Diurnal Cycle of Surface Air Temperature within China in Current Reanalyses: Evaluation and Diagnostics

JIZENG DU AND KAICUN WANG

State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

JIANKAI WANG

Chinese Meteorological Administration, Beijing, China

SHAOJING JIANG AND CHUNLÜE ZHOU

State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

(Manuscript received 12 November 2017, in final form 3 March 2018)

ABSTRACT

Diurnal cycle of surface air temperature T is an important metric indicating the feedback of land–atmospheric interaction to global warming, whereas the ability of current reanalyses to reproduce its variation had not been assessed adequately. Here, we evaluate the daily maximum temperature $T_{\max}$, daily minimum temperature $T_{\min}$, and diurnal temperature range (DTR) in five reanalyses based on observations collected at 2253 weather stations over China. Our results show that the reanalyses reproduce $T_{\min}$ very well; however, except for Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2), they substantially underestimate $T_{\max}$ and DTR by 1.21°–6.84°C over China during the period of 1980–2014. MERRA-2 overestimates $T_{\max}$ and DTR by 0.35° and 0.81°C, which are closest with observation. The reanalyses are skillful in reproducing the interannual variability of $T_{\max}$ and $T_{\min}$ but relatively poor for DTR. All reanalyses underestimate the warming trend of $T_{\min}$ by 0.13°–0.17°C (10 yr)^{−1} throughout China during 1980–2014, and underestimate the warming trend of $T_{\max}$ by 0.24°–0.40°C (10 yr)^{−1} in northwestern China while overestimating this quantity by 0.18°–0.33°C (10 yr)^{−1} in southeastern China. These trend biases in $T_{\max}$ and $T_{\min}$ introduce a positive trend bias in DTR of 0.01°–0.26°C (10 yr)^{−1} within China, especially in the north China plain and southeastern China. In the five reanalyses, owing to the sensitivity discrepancies and trend biases, the surface solar radiation R_s and precipitation frequency (PF) are notable deviation sources of the diurnal cycle of air temperature, which explain 31.0%–38.7% (31.9%–37.8%) and 9.8%–22.2% (7.4%–15.3%) of the trend bias in $T_{\max}$ (DTR) over China, respectively.

1. Introduction

Reanalyses provide temporally and spatially continuous estimates of global atmospheric quantities by assimilating available observations into dynamically consistent models

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-17-0773.s1>.

Corresponding author: Kaicun Wang, kcwang@bnu.edu.cn

(Kalnay and Cai 2003; Kalnay et al. 1996). Reanalysis products have been employed in many applications, such as initializing climate models and reproducing atmospheric processes (Dhomse et al. 2011; Jones et al. 2017; Tang et al. 2003). However, reanalysis products also face ongoing challenges, such as the limitations of the numerical models and observations used in assimilation (Compo et al. 2011; Fujiwara et al. 2017). For example, Rienecker et al. (2011) suggested that the assimilated aircraft observations caused a dramatic rise in upper-tropospheric temperature biases in NCEP–NCAR Reanalysis 1 (NCEP R1) during the 1990s.

Reanalysis products must be evaluated before they are applied. Comparisons of reanalyses and observations are effective in evaluating the performance of reanalyses

DOI: 10.1175/JCLI-D-17-0773.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

(Basist and Chelliah 1997; Friedrich et al. 2017; Mooney et al. 2011; Rusticucci and Kousky 2002; Zhou and Wang 2016). Existing studies have demonstrated that the annual mean air temperature anomalies produced by reanalyses agree well with observations over large scales; however, there are obvious systematic errors in the actual values and uncertainties in the long-term trends, especially in regional scales (Dee et al. 2011a; Smith et al. 2001; Thorne and Vose 2010). Moreover, the reliability of reanalyses is vulnerable to topographic effects, insufficient data for use in assimilation, land cover changes, and model parameterizations (Landelius et al. 2016; Rusticucci and Kousky 2002; Whitaker et al. 2009).

Kalnay and Cai (2003) compared the surface mean air temperatures from reanalyses with the corresponding observations within the United States and found that the reanalyses generally reproduce the interannual variability well but significantly underestimate the warming trends. The mean air temperatures collected at flux towers, which are not assimilated into reanalyses, have been used to evaluate the reanalyses. Betts et al. (2006) assessed the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) based on three flux sites in central Saskatchewan and suggested that possible errors in land surface model and the atmospheric cloud parameterizations caused significant biases in near-surface air temperature and relative humidity. Decker et al. (2012) evaluated surface variables and fluxes in six popular reanalyses based on 33 flux sites over the world, and the results show that the ECMWF interim reanalysis (ERA-Interim) is in better agreement with the observations and indicated the biases in near-surface air temperature could be partly attributed to the biases in sensible and latent fluxes.

Ma et al. (2008) assessed the consistency of surface air temperatures between observations and reanalyses in China and found that there are significant cool biases in the reanalyses, especially in western China. Hu et al. (2010) compared the near-surface air temperatures between observations and reanalyses in eastern China and determined that land surface changes explain approximately one-third of the underestimation of the increasing trend in annual warm nights and half of the underestimation of the decreasing trend in cold nights. Wang et al. (2012) explored the reliability of reanalyses within the Tibetan Plateau and suggested that ERA-Interim displays the best performance in reproducing mean air temperatures, whereas the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA-2) yields the highest correlation with observations.

However, existing studies focus primarily on the mean air temperatures produced by reanalyses and generally neglect the diurnal cycle of temperature. The Japanese

55-year Reanalysis (JRA-55) and ERA-Interim have assimilated the 3- and 6-h observed surface air temperatures to minimize the difference between model forecast and observation (Kobayashi et al. 2015; Simmons et al. 2010); this information is handled separately from the atmospheric analysis component, and these temperatures are not used as initial conditions for the atmospheric forecast. This practice has an important impact on the mean air temperatures produced by JRA-55 and ERA-Interim. In contrast, no reanalysis assimilates observational values of daily maximum temperature $T_{\max}$, daily minimum temperature $T_{\min}$, and diurnal temperature range (DTR). The $T_{\max}$, $T_{\min}$, and DTR values from all of the available reanalyses represent pure model output. Therefore, $T_{\max}$, $T_{\min}$, and DTR are better quantities to use in the evaluation of reanalyses, because it is unfair to examine only the mean air temperature for the reanalyses other than JRA-55 and ERA-Interim. More importantly, $T_{\max}$, $T_{\min}$, and DTR depend on land-atmosphere interactions to a greater degree than the mean temperature; the important role of these interactions in controlling regional climate changes has recently been thoroughly characterized by the climate science community (Dai et al. 2004; Du et al. 2017). Such evaluations using $T_{\max}$, $T_{\min}$, and DTR can shed light on the performance of the reanalyses in reproducing land-atmosphere interactions (Kleidon and Renner 2017).

The dominant physical factors underlying the variability in near-surface temperatures differ significantly with spatial scale (Duethmann et al. 2013; Moalafhi et al. 2017). In global scales, the internal variability of the atmospheric system and increases in greenhouse gas concentrations are the main causes of the observed variability and long-term trends at continental (Dai et al. 2015; Lashof and Ahuja 1990). However, at regional scales, land-atmosphere interactions play a more important role in producing climate change (Dai et al. 1997; Mearns et al. 2012; Muller and O’Gorman 2011). Hence, analyzing the impact of land-atmosphere energy budgets on the variability in near-surface temperatures is essential to determine the likely causes of the biases in the reanalyses in reproducing regional climate changes and to improve their accuracy (Berbery et al. 1999; Betts and Beljaars 2017; Seneviratne et al. 2006).

Surface solar radiation R_s and precipitation P are the main elements that control the land-atmosphere energy budget (Du et al. 2017; Trenberth et al. 2009; Zhou et al. 2017b). After receiving solar radiation, the heated land surface transmits energy to the atmosphere via latent and sensible heat fluxes, and the sensible heat warms the near-surface air directly. The variability in P , especially in terms of precipitation frequency (PF), alters soil

moisture, which determines the fractions of energy partitioned into latent and sensible heat (Tang and Leng 2013; Trenberth and Shea 2005; Zhou et al. 2017b). Hence, the accuracy of the inputs R_s and P and the model parameters that control the relationship between R_s and P and air temperatures have strong effects on the near-surface air temperatures produced by reanalyses.

The R_s and P from reanalysis products are produced by forecast models and adjusted by assimilated observations, and attempt to obtain the optimal estimation of their previous value (Zhang et al. 2016). However, previous studies indicated that there are various deficiencies in R_s and P of reanalysis products because of the limitations in parameterization schemes and the observing system. For example, Zhang et al. (2016) found that most reanalyses [NCEP–DOE Reanalysis 2 (NCEP R2), Climate Forecast System Reanalysis (CFSR), and JRA-55] overestimated the R_s by 3.04–6.08 W m^{-2} , while ERA-Interim underestimated the R_s by -3.26 W m^{-2} over the land. Similarly, Lin et al. (2014) found that NCEP R2 and ERA-Interim overestimated precipitation in the intertropical convergence zone by approximately 4 mm day^{-1} , while JRA-25 and MERRA significantly underestimated precipitation in South America by approximately 3 mm day^{-1} . These deficiencies in R_s and P are likely to affect surface air temperature in reanalyses.

In this study, we evaluate the $T_{\max}$, $T_{\min}$, and DTR of the reanalyses, in addition to the commonly evaluated mean air temperature. We explore the underlying physical causes of the biases in the diurnal cycle of surface air temperature. Observations collected at 2479 weather stations from 1980 to 2014 are used to evaluate five current reanalysis products that provide $T_{\max}$, $T_{\min}$, and DTR. We then assess whether the reanalyses capture the effects of R_s and P on $T_{\max}$ and DTR; according to a previous study, R_s and P have no significant effect on $T_{\min}$ (Du et al. 2017). Further, we quantitatively analyze the contributions of R_s and P to the trend biases of $T_{\max}$ and DTR in reanalyses. Moreover, we have discussed the performance of reanalyses in warm seasons (May–October) and cold seasons (from November to the next April) to assess the seasonal dependence of the detection and attribution in this study.

