Physica A 523 (2019) 1387–1399

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Comparative study of multiple measures on temporal irreversibility of daily air temperature anomaly variations over China

Boer Zhang [a], Fenghua Xie [b], Zunhai Fu [c], Zuntao Fu [a,*]

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

[b] Department of Atmospheric Science, School of Environmental Studies, China University of Geosciences, Wuhan, 430074, China

[c] Weihe Forestry Bureau Middle School, Shangzhi, Heilongjiang Province, 150623, China

HIGHLIGHTS

- Five TI measures reach consistent conclusions to daily temperature observations.
- Novel mechanism is proposed to explain this dominated TI.
- Seasonal TI patterns are consistent with the overall TI features.

ARTICLE INFO

Article history:
Received 13 August 2018
Received in revised form 24 January 2019
Available online 24 April 2019

Keywords:
Daily air temperature anomaly variations
Temporal irreversibility
Nonlinear measure
Extreme air temperature variations

ABSTRACT

As a fundamental property of nonlinear time series, temporal irreversibility (TI) is a hot topic in various fields. Among these studies, many methods and measures have been developed. However, only a few studies contribute to the performance comparison of these methods and measures. Taking the daily mean air temperature from observations over China as an example, the performance of five well developed measures was compared in this study. Three of them are directly related to air temperature increments with A defined from sign ratio, G from magnitude ratio and E from skewness, and the other two measures are L_2 from directed horizontal visibility graph (DHVG for short) and L_1 from consecutive increasing and decreasing steps (CIDS for short). The results show that five measures reach almost consistent conclusions to TI of daily mean air temperature from observations over China, but with different statistical significance and sensitivity to low-frequency trend and measured noise. They also show that the higher TI occurs over southern China, and lower TI over northeast China and southwest China close to Basin of Sichuan and Yungui-Tibetan Plateau. Further studies indicate that this dominated TI behavior is closely related to the number difference between extreme decreasing temperature variations and extreme increasing temperature variations. When conditional threshold exceeds certain critical values, this marked TI can be well quantified by the extreme decreasing but not increasing temperature variations. At last, combining all aspects considered, measure from CIDS is the best among these five TI measures due to its robustness to low-frequency trend and measured noise.

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

https://doi.org/10.1016/j.physa.2019.04.131
0378-4371/

1. Introduction

If for any given n, the series $\{x_1, \ldots, x_n\}$ and $\{x_n, \ldots, x_1\}$ have the same joint probability distributions (i.e. for a given $s > 0, P(X(t), X(t+s)) = P(X(t+s), X(t))$), then a stationary process $X(t)$ is defined as time reversible [1] and it is invariant under the reversal arrow of time. Gaussian linear processes are taken to be this kind of reversible processes. Contrary to this kind of time reversible property, temporal irreversibility (TI) has been taken as an indicator of nonlinearities in the underlying dynamics [2,3]. Studying TI of time series is an important indirect quantitative assessment way of nonlinearity [4,5]. From this TI view point to quantify nonlinearity hidden in observational time series has been attractive studies in various fields, such as physiological series [6–10], atmospheric and oceanic records [11–14].

How to measure this kind of TI behavior in time series plays a crucial role in these studies. Several statistical tests have been developed to detect and quantify TI in time series [3,8–10,12,15–18]. Among these methods, most of them firstly perform a time series symbolization and then analyze symbol strings occurrence in the forward and backward series [16]. This extra amount of ad hoc information for symbolization procedure may bias the correct TI's quantification [3], for example, range partitioning or size of the symbol alphabet may bias the TI's quantification depending on the choice of extra parameters. And among these methods, five measures [3,12,15,17,18] do not consider this kind of ad hoc symbolization procedure. Measure L_2 is developed by Lacasa et al. [3] based on horizontal visibility graph (HVG) [19]. HVG is an algorithmic variant of visibility graphs [20], and it can bypass ad hoc symbolization procedure. In this method, an asymmetry can be derived from the probability distributions of the numbers of incoming and outgoing links by defining the directed horizontal visibility graphs (DHVGs) in any given time series [3,10]. And this ad hoc symbolization procedure problem can also be solved by considering directly the natural local ordinal structures in any given series with its distribution differences between consecutive increasing and decreasing steps (CIDSs), which defines measure L_1 [12]. Other three measures among these five measures are all related to series' increments. Given a time series x_i, its increment time series are $dx_i = x_{i+1} - x_i$. And then measure A is defined from sign ratio of dx_i, measure G from magnitude ratio of dx_i and measure E from skewness of dx_i (details about these measures can be found in Methods section). Since these five measures quantify TI by means of different aspects of time series, we will compare their performance in quantifying TI in details.

Since nonlinearity is an intrinsic feature of many real-wold systems, TI should be unveiled from outputs of these systems [21]. As one of the most important variables of natural changes, air temperature records [22–29] are the best choice for TI analysis and its measures' performance test. Previous studies found that there is a marked nonlinear feature [11,12] in daily terrestrial mean air temperature with the number of warming steps significantly different from the number of cooling steps [30,31]. So we will carry out the performance of above mentioned five measures on daily mean air temperature records over China to quantify TI in details. We will answer the following questions through this study. Will these five measures reach consistent TI results? What has contributed to this dominated TI in the daily mean air temperature records? Will this marked TI feature change over different seasons?

The rest of the paper is organized as follows. In Section 2, we will make a brief introduction to the data sets used in this paper. We will briefly explain the five TI quantifying measures and related methods in Section 3. The results from these five measures will be discussed in Section 4. In Section 5, discussions and conclusions are made.

2. Data

In this article, daily mean air temperature records from 176 stations are used for our analysis. The data are obtained from the China meteorological data sharing service system (http://data.cma.cn/en), with length of 56 years, from 1960 to 2015. Homogenized observations [26] from all 176 meteorological stations are taken part in international exchange. Since some measures (for example, moments related measures depend on the estimated probability density function (PDF) of under analyzed series) may suffer from the effect of low-frequency trend, before carrying out analysis, the data were first standardized by removing the seasonal trend through subtracting the annual cycle, as $T_i' = T_i - \langle T_i \rangle$ [22], where T_i is the daily mean air temperature and $\langle T_i \rangle$ is long-time climatological average for each calendar day. Since the analysis in this study is mainly related to the nonlinear measure (usually non-stationary standard deviation is taken as a kind of nonlinearity), so the seasonal standard deviation was not subtracted. In this article, both temperature anomaly series T_i' and their increment series $dT_i' = T_{i+1}' - T_i'$ were analyzed. It should be pointed out that although increment series can also be derived from the original series T_i, however as a high-pass filter, difference filter cannot totally remove seasonal components in the original series T_i [32]. Measures L_1 and L_2 were calculated from T_i', and measures A, G and E from dT_i'.

Surrogates from the observations can be generated based on model-free procedures [33–36] and model-based method [12,23]. Among the model-free procedures, data-driven methods, such as shuffling [35,36] and Fourier-filtering techniques and phase randomize surrogate procedure (PRS) introduced by Schreiber and Schmitz [33,34], were chosen to generate surrogated data with 1000 samples for each group from normalized air temperature anomaly series. These surrogates were applied to calculate the significance threshold of each measure to distinguish linear series from outputs of nonlinear processes. Just as reported in previous studies [12], nearly all estimated measures are TI significant at a given confidence level, such as 95.00% or above for daily mean air temperature over China. However, there may be different conclusions based on different TI measures for a specific air temperature record due to the different aspects each TI measure reflects and the different robustness of each TI measure to low-frequency trends and measured noise (see details in next parts).

