AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2015JD023906

Key Points:

- Most reanalyses reproduce well the observed PW climatology and interannual variations over China
- Few reanalyses capture the observed long-term PW changes due to the wet biases before about 2002
- The PW biases results mainly from the discontinuities contained in reanalysis relative humidity

Correspondence to:
T. Zhao,
zhaotb@tea.ac.cn

Citation:
Zhao, T., J. Wang, and A. Dai (2015), Evaluation of atmospheric precipitable water from reanalysis products using homogenized radiosonde observations over China, *J. Geophys. Res. Atmos.*, *120*, 10,703–10,727, doi:10.1002/2015JD023906.

Received 8 JUL 2015
Accepted 24 SEP 2015
Accepted article online 29 SEP 2015
Published online 23 OCT 2015

©2015. American Geophysical Union. All Rights Reserved.

Evaluation of atmospheric precipitable water from reanalysis products using homogenized radiosonde observations over China

Tianbao Zhao[1], Juanhuai Wang[2], and Aiguo Dai[3,4]

[1]Key Laboratory of Regional Climate-Environment Research for East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [2]Guangdong Climate Center, Guangzhou, China, [3]Department of Atmospheric and Environmental Sciences, University at Albany, SUNY, Albany, New York, USA, [4]National Center for Atmospheric Research, Boulder, Colorado, USA

Abstract Many multidecadal atmospheric reanalysis products are available now, but their consistencies and reliability are far from perfect. In this study, atmospheric precipitable water (PW) from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR), NCEP/Department of Energy (DOE), Modern Era Retrospective-Analysis for Research and Applications (MERRA), Japanese 55 year Reanalysis (JRA-55), JRA-25, ERA-Interim, ERA-40, Climate Forecast System Reanalysis (CFSR), and 20th Century Reanalysis version 2 is evaluated against homogenized radiosonde observations over China during 1979–2012 (1979–2001 for ERA-40). Results suggest that the PW biases in the reanalyses are within ~20% for most of northern and eastern China, but the reanalyses underestimate the observed PW by 20%–40% over western China and by ~60% over the southwestern Tibetan Plateau. The newer-generation reanalyses (e.g., JRA25, JRA55, CFSR, and ERA-Interim) have smaller root-mean-square error than the older-generation ones (NCEP/NCAR, NCEP/DOE, and ERA-40). Most of the reanalyses reproduce well the observed PW climatology and interannual variations over China. However, few reanalyses capture the observed long-term PW changes, primarily because they show spurious wet biases before about 2002. This deficiency results mainly from the discontinuities contained in reanalysis relative humidity fields in the middle-lower troposphere due to the wet bias in older radiosonde records that are assimilated into the reanalyses. An empirical orthogonal function (EOF) analysis revealed two leading modes that represent the long-term PW changes and El Niño–Southern Oscillation-related interannual variations with robust spatial patterns. The reanalysis products, especially the MERRA and JRA-25, roughly capture these EOF modes, which account for over 50% of the total variance. The results show that even during the post-1979 satellite era, discontinuities in radiosonde data can still induce large spurious long-term changes in reanalysis PW and other related fields. Thus, more efforts are needed to remove spurious changes in input data for future long-term reanalyses.

1. Introduction

Water vapor is the single most important greenhouse gas in the atmosphere. It plays a crucial role in Earth's energy and water cycle; thus, it has large impacts on weather and climate [*Held and Soden*, 2000; *Zhao et al.*, 2012; *Wang et al.*, 2013]. Because saturation vapor pressure increases with air temperature (*T*) and atmospheric relative humidity (RH) is relatively stable [*Dai*, 2006; *Dai et al.*, 2011; *Zhao et al.*, 2012], water vapor is expected to increase as air temperature rises, thus providing a strong positive feedback for CO_2 and other greenhouse gas-induced global warming [*Held and Soden*, 2000]. Therefore, how water vapor changes in the real world and climate models in response to rising temperatures is of great concern.

Traditionally, global radiosonde data represent an important resource for monitoring and understanding the long-term changes in atmospheric water vapor, and they have been used to quantify its long-term trends [e.g., *Ross and Elliott*, 2001; *Wang et al.*, 2003; *Rowe et al.*, 2008; *Zhao et al.*, 2012; *Wang et al.*, 2013]. Radiosonde observations have also been assimilated into weather forecasting models to produce reanalysis products, often without homogenization [e.g., *Kalnay et al.*, 1996; *Uppala et al.*, 2005]. However, the accuracy of the radiosonde-derived water vapor content is limited by the sensor characteristics that vary in space and time [*Wang and Zhang*, 2008]. In particular, the response time and sensitivity to low humidity of radiosonde sensors have been improved over the years. This has induced spurious long-term trends in radiosonde humidity records

[*Dai et al.*, 2011]. These problems, especially the temporal inhomogeneities, have severely hampered the application of the radiosonde humidity data in climate research [*Dai et al.*, 2011].

Atmospheric reanalyses are constrained mainly by atmospheric radiosonde observations before the satellite era (before about 1979); thereafter, satellite observations, which have their own inhomogeneity problems, have become increasingly important for reanalysis products. Since biases and inhomogeneities in radiosonde data have not been corrected before their assimilation into reanalysis products (except for ERA-Interim, which used homogenized air temperature from radiosondes but unhomogenized data for other variables [*Dee et al.*, 2011]). The errors in these input data induce systematic dry/wet biases and spurious long-term changes in reanalysis water vapor and other related fields (e.g., precipitation and cloudiness). The spurious changes often make the reanalysis data unsuitable for trend analyses [*Bengtsson et al.*, 2004; *Trenberth et al.*, 2005; *Qian et al.*, 2006; *Dai et al.*, 2011, 2013].

Some previous studies have analyzed and evaluated humidity and water vapor trends in reanalysis products. For example, *Paltridge et al.* [2009] found that specific humidity in the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis declined between 1973 and 2007, particularly in the tropical middle and upper troposphere. *Dessler and Davis* [2010] examined this issue for five reanalyses (the NCEP/NCAR, the European Centre for Medium-Range Weather Forecasts (ECMWF) 40 Year reanalysis and the more modern Japanese Reanalysis (JRA), Modern Era Retrospective-Analysis for Research and Applications (MERRA), and European Centre for Medium-Range Weather Forecasts (ECMWF) Interim reanalyses. They found that the NCEP/NCAR reanalysis contains a large bias associated with a decrease in tropical upper tropospheric specific humidity mainly due to inhomogeneities in the radiosonde humidity measurements. *Serreze et al.* [2012] compared three of the recent reanalyses (i.e., MERRA, NCEP Climate Forecast System Reanalysis (CFSR), and ERA-Interim) against radiosonde measurements at nine sites north of 70°N. They found that the reanalyses have moist and warm biases at and near the surface from autumn through spring, with smaller biases in summer, and none of the reanalyses correctly capture the cold season humidity and temperature inversions seen in the radiosonde data, with the MERRA in particular showing evidence of artifacts likely introduced by changes in assimilation data streams. *Dai et al.* [2011] showed that trend patterns over land in 700 hPa specific humidity from 1973 to 2008 from the ERA-40 and ERA-Interim are similar to those in the raw radiosonde data with noisy patterns and negative trends over Africa that are due to spurious discontinuities in the humidity records. Since all the reanalyses assimilated the same (or similar) radiosonde data, common spurious changes, such as a tropospheric cooling trend over central Asia since the early 1960s [*Dai et al.*, 2013], may exist in multiple reanalysis products. Such a common feature should not be interpreted as a robust real climate change signal; rather, it could result from the errors in the common input data.

Over China, previous validation studies have focused mainly on surface air temperature, precipitation, surface humidity, surface winds, and radiation from reanalysis products using surface observations [e.g., *Xu et al.*, 2001; *Shi et al.*, 2006; *Ma et al.*, 2008, 2009; *Zhao et al.*, 2004; *Zhao and Fu*, 2006a, 2006b; *Zhao et al.*, 2008; *Mao et al.*, 2010; *Jia et al.*, 2013; *Lin et al.*, 2013]. For example, *Bao and Zhang* [2013] evaluated the NCEP/NCAR, CFSR, ERA-40, and ERA-Interim reanalysis products using radiosonde observations from 11 stations over the Tibetan Plateau of China from 10 May to 9 August 1998 and found that each of the reanalysis can produce the mean values consistent with the verifying soundings for temperature and horizontal winds but with large biases for relative humidity. Recently, *Lu et al.* [2015] compared the precipitable water (PW) from JRA-55, ERA-40, ERA-Interim, MERRA, NCEP/Department of Energy (DOE), and International Satellite Cloud Climatology Project with the analysis data constructed from station observations and Moderate Resolution Imaging Spectroradiometer estimates over the Tibetan Plateau, and they found that none of the reanalyses are able to show the observed long-term trends and variability in PW over the Tibetan Plateau. Besides these, few studies have examined the reliability of atmospheric water vapor and its long-term trends from reanalysis products over China.

Atmospheric reanalysis products have been increasingly used to evaluate climate models [e.g., *Dai et al.*, 2001] and study climate variability and change [e.g., *Trenberth et al.*, 2005]. Therefore, validation of the reanalysis products is crucial for their proper applications. The goal of this paper is to examine how well reanalysis products capture observed variations and long-term changes in atmospheric water vapor and humidity over China since 1979 utilizing the recently homogenized radiosonde humidity data produced by *Dai et al.* [2011] and used by *Zhao et al.* [2012]. Since water vapor affects many other fields such as precipitation and clouds, the issues

found here with the PW are also relevant to the other fields in the reanalysis products. Results of this study can help us better understand the strength and weakness of the commonly used reanalysis products and thus facilitate their applications.

This paper is organized as follows. Section 2 describes the observational and reanalysis products used in this study. Section 3 presents the analysis method. In section 4, we compare atmospheric water vapor contents from nine reanalysis products with the observations in terms of the biases or errors in long-term mean, variations, and long-term changes and discuss the possible causes for the water vapor biases. A summary and concluding remarks are given in section 5.

2. Data Sets

2.1. Observational Data

Dai et al. [2011] developed a new approach and applied it to homogenize twice-daily radiosonde humidity records from the global radiosonde network, including about 160 Chinese stations. After this homogenization, atmospheric humidity variables exhibited spatially more coherent long-term trends and compare more favorably with GPS PW than those without homogenization [*Dai et al.*, 2011; *Zhao et al.*, 2012; *Wang et al.*, 2013]. *Zhao et al.* [2012] used this homogenized data set to study the long-term trends of tropospheric water vapor and humidity during 1970–2008 over China and found that the homogenized humidity not only correlates better with other independent measurements but also shows spatially more coherent long-term trends than the raw data over China. The homogenized data over China are now updated to 2012, with around 100 stations with humidity records that are more than 80% complete since 1970. In this paper, we used the homogenized sounding data from 89 Chinese stations that have valid records over 85% of the time during 1970–2012. Figure 1a shows the distribution of these radiosonde stations, which cover most of China except the western Tibetan Plateau. We realize that the homogenized PW data used in this analysis likely still contain some small discontinuities [*Dai et al.*, 2011] and mean biases associated with, for example, the dry bias of the current Vaisala RS92 radiosondes [*Wang et al.*, 2013]. However, it is difficult to quantify these error bars, and any mean biases in the most recent radiosonde records (used by *Dai et al.* [2011] as the reference) would exist in both our homogenized data and the reanalyses products; thus, they unlikely have a large impact on the differences examined here.

In this study, we focus on the differences (referred to as the biases) in the long-term mean, variations, and changes in column-integrated precipitable water (PW) between reanalysis products and the homogenized radiosonde observations, although specific and relative humidity at some individual levels is also examined for diagnostic purpose. An advantage of PW is that it reflects the total water vapor content for the entire tropospheric column (albeit it is dominated by the lower troposphere), thereby avoiding comparisons at individual pressure levels. The PW in this study was calculated by integrating specific humidity from radiosonde observations or reanalysis data from the surface to 300 hPa using $\mathrm{PW} = \frac{0.1}{g}\int_{300}^{p_s} q\,dp$, where PW is mm, $g = 9.8\,\mathrm{m\,s^{-2}}$ is the acceleration of gravity, p_s is surface pressure in hPa, q is specific humidity in $\mathrm{g\,kg^{-1}}$, and p is air pressure in hPa.

In addition, the monthly precipitation from the Global Precipitation Climatology Centre (GPCC) full data reanalysis (V6; http://www.esrl.noaa.gov/psd/data/gridded/data.gpcc.html) [*Schneider et al.*, 2013] and the Climatic Research Unit Time Series 3.22 (CRU TS3.22; http://browse.ceda.ac.uk/browse/badc/cru/data/cru_ts/cru_ts_3.22) [*Harris et al.*, 2013], and the total cloud cover from the CRU TS3.22, are used to diagnose whether the spurious changes in the reanalysis PW series are apparent in the precipitation and cloudiness.

2.2. Reanalysis Products

The reanalysis PW from surface to 300 mb was computed using monthly q data obtained from nine reanalysis products (see Table 1). They include the NCEP/NCAR [*Kalnay et al.*, 1996; *Kistler et al.*, 2001], NCEP/DOE [*Kanamitsu et al.*, 2002], NASA Global Modeling and Assimilation Office (GMAO) MERRA [*Rienecker et al.*, 2011], JRA-55 [*Ebita et al.*, 2011], JRA-25 [*Onogi et al.*, 2007], ERA-40 [*Uppala et al.*, 2005], and its updated version ERA-Interim [*Dee et al.*, 2011], CFSR [*Saha et al.*, 2010], and the 20th Century Reanalysis version 2 (20CR) [*Compo et al.*, 2011]. Monthly air temperature and relative humidity at six pressure levels (1000, 850, 700, 500, 400, and 300 hPa), as well as the precipitation and total cloud cover from these reanalysis products, were also used to diagnose the biases in the reanalysis PW.

