On the Suitability of Current Atmospheric Reanalyses for Regional Warming Studies over China

Chunlüe Zhou[1], Yanyi He[1], Kaicun Wang[1*]

[1]College of Global Change and Earth System Science, Beijing Normal University, Beijing, 100875, China

***Corresponding Author**: Kaicun Wang, College of Global Change and Earth System Science, Beijing Normal University. Email: kcwang@bnu.edu.cn; tel.: +86 (10)-58803143; fax: +86 (10)-58800059.

Submitted to *Atmospheric Chemistry and Physics*

May 12, 2018

Abstract

Reanalyses are widely used because they add value to routine observations by generating physically or dynamically consistent and spatiotemporally complete atmospheric fields. Existing studies include extensive discussions of the temporal suitability of reanalyses in studies of global change. This study adds to this existing work by investigating the suitability of reanalyses in studies of regional climate change, in which land-atmosphere interactions play a comparatively important role. In this study, surface air temperatures (T_a) from 12 current reanalysis products are investigated; in particular, the spatial patterns of trends in T_a are examined using homogenized measurements of T_a made at ~2200 meteorological stations in China from 1979 to 2010. The results show that ~80% of the mean differences in T_a between the reanalyses and the *in situ* observations can be attributed to the differences in elevation between the stations and the model grids. Thus, the T_a climatologies display good skill, and these findings rebut previous reports of biases in T_a. However, the biases in the T_a trends in the reanalyses diverge spatially (standard deviation=0.15-0.30 ℃/decade using 1 °×1 ° grid cells). The simulated biases in the trends in T_a correlate well with those of precipitation frequency, surface incident solar radiation (R_s), and atmospheric downward longwave radiation (L_d) among the reanalyses (r=-0.83, 0.80 and 0.77; p<0.1) when the spatial patterns of these variables are considered. The biases in the trends in T_a over southern China (on the order of -0.07 ℃/decade) are caused by biases in the trends in R_s, L_d and precipitation frequency on the order of 0.10 ℃/decade, -0.08 ℃/decade, and -0.06 ℃/decade,

respectively. The biases in the trends in T_a over northern China (on the order of -0.12 ℃/decade) result jointly from those in L_d and precipitation frequency. Therefore, improving the simulation of precipitation frequency and R_s helps to maximize the signal component corresponding to regional climate. In addition, the analysis of T_a observations helps representing regional warming in ERA-Interim and JRA-55. Incorporating vegetation dynamics in reanalyses and the use of accurate aerosol information, as in the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), would lead to improvements in the modelling of regional warming. The use of the ensemble technique adopted in the twentieth-century atmospheric model ensemble ERA-20CM significantly narrows the uncertainties associated with regional warming in reanalyses (standard deviation=0.15 ℃/decade).

1. Introduction

Observations and models are two fundamental approaches used in the understanding of climate change. Observations provide a direct link to the climate system via instruments, whereas models provide an indirect link and include information derived from measurements, prior knowledge and theory.

A large number of meteorological observations have been accumulated. These measurements, which are derived from a variety of sources, such as surface stations, ships, buoys, radiosondes, airplanes and satellites, record quantities that include near-surface and upper-air temperatures, humidity, wind and pressure. They constitute a major source of atmospheric information through the depth of the troposphere but suffer from incomplete spatiotemporal coverage and observation errors, including systematic, random and representation errors. Recent satellite-based observations have much better coverage; however, they suffer from other notable limitations, including temporal inhomogeneities (e.g., satellite drift) and retrieval errors (Bengtsson et al., 2007). These spatiotemporally varying gaps restrict the effective application of observations alone in climate research.

To fill in the gaps in observations, models are needed. Such models can be very simple; examples of simple models include linear interpolation or geo-statistical approaches that are based on the spatial and temporal autocorrelation of the observations. However, these models lack the necessary dynamical or physical mechanisms. Given the steady progress of numerical weather prediction (NWP) models in characterizing the global atmospheric circulation in the early 1980s (Bauer et al., 2015), the first generation of reanalyses was produced by combining observations and dynamic models to provide the first global atmospheric datasets for use in scientific research (Bengtsson et al., 1982a, b).

After realizing the great value of this kind of reanalysis in atmospheric research, a step forward was taken with the suggestion made by Bengtsson and Shukla (1988) and Trenberth and Olson (1988) that most meteorological observations should be optimally assimilated under a fixed dynamical system over a period of time long enough to be useful for climate studies. In this way, available observations are ingested by advanced data assimilation techniques to provide a continuous initial state for an NWP model to produce the next short-term forecast. This procedure thus generates physically consistent and spatiotemporally complete three-dimensional atmospheric fields that are updated in light of observations.

Taking this suggestion as a guide, and given the improvements that have been made since the mid-1990s in the integrity of the observations, the models and the assimilation methods used, successive generations of atmospheric reanalyses established by several institutes have improved in quality. These reanalyses include the first two generations of global reanalyses produced by the National Centers for Environmental Prediction, NCEP-R1 (Kalnay et al., 1996) and NCEP-R2 (Kanamitsu et al., 2002) and the reanalyses produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), ERA-15 (Gibson et al., 1997), ERA-40 (Uppala et al., 2005), and ERA-Interim (Dee et al., 2011b); the Japanese Meteorological Agency, JRA-25 (Onogi et al., 2007) and JRA-55 (Kobayashi et al., 2015); and the National

Aeronautics and Space Administration, the Modern-Era Retrospective Analysis for Research and Applications (MERRA) (Rienecker et al., 2011) and its updated version, MERRA-2 (Reichle et al., 2017).

These reanalyses produce global gridded datasets that cover multiple time scales and include a large variety of atmospheric, oceanic and land surface parameters, many of which are not easily or routinely observed but are dynamically constrained by large numbers of observations from multiple sources assimilated using fixed NWP models. During the data assimilation, prior information on uncertainties in the observations and models are used to perform quality checks, to derive bias adjustments and to assign proportional weights. Therefore, such reanalyses add value to the instrumental record through their inclusion of bias adjustments, their broadened spatiotemporal coverage and their increased dynamical integrity or consistency.

Previous studies have revealed that such reanalyses have contributed significantly to a more detailed and comprehensive understanding of the dynamics of the Earth's atmosphere (Dee et al., 2011b;Kalnay et al., 1996;Nguyen et al., 2013;Kidston et al., 2010;Simmonds and Keay, 2000;Simmons et al., 2010;Mitas and Clement, 2006). Extensive assessment studies have reported that most reanalyses display a certain level of performance in terms of their absolute values (Betts et al., 1996;Zhou and Wang, 2016b;Betts et al., 1998), interannual variability (Lin et al., 2014;Lindsay et al., 2014;Zhou and Wang, 2017a, 2016a;Wang and Zeng, 2012), distributions (Gervais et al., 2014;Heng et al., 2014;Mao et al., 2010) and relationships among variables (Niznik and Lintner, 2013;Cash et al., 2015;Zhou et al., 2017;Zhou and Wang, 2016b;Betts, 2004) over regions worldwide. However, these aspects of reanalyses still contain certain errors that restrict the general use of reanalyses, especially in climate applications.

The errors displayed by reanalysis products arise from three sources: observation error, model error and assimilation error (Thorne and Vose, 2010;Parker, 2016;Lahoz and Schneider, 2014;Dee et al., 2014;Zhou et al., 2017). Specifically, observation error incorporates systematic and random errors in instruments and their replacements, errors in data reprocessing and representation error, which arises due to the spatiotemporal incompleteness of observations (Dee and Uppala, 2009;Desroziers et al., 2005). Model error refers mainly to the inadequate representation of physical processes in NWP models (Peña and Toth, 2014;Bengtsson et al., 2007), such as the lack of time-varying surface conditions, such as vegetation growth (Zhou and Wang, 2016b;Trigo et al., 2015), and incomplete cloud-precipitation-radiation parameterizations (Fujiwara et al., 2017;Dolinar et al., 2016). Assimilation error describes errors that arise in the mapping of the model space to the observation space and errors in the topologies of cost functions (Dee, 2005;Dee and Da Silva, 1998;Lahoz and Schneider, 2014;Parker, 2016).

These reanalyses mentioned above consist of the true climate signal and the nonlinear interactions among the observation error, the model error, and the assimilation error that arise during the assimilation process. These time-varying errors can introduce spurious trends without being eliminated by data assimilation systems. Many spurious variations in climate signals were also identified in the early-generation reanalyses (Bengtsson et al., 2004;Andersson et al., 2005;Chen et al., 2008;Zhou and Wang, 2016a, 2017a;Zhou et al., 2017;Schoeberl et al., 2012;Xu and Powell, 2011;Hines et al., 2000;Cornes and Jones, 2013). Therefore, reanalyses produced using the existing reanalysis strategy may not accurately capture climate trends (Trenberth et al., 2008), even though they may contain relatively accurate estimates of synoptic or interannual variations in the Earth's atmosphere.

An emerging requirement for climate applications of reanalysis data is the accurate representation of decadal variability, further increasing the confidence in the estimation of climate trends. This kind of climate reanalysis is required to be free, to a great extent, from other spurious non-climatic signals introduced by changing observations, model imperfections and assimilation error; that is, they must maintain temporal consistency. Therefore, the extent to which climate trends can be assessed using reanalyses attracts much attention and sparks heated debates (Thorne and Vose, 2010;Dee et al., 2011a;Dee et al., 2014;Bengtsson et al., 2007).

Given the great progress that has been made in climate forecasting models (which provide more accurate representations of climate change and variability) and coupled data assimilation, many efforts have been made by several institutes to build consistent climate reanalyses using the strategy of assimilating a relatively small number of high-quality long-term observational datasets. The climate reanalyses of this new generation extend back to the late nineteenth century and include the Climate Forecast System Reanalysis (CFSR), which is produced by the National Centers for Environmental Prediction (Saha et al., 2010); NOAA 20CRv2c, which is produced by the University of Colorado's Cooperative Institute for Research in Environmental Sciences (CIRES) in cooperation with the National Oceanic and Atmospheric Agency (NOAA) (Compo et al., 2011); and ERA-20C (Poli et al., 2016), ERA-20CM (Hersbach et al., 2015) and CERA-20C (Laloyaux et al., 2016), which are produced by the ECMWF. Compo et al. (2013) suggested that the NOAA 20CRv2c reanalysis can reproduce the trend in global mean surface air temperatures. In addition, the uncertainties estimated from multiple ensembles are provided to increase the confidence of the climate trends (Thorne and Vose, 2010;Dee et al., 2014).

From NWP-like reanalyses to climate reanalyses, existing studies focus mainly on comparing the differences in temporal variability between the reanalyses and observations using some statistical metrics, e.g., the mean values, standard deviations, interannual correlations, probability density functions and trends of surface air temperature over regions worldwide. These evaluations provide insight into the temporal evolution of the Earth's atmosphere. However, they lack the performance evaluations used in reanalyses in representing the spatial patterns of these statistics associated with the role of the coupled land-atmosphere and dynamical processes of the climate system. Moreover, the assessment of these spatial patterns provides a direct means of examining the most prominent advantage of reanalyses over geo-statistical interpolation; thus, the spatial patterns require comprehensive investigation.

This study employs high-density station-based datasets of quantities including surface air temperatures (T_a), the surface incident solar radiation (R_s), the surface downward longwave radiation (L_d), and precipitation measured at ~2200 meteorological stations within China from 1979 to 2010. It provides a quantitative examination of the simulated patterns of variations in T_a in both the NWP-like and climate reanalyses and considers the climatology, the interannual variability, the mutual relationships among relevant quantities, the long-term trends and their controlling factors. The results indicate the strengths and weaknesses of the current reanalyses when applied in regional climate change studies and provide possible ways to improve these reanalyses in the near future.

2. Data and Methods

2.1 Observational Datasets

The latest comprehensive daily dataset (which contains averages at 0, 6, 12, and 18 UTC) of quantities that include T_a, precipitation, sunshine duration, relative humidity, water vapor pressure, surface pressure and the cloud fraction from approximately 2400 meteorological stations in China from 1961 to 2014, of which only approximately 194 participate in global exchanges, is obtained from the China Meteorological Administration (CMA; http://data.cma.cn/data). Approximately 2200 stations with complete and homogeneous data are selected for use in this study (Wang and Feng, 2013;Wang, 2008;Wang et al., 2007). The high density of meteorological stations in China promotes the representation of regional patterns in surface warming by reanalyses and the assessment of the skill of simulations.

R_s values based on the revised Ångström-Prescott equation (Wang et al., 2015;Yang et al., 2006;Wang, 2014) are used in this study. The derived R_s values consider the effects of Rayleigh scattering, water vapor absorption and ozone absorption (Wang et al., 2015;Yang et al., 2006) and can accurately reflect the effects of aerosols and clouds on R_s over China (Wang et al., 2012;Tang et al., 2011). Several intensive studies have reported that the derived R_s values can accurately depict the interannual, decadal and long-term variations in R_s (Wang et al., 2015;Wang, 2014;Wang et al., 2012).

L_d is typically estimated by first determining the clear-sky radiation and atmospheric emissivity (Brunt, 1932;Choi et al., 2008;Bilbao and De Miguel, 2007), and then correcting for the cloud fraction (Wang and Liang, 2009;Wang and Dickinson, 2013). The derived L_d values can directly reflect the greenhouse effect of atmospheric water vapor and clouds. Additionally, precipitation frequency is defined as days in a year with daily precipitation at least 0.1 mm in this study, which has been shown to provide a good indication of the effects of precipitation on the interannual variability and trends in T_a (Zhou et al., 2017). Taken together, the derived R_s and L_d values are able to physically quantify the effects of solar radiation and the greenhouse effect on surface warming. Precipitation frequency can regulate the partitioning of available energy into latent and sensible heat fluxes and thus modulates the variations in T_a (Zhou et al., 2017;Zhou and Wang, 2017a).

2.2 Reanalysis Products

All of the major global atmospheric reanalysis products are included in this study (Table 1). The reanalyses are summarized below in terms of three aspects, i.e., the observations assimilated and the forecast model and assimilation method used. The NWP-like reanalyses assimilate many conventional and satellite datasets from multiple sources (Table 1) to characterize the basic upper-air atmospheric fields; the spatiotemporal errors of these datasets vary with time. In particular, the ERA-Interim and JRA-55 reanalyses incorporate many observations of T_a, and the MERRA2 reanalysis includes aerosol optical depth estimates from satellite retrievals and model simulations based on emission inventories, whereas most of the other reanalyses use climatological aerosols (Table 1). To derive consistent long-term climate signals, the new strategy adopted by climate reanalyses involves the assimilation of a small number of relatively effective observed variables, e.g., surface pressure (Table 1). Except for its lack of the assimilation of surface pressure, ERA-20CM employs the same forecast model and external forcings as ERA-20C (Table 1); thus, the inclusion of ERA-20CM in this study provides a useful benchmark series against which to ascertain the skill that is added by assimilating various observations and to cognize the advantage of ensemble simulations. The reanalyses adopt different sea surface temperatures (SSTs) and sea ice concentrations for different time periods, which may lead to temporal discontinuities in the climate signals derived from the reanalyses (Table 1). To address this issue, the boundary conditions in CFSR are derived from its coupled ocean-sea ice models instead of observations (Table 1). CFSR, NOAA 20CRv2c and NOAA 20CRv2 use monthly greenhouse gases (GHGs) with annual means near those used in CMIP5. On the other hand, in ERA-Interim, the GHGs increase more slowly than in CMIP5 after 2000. Finally, NCEP-R1 and NCEP-R2 adopt constant global mean concentrations of the GHGs (Table 1).

The forecast model is a fundamental component of a reanalysis that provides the background fields to the assimilation system. Different reanalyses produced by a single institute generally use similar physical parameterizations; however, updated versions of these parameterizations and higher spatial resolutions are used in the newer generations of these realizations (Table 1). Note that the CFSR is classified into climate reanalysis in this study, mainly because it adopts a climate forecast system (Table 1). The assimilation methods adopted by the current reanalyses incorporate variational methods (3D-Var and 4D-Var) and the ensemble Kalman filter (EnKF) approach (Table 1).