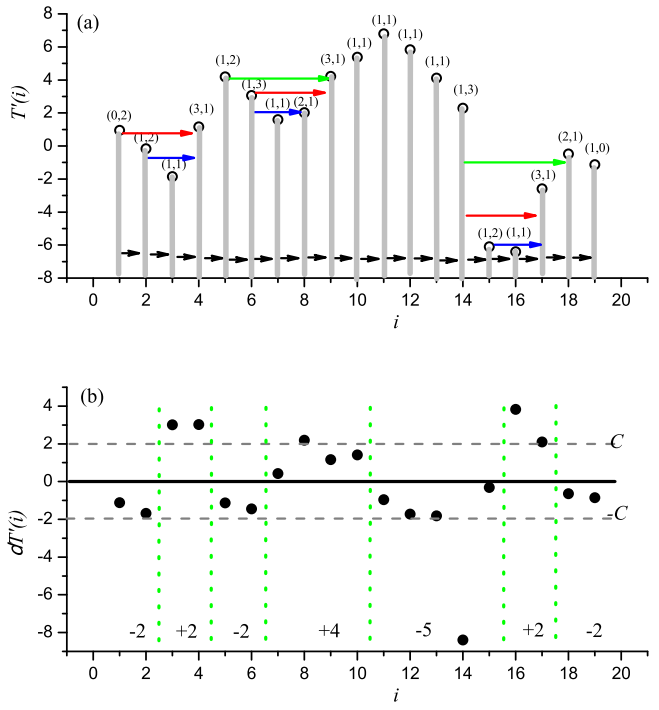

Fig. 1. Segment of typical daily mean air temperature anomaly series T_i' (a), and its corresponding increment series $dT_i' = T_{i+1}' - T_i'$ (b) with slow increasing and rapid decreasing. At the same time, it shows the schematic picture of DHVG (a) and CIDS (b). In DHVG, each data point is taken as a node. The different horizontal arrows in (a) denote the links between both nodes, black for two neighbored nodes, blue for two nodes separated by one node, red for two nodes separated by two nodes and green for two nodes separated by three nodes. For each given node, the links from the past nodes are defined as the ingoing degree and links from the future nodes as the outgoing degree with (k_{in}, k_{out}) shown for each node. The green vertical dash lines in (b) separate the increasing regimes from the decreasing regimes with the blue numbers for each decreasing regime of specific steps and the red numbers for each increasing regime of specific steps, and gray horizontal dash lines denote the given threshold with C times of the standard deviation of dT_i'.

3. Methods

3.1. DHVG and CIDS

For a real-value time series $\{x_t\}_{t=1,\ldots,N}$ of N data points, according to DHVG, each data point is taken as a node and links between two nodes i and j in the graph are calculated according to the following geometrical criterion [3,12,19,37,38]: $x_n < x_i, x_j, \forall \quad n|i < n < j$. This geometrical criterion will lead to that each node has its own degree, see details in Fig. 1a from DHVG schematic picture. For each given node, the links from the past nodes are defined as the ingoing degree and links from the future nodes as the outgoing degree, i.e. the degree of each node can be divided into $k(t) = k_{in}(t) + k_{out}(t)$ with an ingoing degree $k_{in}(t)$ and an outgoing degree $k_{out}(t)$, from which the probability distributions of ingoing degree $k_{in}(t)$ or outgoing degree $k_{out}(t)$ can be calculated as $p_{in}(k)$ and $p_{out}(k)$.

And for this real-value time series $\{x_t\}_{t=1,\ldots,N}$, two states can be classified for each data point: one is decreasing, the other is increasing. And these two states can be seen more clearly in the increment series, see details in Fig. 1b from CIDS schematic picture, where $N-1$ decreasing or increasing states exist. An important feature in Fig. 1b from CIDS schematic picture is that increasing or decreasing state is not isolated but appears in a consecutive way, i.e. an increasing or decreasing state prefers to follow another increasing or decreasing state within a increasing or decreasing regime. It should be pointed out that although increment series is considered in CIDS, not only sign information is included in its definition, certain ordinal pattern information within a regime is also taken into account. Each increasing or decreasing regime size with different steps s can be calculated. For linear processes the maximum increasing or decreasing regime size is almost the same, however, for series from nonlinear processes the maximum increasing regime size is usually different from the maximum decreasing regime size. For example, the maximum increasing regime size is around 12, but the maximum decreasing regime size is around 8 for daily mean temperature anomaly series over China [12]. And then probability distributions of s increasing or decreasing steps can be estimated as $p_i(s) \equiv p(s_i = s)$ and $p_d(s) \equiv p(s_d = s)$ [12].

3.2. Definition for TI measures

Firstly, measures for TI can be defined by absolute distance between two probability distributions $p(x)$ and $q(x)$ as follows [12]:

$$L(p, q) = \sum_{x\in\chi} |p(x)\log p(x) - q(x)\log q(x)|. \tag{1}$$

with $|10^{-m}\log 10^{-m}(m \to +\infty)| = \frac{m}{10^m}(m \to +\infty) \to 0$ and χ is a set of values which x can admit.

For $p_d(s)$ and $p_i(s)$, TI measure is

$$L_1(p_d, p_i) = \sum_{s} |p_d(s)\log p_d(s) - p_i(s)\log p_i(s)|, \tag{2}$$

while for $p_{in}(k)$ and $p_{out}(k)$, it is

$$L_2(p_{in}, p_{out}) = \sum_{k} |p_{in}(k)\log p_{in}(k) - p_{out}(k)\log p_{out}(k)|. \tag{3}$$

The measures (2) and (3) are comparable with two well defined measures [12], the Kullback–Leibler divergence [39,40] used by Lacasa et al. [3] and Euler distance between two distributions [39,40].

Other three TI measures are defined from air temperature anomaly increment time series $dT'_i = T'_{i+1} - T'_i, \quad i = 1, 2, \ldots\ldots N-1$ with T'_i length of N.

Measure A is defined [11] as

$$A = \frac{\sum_{i=1}^{i=N-1}\theta(dT'_i)}{\sum_{i=1}^{i=N-1}\theta(dT'_i) + \sum_{i=1}^{i=N-1}\theta(-dT'_i)}, \tag{4}$$

where $\theta(x) = 1$ for $x > 0$ and is zero otherwise.

If we define

$$N^+ = \sum_{i=1}^{i=N-1}\theta(dT'_i), \quad N^- = \sum_{i=1}^{i=N-1}\theta(-dT'_i), \tag{5}$$

then the measure A can be rewritten as

$$A = \frac{N^+}{N^+ + N^-} = \frac{N^+}{N-1}, \tag{6}$$

From this definition (6), it is obvious that measure A takes only the signs of air temperature anomaly increment time series dT'_i into account, and the effect from the magnitude of dT'_i is omitted. At the same time, the sign regime in CIDS is not taken into account. So if the strength of sign regime is weaker, then A and L_1 will reach the similar results. However, if the strength of sign regime is stronger, A and L_1 may lead to different results.

Taking the magnitude of dT'_i into consideration, the measure G is defined [17,18] as

$$G = \frac{\sum_{i=1}^{i=N-1}[dT'_i|_{dT'_i>0}]^2}{\sum_{i=1}^{i=N-1}(dT'_i)^2}, \tag{7}$$

from which we can learn that measure G is related to the ratio of magnitude of positive air temperature increments to the magnitude of total (positive plus negative) air temperature increments.

The last measure E is defined from the skewness of air temperature increment (dT'_i) [15,18]

$$E = \frac{\sum_{i=1}^{i=N-1}(dT'_i)^3}{[\sum_{i=1}^{i=N-1}(dT'_i)^2]^{\frac{3}{2}}}, \tag{8}$$

and it combines the contribution from both sign and magnitude of dT'_i.

From the definition of TI measures A (6), G (7) and E (8), we know that they quantify different aspects of air temperature increments dT'_i. If A and G are significant departure from $\frac{1}{2}$, it indicates the series under analyzing is irreversible [11]. Similarly, if E is significant departure from 0, it indicates the series under analyzing is also irreversible [15,17,18]. In order to compare these three measures clearly, we will consider $1-G$ and $1-E$ in later results.

