ks} y_j^2)$; (4) estimate DFA exponent α according to the scaling law: $F(s) \propto s^\alpha$. The other two parameters p and q can be determined by the least squares fitting, since the goal of this study is not to derive the optimal model to DTR fluctuations, we set $p = 3$ and $q = 0$ thorough this paper. Actually, the different choice of p and q will not alter the results qualitatively at all (figures not shown here).

It should be noted that the whole DTR fluctuation series were applied to estimate the LTM exponent α and the whole DTR fluctuation series over each station would be divided into two parts of equal data length, one was taken as training series to fit the ARFIMA(p, d, q) model (determining coefficients in AR(3)) and the other was testing series, which would be applied to make one-step ahead prediction with the help of fitted ARFIMA(p, d, q) model.

3.2. Measure on the realizable predictability: forecasting error

The model forecasting error (FE) is the usual measure to whether prediction is successful or not for a specific model, which is taken as the realizable predictability of a specific model to a given underlying process [3]. At the same time, residuals in regression analysis have been applied to check the performance and applicability of fitted model to the underlying analyzed series [48]. So, in this paper, residuals e_i between the testing series DTR'_i and its predicted counterparts \widehat{DTR}'_i are taken to define FE as

$$FE = \frac{\sum_{i=1}^{i=n} e_i^2}{\sum_{i=1}^{i=n} E_i^2} \quad (1)$$

with

$$e_i = DTR'_i - \widehat{DTR}'_i, \quad i = 1, 2, \dots, n \quad (2)$$

where n is the data length of the testing series DTR'_i , and E_i is Gaussian white noise with zero mean and unit standard deviation. From Eq. (1), it is obvious that for a signal with perfect predictive structures, such as sinusoidal signal, the value of FE equals to zero; for a signal with no predictive structure, such as Gaussian white noise with zero mean and unit standard deviation, the value of FE equals to 1.

3.3. Measure on the intrinsic predictability: permutation entropy

It has been pointed out that the system's intrinsic predictability (the highest achievable predictability) is closely related to its system dynamics and it can be quantified by permutation entropy (PE) [22], a model-free, information-theoretic measure of the complexity of a time series [2,22,23,49].

For a given time series outputted from any system $x_i, i = 1, 2, 3, \dots, N$ with data length of N , ordinal patterns within a given fixed window, such as $D = 3, 4, 5, 6 \dots$, can be applied to symbolize this series into corresponding permutation sequence with possible $D!$ kinds of permutations. For each permutation π_j , its relative frequency of occurrence for the whole time series can be derived as $p(\pi_j)$ (see [22] for details) and corresponding PE can be defined as

$$PE = -\frac{1}{\log D!} \sum_{j=1}^{j=D!} p(\pi_j) \log p(\pi_j) \quad (3)$$

and for PE from Eq. (3), its maximal value is 1 (for totally stochastic signals, such as white noise) and minimal value is zero (for totally deterministic signals, such as sinusoidal signal). If more predictive structures are hidden in time series, the value of PE from the corresponding time series will decrease greatly.

4. Results

4.1. Model-based predictability: realizable predictability and its quantification

First of all, the model-based results are shown. Just as we have mentioned in the above sections, the objective in this paper is not to achieve an optimal model to predict the DTR's variations, but to show that there are different predictive structures in DTR's variations over different regions. Therefore, all model-based results given in this paper are from ARFIMA(3, d , 0) model. The same preset parameters can let us discover the different predictive structures in DTR's variations over different regions. This can be directly found in Fig. 1, where a piece of typical DTR anomaly series from a representative station over northern and southern China are presented, one is station 52602(38.45°N, 93.20°E) and the other is station 56964(22.47°N, 100.58°E). Over northern China, much less DTR's variations are caught in predicted series, so more of this structured information is left in the residuals. Contrary to this case, more DTR's variations over southern China are caught in predicted series, less structured information is left in the residuals. Since the realizable predictability is often quantified through the correlation coefficient between observations and predictions [3], we will check whether this measure works well in realizable predictability for DTR's fluctuations. For 22-year testing series over a given station, we can derive predicted one and its residuals. And then, we can calculate a pair of correlation coefficients over each station, one is between testing series and predicted one, and the other between testing series and residuals. At last, 185 paired correlation coefficients were reached. In fact, the different realizable

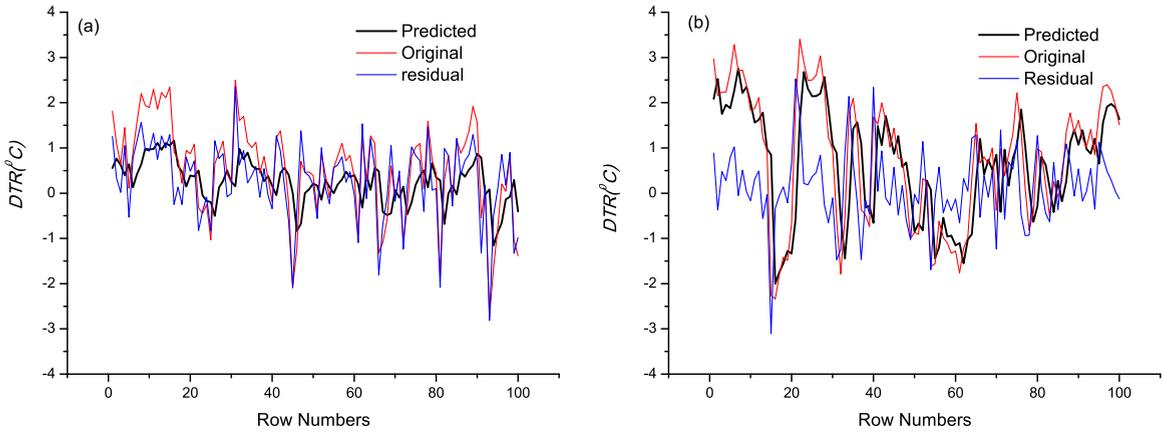


Fig. 1. A piece of typical DTR anomaly series from a representative station (a) 52602 (38.45°N,93.20°E) over northern China, and (b) 56964 (22.47°N,100.58°E) over southern China.

