Evaluation of Re-analyses Over China based on the Temporal Asymmetry of Daily Temperature Variability

Wenmi Chai
Peking University
Yu Huang
Peking University
Lichao Yang
LASG, Institute of Atmospheric Sciences, Chinese Academy of Sciences
Heng Quan
Peking University
Zuntao Fu (fuzt@pku.edu.cn)
Peking University https://orcid.org/0000-0001-9256-8514

Research Article

Keywords: Temporal asymmetry , Re-analysis data, Temperature variability

Posted Date: May 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-493801/v1

Version of Record: A version of this preprint was published at Theoretical and Applied Climatology on November 18th, 2021. See the published version at https://doi.org/10.1007/s00704-021-03839-y.

Evaluation of re-analyses over China based on the temporal asymmetry of daily temperature variability

Wenmi Chai[1], Yu Huang[1], Lichao Yang[2], Heng Quan[1], Zuntao Fu[1,*]

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

[2] LASG, Institute of Atmospheric Sciences, Chinese Academy of Sciences, P.O. Box 9804, Beijing 100029, China

* **Correspondence**: Zuntao Fu, Lab for Climate and Ocean-Atmosphere Studies, Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China. Email: fuzt@pku.edu.cn

Key points:

- Four re-analyses all overestimates the temporal asymmetry of daily maximum and minimum temperature variability over China.
- For overestimated temporal asymmetry, four re-analyses own the similar region-dependent spatial patterns.
- Daily mean air temperature variability of re-analyses is the most suitable for extreme event study.

Abstract

As an intrinsic feature of daily surface air temperature (SAT) variability found in station measurements, temporal asymmetry (TA) can be taken as an evaluation metric to access the quality of SAT re-analysis product. In this study, TA calculated from four SAT variables, i.e. daily mean SAT (T_{mean}), daily maximum SAT (T_{max}), daily minimum SAT (T_{min}) and diurnal temperature range ($T_{DTR}=T_{max}-T_{min}$), is applied to evaluate synoptic-scale performance of four reanalysis products (NCEP-2, JRA-55, ERA-I and ERA-5) over China. The results show that four re-analyses overall overestimate the TA of daily T_{max} and T_{min} variability over China, but with a comparatively consistent estimated TA for T_{mean}. Moreover, the TA of T_{mean} variability for these four re-analyses shares high spatial consistency with those from the observation. However, four re-analyses own the similar region-dependent spatial patterns of overestimated TA for T_{max} and T_{min} variability, especially for T_{max}. Since high TA is an indicator for strong nonlinear feature, only T_{mean} reanalysis is the most suitable to explore synoptic-scale extreme events, such as heat waves and cold waves, which are highly related to the strong nonlinear processes.

Keywords: Temporal asymmetry · Re-analysis data· Temperature variability

1. Introduction

Asymmetric phenomena are ubiquitous in both natural and social sciences (Heinrich 2004; King 1996; Livina et al. 2003; Ashkenazy and Tziperman 2004 ; Lisiecki and Raymo 2005; Bartos and Jánosi 2005; Gyure et al. 2007; Ashkenazy et al. 2008, 2016; Bisgaard and Kulahci 2011; Xie et al. 2016, 2019), and it is an important indicator of nonlinear underlying processes (Schreiber and Schmitz 2000; Bartos and Jánosi 2005; Gyüre et al. 2007; Ashkenazy et al. 2008; Roldan and Parrondo 2010; Lacasa et al. 2012).

As a kind of asymmetry, the temporal asymmetry (TA) in time series, defined by different statistics between forward and backward (reversed) directed series, plays an important role in air temperature variability studies (Bartos and Jánosi 2005; Gyüre et al. 2007; Ashkenazy et al. 2008; Xie et al. 2016, 2019). Previous studies also found that there exists differential TA among different temperature variables' daily fluctuations over China from both station observations and NCEP-2 re-analyses (Xie et al. 2019). For commonly used temperature variables, daily mean temperature (T_{mean}), daily maximum SAT (T_{max}), daily minimum SAT (T_{min}) and diurnal temperature range (T_{DTR}= T_{max}-T_{min}), TA strengths among these four temperature variables are markedly different with the weakest TA for DTR and the strongest TA for T_{mean}. Compared with the TA from station observations, TA from T_{max} and T_{min} in National Centers for Environmental Prediction (NCEP) AMIP-II Reanalysis (hereafter NCEP-2, Kanamitsu et al. 2002) is highly overestimated for most regions over China (Xie et al. 2019).

TA has been found to be an intrinsic feature in daily surface air temperature (SAT) variability in station measurements (Bartos and Jánosi 2005; Gyüre et al. 2007; Xie et al. 2016, 2019; Li et al 2021). Although similar TA behavior was also reported in re-analysis data (Ashkenazy et al. 2008; Xie et al. 2019; Li et al 2021), only a few studies (Xie et al. 2019; Li et al 2021) compared TA difference between station measurement and re-analysis products, detailed and systematic comparative studies are still required. More specifically, is it a universal feature to different kinds of re-analyses for the reported overestimated TA in T_{max} and T_{min} variability from NCEP-2 re-analyses over China (Xie et al. 2019; Li et al 2021)? If the answer is yes, is there any spatial consistency among different reanalysis products?

The reanalyzed products are outputs from the assimilation technology of numerical weather prediction to restore the observation data and compensate for the lack of uneven distribution of weather stations (Bengtsson et al. 1988; Trenberth et al. 1988; Chen et al. 2016). Previous studies used different methods to compare and evaluate reanalysis data at different temporal and spatial scales from the point of view of mean value, standard deviation, long-term trend, long-range correlation (LRC) and so on, and they found re-analyses don't always work well to reproduce consistent results with observations (Flocas et al. 2005; Ma et al. 2008; Pitman and Perkins. 2009; Mao et al. 2010; Marques et al. 2010; Mooney et al. 2011; Alfred et al. 2011; You et al. 2011, 2013; Cornes and Jones. 2013; Chen et al. 2014; Taguchi et al. 2017; Zhu et al. 2017; He et al. 2018; Alghamdi. 2020). For example, He et al. (2018) evaluated the air temperature reanalysis variables

by means of LRC and they found that NCEP-2 reanalysis overestimates LRC over some specific regions. Zhu et al. (2017) revealed that Interim European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-Interim, herein ERA-I, Dee et al. 2011) can capture the intensity indices of the continuous extreme temperature events, but not their frequency indices. Since TA is closely related to the synoptic-scale processes (Ashkenazy et al. 2008; Xie et al. 2016; Li et al. 2021) and nonlinear behaviors (Li et al. 2021), behaviors like extreme temperature events found in ERA-I should also be able to be revealed by TA.

Besides NCEP-2 reanalysis and ERA-I, other commonly used reanalysis products are available: the 55-yr Japanese Project (hereafter JRA-55, Ayataka et al. 2011; Kobayashi et al. 2015), and the Fifth generation ECMWF Re-Analysis (hereafter ERA-5, Radu et al. 2018; Hersbach et al. 2020). These four reanalysis products will be exploited to test whether there are universal conclusions on TA related to SAT variability.

Moreover, Ye and Hsieh (2008) found that increasing nonlinearity in ENSO and Lorenz systems can enhance their predictability by improving the contributions from the low-frequency variations. As one kind of nonlinearity, overestimated TA in the NCEP-2 reanalysis daily minimum and maximum air temperature anomaly series over China is concurrent with overestimated intrinsic predictability, prediction skill, and the occurrence number of extreme events (Li et al. 2021). This finding indicates that overestimated TA in reanalysis products may distort the conclusions on extreme event studies. By comparing the TA calculated from the station observations of different SAT variables to those from different reanalysis products, the most suitable SAT variable and the corresponding reanalysis products to estimate TA can be determined.

The paper is organized as follows. Section 2 summarizes the sources of observation and reanalysis data and the methods used in this study. Section 3 reports the overestimated TA in both T_{max} and T_{min} from observations and re-analyses, contrasting to the comparable consistency in TA for T_{mean}. And quantitative comparison of TA between observations and re-analyses for four different SAT variables is also provided. At last, Section 4 concludes this study with a brief discussion.

2. Data and methods

2.1 Data

2.1.1 Observations

The observational time series (hereafter observations) of T_{mean}, T_{max}, T_{min} and T_{DTR} were downloaded from China Meteorological Administration. After data quality check by removing the time series with missing points, datasets of 643 Chinese meteorological stations were finally selected for this study, covering a total of 40 years from 1979 to 2018.

2.1.2 Reanalysis

For the reanalysis data of 2m air temperature from 1979 to 2018, we selected four commonly used products: NCEP-2 (Kanamitsu et al. 2002), JRA-55 (Ayataka et al. 2011; Kobayashi et al., 2015), ERA-Interim (Dee et al. 2011) and ERA-5 (Radu et al. 2018; Hersbach et al. 2020). NCEP-2 products were acquired from the National Oceanic and Atmospheric Administration (NOAA) and its website at http://www.esrl.noaa.gov/psd. JRA-55 is produced by the Japan Meteorological Agency (JMA) (https://climatedataguide.ucar.edu/climate-data/jra-55). From the ECMWF's website: https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets, both ERA-I and ERA-5 are downloaded. Detailed information about these four reanalysis datasets is shown in **Table 1**.

Table 1 Details of the four reanalysis datasets used in this study

Dataset	Horizontal resolution	Output frequency	Assimilation method	Source
NCEP-2	1.875°×1.904°	6-hourly for analyses	3D-Var	NOAA
JRA-55	1.25°×1.25°	6-hourly for analyses	4D-Var	JMA
ERA-I	0.75°×0.75°	6-hourly for analyses	4D-Var	ECMWF
ERA-5	0.5°×0. 5°	Hourly throughout	4D-Var ensemble	ECMWF

Notes *3D-Var* or *4D-Var*: 3-dimensional or 4-dimensional variational data assimilation

Previous studies found that topography, altitude and sparse station are the main factors affecting the quality of reanalysis data (Rusticucci and Kousky, 2002; Wang et al. 2015; Zhao et al. 2018; He and Zhao 2018). Due to the sparse stations in Tibetan region, we removed the stations over Tibetan regions and only evaluated the quality of the reanalysis data by using observations from the left stations.

In this study, all referred time series are standardized by removing the annual cycle (Koscielny-Bunde et al. 1998) by $T_i' = T_i - \langle T_i \rangle, i = 1, \ldots, N$, where T_i is any given daily air temperature variables, $\langle T_i \rangle$ is its long-time average for each calendar day, T_i' is the corresponding anomaly, and N is the data length.

2.2 Methods

2.2.1 Interpolation for reanalysis

In order to compare TA measures calculated from the four reanalysis products with different spatial resolutions to those from station observations, we interpolated the two-dimensional gridded data into the corresponding observation stations (hereafter interpolated reanalysis). Two widely used methods in previous studies were adopted: Gaussian weight function (Maddox et al. 1981; Xie et al. 2019) and interpolation by directly choosing the closest grid point to represent the targeted station (Diaconescu et al. 2018; Pendergrass and Knutti 2018; Yang et al. 2020).

For Gaussian weight function, the interpolated reanalysis is given by

$$Re_s = \frac{\sum_i \sum_j w(i,j)s(i,j)}{\sum_i \sum_j w(i,j)}, \quad (1)$$

where $s(i,j)$ is the targeted variable over grid (i,j) and the weight function $w(i,j)$ is:

$$w(i,j) = \exp(\frac{-d^2(i,j,k)}{4C^2}), \quad (2)$$

where C is the weighting constant, the Euler distance d is from the grid point (i, j) to the location of the specific station. For the reanalyzed data with different resolutions, detailed numerical calculations found that the smaller value C (for example C=0.5 or C =1.0) within a certain range can produce reliable interpolations.

