Theor Appl Climatol (2015) 119:99–111
DOI 10.1007/s00704-014-1107-1

ORIGINAL PAPER

Comparison of surface air temperature derived from NCEP/DOE R2, ERA-Interim, and observations in the arid northwestern China: a consideration of altitude errors

Shengjie Wang · Mingjun Zhang · Meiping Sun · Baolong Wang · Xiaoyan Huang · Qiong Wang · Fang Feng

Received: 22 June 2013 / Accepted: 26 January 2014 / Published online: 13 February 2014

Abstract The surface air temperatures measured at 68 meteorological stations in the arid northwestern China during 1979–2012 are compared with temperatures interpolated from the National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) Reanalysis 2 (NCEP R2) and the European Center for Medium-Range Weather Forecasts ERA-Interim. The altitude effects on reanalysis temperature errors are discussed, and the interpolated reanalysis data are calibrated by altitude errors between reanalysis and observation. Using a simple correction method with a constant lapse rate, the elevation-related errors can be greatly removed and an improvement is achieved for the interpolated temperature from both NCEP R2 and ERA-Interim. The cold bias of reanalysis data becomes weak after calibration. On an annual basis, root mean square error of temperature derived from NCEP R2 for each stations has decreased from 6.0 (raw data) to 2.6 °C (calibrated data) and that from ERA-Interim has decreased from 3.2 to 1.4 °C. Similarly, correlation coefficients between raw reanalysis-based and observed temperature are 0.191 and 0.709 for NCEP R2 and ERA-Interim, respectively, whereas the correlation coefficients using the calibrated annual data are 0.819 and 0.932 for NCEP R2 and ERA-Interim, respectively. Generally, ERA-Interim is closer to the ground-based observations than NCEP R2. The topographic correction is more effective in summer than in winter, which may be related to the temperature inversion in winter. Evaluation and correction of reanalysis datasets is a crucial work before the gridded data are applied in climate research, and the altitude-related errors should be calibrated especially in the regions with complex topography.

S. Wang · M. Zhang (✉) · M. Sun · B. Wang · Q. Wang
College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China
e-mail: mjzhang2004@163.com

X. Huang
Institute of Arid Meteorology of China Meteorological Administration, Key Laboratory of Arid Climate Change and Reducing Disaster of Gansu Province, Key Open Laboratory of Arid Climate Change and Reducing Disaster of China Meteorological Administration, Lanzhou 730020, China

F. Feng
Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

1 Introduction

The arid northwestern China lies at the center of Eurasia continent. Due to the long distance to the surrounding oceans, the annual precipitation in the arid northwestern China is generally less than 200 mm, and vast desert basins are distributed in this area. In the past several decades, this region has exhibited a significant climatic and environmental change (Yao et al. 2004; Shi et al. 2007; Zhao et al. 2011; Chen et al. 2012; Liu et al. 2013; Wang et al. 2013a, 2013b). During 1960–2010, the regional air temperature averaged from surface stations has significantly increased by 0.34 °C/decade, higher than the average of China (0.25 °C/decade) and the world (0.13 °C/decade) for the same period (Li et al. 2012).

Given the complex landform in the northwestern China, the spatial distribution of in situ meteorological stations is logically uneven. Arithmetic mean value of air temperature measured at ground-based stations may ignore the weight of remote deserts and mountains. Reanalysis datasets provide a gridded climatology with fixed spatial resolution, which can reduce the bias caused by the uneven distribution of stations (Bengtsson et al. 2004). However, it is clear that reanalysis datasets are not always good at simulating conditions on the ground (Kostopoulou et al. 2010; Mooney et al. 2011;

Screen and Simmonds 2011; Wang and Zeng 2012; Cornes and Jones 2013). Validation of reanalysis data against homochronous observations is needed in examining climate change for a certain region.

The reliability of the widely applied National Centers for Environmental Prediction (NCEP) (Kalnay et al. 1996; Kistler et al. 2001; Kanamitsu et al. 2002) and European Center for Medium-Range Weather Forecasts (ECMWF) (Uppala et al. 2005; Dee et al. 2011) reanalysis datasets in China has aroused great attention in climatic diagnosis and assessment, especially for the surface air temperature. Ma et al. (2008) have compared these reanalysis data (ECMWF ERA-40, NCEP/NCAR R1, and NCEP/DOE R2) with the meteorological observations in China, and found that the reliability of reanalysis data in the eastern China is larger than that in the western China. Even in the eastern China, the bias of surface temperature also shows spatial dependency (Liu et al. 2012). The surface temperatures derived from observations and reanalysis datasets over the Tibetan Plateau have been compared by Wang and Zeng (2012) and You et al. (2010, 2013), and the quality of ERA-40 is considered to be better than that of NCEP/NCAR R1. In addition to the assimilation method and data involved in Simmons et al. (2004), the significant bias of reanalysis data in the western China is related with the complex topography and altitude difference between model and observation. Zhao et al. (2008) found that the topographical correction may be one of the critical steps to assess the uncertainties of reanalysis surface temperature, especially in the high mountains and deserts of the western China.

In this study, the climatology derived from long-term meteorological observation and two reanalysis datasets are compared, and the relationship between altitude errors and temperature bias are analyzed. In addition, the seasonal and annual surface temperatures derived from reanalysis are calibrated by altitude correction, and then the calibrated temperature series are compared with the raw data to assess the influence of altitude errors in the reanalyzed datasets.

2 Data and methods

2.1 Data sources

The observed monthly air temperature during 1979–2012 was acquired from the National Meteorological Information Center (NMIC) of China Meteorological Administration (http://cdc.cma.gov.cn). The arid northwestern China lies on the north of the Tibetan Plateau and on the west of the Helan Mountains. Within this boundary, there are 88 stations in the original dataset, and 68 stations without missing data are selected in this paper (Fig. 1; Table 1). All the deleted stations have a proportion of missing data exceeds 10 %. Data quality control was carried out by NMIC, including checks for high-low extreme values and time consistency (Li et al. 2004). In order to remove the effect of data inhomogeneity, penalized maximal F test was applied in this study (Wang 2008; Xu et al. 2013). Possible change points are detected and adjusted with metadata supporting, i.e., Balguntay in 1997 and 2012, Shisanjianfang in 1999, Luntai in 2010, and Barkol in 1985. Station relocation is the main reason of these data inhomogeneity. Then the homogenized series were used in the calculation below.

For NCEP/DOE R2 (NCEP R2 for short), monthly mean 2-m surface temperatures were acquired from the Physical Sciences Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration (http://www.esrl.noaa.gov/psd). This continually updated gridded dataset are available from January 1979 to the present with a spatial resolution of 1.875° (longitude)×1.905° (latitude) (Kanamitsu et al. 2002). Periods of 1979–2012 were selected from NCEP R2 in this study. For ERA-Interim, monthly data of 2-m surface air temperatures were obtained from the ECMWF (http://www.ecmwf.int). The monthly temperatures are acquirable since 1979 with a spatial resolution of 1.5° (longitude)×1.5° (latitude) (Dee et al. 2011). Periods of 1979–2012 were selected from this dataset.