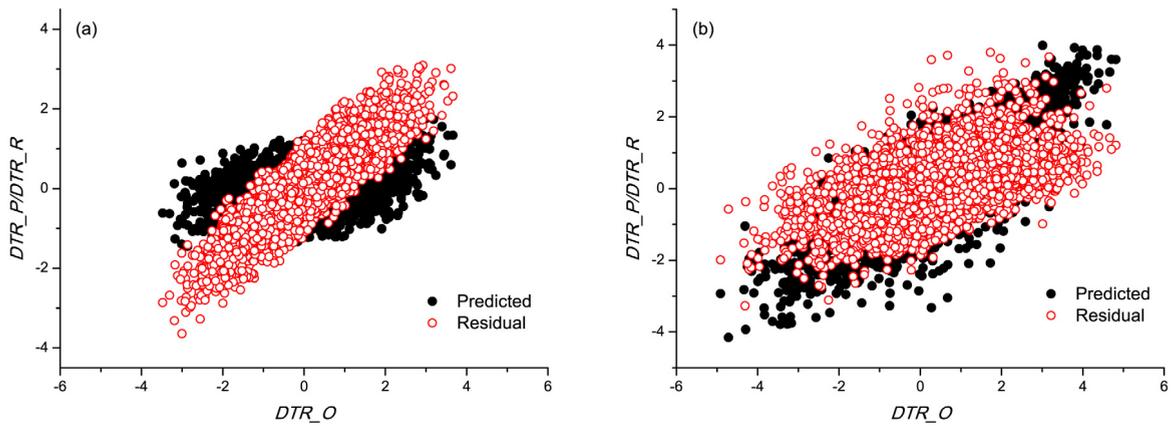


Fig. 2. Scatter plots of 22-year testing DTR anomaly series versus its predicted and residual parts from a representative station (a) 52602 (38.45°N,93.20°E) over northern China, and (b) 56964 (22.47°N,100.58°E) over southern China.

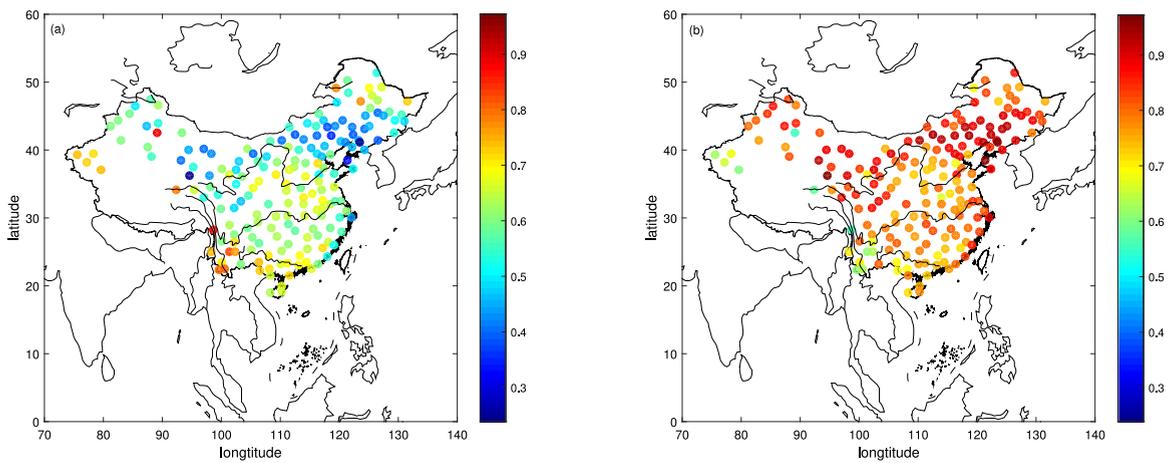


Fig. 3. Spatial distribution of correlation coefficients between 22-year testing DTR anomaly series and its (a) predicted, (b) residual series.

predictability behaviors can be indeed further revealed in the scatter plots for 22-year testing DTR anomaly series versus its predicted and residual parts, as shown in Fig. 2 from above two representative stations 52602(38.45°N, 93.20°E) and

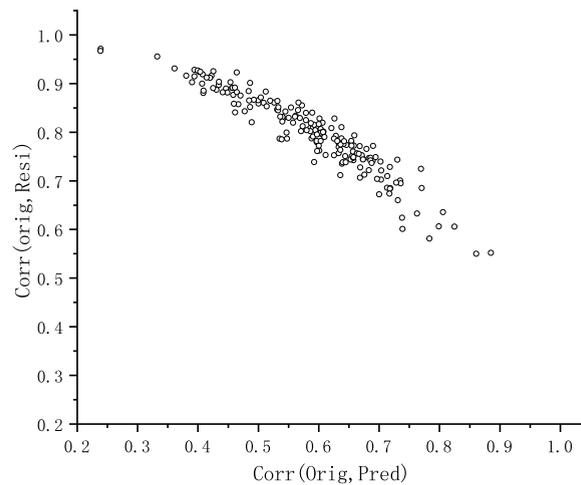


Fig. 4. Scatter plot for correlation coefficients between 22-year testing DTR anomaly series and predicted ones versus correlation coefficients between 22-year testing DTR anomaly series and residuals.