2. Data and methods

a. Observational datasets

The observations of $T_{\max}$, $T_{\min}$, and DTR were obtained from version 3.0 of the Fundamental Elements Datasets of the China National Stations (CNSFED V3.0; Cao et al. 2016; Xu et al. 2013). These data represent the

FIG. 1. A map showing the topography of China and the spatial distribution of selected 2235 meteorological stations. The weather stations cover most regions of China, but their density is extremely uneven; this density is high in southeastern China and sparse in northwestern China and the Tibetan Plateau. To reduce the impact of the spatial heterogeneity, we transform the observation stations into $1^\circ \times 1^\circ$ grid cells, and the value of each grid cell is the average of all of the stations located in that grid cell. All of the calculations in this paper are based on the monthly series of the grid cells. The national means are calculated as the area-weighted average based on the surface area of every grid cell.

latest comprehensive daily datasets and are provided by the National Meteorological Center of China (available online at http://data.cma.cn/en/?r=data/detail&dataCode= SURF_CLI_CHN_MUL_DAY_CES_V3.0). The datasets contain daily historical observations made at 2479 weather stations from 1951 to 2014. There are 2253 (90.9%) stations with more than 30 years of useful values of observations of variables such as $T_{\max}$, $T_{\min}$, daily mean temperature, precipitation, barometric pressure, relative humidity, and sunshine duration from 1980 to 2014 (see Fig. 1). The records in this dataset have been subjected to strict quality controls, including identification of outliers, assessment of spatial and temporal consistency, and correction of suspect and erroneous data.

Although these observations have been checked using the data quality control procedures mentioned above, the inhomogeneity in these observational datasets must be assessed and adjusted before they are used. Inhomogeneities in the records may be caused by the relocation of stations and the replacement of instruments. Such changes result in obvious shifts in the time series of the records and lead to spurious trends (Zhou et al. 2017a). In this study, we detected shifts in the time series by the PMFred algorithm in RHtestsV4, which is established on the penalized maximal F test (PMF) with no need for a reference series (Wang and Feng 2013). Based on the common-trend two-phase regression model, the PMFred algorithm checks for the shifts by detecting any sudden change in the linear trend of time

series (Wang 2008). After checking, for the selected 2253 stations, 1120 (49.7%) stations had a total of 1350 shifts in time series of $T_{\max}$, and 1320 (58.6%) stations had 1752 shifts in time series of $T_{\min}$ during 1980–2014, and all shifts are statistically significant at the 95% confidence level. The spatial and temporal distribution of shifts are shown in Fig. S1 of the supplemental material.

After identifying all of the significant shifts, we used the quantile matching (QM) algorithm in RHtestsV4 to adjust them, which had been applied in the homogenization of meteorological records in recent studies (Aarnes et al. 2015; Dai et al. 2011; Zhou et al. 2017b). The QM adjustment is applied to adjust the base series and make the empirical distributions of all segments in base series match each other (Wang et al. 2010). Note that the base series is the deseasonalized and detrended time series for excluding the impact of the seasonal discontinuity, annual cycle, and the linear trend in base series (Cao et al. 2016). In this study, we set the long-available segment in base series as the datum segment and adjusted the detected shifts at the 95% confidence level. Based on metadata, 412 (36.8%) shifts in $T_{\max}$ and 526 (38.9%) shifts in $T_{\min}$ are related to relocation, which are lower than the shifts percentage caused by relocation (55.0%) assessed by Cao et al. (2016). There were 53 (4.7%) shifts in $T_{\max}$ and 115 (8.5%) shifts in $T_{\min}$ caused by automation during 2003–05. The rest of the shifts are attributed to other causes, such as environmental and undocumented changes. After homogenization, the warming rate of $T_{\max}$ and $T_{\min}$ decreases by $0.04^{\circ}\text{C} (10\text{ yr})^{-1}$ during 1980–2014 over China.

Additionally, most of the weather stations in mountainous areas are located in valleys where they can be built and operated easily; thus, the observed elevation of surface air temperature was much lower than the regional average (Cao et al. 2017; Ma et al. 2008; Zhao et al. 2008). The elevation differences (reanalysis minus observation) between the stations and the model grids result in some cold biases between the reanalyses and the observations. Figure 2 shows that close linear relationships exist between the elevation differences and the biases in $T_{\max}$ and $T_{\min}$. Therefore, we correct the biases in observed $T_{\max}$ and $T_{\min}$ of grids using the elevation differences between observation and reanalyses and the near-surface air temperature lapse rate, which are derived from linear regression coefficients between elevation difference and temperature difference (see Fig. 2).

The variable R_s is the main factor controlling the land surface energy balance during the daytime; however, observations of this quantity are extremely sparse. Yang et al. (2006) proposed an efficient hybrid model that derives R_s from sunshine duration and other

auxiliary input variables (e.g., relative humidity and cloud cover):

$$R_s = [k_0 + k_1 n/N + k_2 (n/N)^2] R_c, \quad (1)$$

where k_0 , k_1 , and k_2 are the fixed coefficients; n and N are the measured sunshine duration and the maximum possible sunshine duration, respectively; and R_c (W m^{-2}) is the surface solar radiation under clear skies, which depends on the gas absorption (e.g., water vapor and ozone), Rayleigh scattering, and aerosol extinction. Previous studies indicate that the sunshine duration–derived R_s values accurately reproduce the internal variability and long-term trends in R_s (Tang et al. 2011; K. C. Wang et al. 2015; Wang 2014). Moreover, the sunshine duration–derived R_s values accurately capture the effects of aerosols and changes in cloud cover on R_s (Magee et al. 2014; Wang et al. 2012).

Precipitation plays an important role in soil moisture and determines the amount of the absorbed R_s that is apportioned to latent heat and sensible heat (Berg et al. 2015). However, studies have shown that PF is more closely related to soil moisture than precipitation amount and is more efficient in indicating the impact of precipitation changes on the energy balance of the land surface (Piao et al. 2009; Wang et al. 2011; Wu et al. 2012; Zhou et al. 2017b). In this study, we set 1 mm day^{-1} as the threshold separating rainy days from rainless days, and the PF was set to the percentage of rainy days in each month.

b. Reanalysis products

To date, five reanalysis products have been developed that provide daily values of $T_{\max}$ and $T_{\min}$, from which DTR can be calculated. These reanalyses include CFSR (Saha et al. 2010), ERA-Interim (Dee et al. 2011b), JRA-55 (Kobayashi et al. 2015), MERRA-2 (Gelaro et al. 2017), and NCEP R2 (Kanamitsu et al. 2002). Table 1 summarizes the characteristics of the five reanalysis products. Note that the CFSR covers only January 1979 through December 2010 and has been extended by the Climate Forecast System, version 2 (CFSv2), to the present. CFSv2 has a higher horizontal resolution than CFSR, but few changes were made to the physical parameterizations; thus, CFSv2 can serve as a continuation of CFSR, and the two datasets are combined for most purposes (Fujiwara et al. 2017; Saha et al. 2014). More information on the horizontal resolution, assimilated scheme, forecast model, and assimilated observations of the five reanalyses is shown in Table 1.

CFSR and NCEP R2 have a standard analysis field of near-surface temperatures, which were derived from a

FIG. 2. A scatterplot showing the relationship between the elevation differences and the temperature biases on the grid scale during the period of 1980–2014 (reanalysis minus obs). Based on the linear regression analysis, the elevation difference is closely related to the biases in T_{max} and T_{min} , but it is not significantly related to DTR. Moreover, the regression coefficients (between -5.04° and $-6.31^{\circ}\text{C km}^{-1}$) are generally comparable to the global mean air temperature lapse rate ($-6.0^{\circ}\text{C km}^{-1}$). The regression coefficients have been used to correct the biases in T_{max} and T_{min} that are caused by the elevation differences.

function of land surface temperature, lowest atmospheric model-level temperature, vertical stability, and surface roughness (Pitman and Perkins 2009; Saha et al. 2010). The near-surface air temperatures in ERA-Interim were interpolated vertically between the lowest level of atmosphere and the land surface by Monin–Obukhov similarity profiles with the parameterization of the surface layer part of the assimilating model (Simmons et al. 2004; Vose et al. 2012). JRA-55 diagnosed the near-surface air temperatures separately from the atmospheric analysis and derived them from a univariate two-dimensional optimal interpolation (Kobayashi et al. 2015). MERRA-2 diagnosed the

near-surface air temperatures in the Goddard Earth Observing System Model, version 5 (GEOS-5), and were interpolated in 2-m height above land surface based on the temperature profile and topography in the analysis model (Molod et al. 2015).

In reanalysis products, R_s is calculated by radiation transfer models based on solar radiation and atmospheric radiative transfer (Slater 2016). The solar radiation is derived from the forecast field in the reanalysis system and the radiative transfer process is determined by the parameterization scheme of atmospheric composition (e.g., ozone and cloud) and the assimilated measurements (e.g., atmospheric humidity profile and

TABLE 1. Detailed information on the five reanalyses used in this paper. In the assimilation column, 3D-Var (4D-Var) denotes the three-dimensional (four-dimensional) variational assimilation scheme, and 3D-FGAT denotes the three-dimensional variational assimilation scheme with the first guess at the appropriate time. In the forecast model column, the year in parentheses denotes the year the version of the operational analysis system was applied in the reanalyses. (GSM is global spectral model.)