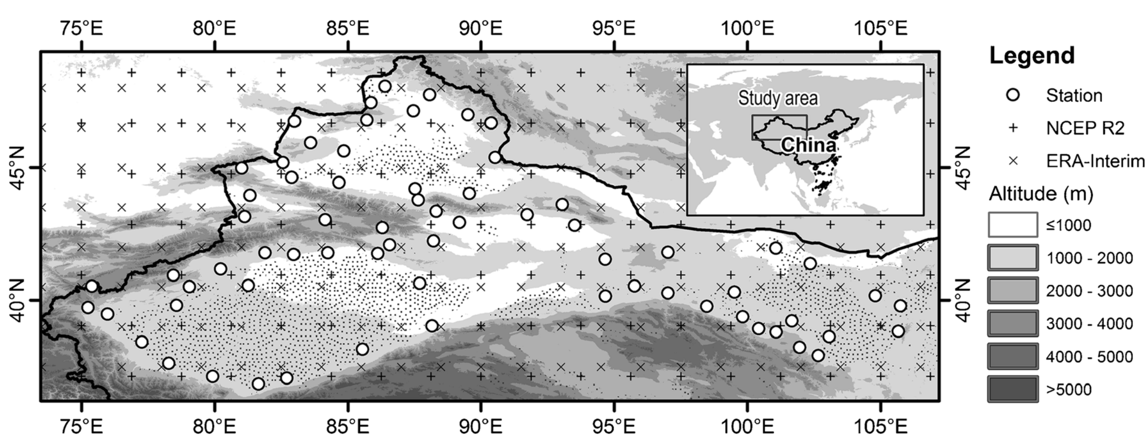

Fig. 1 Map showing meteorological stations and grid boxes of reanalysis datasets in the arid northwestern China. Locations of deserts in China are provided by the Environmental and Ecological Science Data Center for West China (Wang et al. 2005)

Table 1 List of selected meteorological stations with World Meteorological Organization (WMO) index number, latitude, longitude, and altitude

WMO no.	Station name	Latitude (° N)	Longitude (° E)	Altitude (m)
51053	Habahe	48.05	86.40	532.6
51059	Jeminay	47.43	85.87	984.1
51068	Fuhai	47.12	87.47	500.9
51076	Altay	47.73	88.08	735.3
51087	Fuyun	46.98	89.52	807.5
51133	Tacheng	46.73	83.00	534.9
51156	Hoboksar	46.78	85.72	1,291.6
51186	Qinghe	46.67	90.38	1,218.2
51232	Alataw Pass	45.18	82.57	336.1
51241	Toli	45.93	83.60	1,077.8
51243	Karamay	45.62	84.85	449.5
51288	Baytik Shan	45.37	90.53	1,653.7
51330	Wenquan	44.97	81.02	1,357.8
51334	Jinghe	44.62	82.90	320.1
51346	Usu	44.43	84.67	478.7
51365	Caijiahu	44.20	87.53	440.5
51379	Qitai	44.02	89.57	793.5
51431	Yining	43.95	81.33	662.5
51437	Zhaosu	43.15	81.13	1,851.0
51463	Ürümqi	43.78	87.65	935.0
51467	Balguntay	42.73	86.30	1,739.0
51477	Dabancheng	43.35	88.32	1,103.5
51495	Shisanjianfang	43.22	91.73	721.4
51526	Kumux	42.23	88.22	922.4
51542	Bayanbulak	43.03	84.15	2,458.0
51567	Yanqi	42.08	86.57	1,055.3
51573	Turpan	42.93	89.20	34.5
51628	Aksu	41.17	80.23	1,103.8
51633	Baicheng	41.78	81.90	1,229.2
51642	Luntai	41.78	84.25	976.1
51644	Kuqa	41.72	82.97	1,081.9
51656	Korla	41.75	86.13	931.5
51701	Torugart	40.52	75.40	3,504.4
51705	Wuqia	39.72	75.25	2,175.7
51709	Kashi	39.47	75.98	1,289.4
51711	Akqi	40.93	78.45	1,984.9
51716	Bachu	39.80	78.57	1,116.5
51720	Kalpin	40.50	79.05	1,161.8
51730	Aral	40.55	81.27	1,012.2
51765	Tikanlik	40.63	87.70	846.0
51777	Ruoqiang	39.03	88.17	887.7
51811	Shache	38.43	77.27	1,231.2
51818	Pishan	37.62	78.28	1,375.4
51828	Hotan	37.13	79.93	1,375.0
51839	Minfeng	37.07	82.72	1,409.5
51855	Qiemo	38.15	85.55	1,247.2
51931	Yutian	36.85	81.65	1,422.0
52101	Barkol	43.60	93.05	1,677.2
52203	Hami	42.82	93.52	737.2
52267	Ejin Qi	41.95	101.07	940.5

Table 1 (continued)

WMO no.	Station name	Latitude (° N)	Longitude (° E)	Altitude (m)
52313	Hongliuhe	41.53	94.67	1,573.8
52323	Mazongshan	41.80	97.03	1,770.4
52378	Guaizihu	41.37	102.37	960.0
52418	Dunhuang	40.15	94.68	1,139.0
52424	Guazhou	40.53	95.77	1,170.9
52436	Yumenzhen	40.27	97.03	1,526.0
52446	Dingxin	40.30	99.52	1,177.4
52495	Bayan Nuru	40.17	104.80	1,323.9
52533	Jiuquan	39.77	98.48	1,477.2
52546	Gaotai	39.37	99.83	1,332.2
52576	Alxa Youqi	39.22	101.68	1,510.1
52652	Zhangye	38.93	100.43	1,482.7
52661	Shandan	38.80	101.08	1,764.6
52674	Yongchang	38.23	101.97	1,976.9
52679	Wuwei	37.92	102.67	1,531.5
52681	Minqin	38.63	103.08	1,367.5
53502	Jartai	39.78	105.75	1,031.8
53602	Alxa Zuoqi	38.83	105.67	1,561.4

The NASA/NGA SRTM (Shuttle Radar Topography Mission) DEM V4 with spatial resolution of 90 m (http://srtm.csi.cgiar.org) is used to estimate the altitude of actual land surface.

2.2 Methods

In order to compare the observed air temperature with reanalysis-derived temperature, the nearest four grid boxes in NCEP R2 or ERA-Interim were averaged to the location for the given meteorological station. Four deterministic methods for multivariate interpolation were used, including bilinear interpolation, quadratic inverse distance weighted (IDW), cubic IDW, and quartic IDW (Shepard 1968). The interpolated results show that there is no significant difference among these interpolation methods. The correlation coefficients for each station are larger than 0.95, respectively, and the linear regressions are statistically significant at the level of 0.01, respectively. Figure 2 shows an example of the location in Habahe (48.05° N, 86.40° E), and all the correlation coefficients and slopes of linear regression are very close to 1. In this study, the method of cubic IDW are chosen for calculating weighted averages.

In order to remove the errors caused by the elevation differences between reanalysis and observations, the

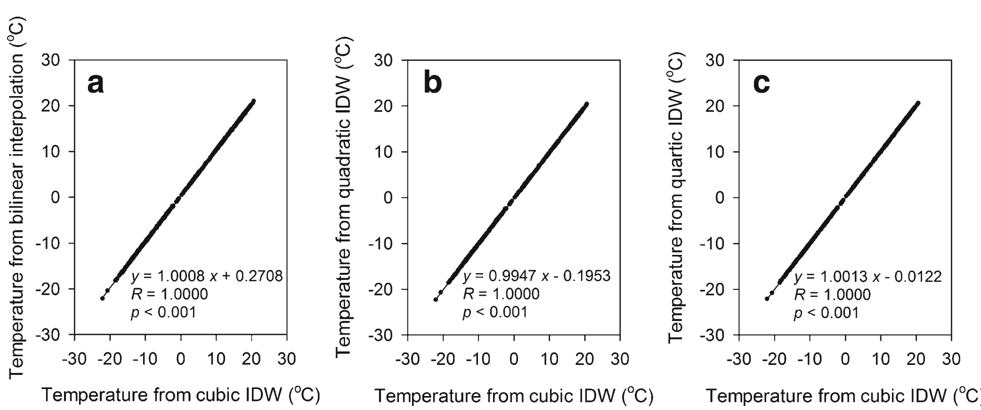

Fig. 2 Relationship between temperatures interpolated from ERA-Interim using cubic IDW and other methods (**a** bilinear interpolation, **b** quadratic IDW, and **c** quartic IDW) for the location of station Habahe (48.05° N, 86.40° E)

interpolated surface temperature was calibrated by using the following method of topographic correction (Cosgrove et al. 2003; Zhao et al. 2008)

$$T_{\mathrm{cal}} = T + \gamma \Delta H \quad (1)$$

where T_{cal} is the calibrated surface temperature, T is the surface temperature interpolated from the nearest four grid boxes using cubic IDW, γ is the lapse rate (assumed to be −0.65 °C/100 m), and ΔH is the altitude difference between the interpolated reanalysis height and observed station height.