56964(22.47°N, 100.58°E) for details. Over northern China, the correlation between the testing DTR anomaly series and its predictions is weaker than that between the testing DTR anomaly series and its residuals, see Fig. 2a. However, over southern China, the correlation between the testing DTR anomaly series and its predictions is higher than that between the testing DTR anomaly series and its residuals, see Fig. 2b. Detailed spatial distribution of correlation coefficient between observations and predictions/residuals is shown in Fig. 3. First of all, we can see linear model as ARFIMA(p, d, q) does not work well for DTR's fluctuations, since the overall correlation coefficients between observations and predictions are all not high, see Fig. 3a. Instead, the correlation coefficients between observations and residuals are higher, and all of the correlation coefficients are almost higher than 0.55. This indicates that most of structured information in original series is left in residuals, and ARFIMA(3, $d, 0$) model fails to catch this information. Secondly, there are dominant region-dependent spatial patterns of the realizable predictability behaviors, and they are overall lower over northern China but higher over southern China. Above results are similar as the findings given in Ref. [36] that the DTR's variations are multi-fractal over southern China, but mono-fractal over northern China. Further analysis shows that there is a well defined monotonic negative correlation pattern between these two correlation coefficients. When the correlation coefficients between testing DTR anomaly series and predictions are lower, correspondingly, the correlation coefficients between testing DTR anomaly series and residuals are higher, vice versa. It should be also pointed out that this monotonic negative correlation is not linear, see Fig. 4. Even when the correlation coefficients between testing DTR anomaly series and predictions are close to 1, the correlation coefficients between testing DTR anomaly series and residuals are still high, above 0.55. However, when the correlation coefficients between testing DTR anomaly series and residuals are close to 1, the correlation coefficients between testing DTR anomaly series and predictions can reach much lower value, only 0.2. This marked asymmetric behavior maybe indicates that DTR anomaly series are not outputs of a linear process.

4.2. Data-driven predictability: intrinsic predictability and its quantification

In order to well understand the dominant region-dependent spatial patterns of the DTR's realizable predictability behaviors, we further carried out studies on DTR's intrinsic predictability. Different from previous studies [2,3] where weighted PE (WPE, amplitude information in time series is incorporated [49]) was adopted to estimate the intrinsic predictability, only PE was chosen in this paper to quantify DTR's intrinsic predictability (we will discuss why we chose PE and it is suitable to cope with DTR's intrinsic predictability in the last section). With the help of Eq. (3), PE is calculated for DTR's fluctuations over each station. Spatial distribution of PE with $D = 4$ for 44-year original series over 185 stations is shown in Fig. 5a. Following the definition given in Ref. [2,3], $1 - PE$ is approximated as the intrinsic predictability of any given time series. Then the first direct impression acquired in Fig. 5a is that the intrinsic predictability is lower for DTR's fluctuations, since all values of $1 - PE$ are much close to zero. The second direct impression is that there is marked contrasting intrinsic predictability over northern and southern China. Long striped regions of the highest PE values are located in or north to the Yellow River regions, which are in well accord with the regions of the lowest correlation efficient values given in Fig. 3a. At the same time, regions of the lower PE values are mainly located over southern China, which are also in well accord with the regions of the higher correlation coefficient values given in Fig. 3a. These results indicate that there is a good correspondence between the intrinsic predictability and realizable predictability.

For better to see the correspondence between the intrinsic predictability and realizable predictability, the whole DTR' fluctuation records are divided into two parts of equal data length, and the first part is taken as training series and the

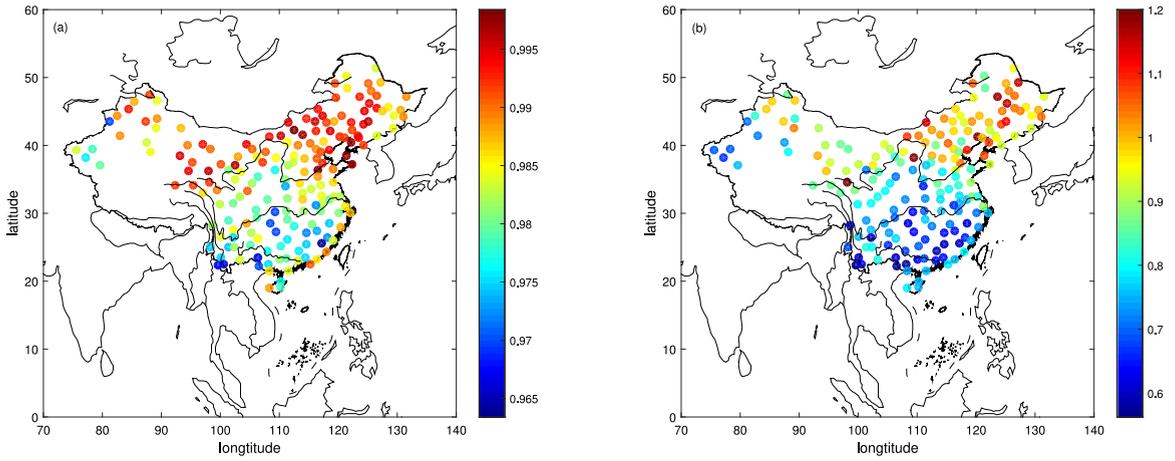


Fig. 5. Spatial distribution of (a) PE with $D = 4$ for 44-year original series, (b) FE for 22-year predicted series.