The 2-m T_a in NCEP-1, NCEP-2, MERRA, MERRA-2, ERA-20C, ERA-20CM, CERA-20C, NOAA 20CRv2c, NOAA 20CRv2 and CFSR are model-derived fields that are functions of the surface skin temperature, the temperature at the lowest model level, the vertical stability and the surface roughness, which are constrained primarily by observations of upper-air variables and the surface pressure (Kanamitsu et al., 2002;Rienecker et al., 2011;Reichle et al., 2017;Poli et al., 2016;Hersbach et al., 2015;Laloyaux et al., 2016;Compo et al., 2011;Saha et al., 2010). However, the T_a in ERA-Interim and JRA-55 are post-processing products by a relatively simple analysis scheme between the lowest model level and the surface and are analysed using ground-based observations of T_a, with the help of Monin-Obukhov similarity profiles consistent with the model's parameterization of the surface layer (Dee et al., 2011b;Kobayashi et al., 2015). Additionally, radiation calculations are diagnostically determined from the prognostic cloud condensate microphysics parameterization, and the cloud macrophysics parameterization assumes a maximum-random cloud overlapping scheme (Saha et al., 2010;Dolinar et al., 2016).

2.3 Method Used to Homogenize the Observed Time Series

Problems related to the observational infrastructure (e.g., instrument ageing and changes in observing practices) and station relocations can also lead to false temporal heterogeneity in time series. Therefore, it is necessary to diminish the impact of data inhomogeneities on the trends in the observed variables during the study period of 1979-2010.

We use the RHtestsV4 software package (Wang and Feng, 2013) to detect and homogenize the breakpoints in the monthly time series. The package includes two algorithms. Specifically, the PMFred algorithm is based on the penalized maximal *F*-test (*PMF*) without a reference series (Wang, 2008), and the PMTred algorithm is based on the penalized maximal *t*-test (*PMT*) with a reference series (Wang et al., 2007).

In this study, we first use the PMFred algorithm to identify potential reference series at the 95% significance level. We then reconstruct homogenous series for each inhomogeneous series using the following steps: 1) horizontal and vertical distances from the inhomogeneous station of less than 110 km and 500 m, respectively, are specified; 2) correlation coefficients between the first-order difference in the homogeneous series with that in the inhomogeneous one exceeding 0.9 are required; and 3) the first ten homogeneous series are averaged using inverse distance weighting to produce a reference series for the inhomogeneous station. Finally, we apply the PMTred algorithm to test all of the inhomogeneous series using the nearby reference series. Several intensive studies have been conducted that indicate the PMTred algorithm displays good performance in detecting change points in inhomogeneous series (Venema et al., 2012;Wang et al., 2007).

If a breakpoint is found to be statistically significant, the quantile-matching (QM) adjustment in RHtestsV4 is recommended for making adjustments to the time series (Wang et al., 2010;Wang and Feng, 2013); in such cases, the longest available segment from 1979 to 2010 is used as the base segment. The QM adjustment aims to match the empirical distributions from all of the detrended segments with that of the specific base segment (Wang et al., 2010). In addition, we replicate the procedures above for the sparsely distributed stations over western China and the Tibetan Plateau. The PMTred algorithm and the QM adjustment have recently been used successfully to homogenize climatic time series (Aarnes et al., 2015;Tsidu, 2012;Dai et al., 2011;Siswanto et al., 2015;Wang and Wang, 2016;Zhou et al., 2017).

As such, the significant breakpoints are detected and adjusted at a confidence level of 95% at 1092 of the 2193 (49.8%) stations for the T_a time series; 1079 of the 2193 (49.2%) stations for the R_s time series; 64 of the 2193 (2.9%) stations for precipitation frequency time series; 971 of the 2193 (44.2%) stations for the L_d time series; 944 of the 2193 (43.0%) stations for the water vapor pressure time series; and 956 of the 2193 (43.6%) stations for the cloud fraction time series.

2.4 Trend Calculations, Partial Linear Regression, and Total Least Squares

The bias, root mean squared error (*RMSE*) and standard deviation are used to assess the absolute value of T_a. The trends in T_a and the relevant variables are calculated using the ordinary least squares method (OLS) and the two-tailed Student's *t*-test. To determine whether the reanalyses contain biases in these trends, the two-tailed Student's *t*-test is also applied to the differences in the time series between the reanalyses and the homogeneous observations.

The partial least squares approach is used to investigate the net relationship between the detrended T_a values and the relevant variables (R_s, L_d and precipitation frequency) after statistically excluding the confounding effects among the relevant variables (Zhou et al., 2017). To evaluate the potential collinearity of independent variables in the regression model, the variance inflation factor (*VIF*) is calculated. The *VIFs* for R_s, precipitation frequency and L_d are less than 4. Specifically, the *VIF* for China of 2.19 is much less than the threshold of 10, above which the collinearity of regression models is bound to adversely affect the regression results (Ryan, 2008).

The Pearson correlation coefficient (r) is used to reveal the spatial relationship between T_a and the relevant variables. To further investigate the relationship between the spatial distributions of the biases in the trends in T_a and the relevant parameters among the twelve reanalysis products, the weighted total least squares (WTLS) is adopted, in which the spatial standard deviations and correlations of pairs of variables on 1°×1° grid cells are included (Reed, 1989;York et al., 2004;Golub and Van Loan, 1980;Hyk and Stojek, 2013;Tellinghuisen, 2010):

$$\omega(x_i) = 1/\hat{\sigma}_{x_i}^2 \qquad (1)$$

$$\omega(y_i) = 1 \Big/ \hat{\sigma}_{y_i}^2 \tag{2}$$

$$W_i = \frac{\omega(x_i)\cdot\omega(y_i)}{\omega(x_i) + b^2\omega(y_i) - 2b\cdot r_i\sqrt{\omega(x_i)\cdot\omega(y_i)}} \tag{3}$$

$$U_i = x_i - \sum_i^n (W_i \cdot x_i) \Big/ \sum_i^n (W_i) \tag{4}$$

$$V_i = y_i - \sum_i^n (W_i \cdot y_i) \Big/ \sum_i^n (W_i) \tag{5}$$

$$\beta_i = W_i[\frac{U_i}{\omega(y_i)} + \frac{b\cdot V_i}{\omega(x_i)} - (b\cdot U_i + V_i)\frac{r_i}{\sqrt{\omega(x_i)\cdot\omega(y_i)}}] \tag{6}$$

$$b = \frac{\sum_{i=1}^n W_i\cdot\beta_i\cdot V_i}{\sum_{i=1}^n W_i\cdot\beta_i\cdot U_i} \tag{7}$$

where x_i and y_i are the median trends in x and y (e.g., T_a and R_s) for the i^{th} reanalysis product; $\hat{\sigma}_{x_i}$, $\hat{\sigma}_{x_i}$ and r_i are the spatial standard deviations and correlations of the trends in x and y for the i^{th} reanalysis product; β_i is the least squares-adjusted value; W_i is the weight of the residual error; and b is the slope estimated by iterative methods with a relative tolerance of 10^{-16}.

The Monte Carlo method with 10000 experiments is applied to estimate the 90% confidence intervals of the slope b. In the Monte Carlo method, the grid index for the 1 °×1 ° grid cells over China, which ranges from 1 to 691, is generated as a random number. On this basis, we can sample the spatial pattern in the biases in the trends in T_a, R_s, L_d and precipitation frequency. We then calculate the median trends and their spatial standard deviations and correlations for each experiment used in the WTLS.

3. Results

3.1 Dependency of Surface Air Temperature Differences on Elevation Differences

Fig. 1 illustrates the differences in T_a from the NWP-like reanalyses and climate reanalyses relative to the homogenized station-based observations over China during the period of 1979-2010. When the T_a values measured at the stations are compared directly with those in the corresponding model grid cells, the results indicate that the reanalysis products underestimate T_a over most of the regions in China (by -0.28 ℃ to -2.56 ℃). These discrepancies are especially pronounced over the Tibetan Plateau and Middle China, where the underestimation ranges from-2.75 ℃ to -7.00 ℃ and from -1.19 ℃ to -2.91 ℃, respectively (Fig. 1 and Table 2). A homogenizing adjustment of 0.03 ℃ from the raw T_a observations is insufficient to cancel the underestimation of T_a by the reanalyses (Fig. 1 and Table 2). Similar biases in T_a within various regions worldwide have been widely reported by previous studies (Mao et al., 2010;Pitman and Perkins, 2009;Reuten et al., 2011;Wang and Zeng, 2012;Zhou et al., 2017;Zhou and Wang, 2016b).

However, we found that the spatial patterns in the differences in T_a are well correlated with the elevation differences between models and stations, as reflected by correlation coefficients (r) of 0.85 to 0.94 (Figs. 2 and S1). These results are in accordance with the reports from NCEP-R1, NCEP-R2 and ERA-40 (You et al., 2010;Ma et al., 2008;Zhao et al., 2008). The elevation differences (ΔHeight; Figs. 2 and S1) between the stations and the model grids consists of the filtering error in the elevations used in the spectral models (Δf) and differences in the site-to-grid elevations (Δs) due to the complexity of the orographic topography. We further quantify the relative contributions of these factors to the T_a differences. The elevation differences can explain approximately 80% of the T_a differences; approximately 74% is produced by the site-to-grid elevation differences, and approximately 6% is produced by the filtering error in the elevations used in the spectral models (Fig. 2).

The regression coefficient of the differences in T_a is approximately 6 ℃/1 km, which is similar to the lapse rate at the surface (Fig. 2). Lapse rate values that exceed 6 ℃/1 km can be seen over the Tibetan Plateau (shown as red dots in Fig. 2). This result is very consistent with the reported lapse rates over China (Li et al., 2015;Fang and Yoda, 1988). In addition, the rate of decrease in the model filtering error is approximately 4 ℃/1 km among the twelve reanalyses (Fig. 2). These results have important implications for the skill of the simulated T_a climatologies of the twelve reanalyses over China.

3.2 Comparison of Regional-scale Surface Air Temperature Series

Fig. 3 shows Taylor diagrams of annual T_a anomalies from the observations and reanalyses over China and its seven subregions. We find that the correlations between the annual T_a anomalies in the twelve reanalysis products and the observations are reasonably strong, as reflected by a median r of 0.95 (Fig. 3), despite the relatively weak correlations over the Tibetan Plateau associated with NCEP-R2 (r=0.24) and CFSR (r=0.53). The simulated time series of T_a anomalies over eastern China are depicted most accurately by the reanalyses (Fig. 3c-g).

Overall, the NWP-like reanalyses (denoted by numbers 3-7) display better skill than the climate reanalyses (denoted by numbers 8-14) in this regard (Fig. 3).

ERA-Interim and JRA-55 display the best performance in the simulated time series of T_a anomalies over China (r=1.00, RMSE=0.05 ℃) and the seven regions (r=0.98, RMSE=0.1 ℃) (Fig. 3), perhaps due to their analysis of surface air temperature observations (Table 1).

Comparing the T_a values from MERRA2 and MERRA shows that MERRA2 displays improved performance over northern China, as reflected by an increase in the correlation coefficient of 0.1 and a reduction in the RMSE of 0.1 ℃ (Fig. 3). This result may occur because MERRA2 includes time-varying aerosol loadings (Balsamo et al., 2015;Reichle et al., 2011). However, the incorporation of this information does not improve the results over Southeast China (Fig. 3h).

CERA-20C displays better performance than ERA-20C and ERA-20CM, perhaps related to the inclusion of coupled climate forecast models and data assimilation, as well as the assimilation of surface pressure data in CERA-20C (Fig. 3 and Table 1). NOAA 20CRv2c and NOAA 20CRv2 display moderate performance in this regard (r=0.8, RMSE=0.3 ℃) (Fig. 3), and the former reanalysis displays no improvement in performance, despite its use of new boundary conditions (Compo et al., 2011).

3.3 Key Factors Regulating Regional Temperature Change

This section discusses key factors that control regional temperature change from the perspective of energy balance and its partitioning. The R_s heats the surface, and the portion of this radiation that becomes the sensible heat flux heats the air near the surface (Zhou and Wang, 2016b;Wang and Dickinson, 2013;Zhou and Wang, 2016c). Part of the energy absorbed by the surface is released back to space as outgoing longwave radiation; some of this radiation is reflected by clouds and is influenced by atmospheric water vapor, further warming the near-surface air (Wang and Dickinson, 2013). This process is known as the greenhouse effect on T_a and is quantified by L_d. Existing studies have suggested that precipitation frequency better represents the interannual variability in soil moisture in China than the precipitation amount (Wu et al., 2012;Piao et al., 2009;Zhou et al., 2017;Zhou and Wang, 2017a); in turn, soil moisture affects vegetation growth and drives changes in surface characteristics (e.g., surface albedo and roughness). These changes alter the partitioning of available energy and thus regulate changes in T_a.

Fig. 4 illustrates the partial relationships between the annual anomalies in T_a and R_s, the precipitation frequency and L_d. The results show that T_a is consistently positively correlated with R_s (except over the Tibetan Plateau) and L_d; however, it is consistently negatively correlated with precipitation frequency in the observations and the twelve reanalysis products (Fig. 4). Based on the observations, the interannual variations in T_a are determined in part by precipitation frequency and L_d in Northeast China and the northern part of Northwest China (Fig. 4). All of the reanalyses roughly capture these factors over these regions, although they display differences in the relative magnitudes (Fig. 4). Specifically, ERA-20CM, NOAA 20CRv2c, NOAA 20CRv2 and CFSR exhibit comparable relationships of T_a with precipitation frequency and L_d; however, MERRA, MERRA2, NCEP-R2, ERA-20C, and CERA-20C overestimate the relationship between T_a and precipitation frequency, and ERA-Interim, JRA-55, and NCEP-R1 overestimate the relationship of T_a with L_d over these regions (Fig. 4).

Over the North China Plain and Middle China, the interannual variations in T_a are partly determined by R_s, precipitation frequency and L_d (Fig. 4). The reanalyses roughly capture the effects of these three factors on T_a, although they display diverse combinations (Fig. 4). Among these combinations, JRA-55, MERRA2, ERA-20CM and ERA-Interim are comparable to the observations over these regions (Fig. 4). Over Southeast China, the interannual variations in T_a are primarily regulated by L_d, precipitation frequency and R_s (Fig. 4). The reanalyses exhibit slightly overestimated relationships of T_a with R_s and underestimated relationships with precipitation frequency (Fig. 4).

Over the Tibetan Plateau, the interannual variations in T_a are regulated by R_s and precipitation frequency (Fig. 4). Most of the reanalyses roughly capture the combinations of these factors but exhibit certain differences in the relative effects of R_s and precipitation frequency on T_a (Fig. 4). MERRA, MERRA2, NOAA 20CRv2c and NOAA 20CRv2 overestimate the relationships of T_a with R_s over the Tibetan Plateau (Fig. 4).

Overall, the spatial patterns of the simulated partial correlation of T_a with R_s in the reanalysis products are significantly correlated with those seen in the observations; r=0.13-0.35 (p<0.05) for the NWP-like reanalyses, and larger values of r=0.24-0.41 (p<0.05) are obtained for the climate reanalyses. Moreover, the spatial patterns in the sensitivity of T_a to R_s exhibit significant correlations (r=0.12-0.17, p<0.05) for most of the climate reanalyses (Table 1). Precipitation frequency displays the largest spatial correlations (r=0.16-0.43, p<0.05) of the sensitivity of T_a with these three relevant parameters in the reanalyses (Table 3). Significant spatial correlations reflecting the relationship (including the partial correlation and sensitivity) of T_a with L_d are also found (Table 1).

3.4 Regional Warming Trend Biases and Their Causes

1) The Whole of China

From 1979 to 2010 over China, T_a exhibits strong warming trends of 0.37 ℃/decade (p<0.05) in the observations and 0.22-0.48 ℃/decade (p<0.05) in the twelve reanalyses (Figs. 5 and S2-S3, Table 2). ERA-Interim and JRA-55 display spatial correlations with the observations (r=0.47 and 0.54, p<0.05) that are due at least partly to the inclusion of some T_a observations, whereas NCEP-R2 and ERA-20C display the worst performance (Figs. S3, Tables 1 and 3). Furthermore, approximately 87% of the observed trends in T_a over China can be explained by the greenhouse effect (i.e., 65% can be explained by the trend in L_d), precipitation frequency (29%) and R_s (-7%, due to the trend in radiative forcing of -1.1 $W \cdot m^{-2}$/decade) (Figs. S3-4). The influence of the greenhouse effect on the observed trends in T_a consists mainly of the trends in the atmospheric water vapor (42%) and the cloud fraction (3%) (Fig. S5). Among the reanalyses, over 90% of the trend in T_a can be explained by the greenhouse effect, precipitation frequency and R_s (Figs. S4-6). Specifically, ERA-Interim, JRA-55, MERRA and MERRA2 display the best ability to capture the contributions of the greenhouse effect (48% to 76%), precipitation frequency (22% to 34%) and R_s (-4% to 13%) to the trend in T_a over China (Figs. S4 and S6). The remaining NWP-like reanalyses (i.e., NCEP-R1 and NCEP-R2) substantially overestimate the contribution of R_s to the trend in T_a, whereas the climate reanalyses overestimate the contribution from L_d (Figs. S4 and S6).