Although these five TI measures are different from their definitions, they may be related to each other. We have mentioned that when the strength of sign regime is weaker, then A and L_1 will reach the similar results. And it was reported [12,38] that when the TI is weaker, both L_1 and L_2 give the consistent results for given series. If stationary condition is satisfied, TI results from measures A (6), G (7) and E (8) are also similar. At the same time, we should still keep in mind that their definitions are different and they reflect different aspects of TI features in underlying series, so we may have different results for the daily mean air temperature anomaly series, more or less.

4. Results

4.1. Consistent temporal irreversibility from different measures

First of all, a segment of typical daily mean air temperature anomaly series with slow increasing and rapid decreasing is shown in Fig. 1. From this piece of series, the TI feature is marked, where the numbers of increasing air temperatures are much larger than the numbers of decreasing air temperatures. Since the mean value of this piece of series is close to zero, above marked imbalance between the numbers of increasing and decreasing air temperatures will lead to stronger magnitudes of decreasing air temperatures than those of increasing air temperatures, especially beyond certain threshold, this magnitude imbalance becomes more predominant. The noticeable asymmetry of signs, magnitudes or both in daily mean air temperature anomaly series makes it a good test bed to compare the performance of above mentioned TI measures.

Consistent with previous findings [12], dominated TI emerges uniformly in nearly all observed daily mean air temperature anomaly variations over China, see Fig. 2a–e. All five measures show concordant spatial patterns. Higher TI occurs over South China, and lower TI over Northeast China and Southwest China close to Basin of Sichuan and Yungui-Tibetan Plateau. Perfect spatial coherence among these five TI measures indicates that the derived TI in the daily mean air temperature is robust to the choice of TI measures and it indeed exists. The specific regional TI patterns suggest that the origin of TI in the daily mean air temperature is closely related to the regional weather and climate conditions, detailed mechanism deserves further studies. This spatial coherence is further revealed in two measures scatter plots. Taking measure A as horizontal ordinate, other four measures as vertical ordinate, consistence between two measures is presented in Fig. 3a–d. All four measures ($1-G$, $1-E$, L_1 and L_2) increase linearly monotonically with increasing A. Nearly perfect monotonous linear variations between measure A and L_1 (or L_2) indicates there is perfect one-to-one correspondence between measure A and L_1 (or L_2) estimated from observed daily mean air temperature anomaly variations over each station. Since certain detailed results for L_1 and L_2 have been reported in literature [12] on the perfect one-to-one correspondence between measure A and L_1 (or L_2), we will only show comparisons between A and $1-G$ or $1-E$ in the rest of parts. Before this comparison, we want to discuss more about the performance of these five TI measures. As we show in the Methods section, these five TI measures reflect different aspects of TI features in the daily mean air temperature variability. There is a close relation between L_1 and A, since both of them is defined from the sign information of the daily mean air temperature increment. Difference between them is that only the sign information of the increment is considered, whereas both the sign information of the increment and its temporal patterns are taken into account in measure L_1, if the strength of temporal patterns in the sign information of the increment is weak, then both L_1 and A will reach the same results. Different from the definition for A, the definition of TI measures $1-G$ or $1-E$ is related to the PDF of the daily mean air temperature increment, more related to the magnitude information of the increment, so $1-G$ and $1-E$ can be different from A if there are more non-stationary trends or non-Gaussian components in the underlying series. From Fig. 3a and Fig. 3b, we can find that although there are one-to-one correspondence between measure A and $1-G$ or $1-E$, the dispersion in the scatter plot is much larger than those for L_1 or L_2. This larger dispersion may be caused by the seasonal standard deviations, which have not been removed from the air temperature anomaly series and they can affect the estimation of $1-G$ or $1-E$, since both of them are closely related the estimation of PDF of the air temperature anomaly increment series. Whereas TI measures A, L_1 and L_2 are insensitive to these non-stationary seasonal standard deviations (Figures not shown here), so there are more consistent results among them.

4.2. Origin for temporal irreversibility

Since TI in daily mean air temperature anomaly series can well be quantified by all five measures, and five measures quantify different aspects of daily mean air temperature anomaly series, what contributes to this consistent TI in daily mean air temperature anomaly series? In order to answer this question, we seek the patterns hidden in the daily mean air temperature anomaly series. From the marked feature presented in Fig. 1, we learn that dominated TI in the daily mean air temperature anomaly series may be related to the day-to-day (inter-diurnal) variations of the daily mean air temperature anomalies. Further studies show that it is directly related to inter-diurnal extreme variations of the daily mean air temperature anomalies.

In order to estimate this kind of inter-diurnal extreme variations, we define numbers of conditional increasing and decreasing as

$$N^{+C} = \sum_{i=1}^{i=N-1} \theta(dT_i' - C\sigma(dT_i')), \quad N^{-C} = \sum_{i=1}^{i=N-1} \theta(-dT_i' - C\sigma(dT_i')), \tag{9}$$

where $\sigma(dT_i') = \sqrt{\frac{1}{N-1}\sum_{i=1}^{i=N-1}(dT_i' - \overline{dT_i'})^2}$ is the standard deviation of temperature increments (dT_i'), and C is constant denoting the strength of a given threshold.