2.2.2 Measuring temporal asymmetry of time series

Different methods have been proposed and applied to measure the TA strength of air temperature variations (Xie et al. 2016, 2019; Zhang et al. 2019; Li et al. 2021). Zhang et al. (2019) compared several TA measure methods and found all of them perform nearly equally well for most of cases. For simplicity, the asymmetry index (A) is adopted to quantify the TA strength in the air temperature anomaly T_i', which is defined as the ratio of positive air temperature variability steps to the total (positive plus negative) steps (Ashkenazy et al. 2008):

$$A = \frac{p}{p+n} = \frac{\sum_{i=1}^{i=N-1} \theta(T'_{i+1} - T'_i)}{\sum_{i=1}^{i=N-1} \theta(T'_{i+1} - T'_i) + \sum_{i=1}^{i=N-1} \theta(-(T'_{i+1} - T'_i))}, \quad (3)$$

where $\theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases}$, N is the length of T_i'. When the value of TA is smaller than 0.5 ($A < 0.5$), it indicates that the air temperature warms rapidly and gradually becomes cold. The value of TA is larger than 0.5 ($A > 0.5$) indicates that the air temperature rapidly cools and gradually warms. When the value of TA is close to 0.5 ($A \approx 0.5$), the air temperature time series are symmetric.

2.2.3 Significance test

To carry out the significance tests, the iterative amplitude Fourier transform (IAAFT) is employed to generate surrogates for each time series by keeping the two-point correlation and probability density function as those from the original series. Taken T_{mean} anomaly time series as an example, we generate 500 surrogate series for it, and calculate A by means of Eq. (3) for each surrogate to obtain 500 values for A. Sorting these 500 values for A can define the lower 1% levels and the upper 99% among these 500 values of A as the significance thresholds. And then two thresholds of A_{c1} and A_{c2} can be obtained. When the calculated value of A for a given series is lower than A_{c1} or higher than A_{c2}, then this given series is taken to be statistically significantly temporal asymmetric. Such significance test through IAAFT surrogate can ensure that our results are not artificially influenced by stochastic effects, the autocorrelation and probability distribution of a time series (Huang et al. 2020).

3. Results

3.1 Overestimated temporal asymmetry in T_{max} and T_{min}

First of all, the estimated temporal asymmetry (TA) for four SAT variables' variability from the observations is compared with those from four kinds of re-analyses, and detailed results are presented in **Fig. 1**. The marked results are that nearly all values of estimated TA in T_{max} and T_{min} from all four re-analyses are overestimated (see **Fig. 1b** and **1c**), especially in T_{max}, only a few from NCEP-2 reanalysis are comparable with or lower than those from observations. This finding is consistent with previous studies in NCEP-2 reanalysis compared with limited station observations (Xie et al. 2019). Different from the results given by only one specific reanalysis product (NCEP-2), the results given here indicate that the overestimated TA in T_{max} and T_{min} from re-analyses may be taken as a shared intrinsic feature to all analyzed reanalysis products.

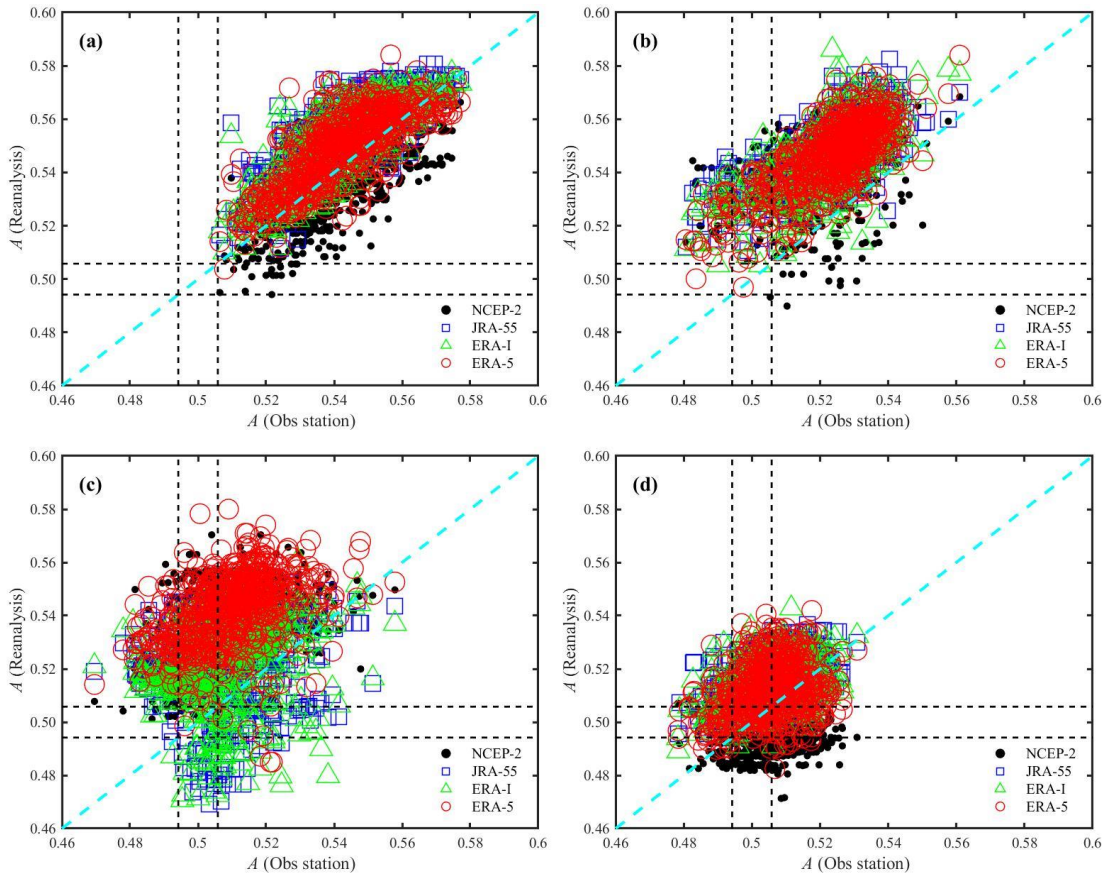

Fig. 1 Temporal asymmetry scatter plots between observations and interpolated re-analyses for **a** T_{mean}, **b** T_{max}, **c** T_{min}, **d** T_{DTR} from 1979 to 2018. The two horizontal and vertical dash black lines denote the two critical thresholds for the confidence interval of A at the significance level of 0.02, respectively. The dash cyan line denotes the 1:1 diagonal line.

Secondly, overestimated TA in both T_{max} and T_{min} from all four kinds of re-analyses has its well-defined spatial patterns (see **Fig. 2** and **Fig. 3**). Especially for T_{max}, four kinds of re-analyses

share very similar spatial distribution of TA difference between observation and interpolated reanalysis, which is defined as $A_{diff} = A_{re} - A_{obs}$ with A_{re} from reanalysis and A_{obs} from observations. There are region-dependent spatial patterns of A_{diff}. Weak overestimated TA occupies most of regions of China. However, there are still some special patterns. The consistent estimation of TA between observation and interpolated reanalysis mainly locates over northeast part of Heilongjiang Province and west part of both Yunnan and Sichuan Provinces. The shared patterns of the strongest overestimated TA among four kinds of re-analyses lie in Guizhou and Chongqing provinces. These regions are also where the TA is not statistically significant in T_{max} for observations (see **Fig. 4f**), but strong TA for four kinds of re-analyses (see **Fig. 4g-4j**). The mechanism behind this phenomenon deserves further study in depth.

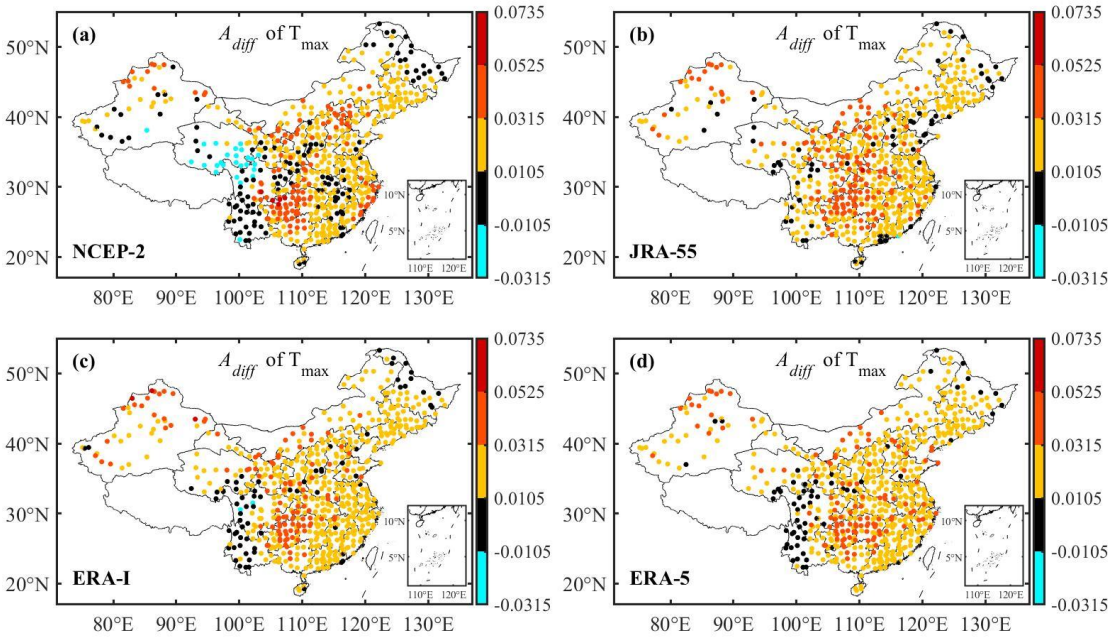

Fig. 2 Spatial distribution of A difference in T_{max} from 1979 to 2018 between observation and interpolated reanalysis for **a** NCEP-2, **b** JRA-55, **c** ERA-I and **d** ERA-5, the black asterisk indicates the confidence interval for the same TA estimation between observation and interpolated reanalysis at the significance level of 0.02.

For T_{min}, no common national-scale spatial pattern of TA difference is shared among four kinds of re-analyses between observation and interpolated reanalysis over China (see **Fig. 3**). Only NCEP-2 and ERA-5 share the similar national-scale spatial distribution of overestimated TA (see **Fig. 3a** and **3d**). However, JRA-55 and ERA-I share the similar spatial distribution of TA difference with an east-west dipole pattern. Over the half part of China east to 110° E, nearly all values of TA are overestimated, whereas over the another half part of China west to 100° E, nearly all values of TA are weakly underestimated. The contrasting spatial patterns of TA in T_{min}

are mainly due to the weak TA strength in observations (see **Fig. 5a**) but strong TA strength in Central parts of China in four kinds of re-analyses (see **Fig. 5b-5e**).