We further quantify the contributions to the biases in the trend in T_a made by those in R_s, L_d and precipitation frequency among the twelve reanalyses over China (Figs. 6-7). Over China, the overestimated R_s trends (by 0.00-3.93 W·m^{-2}/decade; Figs. S8 and S13) increase the trends in T_a (by 0.02-0.16 ℃/decade; Fig. 7) in the twelve reanalyses; the underestimated L_d trends (by -0.25 to -1.61 W·m^{-2}/decade for the NWP-like reanalyses; Figs. S10 and S15) decrease the trends in T_a (by -0.05 to -0.25 ℃/decade for the NWP-like reanalyses; Fig. 7); and the biases in the trends in precipitation frequency (by approximately -1.5 days/decade for the NWP-like reanalyses and approximately 2.6 days/decade for the climate reanalyses; Figs. S9 and S14) decrease the trends in T_a (by 0.01 to 0.05 ℃/decade for the NWP-like reanalyses and -0.01 to -0.06 ℃/decade for the climate reanalyses; Fig. 7). Together, these effects produce an underestimate in the trends in T_a on the order of 0.10 ℃/decade in the reanalyses (Fig. 7 and Table 2).

2) Seven Subregions

Averaged trends over large areas may mask regional differences that reflect diverse regional warming biases and their causes (Figs. 5-7). The mean-adjusted spatial patterns of the biases in the trends in T_a appear to be consistent among the twelve reanalyses (Fig. S7) and mimic the spatial patterns in the overestimated R_s trends over the North China Plain, South China and Northeast China (Fig. S8), given the spatial correlations between these variables in most of the reanalyses (r=0.11-0.42, p<0.05) (Figs. 6 and S7-8, Table 3). However, the reanalyses still underestimate the trends in T_a over most of the regions. The key reason for this underestimation is the increase in precipitation frequency over Northwest China, the Loess Plateau, and Middle China seen in the NWP-like reanalyses and that seen over broader regions in the climate reanalyses (Figs. 5-6 and S9). This relationship is reflected by their negative spatial correlation, which has a maximum value of -0.62 (p<0.05) for MERRA (Table 3). Moreover, the decrease in L_d, which occurs due to the decreases in the atmospheric water vapor and cloud fraction that occur in the NWP-like reanalyses (Figs. S10-12), substantially cancels the warming effect of the overestimation of R_s on T_a over eastern China (Figs. 5 and S7). The opposite changes occur over Southeastern China in the climate reanalyses (Figs. 5 and S10). The effect of the changes in L_d is reflected by its spatial correlations of up to 0.50 (p<0.05) (Table 3).

The corresponding contributions to the biases in the T_a trend from are calculated from those in R_s, L_d and precipitation frequency over seven subregions of China (Figs. 6-7). Over northern China, biases in the trend in T_a result primarily from those in precipitation frequency and L_d (Figs. 6-7). Over Northeast China, the observations exhibit an amplified warming of 0.41 ℃/decade (p<0.05; Fig. 4 and Table 2). This warming is significantly underestimated by NCEP-R1, JRA-55, NOAA 20CRv2 and NOAA 20CRv2c (by on the order of -0.15 ℃/decade) and is overestimated by MERRA and CFSR (by on the order of 0.2 ℃/decade) (Figs. 6-7). These biases in the trends in T_a in the reanalysis are jointly explained by the warming

(0.04-0.48 ℃/decade) induced by the underestimated trends in precipitation frequency and the cooling (-0.04 to -0.42 ℃/decade) induced by the underestimated trends in L_d (Fig. 7).

Over Northwest China, the biases in the trend in precipitation frequency and L_d mainly explain the overestimated warming in NCEP-R2 (by 0.22 ℃/decade) (Fig. 7). The substantially underestimated trend in L_d induced by the decrease in the atmospheric water vapour and cloud fraction (Figs. S9-S12 and S16-17) lead to an underestimate of the warming in MERRA (by -0.22 ℃/decade) (Fig. 7).

Most of the reanalyses display weakened warming over the Tibetan Plateau and the Loess Plateau (Fig. 5 and S3, Table 2). In particular, NCEP-R1 and NCEP-R2 fail to reproduce the warming over the Tibetan Plateau, and MERRA fails to reproduce the warming over the Loess Plateau (Fig. 5 and S3, Table 2). The significant cooling biases in the trends in T_a (by -0.02 to -0.31 ℃/decade) over the Tibetan Plateau and the Loess Plateau result from the underestimated trends in L_d and the overestimated trends in precipitation frequency seen in most of the reanalyses (Figs. 5-7 and S9-12). These cooling biases are further induced by the underestimated trends in R_s (Figs. 5-7 and S8).

Over southern China, the biases in the trend in T_a are regulated by the biases in the trends in R_s, L_d and precipitation frequency (Figs. 6-7). Over Southeast China, the significantly overestimated trends in T_a (by 0.04, 0.02 and 0.17 ℃/decade, respectively) are induced by the overestimated trends in R_s (by 4.25, 3.34 and 6.27 $W \cdot m^{-2}$/decade, respectively) seen in ERA-Interim, JRA-55 and CFSR (Figs. 6-7 and

S8). The underestimated trends in T_a are induced by the overestimated trends in precipitation frequency and L_d in NCEP-R1, MERRA, ERA-20CM, CERA-20C, NOAA 20CRv2 and NOAA 20CRv2c (Figs. 6-7 and S9).

Over Middle China, the significantly overestimated trends in T_a (by 0.04, 0.06, 0.11, 0.03, 0.11 and 0.14 ℃/decade, respectively) are induced by the overestimated trends in R_s (by 2.09, 1.50, 2.59, 1.20 and 4.81 W·m^{-2}/decade, respectively) seen in ERA-Interim, JRA-55, ERA-20C, ERA-20CM, CERA-20C and CFSR (Figs. 6-7 and S8). The overestimated trends in precipitation frequency may lead to cooling in the trends in T_a in the reanalyses, especially for MERRA (which reflects an induced bias in the trend of -0.15 ℃/decade) over Middle China (Figs. 6-7 and S9).

Due to the underestimated trends in the atmospheric water vapor and the cloud fraction (Figs. S11-12), the underestimation of L_d produces a cooling effect on the trend in T_a (by -0.05 to -0.32 ℃/decade) in the reanalyses over the North China Plain (Figs. 6-7 and S10). However, due to the lack of inclusion of plausible trends in aerosol loading, the substantial increases in R_s over the North China Plain (Fig. S8) have strong warming effects on the trends in T_a (by 0.01 to 0.21 ℃/decade) in the reanalyses (Figs. 6-7 and S8). The biases in the trends in precipitation frequency (of approximately -2.5 days/decade for the NWP-like reanalyses and approximately 1.5 days/decade for some of the climate reanalyses) contribute some part of the biases in the trends in T_a (approximately 0.05 ℃/decade for the NWP-like reanalyses and -0.03 ℃/decade for the climate reanalyses).

Overall, the biases in the trends in T_a in the reanalyses can be substantially explained by those in L_d, precipitation frequency and R_s, but this effect varies regionally (Figs. 6-7). Over northern China, the biases in the trend in T_a (which are on the order of -0.12 ℃/decade) result primarily from a combination of those in L_d (which are on the order of -0.10 ℃/decade) and precipitation frequency (which are on the order of 0.05 ℃/decade), with relatively small contributions from R_s (which are on the order of -0.03 ℃/decade). Over southern China, the biases in the trend in T_a (which are on the order of -0.07 ℃/decade) are caused by those in R_s (which are on the order of 0.10 ℃/decade), L_d (which are on the order of -0.08 ℃/decade) and precipitation frequency (which are on the order of -0.06 ℃/decade) (Fig. S18).

3.5 Spatial Linkages of Biases in the Warming Trends in the Twelve Reanalyses

We next integrate the relationships of the spatial patterns in the biases in the trends in T_a with those in R_s, L_d and precipitation frequency over China in the twelve reanalyses (Fig. 8). The results show that the biases in the trends in T_a show significant correlations with R_s (r=0.80, slope=0.06, p=0.09) and precipitation frequency (r=-0.83, slope=-0.04, p=0.02) and L_d (r=0.77, slope=0.10, p=0.10) in the twelve reanalyses if information on these patterns is included. When the spatial patterns of the biases in the trends in these variables are not considered, the biases in the trends in T_a show relatively small correlations with R_s (r=0.32, slope=0.02, p>0.1), precipitation frequency (r=-0.51, slope=-0.02, p=0.09) and L_d (r=0.14, slope=0.02, p>0.1) in the reanalyses (Fig. 8). Similar results are obtained for the atmospheric water vapor (r=0.71, p=0.1) and the cloud fraction (r=-0.74, p=0.09) if their spatial patterns are considered (Figs. S19), and this relationship involving the cloud fraction is very similar to that associated with R_s (Figs. 8 and S19). Within the subregions of China, the biases in the trends in T_a show significant correlations with R_s (r=068 to 0.90, p<0.1), precipitation frequency (r=-0.55 to -0.94, p<0.1) and L_d (r=0.53 to 0.93, p<0.1) when the spatial patterns in the reanalyses are included (Fig. S20). These results provide a novel perspective that can be used to investigate the spatial relationships between biases in the trends in T_a and relevant quantities in reanalyses.

4. Discussion

In this section, we first examine the possible impacts of data homogenization on the trends in T_a. The trends in T_a derived from the original dataset are almost as high as those from the homogenized dataset, especially over the North China Plain and Northwest China (Fig. 5 and Table 2). Homogenization primarily adjusts breakpoints in time series (Wang, 2008), which occur mainly due to station relocation and changes in instruments (Cao et al., 2016;Li et al., 2017;Wang, 2014), and it helps to objectively depict trends in T_a, thus permitting the assessment of the modelled trends in T_a and its spatial patterns that are present in the reanalyses.

We found that the elevation differences between the models and the stations influence the biases in the trends in T_a but cannot explain the spatial patterns in the biases in the trends in T_a (average r=0.11) (Fig. S21). Comparison of the models that use the same grid (NOAA 20CRv2c vs. NOAA 20CRv2, MERRA vs. MERRA2, NCEP-R1 vs. NCEP-R2 and ERA-20C vs. ERA-20CM) shows that the one is correlated with elevation differences, but the other is not, which implies that this statistical correlation does not have physical significance. Nevertheless, the spatial patterns in the normalized trends in T_a (excluding the impacts of the absolute value of temperature on the trends) are very near to those of the trends (Fig. S22), implying that the differences in the absolute value of temperature have an important effect, given that the site-to-grid inconsistency can be neglected.

In the reanalyses, vegetation is only included as climatological information, but the vegetation displays a growth trend during the study period of 1979-2010 within China (Fig. S23). This discrepancy positively enlarges the biases in the trends in T_a due to the vegetation cooling effect (Zeng et al., 2017;Trigo et al., 2015). This effect is reflected by the negative spatial correlation (r=-0.26, p=0.00) between the inverted trend in the NDVI and the biases in the trend in T_a (Fig. S23). The growth of vegetation reduces T_a by regulating surface roughness, surface conductivity, soil moisture and albedo to partition greater amounts of available energy into latent heat fluxes, which leads to the formation of more precipitation (Shen et al., 2015;Spracklen et al., 2013). Thus, the inclusion of vegetation growth will improve the simulation of trends and especially the spatial pattern of T_a in the reanalyses through the incorporation of more complete physical parameterizations (Li et al., 2005;Dee and Todling, 2000;Trigo et al., 2015).

Due to their inclusion of surface air temperature observations, ERA-Interim and JRA-55 display high skill in reproducing the observed patterns; they have near-zero means (0.01 and 0.01 ℃/decade) and the smallest standard deviations (0.16 and 0.15 ℃/decade) of the trend biases among the twelve reanalysis products. However, pattern differences of 37.8% (standard deviation of trend bias/China-averaged trend) are still evident (Figs. 5 and 8). Although it does not incorporate surface air temperature observations, ERA-20CM presents a pattern (with a mean of -0.04 ℃/decade and a standard deviation of 0.15 ℃/decade; Figs. 5 and 8) that is comparable to those of ERA-Interim and JRA-55 and better than that of ERA-20C (mean of -0.08 ℃/decade and standard deviation of 0.20 ℃/decade; Figs. 5 and 8), which uses the same forecast model as ERA-20CM. These results imply that ensemble forecasting could be used to meet important goals. The ensemble simulation technique used in ERA-20CM also displays advantages in that it yields an improved simulated pattern of biases in the trends in R_s (SD=1.84 W m^{-2}/decade, 171%), precipitation frequency (SD=2.78days/decade, 122%) and L_d (SD=1.25 W m^{-2}/decade, 82%) (Fig. 8).

We consider the degree to which the ensemble assimilation technique can improve the spatial patterns of the biases in the trends in T_a in the reanalyses. We find that this technique can detect the biases in the trends in T_a over more another approximately 12% (8%) of the grid cells in CERA-20C, which incorporates 10 ensemble members (NOAA 20CR2vc and NOAA 20CR2v employ 56 ensemble members) (Figs. 5 l-n). However, the biases in the trends in T_a over these grid cells are not significant at a significance level of 0.05, according to Student's *t*-test, implying that the ensemble assimilation technique cannot explain the spatial pattern of the biases in the trends in T_a identified in this study (in Figs. 5 l-n).

To provide a preliminary discussion of the improvements in climate forecast models in reflecting patterns in climate trends, we compare the spatial patterns of the biases in the trends in R_s, precipitation frequency and L_d because observations of these variables are not included in the reanalyses. We find that the climate forecast models, i.e., ERA-20C, ERA-20CM, CERA-20C, NOAA 20CRv2c and NOAA 20CRv2, display better performance in reproducing the pattern of biases in the trends in R_s (mean of 1.36 vs. 2.18 W m^{-2}/decade; SD of 2.04 vs. 2.71 W m^{-2}/decade), precipitation frequency (mean of 1.32 vs. -1.44%/decade; SD of 3.57 vs. 6.14%/decade) and L_d (mean of 0.12 vs. -0.85 W m^{-2}/decade; SD of 1.33 vs. 1.50 W m^{-2}/decade) than the NWP-like models, i.e., ERA-Interim, NCEP-R1, MERRA, JRA-55, NCEP-R2 and MERRA2 (Fig. 8). In addition, because the SST boundary condition evolves freely in CFSR, the patterns of biases in the trends in R_s, precipitation frequency and L_d in CFSR differ substantially from those in the other reanalyses.

We also consider whether the spatial pattern of biases in the trend in T_a is altered by the atmospheric circulation patterns simulated by the ERA-20CM ensemble. In ERA-20CM, the atmospheric circulation patterns are influenced by SSTs and sea ice and then partly mediate the influence of global forcings on the trends in T_a. In ERA-20CM, the probability distribution function of the biases in the trends in T_a from outside the ensemble ranges incorporates that from Student's *t*-test at a significance level of 0.05 (Fig. 5k). This result has important implications in that 1) the climate variability in the ensembles under the different model realizations of SSTs and sea ice cover does not change the pattern of the biases in the trends in T_a (Fig. 5k); moreover,

2) Student's *t*-test exhibits a suitable ability to detect the significance of the biases in the trends in T_a (Fig. 5k) when considering the effects of interannual variability on the trend.

Overall, producing global or regional reanalyses that adequately reflect regional climate is challenging using the current strategy, and further improvements are required. The results and discussion above indicate some potential but challenging approaches that can be used to maximize the signal component corresponding to the regional climate in final reanalyses and robustly narrow the uncertainties in trends.

1) MERRA2's pioneering incorporation of time-varying aerosol loadings provides a way of improving the representation of regional temperature changes over regions such as the North China Plain where the impacts of aerosols on surface temperatures are significant. Thus, we encourage research groups to include accurate aerosol information and improve the skill of simulation of the energy budget and partitioning, especially of regional surface incident solar radiation, in other reanalyses.