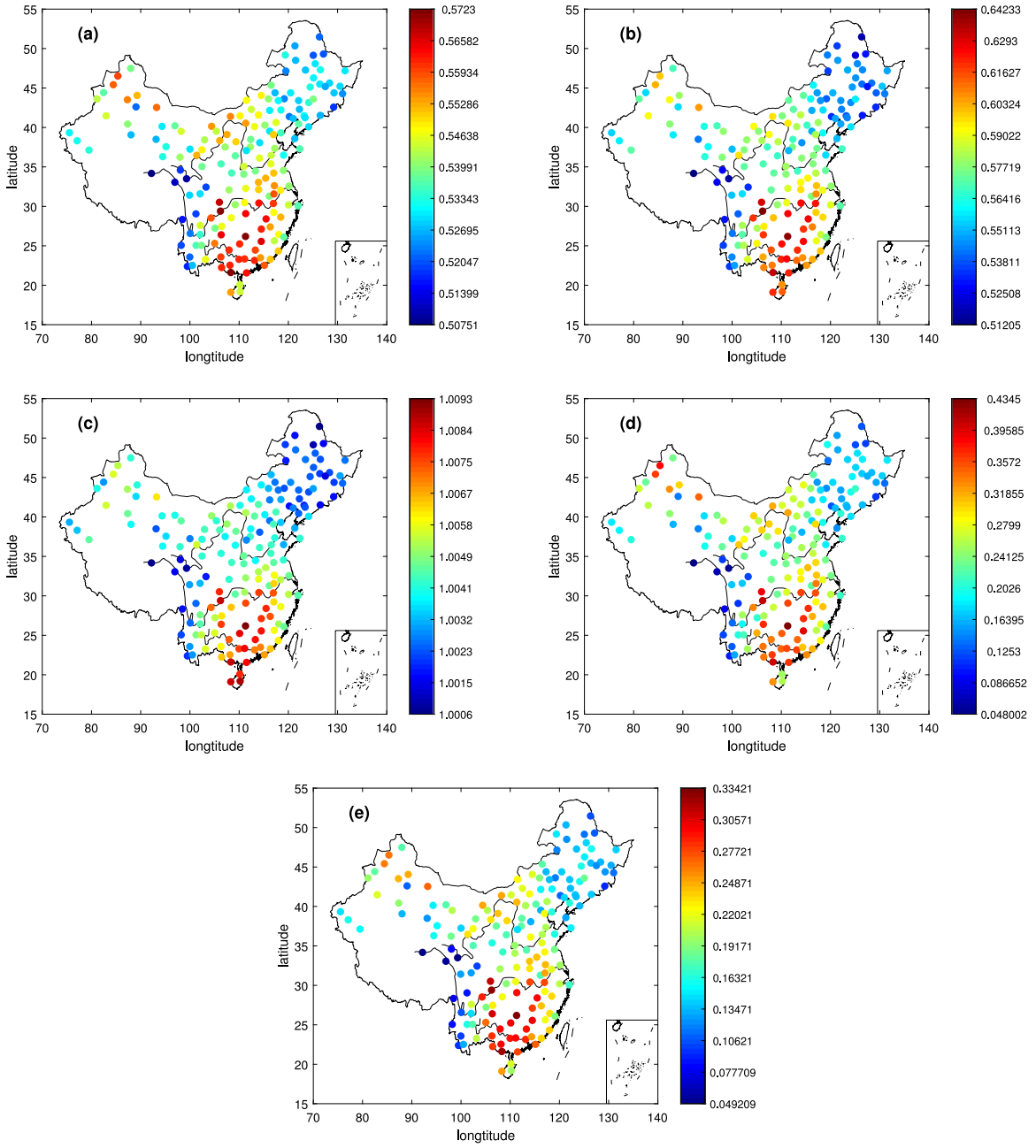

Fig. 2. Spatial distribution for different TI measures (a) A, (b) $1-G$, (c) $1-E$, (d) L_1, (e) L_2 from 176 stations over China.

And then the relative number difference R between conditional decreasing $dT_i' < 0|_{|dT_i'|>C\sigma(dT_i')}$ and conditional increasing $dT_i' > 0|_{|dT_i'|>C\sigma(dT_i')}$ for different thresholds Cs can be defined as

$$R = \frac{N^{-C} - N^{+C}}{N-1}, \tag{10}$$

If $C = 0$, then Eq. (10) can be rewritten as

$$R = 1 - 2A. \tag{11}$$

Eq. (11) indicates that there is a negative correlation between A and R when the threshold $C = 0$. Actually, when the threshold strength C is weak (such as $C = 0.1$ or $C = 0.5$), the spatial distribution of R is contrasted to what has

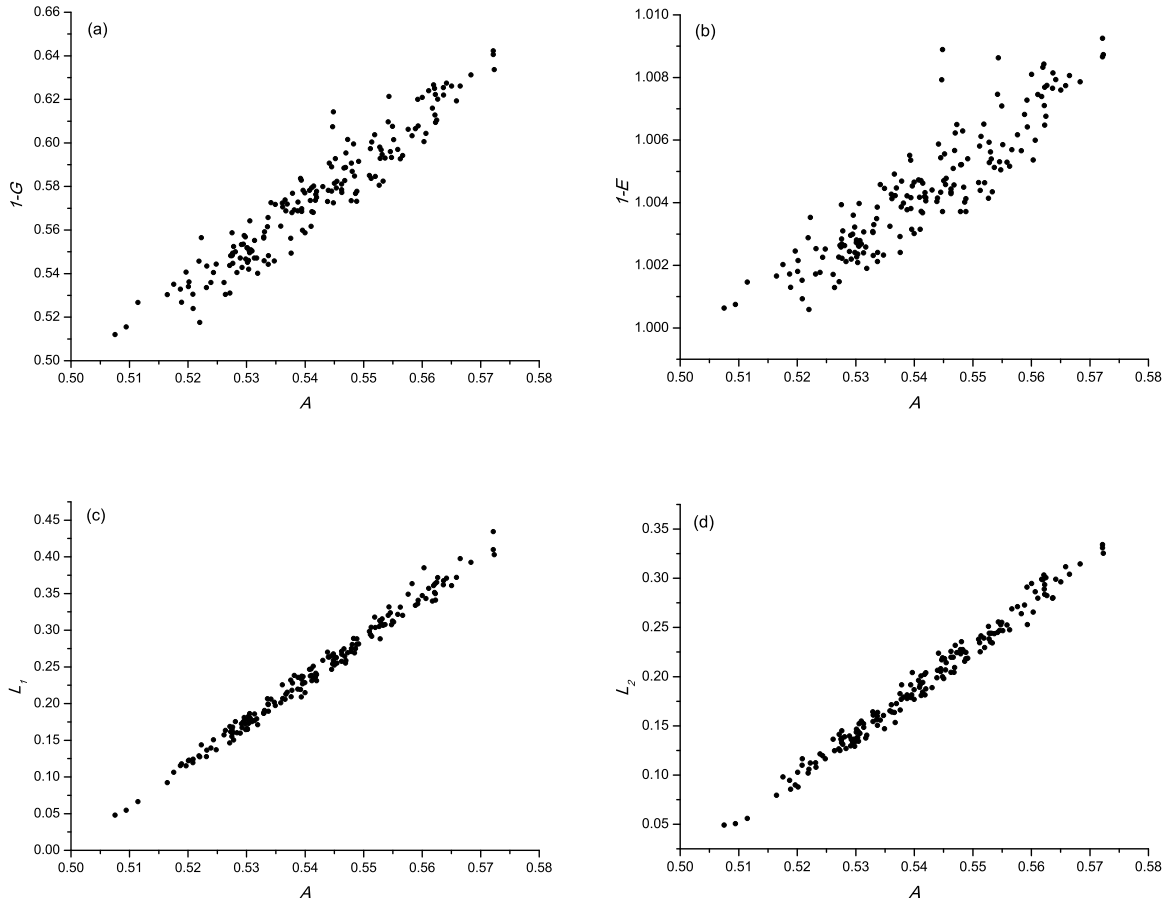

Fig. 3. Scatter plots for (a) $1-G$ V.S. A, (b) $1-E$ V.S. A, (c) L_1 V.S. A, (d) L_2 V.S. A from 176 stations over China.

happened in all five TI measures. The regions of higher (lower) value R correspond to the regions of the lower (higher) value TI measures, see Fig. 4a, b and Fig. 2a–e. This negative correlation is more obvious in the scatter plots of R versus A under different smaller thresholds Cs, see Fig. 5a and b. As the threshold increases further (e.g. $C = 1.0$), the spatial correspondence between R and five TI measures will be lost, see Fig. 4c. Especially the well-defined correlation between R and five TI measures becomes almost completely obscured, see Fig. 5c. As the threshold close to the extreme level (such as $C = 1.5$ or $C = 2.0$), the spatial distribution of R (see Fig. 4d and e) is consistent with what has happened in all five TI measures (see Fig. 2a–e). The regions of higher (lower) value R correspond to the regions of the higher (lower) value TI measures. The negative correlation between R and A found for weak threshold ($C = 0.1$ or $C = 0.5$) changes to positive correlation for stronger threshold ($C = 1.5$ or $C = 2.0$), see Fig. 5d and e. The state is stable for certain stronger thresholds, the spatial distribution of R is nearly saturated and there is only a little change, see Fig. 4d–f for thresholds $C = 1.5$, $C = 2.0$ and $C = 2.5$. Similarly, the positive correlation is stable for certain stronger thresholds, see Fig. 5d–f. All these behaviors found above indicate that the inter-diurnal extreme variations beyond certain critical value (such as 2 standard deviation of air temperature increments) contribute greatly to the dominated TI in the daily mean air temperature anomalies. These findings are interesting and they further confirm that stronger cold or warm fronts are expected to result in stronger asymmetry [11]. In the following subsection, we will show that the stronger asymmetry is closely related to the extreme decreasing variations but not the extreme increasing variations.