The mean bias error (MBE) and root mean square error (RMSE) were used for comparing the observed surface air temperature and reanalysis-interpolated temperature. MBE and RMSE were calculated as

$$\mathrm{MBE} = \frac{\sum_{i=1}^{n} (T_i - t_i)}{n} \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (T_i - t_i)^2}{n}} \quad (3)$$

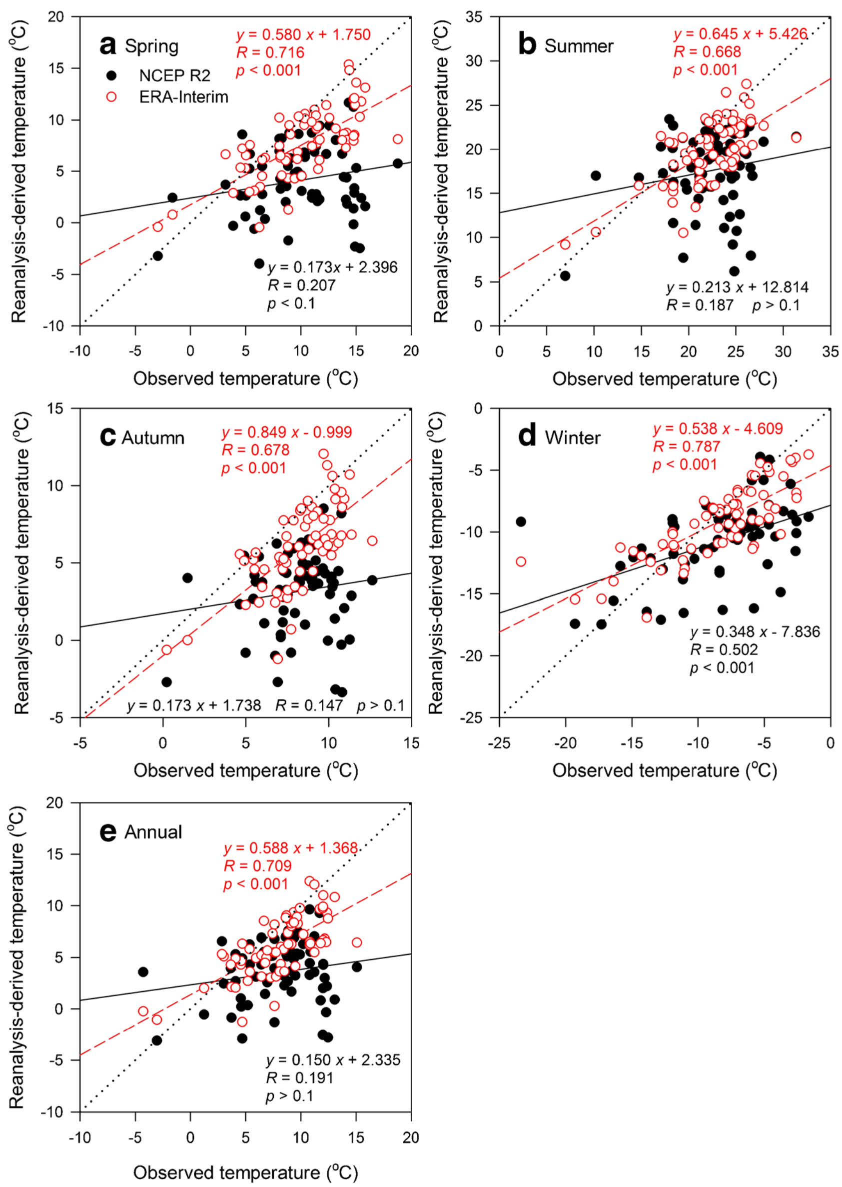

Fig. 3 Correlation between seasonal/annual surface air temperatures derived from observed stations and reanalysis datasets for each station in the arid northwestern China during 1979–2012. The *dotted line* is diagonal line of equality. The mean air temperatures are calculated for each station, respectively

where T_i and t_i are reanalysis-derived and observed seasonal/annual mean surface air temperature, respectively.

In calculation of correlation coefficient, Pearson's correlation and two-tailed t test were applied. All the monthly temperature were calculated on an annual and seasonal basis, i.e., spring, summer, autumn, and winter.

3 Results

3.1 Bias between reanalysis and observations

Figure 3 shows the correlations for seasonal/annual mean temperatures between reanalysis datasets and observed stations, respectively. The mean air temperatures are calculated for each station, respectively. The correlation between annual temperatures from NCEP R2 and observed stations is generally weak, and the correlation coefficient is only 0.191 on an annual basis. In most seasons, the temperatures from NCEP R2 are located to the right of the diagonal line of equality (dotted lines in Fig. 3). This also indicates that NCEP R2 has a cold bias. For REA-Interim, the interpolated temperatures are closer to the diagonal line of equality than those derived from NCEP R2, although the temperatures in most stations are also located to the right of the diagonal line. The difference between observed stations and ERA-Interim is smaller than that between observed stations and NCEP R2.

Figure 4 demonstrates the bias of the two reanalysis datasets. As shown in Fig. 4a, b, for NCEP R2, MBE of 22 (32.4 %) stations range within ±3, and for ERA-Interim, 46 (67.6 %) stations have MBE within ±3. Compared with NCEP R2, the MBE for ERA-Interim is closer to 0, and the median of MBE for NCEP R2 is lower than that for ERA-Interim.