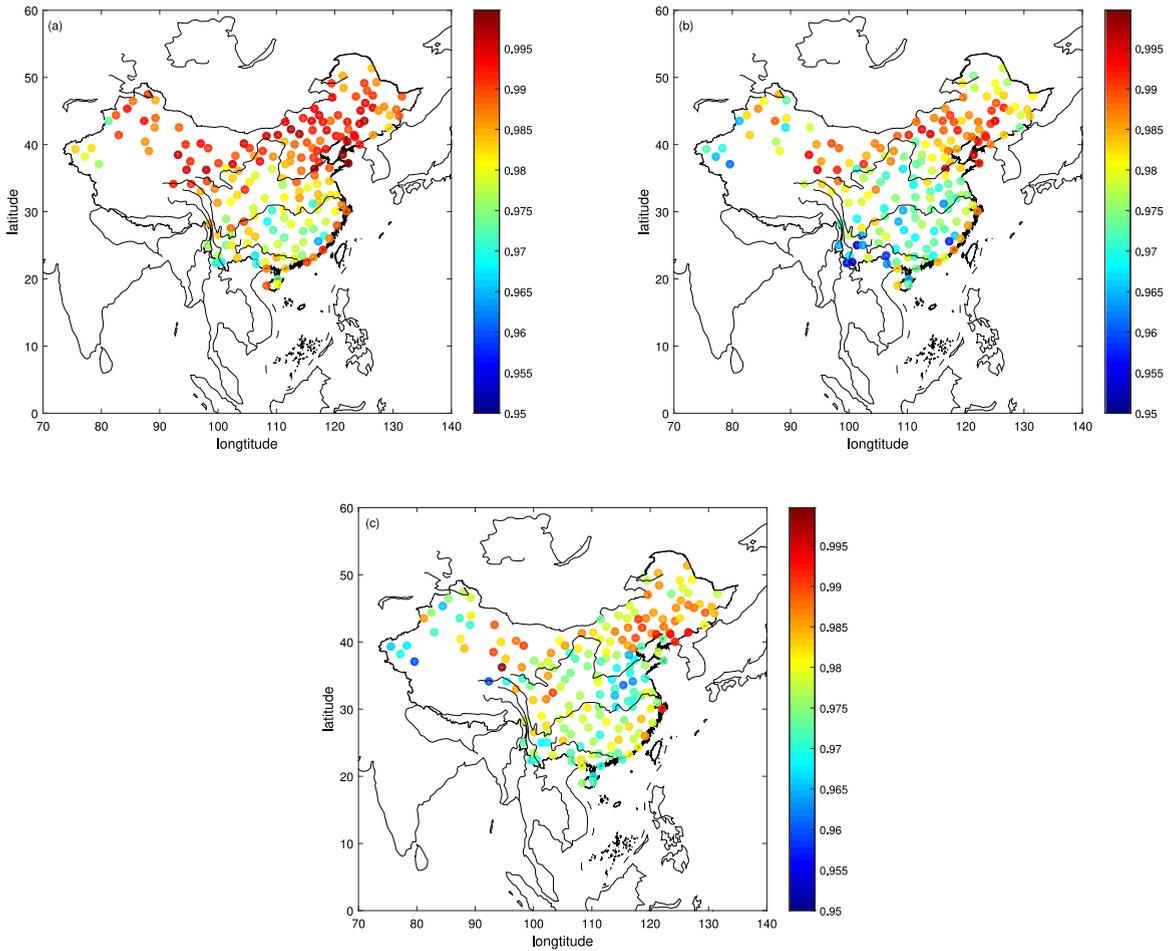


Fig. 6. Spatial distribution of PE with $D = 4$ over 22 years for (a) training (b) testing and (c) predicted DTR anomaly series.

second part as the testing series. At the same time, the predictions from the testing series can be reached based on the fitted ARFIMA(p, d, q) model from the training series. Comparing the PE from both training and testing series, we can see that the overall patterns are similar, and there still exist contrasting intrinsic predictability behaviors over northern and southern China, see Fig. 6a and 6b. However, the intrinsic predictability is enhanced during the second 22 years. The long striped

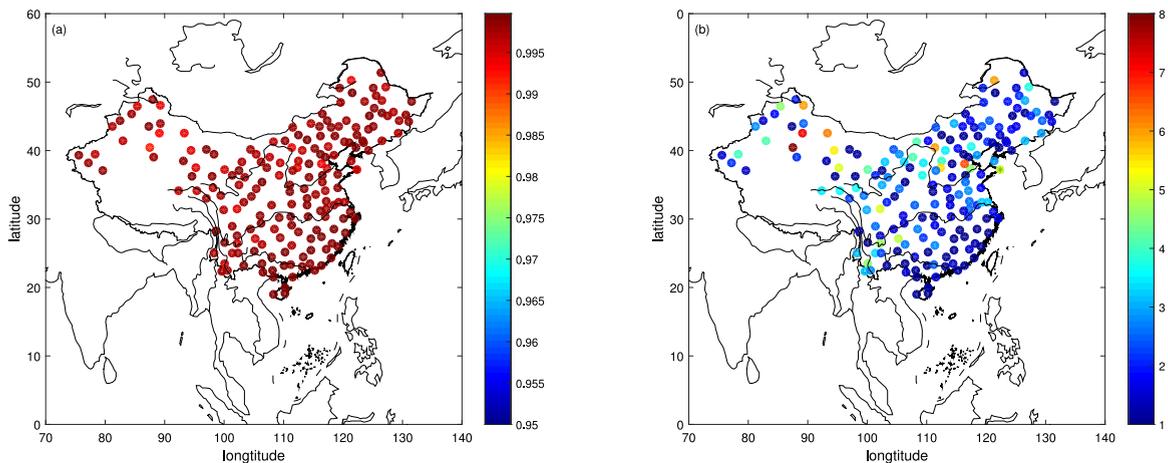


Fig. 7. Spatial distribution of PE with $D = 4$ over 22 years for residuals of the testing DTR anomaly series with (a) the color bar of the same scale as in Fig. 6, (b) the re-scaled color bar to $(1 - PE) * 1000$.