4.3. Seasonal TI patterns from five measures

Actually, there are predominant seasonal variations in surface air temperature series. Although the annual cycle has been removed from the original daily mean air temperature variations, there are still marked seasonal fingerprints. Will these unremoved seasonal components effect the estimated consistent TI from five measures? And will mechanism contributing to TI in daily mean air temperature anomalies change in each season? We will give answer to these questions in the following parts. First of all, measures A, $1-G$ and $1-E$ in each season were calculated. All measures show consistent spatial distribution in each season, see Fig. 6. However, there are still seasonal variations for all measures. Among four

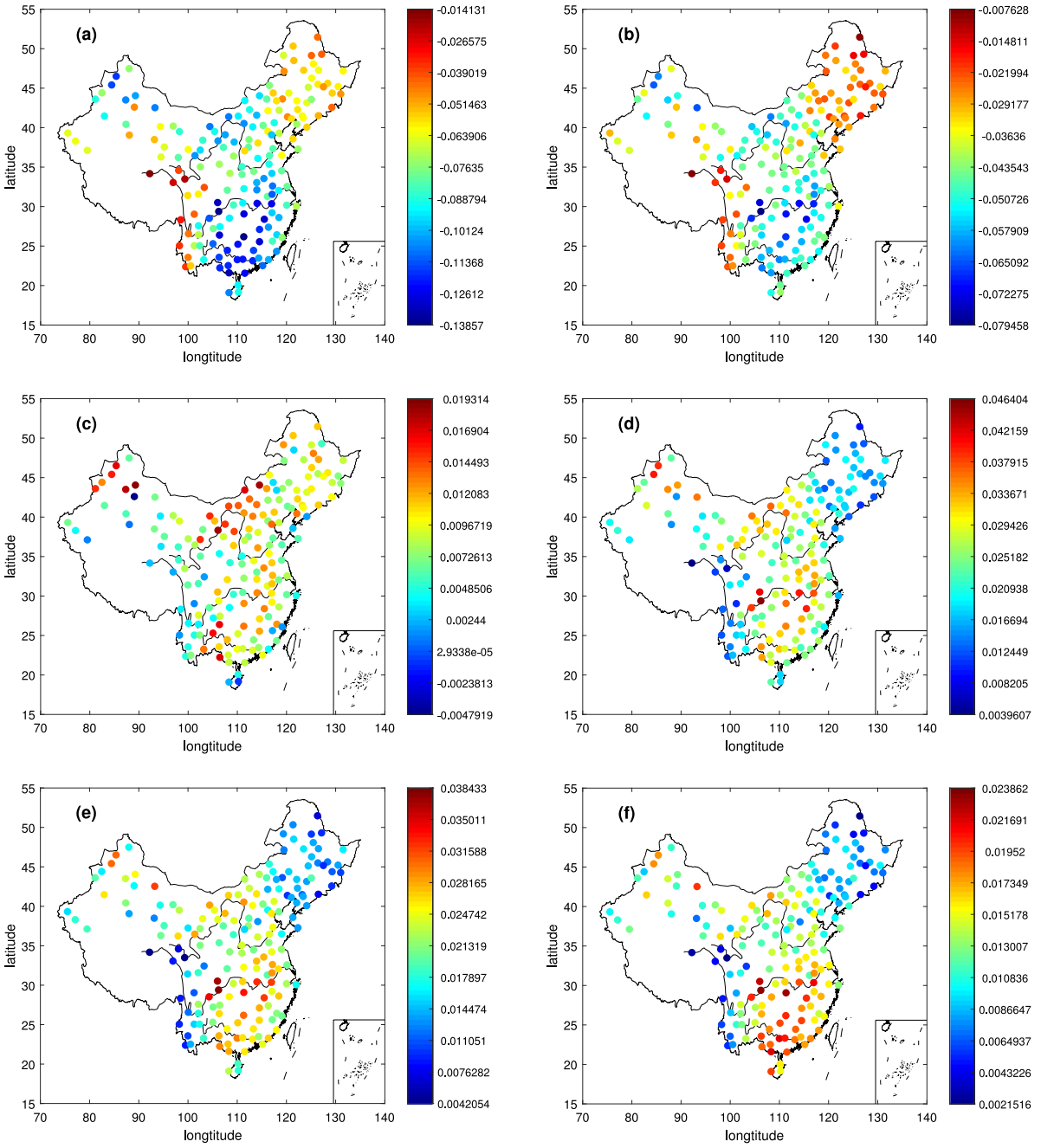

Fig. 4. Spatial distribution for relative number difference R between conditional decreasing and increasing under different thresholds Cs (a) $C = 0.1$, (b) $C = 0.5$, (c) $C = 1.0$, (d) $C = 1.5$, (e) $C = 2.0$, (f) $C = 2.5$ from 176 stations over China.

seasons, TI spatial patterns in the autumn (Fig. 6c, g and k) closely resemble those given from whole year (Fig. 2a–c). There are certain overestimated TI strength over Northeast China in the spring (Fig. 6a, e and i) and underestimated TI over Northwest and Central China in the winter (Fig. 6d, h and l), and very well consistence over other regions for both seasons. Only TI spatial patterns in the summer (Fig. 6b, f and g) overall resemble not so well those given from whole year. These results indicate the unremoved seasonal components affect only a little the estimated consistent TI from five measures.

Similarly, there are seasonal variations of R under extreme conditions, see Fig. 7a–d for $C = 2.5$, and the seasonal spatial patterns of R closely resemble the seasonal spatial patterns of each TI measure. Specifically, the seasonal patterns of R under extreme conditions in each season are mainly related to the spatial patterns of extreme relative decreasing