2) To improve regional climate modelling, forecast output should be produced using a physical ensemble like that employed in ERA-20CM to quantify the uncertainties associated with the relevant parameterizations in the reanalyses, due to the impossibility of optimizing all of the biases. Meanwhile, careful ensemble design would likely yield useful information for use in improving models, assimilation methods and the bias correction of observations by exploring the interdependency among sources of errors. Such designs would undoubtedly have additional benefits for

further development, leading to the next generation of reanalyses.

3) To improve coupled land-atmospheric interactions, the true dynamics of land cover and use should be incorporated. Moreover, the physical parameterizations should be improved, including the responses of surface roughness, surface conductivity and albedo to regional climate. These changes would represent an improvement over the use of constant types and fractions of vegetation, as is done in ERA-Interim (Zhou and Wang, 2016b).

4) Given the implications of the spurious performance of the freely evolving boundary conditions in CFSR, homogeneous and accurate records of SST and sea ice should be produced.

Next-generation reanalyses, including both global and regional reanalyses, will assimilate and analyse *in situ* observations, satellite radiance, and other remote observations. In addition to short-term accuracy and long-term trends, they will also focus on spatial patterns by incorporating or improving accurate representations of land surface conditions and processes within the coupled weather and climate Earth systems. Thus, these reanalyses will advance the simulation of land-atmosphere interactions to yield high skill in studies of regional warming and the detection and attribution of regional climate change using various datasets, which frequently include global and regional reanalyses (Zhou et al., 2018;Zhou and Wang, 2016d;Herring et al., 2018;Trenberth et al., 2015;Stott, 2016;Dai et al., 2017;Zhou and Wang, 2017b). Additionally, the uncertainties associated with regional warming could be ascertained using physics ensembles with various equiprobable realizations of boundary conditions.

5. Conclusions

The reanalyses display differences in T_a when compared to the observations with a range of -10~10 ℃ over China. Approximately 74% and 6% of these differences can be explained by site-to-grid elevation differences and the filtering error in the elevations used in the spectral models. These results imply fairly good skill in the simulation of the climatology of T_a in the twelve reanalyses over China. Moreover, the twelve reanalyses roughly capture the interannual variability in T_a (median r=0.95). In the reanalyses, T_a displays a consistently positive correlation with R_s and L_d and is negatively correlated with precipitation frequency, as seen in observations, despite the evident spatial patterns in their magnitudes over China.

T_a exhibits a strong warming trend of 0.37 ℃/decade (p<0.05) in the observations and 0.22-0.48 ℃/decade (p<0.05) in the twelve reanalyses over China. In the observations, approximately 87% of the observed trend in T_a over China can be explained by the greenhouse effect (i.e., 65% can be explained by the trend in L_d), precipitation frequency (29%) and R_s (-7%, due to the trend in radiative forcing of -1.1 W·m^{-2}/decade).

However, the biases in the trends in T_a seen in the reanalyses relative to the observations display an evident spatial pattern (mean=-0.16~0.11 ℃/decade, SD=0.15-0.30 ℃/decade). The spatial patterns of the biases in the trends in the values of T_a in the reanalyses are significantly correlated with those in R_s (maximum r=0.42,

p<0.05), precipitation frequency (maximum r=-0.62, p<0.05) and L_d (maximum r=0.50, p<0.05). Over northern China, the biases in the trends in T_a (which are on the order of -0.12 ℃/decade) result primarily from a combination of those in L_d (which are on the order of -0.10 ℃/decade) and precipitation frequency (which are on the order of 0.05 ℃/decade), with relatively small contributions from R_s (which are on the order of -0.03 ℃/decade). Over southern China, the biases in the trends in T_a (which are on the order of -0.07 ℃/decade) are regulated by the biases in the trends in R_s (which are on the order of 0.10 ℃/decade), L_d (which are on the order of -0.08 ℃/decade) and precipitation frequency (which are on the order of -0.06 ℃/decade).

If information on spatial patterns is included, the simulated biases in the trends in T_a correlate well with those of precipitation frequency, R_s and L_d in the reanalyses (r=-0.83, 0.80 and 0.77, p<0.1); similar results are obtained for the atmospheric water vapor and the cloud fraction (r=0.71 and -0.74, p<0.1). These results provide a novel perspective that can be used to investigate the spatial relationships between the biases in the trends in T_a and the relevant parameters among the twelve reanalyses. Therefore, improving simulations of precipitation frequency and R_s helps to maximize the signal component corresponding to the regional climate. In addition, the analysis of T_a observations helps to improve the performance of regional warming in ERA-Interim and JRA-55. Incorporating vegetation dynamics in reanalyses and the use of accurate aerosol information, as in MERRA-2, would advance the modelling of regional warming. The ensemble technique adopted in ERA-20CM, a twentieth-century atmospheric model ensemble that does not assimilate observations, significantly narrows the uncertainties of regional warming in the reanalyses (standard deviation=0.15 ℃/decade).

Acknowledgements This study was funded by the National Key R&D Program of China (2017YFA0603601) and the National Natural Science Foundation of China (41525018). The latest observational datasets were obtained from the China Meteorological Administration (CMA; http://www.cma.gov.cn). Considerable gratitude is owed to several reanalysis working teams, including the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the ERA-Interim, ERA-20C, ERA-20CM and CERA-20C data (http://www.ecmwf.int/); the Global Modelling and Assimilation Office (GMAO) at the NASA Goddard Space Flight Center for providing the MERRA and MERRA2 data (http://gmao.gsfc.nasa.gov/); the NOAA Earth System Research Laboratory (ESRL) for providing the NCEP-R1, NCEP-R2, CFSR, NOAA 20CRv2 and NOAA 20CRv2c data (http://www.esrl.noaa.gov/); and the Climate Prediction Division of the Global Environment and Marine Department at the Japan Meteorological Agency for providing the JRA-55 data (http://jra.kishou.go.jp/). We thank the Expert Team on Climate Change Detection and Indices (ETCCDI) for providing the RHtestV4 package (http://etccdi.pacificclimate.org/software.shtml), the United States Geological Survey Earth Resources Observation and Science Data Center for providing the GTOPO30 data (http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html) and the working team of the Global Inventory Monitoring and Modelling System (GIMMS) project (https://ecocast.arc.nasa.gov/data/pub/gimms/). We thank Kevin E. Trenberth for his insightful suggestions.

References

Aarnes, O. J., Abdalla, S., Bidlot, J.-R., and Breivik, Ø.: Marine wind and wave height trends at different ERA-Interim forecast ranges, J. Clim., 28, 819-837, 10.1175/jcli-d-14-00470.1, 2015.

Andersson, E., Bauer, P., Beljaars, A., Chevallier, F., Holm, E., Janiskova, M., Kallberg, P., Kelly, G., Lopez, P., McNally, A., Moreau, E., Simmons, A. J., Thepaut, J. N., and Tompkins, A. M.: Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system, Bull. Am. Meteorol. Soc., 86, 387-402, 10.1175/bams-86-3-387, 2005.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389-407, 10.5194/hess-19-389-2015, 2015.

Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, Nature, 525, 47-55, 10.1038/nature14956, 2015.

Bengtsson, L., Kanamitsu, M., Kallberg, P., and Uppala, S.: FGGE research activities at ECMWF, Bull. Am. Meteorol. Soc., 63, 227-303, 1982a.

Bengtsson, L., Kanamitsu, M., Kallberg, P., and Uppala, S.: FGGE 4-dimensional data assimilation at ECMWF, Bull. Am. Meteorol. Soc., 63, 29-43, 1982b.

Bengtsson, L., and Shukla, J.: Integration of space and in situ observations to study global climate change, Bull. Am. Meteorol. Soc., 69, 1130-1143, 10.1175/1520-0477(1988)069<1130:iosais>2.0.co;2, 1988.

Bengtsson, L., Hagemann, S., and Hodges, K. I.: Can climate trends be calculated from reanalysis data?, J Geophys Res-Atmos, 109, D11111, 10.1029/2004jd004536, 2004.

Bengtsson, L., Haines, K., Hodges, K. I., Arkin, P., Berrisford, P., Bougeault, P., Kallberg, P., Simmons, A. J., Uppala, S., Folland, C. K., Gordon, C., Rayner, N., Thorne, P. W., Jones, P., Stammer, D., and Vose, R. S.: The need for a dynamical climate reanalysis, Bull. Am. Meteorol. Soc., 88, 495-501, 10.1175/bams-88-4-495, 2007.

Betts, A. K., Hong, S.-Y., and Pan, H.-L.: Comparison of NCEP-NCAR reanalysis with 1987 FIFE data, Monthly Weather Review, 124, 1480-1498, 10.1175/1520-0493(1996)124<1480:connrw>2.0.co;2, 1996.

Betts, A. K., Viterbo, P., and Beljaars, A. C. M.: Comparison of the land-surface interaction in the ECMWF reanalysis model with the 1987 FIFE data, Monthly Weather Review, 126, 186-198, 10.1175/1520-0493(1998)126<0186:cotlsi>2.0.co;2, 1998.

Betts, A. K.: Understanding hydrometeorology using global models, Bull. Am. Meteorol. Soc., 85, 1673-1688, 10.1175/bams-85-11-1673, 2004.

Bilbao, J., and De Miguel, A. H.: Estimation of daylight downward longwave atmospheric irradiance under clear-sky and all-sky conditions, J. Appl. Meteor. Climatol., 46, 878-889, 2007.

Brunt, D.: Notes on radiation in the atmosphere. I, Q. J. Roy. Meteorol. Soc., 58, 389-420, 1932.

Cao, L., Zhu, Y., Tang, G., Yuan, F., and Yan, Z.: Climatic warming in China according to a homogenized data set from 2419 stations, Int. J. Climatol., 36, 4384-4392, 10.1002/joc.4639, 2016.

Cash, B. A., III, J. L. K., Adams, J., Altshuler, E., Huang, B., Jin, E. K., Manganello, J., Marx, L., and Jung, T.: Regional structure of the Indian summer monsoon in observations, reanalysis, and simulation, J. Clim., 28, 1824-1841, 10.1175/jcli-d-14-00292.1, 2015.

Chen, J., Del Genio, A. D., Carlson, B. E., and Bosilovich, M. G.: The spatiotemporal structure of twentieth-century climate variations in observations and reanalyses. Part I: Long-term trend, J. Clim., 21, 2611-2633, 2008.

Choi, M., Jacobs, J. M., and Kustas, W. P.: Assessment of clear and cloudy sky parameterizations for daily downwelling longwave radiation over different land surfaces in Florida, USA, Geophys. Res. Lett., 35, L20402, 2008.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, Q. J. Roy. Meteorol. Soc., 137, 1-28, 10.1002/qj.776, 2011.

Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., Brohan, P., Jones, P. D., and McColl, C.: Independent confirmation of global land warming without the use of station temperatures, Geophys. Res. Lett., 40, 3170-3174, 2013.

Cornes, R. C., and Jones, P. D.: How well does the ERA-Interim reanalysis replicate trends in extremes of surface temperature across Europe?, J Geophys Res-Atmos, 118, 10262-10276, 10.1002/jgrd.50799, 2013.

Dai, A., Wang, J., Thorne, P. W., Parker, D. E., Haimberger, L., and Wang, X. L.: A new approach to homogenize daily radiosonde humidity data, J. Clim., 24, 965-991, 10.1175/2010jcli3816.1, 2011.

Dai, A., Rasmussen, R. M., Liu, C., Ikeda, K., and Prein, A. F.: A new mechanism for warm-season precipitation response to global warming based on convection-permitting simulations, Clim. Dyn., published online, 10.1007/s00382-017-3787-6, 2017.

Dee, D. P., and Da Silva, A. M.: Data assimilation in the presence of forecast bias, Q. J. Roy. Meteorol. Soc., 124, 269-295, 1998.

Dee, D. P., and Todling, R.: Data assimilation in the presence of forecast bias: The GEOS moisture analysis, Monthly Weather Review, 128, 3268-3282, 10.1175/1520-0493(2000)128<3268:daitpo>2.0.co;2, 2000.

Dee, D. P.: Bias and data assimilation, Q. J. Roy. Meteorol. Soc., 131, 3323-3343, 2005.

Dee, D. P., and Uppala, S.: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, Q. J. Roy. Meteorol. Soc., 135, 1830-1841, 10.1002/qj.493, 2009.

Dee, D. P., Källén, E., Simmons, A. J., and Haimberger, L.: Comments on “Reanalyses suitable for characterizing long-term trends”, Bull. Am. Meteorol. Soc., 92, 65-70, 10.1175/2010BAMS3070.1, 2011a.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553-597, 10.1002/qj.828, 2011b.

Dee, D. P., Balmaseda, M., Balsamo, G., Engelen, R., Simmons, A. J., and Thépaut, J. N.: Toward a consistent reanalysis of the climate system, Bull. Am. Meteorol. Soc., 95, 1235-1248, 10.1175/bams-d-13-00043.1, 2014.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis‐error statistics in observation space, Q. J. Roy. Meteorol. Soc., 131, 3385-3396, 2005.

Dolinar, E. K., Dong, X., and Xi, B.: Evaluation and intercomparison of clouds, precipitation, and radiation budgets in recent reanalyses using satellite-surface observations, Clim. Dyn., 46, 2123-2144, 10.1007/s00382-015-2693-z, 2016.

Fang, J.-Y., and Yoda, K.: Climate and vegetation in China (I). Changes in the altitudinal lapse rate of temperature and distribution of sea level temperature, Ecol. Res., 3, 37-51, 1988.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Kruger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC reanalysis intercomparison project (S-RIP) and overview of the reanalysis systems, Atmos. Chem. Phys., 17, 1417-1452, 10.5194/acp-17-1417-2017, 2017.

Gervais, M., Gyakum, J. R., Atallah, E., Tremblay, L. B., and Neale, R. B.: How well are the distribution and extreme values of daily precipitation over North America represented in the community climate system model? A comparison to reanalysis, satellite, and gridded station data, J. Clim., 27, 5219-5239, 10.1175/jcli-d-13-00320.1, 2014.

Gibson, J., Kållberg P, Uppala S, Nomura A, Hernandez A, and E., S.: ERA description, ECMWF. ERA-15 Project Report Series 1, European Centre for Medium-range Weather Forecasts, Reading, UK., 1997.

Golub, G. H., and Van Loan, C. F.: An analysis of the total least squares problem, SIAM J. Numer. Anal., 17, 883-893, 1980.

Heng, Z., Fu, Y., Liu, G., Zhou, R., Wang, Y., Yuan, R., Guo, J., and Dong, X.: A study of the distribution and variability of cloud water using ISCCP, SSM/I cloud product, and reanalysis datasets, J. Clim., 27, 3114-3128, 10.1175/jcli-d-13-00031.1, 2014.

Herring, S. C., Christidis, N., Hoell, A., Kossin, J. P., Schreck III, C., and Stott, P. A.: Explaining extreme events of 2016 from a climate perspective, Bull. Am. Meteorol. Soc., 99, S1-S157, 2018.

Hersbach, H., Peubey, C., Simmons, A., Berrisford, P., Poli, P., and Dee, D.: ERA-20CM: a twentieth-century atmospheric model ensemble, Q. J. Roy. Meteorol. Soc., 141, 2350-2375, 10.1002/qj.2528, 2015.

Hines, K. M., Bromwich, D. H., and Marshall, G. J.: Artificial surface pressure trends in the NCEP-NCAR reanalysis over the southern ocean and Antarctica, J. Clim., 13, 3940-3952, 10.1175/1520-0442(2000)013<3940:asptit>2.0.co;2, 2000.

Hyk, W., and Stojek, Z.: Quantifying uncertainty of determination by standard additions and serial dilutions methods taking into account standard uncertainties in both axes, Anal. Chem., 85, 5933-5939, 2013.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., and Woollen, J.: The NCEP/NCAR 40-year reanalysis project, Bull. Am. Meteorol. Soc., 77, 437-471, 1996.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP–DOE AMIP-II Reanalysis (R-2), Bull. Am. Meteorol. Soc., 83, 1631-1643, 10.1175/BAMS-83-11-1631, 2002.