Reanalysis	Horizontal resolution	Assimilation scheme	Forecast model	Assimilated observations of related variables
CSFR	$0.5^\circ \times 0.5^\circ$	3D-FGAT	NCEP CFS (2007)	1) gauge precipitation; 2) satellite radiances/ozone.
ERA-Interim	$0.75^\circ \times 0.75^\circ$	4D-Var	IFS cycle 31r2 (2007)	1) surface air temperature; 2) satellite radiances/ozone; and 3) relative humidity/air pressure.
JRA-55	$1.25^\circ \times 1.25^\circ$	4D-Var	JMA GSM (2009)	1) surface air temperature; 2) gauge precipitation; and 3) satellite radiances/ozone.
MERRA-2	$0.5^\circ \times 0.625^\circ$	3D-FGAT	GEOS-5.12.4 (2015)	1) gauge precipitation; 2) satellite radiances/ozone; and 3) satellite aerosols retrievals.
NCEP R2	$1.875^\circ \times 1.875^\circ$	3D-Var	NCEP MRF (1995)	1) satellite-retrieved radiances

aerosol concentration). There are numerous differences in these aspects between reanalyses. For example, NCEP R2 and ERA-Interim used the climatological ozone field to calculate R_s , while other reanalyses used the ozone analysis field (Dee et al. 2011b; Fujiwara et al. 2017; Kanamitsu et al. 2002). In contrast to other reanalyses, MERRA-2 includes a time-variable aerosol analysis in the atmospheric analysis system (Gelaro et al. 2017). The P in reanalyses are produced by the model forecast based on the related climate variables (e.g., humidity) under a physically consistent framework (Gelaro et al. 2017). For reducing the uncertainty, P in CSFR, JRA-55, and MERRA-2 had assimilated gauge observations, as well as satellite measurements (Kobayashi et al. 2015; Reichle et al. 2017a; Saha et al. 2010).

c. Methods

All of the statistical analyses in this paper are based on monthly anomalies derived from the original daily datasets. Months with more than seven missing days are regarded as missing, and we set 1981–2010 as the reference period for the calculation of anomalies. As shown in Fig. 1, the selected weather stations in China are unevenly distributed, and the site density in the eastern part of the country is much greater than in western China. In particular, there are few stations on the Tibetan Plateau and within the desert areas. To reduce the impacts of this spatial heterogeneity, we average the anomalies from the stations into $1^\circ \times 1^\circ$ grid cells. There are 658 grid cells that contain valid station observations, and these cells account for 68.9% of the total grid cells in China. Table 1 shows that the five reanalyses have different horizontal resolutions that do not match each other. Therefore, all of the variables extracted from the reanalyses are transformed to the same $1^\circ \times 1^\circ$ grid using bilinear interpolation. The national mean is calculated

by the area-weighted average of the grid cells based on the surface area of each grid cell.

We use the regional-mean bias (Bias) and the regional-mean absolute error (MAE) to assess the differences between the reanalyses and the observations in terms of the annual mean and the long-term trends of the relevant variables using the following expressions:

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n w_i x_i \quad \text{and} \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n w_i |x_i|, \quad (3)$$

where n denotes the number of grid cells; $x_i = R_i - O_i$; and w_i , R_i , and O_i denote the area weighting, the reanalysis value, and the observed value in the i th grid cell, respectively. The Bias is used to measure the magnitude of the bias of each reanalysis at the regional scale. Compared with Bias, the MAE measures the regional-average bias without considering its sign and is used to evaluate the magnitude of the bias of reanalyses on the grid scale (Willmott and Matsuura 2005).

The Pearson correlation coefficient r of the detrended time series of monthly anomalies between the reanalyses and the observations is used to assess the ability of the reanalyses to capture the interannual variability of temperatures (see Table S1 in the supplemental material). The trends in this paper are calculated using linear regression based on the time series of monthly anomalies, and the confidence of the trends is examined using the two-tailed Student's t test. Additionally, the autocorrelation functions are used to assess the autocorrelation in time series (de Jong et al. 2011). As shown in Fig. S2 of the supplemental material, using monthly anomalies, the serial correlation in raw time series could be almost eliminated.

To quantitatively evaluate the contribution of R_s and PF to the temperature bias between the reanalysis output and the observations, a multiple regression model is used to calculate the sensitivity of temperature to the variability in R_s and PF:

$$T = \beta_{R_s} R_s + \beta_{PF} PF + c + \varepsilon, \quad (4)$$

where T , R_s , and PF denote the detrended time series of the monthly anomalies of $T_{\max}$ (DTR), R_s , and PF separately; β_{R_s} and β_{PF} are partial regression coefficients that denote the sensitivity of temperature to the forcing from R_s and PF, which represents the change in temperature that would result from a unit change in R_s or PF; c is a constant; and ε indicates the residuals. The coefficient of determination R^2 of Eq. (4) indicates that the proportion of the variance in the dependent variable that can be attributed to the independent variables, which represents the contribution of both R_s and PF to the interannual variability in temperature. Note that Eq. (4) actually applied to variables (T , R_s , and PF) in the reanalysis and in observation separately.

The biases in temperature trends caused by the difference in forcing between the reanalyses and the observations in R_s and PF are calculated using the following expression:

$$\Delta T = F_r \beta_r - F_o \beta_o, \quad (5)$$

where ΔT indicates the temperature trend biases caused by the difference in forcing between the reanalyses and the observations; F_r and F_o denote the forcings (R_s or PF) of the reanalyses and the observations, respectively; and β_r and β_o represent the sensitivity of temperature ($T_{\max}$ and DTR) to the forcing (R_s or PF) in the reanalyses and the observations, respectively.

Based on Eq. (5), we can calculate the biases caused by R_s and PF in each grid cell. The MAE of $T_{\max}$ (DTR) is used to measure the magnitude of the trend bias caused by R_s and PF in reanalyses. For example, during a specific period, for CFSR, the sensitivity of $T_{\max}$ to R_s is β_1 and the trend of R_s is f_1 ; for observation, the sensitivity of $T_{\max}$ to R_s is β_2 , and the trend of R_s is f_2 . The trend of $T_{\max}$ forced by the trend of R_s for CFSR and observation is $\beta_1 \times f_1$ and $\beta_2 \times f_2$, respectively. The difference between $\beta_1 \times f_1$ and $\beta_2 \times f_2$ (ΔT) is the trend biases of $T_{\max}$ caused by R_s in reanalyses, which is determined by both the trend biases of R_s and the sensitivity bias of $T_{\max}$ to R_s . We can calculate the ΔT (positive or negative) in every grid (see Fig. 7). Because the sign of ΔT for grids is not consistent over

China, the Bias of ΔT represents the mean biases on the national scale rather than on the grid scale. The MAE of ΔT can avoid the counteraction between positive ΔT and negative ΔT and is a better metric to measure the magnitude of trend bias caused by R_s and PF in reanalyses on the grid scale.

3. Results

a. Multiyear means

For ERA-Interim and NCEP R2, Fig. 3 shows that $T_{\max}$ is underestimated throughout China, and this bias is stronger over southwestern China. In MERRA-2, $T_{\max}$ displays a small positive bias within the north China plain and the Yungui Plateau (see Fig. 3m). However, the bias in $T_{\max}$ displayed no significant spatial differences in CFSR and JRA-55 (Figs. 3d and 3j, respectively). Compared with $T_{\max}$, the bias in $T_{\min}$ in the reanalyses displays a similar spatial pattern that includes a cold bias in northern China but a warm bias in southern China. The exception is NCEP R2, which displays a cold bias within the Tibetan Plateau and a warm bias in other regions, especially northeastern China (see Fig. 3q). The annual mean bias in DTR displays a spatial pattern that is similar to that of $T_{\max}$.

Table 2 shows that all of the reanalyses except one significantly underestimate the annual mean $T_{\max}$ between -1.21° and -5.66°C , whereas they slightly overestimate the annual mean $T_{\min}$ between 0.14° and 1.19°C within China during the period of 1980–2014. The annual mean DTR in the reanalyses are underestimated between -1.35° and -6.84°C and are mainly attributed to the underestimation in $T_{\max}$. The exception is MERRA-2, which overestimates $T_{\max}$ by 0.35°C but underestimates $T_{\min}$ by -0.46°C , and overestimates DTR by 0.81°C (see Table 2). According to the MAE, on the grid scale, the mean bias values associated with the reanalyses range from 1.20° to 5.82°C . According to MAE, the biases in $T_{\max}$ and DTR are largest in NCEP R2 and ERA-Interim, followed by JRA-55 and CFSR, and they are smallest in MERRA-2. For $T_{\min}$, the bias in NCEP R2 is much larger than in the other reanalyses, which display no obvious differences compared to each other.

For the multiyear mean, MERRA-2 displays better performance in reproducing $T_{\max}$ and DTR than the other reanalyses. NCEP R2 reproduces the diurnal cycle poorly. ERA-Interim displays much worse performance than most of the other latest-generation reanalyses (CFSR, JRA-55, and MERRA-2) with regard to $T_{\max}$ and DTR, whereas it displays better performance in reproducing $T_{\min}$.

FIG. 3. The multiyear mean of (a) T_{max} , (b) T_{min} , and (c) DTR in observation (obs) during the period of 1980–2014. The multiyear mean bias (reanalysis minus obs) in T_{max} , T_{min} , and DTR for (d)–(f) CFSR, (g)–(i) ERA-Interim, (j)–(l) JRA-55, (m)–(o) MERRA-2, (p)–(r) and NCEP R2 after adjusting the impacts of elevation differences between reanalyses and observations.

b. Interannual variability

The correlation coefficients between the reanalyses and the observations are shown in Fig. 4. For T_{max} , the correlations between the reanalyses and the observations are quite significant within China and display no obvious spatial heterogeneities. The exception is NCEP R2, for which the correlation coefficients in southwestern China are much lower than in other regions (see Fig. 4m).

Similar results are obtained for T_{min} . For DTR, the reanalyses yield a similar spatial pattern for the correlation coefficient, which is lowest in the northwestern and northeastern parts of China, whereas it is highest in southeastern China (see Fig. 4).