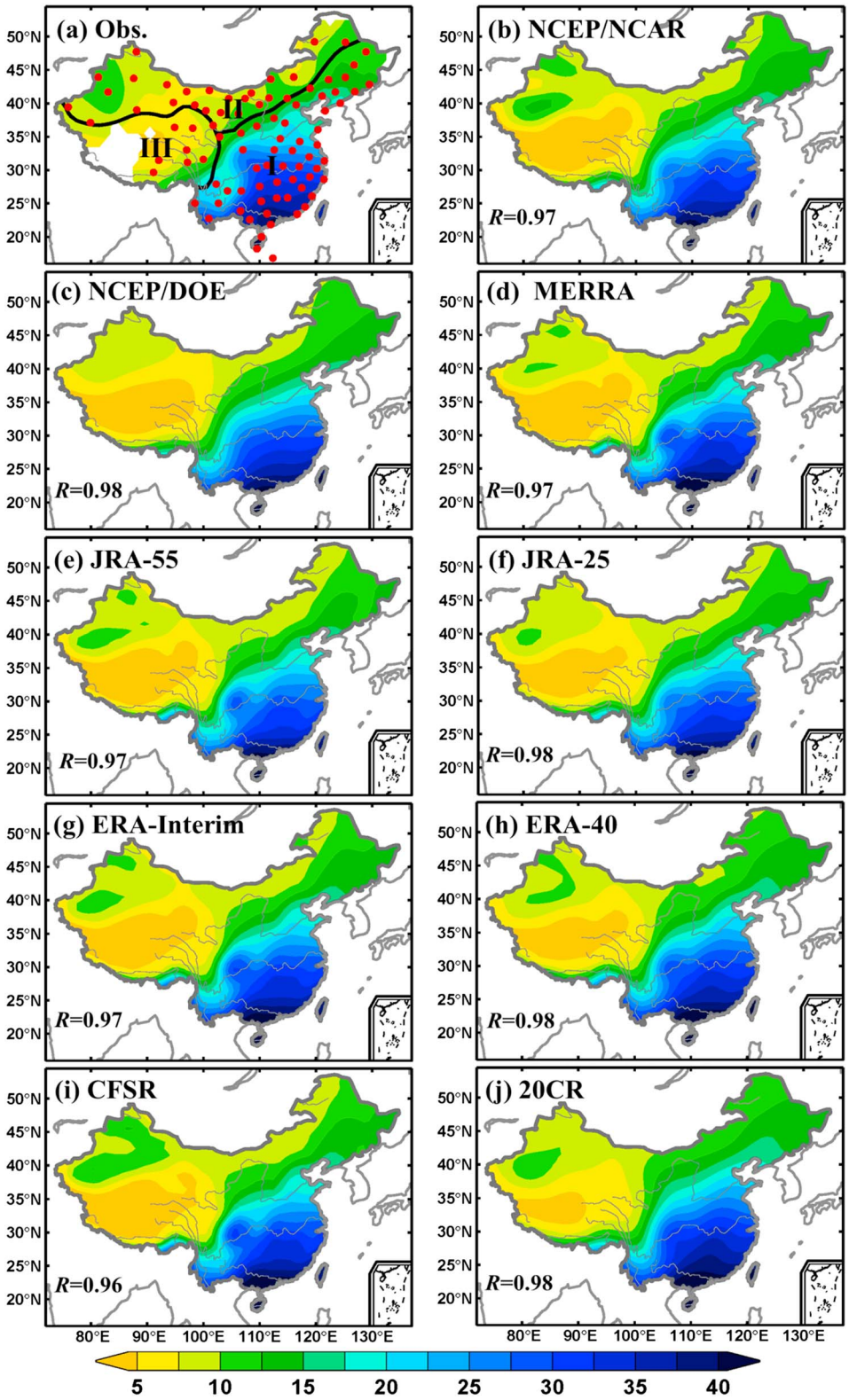

Figure 1. (a–j) The climatological annual mean PW (surface to 300 hPa, mm per month) from the observations and the reanalyses during 1979–2012 (1979–2001 for the ERA-40 in hour). The red dotted circles in Figure 1a indicate the 89 radiosonde stations used in this study, and three subregions are referred to region I as "southeastern China," region II as "northern China," and region III as "southwestern China" (i.e., the Tibetan Plateau with elevation over 3 km). The pattern correlations (*R*) between the reanalysis and the observations (Figure 1a) is shown in Figures 1b–1j.

Table 1. The Nine Reanalyses Used in This Study

Name	Organization	Horizontal resolution	Vertical levels	Temporal coverage	Assimilation
NCEP/NCAR	NCEP/NCAR	2.5° × 2.5°	28	1948 to present	3-D-VAR
NCEP/DOE	NCEP/DOE	2.5° × 2.5°	28	1979 to present	3-D-VAR
MERRA	NASA GMAO	0.5° × 0.67°	72	1979 to present	3-D-VAR, with incremental update
JRA-55	JMA/CRIEPI	1.25° × 1.25°	60	1958 to present	3-D-VAR
JRA-25	JMA/CRIEPI	2.5° × 2.5°	40	1979 to present	3-D-VAR
ERA-Interim	ECMWF	0.7° × 0.7°	60	1979 to present	4-D-VAR
ERA-40	ECMWF	1.125° × 1.125°	60	1957–2002	3-D-VAR
CFSR	NCEP	0.5° × 0.5°	64	1979 to present	3-D-VAR
20CRv2	NOAA/ESRL PSD	2.0° × 2.0°	28	1871–2012	Ensemble Kalman filter

Although each reanalysis was developed to meet specific goals with distinct model physics and resolutions, nearly all of the reanalyses assimilated upper air radiosonde and satellite (mainly after 1979) observations using either a 3-D or 4-D variational assimilation techniques with a constant error covariance matrix for the first-guess fields throughout the reanalysis period. One exception is the 20CR, which assimilated only surface observations of sea surface temperatures (SST), sea level pressure, and sea ice coverage using an ensemble Kalman filter system but no upper air or satellite observations were used [*Compo et al.*, 2011]. Some reanalyses, such as the ERA-Interim, used surface observations of 2 m temperature and humidity in their post processing [*Simmons et al.*, 2010], but most reanalyses (e.g., NCEP/NCAR and NCEP/DOE) did not use surface synoptic observations (of air temperature, humidity, winds, precipitation, and cloudiness). The ERA-Interim employs a 4-D -variational approach; all the others use 3-D -variational data assimilation. All of the reanalyses are forced with observed SSTs with the notable exception of the CFSR, which is a fully coupled land-ocean-atmosphere reanalysis. The newer reanalyses (ERA-interim, MERRA, CFSR, and 20CR) usually include recent advances in modeling and data assimilation; thus, they are expected to be superior to the older ones.

Table 1 summarizes the main features of the nine reanalyses. The 20CRv2 covers the longest period from 1871 to present. The NCEP/NCAR is one of the most widely used reanalyses, and it covers a relatively long period (1948 to present). Other long reanalyses are the ERA-40 (1957–2002) and the JRA-55 (1958 to present). The remaining five reanalyses (NCEP/DOE, JRA-25, ERA-interim, MERRA, and CFSR) all cover the so-called satellite era from 1979 to present. The CFSR has the highest resolution with a grid size of 0.5°, followed by MERRA (0.5° × 0.67°). The NCEP/NAR, NCEP/DOE, and JRA-25 have the lowest resolution with a grid size of 2.5°. To ensure a consistent length of analysis for all the products, we focus on the period from 1979 to 2012 for all the reanalyses, except for ERA-40 for which the period of 1979–2001 is used because it ends in August 2002.

3. Analysis Method

3.1. Data Processing

Following *Zhao et al.* [2012], we required a sampling rate of 50% or higher (i.e., at least half of the days with observations) in deriving the monthly mean value for individual months at the 89 radiosonde stations. Monthly anomalies for the PW and other variables relative to the 1979–2012 and 2003–2012 mean were computed for each month. The monthly anomalies computed as deviations from the 2003–2012 mean were used to illustrate the PW biases associated with the latest instrumental changes taken place in China since about 2002. The reanalysis products were processed similarly, except for the ERA-40 for which the mean of 1979–2001 was used.

To facilitate the comparison, both the monthly mean and anomalies from the radiosonde stations and reanalyses were regridded onto a common 1° × 1° latitude-longitude grid using the Cressman interpolation [*Cressman*, 1959] for the station data and bilinear interpolation for the reanalysis data. We used a maximum search radius of 1000 km in the Cressman method; a station beyond this distance was not used in deriving the estimates for a given grid box. Such a radius represents the typical correlation distance in monthly air temperature and humidity fields and thus is commonly used in gridding temperature and humidity monthly

anomaly data [e.g., *Dai et al.*, 2011]. We emphasize that observations are sparse over some areas, such as the western Tibetan Plateau (Figure 1a); thus, the interpolated PW values are less reliable over these regions.

3.2. Basic Statistics

Using the data on the 1°×1° grid, some basic statistics were calculated to evaluate the performance of each reanalysis product in describing the observed PW climatology, variations, and long-term changes over China. Besides the long-term mean and linear correlations, we calculated the relative bias (RB) and the root-mean-square error (RMSE).

The RB (in % of the observed long-term mean) is defined to assess the relative difference between a reanalysis and the observation averaged during the analysis period at each grid box:

$$\mathrm{RB} = \frac{\frac{1}{n}\sum_{i=1}^{n}\mathrm{PW}_R - \frac{1}{n}\sum_{i=1}^{n}\mathrm{PW}_O}{\frac{1}{n}\sum_{i=1}^{n}\mathrm{PW}_O} \times 100\%, \tag{1}$$

where PW_R and PW_o denote the monthly reanalysis and observed PW, respectively, and n is the number of data points during the time period.

The RMSE measures the average magnitude of the error with a focus on extreme values.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i-1}^{n}(\mathrm{PW}_R - \mathrm{PW}_O)^2}. \tag{2}$$

To minimize the influence of autocorrelation and outliers in data series on long-term trend estimation, the PW trend was estimated using the Sen's nonparametric method [*Sen*, 1968] as

$$b = \mathrm{median}\left(\frac{x_j - x_i}{j - i}\right), \tag{3}$$

where b is an estimate of the slope of a trend and x_i and x_j are the data values at time i and j ($j > i$), respectively. For a relatively small sample ($N < 50$), this approach provides a more robust slope estimate than the least squares method because it is insensitive to outliers or extreme values and competes well against simple least squares even for normally distributed data [*Fan and Yao*, 2003].

The statistical significance of a trend for monthly or annual series at each grid was tested using the Mann-Kendall Tau nonparametric technique [*Mann*, 1945; *Kendall*, 1975]. The Mann-Kendall statistic (S) was calculated as

$$S = \sum_{i-1}^{n-1}\sum_{j=i+1}^{n}\left(\mathrm{sgn}\left(x_j - x_i\right)\right) \tag{4}$$

where n is the number of data points and $\mathrm{sgn}(x_j - x_i)$ is the sign function:

$$\mathrm{sgn}\left(x_j - x_i\right) = \begin{cases} 1, & \text{if } x_j - x_i > 0 \\ 0, & \text{if } x_j - x_i = 0 \\ -1, & \text{if } x_j - x_i < 0 \end{cases}. \tag{5}$$

The S has a variance as

$$V(S) = \frac{n(n-1)(2n+5) - \sum_{k=1}^{m} t_k(t_k - 1)(2t_k + 5)}{18} \tag{6}$$

In equation (6), m is the number of tied groups and t_k denotes the number of ties for group k. A tied group is a set of sample data having the same value. In cases where the sample size $n > 10$, the standard normal test statistic Z_S is computed using equation (6):

$$Z_S = \begin{cases} \dfrac{S-1}{\sqrt{V(s)}}, & \text{if } S > 0 \\ 0, & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{V(s)}}, & \text{if } S < 0 \end{cases}. \tag{7}$$

Positive and negative values of Z_S indicate increasing trends and decreasing trends, respectively. At the 5% and 1% significance level, the null hypothesis of no trend is rejected if $|Z_S| > 1.96$ and $|Z_S| > 2.576$, respectively.

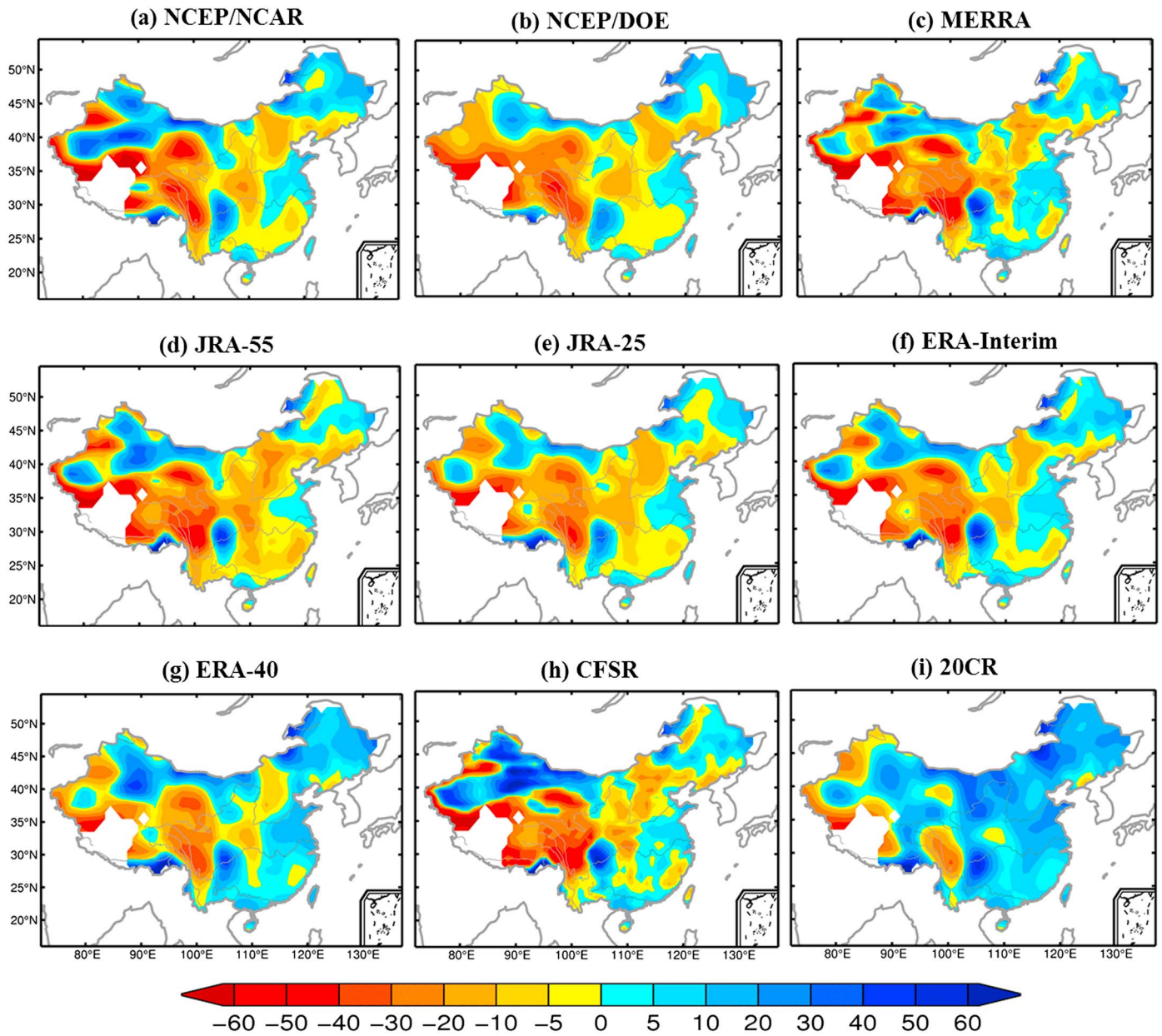

Figure 2. (a–h) The relative bias (in % of the observed mean PW) of the monthly PW from the reanalyses compared with the radiosonde observations during 1979–2012 (1979–2001 for the ERA-40 case in Figure 2g).