Kidston, J., Frierson, D. M. W., Renwick, J. A., and Vallis, G. K.: Observations, simulations, and dynamics of jet stream variability and annular modes, J. Clim., 23, 6186-6199, 10.1175/2010jcli3235.1, 2010.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 reanalysis: general specifications and basic characteristics, J. Meteorol. Soc. Jpn., 93, 5-48, 10.2151/jmsj.2015-001, 2015.

Lahoz, W. A., and Schneider, P.: Data assimilation: making sense of Earth Observation, Front.Environ.Sci., 2, 1-28, 10.3389/fenvs.2014.00016, 2014.

Laloyaux, P., Balmaseda, M., Dee, D., Mogensen, K., and Janssen, P.: A coupled data assimilation system for climate reanalysis, Q. J. Roy. Meteorol. Soc., 142, 65-78, 10.1002/qj.2629, 2016.

Li, H. B., Robock, A., Liu, S. X., Mo, X. G., and Viterbo, P.: Evaluation of reanalysis soil moisture simulations using updated Chinese soil moisture observations, J. Hydrometeorol., 6, 180-193, 10.1175/jhm416.1, 2005.

Li, Q., Zhang, L., Xu, W., Zhou, T., Wang, J., Zhai, P., and Jones, P.: Comparisons of Time Series of Annual Mean Surface Air Temperature for China since the 1900s: Observations, Model Simulations, and Extended Reanalysis, Bull. Am. Meteorol. Soc., 98, 699-711, 10.1175/bams-d-16-0092.1, 2017.

Li, Y., Zeng, Z. Z., Zhao, L., and Piao, S. L.: Spatial patterns of climatological temperature lapse rate in mainland China: A multi-time scale investigation, J Geophys Res-Atmos, 120, 2661-2675, Doi 10.1002/2014jd022978, 2015.

Lin, R., Zhou, T., and Qian, Y.: Evaluation of global monsoon precipitation changes based on five reanalysis datasets, J. Clim., 27, 1271-1289, doi:10.1175/JCLI-D-13-00215.1, 2014.

Lindsay, R., Wensnahan, M., Schweiger, A., and Zhang, J.: Evaluation of seven different atmospheric reanalysis products in the Arctic, J. Clim., 27, 2588-2606, 10.1175/jcli-d-13-00014.1, 2014.

Ma, L., Zhang, T., Li, Q., Frauenfeld, O. W., and Qin, D.: Evaluation of ERA-40, NCEP-1, and NCEP-2 reanalysis air temperatures with ground-based measurements in China, J. Geophys. Res. D Atmos., 113, D15115, 10.1029/2007JD009549, 2008.

Mao, J., Shi, X., Ma, L., Kaiser, D. P., Li, Q., and Thornton, P. E.: Assessment of reanalysis daily extreme temperatures with china's homogenized historical dataset during 1979-2001 using probability density functions, J. Clim., 23, 6605-6623, 10.1175/2010jcli3581.1, 2010.

Mitas, C. M., and Clement, A.: Recent behavior of the Hadley cell and tropical thermodynamics in climate models and reanalyses, Geophys. Res. Lett., 33, L01810, 10.1029/2005gl024406, 2006.

Nguyen, H., Evans, A., Lucas, C., Smith, I., and Timbal, B.: The Hadley circulation in reanalyses: climatology, variability, and change, J. Clim., 26, 3357-3376, 10.1175/jcli-d-12-00224.1, 2013.

Niznik, M. J., and Lintner, B. R.: Circulation, moisture, and precipitation relationships along the south Pacific convergence zone in reanalyses and CMIP5 models, J. Clim., 26, 10174-10192, 10.1175/jcli-d-13-00263.1, 2013.

Onogi, K., Tslttsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kaalhori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 reanalysis, J. Meteorol. Soc. Jpn., 85, 369-432, Doi 10.2151/Jmsj.85.369, 2007.

Parker, W. S.: Reanalyses and observations: What's the difference?, Bull. Am. Meteorol. Soc., 97, 1565-1572, 10.1175/bams-d-14-00226.1, 2016.

Peña, M., and Toth, Z.: Estimation of analysis and forecast error variances, Tellus A, 66, 21767, 2014.

Piao, S. L., Yin, L., Wang, X. H., Ciais, P., Peng, S. S., Shen, Z. H., and Seneviratne, S. I.: Summer soil moisture regulated by precipitation frequency in China, Environ. Res. Lett., 4, 044012, 10.1088/1748-9326/4/4/044012, 2009.

Pitman, A. J., and Perkins, S. E.: Global and regional comparison of daily 2-m and 1000-hpa maximum and minimum temperatures in three global reanalyses, J. Clim., 22, 4667-4681, 10.1175/2009jcli2799.1, 2009.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An atmospheric reanalysis of the twentieth century, J. Clim., 29, 4083-4097, 10.1175/JCLI-D-15-0556.1, 2016.

Reed, B. C.: Linear least‐squares fits with errors in both coordinates, Am. J. Phys., 57, 642-646, 1989.

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Liu, Q., Mahanama, S. P. P., and Touré, A.: Assessment and enhancement of MERRA land surface hydrology estimates, J. Clim., 24, 6322-6338, 10.1175/JCLI-D-10-05033.1, 2011.

Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P. P., and Partyka, G. S.: Land surface precipitation in MERRA-2, J. Clim., 30, 1643-1664, 10.1175/jcli-d-16-0570.1, 2017.

Reuten, C., Moore, R. D., and Clarke, G. K. C.: Quantifying differences between 2-m temperature observations and reanalysis pressure-level temperatures in northwestern North America, J. Appl. Meteor. Climatol., 50, 916-929, 10.1175/2010jamc2498.1, 2011.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Clim., 24, 3624-3648, 10.1175/JCLI-D-11-00015.1, 2011.

Ryan, T. P.: Modern regression methods, John Wiley & Sons, 2008.

Saha, S., Moorthi, S., Pan, H. L., Wu, X. R., Wang, J. D., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H. X., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y. T., Chuang, H. Y., Juang, H. M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H. L., Yang, R. Q., Lord, S., Van den Dool, H., Kumar, A., Wang, W. Q., Long, C., Chelliah, M., Xue, Y., Huang, B. Y., Schemm, J. K., Ebisuzaki, W., Lin, R., Xie, P. P., Chen, M. Y., Zhou, S. T., Higgins, W., Zou, C. Z., Liu, Q. H., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, Bull. Am. Meteorol. Soc., 91, 1015-1057, 10.1175/2010BAMS3001.1, 2010.

Schoeberl, M. R., Dessler, A. E., and Wang, T.: Simulation of stratospheric water vapor and trends using three reanalyses, Atmos. Chem. Phys., 12, 6475-6487, 10.5194/acp-12-6475-2012, 2012.

Shen, M., Piao, S., Jeong, S.-J., Zhou, L., Zeng, Z., Ciais, P., Chen, D., Huang, M., Jin, C.-S., and Li, L. Z.: Evaporative cooling over the Tibetan Plateau induced by vegetation growth, Proc. Nat. Acad. Sci. U.S.A., 112, 9299-9304, 2015.

Simmonds, I., and Keay, K.: Mean Southern Hemisphere extratropical cyclone behavior in the 40-year NCEP-NCAR reanalysis, J. Clim., 13, 873-885, 10.1175/1520-0442(2000)013<0873:mshecb>2.0.co;2, 2000.

Simmons, A. J., Willett, K. M., Jones, P. D., Thorne, P. W., and Dee, D. P.: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets, J. Geophys. Res. D Atmos., 115, D01110, 10.1029/2009JD012442, 2010.

Siswanto, S., Oldenborgh, G. J., Schrier, G., Jilderda, R., and Hurk, B.: Temperature, extreme precipitation, and diurnal rainfall changes in the urbanized Jakarta city during the past 130 years, Int. J. Climatol., 36, 3207-3225, 2015.

Spracklen, D. V., Arnold, S. R., and Taylor, C. M.: Observations of increased tropical rainfall preceded by air passage over forests, Nature, 494, 390-390, 10.1038/nature11904, 2013.

Stott, P.: How climate change affects extreme weather events, Science, 352, 1517-1518, 10.1126/science.aaf7271, 2016.

Tang, W.-J., Yang, K., Qin, J., Cheng, C., and He, J.: Solar radiation trend across China in recent decades: a revisit with quality-controlled data, Atmos. Chem. Phys., 11, 393-406, 2011.

Tellinghuisen, J.: Least-squares analysis of data with uncertainty in x and y: A Monte Carlo methods comparison, Chemom. Intell. Lab. Syst., 103, 160-169, 2010.

Thorne, P., and Vose, R.: Reanalyses suitable for characterizing long-term trends: Are they really achievable?, Bull. Am. Meteorol. Soc., 91, 353-361, 2010.

Trenberth, K. E., and Olson, J. G.: An evaluation and intercomparison of global analyses from the National-Meteorological-Center and the European-Centre-for-Medium-Range-Weather-Forecasts, Bull. Am. Meteorol. Soc., 69, 1047-1057, Doi 10.1175/1520-0477(1988)069<1047:Aeaiog>2.0.Co;2, 1988.

Trenberth, K. E., Koike, T., and Onogi, K.: Progress and prospects for reanalysis for weather and climate, Eos Trans. Am. Geophys. Union, 89, 234-235, 10.1029/2008EO260002, 2008.

Trenberth, K. E., Fasullo, J. T., and Shepherd, T. G.: Attribution of climate extreme events, Nature Clim. Change, 5, 725-730, 10.1038/nclimate2657, 2015.

Trigo, I., Boussetta, S., Viterbo, P., Balsamo, G., Beljaars, A., and Sandu, I.: Comparison of model land skin temperature with remotely sensed estimates and assessment of surface - atmosphere coupling, J. Geophys. Res. D Atmos., 120, D023812, 10.1002/2015JD023812, 2015.

Tsidu, G. M.: High-resolution monthly rainfall database for Ethiopia: Homogenization, reconstruction, and gridding, J. Clim., 25, 8422-8443, 10.1175/jcli-d-12-00027.1, 2012.

Uppala, S. M., KÅllberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q. J. Roy. Meteorol. Soc., 131, 2961-3012, 10.1256/qj.04.176, 2005.

Venema, V., Mestre, O., Aguilar, E., Auer, I., Guijarro, J., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., and Zahradnicek, P.: Benchmarking homogenization algorithms for monthly data, Clim. Past, 8, 89-115, 2012.

Wang, A., and Zeng, X.: Evaluation of multireanalysis products with in situ observations over the Tibetan Plateau, J. Geophys. Res. D Atmos., 117, D05102, 10.1029/2011JD016553, 2012.

Wang, K., and Liang, S.: Global atmospheric downward longwave radiation over land surface under all‐sky conditions from 1973 to 2008, J. Geophys. Res. D Atmos., 114, D19101, 2009.

Wang, K., Dickinson, R., Wild, M., and Liang, S.: Atmospheric impacts on climatic variability of surface incident solar radiation, Atmos. Chem. Phys., 12, 9581-9592, 2012.

Wang, K., and Dickinson, R. E.: Global atmospheric downward longwave radiation at the surface from ground-based observations, satellite retrievals, and reanalyses, Rev. Geophys., 51, 150-185, 10.1002/rog.20009, 2013.

Wang, K.: Measurement biases explain discrepancies between the observed and simulated decadal variability of surface incident solar radiation, Sci. Rep., 4, 6144, 10.1038/srep06144, 2014.

Wang, K. C., Ma, Q., Li, Z. J., and Wang, J. K.: Decadal variability of surface incident solar radiation over China: Observations, satellite retrievals, and reanalyses, J Geophys Res-Atmos, 120, 6500-6514, 10.1002/2015JD023420, 2015.

Wang, X., and Wang, K.: Homogenized variability of radiosonde-derived atmospheric boundary layer height over the global land surface from 1973 to 2014, J. Clim., 29, 6893-6908, 10.1175/JCLI-D-15-0766.1, 2016.

Wang, X. L., Wen, Q. H., and Wu, Y.: Penalized maximal t test for detecting undocumented mean change in climate data series, J. Appl. Meteor. Climatol., 46, 916-931, 2007.

Wang, X. L.: Penalized maximal F test for detecting undocumented mean shift without trend change, J. Atmos. Oceanic Technol., 25, 368-384, 2008.

Wang, X. L., Chen, H., Wu, Y., Feng, Y., and Pu, Q.: New techniques for the detection and adjustment of shifts in daily precipitation data series, J. Appl. Meteor. Climatol., 49, 2416-2436, 2010.

Wang, X. L., and Feng, Y.: RHtestsV4 user manual, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada. 28 pp. available at http://etccdi.pacificclimate.org/software.shtml, 2013.

Wu, C., Chen, J. M., Pumpanen, J., Cescatti, A., Marcolla, B., Blanken, P. D., Ardö, J., Tang, Y., Magliulo, V., and Georgiadis, T.: An underestimated role of precipitation frequency in regulating summer soil moisture, Environ. Res. Lett., 7, 024011, 2012.

Xu, J., and Powell, A. M., Jr.: Uncertainty of the stratospheric/tropospheric temperature trends in 1979-2008: multiple satellite MSU, radiosonde, and reanalysis datasets, Atmos. Chem. Phys., 11, 10727-10732, 10.5194/acp-11-10727-2011, 2011.

Yang, K., Koike, T., and Ye, B.: Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets, Agr. Forest Meteorol., 137, 43-55, 10.1016/j.agrformet.2006.02.001, 2006.

York, D., Evensen, N. M., Martínez, M. L., and Delgado, J. D. B.: Unified equations for the slope, intercept, and standard errors of the best straight line, Am. J. Phys., 72, 367-375, 10.1119/1.1632486, 2004.

You, Q., Kang, S., Pepin, N., Flügel, W.-A., Yan, Y., Behrawan, H., and Huang, J.: Relationship between temperature trend magnitude, elevation and mean temperature in the Tibetan Plateau from homogenized surface stations and reanalysis data, Global Planet. Change, 71, 124-133, 10.1016/j.gloplacha.2010.01.020, 2010.

Zeng, Z., Piao, S., Li, L. Z., Zhou, L., Ciais, P., Wang, T., Li, Y., Lian, X., Wood, E. F., and Friedlingstein, P.: Climate mitigation from vegetation biophysical feedbacks during the past three decades, Nature Clim. Change, 7, 432-436, 2017.

Zhao, T., Guo, W., and Fu, C.: Calibrating and evaluating reanalysis surface temperature error by topographic correction, J. Clim., 21, 1440-1446, 10.1175/2007jcli1463.1, 2008.

Zhou, C., and Wang, K.: Land surface temperature over global deserts: means, variability and trends, J. Geophys. Res. D Atmos., 121, 2016JD025410, 10.1002/2016JD025410, 2016a.

Zhou, C., and Wang, K.: Evaluation of surface fluxes in ERA-Interim using flux tower data, J. Clim., 29, 1573-1582, 10.1175/JCLI-D-15-0523.1, 2016b.

Zhou, C., and Wang, K.: Biological and environmental controls on evaporative fractions at ameriflux sites, J. Appl. Meteor. Climatol., 55, 145-161, 10.1175/JAMC-D-15-0126.1, 2016c.

Zhou, C., and Wang, K.: Spatiotemporal divergence of the warming hiatus over land based on different definitions of mean temperature, Sci. Rep., 6, 31789, 10.1038/srep31789, 2016d.

Zhou, C., and Wang, K.: Contrasting daytime and nighttime precipitation variability between observations and eight reanalysis products from 1979 to 2014 in China, J. Clim., 30, 6443-6464, 10.1175/JCLI-D-16-0702.1, 2017a.

Zhou, C., and Wang, K.: Quantifying the sensitivity of precipitation to the long-term warming trend and interannual-decadal variation of surface air temperature over China, J. Clim., 30, 3687-3703, 10.1175/jcli-d-16-0515.1, 2017b.

Zhou, C., Wang, K., and Ma, Q.: Evaluation of eight current reanalyses in simulating land surface temperature from 1979 to 2003 in China, J. Clim., 30, 7379-7398, 10.1175/jcli-d-16-0903.1, 2017.

Zhou, C., Wang, K., and Qi, D.: Attribution of the July 2016 extreme precipitation event over China's Wuhan, Bull. Am. Meteorol. Soc., 99, S107-S112, 2018.

Table 1. Summary information on the twelve reanalysis products, including institution, model resolution, assimilation system, surface observations included associated with surface air temperatures, sea ice and sea surface temperatures (SSTs) and greenhouse gas (GHG) boundary conditions. The number in the parentheses in the Model Name column is the year of the version of the forecast model used. More details on each product can be found in the associated reference.