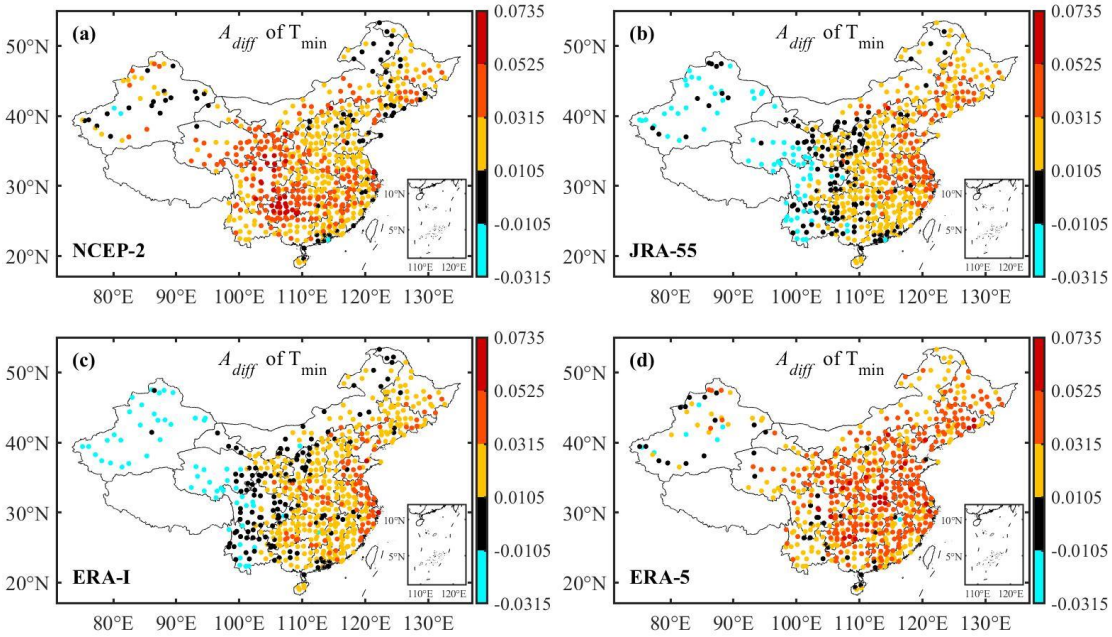

Fig. 3 Same as Fig.2 but for T_{min}.

Besides the almost identical patterns of TA difference for T_{min}, JRA-55 and ERA-I also share almost the same TA patterns for all temperature variables (T_{mean}, T_{max}, T_{min} and T_{DTR}), details can be found in **Fig. 4c, 4d, 4h, 4j** and **Fig. 5c, 5d, 5h, 5j**. The similarity between ERA-I and JRA-55 re-analyses on synoptic-scale phenomena has been reported in the literature. When Pinheiro et al. (2020) studied subtropical cut-off lows in the southern hemisphere, they found that the differences of track density between ERA-I and JRA-55 are relatively small. They explained this similarity between ERA-I and JRA-55 as both of them used the same data assimilation systems (Pinheiro et al. 2020).

3.2 Comparable estimation of temporal asymmetry in T_{mean}

Different from the markedly overestimated TA found in T_{max} and T_{min}, more consistent TA estimations are revealed in T_{mean} between observations and re-analyses (see **Fig. 6**). There are only a few stations (less than 1%) with strong overestimated TA and nearly 50% stations have the same TA estimation for both observations and re-analyses from NCEP-2 and ERA-5 (see **Fig. 6a** and **6d**). Moreover, nearly all values of estimated TA in T_{mean} from both observations and re-analyses are statistically significant at the significance level of 0.02 (only less than 2% stations in NCEP-2 have insignificant TA estimation). For clear view, the same critical thresholds are marked with vertical and horizontal black dotted lines in the scatter plot, as shown in **Fig. 1.** The lower value of A_{c1} is 0.4942, and the upper value A_{c2} is 0.5058. This indicates that nearly all the daily mean

SAT variability takes the temporal asymmetric behaviors over China.

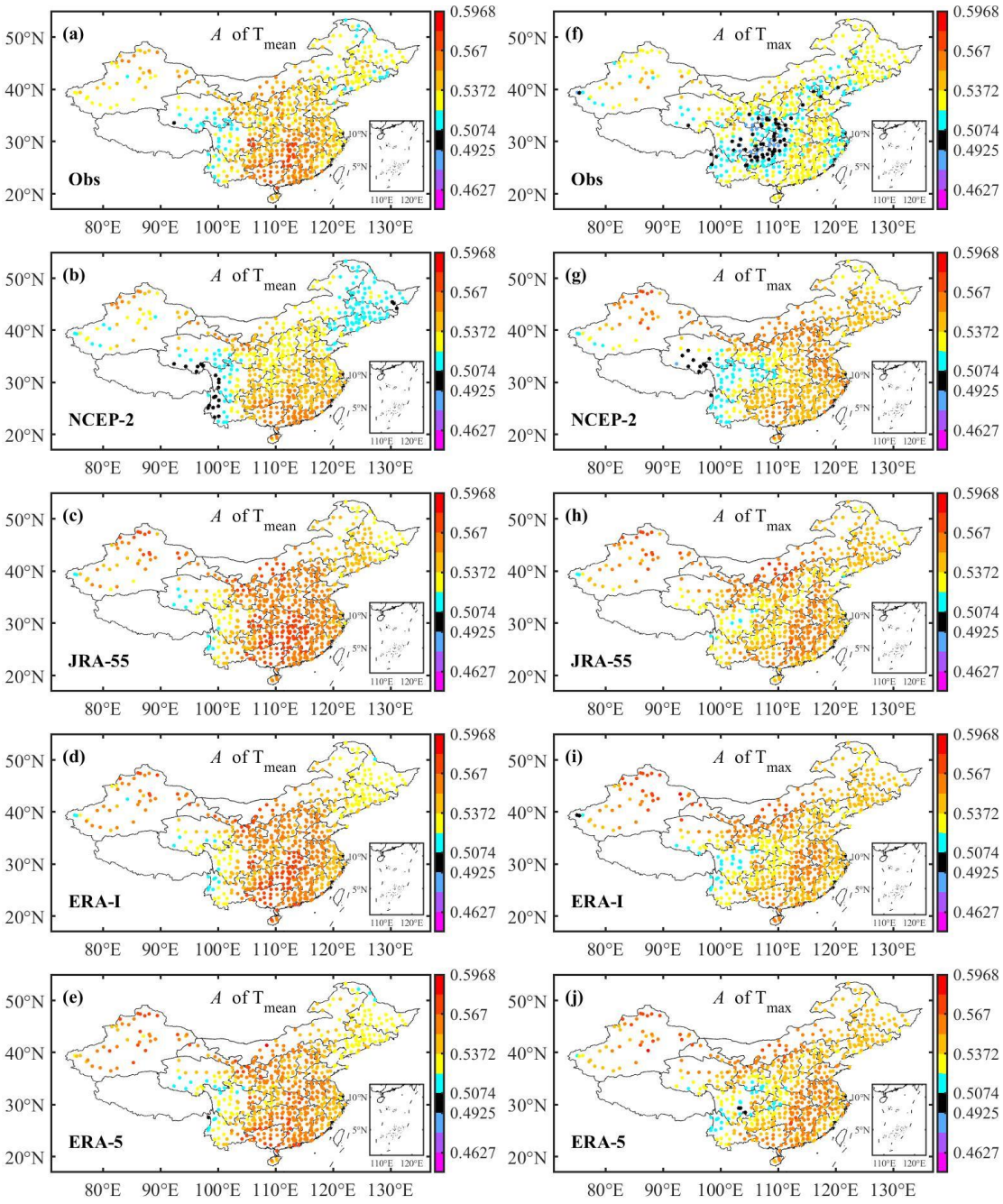

Fig. 4 Spatial distribution of A difference in T_{mean} (left column) and T_{max} (right column) from 1979 to 2018 between observation (**a**, **f**) and interpolated reanalysis: NCEP-2(**b**, **g**), JRA-55 (**c**, **h**), ERA-I (**d**, **i**) and ERA-5 (**e**, **j**). The black asterisk indicates the confidence interval with two critical thresholds A_{C1}=0.4925 and A_{c2}=0.5074 for TA estimation at the significance level of 0.02.

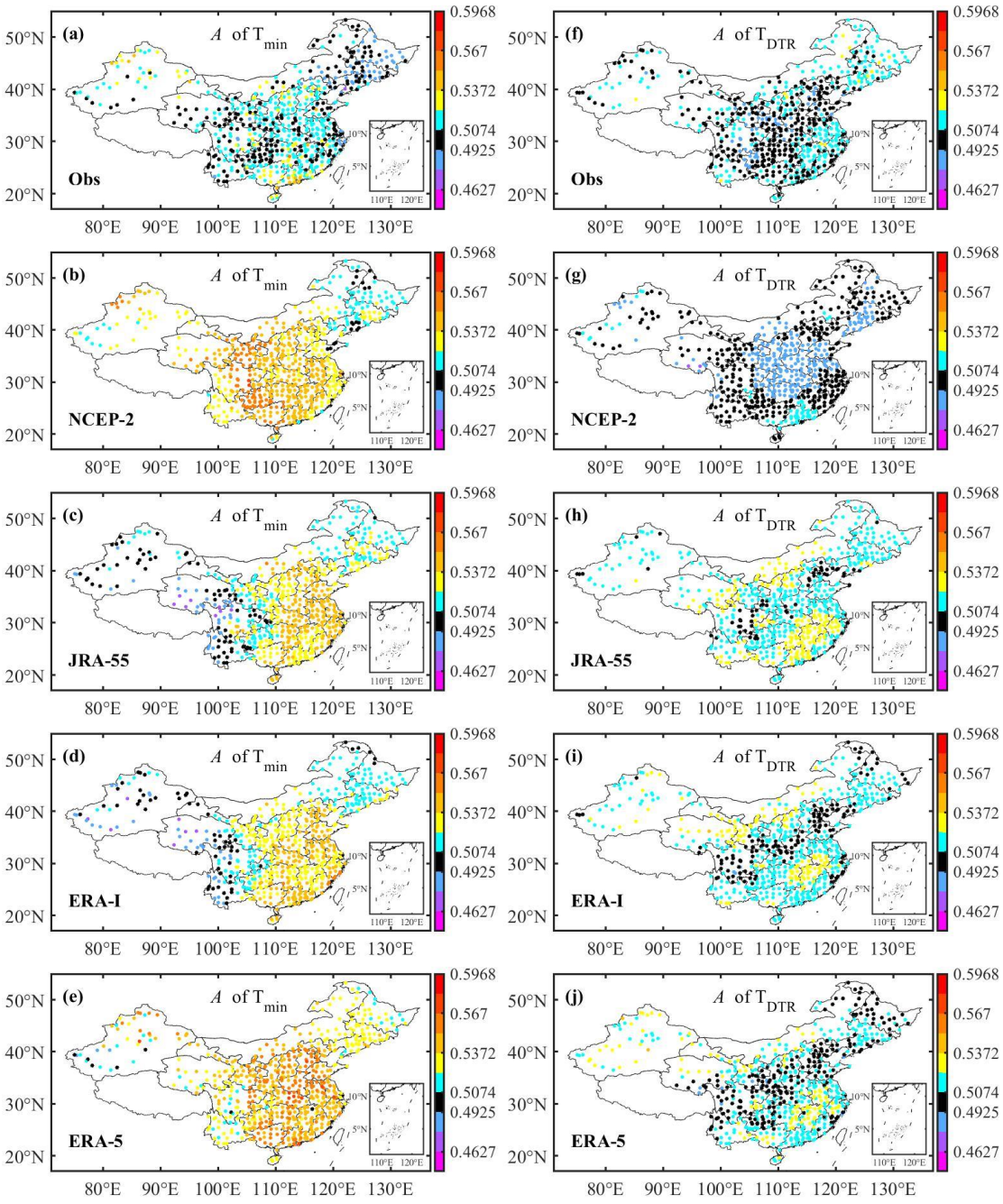

Fig. 5 Same as Fig. 4 but for T_{min} (left column) and T_{DTR} (right column).