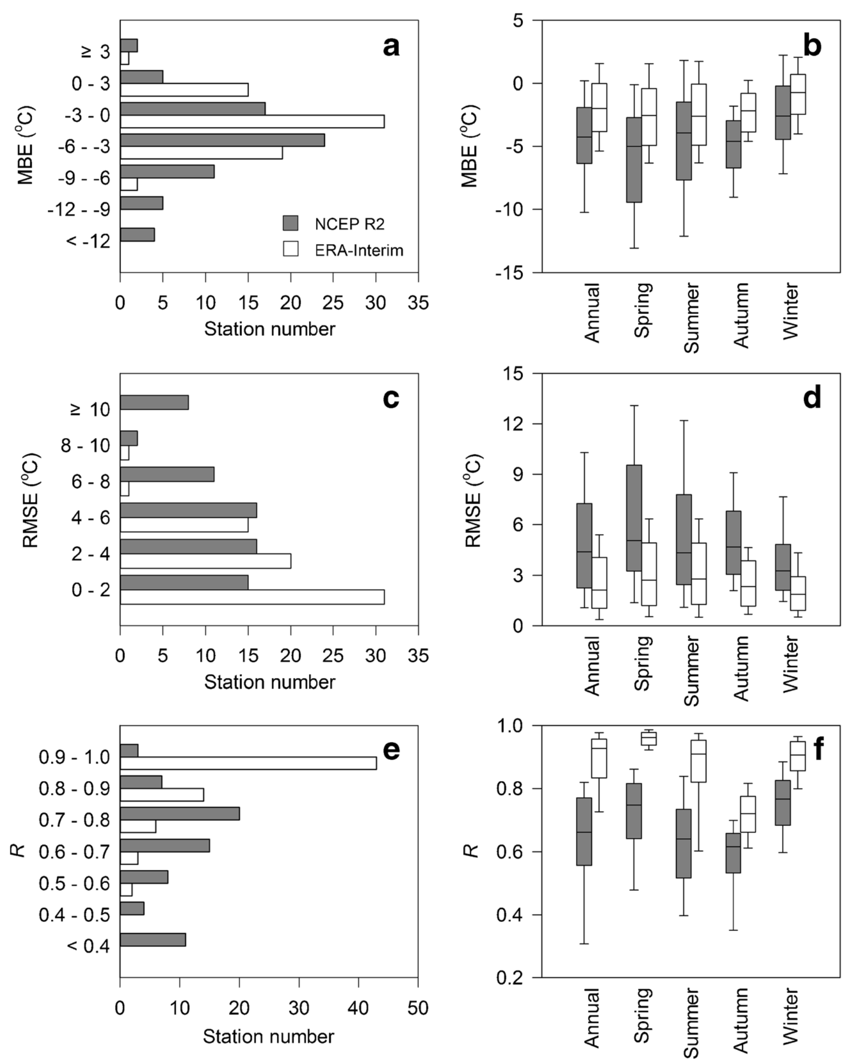

Fig. 4 Distribution (*left column*) and box plot (*right column*) of annual/seasonal MBE (**a**, **b**), RMSE (**c**, **d**), and correlation coefficient (**e**, **f**) for NCEP R2 and ERA-Interim in the arid northwestern China during 1979–2012. For the box plot, the *lower boundary of the box* indicates the 25th percentile, a *line within the box* marks the 50th percentile (median), and the *upper boundary of the box* indicate the 75th percentile; *whiskers* (*error bars*) above and below the box indicate the 90th and 10th percentiles

Figure 4c, d also indicates that ERA-Interim is good at describing the surface air temperature. Approximately half stations (45.6 %) derived from ERA-Interim have a low RMSE less than 2 °C, but the proportion from NCEP R2 is only 22.1 %. There is no station from ERA-Interim with RMSE larger than 10 °C, but 11.8 % stations from NCEP R2 have RMSE no less than 10 °C. On a seasonal basis, spring usually has the largest RMSE, and winter have the least. Shown in Fig. 4e, f, most stations (63.2 %) have correlation coefficient larger than 0.9 for ERA-Interim, and the proportion for NCEP R2 is only 4.4 %; 33.8 % stations for NCEP R2 have correlation coefficient less than 0.6, but only 2.9 % stations interpolated from ERA-Interim have correlation coefficient no more than 0.6.

3.2 Influence of altitude error in reanalysis

Difference of altitudes between reanalysis model and actual land surface are considered as an important factor influencing the reliability of surface air temperature in reanalysis data. Figure 5a, b shows the spatial distribution of altitude errors between the reanalysis and observations. For NCEP R2, the maximum of altitude errors for the locations of observation stations is up to 2,676.9 m, and the minimum is −1,020.8 m. For ERA-Interim, the altitude errors range from −553.5 to 1,205.1 m, and the mean value is 384.6 m. Figure 5c, d shows a wider scope of the altitude errors in the arid northwestern China. For NCEP R2, positive altitude errors larger than 1,500 m can be found at the southern margin of Tarim Basin (Taklimakan Desert) and other low-lying area near the Qilian Mountains and Tianshan Mountains. The negative altitude errors are widely distributed over the high-elevation areas especially the Tianshan Mountains. Compared with NCEP R2, the altitude of ERA-Interim is more similar with the actual land surface. The area with altitude errors larger than 1,500 m and less than −1,000 in ERA-Interim dataset is much less than that in NCEP R2.

Figure 6 demonstrates the relationship between the altitude errors and seasonal/annual MBE, and the negative correlation can be found for each seasonal/annual series. Positive altitude error usually corresponds to negative MBE, and MBE for the stations with negative altitude error usually have positive values. All the correlations shown in Fig. 6 are statistically significant at the 0.01 level. The slopes of linear regression on annual basis are −0.48 °C/100 m and −0.56 °C/100 m for NCEP R2 and ERA-Interim, respectively. For the seasonal series, the slopes of linear regression range from −0.26 °C/100 m to −0.74 °C/100 m, and winter usually has the larger slope and summer has the lower. For the stations with negative altitude error, annual MBE for NCEP R2 and ERA-Interim are 0.6 and 1.3 °C, respectively (Table 2). However, for the stations with altitude errors larger than 500 m, the annual MBE for NCEP R2 is up to −6.6 °C and that for ERA-Interim is −4.2 °C.

3.3 Calibrated reanalysis by altitude correction

According to the abovementioned relationship between MBE and altitude error, the monthly reanalysis data can be calibrated by altitude correction. Based on the altitude difference between interpolated reanalysis model and corresponding

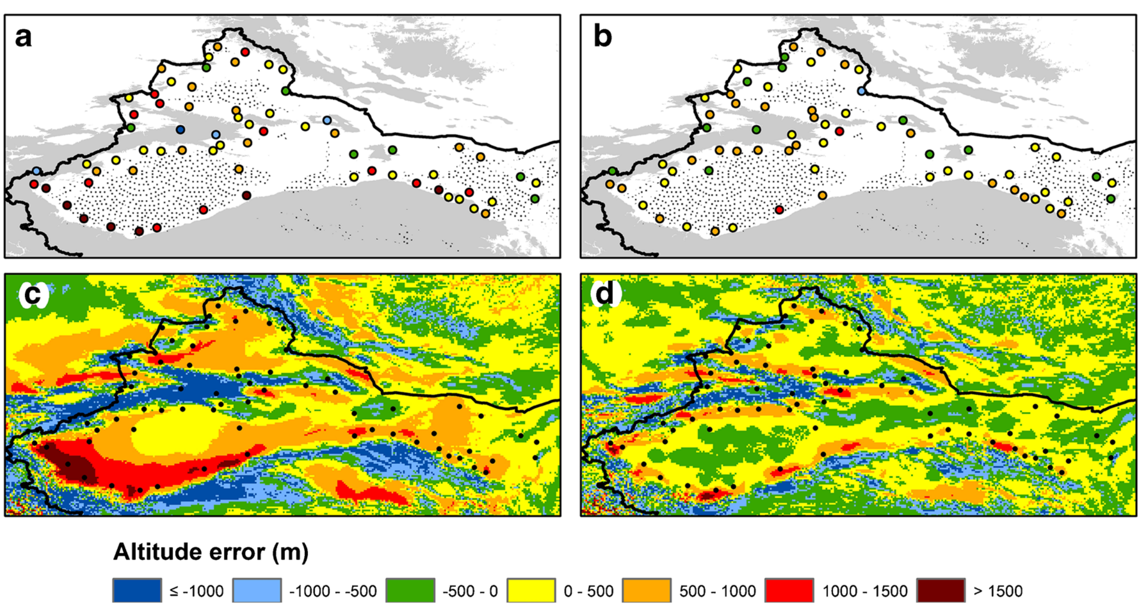

Fig. 5 Spatial distribution of altitude error calculated by using interpolated reanalysis height (*left column*, NCEP R2; *right column*, ERA-Interim) minus topography height (**a**, **b** observed stations and **c**, **d** SRTM DEM) in the arid northwestern China. *Gray shadows* in (**a**) and (**b**) denote altitude above 2,000 m