Reanalysis	Institution	Model Name	Model Resolution	Period	Assimilation System
ERA-Interim	ECMWF	IFS version Cy31r2 (2007)	T255 ~80 km, 60 levels	1979 onwards	4D-VAR
JRA-55	JMA	JMA operational numerical weather prediction system (2009)	T319 ~55 km, 60 levels	1958-2013	4D-VAR
NCEP-R1	NCEP/NCAR	NCEP operational numerical weather prediction system (1995)	T62 ~210 km, 28 levels	1948 onwards	3D-VAR
NCEP-R2	NCEP/DOE	Modified NCEP-R1 model (1998)	T62 ~210 km, 28 levels	1979 onwards	3D-VAR
MERRA	NASA/GMAO	GEOS-5.0.2 atmospheric general circulation model (2008)	0.5°× 0.667° ~55 km, 72 levels	1979 onwards	3D-VAR with incremental updating (GEOS IAU)
MERRA-2	NASA/GMAO	Updated version of GEOS-5.12.4 used in MERRA; its land model is similar to that of MERRA (2015)	0.5°× 0.625° ~55 km, 72 levels	1980 onwards	3D-VAR with incremental updating (GEOS IAU)
ERA-20C	ECMWF	IFS version Cy38r1 (2012), coupled atmosphere-land-ocean-waves system	T159 ~125 km, 91 levels	1900-2010	4D-VAR
ERA-20CM	ECMWF	Similar to that used in ERA-20C (2012)	T159 ~125 km, 91 levels	1900-2010	-
CERA-20C	ECMWF	IFS version Cy41r2 (2016), coupled atmosphere-ocean-land-waves-sea ice system	T159 ~125 km, 91 levels	1901-2010	CERA ensemble assimilation technique
NOAA 20CRv2c	NOAA/ESRL PSD	NCEP GFS (2008), an updated version of the NCEP Climate Forecast System (CFS) coupled atmosphere-land model	T62 ~210 km, 28 levels	1851-2014	Ensemble Kalman filter
NOAA 20CRv2	NOAA/ESRL PSD	Same model as NOAA 20CRv2c (2008)	T62 ~210 km, 28 levels	1871-2012	Ensemble Kalman filter
CFSR	NCEP	NCEP CFS (2011) coupled atmosphere-ocean-land-sea ice model	T382 ~38 km, 64 levels	1979-2010	3D-VAR

Table 1. Continued from right column.

Related Assimilated and Analysed Observations	Sea Ice and SSTs	GHG Forcing	Reference
1) Includes *in situ* observations of near-surface air temperature/pressure/relative humidity 2) Assimilates upper-air temperatures/wind/specific humidity 3) Assimilates rain-affected SSM/I radiances	A changing suite of SST and sea ice data from observations and NCEP	Interpolation by 1.6 ppmv/year from the global mean CO_2 in 1990 of 353 ppmv	(Dee et al., 2011b)
1) Analyses available near-surface observations 2) Assimilates all available traditional and satellite observations	*In situ* observation-based estimates of the COBE SST data and sea ice	Same as CMIP5	(Kobayashi et al., 2015)
1) Initiated with weather observations from ships, planes, station data, satellite observations and many more sources 2) No inclusion of near-surface air temperatures 3) Uses observed precipitation to nudge soil moisture 4) No information on aerosols	Reynolds SSTs for 1982 on and the UKMO GISST data for earlier periods; sea ice from SMMR/SSMI	Constant global mean CO_2 of 330 ppmv; no other trace gases	(Kalnay et al., 1996)
1) No inclusion of near-surface air temperatures 2) No information on aerosols	AMIP-II prescribed	Constant global mean CO_2, 350 ppmv; no other trace gases	(Kanamitsu et al., 2002)
1) Neither MERRA nor MERRA-2 analyse near-surface air temperature, relative humidity, or other variables 2) Radiosondes do provide some low-level observations	Reynolds SSTs prescribed	Same as CMIP5	(Rienecker et al., 2011)
1) Includes newer observations (not included in MERRA) after the 2010s 2) Includes aerosols from MODIS and AERONET measurements over land after the 2000s and from the GOCART model before the 2000s 3) Assimilates observation-corrected precipitation to correct the model-generated precipitation before reaching the land surface	AMIP-II and Reynolds SSTs	Same as CMIP5	(Reichle et al., 2017)
1) Assimilates surface pressures from ISPDv3.2.6 and ICOADSv2.5.1 and surface marine winds from ICOADSv2.5.1 2) Uses monthly climatology of aerosols from CMIP5	SSTs and sea ice from HadISST2.1.0.0	Same as CMIP5	(Poli et al., 2016)
Assimilates no data and includes radiative forcings from CMIP5	SSTs and sea ice realizations from HadISST2.1.0.0 used in 10 members	Same as CMIP5	(Hersbach et al., 2015)
1) Assimilates surface pressures from ISPDv3.2.6 and ICOADSv2.5.1 and surface marine winds from ICOADSv2.5.1 2) Assimilates no data in the land, wave and sea ice components but uses the coupled model at each time step	SSTs from HadISST2.1.0.0	Same as CMIP5	(Laloyaux et al., 2016)
Assimilates only surface pressure and sea level pressure	SSTs from HadISST1.1 and sea ice from COBE SST	Monthly 15 °gridded estimates of CO_2 from WMO observations	(Compo et al., 2011)
Same as NOAA 20CRv2c	SSTs and sea ice from HadISST1.1	Monthly 15 °gridded estimates of CO_2 from WMO observations	(Compo et al., 2011)
1) Assimilates all available conventional and satellite observations but not near-surface air temperatures 2) Atmospheric model contains observed changes in aerosols 3) Uses observation-corrected precipitation to force the land surface analysis	Generated by coupled ocean-sea ice models; evolves freely during the 6-h coupled model integration	Monthly 15 °gridded estimates of CO_2 from WMO observations	(Saha et al., 2010)

Table 2. Differences (unit: ℃) relative to the homogenized observations and trends (unit: ℃/decade) in surface air temperatures (T_a) from 1979 to 2010 over China and its seven subregions. The bold and italic bold fonts indicate results that are significant according to two-tailed Student's t-tests with significance levels of 0.05 and 0.1, respectively.

Region	China		Tibetan Plateau		Northwest China		Loess Plateau		Middle China		Northeast China		North China Plain		Southeast China	
	Diff.	Trend	Diff.	Trend	Diff.	Trend	Diff.	Trend	Diff.	Trend	Diff.	Trend	Diff.	Trend	Diff.	Trend
ERA-Interim	-0.87	**0.38**	-3.49	**0.33**	-1.82	**0.37**	-0.32	**0.50**	-1.19	**0.28**	-0.03	**0.42**	-0.02	**0.45**	-0.03	**0.37**
NCEP-R1	-2.56	**0.23**	-6.80	0.11	-4.45	**0.39**	-1.77	**0.21**	-2.91	**0.23**	-1.28	**0.27**	-1.21	**0.23**	-1.33	**0.22**
MERRA	-0.48	**0.25**	-3.48	**0.33**	0.95	0.14	1.14	0.09	-1.35	**0.12**	-0.22	**0.52**	0.67	**0.26**	-0.27	**0.24**
JRA-55	-1.10	**0.38**	-3.49	**0.42**	-1.70	**0.39**	-0.58	**0.52**	-1.61	**0.30**	-0.25	**0.37**	-0.26	**0.41**	-0.50	**0.34**
NCEP-R2	-2.10	**0.25**	-5.76	-0.07	-4.29	**0.58**	-1.33	0.10	-2.80	**0.20**	-0.51	**0.36**	-0.38	**0.23**	-1.14	**0.36**
MERRA2	-0.91	**0.28**	-3.41	**0.35**	0.34	**0.32**	0.12	**0.19**	-1.35	**0.23**	-0.73	**0.41**	-0.24	0.18	-0.64	**0.25**
ERA-20C	-1.42	**0.29**	-6.56	**0.33**	-1.95	**0.31**	0.03	0.21	-2.01	**0.35**	-0.19	**0.32**	1.05	0.19	-0.47	**0.28**
ERA-20CM	-1.48	**0.32**	-5.93	**0.28**	-1.39	**0.38**	-0.36	**0.33**	-2.13	**0.27**	-0.23	**0.41**	-0.31	**0.34**	-0.51	**0.29**
CERA-20C	-2.06	**0.34**	-7.00	**0.41**	-2.15	**0.38**	-0.78	**0.36**	-2.59	**0.34**	-0.76	**0.43**	-0.40	0.19	-1.20	**0.29**
NOAA 20CRv2c	-0.28	**0.22**	-2.75	**0.39**	-0.01	**0.28**	1.62	0.16	-1.68	**0.18**	-0.16	0.11	1.06	0.15	0.18	**0.22**
NOAA 20CRv2	-0.32	**0.24**	-2.78	**0.33**	-0.01	**0.29**	1.48	0.20	-1.77	**0.19**	-0.07	0.25	0.97	0.21	0.12	**0.19**
CFSR	-1.74	**0.48**	-5.09	**0.46**	-1.03	**0.44**	-0.25	**0.40**	-2.91	**0.37**	-0.49	**0.67**	-0.37	**0.47**	-1.58	**0.51**
Obs-raw	0.03	**0.40**	0.03	**0.46**	0.09	**0.44**	0.01	**0.52**	0.05	**0.30**	0.00	**0.40**	0.05	**0.42**	0.03	**0.36**
Obs-homogenized		**0.37**		**0.44**		**0.36**		**0.50**		**0.24**		**0.41**		**0.38**		**0.33**

Table 3. Spatial pattern correlation (unit: 1) of three groups: partial relationships, trends and simulated biases in the trends in surface air temperature (T_a) against surface incident solar radiation (R_s), precipitation frequency (PF) and surface downward longwave radiation (L_d). The bold and italic bold fonts indicate results that are significant according to two-tailed Student's t-tests with significance levels of 0.05 and 0.1, respectively.

Pattern Correlation	Partial Relationship						Trend				Trend Bias		
	(T_a, R_s)		(T_a, PF)		(T_a, L_d)		(T_a, T_a)	(T_a, R_s)	(T_a, PF)	(T_a, L_d)	(T_a, R_s)	(T_a, PF)	(T_a, L_d)
	Corr.	Slope	Corr.	Slope	Corr.	Slope							
ERA-Interim	**0.29**	0.01	0.03	**0.31**	**0.21**	**0.25**	**0.47**	**-0.11**	-0.04	**0.33**	**0.26**	**-0.12**	**0.10**
NCEP-R1	**0.30**	***0.06***	**0.18**	**0.30**	**0.36**	0.00	0.02	**-0.36**	-0.02	**0.62**	-0.03	-0.04	**0.43**
MERRA	**0.29**	0.06	**0.13**	**0.39**	0.05	**0.20**	**0.21**	**0.66**	**-0.81**	**-0.53**	**0.42**	**-0.62**	-0.05
JRA-55	**0.35**	**0.21**	**0.22**	**0.16**	**0.29**	**0.27**	**0.54**	**-0.33**	**0.31**	**0.57**	0.00	**0.14**	**0.29**
NCEP-R2	**0.22**	0.03	**0.20**	**0.36**	**0.27**	0.04	**-0.08**	**0.18**	**-0.29**	**0.28**	**0.15**	**-0.14**	**0.35**
MERRA2	**0.13**	0.05	**0.26**	**0.43**	**0.09**	**0.30**	**0.22**	**0.30**	**-0.11**	**0.11**	-0.02	**-0.12**	**0.28**
ERA-20C	**0.28**	***-0.07***	***-0.07***	**0.43**	**0.19**	0.02	-0.07	**0.18**	**-0.33**	0.03	**0.11**	**-0.25**	**0.31**
ERA-20CM	**0.24**	-0.04	-0.03	**0.32**	**0.26**	**0.18**	**0.28**	**-0.32**	**0.31**	**0.83**	-0.02	**0.12**	**0.34**
CERA-20C	**0.41**	**0.17**	**0.10**	**0.37**	**0.08**	***0.07***	**0.29**	**0.50**	**-0.58**	**-0.07**	-0.01	**-0.22**	**0.23**
NOAA 20CRv2c	**0.39**	**0.15**	**-0.22**	**0.25**	**0.14**	**0.15**	**0.08**	***-0.07***	**-0.11**	**0.55**	**-0.25**	***-0.05***	**0.50**
NOAA 20CRv2	**0.38**	**0.15**	**-0.21**	**0.18**	**0.14**	**0.23**	**0.19**	-0.02	**-0.20**	**0.56**	**-0.18**	**0.11**	**0.47**
CFSR	**0.33**	**0.12**	**0.10**	**0.19**	**0.37**	**0.21**	**0.19**	**0.11**	**-0.26**	**0.07**	**0.31**	**-0.08**	**0.15**
Obs-raw								***-0.07***	**0.27**	**0.50**			
Obs-homogenized								**-0.09**	**0.35**	**0.32**			

Figure Captions:

Figure 1. The multiyear-averaged differences in surface air temperatures (T_a, unit: ℃) during the period of 1979-2010 from the twelve reanalysis products relative to the homogenized observations over China. The reanalysis products are (**a**) ERA-Interim, (**b**) NCEP-R1, (**c**) MERRA, (**d**) JRA-55, (**e**) NCEP-R2, (**f**) MERRA2, (**g**) ERA-20C, (**h**) ERA-20CM, (**i**) CERA-20C, (**j**) NOAA 20CRv2c, (**k**) NOAA 20CRv2 and (**l**) CFSR. The mainland of China is divided into seven regions (shown in Fig. 1c), specifically ① the Tibetan Plateau, ② Northwest China, ③ the Loess Plateau, ④ Middle China, ⑤ Northeast China, ⑥ the North China Plain and ⑦ South China.

Figure 2. The impact of inconsistencies between station and model elevations on the simulated multiyear-averaged differences in surface air temperatures (T_a, unit: ℃) during the study period of 1979-2010 over China. The elevation difference (ΔHeight) between the stations and the models consists of the filtering error in the elevations used in the spectral models (Δf) and the difference in site-to-grid elevations (Δs) due to the complexity of orographic topography. Δf is derived from the model elevations minus the 'true' elevations in the corresponding model grid cells from GTOPO30. The GTOPO30 orography is widely used in reanalyses, e.g., by ECMWF. The colour bar denotes the station elevations (unit: m). The relationship of the T_a differences is regressed on ΔHeight (shown at the bottom of each subfigure) or Δf and Δs (shown at the top of each subfigure); the corresponding explained variances are shown.

Figure 3. Taylor diagrams for annual time series of the observed and reanalysed surface air temperature anomalies (T_a, unit: ℃) from 1979 to 2010 in **(a)** China and

(b-h) the seven subregions. The correlation coefficient, standard deviation and root mean squared error (RMSE) are calculated against the observed homogenized T_a anomalies.

Figure 4. Composite map of partial correlation coefficients of the detrended surface air temperature (T_a, unit: ℃) against surface incident solar radiation (R_s), precipitation frequency (PF) and surface downward longwave radiation (L_d) during the period of 1979-2010 from observations and the twelve reanalysis products. The marker '+' denotes the negative partial correlations of T_a with R_s over the Tibetan Plateau in NCEP-R2, ERA-20C and ERA-20CM.

Figure 5. (**a**, **b**) The observed trends in surface air temperature (T_a, unit: ℃/decade) and the simulated biases in the trends in T_a (unit: ℃/decade) during the period of 1979-2010 from (**c**) raw observations and **(d-o)** the twelve reanalysis products over China with respect to the homogenized observations. The squares denote the original homogeneous time series, and the dots denote the adjusted homogeneous time series. The probability distribution functions of all of the biases in the trends are shown as coloured histograms, and the black stairs are integrated from the trend biases with a significance level of 0.05 (based on two-tailed Student's *t*-tests). The cyan and green stars in (k-n) represent estimates of the biases in the trends outside the ensemble ranges whose locations are denoted by the black dots shown in (k-n).

Figure 6. Composite map of the contributions (unit: ℃/decade) of the biases in the trends in three relevant parameters, surface incident solar radiation (R_s, in red), surface downward longwave radiation (L_d, in green) and precipitation frequency (in

blue) to the biases in the trends in surface air temperature (T_a) during the study period of 1979-2010, as estimated using the twelve reanalysis products over China.