4. Results

4.1. Bias and Error Assessments

Figure 1 shows the spatial distribution of the climatological annual mean PW from the radiosonde observations and the reanalyses during 1979–2012 (1979–2001 for the ERA-40 case) over China. It can be seen that most reanalyses reproduce well the observed mean PW and its spatial variations over China, especially over eastern China. The pattern correlation between the reanalyses and the observations is 0.96 or higher. The RB (relative bias) of the reanalysis PW compared with the radiosonde PW during 1979–2012 (1979–2001 for the ERA-40) is shown in Figure 2. For most of the reanalyses, the RB is within ~10% over most northern and eastern China, but many of the reanalyses underestimate the observed PW by ~20–60% over most western China. In particular, the RB reaches −40 to −60% over the western Tibetan Plateau, where the large negative RB results partly from the fact that the radiosonde PW was interpolated from the surrounding stations with lower elevations (thus with a deeper atmosphere and higher PW). This suggests that the elevation difference between the

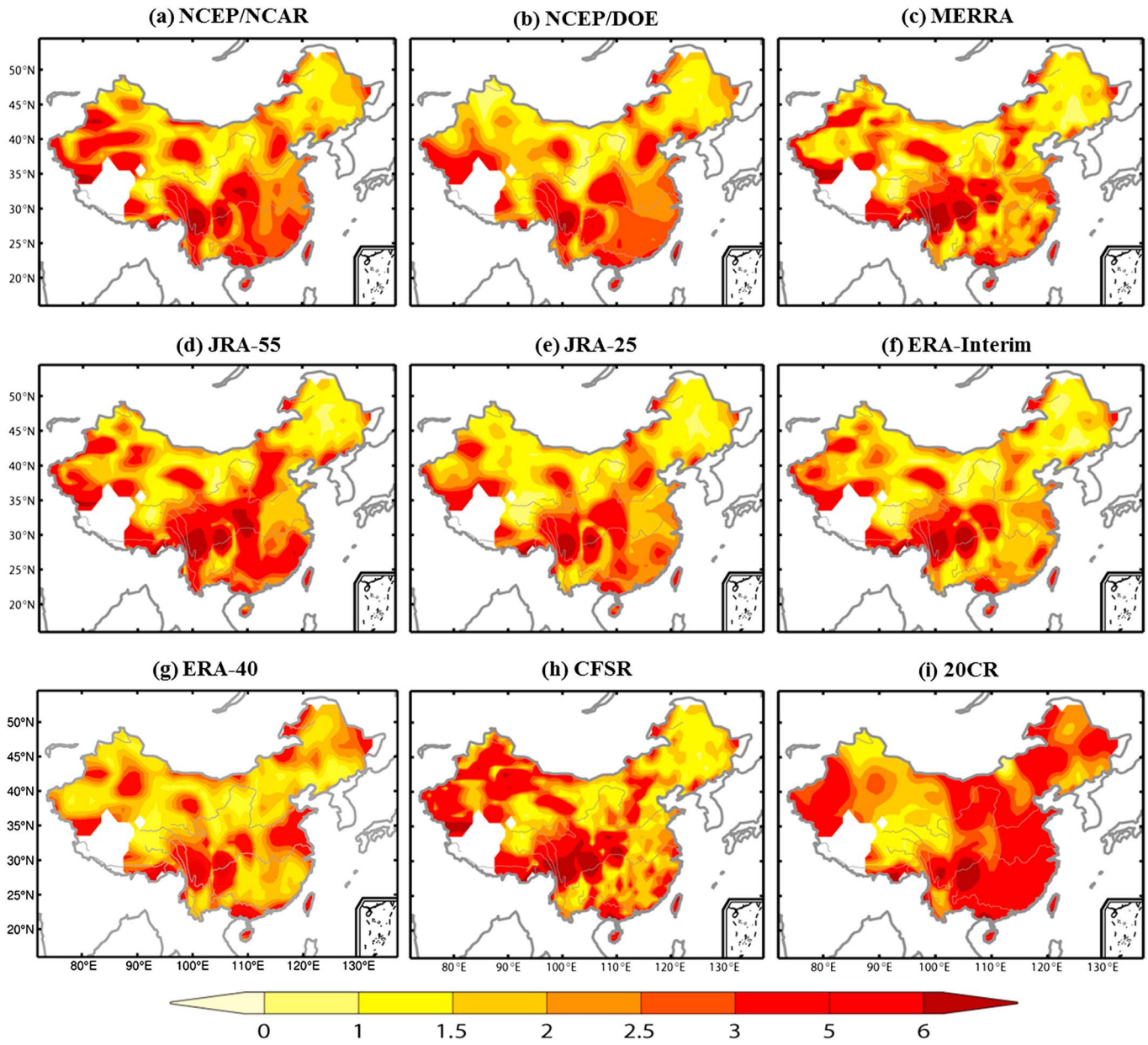

Figure 3. (a–h) The root-mean-square error (RMSE) of reanalysis monthly PW (mm) relative to the PW derived from radiosonde observations over China during 1979–2012 (1979–2001 for Figure 3g).

reanalyses and the radiosonde stations can induce a large RB in the long-term mean PW. The 20CR PW is about 5%–20% higher than the radiosonde PW over most of China except the western and eastern parts of the Tibetan Plateau (Figure 3i), while the CFSR overestimates the PW by 20%–60% over Xinjiang province (the Northwest) of China (Figure 3h). The ERA-40 also shows positive RB (5–20%) over many areas outside the Tibetan Plateau (Figure 3g). Although the CFSR, MERRA, and ERA-Interim have higher spatial resolution than the others (Table 1), their PW shows RBs with spatial patterns and magnitude similar to the other products. Although the large dry biases over the Tibetan Plateau may partly result from the elevation difference mentioned above, the relative biases over the rest of the country are likely due to systematic errors induced mainly by model and observational biases in the reanalysis products. Figure 2 suggests that despite of using improved assimilation systems the newer-generation reanalyses do not clearly show reduced biases in their mean PW over China.

The RMSE (Figure 3) of the reanalysis PW with respect to the radiosonde observations is about 1–3 mm for most of northern China (north of ~35°N), but it is over 3 mm in most northwestern and southwestern China,

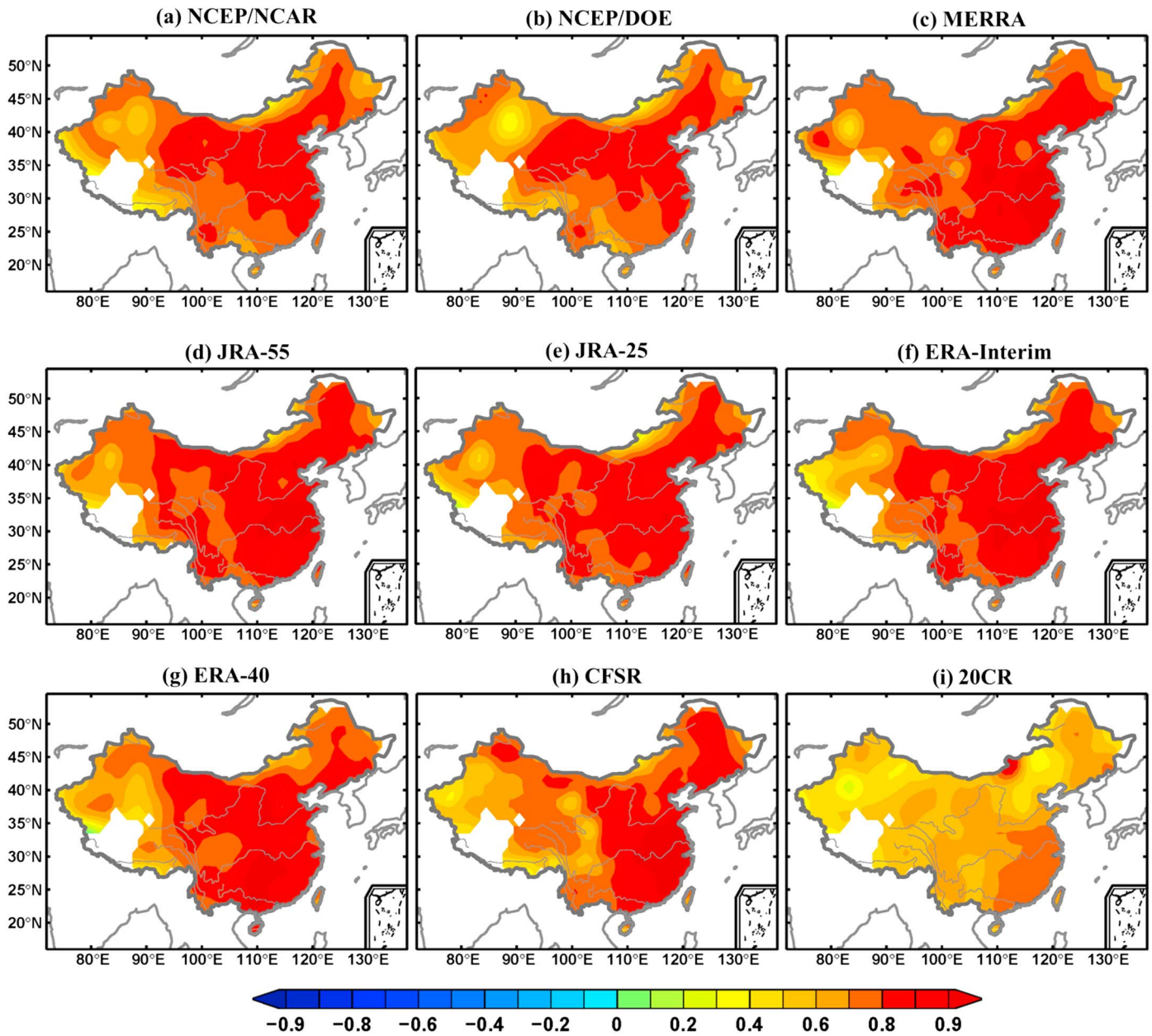

Figure 4. (a–h) Spatial distributions of the correlation coefficient between the observed and reanalysis monthly PW anomalies during 1979–2012 (1979–2001 for Figure 4g) over China. Essentially, all the areas are statistically significant at the 5% level.

where the radiosonde PW is less reliable due to sparse observations there. In particular, the RMSE reaches 6 mm or more over the Sichuan Basin and southwestern part of the Tibetan Plateau. The NCEP/NCAR and NCEP/DOE, two earlier-generation reanalyses, and the JRA-55 show larger RMSE (>2 mm) over most southeastern China than the other products, while the 20CR (Figure 3i) exhibits a RMSE over 2.5 mm over eastern China (east of 100°E.) and northwestern part of the Tibetan Plateau. The large RMSE in the 20CR PW is likely because of no humidity, and other upper air observations were assimilated into this product [*Compo et al.*, 2011]. The newer reanalysis products (ERA-interim, MERRA, and CFSR) generally show smaller RMSE than the other products (Figures 3c, 3f, and 3h).

4.2. Correlation and Trend Analysis

Figure 4 shows the correlation maps between the observed and reanalysis monthly PW anomalies during 1979–2012 (1979–2001 for ERA-40). It can be seen that most reanalysis products, except the 20CR, are highly

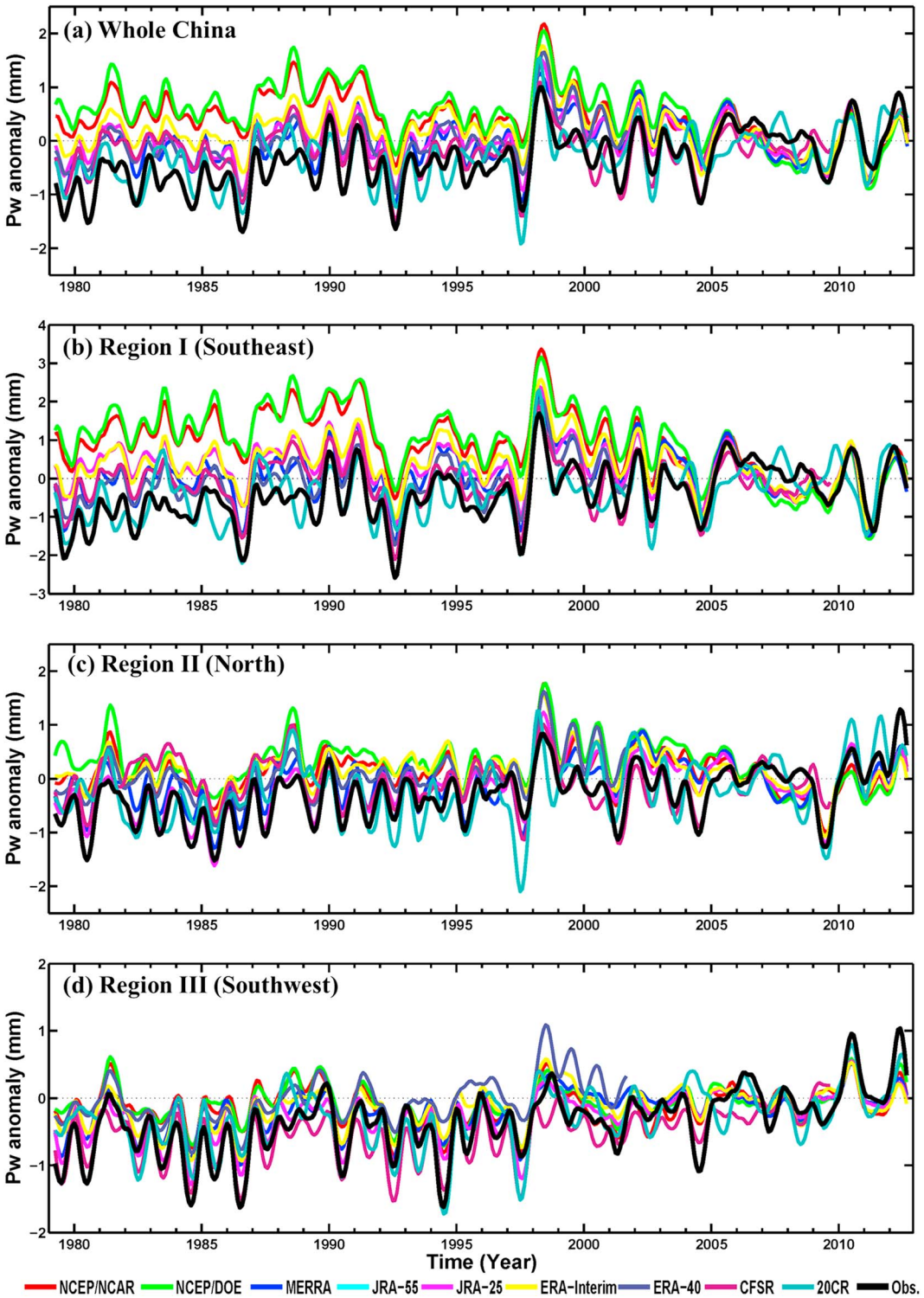

Figure 5. Time series of 9-point-moving-averaged monthly PW anomalies (mm, relative to the 2003–2012 mean, but the 1979–2001 mean for ERA-40) averaged over (a) whole China and the (c and d) three subregions outlined in Figure 1a from 1979 to 2012. The year label on the *x* axis marks the start?? of the year in this and all other time series plots of this paper.