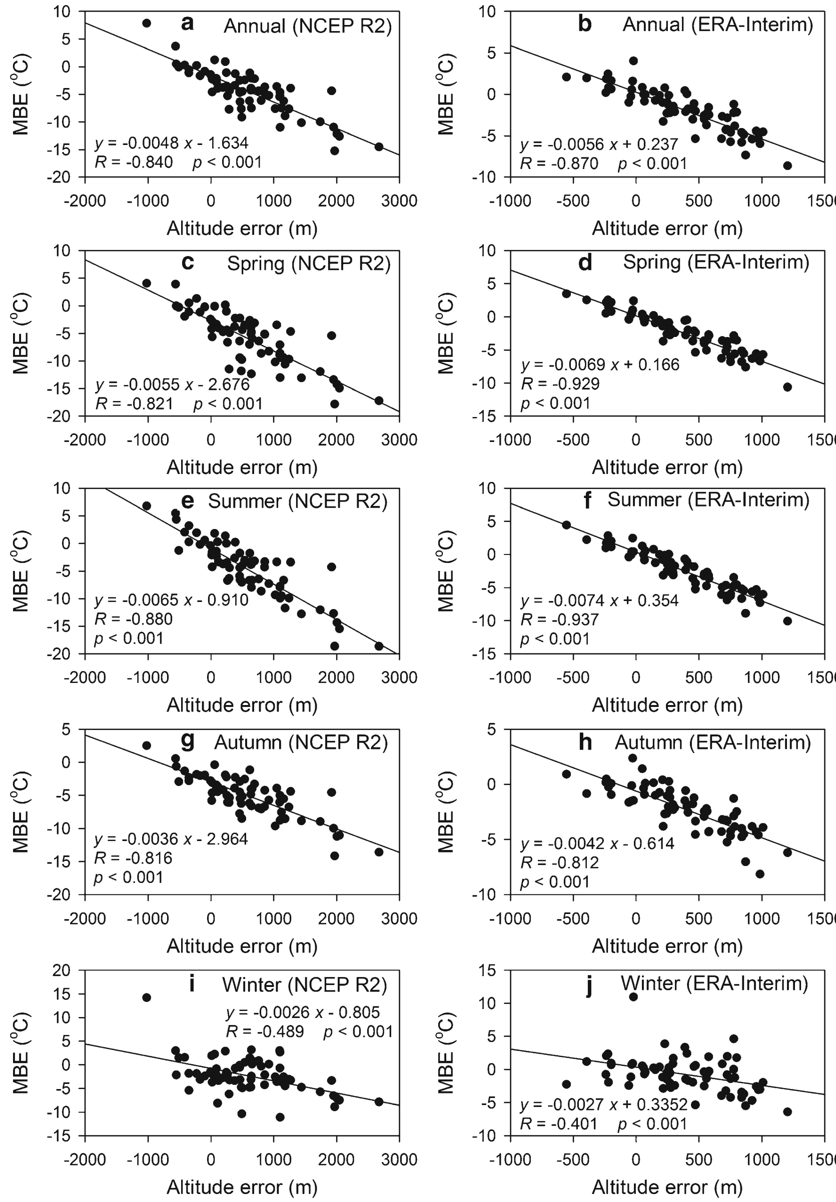

Fig. 6 Correlation between annual/seasonal MBE and altitude error (reanalysis minus observation) for NCEP R2 (*left column*) and ERA-Interim (*right column*) in the arid northwestern China during 1979–2012

observation stations, a constant lapse rate of air temperature (–0.65 °C/100 m) are assumed to calibrate the reanalysis data. Similar as Fig. 3, Fig. 7 demonstrates correlations between calibrated reanalysis-derived temperature and observed temperature. The fit lines after calibration are generally closer to diagonal line of equality. Detail information about the raw and calibrated temperature series is listed in Table 3. After calibration, the interpolated reanalysis data has greatly improved. For the annual series, MBE from NCEP R2 has increased from −4.5 (raw data) to −0.6 °C (calibrated data), and that from ERA-Interim has increased from −1.9 to 0.6 °C. RMSE has also reduced after altitude correction. The annual RMSE from NCEP R2 is 6.0 °C in raw data and 2.6 °C in calibrated data,

Table 2 MBE of annual/seasonal surface air temperature interpolated from NCEP R2 and ERA-Interim at different altitude error ranges in the arid northwestern China during 1979–2012

Reanalysis data	Altitude error (m)	Station number	MBE (°C)				
			Annual	Spring	Summer	Autumn	Winter
NCEP R2	≤0	11	0.6	0.2	2.1	−1.4	0.1
	0 to 500	23	−3.7	−5.0	−3.1	−4.4	−2.5
	>500	34	−6.6	−8.5	−8.0	−6.7	−3.0
ERA-Interim	≤0	13	1.3	1.5	1.6	−0.2	0.8
	0 to 500	28	−1.3	−1.7	−1.5	−1.5	−0.4
	>500	27	−4.2	−5.2	−5.4	−4.0	−1.8

and that from ERA-Interim is 3.2 °C in raw data and 1.4 °C in calibrated data. The correlation coefficient between raw temperature derived from NCEP R2 and observed temperature is only 0.191, and the coefficient has risen up to 0.819 after correction. The correlation coefficient for ERA-Interim has also increased from 0.709 (raw data) to 0.932 (calibrated data). In addition, the slope of linear regression using the calibrated data is closer to 1, compared with that using the raw data. On a annual basis, the raw slope from NCEP R2 is 0.150 and the calibrated slope is up to 1.018, and the slope from ERA-Interim changes from 0.588 (raw data) to 0.921 (calibrated data). Similar to the annual series, the seasonal series from reanalysis generally shows a significant improvement after altitude correction, except that in winter.

The improvement of calibrated temperature can also be found in the box plot shown in Fig. 8. After calibration, both the median MBEs of NCEP R2 and ERA-Interim are closer to 0, which is obvious in all the annual/seasonal series except winter. Similarly, the median RMSE of NCEP R2 and ERA-Interim have also decreased after calibration except in winter. Compared with NCEP R2, the calibrated RMSE of ERA-Interim are still much lower. Among the four seasons, winter usually have the most significant fluctuation.