Figure 7. Contribution s(unit: ℃/decade) of the biases in the trends in surface air temperatures (T_a) from three relevant parameters, surface incident solar radiation (R_s, in brown), surface downward longwave radiation (L_d, in light blue) and precipitation frequency (PF, in deep blue) during the study period of 1979-2010 from the twelve reanalysis products over China and its seven subregions.

Figure 8. Spatial associations of the simulated biases in the trend in surface air temperature (T_a) versus three relevant parameters among the twelve reanalysis products (solid lines indicate the NWP-like reanalyses, and dashed lines indicate the climate reanalyses). The probability density functions (unit: %) of these biases in the trends are estimated from approximately 700 1°×1° grid cells that cover China. The median values (coloured dots with error bars of spatial standard deviations) of the biases in the trends in T_a (unit: ℃/decade) in the twelve reanalyses are regressed onto those of (**a**) the surface incident solar radiation (R_s, unit: W m^{-2}/decade), (**b**) precipitation frequency (unit: days/decade) and (**c**) the surface downward longwave radiation (L_d, unit: W m^{-2}/decade) using the ordinary least squares method (OLS, denoted by the dashed grey lines) and the weighted total least squares method (WTLS, denoted by the solid black lines). The 5-95% confidence intervals of the regressed slopes obtained using WTLS are shown as shading. The regressed correlations and slopes are shown as grey and black text, respectively.

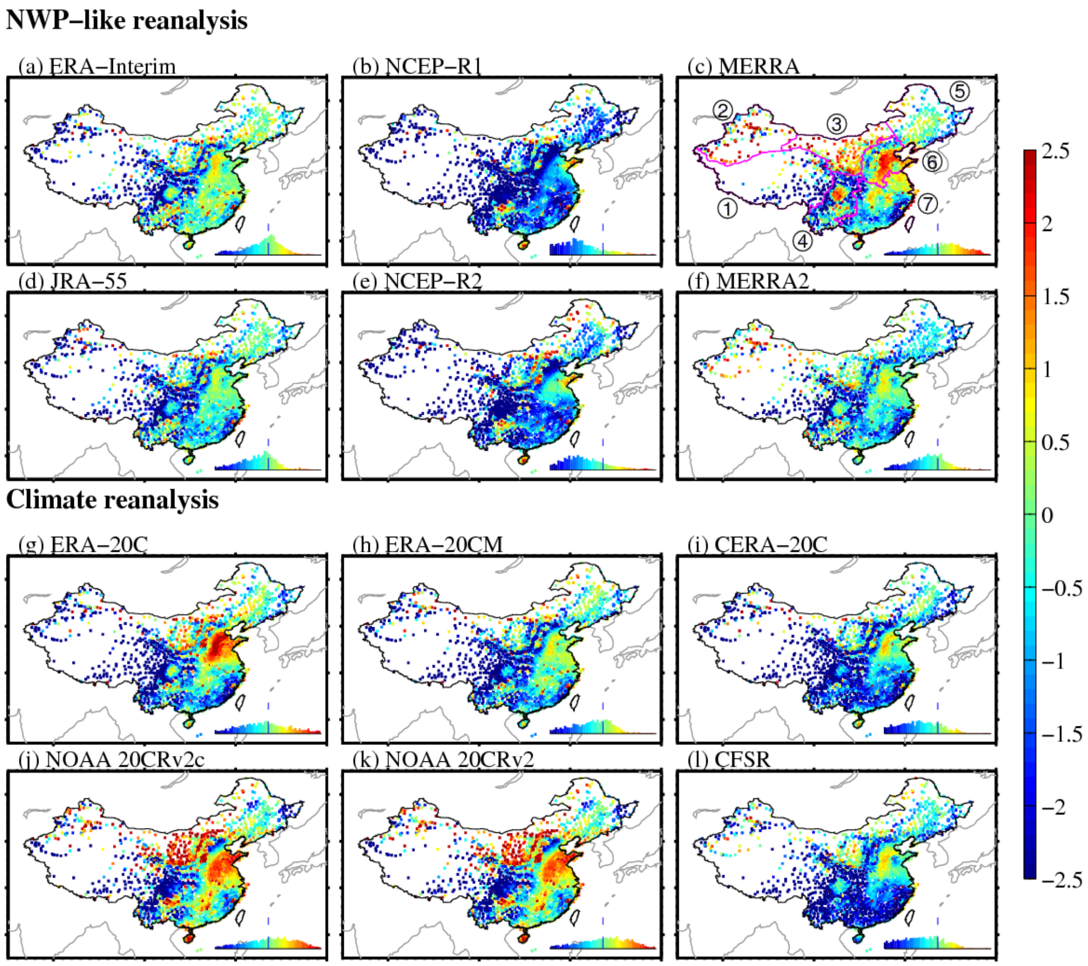

Figure 1. The multiyear-averaged differences in surface air temperatures (T_a, unit: ℃) during the period of 1979-2010 from the twelve reanalysis products relative to the homogenized observations over China. The reanalysis products are (**a**) ERA-Interim, (**b**) NCEP-R1, (**c**) MERRA, (**d**) JRA-55, (**e**) NCEP-R2, (**f**) MERRA2, (**g**) ERA-20C, (**h**) ERA-20CM, (**i**) CERA-20C, (**j**) NOAA 20CRv2c, (**k**) NOAA 20CRv2 and (**l**) CFSR. The mainland of China is divided into seven regions (shown in Fig. 1c), specifically ① the Tibetan Plateau, ② Northwest China, ③ the Loess Plateau, ④ Middle China, ⑤ Northeast China, ⑥ the North China Plain and ⑦ South China.

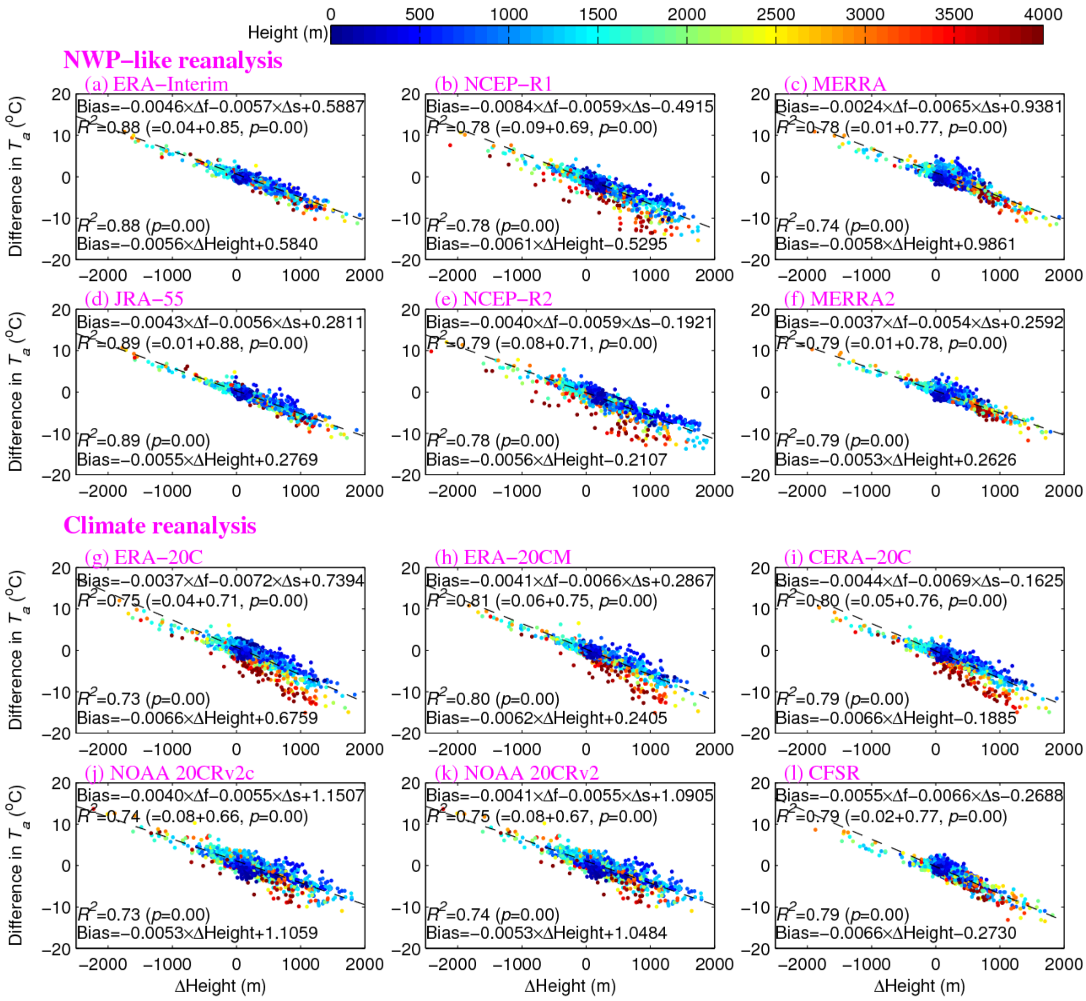

Figure 2. The impact of inconsistencies between station and model elevations on the simulated multiyear-averaged differences in surface air temperatures (T_a, unit: ℃) during the study period of 1979-2010 over China. The elevation difference (ΔHeight) between the stations and the models consists of the filtering error in the elevations used in the spectral models (Δf) and the difference in site-to-grid elevations (Δs) due to the complexity of orographic topography. Δf is derived from the model elevations minus the 'true' elevations in the corresponding model grid cells from GTOPO30. The GTOPO30 orography is widely used in reanalyses, e.g., by ECMWF. The colour bar denotes the station elevations (unit: m). The relationship of the T_a differences is regressed on ΔHeight (shown at the bottom of each subfigure) or Δf and Δs (shown at

the top of each subfigure); the corresponding explained variances are shown.

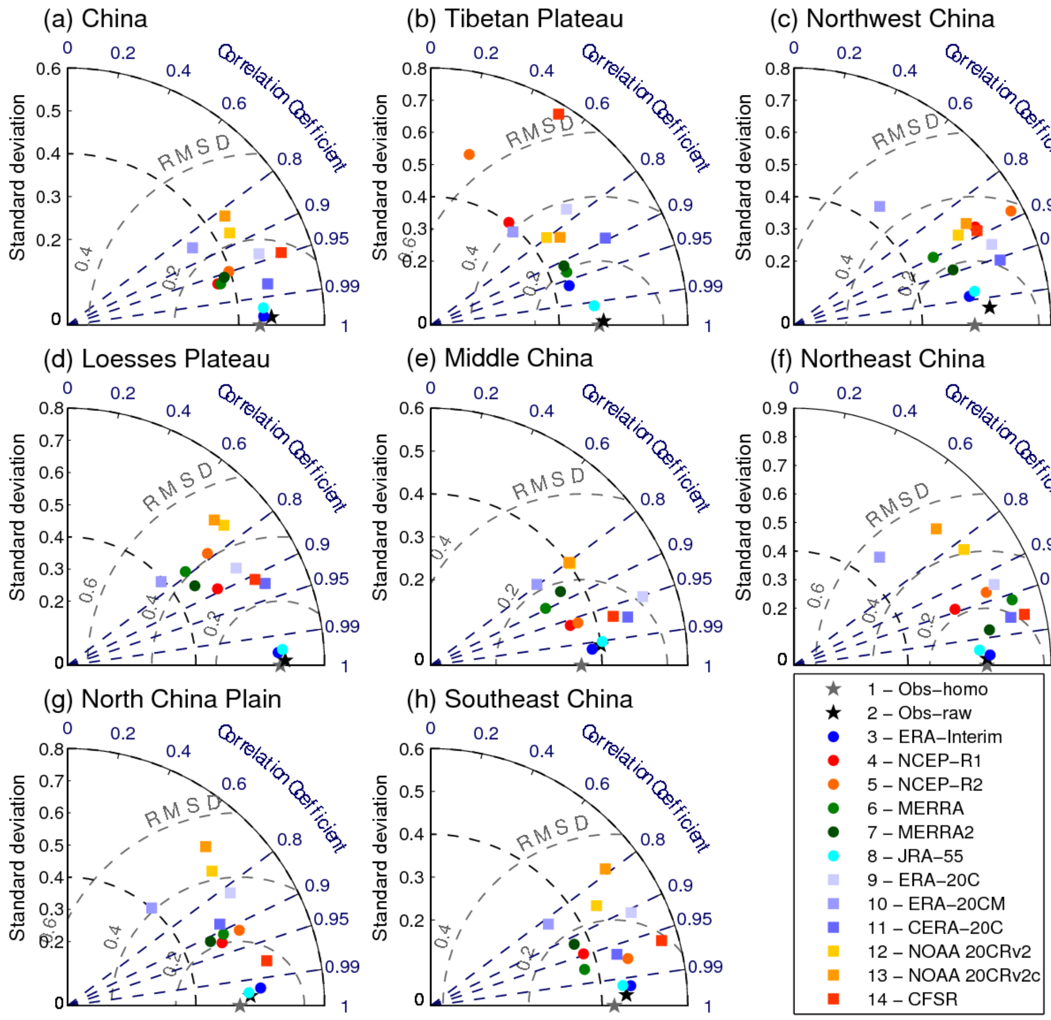

Figure 3. Taylor diagrams for annual time series of the observed and reanalysed surface air temperature anomalies (T_a, unit: ℃) from 1979 to 2010 in **(a)** China and **(b-h)** the seven subregions. The correlation coefficient, standard deviation and root mean squared error (RMSE) are calculated against the observed homogenized T_a anomalies.

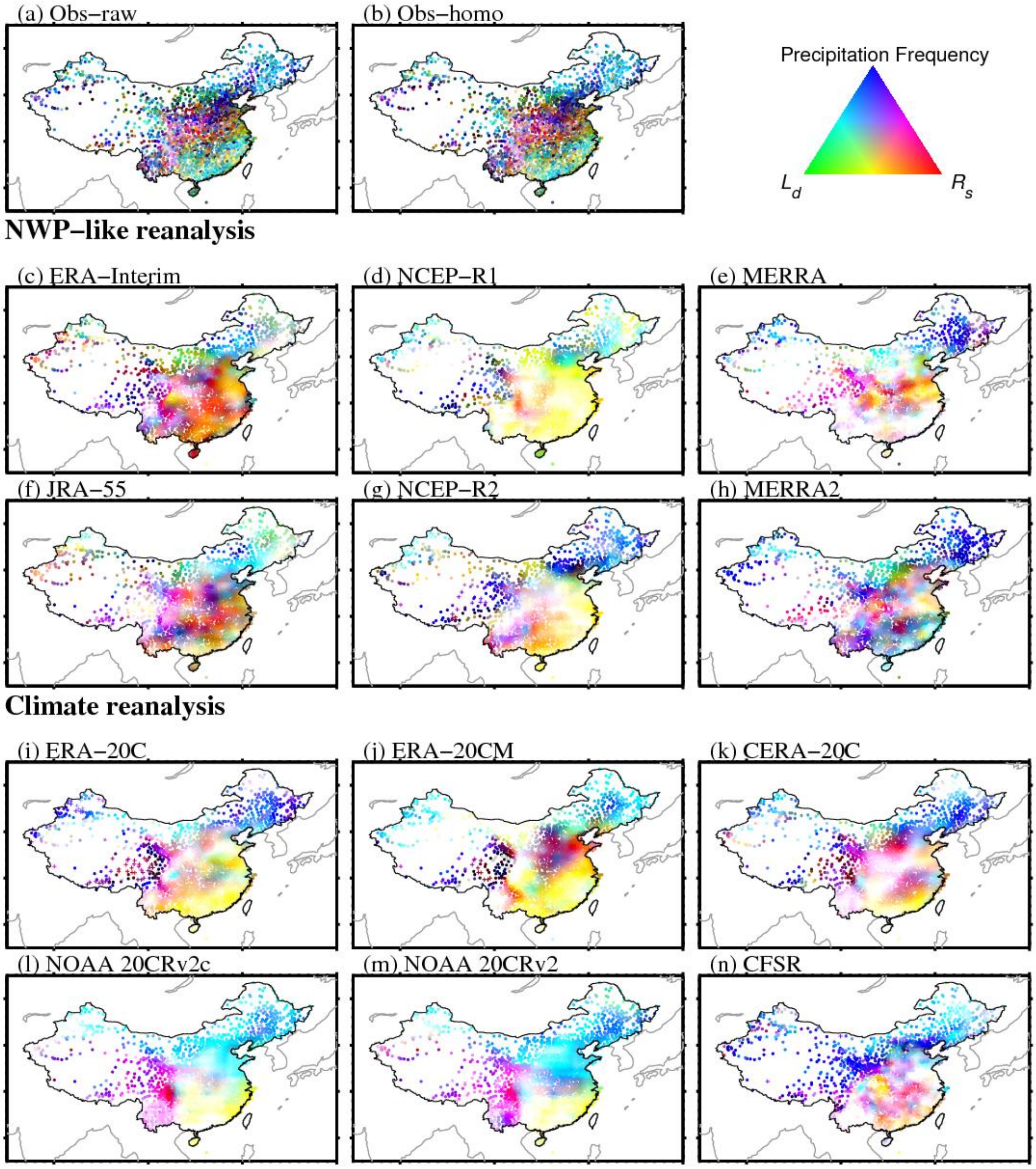

Figure 4. Composite map of partial correlation coefficients of the detrended surface air temperature (T_a, unit: ℃) against surface incident solar radiation (R_s), precipitation frequency (PF) and surface downward longwave radiation (L_d) during the period of 1979-2010 from observations and the twelve reanalysis products. The marker '+' denotes the negative partial correlations of T_a with R_s over the Tibetan Plateau in NCEP-R2, ERA-20C and ERA-20CM.