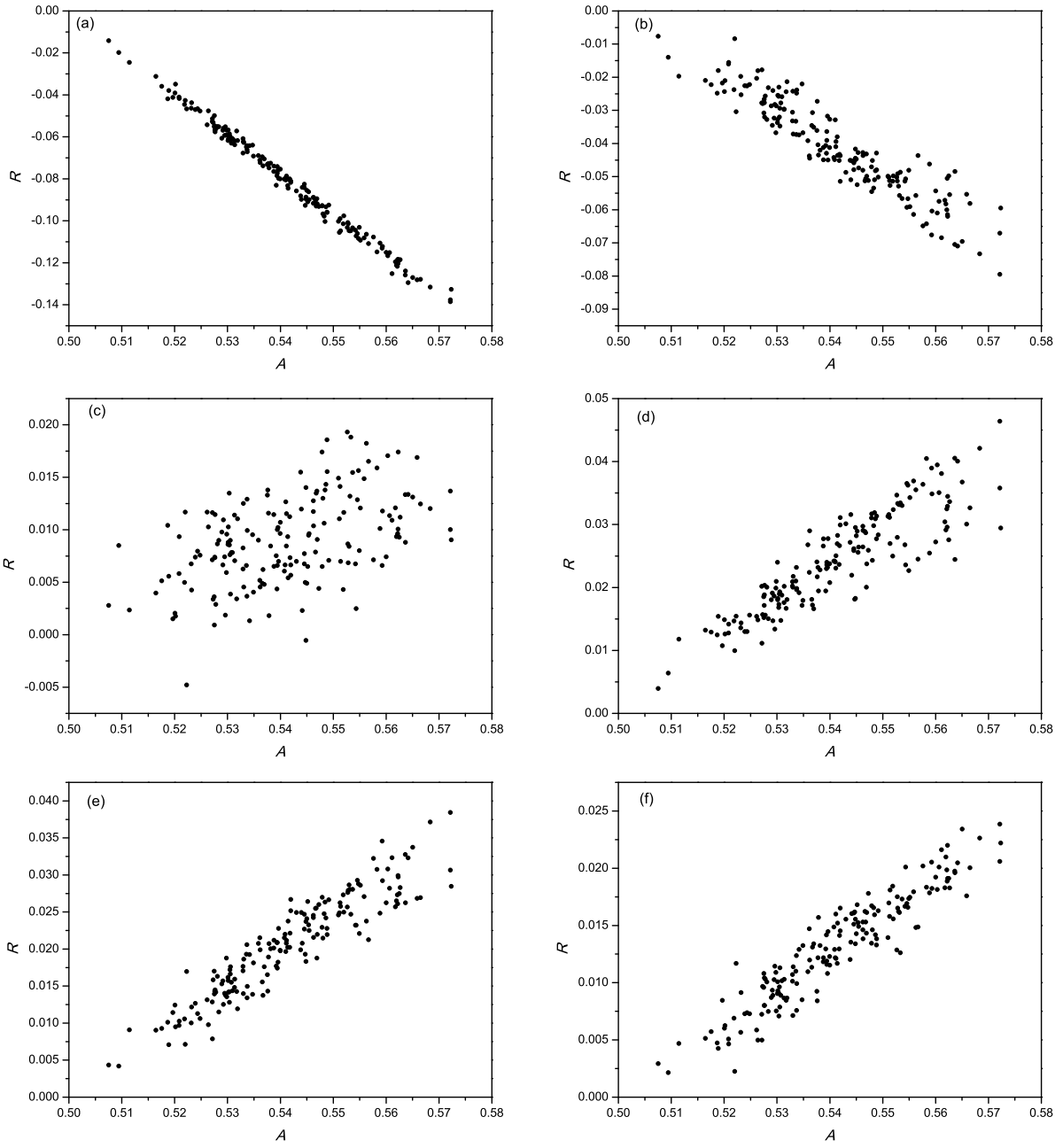

Fig. 5. Scatter plots for relative conditional decreasing and increasing number difference R versus TI measure A under different thresholds Cs (a) $C = 0.1$, (b) $C = 0.5$, (c) $C = 1.0$, (d) $C = 1.5$, (e) $C = 2.0$, (f) $C = 2.5$ from 176 stations over China.

air temperature numbers $N^{-C}/(N-1)$ under the same condition, see Fig. 7a–d and Fig. 7e–h. The spatial distributions of extreme relative increasing air temperature numbers $N^{+C}/(N-1)$ under the same condition are nearly homogeneous with much lower values in each season, see Fig. 7i–l, so the contribution of extreme increasing air temperature numbers N^{+C} under this extreme condition to seasonal variations in R under the same conditions is negligible.

Since the seasonal variations in R under extreme conditions are closely related to the seasonal spatial patterns of each TI measure, this coherence can be revealed in the scatter plots between R and each TI measure in each season, this is confirmed in Fig. 8. There is good linear positive correlation between R and each TI measure in the summer, autumn and winter, but positive correlation with weak nonlinear feature between R and each TI measure in spring.

All these findings show that the seasonal variations for each TI measure are also consistent and they are all related to the seasonal variations of R under extreme conditions.

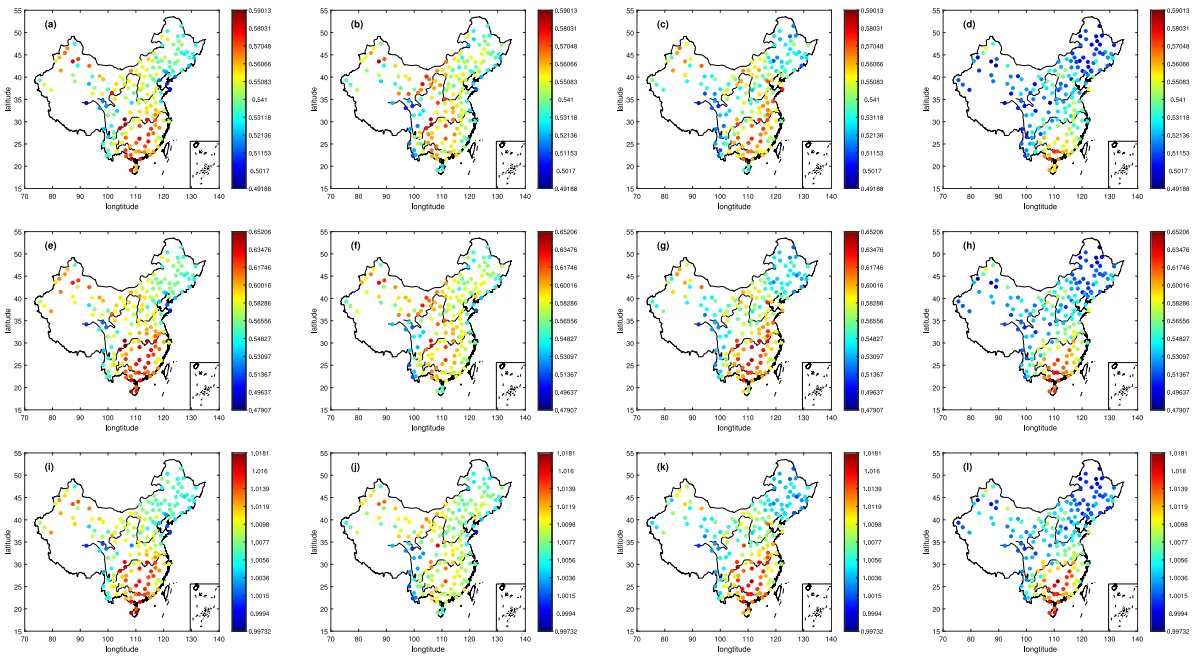

Fig. 6. Spatial distribution for different TI measures over four seasons. The first column for spring, the second column for summer, the third for autumn and the last column for winter; the first row for A (a, b, c, d), the second row for $1-G$ (e, f, g, h) and the last row for $1-E$ (i, j, k, l) from 176 stations over China.

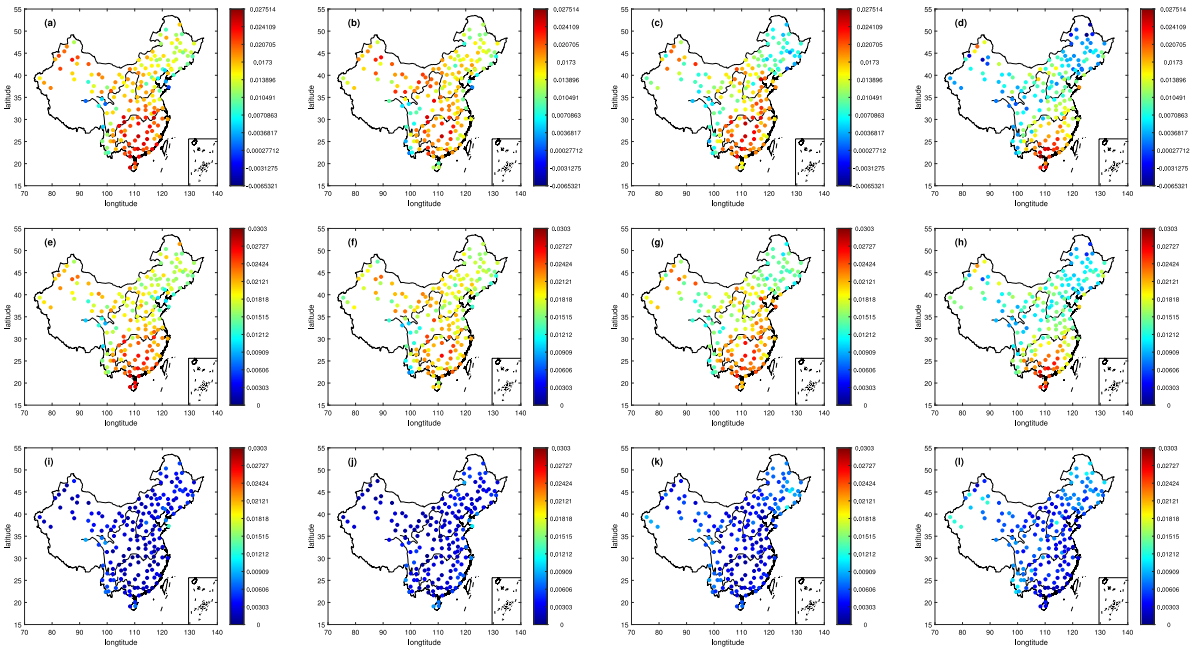

Fig. 7. Spatial distribution for R, relative decreasing number and relative increasing number under threshold $C = 2.5$ over four seasons. The first column for spring, the second column for summer, the third for autumn and the last column for winter; the first row for R (a, b, c, d), the second row for relative decreasing number (e, f, g, h) and the last row for relative increasing number (i, j, k, l) from 176 stations over China.