As shown in Table S1, for T_{max} and T_{min} , the correlation coefficients describing the relationship between NCEP R2 ($r = 0.71$ for T_{max} and $r = 0.68$ for T_{min}) and the observations are significantly lower than those

TABLE 2. The biases between the reanalyses and the observations (reanalysis minus obs) in the multiyear mean values and the long-term trends ($\pm 95\%$ confidence interval) within China during the period of 1980–2014 after adjusting for elevation differences. Positive values represent overestimates, and negative values represent underestimates. MAE is also shown in the table. The annual mean values are multiyear means of the actual values during the period of 1980–2014. The long-term biases in the trends are calculated using the monthly anomalies of the variables.

Data	Index	Multiyear mean value ($^{\circ}\text{C}$)			Long-term trend [$^{\circ}\text{C} (10\text{ yr})^{-1}$]		
		$T_{\max}$	$T_{\min}$	DTR	$T_{\max}$	$T_{\min}$	DTR
Obs	Mean	17.61	6.56	11.05	0.28 ± 0.14	0.34 ± 0.11	-0.06 ± 0.11
CFSR – obs	Mean	–1.21	0.14	–1.35	0.13 ± 0.08	-0.13 ± 0.08	0.26 ± 0.09
	MAE	1.63	1.33	1.90	0.26 ± 0.08	0.23 ± 0.08	0.36 ± 0.09
ERA-Interim – obs	Mean	–5.10	0.52	–5.61	-0.11 ± 0.08	-0.16 ± 0.06	0.05 ± 0.09
	MAE	5.12	1.19	5.61	0.20 ± 0.08	0.21 ± 0.06	0.18 ± 0.09
JRA-55 – obs	Mean	–1.31	0.62	–1.93	-0.02 ± 0.07	-0.15 ± 0.07	0.13 ± 0.08
	MAE	1.57	1.37	2.17	0.16 ± 0.07	0.20 ± 0.07	0.22 ± 0.08
MERRA-2 – obs	Mean	0.35	–0.46	0.81	-0.12 ± 0.08	-0.13 ± 0.07	0.01 ± 0.09
	MAE	1.20	1.20	1.74	0.19 ± 0.08	0.19 ± 0.07	0.20 ± 0.09
NCEP R2 – obs	Mean	–5.66	1.19	–6.84	-0.09 ± 0.10	-0.17 ± 0.10	0.09 ± 0.09
	MAE	5.82	2.90	6.84	0.26 ± 0.10	0.26 ± 0.10	0.21 ± 0.09

associated with the other reanalyses ($r = 0.81$ – 0.85 for $T_{\max}$ and $r = 0.76$ – 0.79 for $T_{\min}$). For DTR, the correlation coefficients of NCEP R2 ($r = 0.51$) are the lowest of the reanalyses, followed by ERA-Interim ($r = 0.56$); the other reanalyses displayed no significant differences relative to each other ($r = 0.61$ – 0.63). Overall, JRA-55 displays the best performance in reproducing the interannual variability in $T_{\max}$, $T_{\min}$, and DTR within China and is consistent with previous studies that indicate that JRA-55 displays good performance in reproducing the variability in mean temperature in East Asia (Chen et al. 2014). The correlation coefficients between the reanalyses and the observations are significant at the 95% confidence level in all of the grid cells. By contrast, all of the reanalyses are skillful in capturing the interannual variability in $T_{\max}$ and $T_{\min}$ on the grid scale, but are relatively poor in reproducing the interannual variability in DTR.

c. Long-term trends

Figure 5 shows the long-term trends of $T_{\max}$, $T_{\min}$, and DTR derived from observations. The warming trends in these variables display clear spatial heterogeneity. The warming rate of $T_{\max}$ is comparatively high in northwestern and southeastern China and relatively low in the north China plain and the Pearl River basin (see Fig. 5a). The warming trend in $T_{\min}$ is relatively high in northern China and relatively low in southern China (see Fig. 5b). Most regions of China, except from the Sichuan basin and Guanzhong Plain, display decreasing trends in DTR (see Fig. 5c).

The observed nationally averaged rates of warming in $T_{\max}$ and $T_{\min}$ from 1980 to 2014 are $0.28^{\circ} \pm 0.11^{\circ}\text{C} (10\text{ yr})^{-1}$ and $0.34^{\circ} \pm 0.09^{\circ}\text{C} (10\text{ yr})^{-1}$ (see Table 2). For $T_{\max}$, the warming rate of JRA-55 is closest to that

derived from the observations, whereas CFSR overestimates the warming rate by $0.13^{\circ}\text{C} (10\text{ yr})^{-1}$, and ERA-Interim, MERRA-2, and NCEP R2 underestimate the warming trend by similar amounts (see Table 2). All of the reanalyses consistently underestimate the warming trend in $T_{\min}$ (see Table 2). The DTR extracted from the observations display a decreasing trend of $-0.06^{\circ} \pm 0.05^{\circ}\text{C} (10\text{ yr})^{-1}$. All of the reanalyses display significant positive biases [0.01° – $0.26^{\circ}\text{C} (10\text{ yr})^{-1}$] in the nationally averaged trend in DTR over China from 1980 to 2014.

In spatial terms, the current reanalyses overestimate the warming trend in $T_{\max}$ in southeastern China by 0.18° – $0.34^{\circ}\text{C} (10\text{ yr})^{-1}$ and underestimate that in northwestern China between -0.24° and $-0.40^{\circ}\text{C} (10\text{ yr})^{-1}$, especially within the Loess Plateau. However, MERRA-2 underestimates the warming trend in $T_{\max}$ over most of China by $-0.25^{\circ}\text{C} (10\text{ yr})^{-1}$, except in northeastern China. The underestimation of the warming trend in $T_{\min}$ is relatively pronounced in northern China. The patterns of the trend biases in DTR from the reanalyses share a similar pattern with those of $T_{\max}$.

d. The contributions of biases in the trends in R_s (PF) to those in $T_{\max}$ (DTR)

The sensible heat flux from the land surface is the main energy source for heating the near-surface air. Thus, exchanges of energy between the land surface and the atmosphere have substantial effects on regional warming (Betts 2009). Studies indicate that there are significant regional differences in the variability in R_s and P , and these differences resemble the spatial pattern of regional warming (Tang et al. 2011; Wild et al. 2005; Zhai et al. 2005). Our previous work shows that the land–atmosphere energy budget is an essential input in

FIG. 4. The spatial patterns of r between reanalyses and observations for $T_{\max}$, $T_{\min}$, and DTR calculated using (a)–(c) CFSR, (d)–(f) ERA-Interim, (g)–(i) JRA-55, (j)–(l) MERRA-2, and (m)–(o) NCEP R2 in annual during the period of 1980–2014 within China. The correlation coefficients are calculated using the detrended monthly anomalies of the variables. The reanalyses are sampled to have the same spatial and temporal coverage as the observations.

reanalyses that enables them to capture the interannual variability and long-term trends in near-surface air temperatures (Du et al. 2017).

The sensitivity of $T_{\max}$ and DTR to R_s in the reanalyses is close to that in the observations, whereas the sensitivity of $T_{\max}$ to R_s is significantly underestimated in the Tibetan Plateau and northern China (see Fig. S3 in the supplemental material). The national mean sensitivities of both $T_{\max}$ and DTR to R_s are approximately $0.04^\circ\text{C} (\text{W m}^{-2})^{-1}$ in the observations. The sensitivities to R_s in the reanalyses range from 0.01° to $0.04^\circ\text{C} (\text{W m}^{-2})^{-1}$ for $T_{\max}$ and from 0.02° to $0.04^\circ\text{C} (\text{W m}^{-2})^{-1}$ for DTR. Most of the reanalyses underestimate the nationally averaged sensitivity of $T_{\max}$ and DTR to R_s , except for MERRA-2, which is closest to the observations (see Fig. S3m).

The sensitivities of $T_{\max}$ and DTR to PF display an obvious spatial pattern in the observations. For $T_{\max}$, the current reanalyses display a spatial pattern that is similar to that seen in the observations, and slightly underestimate the sensitivity in southern China (see Fig. S4 in the supplemental material). The national mean sensitivities of $T_{\max}$ and DTR to PF are -0.05° and $-0.04^\circ\text{C } \%^{-1}$, and those from the reanalyses are between -0.03° and $-0.05^\circ\text{C } \%^{-1}$ and between -0.01° and $-0.04^\circ\text{C } \%^{-1}$, respectively (see Fig. S4m). CFSR yields the estimates of the sensitivities of $T_{\max}$ and DTR to PF that are closest to those obtained from the observations.

During the period of 1980–2014, the national mean R_s displays a slight decreasing trend of $-0.60 \pm 3.14 \text{ W m}^{-2} (10 \text{ yr})^{-1}$ (see Table 3). Within in the north China plain,

FIG. 5. The long-term trends in $T_{\max}$, $T_{\min}$, and DTR for the (a)–(c) observations, and the biases in the long-term trends between the observations and (d)–(f) CFSR, (g)–(i) ERA-Interim, (j)–(l) JRA-55, (m)–(o) MERRA-2, (p)–(r) and NCEP R2 in annual during the period of 1980–2014 within China. The long-term trends are calculated by linear regression based on the time series of monthly anomalies, and the confidence of the trends are assessed using the two-tailed Student's t test. The dark areas indicate the grid cells where the trends are significant at the 95% confidence level.

R_s decreases significantly (see Fig. 6a). By contrast, the national mean trend of R_s is underestimated by $-0.36 \text{ W m}^{-2} (10 \text{ yr})^{-1}$ in MERRA-2, while overestimated by $0.86\text{--}3.04 \text{ W m}^{-2} (10 \text{ yr})^{-1}$ in other reanalyses. In spatial, most reanalyses overestimate the trend of R_s in eastern China, except MERRA-2, which only overestimates that in northeastern China. All reanalyses underestimate the trend of R_s over western China (see Figs. 6b–f). To date, MERRA-2 is the only reanalysis that assimilates aerosol observations (Gelaro et al. 2017).