The regions of the strongest TA strength lie over central and southeast regions of China, especially to the south of Qinling Mountains and to east of 105° E (see **Fig. 4a**). Over these regions, the variations of T_{mean} gradually warm and rapidly cool. Such a temporal asymmetry phenomenon occurring at mid-latitudes has been attributed to the different contributions from the warm and cold fronts (Ashkenazy et al. 2008). Lower A is found in the Basin of Tarim, Qaidam and Sichuan, Yun-Gui Plateau and northeast of China (see **Fig. 4a**). Surprisingly, all four re-analyses can reproduce the high and low values of TA over these regions (see **Fig. 6**), and this is why there are more consistent TA estimations in T_{mean} (see **Fig. 1a**).

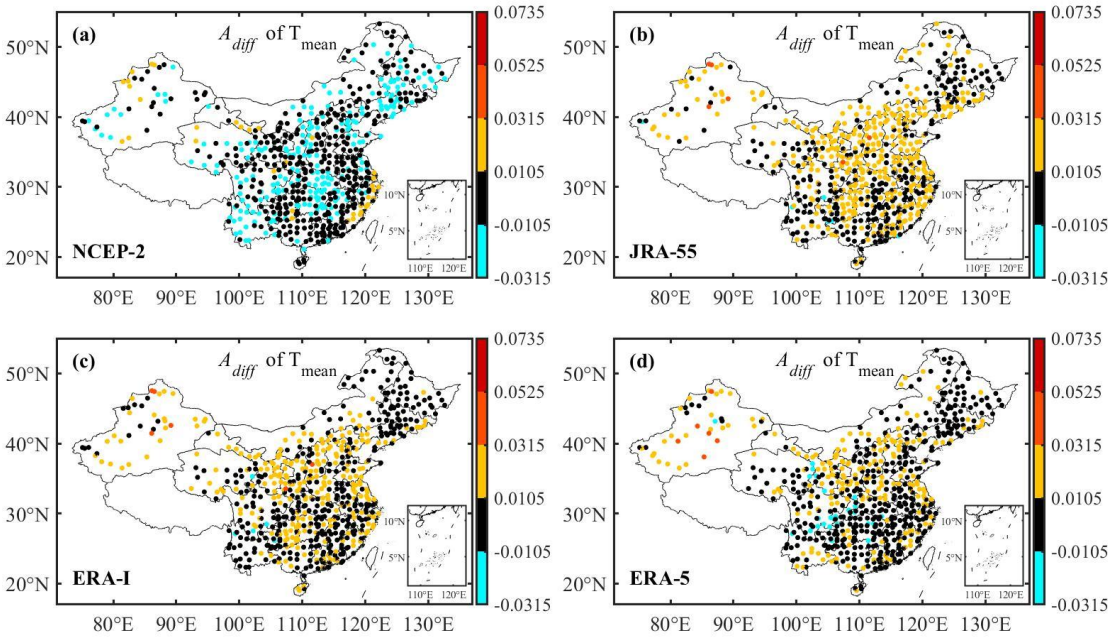

Fig. 6 Same as Fig. 2 but for T_{mean}.

3.3 TA spatial pattern similarity quantification

In order to quantify the TA spatial pattern similarity between observations and interpolated re-analyses, we adopt the Taylor diagram (Taylor. 2001) to compare it quantitatively. Three statistics most often used to quantify pattern similarity in Taylor diagram are the correlation coefficient between observations and interpolated re-analyses, defined as

$$RCC = \frac{\sum_{i=1}^{n}(A_{Ra}-\overline{A_{Ra}})(A_o-\overline{A_o})}{n\sigma_{A_{Ra}}\sigma_{A_o}} , \quad (4)$$

the root-mean-square (RMS) difference between observations and interpolated re-analyses standardized by the results from observations, defined as

$$SRMS = \sqrt{\frac{\sum_{i=1}^{n}\left[(A_{Ra}-\overline{A_{Ra}})-(A_o-\overline{A_o})\right]^2}{n\sigma_{A_o}^2}} , \quad (5)$$

and the standard deviation of A_{Ra} standardized by A_o

$$SSTD = \frac{\sigma_{A_{Ra}}}{\sigma_{A_o}} , \quad (6)$$

with n the number of stations, σ_{A_o} and $\sigma_{A_{Ra}}$ are standard deviation for A from observations and a specific interpolated reanalysis. For perfectly consistent TA estimations between observations and interpolated reanalysis over all considered regions over China, $RCC = 1$, $SRMS = 0$ and $SSTD = 1$.

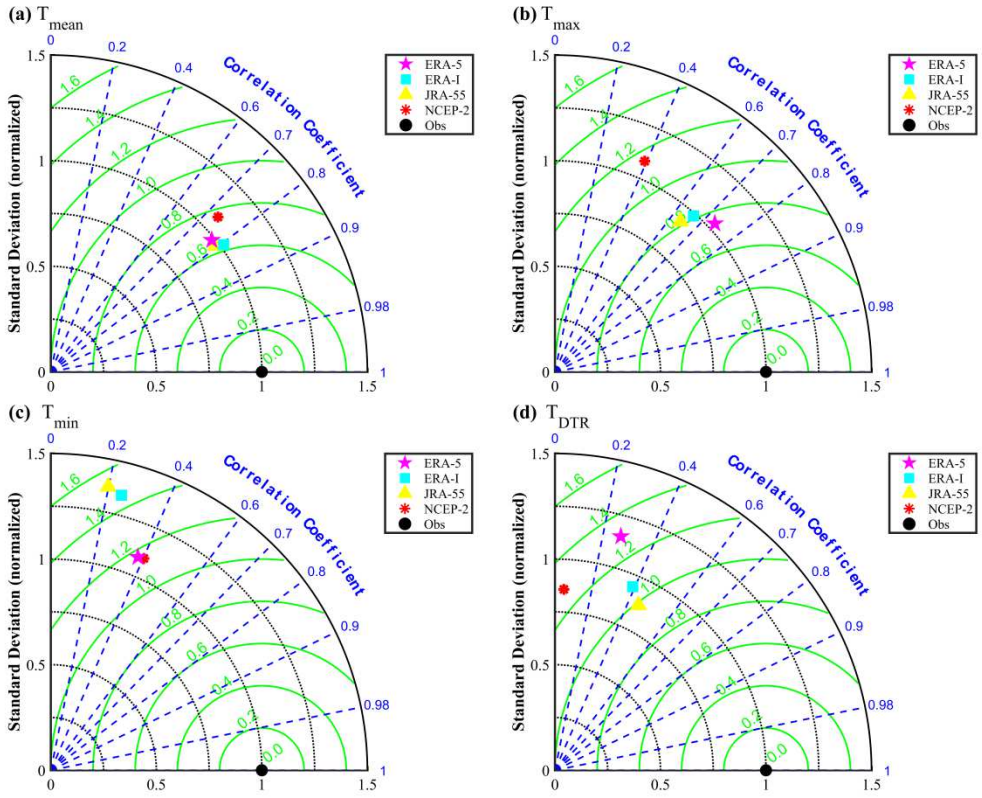

Fig. 7 Taylor diagram for **a** T_{mean}, **b** T_{max}, **c** T_{min}, **d** T_{DTR} of A from observation and interpolated reanalysis. The black dot stands for the results calculated from observations, which serves as the reference. The radial distance from the origin is proportional to the standard deviation of a pattern normalized by reference pattern, the centered root mean square (RMS) difference between the reference and re-analyses is proportional to their distance apart and the correlation between the reference and re-analyses is given by the azimuthal position of a given reanalysis.

Taylor diagrams constructed from four kinds of interpolated re-analysis and observations for TA of four SAT variability could more accurately evaluate the quality of different reanalysis products for different SAT variables (**Fig. 7**). First of all, it should be pointed out that synoptic-scale performance of all analyzed re-analyses is not comparable to their climate-scale performance. For T_{mean}, the four reanalysis products all perform the best among these four SAT variables, and they are all located in a limited region in Taylor diagram (see **Fig. 7a**). Especially ERA-I, ERA-5 and JRA-55 are nearly of the equally well performance with the highest RCC value (around 0.8) and the lowest SRMS value (around 0.62), only NCEP-2 a little bit worse with the RCC value around 0.73 and SRMS value around 0.78. Four reanalysis products perform second well for T_{max} among these four SAT variables (see **Fig. 7b**), among them, ERA-5 is the best with the RCC value around 0.73 and SRMS value around 0.74 and NCEP-2 the worst with the RCC value around 0.4 and SRMS value around 1.14. There are great discrepancies in TA estimation between the observation and the interpolated re-analyses for both T_{min} and T_{DTR} (**Fig. 7c** and **7d**), for all re-analyses, the RCC values are low (less than 0.5) and SRMS values are all

high (larger than 0.99).

Moreover, Ye and Hsieh found (2008) that increasing nonlinearity in ENSO and Lorenz systems can enhance their predictability by improving the contributions from the low-frequency variations. As a kind of nonlinearity, temporal asymmetry in daily SAT variability is closely related to extreme events and some small-scale phenomena (Raghavendra et al. 2018; Li et al. 2021). More comparable TA estimation between observations and re-analyses leads daily mean surface temperature reanalysis products to be the most suitable choice to synoptic-scale extreme event study.

4. Discussion and conclusion

In order to directly compare the TA from observation with those from reanalysis, the grid reanalysis data have been interpolated into the targeted station to reach the interpolated reanalysis data. In previous studies, a number of interpolation methods were proposed and applied to test the impacts of interpolations on the derived results (Maddox et al. 1981; Xie et al. 2019). Xie et al. (2019) found that if the suitable choice is made to the interpolation distance parameter C, interpolations don't change the TA calculations too much. We compared the effects from different methods and different choices of interpolation distance parameter C on the calculations of TA from the original and four interpolated reanalysis data, detailed results can be found in **Fig. 8** and were summarized in **Table 2**. It is confirmed that if the suitable distance parameter is chosen, the estimation of TA is insensitive to the interpolations. Also the different interpolation methods do not change the TA calculations too much. Especially, the estimation of TA from ERA-I and JRA-55 is more robust to the interpolation methods and choices of distance parameter (**Fig. 8** and **Table 2**).

Table 2 Correlation coefficient between the probability density function of A in T_{mean} from the original and those from interpolated reanalysis (NCEP-2, JRA-55, ERA-I and ERA-5)

Original reanalysis (T_{mean})	Interpolated reanalysis			
	Method-2	Gauss(C=0.5)	Gauss(C=1.0)	Gauss(C=1.5)
NCEP-2	0.960	0.803	0.710	0.638
JRA-55	0.941	0.898	0.886	0.864
ERA-I	0.937	0.885	0.839	0.878
ERA-5	0.918	0.880	0.850	0.757

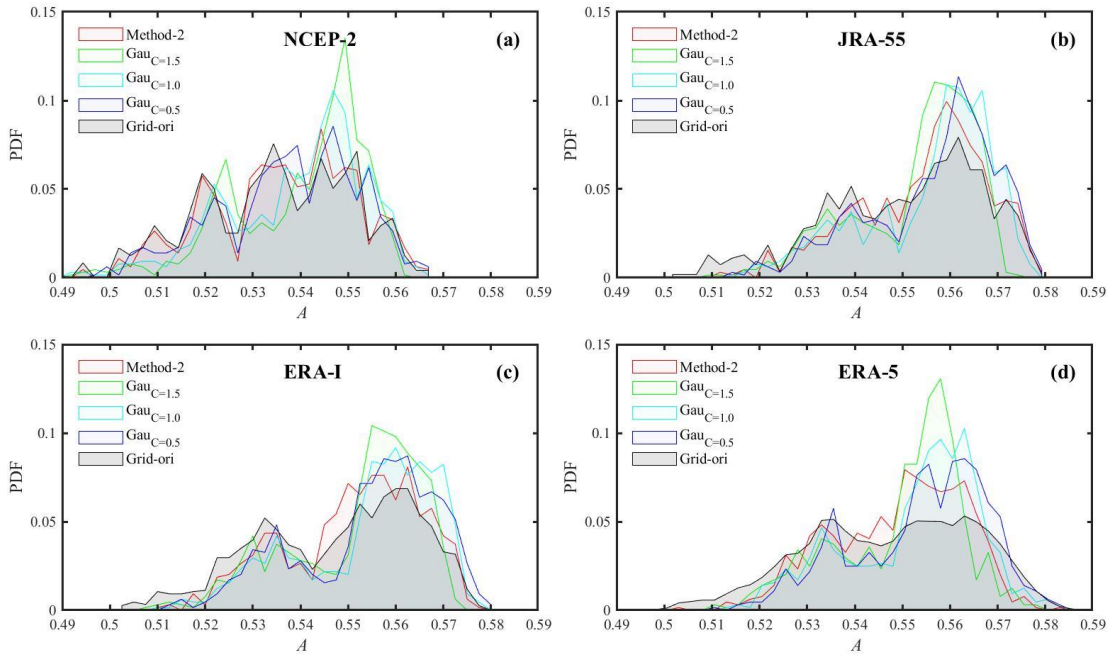

Fig. 8 Probability density function (PDF) for A of T_{mean} from **a** NCEP-2, **b** JRA-55, **c** ERA-I and **d** ERA-5 by different interpolation ways. Gau means the Gaussian weight function interpolation with different values of C (C=1.5, green; C=1.0, cyan; C=0.5, blue), Method-2 represents interpolation by the closest points to the stations (red) and Grid-ori from original reanalysis (grey) shadow.