5. Conclusion and discussion

As a fundamental property of nonlinear time series, temporal irreversibility (TI) behaviors of daily mean air temperature from observations over China were quantified by five measures (A, G, E, L_1 and L_2). The results show that five measures

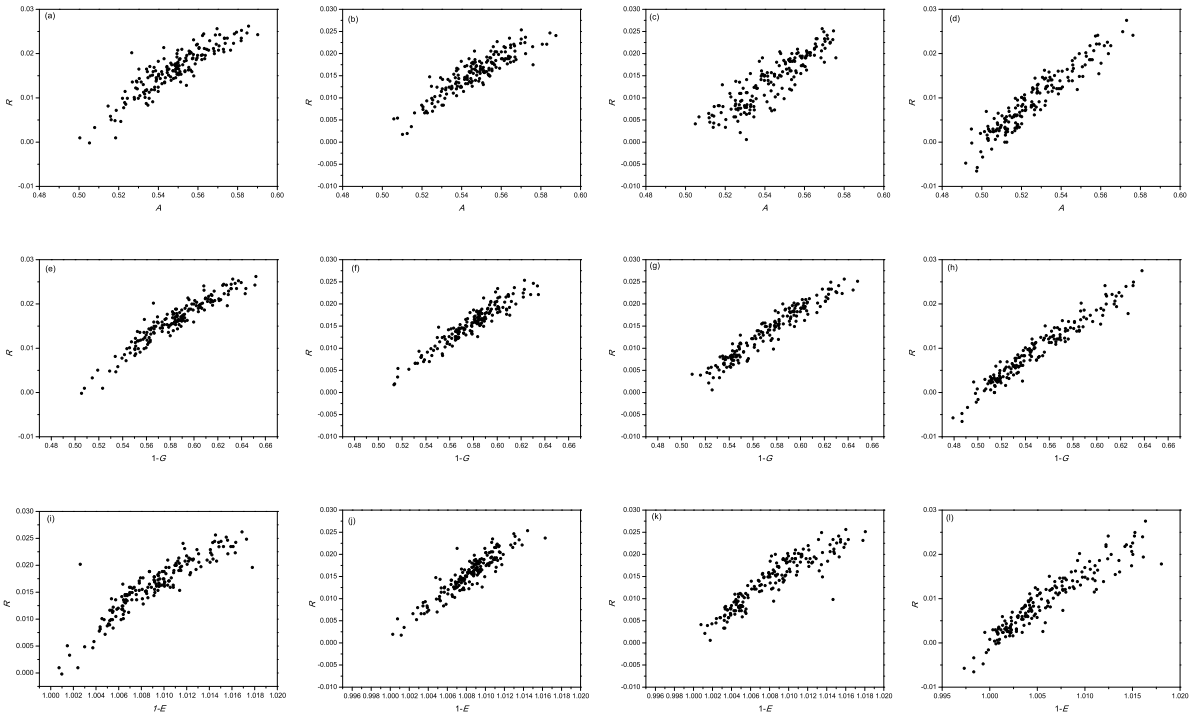

Fig. 8. Scatter plots for R versus different TI measures under threshold $C = 2.5$ over four seasons. The first column for spring, the second column for summer, the third for autumn and the last column for winter; the first row for A (a, b, c, d), the second row for $1 - G$ (e, f, g, h) and the last row for $1 - E$ (i, j, k, l) from 176 stations over China.

all reach consistent conclusions to TI of daily mean air temperature from observations over China. This finding indicates that all five TI measures have the similar performance, and they all can be applied to quantify the TI behaviors hidden in time series from various fields. Since time series from natural processes inevitable suffer from non-stationarity, results under these cases from measures A, G and E may be biased. While measures L_1 and L_2 are robust to non-stationarity [3,12], therefore, if we are not sure that the time series under analysis are stationary or not, measure L_1 or L_2 is recommended.

Daily mean air temperature from both NCEP reanalysis [11] and observations [12] has been reported on TI or asymmetry studies, however, only a few detailed regional studies of TI or asymmetry have been carried out. The results given in this paper from five TI measures all show that there are dominated regional TI features. The higher TI occurs over South China, and lower TI over Northeast China and Southwest China close to Basin of Sichuan and Yungui-Tibetan Plateau. It has been suggested that cooling rapidly and warming gradually asymmetry in daily mean air temperature over mid-latitudes is partially related to cold fronts [11]. However, distinct temporal irreversible structures over the same latitudes indicate that region-dependent TI is also related directly to the regional weather and climate conditions. So the strength of TI is process-dependent, incorporating nonlinearity to correctly model this process is of great importance to understand the nature of time irreversibility or temporal asymmetry.

Further studies indicate that the dominated TI behavior is related to the number difference between extreme decreasing air temperature and extreme increasing air temperature variations. When conditional threshold exceeds certain critical values, this marked TI can be well quantified by the relative extreme decreasing and increasing number difference. At the same time, different TI measures over four seasons show the similar spatial patterns to those of difference between relative decreasing number and relative increasing number under certain threshold, which are nearly totally controlled by spatial distributions of extreme relative decreasing numbers. This new finding suggests that TI in the daily mean air temperature anomalies is closely related to the extreme air temperature variations. And this finding is also contrary to the statistical findings in the literature that fluctuations of many atmospheric variables can be well approximated by low order autoregressive (AR) processes on small time scales [41] and higher than the second order AR models are rarely needed for daily surface air temperature records after proper detrending [30]. However, the results shown in this paper indicate that the daily mean air temperature series is TI and it is output from nonlinear but not linear process. Since so many daily mean air temperature anomaly series from both observations and reanalysis are time series irreversible or temporal asymmetry, in order to properly describe the nonlinear features of daily mean temperature anomaly variations, a new statistical model is required. At the same time, quality assessment of reanalysis data is a hot topic. Many new measures have been applied to evaluate reanalysis data from different aspects, such as long-range correlation [42,43]. TI can be taken as a novel measure to evaluate the quality of reanalysis based on nonlinear features, and this will be reported in the later studies.

At last, we should stress that although the overall TI feature from five TI measures are consistent. However, there are still some differences among them since their definitions are different and they reflect different aspects of TI in series. First of all, L_1, L_2 and A is more robust to non-stationary low-frequency trends in underlying series, but $1-G$ and $1-E$ is not, since both of them are closely related to the estimated PDF of underlying series. For L_1 and L_2, L_1 is not so sensitive to the extreme values if the local ordinal patterns are unchanged, but L_2 not [38]. Secondly, L_1 is more robust to the measured noise than L_2 (figure not shown here). Thirdly, there is a close relation between L_1 and A, since both of them is defined from the sign information of the daily mean air temperature increment. Difference between them is that only the sign information of the increment is considered, whereas both the sign information of the increment and its temporal patterns are taken into account in measure L_1, if the strength of temporal patterns in the sign information of the increment is stronger, then both L_1 can reveal more detailed TI features than A does. So, combining all aspects considered, L_1 from CIDS is the best among these TI measures.

Acknowledgments

Many thanks are due to valuable suggestions from anonymous reviewers and supports from National Natural Science Foundation of China (No. 41475048 and No. 41705041).