Table 2. The Correlation Coefficients of Regional-Mean PW, Temperature (*T*), and Relative Humidity (RH) Anomalies Between the Reanalyses and Observations During 1979–2012 (1979–2001 for ERA-40)[a]

		NCEP/NCAR	NCEP/DOE	MERRA	JRA-55	JRA-25	ERA-Int.	ERA-40	CFSR	20CR
Nationwide	**PW**	**0.73**	**0.66**	**0.91**	**0.92**	**0.86**	**0.85**	**0.88**	**0.91**	**0.77**
	T	**0.96**	**0.95**	**0.98**	**0.98**	**0.97**	**0.97**	**0.95**	**0.93**	**0.88**
	RH	**0.60**	**0.67**	**0.76**	**0.82**	**0.68**	**0.75**	**0.82**	**0.60**	**0.54**
Region I (Southeast)	**PW**	**0.75**	**0.70**	**0.93**	**0.90**	**0.85**	**0.88**	**0.94**	**0.92**	**0.78**
	T	**0.98**	**0.98**	**0.99**	**0.99**	**0.99**	**0.99**	**0.94**	**0.98**	**0.88**
	RH	**0.58**	**0.58**	**0.79**	**0.80**	**0.60**	**0.78**	**0.88**	**0.77**	**0.70**
Region II (North)	**PW**	**0.79**	**0.68**	**0.89**	**0.91**	**0.87**	**0.82**	**0.81**	**0.83**	**0.63**
	T	**0.98**	**0.98**	**0.99**	**0.99**	**0.98**	**0.98**	**0.93**	**0.96**	**0.91**
	RH	**0.68**	**0.71**	**0.83**	**0.84**	**0.77**	**0.79**	**0.82**	**0.67**	**0.50**
Region III (Southwest)	**PW**	**0.78**	**0.84**	**0.82**	**0.83**	**0.84**	**0.82**	**0.73**	**0.77**	**0.69**
	T	**0.88**	**0.83**	**0.94**	**0.91**	**0.90**	**0.92**	**0.91**	**0.80**	**0.71**
	RH	**0.76**	**0.76**	**0.74**	**0.69**	**0.69**	**0.74**	**0.74**	**0.47**	**0.37**

[a]Numbers in boldface are statistically significant at the 5% level. The *T* and RH were averaged from 1000 to 300 mb using pressure as weighting.

correlated ($r > 0.7$) with the observations over most China, especially east of ~90°E. In particular, the MERRA, JRA-55, ERA-Interim, ERA-40, and the CFSR are more strongly correlated ($r \geq 0.9$) with the observations over many areas of eastern China than the other products. Again, the 20CR shows noticeably lower correlations as its humidity fields over land are not directly constrained by observations.

To examine the interannual variations and long-term changes in regional PW, Figure 5 shows the time series of monthly PW anomalies (relative to 2003–2012 mean, which is considered more reliable than older records) from the reanalysis products and the radiosonde observations averaged over the three subregions (cf. Figure 1a) and China as a whole. It is clear that the PW interannual variations in most reanalyses are highly consistent ($r = 0.63$–0.94; Table 2) with the observed PW series over China. The NCEP/NCAR, NCEP/DOE, and 20CR show lower performances than the other reanalyses in depicting the observed PW variations over northern and southwestern China (Table 2). Consistent with Figure 4, relatively low correlations between the reanalyses and the observations are seen over the Tibetan Plateau (i.e., region III; Table 2). For long-term changes, however, there are clear discrepancies between the reanalyses and the observations before around 2003 due to wet biases in the reanalyses in the earlier period. This is especially true over southeastern and northern China (Figures 5b and 5c). This wet bias in monthly mean PW results from the fact that the reanalyses assimilated the unhomogenized radiosonde records that contain a wet sampling bias before around 2003 due to a lack of sampling of dry conditions by and a slow response time of older humidity sensors [*Dai et al.*, 2011]. These biases have been corrected in the homogenized data used here. Some studies [e.g., *Wang and Zhang*, 2008; *Bian et al.*, 2010] also confirmed that systematic replacement of older Shang-M (slow response with wet biases) by newer Shang-E humidity sensors has taken place in Chinese radiosonde network around the year of 2002, which can lead to a spurious downward trend in tropospheric humidity at most Chinese stations. The use of the 2003–2012 mean as the baseline forces all the lines in Figure 5 (and Figures 7 and 8 shown below) to have the same zero mean and thus creates an apparent better match during 2003–2012 than during earlier years. The large spread among the lines before 2003 may also give a false impression that the variations for individual lines are larger before 2003 than during the later period in these figures.

The spatial patterns of the long-term PW trends from the reanalyses and observations during the study period are shown in Figure 6. It can be seen that none of the reanalyses, besides the ERA-40 and 20CR, can reproduce the observed trend pattern, and the reanalyses themselves differ substantially among them. In contrast to the large upward trends in the observations, spurious large downward trends over Southeast China are seen in the NCEP/NCAR, NCEP/DOE, JRA-55, and the JRA-25. The newer-generation reanalyses (the ERA-Interim, JRA-55, and MERRA) also perform poorly in capturing the long-term trends, even though they do well in depicting the observed PW variations over China (Figures 4 and 5). The ERA-40 and 20CR roughly capture the observed PW trend patterns with comparable magnitudes of upward trends over most China (with a significant pattern correlation of 0.77), and for China as a whole and for the subregions (Table 3). The good performance of the 20CR is partly because this reanalysis assimilated only the observations of surface pressure and SSTs [*Compo et al.*, 2011], and thus, it is not affected by the inhomogeneities in radiosonde and satellite records.

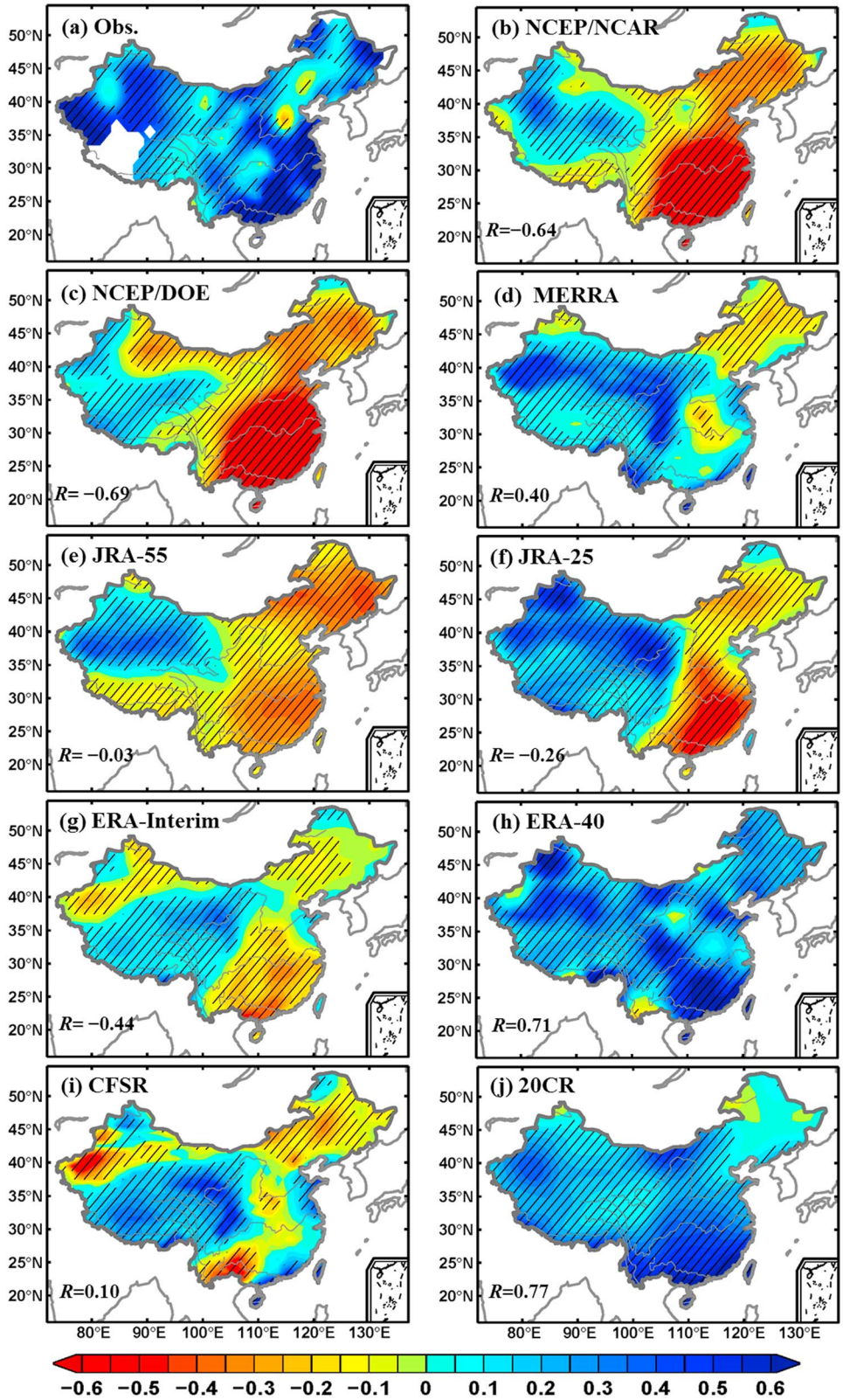

Figure 6. (a–j) Maps of linear trends (mm decade^{-1}) for monthly PW anomalies estimated from the observations and the reanalyses over China during 1979–2012 (1979–2001 for the ERA-40 in Figure 6h). ***R*** in Figures 6b–6j is the pattern correlations between the reanalyses and the observations in Figure 6a.The stippling indicates that the trend is statistically significant at the 5% level.

Table 3. The PW Liner Trends (mm decade^{-1}) Derived From Regional-Mean Monthly Anomalies for the Reanalyses and the Observations[a]

	Obs.	NCEP/NCAR	NCEP/DOE	MERRA	JRA-55	JRA-25	ERA-Int.	ERA-40	CFSR	20CR
Nationwide	**0.33**	**−0.23**	**−0.28**	**0.11**	0.05	0.01	−0.03	**0.25**	0.01	**0.23**
Region I (Southeast)	**0.38**	**−0.49**	**−0.56**	**0.07**	**−0.15**	**−0.21**	**−0.12**	**0.31**	−0.03	**0.31**
Region II (North)	**0.30**	−0.02	**−0.10**	**0.15**	**0.25**	**0.20**	−0.01	**0.24**	**−0.06**	**0.20**
Region III (Southwest)	**0.27**	**0.04**	**0.04**	**0.16**	**0.19**	**0.20**	**0.13**	**0.24**	**0.20**	**0.18**

[a]The time periods are the same as Table 2. Numbers in boldface are statistically significant at the 5% level.

In summary, most reanalysis products can depict observed interannual PW variations over eastern and northern China but show relatively limited performance over western China, especially over the Tibetan Plateau. Few reanalyses besides the ERA-40 and 20CR can capture the observed long-term PW trends, and their PW trends differ substantially from each other. This poor performance results from the spurious changes in the raw humidity records from radiosondes and the varying role of satellite observations in individual reanalyses. In general, the newer-generation reanalyses (e.g., ERA-40, ERA-Interim, JRA-25, JRA-55, MERRA, and CFSR) perform relatively better than the earlier ones (e.g., the NCEP/NCAR and NCEP/DOE) in describing the observed PW variations over China, but the 20CR and ERA-40 are the only products that broadly capture the observed PW trends over China since the 1979.

4.3. Contributions of T and RH to PW Biases

Since specific humidity (q) is equal to saturation specific humidity ($q_s(T)$) times RH, the long-term PW change can be attributed to two components associated with air temperature (T) and RH changes. The $q_s(T)$ is expected to increase at approximately 7% per 1 K rise in T following the Clausius-Clapeyron equation [*Trenberth et al.*, 2003, 2005], while RH changes are often small [*Dai*, 2006; *Zhao et al.*, 2012]. Some previous studies have examined the relationship between tropospheric water vapor content and air temperature [*Gaffen et al.*, 1992; *Sun and Oort*, 1995; *Zhai and Eskridge*, 1997; *Wang and Gaffen*, 2001; *Ross et al.*, 2002] based on unhomogenized radiosonde data. *Zhao et al.* [2012] examined the relationships between the PW and air temperature as well as RH changes over China using the homogenized radiosonde data from *Dai et al.* [2011]. They found that PW variations and changes are highly correlated with middle-lower tropospheric mean temperature, with $r = 0.83$ and a slope of ~7.6% K^{-1}, while tropospheric RH only shows small variations and weak trends over China. Here we compare the PW versus T and PW versus RH relationships in the reanalyses with those in homogenized radiosonde data to further quantify any biases existed in the reanalysis products.

