https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

1 km Monthly Precipitation and Temperatures Dataset for China from 1952 to 2019 based on a Brand-New and High-Quality Baseline Climatology Surface

Haibo Gong[1, 2, 3, 4, 5], Xueqiao Xiang[1, 2, 3, 4, 5], Huiyu Liu* [1, 2, 3, 4, 5], Xiaojuan Xu[1, 2, 3, 4, 5], Fusheng Jiao[1, 2, 3, 4, 5], Zhenshan Lin[1, 2, 3, 4, 5]

[1]Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing Normal University, Nanjing, 210023, China

[2]Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China

[3]State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing Normal University, Nanjing, 210023, China

[4]College of Geography Science, Nanjing Normal University, Nanjing 210023, China

[5]Jiangsu Key Laboratory of Environmental Change and Ecological Construction, Nanjing Normal University, Nanjing 210023, China

Correspondence: Dr Huiyu Liu, Nanjing Normal University, No 1 Wenyuan Road, Qixia District, Nanjing, China. Tel: (+86) 18951838599; E-mail: liuhuiyu@njnu.edu.cn

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Abstract

Long-term climate data and high-quality baseline climatology surface with high resolution are highly essential to multiple fields in climatological, ecological, hydrological, and environmental sciences. Here, we created a brand-new baseline climatology surface (ChinaClim_baseline) and developed a 1km monthly precipitation and temperatures dataset in China during 1952-2019 (ChinaClim_timeseries). Thin plate spline (TPS) algorithm in each month with different model formulations by accounting for satellite-driven products, was used to generate ChinaClim_baseline and monthly climate anomaly surface. Meanwhile, climatologically aided interpolation (CAI) was used to superimpose monthly anomaly surface with ChinaClim_baseline to generate ChinaClim_timeseries. Our results showed that ChinaClim_baseline exhibited very high performance. For precipitation estimation, the values of all R^2 were over 0.860, and the values of *RMSEs* and *MAEs* were 8.149 mm~21.959 mm and 2.787~14.125 mm, respectively. Annual average, maximum and minimum temperature had average R^2 of 0.967~0.992, *MAEs* of 0.321~0.785 °C, and *RMSEs* between 0.485 °C and 1.233 °C for all months. ChinaClim_baseline performed much better than WorldClim2 and CHELSA, especially in summer months and the regions with low-density weather stations in temperate continental and high cold Tibetan Plateau climate zones. For ChinaClim_timeseries, precipitation had an average R^2 of 0.699~0.923, an average *RMSE* between 7.449 mm and 56.756 mm, and an average of *MAE* of 4.263~40.271 mm for all months. Temperature elements had an average R^2 of 0.936~0.985, an average *RMSE* between 0.807 °C and 1.766 °C, and an average *MAE* of 0.548~1.236 °C for all months. Compared with Peng's climate surface and CHELSAcruts, R^2 increased by approximately 6 %, *RMSE* and *MAE* decreased by approximately 15 % for precipitation; R^2 of temperatures had no obvious changes, but *RMSE* and *MAE* decreased by 8.37~34.02 %. The results showed that the performance of ChinaClim_timeseries in interannual variations performed much better than other datasets, thanks to the help of ChinaClim_baseline and satellite-driven products. Remarkably, ChinaClim_baseline did not greatly improve precipitation estimation, but it deeply improved temperature estimation; the satellite-driven TRMM3B43 anomaly can greatly improve precipitation estimation, while the LST anomaly did not substantially improve temperature estimation. ChinaClim_baseline can be used as an excellent baseline

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

climatology surface for obtaining high-quality and long-term climate datasets from past to future. In the meantime, ChinaClim_timeseries of 1km spatial resolution based on ChinaClim_baseline, is very suitable for investigating the spatial-temporal climate changes and their impacts on eco-environmental systems in China. Now, ChinaClim_baseline is available at https://doi.org/10.5281/zenodo.4287824 (Gong, 2020a), ChinaClim_timeseries of precipitation is available at https://doi.org/10.5281/zenodo.4288388 (Gong, 2020b), ChinaClim_timeseries of maximum temperature is available at https://doi.org/10.5281/zenodo.4288390 (Gong, 2020c) and ChinaClim_timeseries of minimum temperature is available at https://doi.org/10.5281/zenodo.4288392 (Gong, 2020d).