The national mean PF declined by $-0.44\% \pm 0.59\% (10 \text{ yr})^{-1}$ from 1980 to 2014 and decreased significantly in southern China (see Table 3 and Fig. 6g). The national mean trends of PF are underestimated in CFSR and JRA-55 by -0.56% and $-0.06\% (10 \text{ yr})^{-1}$, but are overestimated in MERRA-2 and NCEP R2 by 0.07% and $0.29\% (10 \text{ yr})^{-1}$, and there is no significant bias in ERA-Interim. The spatial pattern of trend in PF in MERRA-2 is similar to that seen in the observations, which is consistent with the use of observed precipitation in

TABLE 3. The national mean biases (reanalysis minus obs) of R_s [$\text{W m}^{-2} (10 \text{ yr})^{-1}$] and PF [$\% (10 \text{ yr})^{-1}$] in the long-term trends ($\pm 95\%$ confidence interval) and their impacts on the long-term trends in $T_{\max}$ [$^{\circ}\text{C} (10 \text{ yr})^{-1}$] and DTR [$^{\circ}\text{C} (10 \text{ yr})^{-1}$] within China. The biases in the long-term trends are calculated using the monthly anomalies on the grid scale, and the national average-based area weighting is then calculated. The R_s -forced and PF-forced biases and the biases in the relevant trends are calculated using Eq. (4).

Data	Index	R_s	PF	R_s forced		PF forced	
				$T_{\max}$	DTR	$T_{\max}$	DTR
Obs	Trend	-0.64 ± 1.58	-0.44 ± 0.76	-0.01	-0.01	0.07	0.09
CFSR – obs	Trend	3.04 ± 1.20	-0.56 ± 0.85	0.11	0.12	0.02	0.04
	MAE	4.23 ± 1.20	1.53 ± 0.85	0.16	0.16	0.08	0.06
ERA-Interim – obs	Trend	1.62 ± 1.04	0.00 ± 0.82	0.03	0.04	-0.01	0.00
	MAE	2.59 ± 1.04	0.80 ± 0.82	0.09	0.09	0.04	0.02
JRA-55 – obs	Trend	0.86 ± 1.10	-0.06 ± 0.94	0.03	0.03	-0.02	-0.01
	MAE	2.26 ± 1.10	1.19 ± 0.94	0.10	0.09	0.06	0.03
MERRA-2 – obs	Trend	-0.36 ± 1.15	0.29 ± 0.63	-0.01	-0.01	-0.01	-0.01
	MAE	2.40 ± 1.15	0.56 ± 0.63	0.09	0.10	0.02	0.02
NCEP R2 – obs	Trend	2.95 ± 1.28	0.07 ± 0.88	0.03	0.06	-0.01	0.00
	MAE	3.60 ± 1.28	1.12 ± 0.88	0.14	0.10	0.05	0.02

MERRA-2 (Reichle et al. 2017b). However, CFSR, ERA-Interim, JRA-55, and NCEP R2 display a spurious increasing trend in western China and especially within the Loess Plateau, whereas they overestimate the decreasing trend within eastern China (see Figs. 6h–k).

Variables R_s and PF are the two most important forcing factors for air temperature calculations in the forecast model: the former determines the energy source and the latter determines the distribution of the energy absorbed by land surface. In addition, the biases caused by R_s and PF are affected by the sensitivity of air temperature to them in reanalyses, which depends on the parameterization scheme of land–atmospheric interaction. Hence, based on the above analyses, we quantify the effects of biases in R_s and PF on the long-term trends in $T_{\max}$ and DTR in the reanalyses. As shown in Fig. 7, except for MERRA-2, the biases in the R_s forcing in most of the reanalyses lead to positive biases in the trends in $T_{\max}$ and DTR in southeastern China and especially in the north China plain, as well as a slight negative bias in this trend in northwestern China during 1980–2014. In MERRA-2, the biases in R_s forcing lead to positive biases in these trends in the Tibetan Plateau and northeastern China, whereas they lead to negative biases in these trends in other regions (see Fig. 7d). The R_s explains 39.5%–47.2% of the bias in the trend in $T_{\max}$ and 39.6%–48.2% of the bias in the trend in DTR in the reanalyses (see Figs. 7f,i).

Compared to those of R_s , the bias in the trends in $T_{\max}$ and DTR caused by PF are relatively small (see Fig. 8 and Table 3). In terms of the national average, PF explains 7.2%–14.6% of the trend bias in $T_{\max}$ and 4.3%–19.7% of the trend bias in DTR. The trend bias in $T_{\max}$ caused by PF is positive in northeastern China and negative in other regions, especially southeastern China.

The bias in PF forcing causes a significant positive bias in the trend in DTR in eastern China and southern China, whereas the bias in this trend is negative in the western Loess Plateau.

The trend biases in $T_{\max}$ caused by both R_s and PF in CFSR and NCEP R2 are much larger than those estimated from the other reanalyses. However, the trend biases in DTR attributed to R_s and PF in CFSR are far higher than those of the other reanalyses. MERRA-2 displays the lowest biases in the trends in both $T_{\max}$ and DTR attributed to PF. Except for MERRA-2, most of the reanalyses overestimate the warming trend in $T_{\max}$ and DTR in the north China plain because they significantly overestimate the trend in R_s (see Figs. 6 and 7). MERRA-2 assimilates space-based aerosol measurements (Gelaro et al. 2017; Reichle et al. 2017a), which may explain its better performance in reproducing the impacts of R_s and PF on the long-term trends in $T_{\max}$ and DTR (Gryspeerd et al. 2014; Jung et al. 2016; Pöschl et al. 2010).

e. The seasonality of evaluation and diagnostics

Since in a typical monsoon climate zone, the temperature and precipitation over China have significant seasonal differences. Hence, whether the performance of reanalyses has significant seasonality is worth discussing. As shown in Figs. S5 and S6 of the supplemental material, the spatial patterns of multiyear mean biases for $T_{\max}$, $T_{\min}$, and DTR had no significant seasonal difference, apart from $T_{\max}$ and DTR having a larger negative bias in cold seasons compared to that in warm seasons. On the national scale, the area-weighted biases for DTR in cold seasons are higher than those in warm seasons by 0.02° – 1.30°C . On the grid scale, the biases of the diurnal cycle of surface air temperature are larger in

FIG. 6. (a) The trend in R_s of the observations and the bias in the trend between the observations and the reanalysis datasets (b) CFSR, (c) ERA-Interim, (d) JRA-55, (e) MERRA-2, and (f) NCEP R2 in years during 1980–2014. (g) The trend in PF of the observations and the bias in the trend between the observations and the reanalysis datasets (h) CFSR, (i) ERA-Interim, (j) JRA-55, (k) MERRA-2, and (l) NCEP R2. The dark areas indicate the grid cells where the trends are significant at the 95% confidence level.

cold seasons than in warm seasons (contrast the MAE in Tables S2 and S3 in the supplemental material). Reanalyses are skillful in reproducing the interannual variability of $T_{\max}$ and $T_{\min}$ in both warm seasons and cold seasons (see Table S1). However, the correlation coefficients for DTR in northern China during cold seasons were much larger than those during warm seasons (see Figs. S7 and S8). For long-term trends, the spatial patterns in warm seasons and cold seasons have no significant difference, being similar to the annual scale (Figs. S9 and S10). However, the trend biases of $T_{\max}$, $T_{\min}$, and DTR during cold seasons are 1.4%–39.2% larger than that during warm seasons on the grid scale (see Tables S2 and S3).

Contrary to the seasonality of the biases in surface air temperatures, for all reanalyses, the trend biases for R_s and PF during warm seasons were larger than those during cold seasons. The national mean trend biases for R_s and PF during warm seasons are $2.94\text{--}5.71 \text{ W m}^{-2} (10 \text{ yr})^{-1}$ and $0.83\%\text{--}1.85\% (10 \text{ yr})^{-1}$, respectively (see Tables S4 and S5 in the supplemental material), and larger than those during cold seasons [$1.61\text{--}3.04 \text{ W m}^{-2} (10 \text{ yr})^{-1}$ and $0.43\%\text{--}1.54\% (10 \text{ yr})^{-1}$]. The spatial heterogeneity of trend biases for both R_s and PF during warm seasons is much larger than that during cold seasons (see Figs. S11 and S12 in the supplemental material). During warm seasons, the trend bias in R_s (PF) can explain 33.0%–37.8% (12.6%–20.5%) of the trend bias of $T_{\max}$ and 33.9%–39.0% (9.4%–13.7%)

FIG. 7. The contributions of R_s to the biases in the trends in $T_{\max}$ and DTR during the period of 1980–2014. The reanalysis datasets are (a),(g) CFSR, (b),(h) ERA-Interim, (c),(i) JRA-55, (d),(j) MERRA-2, and (e),(k) NCEP R2. (f),(l) Histograms show the national area-weighted mean of the contributions of the biases in the trends in R_s to the biases in the trends in $T_{\max}$ and DTR, respectively.

of the trend bias of DTR. During cold seasons, the trend bias in R_s (PF) can explain 25.0%–30.1% (8.2%–22.2%) of the trend bias of $T_{\max}$ and 21.2%–31.0% (5.7%–13.3%) of the trend bias of DTR (see Figs. S13 and S14). The proportion of the trend bias in $T_{\max}$ and DTR attributed to R_s and PF is larger during warm seasons than during cold seasons, especially to R_s .

4. Conclusions and discussion

In this paper, we use the observations from 2479 weather stations in China to evaluate the diurnal cycle of surface air temperature in five current reanalyses.