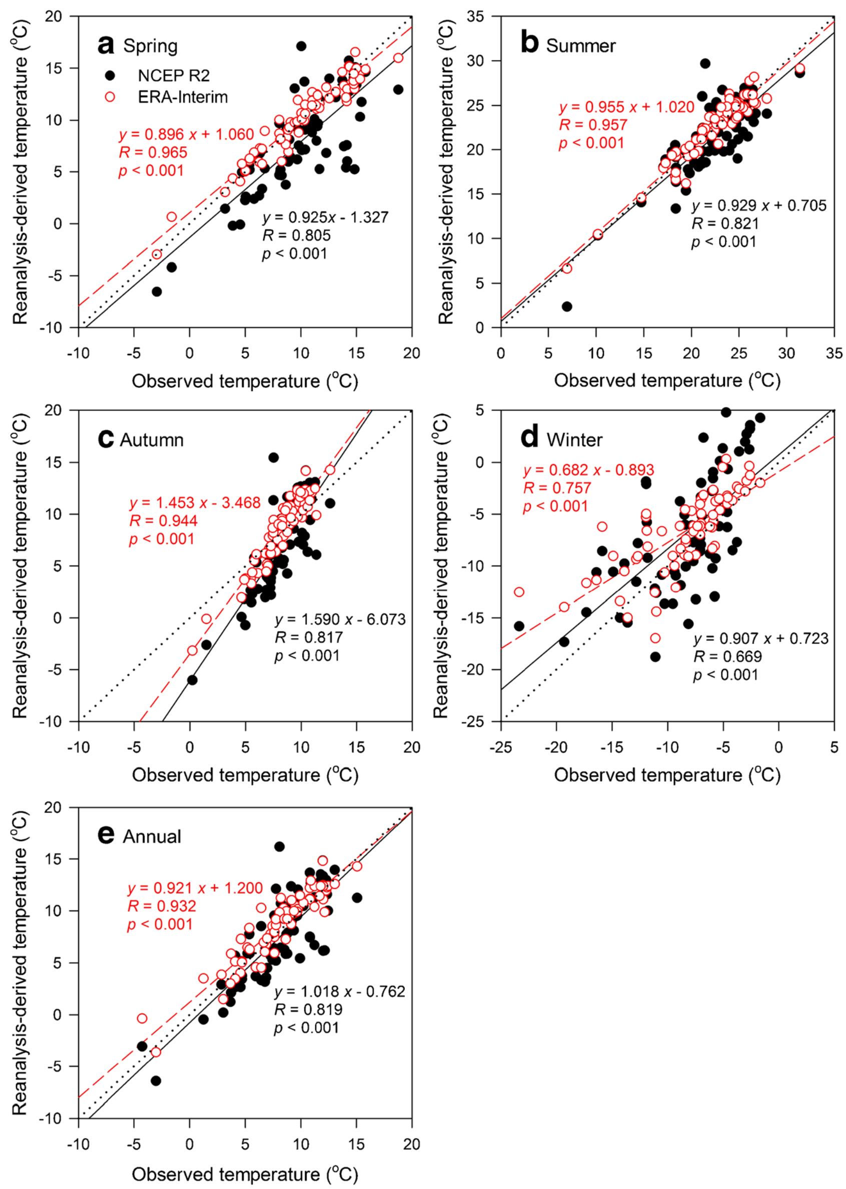

Fig. 7 Correlation between seasonal/annual surface air temperatures derived from observed stations and calibrated reanalysis datasets for each station in the arid northwestern China during 1979–2012. The *dotted line* is diagonal line of equality. The mean air temperatures are calculated for each station, respectively

Table 3 Comparison of annual/seasonal MBE, RMSE, correlation coefficient and slope of linear regression between raw and calibrated surface air temperature interpolated from NCEP R2 and ERA-Interim in the arid northwestern China during 1979–2012

Reanalysis data	Season	MBE (°C)		RMSE (°C)		*R*		Slope	
		Raw	Calibrated	Raw	Calibrated	Raw	Calibrated	Raw	Calibrated
NCEP R2	Annual	−4.5	−0.6	6.0	2.6	0.191	0.819	0.150	1.018
	Spring	−5.9	−2.1	7.6	3.5	0.207	0.805	0.173	0.925
	Summer	−4.7	−0.9	7.0	2.6	0.187	0.821	0.213	0.929
	Autumn	−5.1	−1.2	5.9	3.0	0.147	0.817	0.173	1.590
	Winter	−2.3	1.5	4.4	4.5	0.502	0.669	0.348	0.907
ERA-Interim	Annual	−1.9	0.6	3.2	1.4	0.709	0.932	0.588	0.921
	Spring	−2.5	0.0	3.8	1.1	0.716	0.965	0.580	0.896
	Summer	−2.5	0.0	3.9	1.1	0.668	0.957	0.645	0.955
	Autumn	−2.2	0.3	3.0	1.5	0.678	0.944	0.849	1.453
	Winter	−0.7	1.8	2.7	3.3	0.787	0.757	0.538	0.682

As is shown in Fig. 9, most stations show a trend that calibrated MBE is closer to 0. The proportions of stations with improved MBE (closer to 0) for annual, spring, summer, autumn and winter are 77.9, 83.8, 80.9, 80.9, and 51.5 % for NCEP R2, respectively, and 73.5, 88.2, 83.8, 67.6, and 42.6 % for ERA-Interim, respectively. The annual MBE of NCEP R2 ranges from −15.3 to 7.8 °C before altitude correction, but the calibrated data range between −6.0 and 8.1 °C. For ERA-Interim, the raw annual MBE ranges from −8.6 to 4.0 °C, and the calibrated MBE ranges from −2.3 to 3.9 °C.

In Fig. 10, RMSE at most stations also show a great improvement after calibration. On an annual basis, the proportions of stations with decreased RMSE are 77.9 and 73.5 % for NCEP R2 and ERA-Interim, respectively. Before

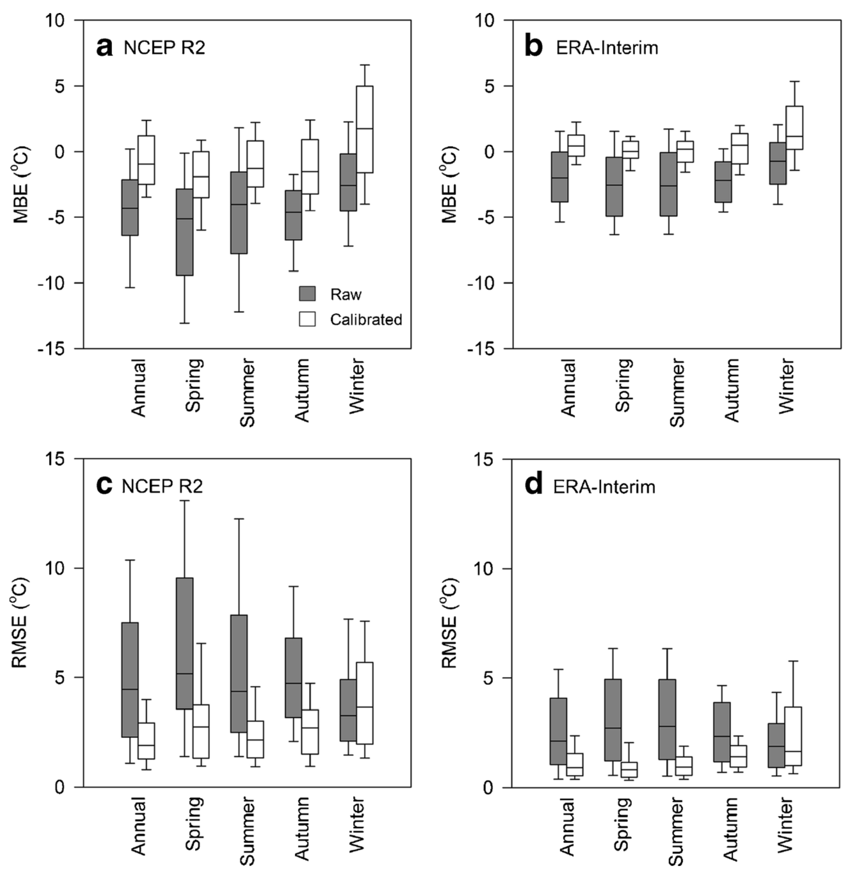

Fig. 8 Box plot of MBE (**a**, **b**) and RMSE (**c**, **d**) of raw and calibrated surface air temperature interpolated from NCEP R2 (*left column*) and ERA-Interim (*right column*) in the arid northwestern China during 1979–2012. The *lower boundary of the box* indicates the 25th percentile, a *line within the box* marks the 50th percentile (median), and the *upper boundary of the box* indicate the 75th percentile; *whiskers* (*error bars*) above and below the box indicate the 90th and 10th percentiles