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

1 Introduction

Long-term information on climatic conditions is pivotal for understanding global changes including atmospheric movements, vegetation dynamics, soil moisture, and other related scientific and application fields which are conducted at a resolution of ~1 km (Chaney et al., 2014; Gao et al., 2018; Hijmans et al., 2005; Karger et al., 2017; Liu et al., 2016; New et al., 2002; Pfister et al., 2020; Wagner and Wolfgang, 2003). However, global climate datasets often only represent climatic variation at spatial resolutions of 0.25~1 degree, such as Climatic Research Unit: CRU (Harris et al., 2014), The European Centre for Medium-Range Weather Forecast (ECWMF) Climatic reanalysis: ERA (Sterl et al., 1998), Global Precipitation and temperature: UDEL (Lawrimore et al., 2011), The Berkeley Earth Surface Temperatures: BEST (Muller et al., 2013), Global Precipitation Climatology Centre: CPCC (Becker et al., 2013). Thus, high resolution gridded climate data is urgently needed for studying global and regional climate change and its influences (Hamann et al., 2015; Hijmans et al., 2005; Karger et al., 2017).

A variety body of work was motivated to obtain high resolution gridded climate data with spatial interpolation methods and statistical downscaling (Li and Shao, 2010; Wu and Li, 2013; Hartkamp et al., 1999; Boer et al., 2001). Spatial interpolation methods including Kriging, Inverse Distance Weighting and Spline were widely applied in the estimates of climate elements (temperature, precipitation, vapor pressure, solar radiation and wind speed) at arbitrary spatial resolution. Among them, thin plate spline (TPS) interpolation was considered to perform well in generating grids of climate elements (Boer et al., 2001; Hartkamp et al., 1999; Hijmans et al., 2005; Hutchinson, 1995; Fick et al., 2017). However, for the estimates of long-term monthly climate surface, recent studies have shown that climatologically aided interpolation (CAI) employing the temporal anomaly surface and an accurate baseline climatology surface with high resolution, is well suited for producing long-term climate datasets than direct interpolation using original weather stations (Abatzoglou et al., 2018; Becker et al., 2013; C. Vega et al., 2017; Karger et al., 2017; Mosier et al., 2014; Peng et al., 2019; Willmott and Robeson, 2010). Remarkably, the quality of monthly climate surface, generated by CAI method, was thought to be determined by the baseline climatology surface (Gao et al., 2018; Peng et al., 2019). Baseline climatology surface, also called 30-Year

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Normals, described the average monthly conditions over the most recent three full decades. Previous efforts have developed many high-quality baseline climatology surfaces with a resolution of ~1km, such as WorldClim (Hijmans et al., 2005), WorldClim2 (Fick et al., 2017) and CHELSA (Karger et al., 2017) for global land surface, PRISM (Daly et al., 2002; Daly et al., 2008) and Daymet (Thornton et al., 1997) for North America. Although these baseline climatology surfaces are widely used for basic and applied studies such as climatological, ecological, hydrological, and environmental fields (Belda et al., 2017; Ray et al., 2015), a gap between these gridded climate datasets and weather stations was still observed in many areas with complex topography and sharp gradient changes due to lacking of sufficient weather stations information (New et al., 2002; Fick et al., 2017). Data quality of WorldClim was thought to depend on local climate variability, quality and density of observations, and the degree of the fitted spline (Hijmans et al.,2005). Unfortunately, for currently available high-quality baseline climatology surface with high-resolution covering China like WorldClim2 and CHELSA, only a small part of weather stations (323 and 228 stations for WorldClim2 and CHELSA respectively) were employed to generate baseline climatology surface. Weather stations are the most reliable source of the estimation of temperatures and precipitation, and thus more weather stations can provide more accurate point measure information. In fact, we can use a dataset of 30-year average climate (1980-2010) containing more than 2000 weather stations from China Meteorological Data Service Center (CMD: http://cdc.nmic.cn) and Central Weather Bureau (www.cwb.gov.tw),which are bound to greatly improve the quality of the baseline climatology surface for China.