As an important nonlinear indicator, temporal asymmetry of time series can be taken as an intrinsic feature of nonlinear time series. TA calculation from observation in daily mean surface air temperature variability over China, (**Fig. 1a**), having nearly totally statistically significant temporal asymmetry lead it to be an intrinsic feature of daily surface air temperature variability. It can be taken as a metric to access the quality of different daily SAT reanalysis products. Taking TA as an evaluation measure, four daily SAT reanalysis products (NCEP-2, ERA-I, ERA-5, JRA-55) are accessed. Compared with the observations, the four re-analyses can consistently reproduce the TA in T_{mean} over South China. NCEP-2 underestimates the TA in T_{mean} over the northeast and central regions of China. JRA-55, ERA-I and ERA-5 overestimate the TA in T_{mean} over the northwest and central regions. However, all four re-analyses universally overestimate the TA in T_{max} and T_{min}. These results are different from the findings by previous studies based on the linear view (Zhao et al. 2018; He and Zhao 2018). Due to the weak TA in the T_{DTR} (see **Fig. 1d** and **Fig. 5f-5J**), there are larger relative uncertainties in calculation and comparison of TA from the observations and reanalysis, so no conclusive results can be reached on T_{DTR}.

It was reported that the modeled daily maximum temperature and daily minimum temperature are unsuitable for the study of extreme events such as heat waves due to the underestimated daily maximum or overestimated daily minimum temperature compared with observations (Raghavendra et al. 2018). Based on the results of TA estimation from all four

reanalysis products, we confirm that the nonlinear strength is highly overestimated in T_{max} and T_{min} from the re-analyses. Since there is a close relation between nonlinearity and extreme events (Ye and Hsieh 2008; Li et al. 2021), results from daily maximum or minimum temperature re-analyses may distort conclusions on extreme events such as heat waves and cold waves. On the contrary, the most comparable consistency of TA estimation in daily mean surface air temperature variability from both observations and re-analyses make it to be a reasonable choice.

Author contributions

Study design: Zuntao Fu, Wenmi Chai; data collection: Wenmi Chai; statistical analysis: Wenmi Chai; result interpretation: Zuntao Fu, Wenmi Chai; manuscript preparation and review: Wenmi Chai, Yu Huang, Lichao Yang, Heng Quan and Zuntao Fu.

Funding

This research was supported by the National Natural Science Foundation of China through Grants (No. 41475048 and 41975059).

Code availability

Not applicable.

Data availability

All observed records for these four temperature variables were downloaded from the China meteorological data sharing service system (http://cdc.cma.gov.cn). Reanalysis NCEP-2 products were acquired from the National Oceanic and Atmospheric Administration (NOAA) and its website at http://www.esrl.noaa.gov/psd. JRA-55 is produced by the Japan Meteorological Agency (JMA) (https://climatedataguide.ucar.edu/climate-data/jra-55). From the ECMWF's website: https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets, both ERA-I and ERA-5 are downloaded.

Declarations

Ethics approval

We confirm that this article is an original research and has not been published or presented previously in any journal or conference in any language (in whole or in part).

Consent to participate and consent for publication

We have consent to participate and publish.

Conflict of interest/competing interests

The authors declare no competing interest.

References

Alfred M, Powell Jr, Xu JJ (2011) A new assessment of the mid-1970s abrupt atmospheric temperature change in the NCEP/NCAR reanalysis and associated solar forcing implications. Theor Appl Climatol 104: 443-458.

Alghamdi AS (2020) Evaluation of four reanalysis datasets against radiosonde over Southwest Asia. Atmosphere 11: 402.

Ashkenazy Y, Feliks Y, Gildor H, Tziperman E (2008) Asymmetry of daily temperature records. J Atmos Sci 65: 3327.

Ashkenazy Y, Fredj E et al (2016) Current temporal asymmetry and the role of tides: Nan-Wan bay vs. the gulf of Elat. Ocean Sci 12: 733

Ashkenazy Y, Tziperman E (2004) Are the 41 kyr glacial oscillations a linear response to Milankovitch forcing? Quat Sci Rev 23: 1879-1890.

Ayataka E et al (2011) The Japanese 55-year reanalysis “JRA-55”: An interim report. SOLA, 7: 149-152.

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

Bengtsson L, Shukla J (1988) Integration of space and in situ observations to study global climate change. Bull Am Meteor Soc 69:1130-1143.

Bisgaard S, Kulahci M (2011) Time series analysis and forecasting by example. Wiley, New York.

Chen B, Liu Z (2016) Global water vapor variability and trend from the latest 36 year (1979 to 2014) data of ECMWF and NCEP reanalyses, radiosonde, GPS, and microwave satellite. J Geophys Res 121: 11442-11462.

Chen GX, Iwasaki T (2014) Evaluation of the warm-season diurnal variability over East Asia in recent reanalyses JRA-55, ERA-Interim, NCEP CFSR, and NASA MERRA. J Climate 27: 5517-5537.

Cornes RC, Jones P D (2013) How well does the ERA-Interim reanalysis replicate trends in extremes of surface temperature across Europe? J Geophys Res 118: 10262-10276.

Dee DP, and Coauthors (2011) The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quart J Roy Meteor Soc 137: 553-597.

Diaconescu EP, Mailhot A, Brown R, Chaumont D (2018) Evaluation of CORDEX-Arctic daily precipitation and temperature based climate indices over Canadian Arctic land areas. Climate Dynamics 50(5-6): 2061-2085.

Flocas HA, Tolika K, Anagnostopoulou C, Patrikas I, Maheras P, Vafiadis M (2005) Evaluation of maximum and minimum temperature of NCEP-NCAR reanalysis data over Greece. Theor Appl Climatol 80: 49-65.

Gyure B, Bartos I, Janosi IM (2007) Nonlinear statistics of daily temperature fluctuations reproduced in a laboratory experiment. Phys Rev E 76: 037301.

He WP, Zhao SS (2018) Assessment of the quality of NCEP-2 and CFSR reanalysis daily

temperature in China based on long-range correlation. Climate Dynamics 50: 493.
Heinrich H (2004) Origin and consequences of cyclic ice rafting in the Northeast Atlantic Ocean during the past 130,000 years. Quat Res 29: 142-152.
Hersbach H, Bell B, Berrisford P, et al (2020) The ERA5 global reanalysis. Quart J Roy Meteor Soc 146: 1999-2049.
Huang Y, et al (2020) Systematic identification of causal relations in high-dimensional chaotic systems: application to stratosphere-troposphere coupling. Climate Dynamics 55: 2469-2481.
Kanamitsu M et al (2002) NCEP-DOE AMIP-II reanalysis (R-2). Bull Am Meteorol Soc 83: 1631.
King T (1996) Quantifying nonlinearity and geometry in time series of climate. Quat Sci Rev 15: 247-266.
Kobayashi S, Ota Y, Harada Y, Ebita A, Moriya M, Onoda H, Onogi K, Kamahori H, Kobayashi C, Endo H, Miyaoka K, Takahashi K (2015) The JRA-55 reanalysis: general specifications and basic characteristics. J Metorol Soc Jpn 93: 5-48.
Koscielny-Bunde E, Bunde A, Havlin S, Roman HE (1998) Indication of a universal persistence law governing atmospheric variability. Phys Rev Lett 81: 729.
Lacasa L, Nunez A, Roldan E, Parrondo JMR, Luque B (2012) Time series irreversibility: a visibility graph approach. Eur Phys J B 85: 217.
Li RC, Huang Y, Xie FH, Fu ZT (2021) Discrepancies in surface temperature between Reanalysis and station observations over China and their implications. Atmos Oceanic Sci Lett 14:100008
Lisiecki LE, Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic delta O-18 records. Paleoceanography 20: PA1003.
Livina V, Ashkenazy Y, Kizner Z, Strygin V, Bunde A, Havlin S (2003) A stochastic model of river discharge fluctuations. Phys A 330: 283-290.
Ma L, Zhang T, Li Q, Frauenfeld OW, Qin D (2008) Evaluation of ERA-40, NCEP-1,and NCEP-2 reanalysis air temperatures with ground-based measurements in China. J Geophys Res 113: D15115.
Mao J, Shi X, Ma L, Kaiser D P, Li Q, Thornton P E (2010) Assessment of reanalysis daily extreme temperatures with China's homogenized historical dataset during 1979-2001 using probability density functions. J Clim 23: 6605-6623.
Marques CAF, Rocha A, Corte-Real J (2010) Comparative energetics of ERA-40, JRA-25 and NCEPR2 reanalysis, in the wave number domain. Dyn Atmos Oceans 50: 375-399.
Mooney PA, Mulligan FJ, Fealy R (2011) Comparison of ERA-40, ERA-interim and NCEP/NCAR reanalysis data with observed surface air temperature over Ireland. Int J Climatol 31: 545-557.
Pendergrass AG, Knutti R (2018) The uneven nature of daily precipitation and its change. Geophysical Research Letters 45(21): 11980-11988.
Pinheiro HR, et al (2020) An intercomparison of subtropical cut-off lows in the Southern Hemisphere using recent reanalyses: ERA-Interim, NCEP-CFRS, MERRA-2, JRA-55, and