Figure 7 shows the time series of tropospheric mean monthly T and RH anomalies from the reanalyses and the radiosonde observations averaged over whole China and the three subregions. It can be seen that the variations and long-term changes in reanalysis T are highly consistent with the observations, especially over Southeast and North China (Figures 7c and 7e). However, the RH from the reanalyses agrees well with the observations only after about 2003, before which most of the reanalyses show dry biases, especially over Southeast China (Figure 7d). The RH biases are largely consistent with the PW biases shown in Figure 5, which suggests that most of the PW biases result from the biases in RH. Table 2 summarizes the correlations of T and RH between the reanalyses and the observations over whole China and the three subregions. It is clear that the correlations for T are higher than those for RH for all the regions, including the Tibetan Plateau (Southwest China). This implies that the reanalysis RH contains larger discontinuities than T records, especially since about 2003 (Figure 7).

Figure 8 shows the time series of national mean PW, T, and RH anomalies relative to the 2003–2012 mean shown in Figure 5a and Figures 7a and 7b. It can be seen that the PW is more consistent with the T than with the RH in the variations and long-term changes for the observations, whereas it is the opposite for most of the reanalyses. Except for the ERA-40, CFSR, and 20CR, all other reanalyses show that the PW decline since about 2005 agrees well with the RH decrease, especially in the NCEP/NCAR and NCEP/DOE. In contrast to the observations, the PW versus RH correlations are stronger than the PW versus T correlations in the NCEP/NCAR, NCEP/DOE, JRA-55, JRA-25, and ERA-Interim in Southeast and North China (Table 4).

Figures 9 and 10 compare the long-term trends of tropospheric mean T and RH monthly anomalies from the reanalyses and the observations. Most of the reanalyses show significant warming trends as seen in the observations over most China, although large differences exist in the spatial pattern and magnitude. For example, the NCEP/DOE shows cooling over parts of the Tibetan Plateau where the observations suggest

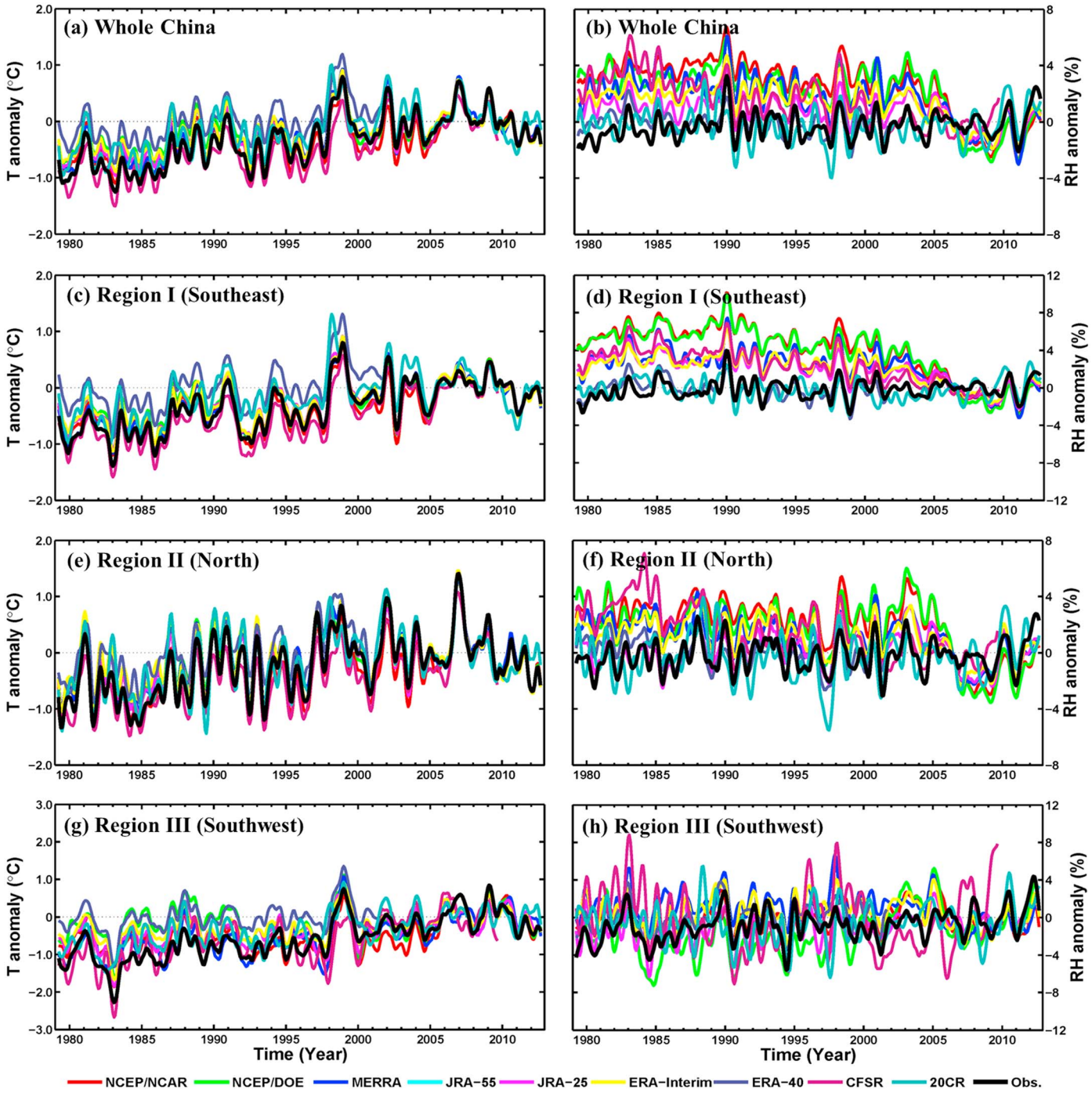

Figure 7. Same as Figure 5 but for temperature anomalies (*T* in °C, left column) and relative humidity anomalies (RH in %, right column) averaged from 1000 to 300 mb using pressure thickness for each level as weighting. The monthly *T* and RH anomalies are computed as deviations from the 2003–2012 mean (the 1979–2001 mean for ERA-40).

large warming. The CFSR and 20CR capture the observed warming patterns and magnitude better than the others (with the pattern correlations of 0.94 and 0.95, respectively). For RH, however, most reanalysis products except 20CR show a significant drying trend over most eastern China and parts of northern China (Figures 10b–10j), in contrast to the weak trends seen in the radiosonde data (Figure 10a). These RH trend patterns from the reanalyses are broadly comparable with their PW trends shown in Figure 6, with a pattern correlation over 0.6 for the NCEP/NCAR, NCEP/DOE, JRA-55, JRA-25, and ERA-Interim. This suggests that most

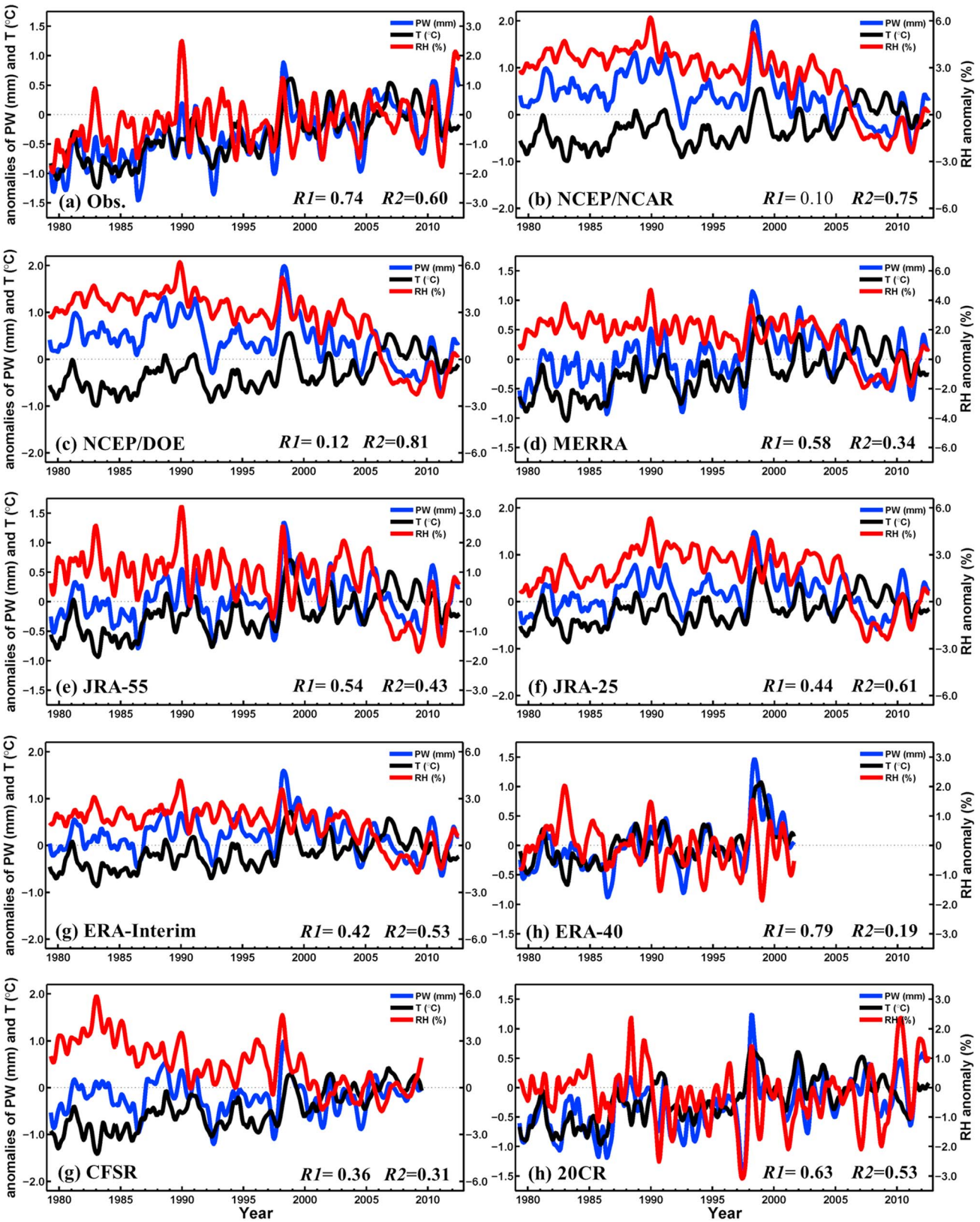

Figure 8. Time series of 9-point-moving-averaged PW anomalies (mm, blue line), the 1000–300 mb mean *T* (°C, black line) and RH anomalies (%, red line) for China. The R1 and R2 are the correlation coefficients between the *T* and PW, and between the RH and PW, respectively. These monthly anomalies are relative to the 2003–2012 mean (the 1979–2001 mean for ERA-40).

Table 4. The Correlation Coefficients of Regional-Mean Monthly Anomalies Between the *T* and PW and Between the RH and PW During 1979–2012 (1979–2001 for ERA-40)[a]

		Obs.	NCEP/NCAR	NCEP/DOE	MERRA	JRA-55	JRA-25	ERA-Int.	ERA-40	CFSR	20CR
Nationwide	***T* versus PW**	**0.72**	0.10	**0.13**	**0.57**	**0.54**	**0.44**	**0.42**	**0.78**	**0.35**	**0.62**
	RH versus PW	**0.61**	**0.75**	**0.80**	**0.36**	**0.44**	**0.61**	**0.53**	**0.21**	**0.32**	**0.53**
Region I (Southeast)	***T* versus PW**	**0.74**	**0.05**	**0.04**	**0.54**	**0.37**	**0.29**	**0.39**	**0.77**	**0.42**	**0.71**
	RH versus PW	**0.60**	**0.81**	**0.84**	**0.38**	**0.55**	**0.65**	**0.58**	**0.27**	**0.32**	**0.40**
Region II (North)	***T* versus PW**	**0.56**	**0.32**	**0.23**	**0.45**	**0.49**	**0.41**	**0.33**	**0.63**	**0.10**	**0.31**
	RH versus PW	**0.53**	**0.59**	**0.76**	**0.38**	**0.56**	**0.57**	**0.51**	**0.16**	**0.51**	**0.64**
Region III (Southwest)	***T* versus PW**	**0.59**	**0.66**	**0.48**	**0.68**	**0.64**	**0.58**	**0.64**	**0.61**	**0.44**	**0.60**
	RH versus PW	**0.66**	**0.46**	**0.45**	**0.29**	**0.59**	**0.68**	**0.43**	**0.46**	**0.36**	**0.69**

[a]Numbers in boldface are statistically significant at the 5% level. The *T* and RH were averaged from 1000 to 300 mb using pressure as weighting.

of the PW changes can be explained by the RH changes in most of the reanalyses, rather than due to *T* changes as in the observations and 20CR. The 20CR stands out with weak RH trends that are comparable to the observations (Figure 10j). We also examined these relationships at some individual pressure levels in the middle-lower troposphere and found similar results.

To further illustrate the contributions of *T* and RH changes to the PW differences between the reanalyses and the observations, Figure 11 shows the difference series of PW, *T*, and RH anomalies between the reanalyses and the observations averaged over China. It is clear that the long-term changes of the PW difference series come mainly from those of the RH difference series for all the reanalyses, with a positive correlation of 0.65–0.98 between them.

4.4. EOF Analysis

To investigate how well the spatiotemporal variations of PW is captured in the reanalyses, we performed an empirical orthogonal function (EOF) analysis of the monthly PW anomalies (normalized by local standard deviation) to identify and compare the leading modes of variability. Figures 12 and 13 show the spatial patterns and the principal components (PCs) for the two leading EOFs, which together explain nearly 50% of the total variance for both the observations and reanalyses. Figure 12 shows that the first EOFs from the reanalyses agree well with EOF 1 from the observations, with a pattern correlation of 0.73–0.88 and a temporal correlation of 0.18–0.77. This EOF represents a pattern with the same sign over whole China, with the largest contribution from central East China. Up to around year 2005, the PCs show an upward trend for the observations and most of the reanalyses. After 2005, all the reanalyses, except the 20CR and CFSR, show decreases (Figure 12j). As mentioned above, the recent changes in radiosonde types in the early 2000s in China induced a dry biases (relative to previous records) in radiosonde humidity data for the recent years, and this change appears to have had large impacts on the PW of the last 10 years in many of the reanalyses, including NCEP/NCAR, NCEP/DOE, MERRA, JRA-25, JRA-50, and ERA-Interim (Figure 12j). The MERRA performed best among the reanalyses in reproducing EOF 1 (a long-term change mode), with a pattern correlation of 0.88 and temporal correlation of 0.77. Seasonal analyses revealed (not shown) that most reanalyses seem to perform better in winter than in summer in reproducing the first EOF.