Previous efforts have shown that the estimate of climate elements is likely to be improved by using satellite-driven products in the regions with insufficient weather station density (or quality) (Deblauwe et al., 2016; Jin and Dickinson, 2010; Mildrexler et al., 2011). With the development of remote sensing and geographic information technology, satellite-driven climate grid products become the optimum source in measuring climate elements at regional and global scales (Huffman et al., 2010; Michaelides et al., 2009; Siuki et al., 2017). The TRMM Multisatellite Precipitation Analysis (TMPA) monthly 3B43 products have been utilized extensively to provide valuable precipitation information in areas with sparse weather stations over the last two decades (Biasutti et al., 2012; Huffman et al., 2010; Simpson et al., 1996). Land surface temperature (LST) is now

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

available from satellite-borne instruments, which is widely incorporated in estimating air temperature (Kilibarda et al., 2014, Yao et al., 2020). Both WorldClim2 and CHELSA have not considered satellite-driven precipitation products and CHELSA have not considered satellite-driven temperature products. Despite many studies have shown these TRMM3B43 and LST products can increase the accuracy of the estimates of precipitation and temperature (Kilibarda et al., 2014; Kolios and Kalimeris, 2020; Yao et al., 2020), they are only available after 1997 and 2000 respectively, which is not long enough for the long-term ecological and hydrological analyses and modeling. Therefore, there is an urgent need to combine satellite-driven TRMM3B43 and LST in climate interpolation to generate a brand-new and higher-quality baseline climatology surface (ChinaClim_baseline), and further to combine ChinaClim_baseline to create a high-quality monthly time series of precipitation and temperatures dataset with high spatial resolution for China (ChinaClim_timeseries) from 1952 to 2019 with CAI method.

Specifically, the objectives of this work are: (1) to create a brand-new and higher-quality baseline climatology surface for China (ChinaClim_baseline). (2) to generate a 1 km monthly temperatures and precipitation dataset in China for the period of 1952-2019 (ChinaClim_timeseries).

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

2 Data

2.1 Weather observation stations

Dataset of 30-year average climate (1980-2010) was obtained from two sources, 2438 weather stations from CMD and 25 weather stations from Central Weather Bureau. Dataset of monthly surface observation values drawn from 613 weather stations for the period of 1952-2019 was collected from CMD. Influenced by the monsoon and Tibetan Plateau, four climate zones (Fig.1: Temperate continental, Temperate monsoonal, High cold Tibetan Plateau, and Subtropical-tropical monsoonal climate zones) have experienced various climate changes in both precipitation and temperature (He et al 2018), and so, weather stations were divided into four zones to assess the accuracy of data products in the areas with sparse and dense weather stations.

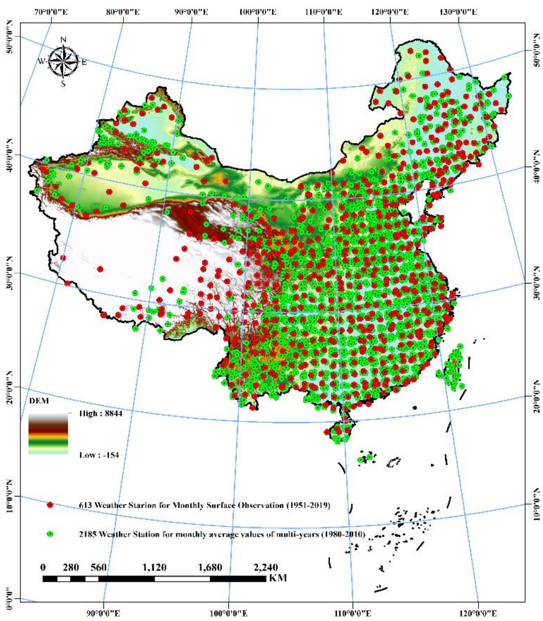

Figure1. The spatial distribution of weather stations in four climate zones (i.e. Temperate continental, Temperate monsoonal, High cold Tibetan Plateau, and Subtropical-tropical monsoonal climate zones) of China. (Map created by myself)