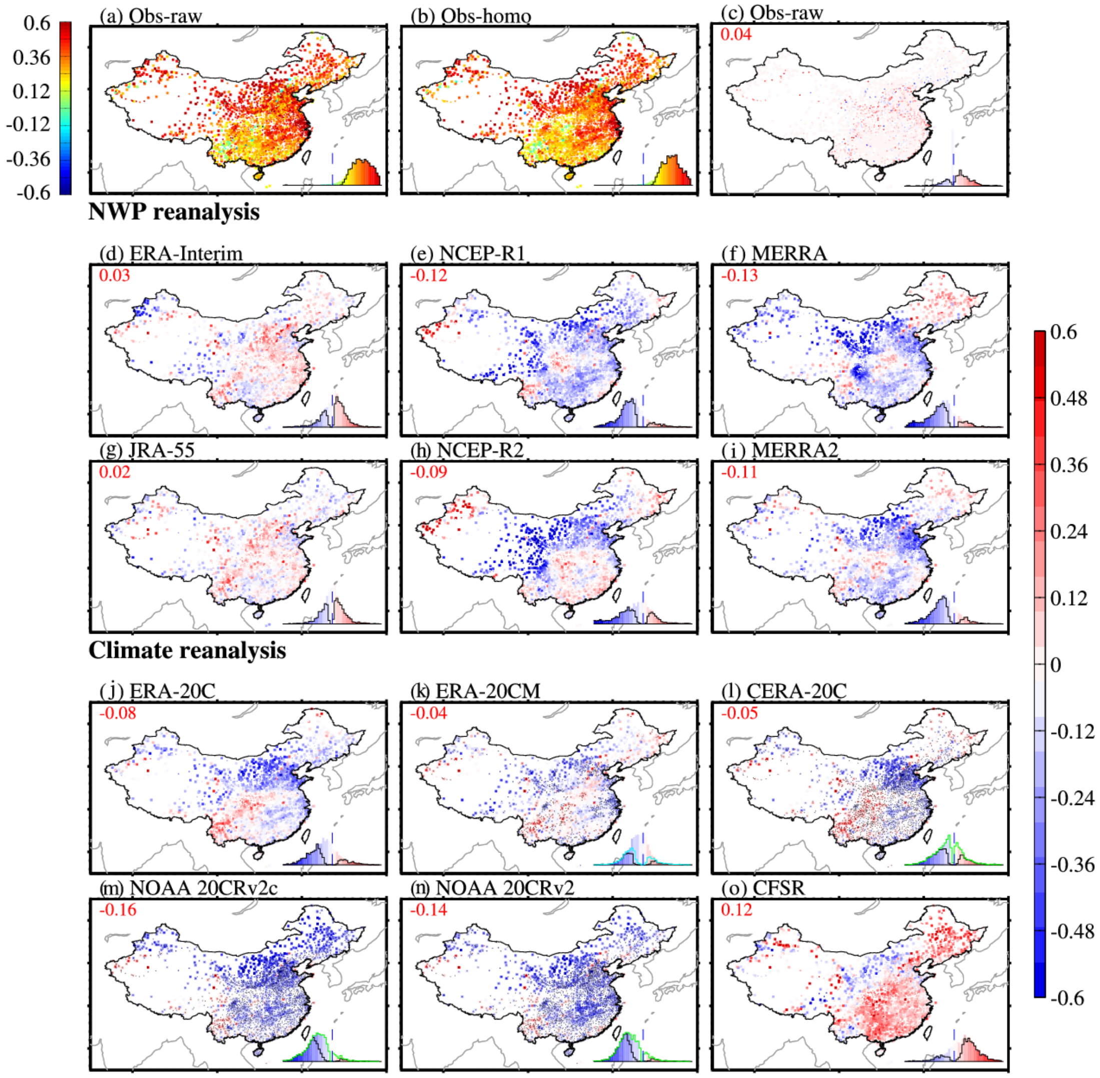

Figure 5. (**a**, **b**) The observed trends in surface air temperature (T_a, unit: ℃/decade) and the simulated biases in the trends in T_a (unit: ℃/decade) during the period of 1979-2010 from (**c**) raw observations and **(d-o)** the twelve reanalysis products over China with respect to the homogenized observations. The squares denote the original homogeneous time series, and the dots denote the adjusted homogeneous time series. The probability distribution functions of all of the biases in the trends are shown as coloured histograms, and the black stairs are integrated from the trend biases with a significance level of 0.05 (based on two-tailed Student's *t*-tests). The cyan and green stars in (k-n) represent estimates of the biases in the trends outside the ensemble

ranges whose locations are denoted by the black dots shown in (k-n).

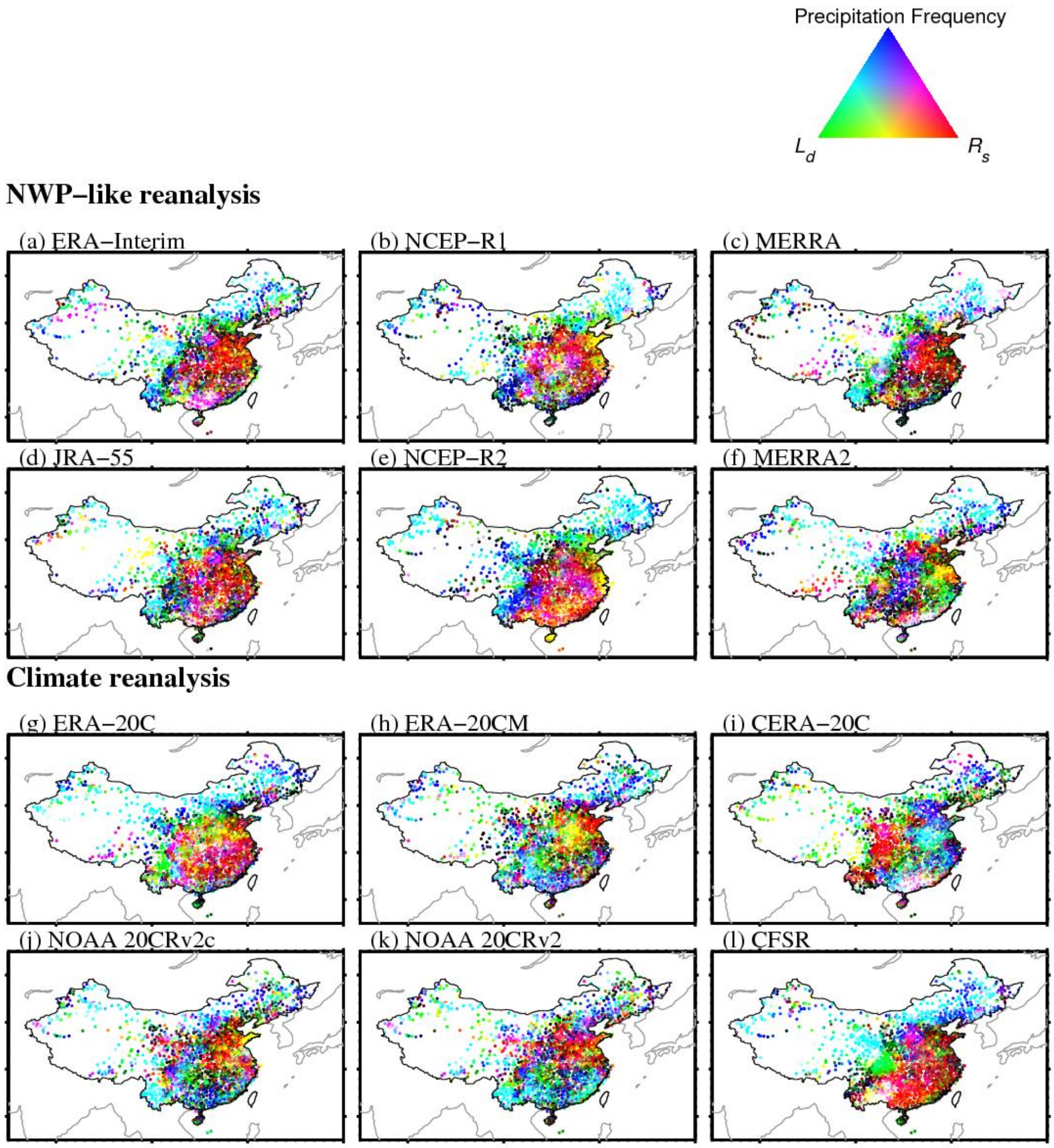

Figure 6. Composite map of the contributions (unit: ℃/decade) of the biases in the trends in three relevant parameters, surface incident solar radiation (R_s, in red), surface downward longwave radiation (L_d, in green) and precipitation frequency (in blue) to the biases in the trends in surface air temperature (T_a) during the study period of 1979-2010, as estimated using the twelve reanalysis products over China.

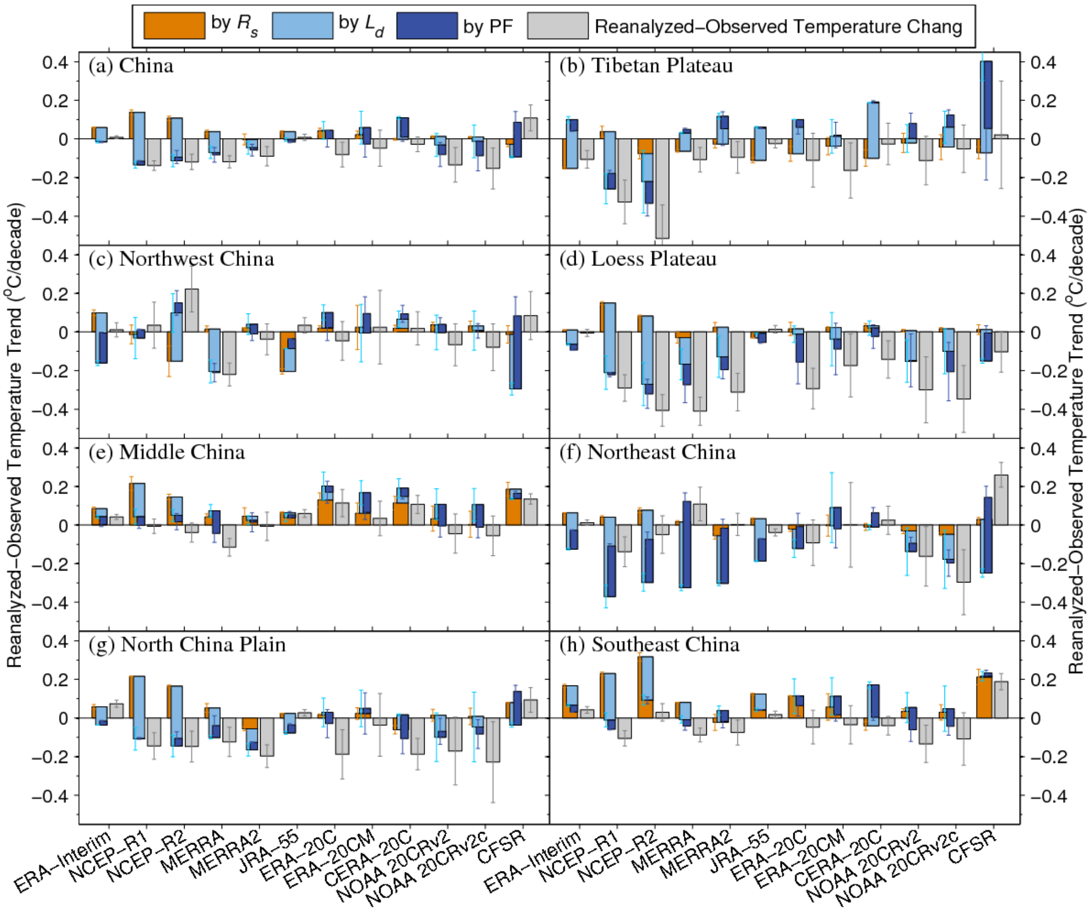

Figure 7. Contribution s(unit: ℃/decade) of the biases in the trends in surface air temperatures (T_a) from three relevant parameters, surface incident solar radiation (R_s, in brown), surface downward longwave radiation (L_d, in light blue) and precipitation frequency (PF, in deep blue) during the study period of 1979-2010 from the twelve reanalysis products over China and its seven subregions.

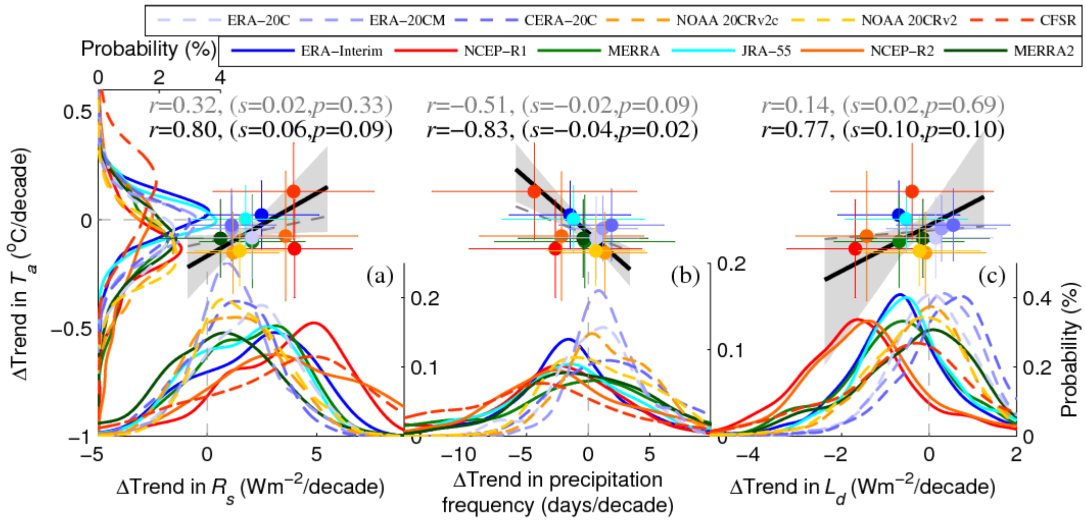

Figure 8. Spatial associations of the simulated biases in the trend in surface air temperature (T_a) versus three relevant parameters among the twelve reanalysis products (solid lines indicate the NWP-like reanalyses, and dashed lines indicate the climate reanalyses). The probability density functions (unit: %) of these biases in the trends are estimated from approximately 700 1°×1° grid cells that cover China. The median values (coloured dots with error bars of spatial standard deviations) of the biases in the trends in T_a (unit: ℃/decade) in the twelve reanalyses are regressed onto those of (**a**) the surface incident solar radiation (R_s, unit: W m^{-2}/decade), (**b**) precipitation frequency (unit: days/decade) and (**c**) the surface downward longwave radiation (L_d, unit: W m^{-2}/decade) using the ordinary least squares method (OLS, denoted by the dashed grey lines) and the weighted total least squares method (WTLS, denoted by the solid black lines). The 5-95% confidence intervals of the regressed slopes obtained using WTLS are shown as shading. The regressed correlations and slopes are shown as grey and black text, respectively.