References

[1] G. Weiss, Time-reversibility of linear stochastic processes, J. Appl. Probab. 12 (1975) 831–836.
[2] E. Roldan, J.M.R. Parrondo, Estimating dissipation from single stationary trajectories, Phys. Rev. Lett. 105 (2010) 15.
[3] L. Lacasa, A. Nuñez, E. Roldan, J.M.R. Parrondo, B. Luque, Time series irreversibility: A visibility graph approach, Eur. Phys. J. B 85 (2012) 217.
[4] C. Diks, J.C. van Houwelingen, F. Takens, J. DeGoede, Reversibility as a criterion for discriminating time series, Phys. Lett. A 201 (1995) 221–228.
[5] L. Stone, G. Landan, R.M. May, Detecting time's arrow: A method for identifying nonlinearity and deterministic chaos in time-series data, Proc. R. Soc. Lond. B 263 (1996) 1509–1513.
[6] A.C.C. Yang, S.S. Huseu, H.W Yien, A.L. Goldberger, C.K. Peng, Linguistic analysis of the human heartbeat using frequency and rank order statistics, Phys. Rev. Lett. 90 (2003) 108103.
[7] M. Costa, A.L. Goldberger, C.K. Peng, Broken asymmetry of the human heartbeat: Loss of time irreversibility in aging and disease, Phys. Rev. Lett. 95 (2005) 198102.
[8] C. Cammarota, E. Rogora, Time reversal, symbolic series and irreversibility of human heartbeat, Chaos Solitons Fractals 32 (2006) 1649–1654.
[9] M. Costa, C.K. Peng, A.L. Goldberger, Multiscale analysis of heart rate dynamics: Entropy and time irreversibility measures, Cardiovasc. Eng. (2008) http://dx.doi.org/10.1007/s10558-007-9049-1.
[10] J.F. Donges, R.V. Donner, J. Kurths, Testing time series irreversibility using complex network methods, Europhys. Lett. 102 (2013) 10004.
[11] Y. Ashkenazy, Y. Feliks, H. Gildor, E. Tziperman, Asymmetry of daily temperature records, J. Atmos. Sci. 65 (2008) 3327.
[12] F.H. Xie, Z.T. Fu, L. Piao, J.Y. Mao, Time irreversibility of mean temperature anomaly variations over China, Theor. Appl. Clim. 123 (2016) 161.
[13] Z.T. Fu, L. Shi, F.H. Xie, L. Piao, Nonlinear features of Northern annular mode variability, Physica A 449 (2016) 390.
[14] Y. Ashkenazy, E. Fredj, et al., Current temporal asymmetry and the role of tides: Nan-Wan Bay Vs the Gulf of Elat, Ocean. Sci. 12 (2016) 733.
[15] C.L. Ehlers, J. Havstad, D. Prichard, J. Theiler, Low doses of ethanol reduce evidence for nonlinear structure in brain activity, J. Neurosci. 18 (1998) 7474–7486.
[16] C.S. Daw, C.E.A. Finney, M.B. Kennel, Symbolic approach for measuring temporal irrevetsibility, Phys. Rev. E 62 (2000) 1912–1921.
[17] P. Guzik, J. Piskorski, T. Krauze, A. Wykretowicz, H. Wysocki, Heart rate asymmetry by Poincare plots of RR intervals, Biomed. Tech. 51 (2006) 272–275.
[18] A. Porta, K.R. Casali, A.G. Casali, T. Gnecchi-Ruscone, E. Tobaldini, N. Montano, S. Lange, D. Geue, D. Cysarz, P. Van Leeuwen, Temporal asymmetries of short-term heart period variability are linked to autonomic regulation, Am. J. Physiol. Regul. Integr. Comp. Physiol. 295 (2008) R550–R557.
[19] B. Luque, L. Lacasa, F. Ballesteros, J. Luque, Horizontal visibility graphs: Exact results for random time series, Phys. Rev. E 80 (2009) 046103.
[20] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, J.C. Nuño, From time series to complex networks: The visibility graph, Proc. Natl. Acad. Sci. 105 (2008) 4972–4975.
[21] A. Burykin, M. Costa, C.K. Peng, A.L. Goldberger, T.G. Buchman, Generating signals with multiscale time irreversibility: The asymmetric weierstrass function, Complexity 16 (2011) 29.
[22] E. Koscielny-Bunde, A. Bunde, S. Havlin, H.E. Roman, Indication of a universal persistence law governing atmospheric variability, Phys. Rev. Lett. 81 (1998) 729.
[23] A. Király, I.M. Jánosi, Stochastic modeling of daily temperature fluctuations, Phys. Rev. E 65 (2002) 051102.
[24] A. Bunde, J.F. Eichner, J.W. 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.
[25] P.M. Zhai, X.L. Pan, Trends in temperature extremes during 1951∼1999 in China, Geophys. Res. Lett. 30 (2003) 1913.
[26] 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.
[27] N.M. Yuan, Z.T. Fu, J.Y. Mao, Different scaling behaviors in daily temperature records over China, Physica A 389 (2010) 4087.
[28] N.M. Yuan, Z.T. Fu, J.Y. Mao, Different multi-fractal behaviors of diurnal temperature range over the north and the south of China, Theor. Appl. Climatol. 112 (2013) 673.
[29] N.M. Yuan, Z.T. Fu, Century-scale intensity modulation of large-scale variability in long historical temperature records, J. Clim. 27 (2014) 1742.
[30] I. Bartos, I.M. Jánosi, Atmospheric response function over land: Strong asymmetries in daily temperature fluctuations, Geophys. Res. Lett. 32 (2005) L23820.
[31] B. Gyüre, I. Bartos, I.M. Jánosi, Nonlinear statistics of daily temperature fluctuations reproduced in a laboratory experiment, Phys. Rev. E 76 (2007) 037301.
[32] F.Y. Lu, N.M. Yuan, Z.T. Fu, J.Y. Mao, Universal scaling behaviors of meteorological variables' volatility and relations with original records, Physica A 391 (2012) 4953.
[33] T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests, Phys. Rev. Lett. 77 (1996) 635–638.
[34] J.F. Eichner, J.W. Kantelhardt, A. Bunde, S. Havlin, Statistics of return intervals in long-term correlated records, Phys. Rev. E 75 (2007) 011128.
[35] H.A. Makse, S. Havlin, M. Schwartz, H.E. Stanley, Method for generating long-range correlations for large systems, Phys. Rev. E 53 (1996) 5445.

[36] R.B. Govindan, J.D. Wilson, H. Preißl, H. Eswaran, J.Q. Campbell, C.L. Lowery, Detrended fluctuation analysis of short datasets: An application to fetal cardiac data, Physica D 226 (2007) 23.
[37] R.V. Donner, Y. Zou, J.F. Donges, N. Marwan, J. Kurths, Recurrence networks: A novel paradigm for nonlinear time series analysis, New J. Phys. 12 (2010) 033025.
[38] F.H. Xie, D. Nian, Z.T. Fu, Differential temporal asymmetry among different temperature variables' daily fluctuations, Clim. Dyn. (2019) http://dx.doi.org/10.1007/s00382-018-04603-1.
[39] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley & Sons, Inc., New Jersey, 2006.
[40] A.M. Kowalski, M.T. Martin, A. Plastino, O.A. Rosso, M. Casas, Distances in probability space and the statistical complexity setup, Entropy 13 (2011) 1055–1075.
[41] H. von Storch, F.W. Zwiers, Statistical Analysis in Climate Research, Cambridge Univ. Press, Cambridge, 1999.
[42] S.S. Zhao, W.P. He, Y.D. 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.
[43] W.P. He, S.S. Zhao, Assessment of the quality of NCEP-2 and CFSR reanalysis daily temperature in China based on long-range correlation, Clim. Dyn. 50 (2018) 493.