2.2 Version 7 TRMM3B43 datasets and Land Surface Temperature

The Tropical Rainfall Measuring Mission (TRMM), a joint project by the National Aeronautics and

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Space Administration (NASA) of the United States and the Japan Aerospace Exploration Agency (JAXA), was launched in November 1997 to monitor and investigate the tropical and subtropical rain system (Huffman et al., 2010; Simpson et al., 1996). The Version 7 monthly TRMM3B43 in NetCDF format was downloaded from https://mirador.gsfc.nasa.gov, with a spatial resolution of 0.25 degree over a latitude range from 50°S to 50°N during 1998-2019. It was resampled to ~1km spatial resolution via bilinear interpolation, and was averaged to get monthly and yearly TRMM3B43. Land surface temperature (LST) was compiled from Moderate Resolution Imaging Spectroradiometer (MODIS). Mean night and day LST values were extracted from ~1 km resolution MOD11A2 images, averaged by month and year from 2001 to 2019. The MOD11A2 images can be freely available at https://ladsweb.modaps.eosdis.nasa.gov.

2.3 Elevation and distance to the nearest coast

Elevation data with a spatial resolution of 30 m from Shuttle Rader Topography Mission (STRM) (data available at http://srtm.csi.cgiar.org/) was aggregated to ~1km spatial resolution. Coastline dataset was downloaded from https://www.ngdc.noaa.gov/mgg/shorelines/. We calculated the distance to the nearest coast using Euclidean distance in ArcGIS 10.2 with the fine coastline datasets.

2.4 Baseline climatology surfaces and monthly climate datasets used for comparison

Two baseline climatology surfaces as WorldClim2 (Fick et al., 2017) and CHELSA (Karger et al., 2017) with 1km spatial resolution were used to compare the accuracy of ChinaClim_baseline. WorldClim2 was interpolated with ANUSPLIN (Hutchinson, 1995), a method that fits thin plate splines through station data in three dimensions: latitude, longitude, and elevation. WorldClim2 can be accessed online at www.worldclim.org. CHELSA is essentially a quasi-mechanistical statistical downscaling of the ERA-Interim reanalysis, with the temperature downscaling based on mean lapse rates and elevation, and the precipitation algorithm using geographic predictors including wind fields, exposure, and boundary layer height (Karger et al., 2017). CHELSA can be freely available at www.chelsa-climate.org.

We also collected two long-term climate datasets with high resolution. One is the recently published Peng's climate surfaces (Peng et al., 2019). This climate dataset was spatially downscaled from 30' Climatic Research Unit (CRU) time series dataset with the baseline climatology surface of WorldClim2 using CAI. This is a 1km dataset of monthly air temperatures at 2m and precipitation for China in the period of 1901-2017. Peng's climate surface can be freely available at www.zenodo.org. The other is CHELSAcruts, a delta changes monthly climate dataset for the years 1901-2016 including mean monthly maximum temperatures, mean monthly minimum temperatures, and monthly precipitation sum. Anomalies of the CRU TS 4.01 dataset were interpolated between all CRU TS grid cells and are then added (for temperature variables) or multiplied (in case of precipitation) to high resolution climate data from CHELSA (Karger et al., 2017). CHELSAcruts can be freely available at www.chelsa-climate.org.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