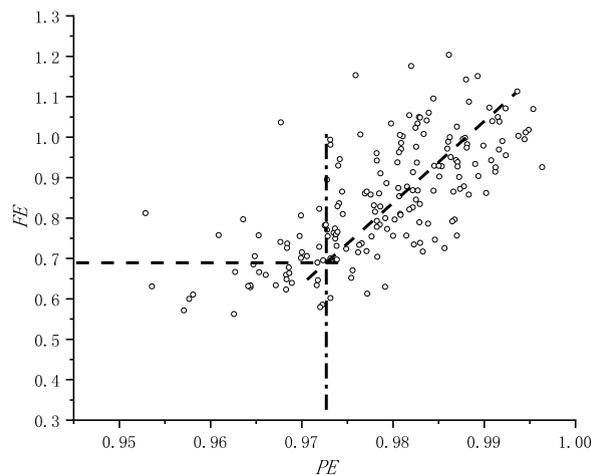


Fig. 8. Scatter plot for PE from the 22-year testing series v.s. FE from predicted series of the corresponding 22-year testing series. The dash black lines provide a visual guide for different regimes of overall monotonic behavior and the vertical dash-dot black line for visual guide of regime dividing.

regions of the highest PE values are shrunk with the decreased intrinsic predictability strength and the regions of the lower PE values are expanded with enhanced intrinsic predictability strength. This indicates that intrinsic predictability strength may change with time, and instantaneous or long-term variations of the DTR's intrinsic predictability should be also carefully checked. The PE of corresponding predictions shares the quite similar features with that of the training series, both their strength and spatial patterns, see Fig. 6b and 6c. The resemblance between the training series and predictions can be further confirmed in the results from corresponding residuals. The strength of PE over all the stations is close to the upper limit of PE , and the spatial distribution of PE over all the stations is uniform, see Fig. 7a, although minor difference can be still found in re-scaled case, see Fig. 7b. At last, we want to reemphasize an important point that the resemblance between predictability and multi-fractal strength North–South contrasting spatial pattern indicates a close relationship between them. When the multi-fractal strength is stronger over southern China, both the intrinsic and realizable predictability are also higher, on the contrary, when the multi-fractal strength is weaker over northern China, both the intrinsic and realizable predictability are also lower. This may indicate that the enhanced multi-fractal strength favors the increase in predictability, both intrinsic and realizable. The predictive structures in DTR fluctuations can be effectively quantified by both PE and FE , especially spatial patterns of predictive structures in DTR fluctuations can be effectively quantified by both PE and FE .

4.3. Relation between PE and FE in DTR

Previous studies found there is a well defined correlation pattern between WPE and FE [2,3] and the intrinsic predictability is positively correlated with the realizable predictability. In the above section, we have shown that PE is in well accord with the correlation coefficient between testing series and predictions.

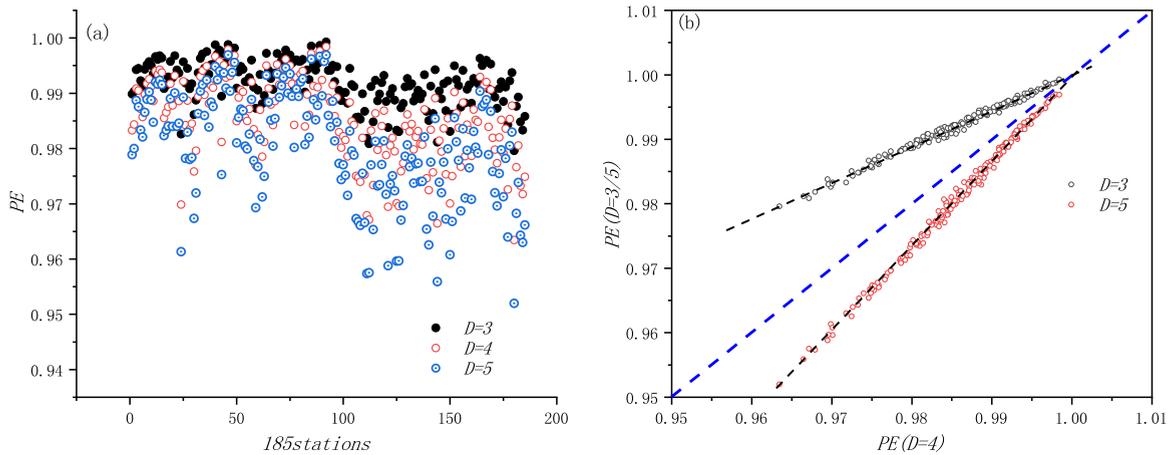


Fig. 9. The effect of D on estimation of PE , (a) values of PE for $D = 3, 4, 5$ over each station and (b) scatter plots of PE for $D = 4$ V.S. $D = 3, 5$ from 1970 to 2013 over 185 stations. The dash black lines provide a visual guide for monotonic line relation between $D = 4$ and $D = 3, 5$ and the dash blue line is 1 : 1 diagonal line for visual guide.

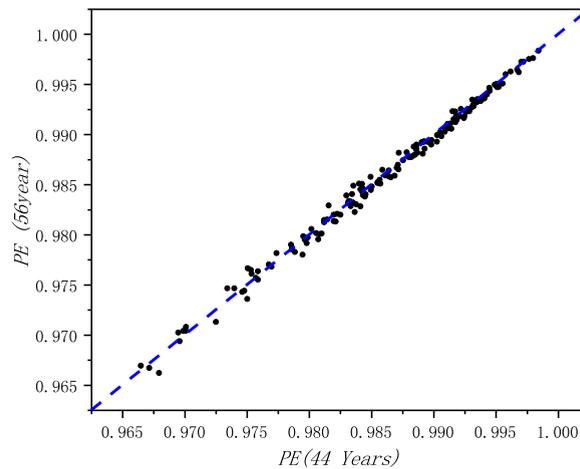


Fig. 10. Scatter plot for PE from 1970 to 2013 versus 1960 to 2015 over 185 stations. The dash blue line is 1 : 1 diagonal line for visual guide.