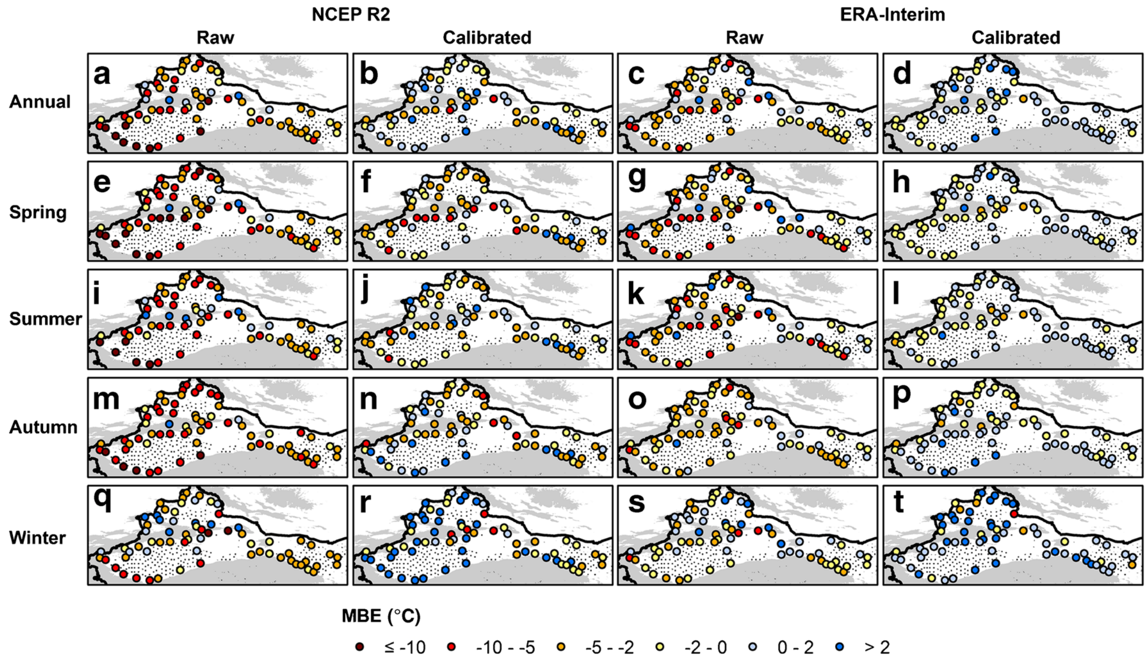

Fig. 9 Spatial distribution of annual/seasonal MBE of raw (*1st and 3rd columns*) and calibrated (*2nd and 4th columns*) surface air temperature interpolated from NCEP R2 (*left two columns*) and ERA-Interim (*right two columns*) in the arid northwestern China during 1979–2012. *Gray shadows* denote altitude above 2,000 m

calibration, the proportions of stations with RMSE less than 2 °C for annual, spring, summer, autumn, and winter are 20.6, 13.2, 17.6, 8.8, and 23.5 % for NCEP R2, respectively, and 45.6, 32.4, 38.2, 44.1, and 52.9 % for ERA-Interim, respectively. After calibration, the proportions have changed to 54.4, 38.2, 45.6, 32.4, and 25.0 % for NCEP R2, respectively, and 85.3, 88.2, 92.6, 79.4, and 54.4 % for ERA-Interim, respectively. Generally, after calibration, reanalysis-interpolated data at most stations have significantly improved.

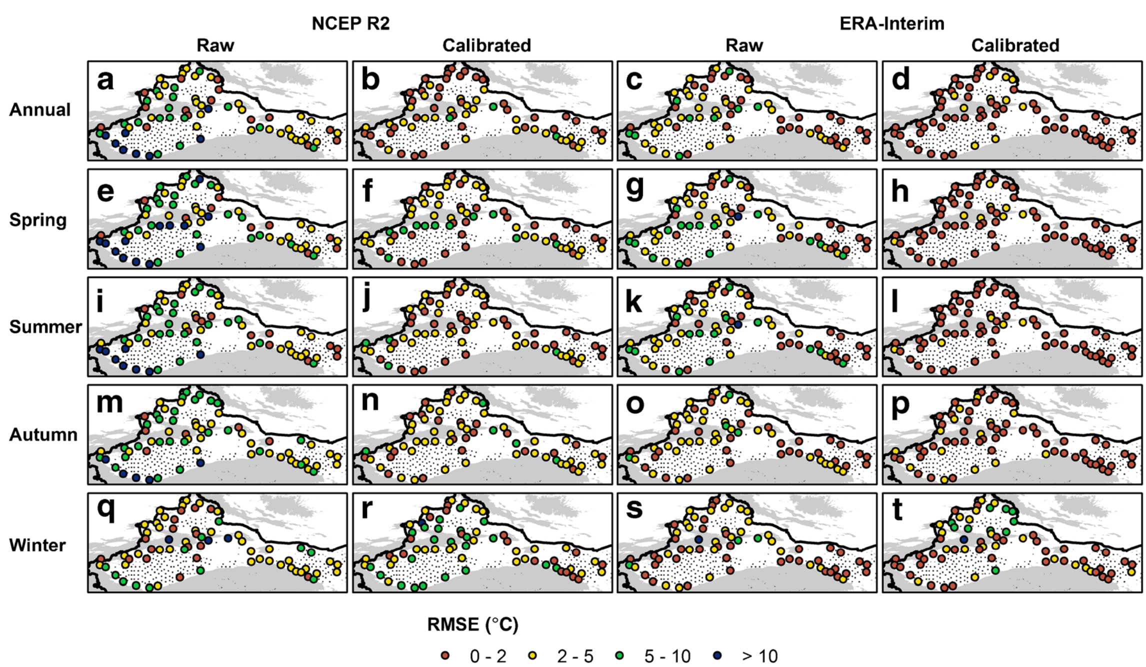

Fig. 10 Spatial distribution of annual/seasonal RMSE of raw (*1st and 3rd columns*) and calibrated (*2nd and 4th columns*) surface air temperature interpolated from NCEP R2 (*left two columns*) and ERA-Interim (*right two columns*) in the arid northwestern China during 1979–2012. *Gray shadows* denote altitude above 2,000 m

4 Discussion and conclusions

Based on the monthly surface air temperature derived from observation and reanalysis datasets (NCEP R2 and ERA-Interim) during 1979–2012, the altitude effects on reanalysis data errors are discussed, and the interpolated reanalysis data are calibrated by altitude errors. Previous studies (Zhao et al. 2008) indicated that the effect of the topographic correction for reanalysis dataset is more effective in the regions with complex altitude difference. In the arid northwestern China, most meteorological stations are located at the flat oasis belts, between the desert basins and high mountains. Generally, the altitude of observational stations is lower than the model altitude of reanalysis data at the same location, as is shown in Fig. 5. Before the calibration, both NCEP R2 and ERA-Interim have significant cold biases, compared with observation. Altitude difference between reanalysis model and observed station negatively correlates with the bias of surface air temperature. The cold bias in interpolated stations is related by the altitude differences between model and observation. Topographic adjustment is needed for minimize the elevation-related bias.

Using a simple correction method with a constant lapse rate, the elevation-related errors can be greatly removed and an improvement can be achieved for the interpolated temperature. The cold bias becomes weak after calibration, and the significant improvements for both NCEP R2 and ERA-Interim can be clearly seen from the MBE and RMSE. It should be noticed that the lapse rate of air temperature are assumed to the global mean (0.65 °C/100 m) in this study. The lapse rate is not always the same, and seasonality and spatial dependency of lapse rate have been widely reported (Fang and Yoda 1988; Pepin et al. 1999; Pepin 2001; Rolland 2003; Holden and Rose 2011). However, the calibrated temperature series using the constant lapse rate is generally acceptable, and much closer to the observations than the raw data.

For current climatology, compared with NCEP R2, ERA-Interim is generally closer to ground-based observation on a seasonal and annual basis. This is consistent with previous researches in China (Ma et al. 2008; Zhao et al. 2008; You et al. 2010, 2013). The difference between NCEP R2 and ERA-Interim is mainly attributed to the assimilation method for the two reanalysis datasets, and the different treatment for surface observations may create significant inconsistencies; 2-m temperature from ERA-Interim is derived by analyzing surface synoptic observations, and ERA-Interim incorporates the surface meteorological variables such as temperature in assimilation system when it is available. However, NCEP R2 is more dependent on free-air forcing, and does not assimilate surface observations (Simmons et al. 2004). Compared with ERA-Interim, NCEP R2 is not good at representing surface air temperatures but is more representative of free atmosphere temperature.