3 Method

3.1 Creation of baseline climatology surface over China (ChinaClim_baseline)

The monthly averaged precipitation and temperatures of multi-years (1980-2010) were interpolated with the thin plate spline (TPS) from R packages "fields". Spline models for the *N* observed data values z_i are fit as the following:

$$z_i = f(x_i) + a^T y_i + \lambda \qquad (i = 1, \ldots, N) \tag{1}$$

Where f is a smooth function of the spline independent variable x_i, a is a vector of linear coefficients for the independent covariates y_i. In this study, we considered longitude, latitude, elevation, distance to the nearest coast and satellite-driven variables to construct TPS model. We listed climate elements and variables used in TPS model for estimating ChinaClim_baseline in Table 1. It is worth noting that longitude, latitude and elevation were set as spline independent variables and the other variables were used as either independent spline variables or linear covariates. Especially, Elevation (m) was divided by 1000 following scaling recommendations by Hutchinson (1995). Precipitation values were square root transformed prior to fitting following recommendations by Hutchinson and Xu (2013). Moreover, TRMM3B43 contained a latitude range from 50°S to 50°N, so we constructed TPS model including TRMM3B43 in the area south of 50°N and constructed TPS models without TRMM3B43 in the area north of 49°N. The 1° overlap area ensures that baseline climatology surface of the two areas can be better merged by weighting estimates inversely proportional to distance from each region's border (Hijmans et al., 2005; New et al., 2002).

Specifically, the process for generating ChinaClim_baseline can be described as follows (Fig.2):

(1) After removing duplicate and invalid weather stations, the remaining were split into 10 folds in each climate zones to assure that there was enough train and test data for each climate zones to build and verify the model, and thus to avoid spatial autocorrelation.

(2) We randomly extracted 9 folds weather stations in each climate zones and combined them into a new train data set. The remained were combined as test data set to valid the accuracy of model.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

(3) 14 model formulations for each month were tried using different combinations of variables to generate baseline climatology surface (Model formulations about Longitude, Latitude, Elevation, Distance to the nearest coast and Satellite-driven TRMM3B43 and LST described in Table S1).

(4) Each surface for each month was created by selecting only the model with the highest average R^2 value.

(5) Repeat steps 2 to 4 for 10 times, and final baseline climatology surface (ChinaClim_baseline) was created by averaging ten surfaces.

Table 1. Climate elements and variables used in TPS model for creating baseline climatology and anomaly surface.

Climate elements	Unit	Variables used in TPS models
Precipitation	mm	Lon, Lat, Elev, Coast,Trmm_m,Trmm_y
Minimum temperature	℃	Lon, Lat, Elev, Coast, Lst_nm, Lst_ny
Maximum temperature	℃	Lon, Lat, Elev, Coast, Lst_dm, Lst_dy
Average temperature	℃	Lon, Lat, Elev, Coast, Lst_am, Lst_ay
Precipitation anomaly	mm	Lon, Lat, Elev, Coast, Trmm_a(1998-2019), Base_prep
Minimum temperature anomaly	℃	Lon, Lat, Elev, Coast, Lst_na(2001-2019), Base_tmin
Maximum temperature anomaly	℃	Lon, Lat, Elev, Coast, Lst_da(2001-2019), Base_tmax
Average temperature anomaly	℃	Lon, Lat, Elev, Coast, Lst_aa(2001-2019), Base_tavg

Notes: Variables include longitude (Lon), latitude (Lat), elevation (Elev), distance to the nearest coast (Coast), averaged monthly (Trmm_m) and yearly (Trmm_y) TRMM3B43 during 1998-2019, monthly TRMM anomaly (Trmm_a), MOD11A2 land surface temperature (the day LST, the night LST, and the average of the day and nght LST) during 2001-2019 averaged by month (Lst_dm, Lst_nm, Lst_am) and year (Lst_dy Lst_ny, Lst_ay), MOD11A2 land surface temperature anomaly during 2001-2019 (Lst_da, Lst_na, Lst_aa), Baseline precipitation surface (Base_prep), Baseline temperatures surface (Base_tmin, Base_tavg, Base_tmax). Lon, Lat and Elev were set as spline independent variables and the other variables were set as either independent spline variables or linear covariates