The PC1 time series of the observed PW (Figure 12j) shows apparent stepwise jumps around 1986/1987 and 1997/1998. To a lesser degree, these apparent jumps are also evident in the PW time series for whole China in Figure 5a, and also in *Zhao et al.* [2012, Figure 14a]. These and other PW changes are actually closely associated with similar changes in surface (T_s) and lower tropospheric (T_m) air temperature changes, rather than the homogenized RH from *Dai et al.* [2011] (Figures 7a and 7b, also see Figures 14a and 14b of *Zhao et al.* [2012]). It is possible that the homogenized T_m data from *Haimberger et al.* [2008] (used by *Dai et al.* [2011]) might still contain some discontinuities that could be propagated into the PW (which was derived using the T_m and RH data). However, Figure 14a of *Zhao et al.* [2012] shows that the T_m (from radiosonde observations) and T_s (from surface station observations) over whole China are quite consistent with each other. This suggests that most of the T_m variations (including those around 1986/1987 and 1997/1998) are likely to be real. Thus, we believe that the whole China PW changes shown in Figure 5a of this paper and Figure 14a of *Zhao et al.* [2012] are likely to be real, and the changes around 1997/1998 is associated with the strong El Niño at that time.

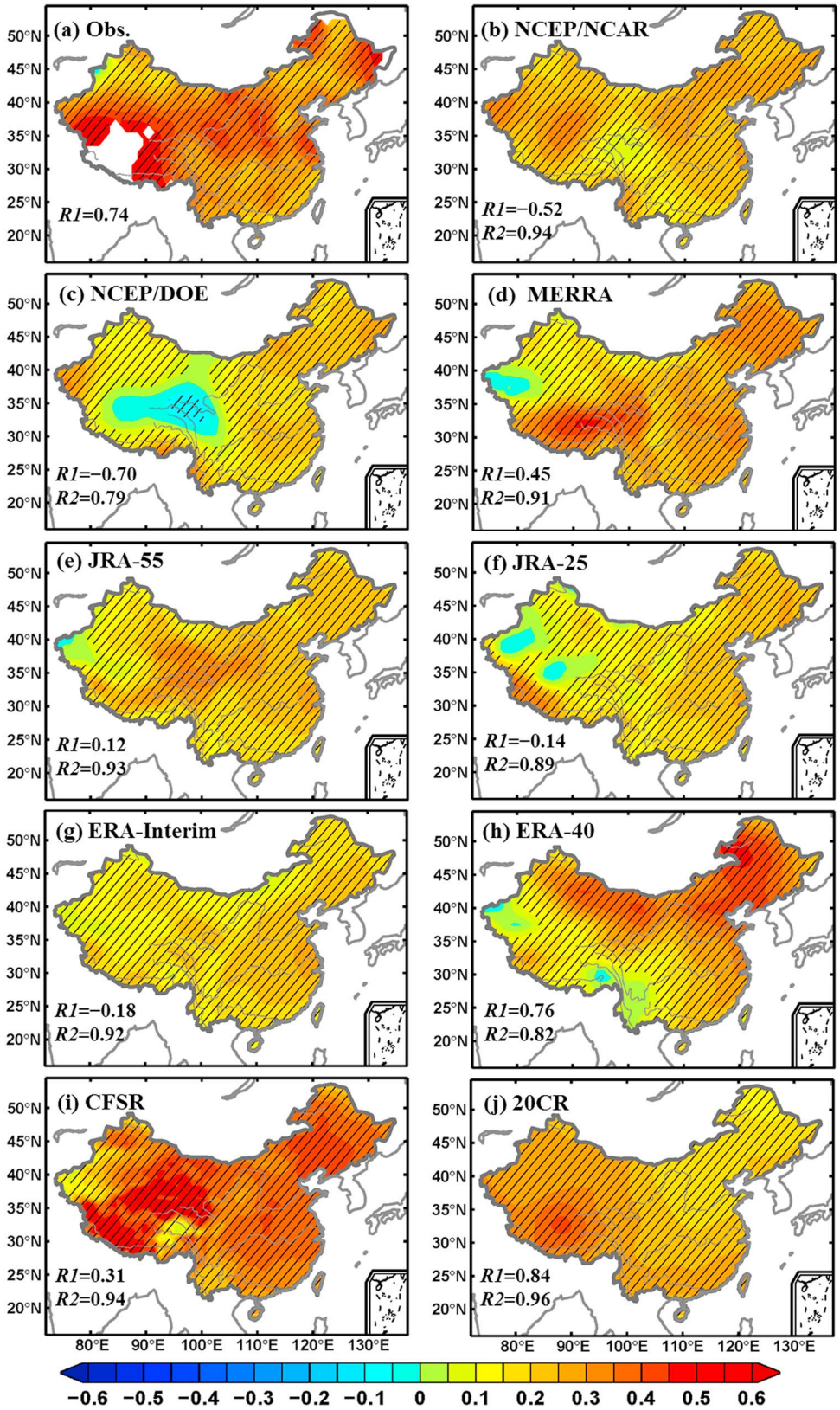

Figure 9. (a–j) Same as Figure 6, but for the linear trends of tropospheric mean T monthly anomalies (°C decade^{-1}) averaged from surface to 300 mb. R1 in each panel is the pattern correlation of the *T* trend with that of the PW trend shown in Figure 6, and R2 in Figures 9b–9j is the pattern correlation of the T trends between the reanalysis and the observations shown in (a).

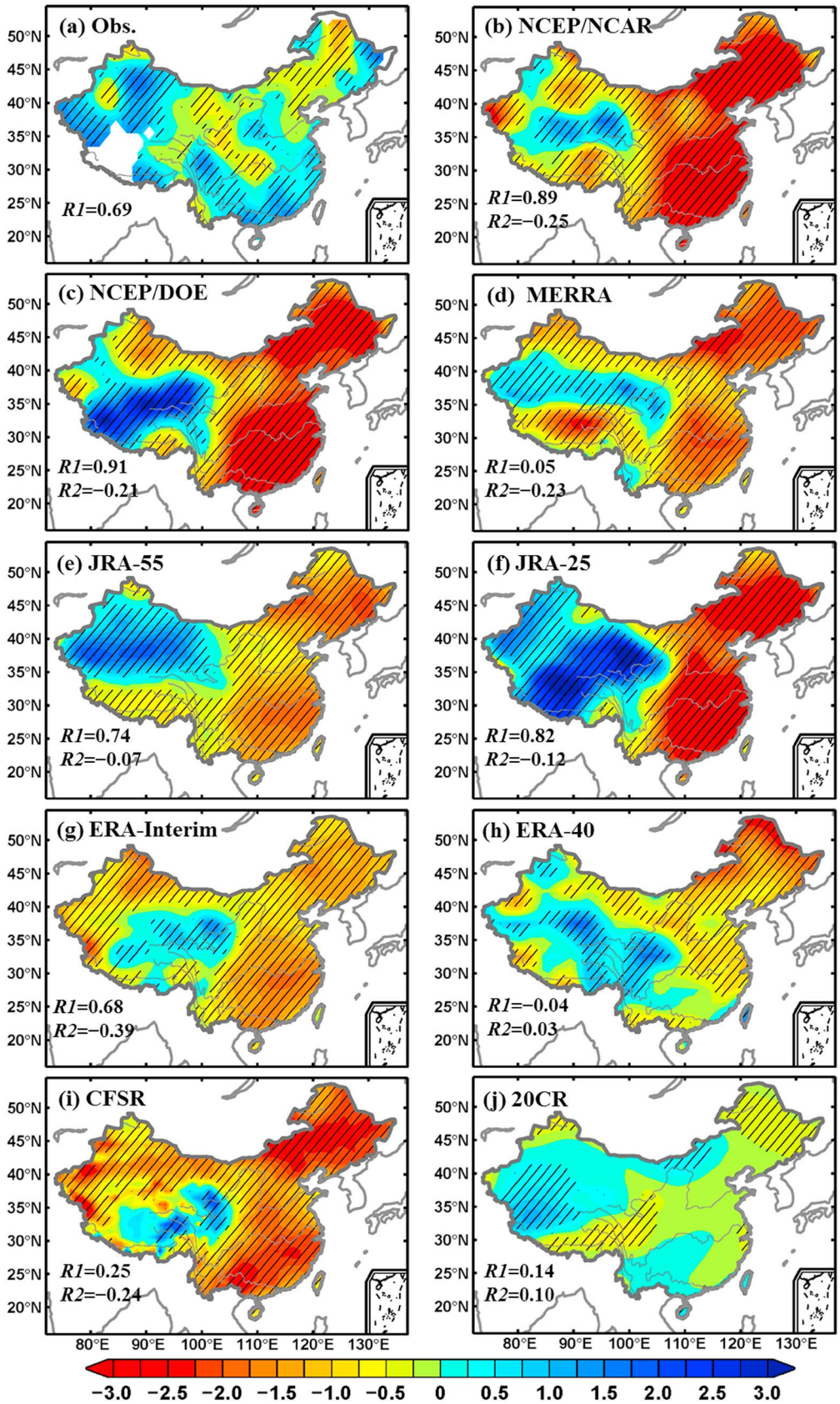

Figure 10. As in Figure 9, but for the RH linear trends (% decade^{-1}).

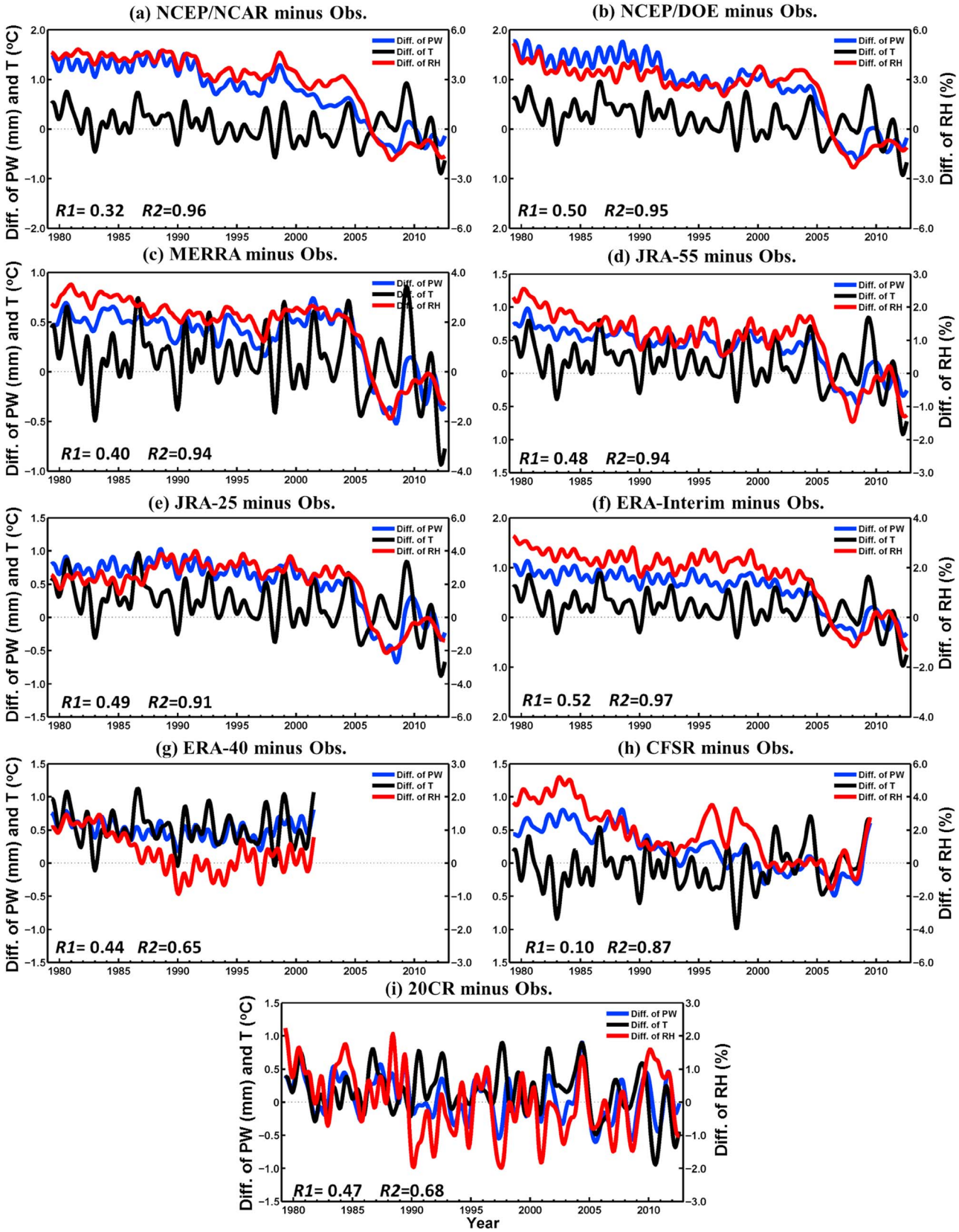

Figure 11. Difference series (9-point-moving-averaged) of PW anomalies (dPW, mm, blue line), *T* anomalies (dT, °C, black line), and RH anomalies (dRH, %, red line) between the reanalyses and the observations for China. The R1 (R2) is the correlation coefficient between the dT and dPW (dRH and dPW).