The effects of topographic correction also show a seasonal difference. The effects of topographic correction are better in summer than in winter. After calibration, the improvements in winter are very limited for both NCEP R2 and ERA-Interim. Inverse of temperature in the arid northwestern China (You and Liu 1995; Mamtimin and Meixner 2011) may be an important factor of this issue. In the reanalysis datasets, temperature inverse in winter are not well simulated. Hence, warmer winter can be found in the raw data, and the topographical corrections have limited influence on the winter temperature.

Generally, evaluation and correction of reanalysis datasets is a crucial work before the gridded data are applied in climate research. The calibration of topographical errors is a necessary step to evaluate the uncertainties of reanalysis surface temperature, especially for the regions with complex landform.

Acknowledgments This study was supported by the National Basic Research Program of China (973 Program) (no. 2013CBA01801) and the National Natural Science Foundation of China (no. 41161012). The authors are very grateful to the editor and other anonymous reviewer for their constructive suggestions.

References

Bengtsson L, Hagemann S, Hodges KI (2004) Can climate trends be calculated from reanalysis data? J Geophys Res 109. doi:10.1029/2004JD004536

Chen Z, Chen Y, Li W (2012) Response of runoff to change of atmospheric 0 °C level height in summer in arid region of Northwest China. Sci China Earth Sci 55:1533–1544. doi:10.1007/s11430-012-4472-6

Cornes RC, Jones PD (2013) How well does the ERA-Interim reanalysis replicate trends in extremes of surface temperature across Europe? J Geophys Res 118:10262–10276. doi:10.1002/jgrd.50799

Cosgrove BA, Lohmann D, Mitchell KE, Houser PR, Wood EF, Schaake JC, Robock A, Marshall C, Sheffield J, Duan Q, Luo L, Wayne Higgins R, Pinker RT, Dan Tarpley J, Meng J (2003) Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. J Geophys Res 108. doi:10.1029/2002JD003118

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

Fang J-Y, Yoda K (1988) Climate and vegetation in China (I). Changes in the altitudinal lapse rate of temperature and distribution of sea level temperature. Ecol Res 3:37–51. doi:10.1007/BF02348693

Holden J, Rose R (2011) Temperature and surface lapse rate change: a study of the UK's longest upland instrumental record. Int J Climatol 31:907–919. doi:10.1002/joc.2136

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds

R, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77: 437–471. doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

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

Kistler R, Kalnay E, Collins W, Saha S, White G, Woollen J, Chelliah M, Ebisuzaki W, Kanamitsu M, Kousky V, Van den Dool H, Jenne R, Fiorino M (2001) The NCEP-NCAR 50-year reanalysis: monthly means CD-ROM and documentation. Bull Am Meteorol Soc 82:247–267. doi:10.1175/1520-0477(2001)082<0247:TNNYRM> 2.3.CO;2

Kostopoulou E, Giannakopoulos C, Holt T, Le Sager P (2010) Assessment of interpolated ERA-40 reanalysis temperature and precipitation against observations of the Balkan Peninsula. Theor Appl Climatol 102:115–124. doi:10.1007/s00704-009-0249-z

Li Q, Liu X, Zhang H, Peterson TC, Easterling DR (2004) Detecting and adjusting temporal inhomogeneity in Chinese mean surface air temperature data. Adv Atmos Sci 21:260–268. doi:10.1007/BF02915712

Li B, Chen Y, Shi X (2012) Why does the temperature rise faster in the arid region of northwest China. J Geophys Res 117. doi:10.1029/2012JD017953

Liu Z, Xu Z, Yao Z, Huang H (2012) Comparison of surface variables from ERA and NCEP reanalysis with station data over eastern China. Theor Appl Climatol 107:611–621. doi:10.1007/s00704-011-0501-1

Liu X, Zhang D, Luo Y, Liu C (2013) Spatial and temporal changes in aridity index in northwest China: 1960 to 2010. Theor Appl Climatol 112:307–316. doi:10.1007/s00704-012-0734-7

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

Mamtimin B, Meixner FX (2011) Air pollution and meteorological processes in the growing dryland city of Urumqi (Xinjiang, China). Sci Total Environ 409:1277–1290. doi:10.1016/j.scitotenv.2010.12.010

Mooney PA, Mulligan FJ, Fealy R (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. doi:10.1002/joc.2098

Pepin N (2001) Lapse rate changes in northern England. Theor Appl Climatol 68:1–16. doi:10.1007/s007040170049

Pepin N, Benham D, Taylor K (1999) Modeling lapse rates in the maritime uplands of Northern England: implications for climate change. Arct Antarct Alp Res 31:151–164. doi:10.2307/1552603

Rolland C (2003) Spatial and seasonal variations of air temperature lapse rates in alpine regions. J Clim 16:1032–1046. doi:10.1175/1520-0442(2003)016<1032:SASVOA>2.0.CO;2

Screen JA, Simmonds I (2011) Erroneous Arctic temperature trends in the ERA-40 Reanalysis: a closer look. J Clim 24:2620–2627. doi:10.1175/2010JCLI4054.1

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 ACM National Conference. 517–524. doi:10.1145/800186.810616

Shi Y, Shen Y, Kang E, Li D, Ding Y, Zhang G, Hu R (2007) Recent and future climate change in northwest China. Clim Chang 80:379–393. doi:10.1007/s10584-006-9121-7

Simmons AJ, Jones PD, Bechtold VD, Beljaars ACM, Kallberg PW, Saarinen S, Uppala SM, Viterbo P, Wedi N (2004) Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analyses of surface air temperature. J Geophys Res 109. doi:10.1029/2004JD005306

Uppala SM, KÅllberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, Mcnally AP, Mahfouf J-F, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131:2961–3012. doi:10.1256/qj.04.176

Wang XL (2008) Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal *t* or *F* test. J Appl Meteorol Climatol 47:2423–2444. doi:10.1175/2008JAMC1741.1

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

Wang Y, Wang J, Qi Y, Yan C (2005) Dataset of desert distribution in China (1:100,000). Environmental and Ecological Science Data Center for West China. National Natural Science Foundation of China, Lanzhou, China. doi:10.3972/westdc.006.2013.db

Wang H, Chen Y, Chen Z (2013a) Spatial distribution and temporal trends of mean precipitation and extremes in the arid region, northwest of China, during 1960–2010. Hydrol Process 27:1807–1818. doi:10.1002/hyp.9339

Wang H, Chen Y, Xun S, Lai D, Fan Y, Li Z (2013b) Changes in daily climate extremes in the arid area of northwestern China. Theor Appl Climatol 112:15–28. doi:10.1007/s00704-012-0698-7

Xu W, Li Q, Wang XL, Yang S, Cao L, Feng Y (2013) Homogenization of Chinese daily surface air temperatures and analysis of trends in the extreme temperature indices. J Geophys Res 118. doi:10.1002/jgrd.50791

Yao T, Wang Y, Liu S, Pu J, Shen Y, Lu A (2004) Recent glacial retreat in High Asia in China and its impact on water resource in Northwest China. Sci China Ser D Earth Sci 47:1065–1075. doi:10.1360/03yd0256

You L, Liu Y (1995) Some microphysical characteristics of cloud and precipitation over China. Atmos Res 35:271–281. doi:10.1016/0169-8095(94)00023-7

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

You Q, Fraedrich K, Ren G, Pepin N, Kang S (2013) Variability of temperature in the Tibetan Plateau based on homogenized surface stations and reanalysis data. Int J Climatol 33:1337–1347. doi:10.1002/joc.3512

Zhao T, Guo W, Fu C (2008) Calibrating and evaluating reanalysis surface temperature error by topographic correction. J Clim 21: 1440–1446. doi:10.1175/2007JCLI1463.1

Zhao X, Tan K, Zhao S, Fang J (2011) Changing climate affects vegetation growth in the arid region of the northwestern China. J Arid Environ 75:946–952. doi:10.1016/j.jaridenv.2011.05.007