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

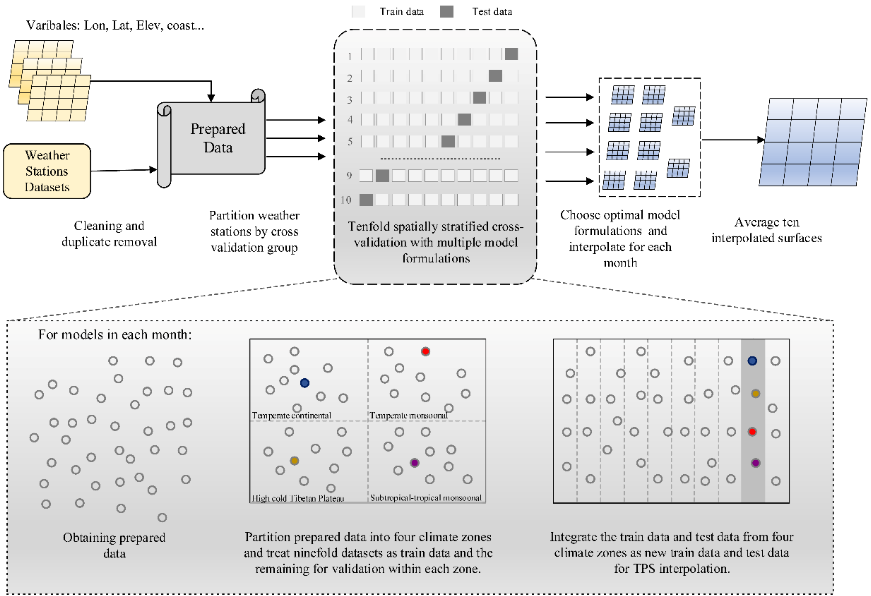

Figure2. Workflow for creating baseline climatology surface (ChinaClim_baseline) and monthly anomaly surface of China (adapted from Fick et al., 2017)

3.2 Generation of monthly precipitation and temperatures surface for China (ChinaClim_timeseries)

CAI method was used to superimpose monthly anomaly surface and baseline climatology surface (ChinaClim_baseline) to produce monthly precipitation and temperatures surface during 1952.01-2019.12 in China (ChinaClim_timeseries) as the following (Fig.3) ：

Firstly, the anomaly time series was calculated by the difference between the original time series from weather stations and ChinaClim_baseline described in Chapter 3.1.

Secondly, similar to the way to obtain ChinaClim_baseline (Fig.2), we applied TPS model to generate monthly anomaly surface from 1952.01 to 2019.12 with a segmented strategy. For monthly anomaly surface (Precipitation: 1952-1997; Temperature: 1952-2000), 7 model formulations were built by using different combinations of variables (Longitude, Latitude, Elevation, Distance to the nearest coast and ChinClim_baseline described in Table S7: Model1-7). Similarly, for monthly anomaly surface during 1998-2019 and 2001-2019 for Precipitation and Temperature, respectively, 14 model formulations were constructed using different combinations of variables (Longitude,

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Latitude, Elevation, Distance to the nearest coast, ChinaClim_baseline and Satellite-driven TRMM and LST anomaly described in Table S7: Model1-14). Monthly anomaly surface during 1952-2019 was created by selecting only the model with the highest average R^2 value.

Finally, ChinaClim_timeseries was generated by superimposing monthly anomaly surface and ChinaClim_baseline from 1952.01 to 2019.12.