Compared to daily mean temperature, the evaluations based on diurnal cycle are more sensitive to physical rationality of reanalysis systems and are less affected by assimilated observations. Additionally, we examine the performance of the reanalyses in capturing the impacts of R_s and PF on $T_{\max}$ and DTR and quantify the contributions of the biases in the trends in R_s and PF to the biases in the trends in $T_{\max}$ and DTR.

The results show that, except for MERRA-2, all of the reanalyses underestimate the multiyear mean $T_{\max}$ between -5.66° and -1.21°C , which results in the reanalyses underestimating DTR between -6.84° and -1.35°C during the period of 1980–2014 within China.

FIG. 8. As in Fig. 7, but for the contributions of PF.

MERRA-2 overestimates multiyear mean $T_{\max}$ and DTR by 0.36° and 0.81°C , respectively, and displays the best agreement with the values extracted from the observations. All of the reanalyses accurately reproduce the observed multiyear mean of $T_{\min}$.

These current reanalyses are skillful in reproducing the interannual variability in the diurnal cycles of $T_{\max}$ ($r = 0.71\text{--}0.85$) and $T_{\min}$ ($r = 0.68\text{--}0.79$), while they had a relatively poor performance in DTR ($r = 0.51\text{--}0.63$) during the period 1980–2014 over China. The reanalyses capture the interannual variability in $T_{\max}$ and $T_{\min}$ very well in most regions, but they show a relatively poor performance in the Tibetan Plateau, especially for $T_{\min}$. For DTR, the performance of the reanalyses is far weaker than

that of $T_{\max}$ and $T_{\min}$, especially in the Tibetan Plateau and northern China.

The warming trend in $T_{\max}$ is overestimated in southeastern China between 0.18° and $0.34^{\circ}\text{C } (10\text{yr})^{-1}$, whereas it is underestimated in northeastern China between -0.24° and $-0.40^{\circ}\text{C } (10\text{yr})^{-1}$, particularly within the Loess Plateau. The exception is MERRA-2, which underestimates the warming rate of $T_{\max}$ and DTR in southeastern China. The warming trend in $T_{\min}$ over China as a whole is underestimated between -0.13° and $-0.17^{\circ}\text{C } (10\text{yr})^{-1}$ in the reanalyses, and the degree of underestimation in northern China is higher than that in southern China. The nationally averaged warming trend in DTR is overestimated between 0.01° and 0.26°C

$(10\text{ yr})^{-1}$ and displays a similar pattern as the bias in the trend in $T_{\max}$ (see Fig. 3).

MERRA-2 had the best overall performance in capturing the diurnal cycle of surface air temperatures, while the overall performance of NCEP R2 is much weaker than other reanalyses. Except for NCEP R2, the multiyear biases of $T_{\max}$ and DTR in ERA-Interim were much larger than other reanalyses, whereas ERA-Interim showed a better performance in daily mean temperature than previous studies (Simmons et al. 2010; Wang and Zeng 2012; S. Wang et al. 2015).

Compared with the observations, all reanalyses slightly underestimate the sensitivities of $T_{\max}$ and DTR to R_s , except MERRA-2, which produces the results that are closest to the observations. The sensitivities of $T_{\max}$ and DTR to PF are slightly overestimated in CFSR but are greatly underestimated in the other four reanalyses. Additionally, the trends of R_s in reanalyses are overestimated in southeastern China and underestimated in northwestern China, except for MERRA-2, in which the trends of R_s are underestimated in most regions, excluding northeastern China. The trend in PF in MERRA-2 is similar to that seen in the observations, but other reanalyses display a spurious increasing trend in western China and overestimate the decreasing trend within eastern China. The biases in sensitivities and long-term trends of R_s and PF caused a significant bias in the diurnal cycle of surface air temperature. Based on the quantitative attribution, R_s explains 31.0%–38.7% of the bias in the trend in $T_{\max}$ and 31.9%–37.8% of the bias in the trend in DTR in the reanalyses. PF explains 9.8%–22.2% of the bias in the trend in $T_{\max}$ and 7.4%–15.3% of the bias in the trend in DTR.

Compared with observations, reanalyses have many advantages: first, the physical consistency of reanalyses can reduce the impact of local nonclimatic factors on homogenization (Kalnay and Cai 2003; Yang et al. 2011; Zhou et al. 2004), and second, complete spatial and temporal coverage of reanalyses can provide more accurate detection of global warming (Huang et al. 2017). However, its suitability for characterizing long-term climate trends and its suitability for regional climate change study still are controversial issues. Thorne and Vose (2010) stated that the primary aim of reanalyses are to provide the best possible analysis at each time step; hence, reanalyses may be not homogeneous for assimilating short or intermittent records. By contrast, Dee et al. (2011a) considered that the reanalyses are primarily constructed to long-term homogeneous records, and artificial biases and spurious trends in reanalyses were being corrected by increasing assimilated observations and more advanced assimilated schemes (Dee and Uppala 2009).

Additionally, the reliability of observed records is essential for the confidence of the evaluation and diagnostics of reanalysis products. Compared to raw records, the warming rates of $T_{\max}$ and $T_{\min}$ in homogenized records decreased by $0.04^{\circ}\text{C} (10\text{ yr})^{-1}$ during 1980–2014, which is much smaller than the trend biases between reanalyses and observations [0.16° – $0.26^{\circ}\text{C} (10\text{ yr})^{-1}$]. Although the inhomogeneity in observation has little effect on the assessment of trend biases in reanalysis products, the homogenization of observed records can help to improve the accuracy of quantitative detection and attribution.

In this study, the results show there are significant trend biases of surface air temperatures, R_s , and PF in reanalyses, and the trend biases can be transmitted among variables by physical processes. Hence, it is essential to analyze the spatiotemporal characteristics and physical origins of the biases in the reanalyses, based on the land–atmosphere processes that they resolve (Betts et al. 2009). Compared with the variable-based evaluation, the process-based evaluation supports more efficient modification of the reanalyses while maintaining their physical consistency (Xie et al. 2015); for example, no less than half of the trend biases of $T_{\max}$ and DTR are attributed to the biases of R_s and PF in this study. Investigations of the biases in the diurnal cycle and exploration of the contributions of R_s and PF are useful in understanding the limitations of the reanalyses and updating them for use in studies on long-term climate changes.

Acknowledgments. This study was funded by the National Basic Research Program of China (2017YFA0603601), the National Natural Science Foundation of China (41525018), and the Subject Construction Fund (312231103). The latest observational data were obtained from the China Meteorological Administration (CMA; http://data.cma.cn/en/?r=data/detail&dataCode=SURF_CLI_CHN_MUL_DAY_CES_V3.0). The authors thank the reanalysis groups that produced the datasets that we downloaded and compared. These groups are the ECMWF, which provided the ERA-Interim data (<http://www.ecmwf.int/>); the Global Modeling and Assimilation Office (GMAO) at the NASA Goddard Space Flight Center, which provided the MERRA and MERRA-2 data (<http://gmao.gsfc.nasa.gov/>); the National Oceanic and Atmospheric Administration (NOAA)'s Earth System Research Laboratory (ESRL), which provided the NCEP R2 and CFSR data (<http://www.esrl.noaa.gov/>); and the Climate Prediction Division of the Global Environment and Marine Department at the Japan Meteorological Agency (JMA), which provided the JRA-55 data (<http://jra.kishou.go.jp/>).