To better confirm this close relationship, we calculate the FE based on the definition given in Eq. (1) from the residuals. The spatial distribution of FE shows much clearer North–South contrasting behaviors, see Fig. 5b. The latitude of around 37°N is the dividing line, where the FE is below 1 over regions south to this line, and FE above 1 over regions north to this line. And the smaller values of FE over southern China indicate there are more deterministic predictive structures in the testing DTR's fluctuations. Just as we have mentioned in previous section, when the training series are sinusoidal signals, the calculated FE should be the bottom limit zero. Due to higher frequency fluctuations or measured noise in DTR's fluctuations, the calculated FE cannot reach the bottom limit. On the contrary, there are less deterministic predictive structures in the testing DTR's fluctuations over northern China, therefore, the estimated FE is higher than 1. And most importantly, the spatial patterns of FE resemble very well those of PE given in Fig. 5a, though the detailed strength is different between them. Actually, this inconsistency in strength between them is maybe from the fact that the relationship between PE and FE is not totally linear, see Fig. 8. From the scatter plot for PE from the 22-year testing series versus FE from predicted series of the corresponding 22-year testing series, regime dependent relations can be revealed. When PE is less than 0.973, FE is overall unchanged, but when PE is larger than 0.973, FE overall linearly increases as PE increases. However, the details shown in Fig. 8 also indicate that there maybe exist still other factors contributing to this relationship. At the same time, higher PE for DTR anomaly indicates that structured information, the nonlinear and/or deterministic strength, in DTR anomaly is much weaker, so there is no too much predictability ($1-PE$ is close to zero) in DTR anomaly. The physical meaning for this lower value is that even we can build the perfect model, the realizable predictability is still very low for DTR anomaly.

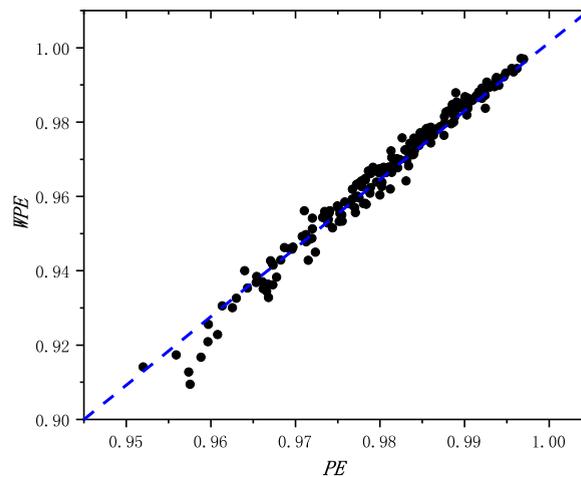


Fig. 11. Scatter plot for PE v.s. WPE for DTR from 1970 to 2013 over 185 stations. The dash blue line provides a visual guide for monotonic line relation between PE and WPE .

5. Conclusion and discussion

In this paper, both realizable and intrinsic predictability of DTR have been quantified by model-based measure FE and data-driven measure PE , and it is confirmed that there is still well defined relation between FE and PE in DTR fluctuations. This is an interesting finding since the well defined relation between FE and PE has only been reported in some idealized series and empirical series [2,3]. This is the first result found in the climate field and there is a regime dependent relation between FE and PE .

There are some issues that we should must do to check whether these findings are robust to the parameter choice.

Firstly, we check the choice of PE dimension D . All results given in this paper are mainly based on $D = 4$, we should validate that the results are robust to different choices of D . Careful calculations for different choices of D show that the results are qualitatively insensitive to the choice of D . This can be revealed in Fig. 9, where PE are calculated for $D = 3$, $D = 4$ and $D = 5$. We can see that although the specific values of PE over each station for changing D are different, the qualitative behaviors over all stations are similar for different choices of D . And this qualitative consistence can be further confirmed in the scatter plot of PE for $D = 4$ versus those for $D = 3$ or $D = 5$, where PE over all stations for different choices of D is nearly in a straight line. This further indicates that the results given in this paper are insensitive to the choice of D .

The second issue should be considered is the finite length effect in the analysis. We have used data with different lengths in this paper for different purposes, we should make sure that this will not cause any serious problem in reaching our basic conclusion. Due to the advantage of permutation method [22,23], the finite size effect is minor even when the short series is considered. The recommended data length L given in Ref. [23] is $L \geq 5D!$, we can see that this can be easily satisfied in our calculations. For comparison, we show the results from different spans used in this paper in Fig. 10. It is obvious that both results perfectly lie in the 1 : 1 line, and this indicates that results from different spans are nearly the same. Actually, even much shorter (but larger than two years) data size is chosen, the results are also not altered qualitatively (figures are not shown here).

In this paper, permutation entropy not weighted permutation entropy [2,3] has been adopted to quantify the intrinsic predictability in time series. Will this choice distort the results given in this paper? The answer is surely not. Calculations of both permutation entropy and weighted permutation entropy over all 185 stations show that the results from both measures for DTR over 185 stations are qualitatively similar, and this can be found in Fig. 11. Further studies show that the DTR fluctuations are almost Gaussian distributed, there are no dominated amplitude dependent structures in the DTR fluctuations over all 185 stations, so the contributions to the permutation entropy from amplitude effect are minor in the DTR fluctuations. For simplicity, we choose PE rather than WPE to quantify the intrinsic predictability in the DTR fluctuations. If there are dominated amplitude dependent structures in the underlying analyzed series, WPE should be chosen as the quantitative measure on intrinsic predictability, just as suggested in the literature [2,3].