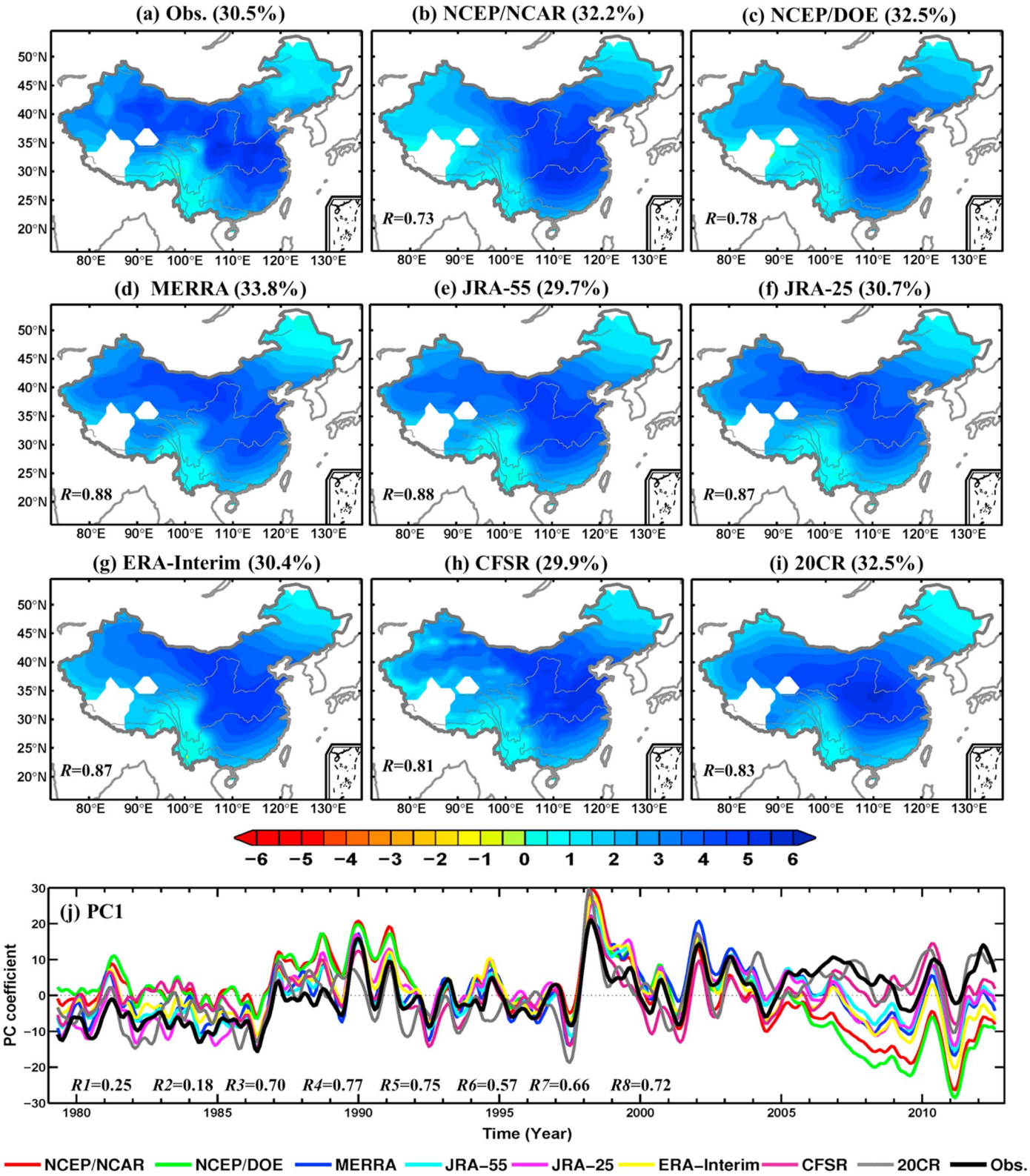

Figure 12. (a–i) The first leading EOF and (j) the corresponding 7-point-moving averaged PC time series derived from observed and reanalysis PW monthly anomalies during 1979–2012. The monthly PW anomalies were normalized by local standard deviation and multiplied by the square root of cosine of the latitude at each grid box before the EOF analysis. The explained percentage variance is also shown in parentheses on top of Figures 12a–12i. The pattern correlation (R) between the reanalysis and the observations (Figure 12a) is shown in Figures 12b–12i, and the PC correlation (R1–R8) between the observation and corresponding reanalysis is also shown in Figure 12j.

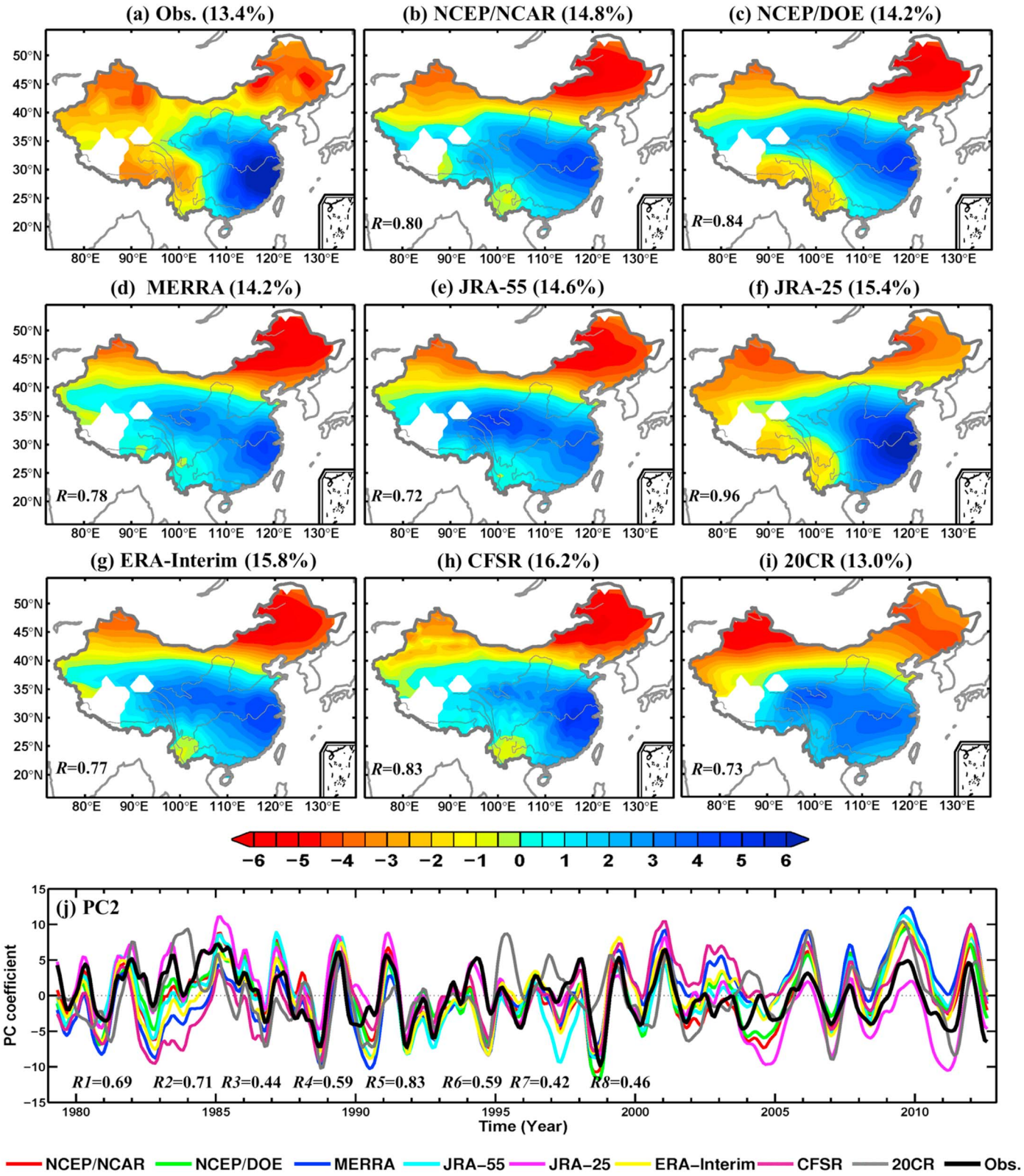

Figure 13. (a–j) Same as in Figure 12, but for the second leading EOF and its PC time series. The black line (Obs.) in Figure 13j has a maximum correlation of ~0.3 with the Niño3.4 SST index when the PC lags the SST index by 9 months.

The second EOF of PW in the observations shows a dipole (i.e., anticorrelated) mode between Southeast and the rest of China, with the temporal coefficient (PC2) showing mostly multiyear variations (Figures 13a and 13j). This observed mode is broadly comparable to the second EOF of the normalized PW anomalies (in %) shown in Figure 15 of *Zhao et al.* [2012]. It has a weak correlation of 0.3 with the El Niño–Southern Oscillation (ENSO) with a nine month lag by PC2. This mode suggests that during warm El Niño events, water vapor content tends to be above normal over Southeast China and below normal over the rest of China. The EOF 2 pattern is linked to a large-scale dipole pattern of anomaly vertical motion [*Zhao et al.*, 2012]. The corresponding EOF 2 derived from the reanalyses roughly captures the spatial and temporal variations represented by the EOF 2 from the observations, with significant pattern correlations of 0.72–0.96 with Figure 13a. The PC2 from the reanalyses exhibit similar multiyear variations as seen in the PC2 of the observations with significant correlations of 0.42–0.83 (Figure 13j). Among the reanalyses, the JRA-25 performed best in describing the spatiotemporal variations revealed by EOF 2 from the observations, with a strong pattern correlation of 0.96 and temporal correlation of 0.83.

5. Summary and Concluding Remarks

In this study, the interannual variations and long-term changes of precipitable water (PW) derived from nine reanalysis products (i.e., NCEP/NCAR, NCEP/DOE, MERRA, JRA-55, JRA-25, ERA-Interim, ERA-40, CFSR, and 20CR) are evaluated against the homogenized radiosonde data over China from 1979 to 2012 (1979–2001 for ERA-40). The possible causes for the PW biases in the reanalyses are also discussed by analyzing the tropospheric mean air temperature (*T*) and relative humidity (RH). The main findings are summarized below.

Results show that most of the reanalyses reproduce well the observed PW climatology (i.e., mean values and their spatial variations) over China, with a pattern correlation of 0.96 or higher. Their relative biases are within ~20% over most of China, but can be 20%–40% or higher over western China, such as the Tibetan Plateau, the Taklimakan desert, and the Sichuan Basin. The largest dry bias (>60%) occurs over the northwestern part of the Tibetan Plateau, partly due to the elevation differences between the radiosonde stations and reanalysis grid boxes. The newer-generation reanalyses (e.g., MERRA, JRA25, JRA55, CFSR, and ERA-Interim) generally have smaller root-mean-square-error (RMSE) than the older-generation ones (e.g., NCEP/NCAR, NCEP/DOE, and ERA-40).

Most of the reanalyses depict well the observed PW interannual variations over eastern and northern China but show relatively limited performance over western China, especially over the Tibetan Plateau. Most of the reanalyses are unable to capture the observed PW upward trends since 1979 by showing spurious wet biases before about 2003, and their PW trends differ substantially from each other. In particular, the ERA-Interim, NCEP/NCAR, NCEP/DOE, JRA-25, and the JRA-55 show significantly downward trends (contrary to the observations) over southeastern China during 1979–2012. Such erroneous long-term trends in reanalysis PW result mainly from the spurious changes in the assimilated raw humidity data from radiosondes and the varying role of satellite observations in each reanalysis. As an exception, the 20CR roughly reproduces the observed PW increases from 1979 to 2012 over most of China, mainly because this product did not assimilate upper air radiosonde and satellite observations and thus was not affected by inhomogeneities in these observational data.

Further analyses revealed that the wet PW bias before about 2003 in most of the reanalyses resulted mainly from the wet bias in RH before the early 2000s, rather than due to the errors in tropospheric *T*. This RH wet bias is linked to a change from the old humidity sensor Shang-M to the new Shang-E around 2002 over China [*Wang and Zhang*, 2008; *Bian et al.*, 2010]. Since the old sensor had a longer response time and lower sensitivity to dry conditions, the monthly mean RH from the earlier record contains a wet bias relative to the recent record. These radiosonde biases were assimilated into the reanalyses and apparently had large impacts on reanalysis PW and humidity fields, despite the assimilation of satellite observations since 1979. Thus, most reanalysis products, except the 20CR, should not be used for estimating long-term changes in atmospheric water vapor over China.

An EOF analysis of the monthly PW anomalies during 1979–2012 shows that the two leading EOFs and their corresponding PCs, which account for nearly 50% of the total variance, derived from the reanalysis are comparable to those from the observations, with the MERRA and JRA-25 performing best in reproducing the EOF 1 and EOF 2, respectively. The first EOF depicts a pattern of the same sign with the largest contribution from Southeast China that reflects primarily the national mean PW changes, while the second EOF shows a dipole pattern between Southeast China and the rest of the country that exhibits mostly interannual variations correlated with ENSO, with higher PW during El Niño events over Southeast China and lower PW over the rest of the country.

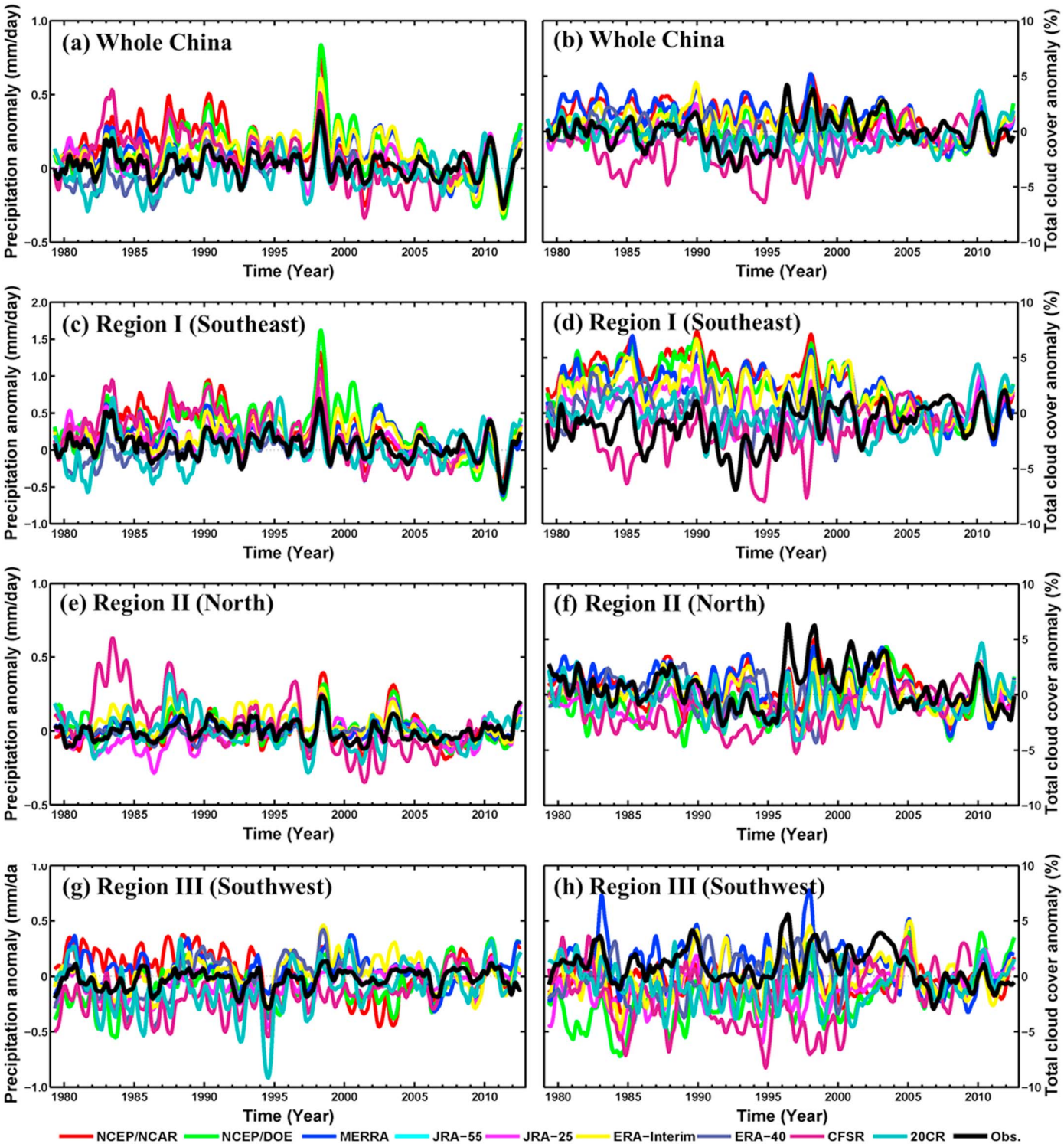

Figure 14. Same as Figure 5 but for precipitation anomalies (mm/d, left column) and total cloud cover anomalies (%, right column). Note that the observed precipitation is obtained by merging the CRU TS3.22 (2011–2012) into the GPCC V6 (1901–2010), and the observed total cloud cover is obtained from the CRU TS3.22.