REFERENCES

- Aarnes, O. J., S. Abdalla, J.-R. Bidlot, and Ø. Breivik, 2015: Marine wind and wave height trends at different ERA-Interim forecast ranges. *J. Climate*, **28**, 819–837, <https://doi.org/10.1175/JCLI-D-14-00470.1>.
- Basist, A. N., and M. Chelliah, 1997: Comparison of tropospheric temperatures derived from the NCEP/NCAR reanalysis, NCEP operational analysis, and the Microwave Sounding Unit. *Bull. Amer. Meteor. Soc.*, **78**, 1431–1447, [https://doi.org/10.1175/1520-0477\(1997\)078<1431:COTTFD>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<1431:COTTFD>2.0.CO;2).
- Berbery, E. H., K. E. Mitchell, S. Benjamin, T. Smirnova, H. Ritchie, R. Hogue, and E. Radeva, 1999: Assessment of land-surface energy budgets from regional and global models. *J. Geophys. Res.*, **104**, 19 329–19 348, <https://doi.org/10.1029/1999JD900128>.
- Berg, A., and Coauthors, 2015: Interannual coupling between summertime surface temperature and precipitation over land: Processes and implications for climate change. *J. Climate*, **28**, 1308–1328, <https://doi.org/10.1175/JCLI-D-14-00324.1>.
- Betts, A. K., 2009: Land-surface-atmosphere coupling in observations and models. *J. Adv. Model. Earth Syst.*, **1**, 1–18, <https://doi.org/10.3894/JAMES.2009.1.4>.
- , and A. C. M. Beljaars, 2017: Analysis of near-surface biases in ERA-Interim over the Canadian prairies. *J. Adv. Model. Earth Syst.*, **9**, 2158–2173, <https://doi.org/10.1002/2017MS001025>.
- , J. H. Ball, A. G. Barr, T. A. Black, J. H. McCaughey, and P. Viterbo, 2006: Assessing land-surface-atmosphere coupling in the ERA-40 reanalysis with boreal forest data. *Agric. For. Meteorol.*, **140**, 365–382, <https://doi.org/10.1016/j.agrformet.2006.08.009>.
- , M. Köhler, and Y. Zhang, 2009: Comparison of river basin hydrometeorology in ERA-Interim and ERA-40 reanalyses with observations. *J. Geophys. Res.*, **114**, D010761, <https://doi.org/10.1029/2008JD010761>.
- Cao, B., S. Gruber, and T. Zhang, 2017: REDCAPP (v1.0): Parameterizing valley inversions in air temperature data downscaled from reanalyses. *Geosci. Model Dev.*, **10**, 2905–2923, <https://doi.org/10.5194/gmd-10-2905-2017>.
- Cao, L., Y. Zhu, G. Tang, F. Yuan, and Z. Yan, 2016: Climatic warming in China according to a homogenized data set from 2419 stations. *Int. J. Climatol.*, **36**, 4384–4392, <https://doi.org/10.1002/joc.4639>.
- Chen, G., T. Iwasaki, H. Qin, and W. Sha, 2014: Evaluation of the warm-season diurnal variability over East Asia in recent reanalyses JRA-55, ERA-Interim, NCEP CFSR, and NASA MERRA. *J. Climate*, **27**, 5517–5537, <https://doi.org/10.1175/JCLI-D-14-00005.1>.
- Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28, <https://doi.org/10.1002/qj.776>.
- Dai, A., A. D. Del Genio, and I. Y. Fung, 1997: Clouds, precipitation and temperature range. *Nature*, **386**, 665–666, <https://doi.org/10.1038/386665b0>.
- , K. E. Trenberth, and T. T. Qian, 2004: A global dataset of Palmer Drought Severity Index for 1870–2002: Relationship with soil moisture and effects of surface warming. *J. Hydrometeorol.*, **5**, 1117–1130, <https://doi.org/10.1175/JHM-386.1>.
- , J. Wang, P. W. Thorne, D. E. Parker, L. Haimberger, and X. L. Wang, 2011: A new approach to homogenize daily radiosonde humidity data. *J. Climate*, **24**, 965–991, <https://doi.org/10.1175/2010JCLI3816.1>.
- , J. C. Fyfe, S.-P. Xie, and X. G. Dai, 2015: Decadal modulation of global surface temperature by internal climate variability. *Nat. Climate Change*, **5**, 555–559, <https://doi.org/10.1038/nclimate2605>.
- Decker, M., M. A. Brunke, Z. Wang, K. Sakaguchi, X. Zeng, and M. G. Bosilovich, 2012: Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. *J. Climate*, **25**, 1916–1944, <https://doi.org/10.1175/JCLI-D-11-00004.1>.
- Dee, D. P., and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 1830–1841, <https://doi.org/10.1002/qj.493>.
- , E. Källén, A. J. Simmons, and L. Haimberger, 2011a: Comments on “Reanalyses suitable for characterizing long-term trends.” *Bull. Amer. Meteor. Soc.*, **92**, 65–70, <https://doi.org/10.1175/2010BAMS3070.1>.
- , and Coauthors, 2011b: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- de Jong, R., S. de Bruin, A. de Wit, M. E. Schaepman, and D. L. Dent, 2011: Analysis of monotonic greening and browning trends from global NDVI time-series. *Remote Sens. Environ.*, **115**, 692–702, <https://doi.org/10.1016/j.rse.2010.10.011>.
- Dhomse, S., M. P. Chipperfield, W. Feng, and J. D. Haigh, 2011: Solar response in tropical stratospheric ozone: A 3-D chemical transport model study using ERA reanalyses. *Atmos. Chem. Phys.*, **11**, 12 773–12 786, <https://doi.org/10.5194/acp-11-12773-2011>.
- Du, J., K. Wang, J. Wang, and Q. Ma, 2017: Contributions of surface solar radiation and precipitation to the spatiotemporal patterns of surface and air warming in China from 1960 to 2003. *Atmos. Chem. Phys.*, **17**, 4931–4944, <https://doi.org/10.5194/acp-17-4931-2017>.
- Duethmann, D., J. Zimmer, A. Gafurov, A. Güntner, D. Kriegel, B. Merz, and S. Vorogushyn, 2013: Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling. *Hydrol. Earth Syst. Sci.*, **17**, 2415–2434, <https://doi.org/10.5194/hess-17-2415-2013>.
- Friedrich, L. S., A. J. McDonald, G. E. Bodeker, K. E. Cooper, J. Lewis, and A. J. Paterson, 2017: A comparison of Loon balloon observations and stratospheric reanalysis products. *Atmos. Chem. Phys.*, **17**, 855–866, <https://doi.org/10.5194/acp-17-855-2017>.
- Fujiwara, M., and Coauthors, 2017: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmos. Chem. Phys.*, **17**, 1417–1452, <https://doi.org/10.5194/acp-17-1417-2017>.
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Gryspeerdt, E., P. Stier, and D. G. Partridge, 2014: Links between satellite-retrieved aerosol and precipitation. *Atmos. Chem. Phys.*, **14**, 9677–9694, <https://doi.org/10.5194/acp-14-9677-2014>.
- Hu, Y., W. Dong, and Y. He, 2010: Impact of land surface forcings on mean and extreme temperature in eastern China. *J. Geophys. Res.*, **115**, D19117, <https://doi.org/10.1029/2009JD013368>.
- Huang, J., and Coauthors, 2017: Recently amplified Arctic warming has contributed to a continual global warming trend. *Nat. Climate Change*, **7**, 875–879, <https://doi.org/10.1038/s41558-017-0009-5>.

- Jones, P. D., C. Harpham, A. Troccoli, B. Gschwind, T. Ranchin, L. Wald, C. M. Goodess, and S. Dorling, 2017: Using ERA-Interim reanalysis for creating datasets of energy-relevant climate variables. *Earth Syst. Sci. Data*, **9**, 471–495, <https://doi.org/10.5194/essd-9-471-2017>.
- Jung, E., B. A. Albrecht, G. Feingold, H. H. Jonsson, P. Chuang, and S. L. Donaher, 2016: Aerosols, clouds, and precipitation in the North Atlantic trades observed during the Barbados aerosol cloud experiment—Part 1: Distributions and variability. *Atmos. Chem. Phys.*, **16**, 8643–8666, <https://doi.org/10.5194/acp-16-8643-2016>.
- Kalnay, E., and M. Cai, 2003: Impact of urbanization and land-use change on climate. *Nature*, **423**, 528–531, <https://doi.org/10.1038/nature01675>.
- , and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643, <https://doi.org/10.1175/BAMS-83-11-1631>.
- Kleidon, A., and M. Renner, 2017: An explanation for the different climate sensitivities of land and ocean surfaces based on the diurnal cycle. *Earth Syst. Dyn.*, **8**, 849–864, <https://doi.org/10.5194/esd-8-849-2017>.
- Kobayashi, S., and Coauthors, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, <https://doi.org/10.2151/jmsj.2015-001>.
- Landelius, T., P. Dahlgren, S. Gollvik, A. Jansson, and E. Olsson, 2016: A high-resolution regional reanalysis for Europe. Part 2: 2D analysis of surface temperature, precipitation and wind. *Quart. J. Roy. Meteor. Soc.*, **142**, 2132–2142, <https://doi.org/10.1002/qj.2813>.
- Lashof, D. A., and D. R. Ahuja, 1990: Relative contributions of greenhouse gas emissions to global warming. *Nature*, **344**, 529–531, <https://doi.org/10.1038/344529a0>.
- Lin, R., T. Zhou, and Y. Qian, 2014: Evaluation of global monsoon precipitation changes based on five reanalysis datasets. *J. Climate*, **27**, 1271–1289, <https://doi.org/10.1175/JCLI-D-13-00215.1>.
- Ma, L., T. Zhang, Q. Li, O. W. Frauenfeld, and D. Qin, 2008: Evaluation of ERA-40, NCEP-1, and NCEP-2 reanalysis air temperatures with ground-based measurements in China. *J. Geophys. Res.*, **113**, D15115, <https://doi.org/10.1029/2007JD009549>.
- Magee, N. B., E. Melaas, P. M. Finocchio, M. Jardel, A. Noonan, and M. J. Iacono, 2014: Blue Hill Observatory sunshine: Assessment of climate signals in the longest continuous meteorological record in North America. *Bull. Amer. Meteor. Soc.*, **95**, 1741–1751, <https://doi.org/10.1175/BAMS-D-12-00206.1>.
- Mearns, L. O., and Coauthors, 2012: The North American Regional Climate Change Assessment Program: Overview of phase I results. *Bull. Amer. Meteor. Soc.*, **93**, 1337–1362, <https://doi.org/10.1175/BAMS-D-11-00223.1>.
- Moalafhi, D. B., J. P. Evans, and A. Sharma, 2017: Influence of reanalysis datasets on dynamically downscaling the recent past. *Climate Dyn.*, **49**, 1239–1255, <https://doi.org/10.1007/s00382-016-3378-y>.
- Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339–1356, <https://doi.org/10.5194/gmd-8-1339-2015>.
- Mooney, P. A., F. J. Mulligan, and R. Fealy, 2011: Comparison of ERA-40, ERA-Interim and NCEP/NCAR reanalysis data with observed surface air temperatures over Ireland. *Int. J. Climatol.*, **31**, 545–557, <https://doi.org/10.1002/joc.2098>.
- Muller, C. J., and P. A. O’Gorman, 2011: An energetic perspective on the regional response of precipitation to climate change. *Nat. Climate Change*, **1**, 266–271, <https://doi.org/10.1038/nclimate1169>.
- Piao, S., L. Yin, X. Wang, P. Ciais, S. Peng, Z. Shen, and S. I. Seneviratne, 2009: Summer soil moisture regulated by precipitation frequency in China. *Environ. Res. Lett.*, **4**, 044012, <https://doi.org/10.1088/1748-9326/4/4/044012>.
- Pitman, A. J., and S. E. Perkins, 2009: Global and regional comparison of daily 2-m and 1000-hpa maximum and minimum temperatures in three global reanalyses. *J. Climate*, **22**, 4667–4681, <https://doi.org/10.1175/2009JCLI2799.1>.
- Pöschl, U., and Coauthors, 2010: Rainforest aerosols as biogenic nuclei of clouds and precipitation in the Amazon. *Science*, **329**, 1513–1516, <https://doi.org/10.1126/science.1191056>.
- Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S. Partyka, 2017a: Land surface precipitation in MERRA-2. *J. Climate*, **30**, 1643–1664, <https://doi.org/10.1175/JCLI-D-16-0570.1>.
- , C. S. Draper, Q. Liu, M. Girotto, S. P. P. Mahanama, R. D. Koster, and G. J. M. De Lannoy, 2017b: Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, **30**, 2937–2960, <https://doi.org/10.1175/JCLI-D-16-0720.1>.
- Rienecker, M. M., and Coauthors, 2011: MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, <https://doi.org/10.1175/JCLI-D-11-00015.1>.
- Rusticucci, M. M., and V. E. Kousky, 2002: A comparative study of maximum and minimum temperatures over Argentina: NCEP–NCAR reanalysis versus station data. *J. Climate*, **15**, 2089–2101, [https://doi.org/10.1175/1520-0442\(2002\)015<2089:ACSOMA>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2089:ACSOMA>2.0.CO;2).
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, <https://doi.org/10.1175/2010BAMS3001.1>.
- , and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, <https://doi.org/10.1175/JCLI-D-12-00823.1>.
- Seneviratne, S. I., D. Lüthi, M. Litschi, and C. Schär, 2006: Land–atmosphere coupling and climate change in Europe. *Nature*, **443**, 205–209, <https://doi.org/10.1038/nature05095>.
- Simmons, A. J., and Coauthors, 2004: Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. *J. Geophys. Res.*, **109**, D24115, <https://doi.org/10.1029/2004JD005306>.
- , K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee, 2010: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets. *J. Geophys. Res.*, **115**, D01110, <https://doi.org/10.1029/2009JD012442>.
- Slater, A. G., 2016: Surface solar radiation in North America: A comparison of observations, reanalyses, satellite, and derived products. *J. Hydrometeorol.*, **17**, 401–420, <https://doi.org/10.1175/JHM-D-15-0087.1>.
- Smith, S. R., D. M. Legler, and K. V. Verzone, 2001: Quantifying uncertainties in NCEP reanalyses using high-quality research vessel observations. *J. Climate*, **14**, 4062–4072, [https://doi.org/10.1175/1520-0442\(2001\)014<4062:QUINRU>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<4062:QUINRU>2.0.CO;2).
- Tang, Q., and G. Leng, 2013: Changes in cloud cover, precipitation, and summer temperature in North America from