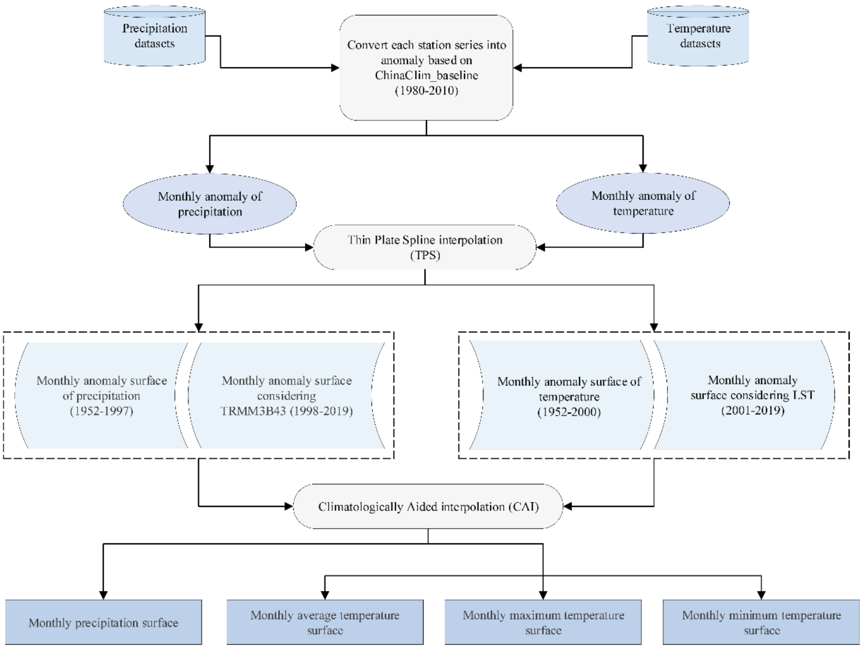

Figure3. Workflow for ChinaClim_timeseries generation based on climatologically aided interpolation (CAI).

3.3 Evaluation metrics

Three statistic indices including the root mean square error (*RMSE*), mean absolute error (*MAE*) and coefficients of determination (R^2) are examined to evaluate the performance of ChinaClim_baseline and ChinaClim_timeseries.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - M_i)^2}{n}} \tag{2}$$

$$MAE = \frac{\sum_{i=1}^{n}|P_i - M_i|}{n} \tag{3}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n} (M_i - \overline{M})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{n} (M_i - \overline{M})^2 (P_i - \overline{P})^2}} \right)^2 \quad (4)$$

Where *Pi* is the estimates like ChinaClim_baseline/ChinaClim_timeseries in the *i*th weather station; *Mi* is the measured value from the *i*th weather station; *n* is the number of weather stations; $\overline{P}$ is the average of the estimates like ChinaClim_baseline/ChinaClim_timeseries from *n* weather stations; $\overline{M}$ is the average of the measured value from *n* weather stations.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

4 Results

4.1 A brand-new and high-quality baseline climatology surface for China (ChinaClim_baseline)

4.1.1 The optimal models and its overall accuracy for ChinaClim_baseline

For the estimates of precipitation (Table S2), the models with the highest R^2 in each month all employed satellite-driven TRMM3B43, especially that monthly averaged TRMM (TRMM_m) improve the accuracy in all months, which indicated that TRMM3B43 improved the estimates of precipitation. For all temperature elements included average, maximum and minimum temperature (Table S3-5), models considering LST were superior in most months especially in summer months. Thus, LST could improve the interpolation of temperatures, while LST's ability to improve temperature estimation might be restricted in winter months, especially for average and minimum temperature. Overall, satellite-driven data can improve the estimates of precipitation and temperatures.

As shown from Table 3, ChinaClim_baseline exhibited very high performance over independent weather stations. Specifically, for precipitation estimation, the lowest value of R^2 was 0.860 in December, and the highest value was close to 0.98 in March and April, and the values of *RMSEs* and *MAEs* were 8.149~21.959 mm and 2.787~14.125 mm, respectively. Temperature elements had an average R^2 of 0.967~0.992, an average *RMSEs* between 0.485 and 1.233 °C, and an average *MAEs* of 0.321~0.785 °C for all months. Specifically, R^2 of all temperature elements were very high, but the *MAE* and *RMSE* of the average temperature were the smallest, followed by the maximum and minimum temperature. Moreover, the temperature estimation performed much better in summer months than in winter months with lower *RMSE* and *MAE*.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Table 3. Tenfold cross-validation statistics for ChinaClim_baseline. Coefficients of determination (R^2), root mean square error (*RMSE*) and mean absolute error (*MAE*) between observed and baseline climatology surface over independent weather stations.