It should be pointed out why we choose DTR to show the relation between the realizable predictability and the intrinsic predictability. Previous studies [36] show that there are contrasting multi-fractal behaviors in DTR over southern and northern China, so we want to know whether there are also contrasting predictability behaviors in DTR over southern and northern China. The answer to this question is yes. There are region-specific predictability behaviors in DTR over different regions of China, and most importantly, there are overall contrasting predictability behaviors in DTR over southern and northern China, both realizable predictability and the intrinsic predictability. Why are there so different predictability behaviors in DTR over southern and northern China? The detailed answer to this question deserves further studies. Even so, we can conjecture that the different multi-fractal behaviors in DTR over southern and northern China contribute greatly to

this dominated spatial patterns of predictability behaviors. We can also find some clues in this direction from the literature. Previous studies show that the predictability can be enhanced by increasing nonlinearity in ENSO and Lorenz systems [9]. As one of important nonlinear measures, the increasing multi-fractal strength in DTR fluctuations maybe also enhances their predictability, both realizable predictability and the intrinsic predictability. At the same time, we should note that the multi-fractal strength is not a sufficient and necessary condition to nonlinearity, and mono-fractal series can also be nonlinear [50]. This can be also taken as an explanation why there is no fully South–North contrasting behavior in DTR over China.

At last, it should be noted that the purpose of this paper is not to optimize a realizable model to reach the maximal realized predictability in DTR fluctuations, not to make a perfect prediction of DTR fluctuations over multiple steps. So we only consider linear model (ARFIMA(p, d, q) model) and one-step ahead prediction to show the contrasting realizable predictability behaviors in DTR over southern and northern China. The choice indeed helps us to reach this goal, but with a lower realized predictability. In fact, if we consider more complicated predication model, such as nonlinear prediction model, much higher prediction can be indeed realized over multiple steps (For example, if we adopt the nonlinear prediction based on the Lorenz method of analogues in embedded space [2], the realizable predictability in DTR can be improved greatly. Detailed results are not shown here). At the same time, we should point out that only station measured DTR data have been analyzed in this paper. If we want to reach more detailed spatial patterns of both realizable predictability and the intrinsic predictability in DTR fluctuations, reanalysis and outputs of models may be of great importance [51,52], where more studies can be done, such as DTR reanalysis quality evaluation based on data-driven method, the mechanisms causing DTR South–North asymmetric predictability from outputs of models and so on.

Acknowledgment

This research was supported by the National Natural Science Foundation of China through Grants (No.41675049 and No.41475048).

References

- [1] J Hacker, J Hansen, J Berner, et al., Predictability, *BAMS* 86 (2005) 1733–1737.
- [2] J Garland, R James, E Bradley, Model-free quantification of time-series predictability, *Phys. Rev. E* 90 (2014) 052910.
- [3] F Pennekamp, AC Iles, J Garland, et al., The intrinsic predictability of ecological time series and its potential to guide forecasting, *bioRxiv*, 2018, <http://dx.doi.org/10.1101/350017>.
- [4] LW Salvino, R Cawley, C Grebogi, JA Yorke, Predictability in time series, *Phys. Lett. A* 209 (1995) 327–332.
- [5] G Rangarajan, DA Sant, A climate predictability index and its applications, *Geophys. Res. Lett.* 24 (1997) 1239–1242.
- [6] DJ Patil, BR Hunt, E Kalnay, JA Yorke, E Ott, Local low dimensionality of atmospheric dynamics, *Phys. Rev. Lett.* 86 (2001) 5878–5881.
- [7] GC Yuan, MS Lozier, LJ Pratt, CKRT. Jones, KR Helfrich, Estimating the predictability of an oceanic time series using linear and nonlinear methods, *J. Geophys. Res.* 109 (2004) C08002, <http://dx.doi.org/10.1029/2003JC002148>.
- [8] W von Bloh, MC Romano, M Thiel, Long-term predictability of mean daily temperature data, *Nonlinear Process. Geophys.* 12 (2005) 471–479.
- [9] Z Ye, WW Hsieh, Enhancing predictability by increasing nonlinearity in ENSO and Lorenz systems, *Nonlinear Processes Geophys.* 15 (2008) 793–801.
- [10] T DelSole, MK Tippett, Predictability: Recent insights from information theory, *Rev. Geophys.* 45 (2010) RG4002.
- [11] RQ Ding, JP Li, KH Seo, Predictability of the Madden-Julian oscillation estimated using observational data, *Mon. Wea. Rev.* 138 (2010) 1004–1013.
- [12] RQ Ding, JP Li, KH Seo, Estimate of the predictability of boreal summer and winter intraseasonal oscillations from observations, *Mon. Wea. Rev.* 139 (2011) 2421–2438.
- [13] C Franzke, T Woollings, On the persistence and predictability properties of North Atlantic climate variability, *J. Clim.* 24 (2011) 466–472.
- [14] JP Li, RQ Ding, Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogues, *Mon. Wea. Rev.* 139 (2011) 3265–3283.
- [15] ZS Zhang, ZQ Gong, R Zhi, GL Feng, JG Hu, Preliminary research on the relationship between long-range correlations and predictability, *Chin. Phys. B* 20 (2011) 019201.
- [16] L Karthikeyan, DN Kumar, Predictability of nonstationary time series using wavelet and EMD based ARMA models, *J. Hydr.* 502 (2013) 103–119.
- [17] JP Li, RQ Ding, Temporal-spatial distribution of the predictability limit of monthly sea surface temperature in the global oceans, *Int. J. Climatol.* 33 (2013) 1936–1947.
- [18] RQ Ding, JP Li, F Zheng, J Feng, DQ Liu, Estimating the limit of decadal-scale climate predictability using observational data, *Clima. Dyn.* 46 (2016) 1563–1580.
- [19] ZL Hou, JP Li, RQ Ding, J Feng, WS Duan, The application of nonlinear local Lyapunov vectors to the Zebiak-Cane model and their performance in ensemble prediction, *Clima. Dyn.* 51 (2017) 283–304.
- [20] JP Li, J Feng, RQ Ding, Attractor radius and global attractor radius and their application to the quantification of predictability limits, *Clima. Dyn.* 51 (2018) 2359–2374.
- [21] NM Yuan, Y Huang, JP Duan, et al., On climate prediction: How much can we expect from climate memory? *Clima. Dyn.* (2018) <https://doi.org/10.1007/s00382-018-4168-5>.
- [22] C Bandt, B Pompe, Permutation entropy: A natural complexity measure for time series, *Phys. Rev. Lett.* 88 (2002) 174102.
- [23] M Riedl, A Müller, N Wessel, Practical considerations of permutation entropy, *Eur. Phys. J. Spec. Top.* 222 (2013) 249–262.
- [24] QL Li, ZT Fu, NM Yuan, Beyond Benford's Law: Distinguishing Noise from chaos, *PLoS ONE* 10 (2015) e0129161.
- [25] T King, Quantifying nonlinearity and geometry in time series of climate, *Quat. Sci. Rev.* 15 (1996) 247–266.
- [26] I Bartos, IM Jánosi, Atmospheric response function over land: Strong asymmetries in daily temperature fluctuations, *Geophys. Res. Lett.* 32 (2005) L23820.
- [27] B Gyüre, I Bartos, IM Jánosi, Nonlinear statistics of daily temperature fluctuations reproduced in a laboratory experiment, *Phys. Rev. E* 76 (2007) 037301.
- [28] Y Ashkenazy, Y Feliks, H Gildor, E Tziperman, Asymmetry of daily temperature records, *J. Atmos. Sci.* 65 (2008) 3327.
- [29] FH Xie, ZT Fu, L Piao, JY Mao, Time irreversibility of mean temperature anomaly variations over China, *Theor. Appl. Climat.* 123 (2016) 161.
- [30] TR Karl, G Kukla, et al., Global warming: Evidence for asymmetric diurnal temperature change, *Geophys. Res. Lett.* 18 (1991) 2253.
- [31] TR Karl, PD Jones, et al., A new perspective on recent global warming: Asymmetric trends of daily maximum and minimum temperature, *Bull. Americ. Meteorol. Soc.* 74 (1993) 1007.