- 1982 to 2009. *J. Climate*, **26**, 1733–1744, <https://doi.org/10.1175/JCLI-D-12-00225.1>.
- Tang, W.-J., K. Yang, J. Qin, C. C. K. Cheng, and J. He, 2011: Solar radiation trend across China in recent decades: A revisit with quality-controlled data. *Atmos. Chem. Phys.*, **11**, 393–406, <https://doi.org/10.5194/acp-11-393-2011>.
- Tang, Y. M., R. Kleeman, A. M. Moore, A. Weaver, and J. Vialard, 2003: The use of ocean reanalysis products to initialize ENSO predictions. *Geophys. Res. Lett.*, **30**, 1694, <https://doi.org/10.1029/2003GL017664>.
- Thorne, P. W., and R. S. Vose, 2010: Reanalyses suitable for characterizing long-term trends. *Bull. Amer. Meteor. Soc.*, **91**, 353–361, <https://doi.org/10.1175/2009BAMS2858.1>.
- Trenberth, K. E., and D. J. Shea, 2005: Relationships between precipitation and surface temperature. *Geophys. Res. Lett.*, **32**, L14703, <https://doi.org/10.1029/2005GL022760>.
- , J. T. Fasullo, and J. Kiehl, 2009: Earth's global energy budget. *Bull. Amer. Meteor. Soc.*, **90**, 311–323, <https://doi.org/10.1175/2008BAMS2634.1>.
- Vose, R. S., S. Applequist, M. J. Menne, C. N. Williams Jr., and P. Thorne, 2012: An intercomparison of temperature trends in the U.S. Historical Climatology Network and recent atmospheric reanalyses. *Geophys. Res. Lett.*, **39**, L10703, <https://doi.org/10.1029/2012GL051387>.
- Wang, A., and X. Zeng, 2012: Evaluation of multireanalysis products with in situ observations over the Tibetan Plateau. *J. Geophys. Res.*, **117**, D05102, <https://doi.org/10.1029/2011JD016553>.
- Wang, K. C., 2014: Measurement biases explain discrepancies between the observed and simulated decadal variability of surface incident solar radiation. *Sci. Rep.*, **4**, 6144, <https://doi.org/10.1038/srep06144>.
- , R. E. Dickinson, M. Wild, and S. Liang, 2012: Atmospheric impacts on climatic variability of surface incident solar radiation. *Atmos. Chem. Phys.*, **12**, 9581–9592, <https://doi.org/10.5194/acp-12-9581-2012>.
- , Q. Ma, Z. Li, and J. Wang, 2015: Decadal variability of surface incident solar radiation over China: Observations, satellite retrievals, and reanalyses. *J. Geophys. Res. Atmos.*, **120**, 6500–6514, <https://doi.org/10.1002/2015JD023420>.
- Wang, S., M. Zhang, M. Sun, B. Wang, X. Huang, Q. Wang, and F. Feng, 2015: Comparison of surface air temperature derived from NCEP/DOE R2, ERA-Interim, and observations in the arid northwestern China: A consideration of altitude errors. *Theor. Appl. Climatol.*, **119**, 99–111, <https://doi.org/10.1007/s00704-014-1107-1>.
- Wang, X. L., 2008: Penalized maximal F test for detecting undocumented mean shift without trend change. *J. Atmos. Oceanic Technol.*, **25**, 368–384, <https://doi.org/10.1175/2007JTECHA982.1>.
- , and Y. Feng, 2013: RHtestsV4 user manual. Climate Research Division, Environment Canada Toronto, Ontario, Canada, 29 pp., http://etccdi.pacificclimate.org/RHtest/RHtestsV4_UserManual_10Dec2014.pdf.
- , H. Chen, Y. Wu, Y. Feng, and Q. Pu, 2010: New techniques for the detection and adjustment of shifts in daily precipitation data series. *J. Appl. Meteor. Climatol.*, **49**, 2416–2436, <https://doi.org/10.1175/2010JAMC2376.1>.
- Wang, Y., and Coauthors, 2011: Precipitation frequency controls interannual variation of soil respiration by affecting soil moisture in a subtropical forest plantation. *Can. J. For. Res.*, **41**, 1897–1906, <https://doi.org/10.1139/x11-105>.
- Whitaker, J. S., G. P. Compo, and J.-N. Thépaut, 2009: A comparison of variational and ensemble-based data assimilation systems for reanalysis of sparse observations. *Mon. Wea. Rev.*, **137**, 1991–1999, <https://doi.org/10.1175/2008MWR2781.1>.
- Wild, M., and Coauthors, 2005: From dimming to brightening: Decadal changes in solar radiation at Earth's surface. *Science*, **308**, 847–850, <https://doi.org/10.1126/science.1103215>.
- Willmott, C. J., and K. Matsuura, 2005: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Res.*, **30**, 79–82, <https://doi.org/10.3354/cr030079>.
- Wu, C., and Coauthors, 2012: An underestimated role of precipitation frequency in regulating summer soil moisture. *Environ. Res. Lett.*, **7**, 024011, <https://doi.org/10.1088/1748-9326/7/2/024011>.
- Xie, S.-P., and Coauthors, 2015: Towards predictive understanding of regional climate change. *Nat. Climate Change*, **5**, 921–930, <https://doi.org/10.1038/nclimate2689>.
- Xu, W., Q. Li, X. L. Wang, S. Yang, L. Cao, and Y. Feng, 2013: Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices. *J. Geophys. Res. Atmos.*, **118**, 9708–9720, <https://doi.org/10.1002/jgrd.50791>.
- Yang, K., T. Koike, and B. Ye, 2006: Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets. *Agric. For. Meteorol.*, **137**, 43–55, <https://doi.org/10.1016/j.agrformet.2006.02.001>.
- Yang, X., Y. Hou, and B. Chen, 2011: Observed surface warming induced by urbanization in east China. *J. Geophys. Res.*, **116**, D14113, <https://doi.org/10.1029/2010JD015452>.
- Zhai, P., X. Zhang, H. Wan, and X. Pan, 2005: Trends in total precipitation and frequency of daily precipitation extremes over China. *J. Climate*, **18**, 1096–1108, <https://doi.org/10.1175/JCLI-3318.1>.
- Zhang, X., S. Liang, G. Wang, Y. Yao, B. Jiang, and J. Cheng, 2016: Evaluation of the reanalysis surface incident shortwave radiation products from NCEP, ECMWF, GSFC, and JMA using satellite and surface observations. *Remote Sens.*, **8**, 225, <https://doi.org/10.3390/rs8030225>.
- Zhao, T., W. Guo, and C. Fu, 2008: Calibrating and evaluating reanalysis surface temperature error by topographic correction. *J. Climate*, **21**, 1440–1446, <https://doi.org/10.1175/2007JCLI1463.1>.
- Zhou, C. L., and K. C. Wang, 2016: Evaluation of surface fluxes in ERA-Interim using flux tower data. *J. Climate*, **29**, 1573–1582, <https://doi.org/10.1175/JCLI-D-15-0523.1>.
- , Y. He, and K. Wang, 2017a: On the suitability of current atmospheric reanalyses for regional warming studies over China. *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2017-966>.
- , K. Wang, and Q. Ma, 2017b: Evaluation of eight current reanalyses in simulating land surface temperature from 1979 to 2003 in China. *J. Climate*, **30**, 7379–7398, <https://doi.org/10.1175/JCLI-D-16-0903.1>.
- Zhou, L., R. E. Dickinson, Y. Tian, J. Fang, Q. Li, R. K. Kaufmann, C. J. Tucker, and R. B. Myneni, 2004: Evidence for a significant urbanization effect on climate in China. *Proc. Natl. Acad. Sci. USA*, **101**, 9540–9544, <https://doi.org/10.1073/pnas.0400357101>.