Despite the uneven and sparse observations over western China, especially over the Tibetan Plateau and the Taklimakan desert, our results strongly suggest that discontinuities in radiosonde observations can lead to spurious changes in reanalysis products even for the post-1979 satellite era, when satellite observations become increasingly important for constraining upper air conditions. Furthermore, because water vapor

affects precipitation, cloudiness, and other related fields, it is possible that the errors and biases in PW may propagate into the other fields in the reanalysis products, as is the case for precipitation and cloud amount over China (Figure 14). Thus, the findings regarding the PW biases reported here have important implications for other water vapor-related fields in reanalysis products. Therefore, more efforts should be devoted to generate homogenized input data for future reanalysis products. This has become an increasingly urgent issue as many groups have been trying to produce increasingly longer reanalysis products for use in long-term climate change analyses. Without the input of homogenized observations, long-term reanalysis products will likely produce spurious and misleading changes.

Acknowledgments
The NCEP/NCAR, NCEP/DOE, and 20CR were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, and are available at http://www.esrl.noaa.gov/psd/. The CFSR, JRA-55, and JRA-25 were obtained from the CISL Research Data Archive (http://rda.ucar.edu/) at NCAR in Boulder, Colorado, and the NCAR is supported by grants from the National Science Foundation. The ERA-40 and ERA-interim were obtained from the ECMWF Data Server at http://apps.ecmwf.int/datasets/. The MERRA was obtained from the Goddard Earth Sciences Data and Information Services Center, Greenbelt, Maryland, from their website at http://disc.sci.gsfc.nasa.gov/mdisc. This work was supported by the National Basic Research Program of China (grant 2012CB956203), the China Meteorological Administration R&D Special Fund for Public Welfare (Meteorology) (grant GYHY201506002), the National Key Technologies R&D Program of China (grant 2012BAC22B04), the Climate Change Special Fund of the China Meteorological Administration: Atmospheric Water Vapor Changes in China and Its Causes (grant CCSF201330), and the Beijing Nova Program (grant 2008B40). A. Dai is supported by U.S. National Science Foundation (grant AGS-1353740), U.S. Department of Energy's Office of Science (award DE-SC0012602), and the U.S. National Oceanic and Atmospheric Administration (award NA15OAR4310086).

References

Bao, X., and F. Zhang (2013), Evaluation of NCEP–CFSR, NCEP–NCAR, ERA-Interim, and ERA-40 reanalysis datasets against independent sounding observations over the Tibetan Plateau, *J. Clim.*, *26*, 206–214.

Bengtsson, L., S. Hagemann, and K. I. Hodges (2004), Can climate trends be calculated from reanalysis data, *J. Geophys. Res.*, *109*, D1111, doi:10.1029/2004JD004536.

Bian, J., H. Chen, H. Vomel, Y. Duan, Y. Xua, and D. Lü (2010), Intercomparison of humidity and temperature sensors: GTS1, Vaisala RS80 and CFH, *Adv. Atmos. Sci.*, *28*, 139–146, doi:10.1007/s00376-010-9170-8.

Compo, G. P., et al. (2011), The twentieth century reanalysis project, *Q. J. R. Meteorol. Soc.*, *137*, 1–28.

Cressman, G. P. (1959), An operational objective analysis scheme, *Mon. Weather Rev.*, *87*, 367–374.

Dai, A. (2006), Recent climatology, variability and trends in global surface humidity, *J. Clim.*, *19*, 3589–3606.

Dai, A., T. M. L. Wigley, B. A. Boville, J. T. Kiehl, and L. E. Buja (2001), Climates of the 20th and 21st centuries simulated by the NCAR Climate System Model, *J. Clim.*, *14*, 485–519.

Dai, A., 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. Clim.*, *24*, 965–991.

Dai, A., H. Li, Y. Sun, L.-C. Hong, C. C. LinHo, and T. Zhou (2013), The relative roles of upper and lower tropospheric thermal contrasts and tropical influences in driving Asian summer monsoons, *J. Geophys. Res. Atmos.*, *118*, 7024–7045, doi:10.1002/jgrd.205065.

Dee, D. P., et. al. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*, 553–597.

Dessler, A. E., and S. M. Davis (2010), Trends in tropospheric humidity from reanalysis systems, *J. Geophys. Res.*, *115*, D19127, doi:10.1029/2010JD014192.

Ebita, A., et al. (2011), The Japanese 55-year Reanalysis "JRA-55," an interim report, *SOLA*, *7*, 149–152.

Fan, J. Q., and Q. W. Yao (2003), *Nonlinear Time Series: Nonparametric and Parametric Methods (Springer Series in Statistics)*, Springer, New York.

Gaffen, D. J., W. P. Elliott, and A. Robock (1992), Relationships between tropospheric water vapor and surface temperature as observed by radiosondes, *Geophys. Res. Lett.*, *19*, 1839–1942, doi:10.1029/92GL02001.

Haimberger, L., C. Tavolato, and S. Sperka (2008), Toward elimination of the warm bias in historic radiosonde temperature records—Some new results from a comprehensive intercomparison of upper air data, *J. Clim.*, *21*, 4587–4606.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister (2013), Updated high-resolution grids of monthly climatic observations – Rhe CRU TS3.10 dataset, *Int. J. Clim.*, *34*, 623–642, doi:10.1002/joc.3711.

Held, I. M., and B. J. Soden (2000), Water vapor feedback and global warming, *Ann. Rev. Energ. Environ.*, *25*, 441–475.

Jia, B., Z. Xie, A. Dai, C. Shi, and F. Chen (2013), Evaluation of satellite and reanalysis products of downward surface solar radiation over East Asia: Spatial and seasonal variations, *J. Geophys. Res. Atmos.*, *118*, 1–16, doi:10.1002/jgrd.50353.

Kalnay, E., et al. (1996), The NCEP/NCAR 40-year reanalysis project, *Bull. Am. Meteorol. Soc.*, *77*, 437–470.

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. Am. Meteorol. Soc.*, *83*, 1631–1643, doi:10.1175/BAMS-83-11-1631.

Kendall, M. G. (1975), *Rank Correlation Methods*, Griffin, London.

Kistler, R., et al. (2001), The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, *Bull. Am. Meteorol. Soc.*, *82*, 247–267.

Lin, C., K. Yang, J. Qin, and R. Fu (2013), Observed coherent trends of surface and upper-air wind speed over China since 1960, *J. Clim.*, *26*, 2891–2903.

Lu, N., K. E. Trenberth, J. Qin, K. Yang, and L. Yao (2015), Detecting long-term trends in precipitable water over the Tibetan Plateau by synthesis of station and MODIS observations, *J. Clim.*, *28*, 1707–1722, doi:10.1175/JCLI-D-14-00303.1.

Ma, L., T. Zhang, Q. X. Li, O. W. Frauenfeld, and D. H. 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, doi:10.1029/2007JD009549.

Ma, L., T. Zhang, O. W. Frauenfeld, B. Ye, D. Yang, and D. Qin (2009), Evaluation of precipitation from the ERA-40, NCEP-1, and NCEP-2 reanalyses and CMAP-1, CMAP-2, and GPCP-2 with ground-based measurements in China, *J. Geophys. Res.*, *114*, D09105, doi:10.1029/2008JD011178.

Mann, H. B. (1945), Nonparametric tests against trend, *Econometrica*, *13*, 245–259.

Mao, J., X. Shi, L. Ma, D. P. Kaiser, Q. Li, and P. E. Thornton (2010), Assessment of reanalysis daily extreme temperatures with China's homogenized historical dataset during 1979–2001 using probability density functions, *J. Clim.*, *23*, 6605–6623, doi:10.1175/2010JCLI3581.1.

Onogi, K., et al. (2007), The JRA-25 Reanalysis, *J. Meteor. Soc. Jpn.*, *85*, 369–432.

Paltridge, G., A. Arking, and M. Pook (2009), Trends in middle-and upper-level tropospheric humidity from NCEP reanalysis data, *Theor. Appl. Clim.*, *98*, 351–359, doi:10.1007/s00704-009-0117-x.

Qian, T., A. Dai, K. E. Trenberth, and K. W. Oleson (2006), Simulation of global land surface conditions from 1948 to 2002: Part I: Forcing data and evaluations, *J. Hydrometeorol.*, *7*, 953–975.

Rienecker, M. M., et al. (2011), MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Clim.*, *24*, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Ross, R. J., and W. P. Elliott (2001), Radiosonde-based Northern Hemisphere tropospheric water vapor trends, *J. Clim.*, *14*, 1602–1612.

Ross, R. J., W. P. Elliott, and D. J. Seidel (2002), Lower-tropospheric humidity–temperature relationships in radiosonde observations and atmospheric general circulation models, *J. Hydrometeorol.*, *3*, 26–38.

Rowe, P. M., L. M. Miloshevich, D. D. Turner, and V. P. Walden (2008), Dry bias in Vaisala RS90 radiosonde humidity profiles over Antarctica, *J. Atmos. Oceanic Technol.*, *25*, 1529–1541.

Saha, S., et al. (2010), The NCEP climate forecast system reanalysis, *Bull. Am. Meteorol. Soc.*, *91*, 1015–1057, doi:10.1175/2010BAMS3001.1.
Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf (2013), GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theor. Appl. Clim.*, doi:10.1007/s00704-013-0860-x.
Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc.*, *63*, 1379–1389.
Serreze, M. C., A. P. Barrett, and J. Stroeve (2012), Recent changes in tropospheric water vapor over the Arctic as assessed from radiosondes and atmospheric reanalyses, *J. Geophys. Res.*, *117*, D10104, doi:10.1029/2011JD017421.
Shi, X. H., X. D. Xu, and L. A. Xie (2006), Reliability analyses of anomalies of NCEP/NCAR reanalysis wind speed and surface temperature in climate change research in China [in Chinese], *Acta Meteorol. Sin.*, *64*, 709–722.
Simmons, A. J., 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, doi:10.1029/2009JD012442.
Sun, D. Z., and A. H. Oort (1995), Humidity–temperature relationships in the tropical troposphere, *J. Clim.*, *8*, 1974–1987.
Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons (2003), The changing character of precipitation, *Bull. Am. Meteorol. Soc.*, *84*, 1205–1217.
Trenberth, K. E., J. Fasullo, and L. Smith (2005), Trends and variability in column-integrated atmospheric water vapor, *Clim Dyn.*, *24*, 741–758.
Uppala, S. M., et al. (2005), The ERA-40 reanalysis, *Q. J. R. Meteorol. Soc.*, *131*, 2961–3012.
Wang, J., and L. Zhang (2008), Systematic errors in global radiosonde precipitable water data from comparisons with ground based GPS measurements, *J. Clim.*, *21*, 2218–2238.
Wang, J., D. J. Carlson, D. B. Parsons, T. F. Hock, D. Lauritsen, H. L. Cole, K. Beierle, and E. Chamberlain (2003), Performance of operational radiosonde humidity sensors in direct comparison with a chilled mirror dew-point hygrometer and its climate implication, *Geophys. Res. Lett.*, *30*(16), 1860, doi:10.1029/2003GL016985.
Wang, J., L. Zhang, A. Dai, F. Immler, M. Sommer, and H. Vömel (2013), Radiation dry bias correction of Vaisala RS92 humidity data and its impacts on historical radiosonde data, *J. Atmos. Oceanic Technol.*, *30*, 197–214, doi:10.1175/JTECH-D-12-00113.1.
Wang, J. X. L., and D. J. Gaffen (2001), Late twentieth-century climatology and trends of surface humidity and temperature in China, *J. Clim.*, *14*, 2833–2845.
Xu, Y., Y. H. Ding, and Z. C. Zhao (2001), Confidence analysis of NCEP/NCAR 50-year global reanalyzed data in climate change research in China [in Chinese], *Q. J. Appl. Meteorol.*, *23*, 337–347.
Zhai, P., and R. E. Eskridge (1997), Atmospheric water vapor over China, *J. Clim.*, *10*, 2643–2652.
Zhao, T., and C. Fu (2006a), Preliminary comparison and analysis between ERA-40, NCEP-2 reanalysis and observations over China [in Chinese], *Clim. Environ. Res.*, *11*, 14–32.
Zhao, T., and C. Fu (2006b), Comparison of products from ERA-40, NCEP-2, and CRU with station data for summer precipitation over China, *Adv. Atmos. Sci.*, *23*, 593–604.
Zhao, T., L. Ai, and J. Feng (2004), An intercomparison between NCEP reanalysis and observed data over China [in Chinese], *Clim. Environ. Res.*, *9*, 278–294.
Zhao, T., W. Guo, and C. Fu (2008), Calibrating and evaluating reanalysis surface temperature error by topographic correction, *J. Clim.*, *21*, 1440–1446.
Zhao, T., A. Dai, and J. Wang (2012), Trends in tropospheric humidity from 1970 to 2008 over China from a homogenized radiosonde dataset, *J. Clim.*, *25*, 4549–4567.