- [32] RO Weber, P Talkner, G Stefanicki, Asymmetric diurnal temperature change in the Alpine region, *Geophys. Res. Lett.* 21 (1994) 673.
- [33] RC Balling, DA Periconi, RS Corveny, Large asymmetric temperature trends at Mount Wilson, California, *Geophys. Res. Lett.* 26 (1999) 2753.
- [34] M Pattanyús-Ábrahám, A Király, IM Jánosi, Nonuniversal atmospheric persistence: Different scaling of daily minimum and maximum temperatures, *Phys. Rev. E* 69 (2004) 021110.
- [35] RG Lauritsen, JC Rogers, U.S. Diurnal temperature range variability and regional causal mechanisms, 1901–2002, *J. Clim.* 25 (2012) 7216.
- [36] NM Yuan, ZT Fu, JY Mao, Different multi-fractal behaviors of diurnal temperature range over the north and the south of China, *Theor. Appl. Climatol.* 112 (2013) 673–682.
- [37] Q Li, H Zhang, J Chen, W Li, X Liu, P Jones, A mainland China homogenized historical temperature dataset for 1951~2004, *Bull. Amer. Meteor.* 90 (2009) 1062.
- [38] QM Deng, D Nian, ZT Fu, The impact of inter-annual variability of annual cycle on long-term persistence of surface air temperature in long historical records, *Clim. Dyn.* 50 (2018) 1091–1100.
- [39] E Koscielny-Bunde, A Bunde, S Havlin, HE Roman, Indication of a universal persistence law governing atmospheric variability, *Phys. Rev. Lett.* 81 (1998) 729.
- [40] H von Storch, FW Zwiers, *Statistical analysis in climate research*, Cambridge Univ. Press, Cambridge, 1999.
- [41] A Király, IM Jánosi, Stochastic modeling of daily temperature fluctuations, *Phys. Rev. E* 65 (2002) 051102.
- [42] S Bisgaard, M Kulahci, *Time series analysis and forecasting by example*, John Wiley & Sons, New York, 2011.
- [43] T Graves, RB Gramacy, C Franzke, NW Watkins, Efficient Bayesian inference for natural time series using ARFIMA processes, *Nonlinear Processes Geophys.* 22 (2015) 679–700.
- [44] CK Peng, SV Buldyrev, S Havlin, M Simons, HE Stanley, AL Goldberger, Mosaic organization of DNA nucleotides, *Phys. Rev. E* 49 (1994) 1685–1689.
- [45] JW Kantelhardt, E Koscielny-Bunde, HHA Rego, S Havlin, A Bunde, Detecting long-range correlations with detrended fluctuation analysis, *Physica A* 295 (2001) 441–454.
- [46] A Bunde, J Eichner, JW Kantelhardt, S Havlin, Long-term memory: A natural mechanism for the clustering of extreme events and anomalous residual times in climate records, *Phys. Rev. Lett.* 94 (2005) 048701.
- [47] NM Yuan, ZT Fu, JY Mao, Different scaling behaviors in daily temperature records over China, *Physica A* 389 (2010) 4087.
- [48] S Chatterjee, As Hadi, *Regression Analysis By Example*, John Wiley & Sons, Inc, New Jersey, 2006.
- [49] B Fadlallah, et al., Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information, *Phys. Rev. E* 87 (2013) 022911.
- [50] RB Govindan, H Kantz, Long-term correlations and multifractality in surface wind speed, *Europhys. Lett.* 68 (2004) 184–190.
- [51] SS Zhao, WP He, YD Jiang, Evaluation of NCEP-2 and CFSR reanalysis seasonal temperature data in China using detrended fluctuation analysis, *Int. J. Climatol.* 38 (2017) <http://dx.doi.org/10.1002/joc.5173>.
- [52] WP He, SS Zhao, Assessment of the quality of NCEP-2 and CFSR reanalysis daily temperature in China based on long-range correlation, *Clima. Dyn.* 50 (2018) 493.