JRA-25. Climate Dynamics 54:777-792.
Pitman AJ, Perkins SE (2009) Global and regional comparison of daily 2-m and 1000-hPa maximum and minimum temperatures in three global reanalysis. J Clim 22: 4667-4681.
Radu CR, Rozum I, Schepers D, Simmons A, Soci C, Dee D, Thépaut JN (2018) ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on < 26-02-2021 >), 10.24381/cds.adbb2d47.
Raghavendra A, Dai AG, Milrad SM, Cloutier-Bisbee SR (2018) Floridian heatwaves and extreme precipitation-future climate projections. Clim Dyn 52: 495-508.
Roldan E, Parrondo JMR (2010) Estimating dissipation from single stationary trajectories. Phys Rev Lett 105: 150607.
Rusticucci MM, Kousky E (2002) A comparative study of maximum and minimum temperatures over Argentina: NCEP-NCAR reanalysis versus station data. J Clim 15: 2089-2101.
Schreiber T, Schmitz A (2000) Surrogate time series. Phys D 142: 346-382.
Taguchi M (2017) Comparison of large-scale dynamical variability in the extratropical stratosphere among the JRA-55 family data sets: impacts of assimilation of observational data in JRA-55 reanalysis data. Atmos Chem Phys 17: 11193-11207.
Taylor Karl E. (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res 106: 7183-7192.
Trenberth KE, Olson JG (1988) An evaluation and intercomparison of global analyses from national meteorological center and the European centre for meddium range weather forecast. Bull Am Meteor Soc 69: 1047-1057
Wang S, Zhang M, Sum M, Wang B, Huang X, Wang Q, Feng F (2015) Comparison of surface air temperature derived from NCEP/DOE R2, ERA-Interim, and observations in the arid northwestern China: a consideration of altitude errors. Theor Appl Climatol 119: 99-111.
Xie FH, Fu ZT, Piao L, Mao JY (2016) Time irreversibility of mean temperature anomaly variations over China. Theor Appl Climatol 123:161.
Xie, FH, Nian D, Fu ZT (2019) Differential temporal asymmetry among different temperature variables' daily fluctuations. Clim Dyn 53: 585-600.
Yang LC, Franzke CL, Fu ZT (2020) Power-law behaviour of hourly precipitation intensity and dry spell duration over the United States. Int J Climatol 40: 2429-2444.
Ye, ZQ, Hsieh WW (2008) Enhancing predictability by increasing nonlinearity in ENSO and Lorenz systems. Nonlin Processes Geophys 15: 793-801.
You Q, Fraedrich K, Min J, Kang S, Zhu X, Ren G, Meng X (2013) Can temperature extremes in China be calculated from reanalysis? Glob Planet Change 111: 268-279.
You Q, Kang S, Aguilar E, Pepin N, Flugel W, Yan Y, Xu Y, Zhang Y, Huang J (2011) Changes in daily climate extremes in China and their connection to the large scale atmospheric circulation during 1961-2003. Clim Dyn 36: 2399-2417.
Zhao SS, He WP, Jiang YD (2018) Evaluation of NCEP-2 and CFSR reanalysis seasonal

510 temperature data in China using detrended fluctuation analysis. Int J Climatol 38: 252-263.
511 Zhang BE, Xie FH, Fu ZH, Fu ZT (2019) Comparative study of multiple measures on temporal
512 irreversibility of daily air temperature anomaly variations over China. Phys A 523: 1387-1399.
513 Zhu J, Huang D, Yan P, Huang Y, Kuang X (2017) Can reanalysis datasets describe the persistent
514 temperature and precipitation extremes over China? Theor Appl Climatol 130: 655-671.

Figures

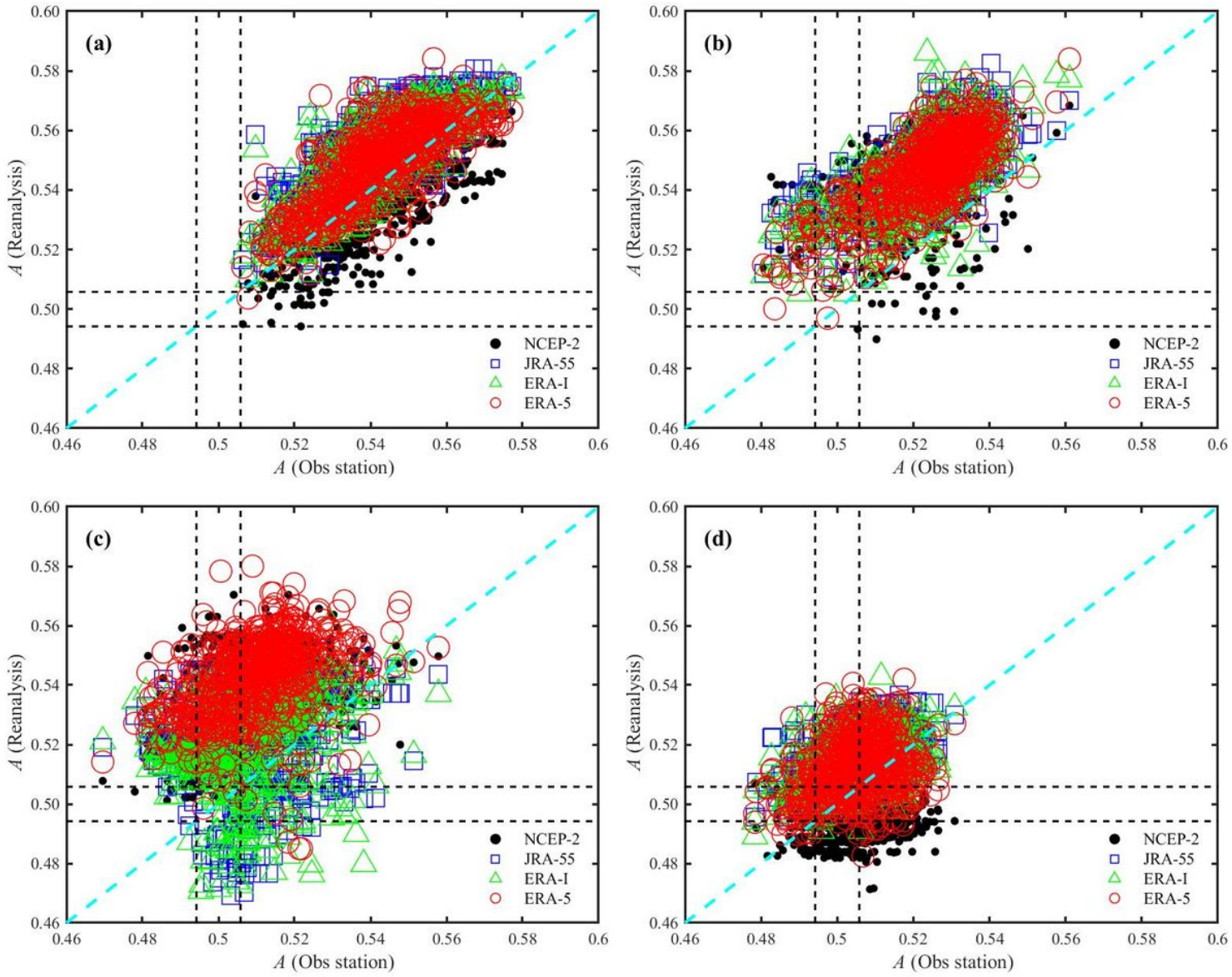

Figure 1

Temporal asymmetry scatter plots between observations and interpolated re-analyses for a Tmean, b Tmax, c Tmin, d TDTR from 1979 to 2018. The two horizontal and vertical dash black lines denote the two critical thresholds for the confidence interval of A at the significance level of 0.02, respectively. The dash cyan line denotes the 1:1 diagonal line.

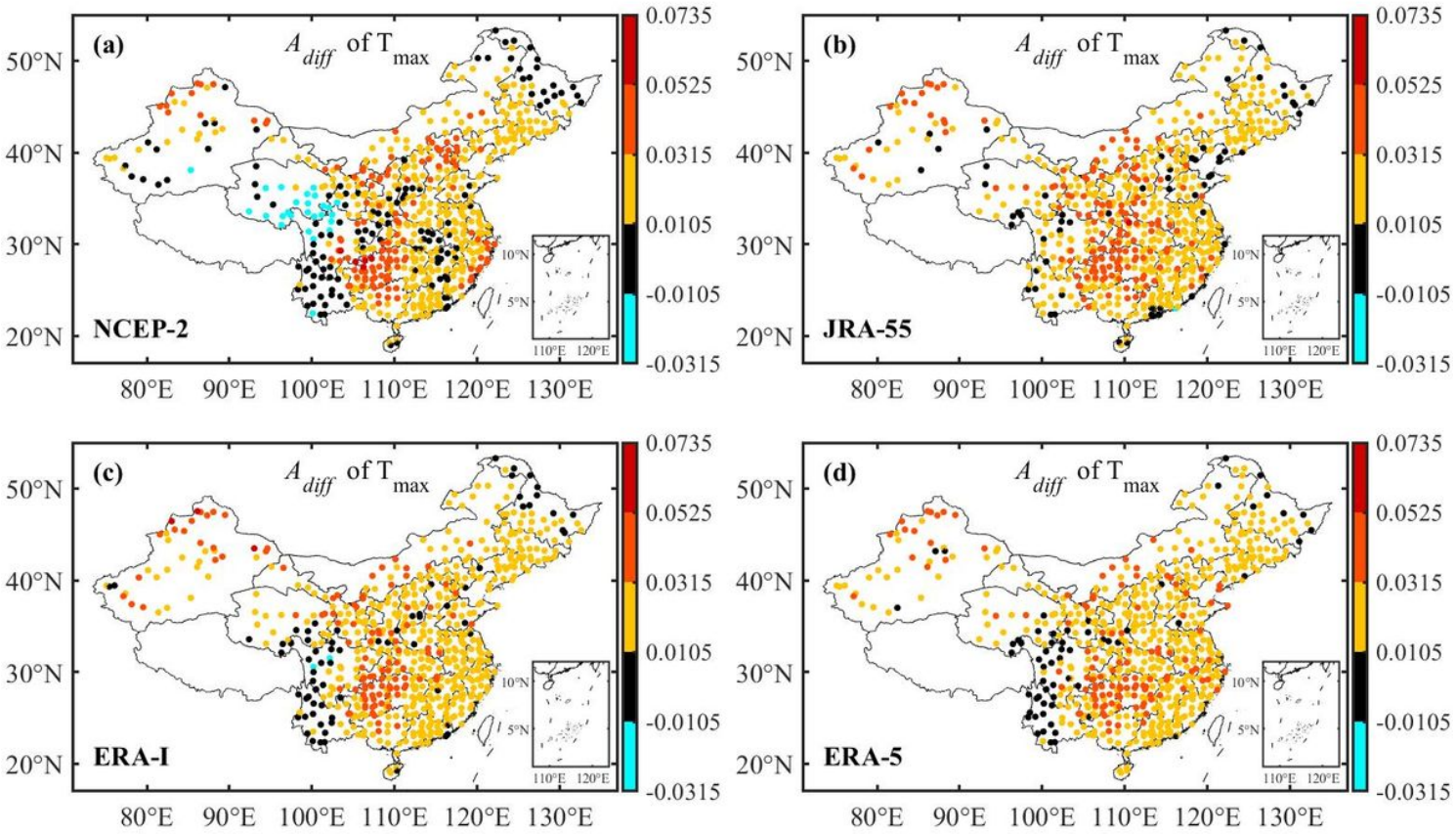

Figure 2

Spatial distribution of A difference in Tmax from 1979 to 2018 between observation and interpolated reanalysis for a NCEP-2, b JRA-55, c ERA-I and d ERA-5, the black asterisk indicates the confidence interval for the same TA estimation between observation and interpolated reanalysis at the significance level of 0.02. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

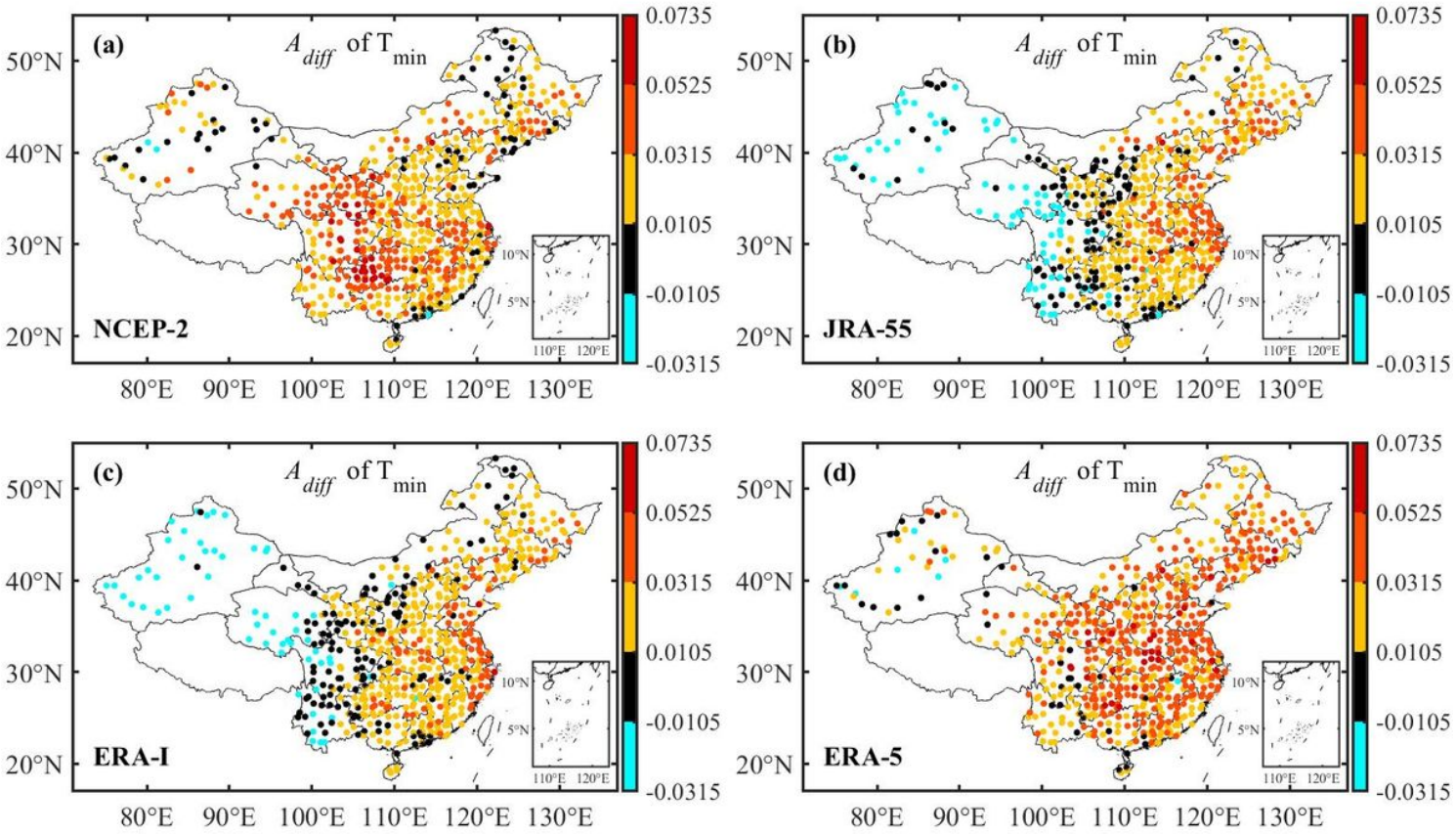

Figure 3

Same as Fig.2 but for Tmin. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

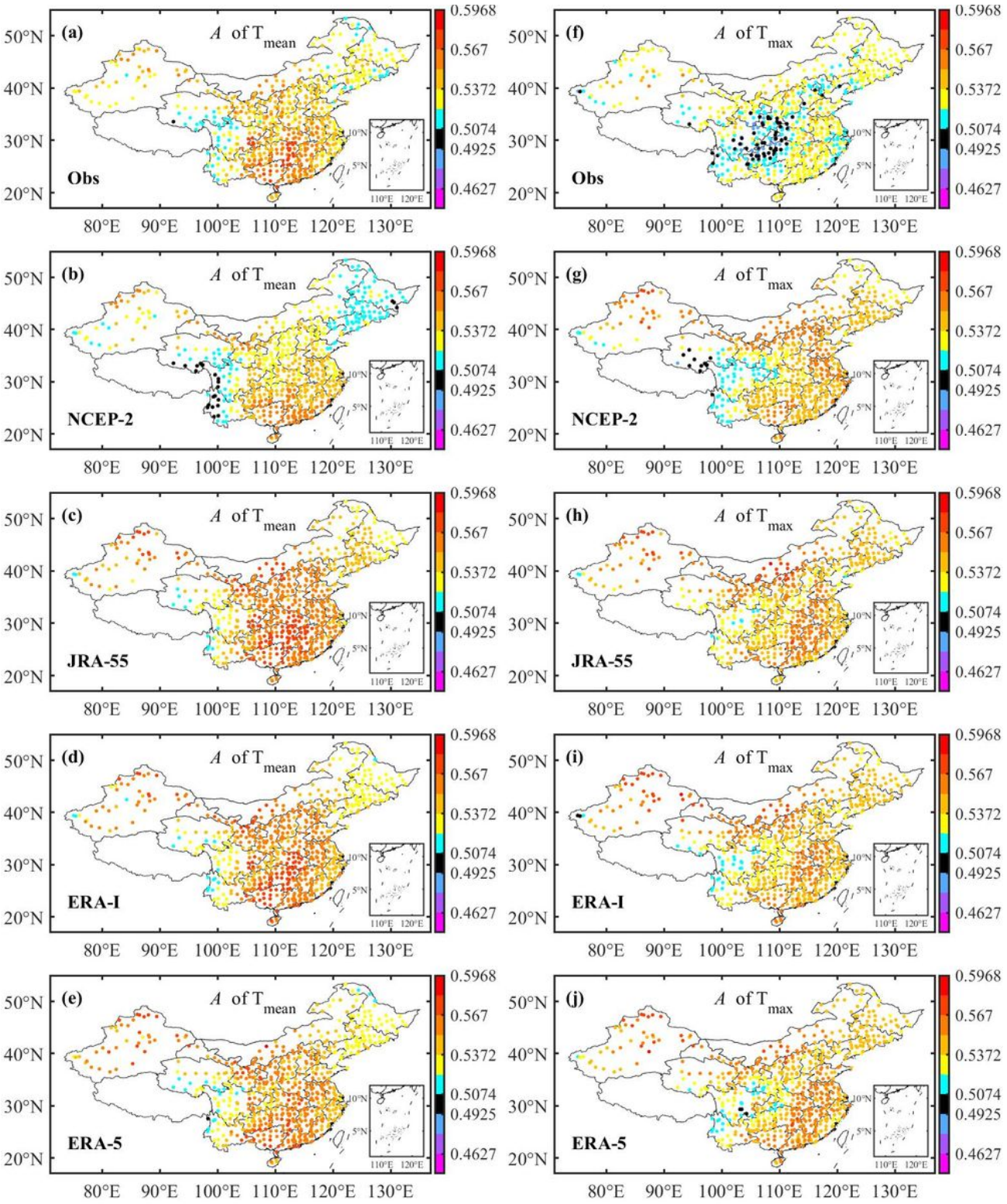

Figure 4

Spatial distribution of A difference in Tmean (left column) and Tmax (right column) from 1979 to 2018 between observation (a, f) and interpolated reanalysis: NCEP-2(b, g), JRA-55 (c, h), ERA-I (d, i) and ERA-5 (e, j). The black asterisk indicates the confidence interval with two critical thresholds A_C1=0.4925 and A_c2=0.5074 for TA estimation at the significance level of 0.02. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the

part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

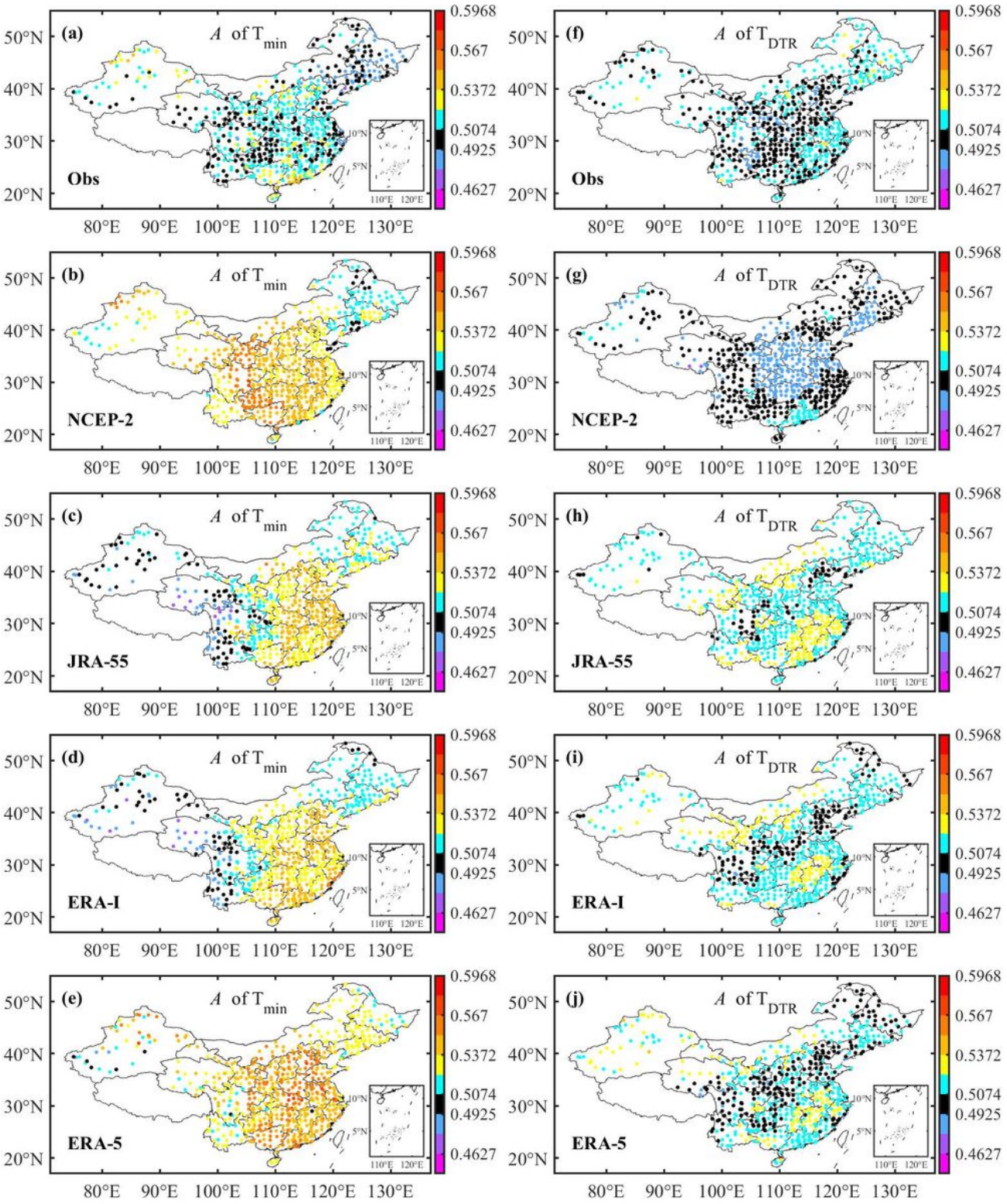

Figure 5

Same as Fig. 4 but for Tmin (left column) and TDTR (right column). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever

on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

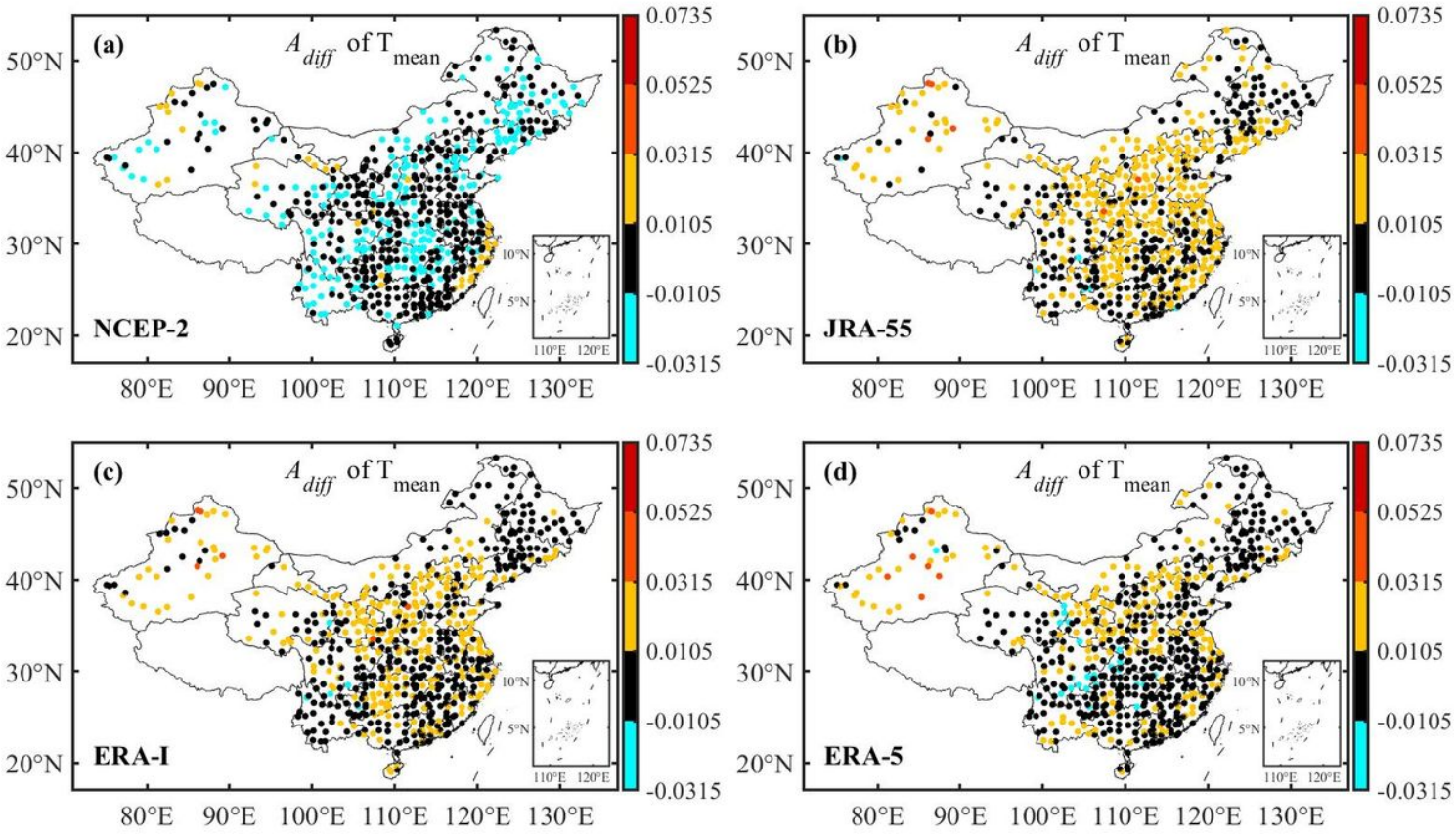

Figure 6

Same as Fig. 2 but for Tmean. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

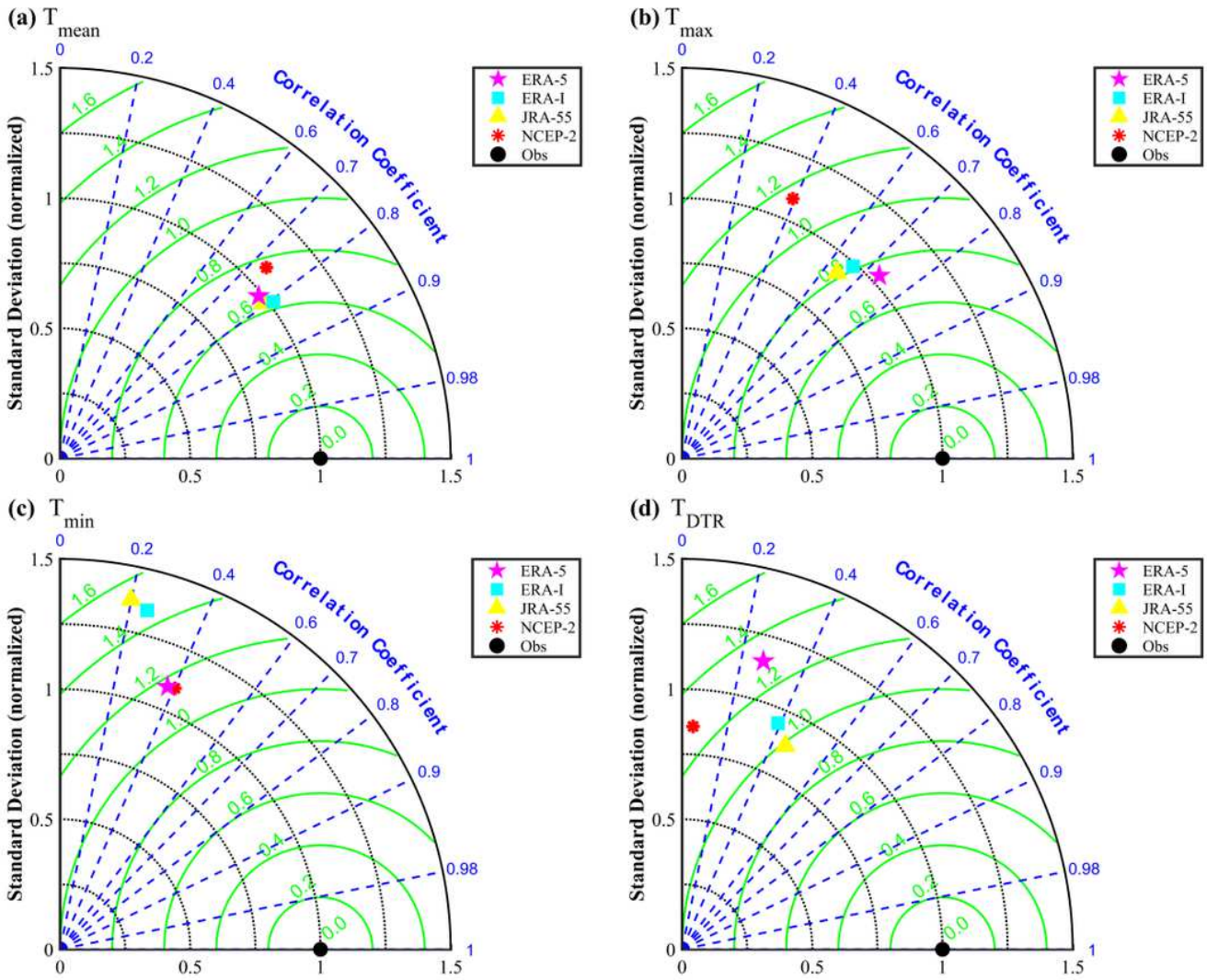

Figure 7

Taylor diagram for a Tmean, b Tmax, c Tmin, d TDTR of A from observation and interpolated reanalysis. The black dot stands for the results calculated from observations, which serves as the reference. The radial distance from the origin is proportional to the standard deviation of a pattern normalized by reference pattern, the centered root mean square (RMS) difference between the reference and re-analyses is proportional to their distance apart and the correlation between the reference and re-analyses is given by the azimuthal position of a given reanalysis.

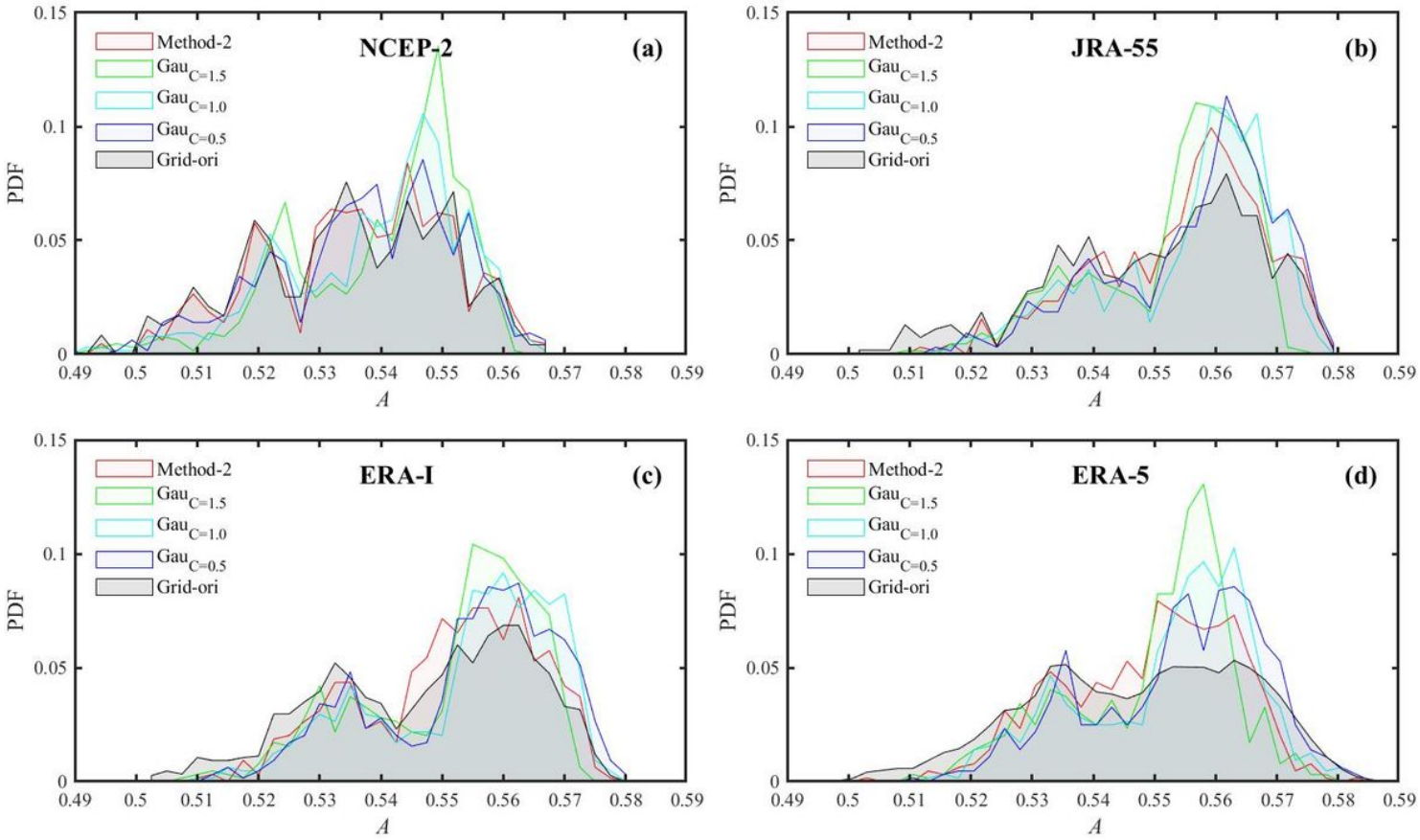

Figure 8

Probability density function (PDF) for A of Tmean from a NCEP-2, b JRA-55, c ERA-I and d ERA-5 by different interpolation ways. Gau means the Gaussian weight function interpolation with different values of C (C=1.5, green; C=1.0, cyan; C=0.5, blue), Method-2 represents interpolation by the closest points to the stations (red) and Grid-ori from original reanalysis (grey) shadow.