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Precipitation	R^2	0.920	0.950	0.975	0.978	0.966	0.953	0.905	0.899	0.928	0.914	0.899	0.860
	RMSE	8.149	8.205	8.547	9.542	14.229	20.378	21.959	21.858	15.543	14.293	11.503	8.801
	MAE	2.787	3.171	4.252	5.392	7.685	11.561	14.125	13.289	8.578	5.732	3.701	2.484
Average temperature	R^2	0.990	0.989	0.989	0.987	0.984	0.982	0.985	0.986	0.988	0.990	0.993	0.991
	RMSE	0.902	0.822	0.676	0.576	0.539	0.520	0.498	0.485	0.519	0.580	0.654	0.809
	MAE	0.558	0.523	0.438	0.378	0.360	0.331	0.321	0.322	0.360	0.400	0.454	0.528
Maximum temperature	R^2	0.990	0.986	0.982	0.974	0.968	0.967	0.974	0.979	0.981	0.986	0.992	0.992
	RMSE	0.815	0.840	0.787	0.715	0.657	0.649	0.598	0.540	0.550	0.600	0.620	0.738
	MAE	0.464	0.482	0.458	0.425	0.382	0.392	0.366	0.334	0.335	0.343	0.379	0.440
Minimum temperature	R^2	0.985	0.985	0.986	0.984	0.982	0.980	0.983	0.984	0.983	0.986	0.987	0.985
	RMSE	1.233	1.106	0.884	0.771	0.719	0.659	0.609	0.618	0.716	0.797	0.934	1.149
	MAE	0.785	0.717	0.603	0.547	0.516	0.447	0.409	0.414	0.496	0.565	0.651	0.773

4.1.2 Comparison of ChinaClim_baseline to WorldClim2 and CHELSA in four climate zones.

To better assess the performance of ChinaClim_baseline, it was compared to two widely recognized baseline climatology surface with same spatial resolution: WorldClim2 (Fick et al., 2017) and CHELSA (Karger et al., 2017). The independent weather stations from a tenfold cross-validation approach were used to diagnose the performance of ChinaClim_baseline, while the independent weather stations extracted from CMD were used to calculate the accuracy of WorldClim2 and CHELSA. Considering that both worldClim2 and CHELSA used elevation to estimate temperature and precipitation, large deviations between the recorded and actual elevation (1 km DEM) in some weather stations might cause large discrepancy in the estimated and actual precipitation (Fick et al., 2017). Thus, only these independent weather stations with small deviations (< 200 m) between the recorded and actual elevation (1 km DEM) were used to assess the accuracy of WorldClim2 and CHELSA. (Figs 4, 6, 8 and 10). Moreover, spatial differences between ChinaClim_baseline and WorldClim2 as well as CHELSA for annual total precipitation, annual average temperature, January minimum temperature, and July maximum temperature were shown in Figs 5, 7, 9, and 11, respectively.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

As shown in Fig.4, despite relatively small differences in precipitation accuracy of three baseline climatology surfaces in Oct-Jun, ChinaClim_baseline greatly improved the accuracy of precipitation in Jul-Sep with lower *RMSE* and *MAE* along with higher R^2 in the all four climate zones. Specifically, in the temperate monsoonal and subtropical-tropical monsoonal zones with high-density weather stations, compared with precipitation accuracy of WorldClim2 and CHELSA, the accuracy of ChinaClim_baseline was much higher in most months. In high cold Tibetan Plateau, the accuracy of CHELSA was the worst with the highest *RMSE* and *MAE* in summer months. Although the *RMSE* and *MAE* of WorldClim2 were slightly lower than ChinaClim_baseline, its R^2 was by far lower than ChinaClim_baseline with huge seasonal variations. There are some spatial differences between ChinaClim_baseline and WorldClim2 and CHELSA for annual total precipitation (Fig.5). WorldClim2 tended to be drier than ChinaClim_baseline in many locations, especially at higher elevations. CHELSA was pretty wetter in high cold Tibetan Plateau and much drier in temperate continental than ChinaClim_baseline.

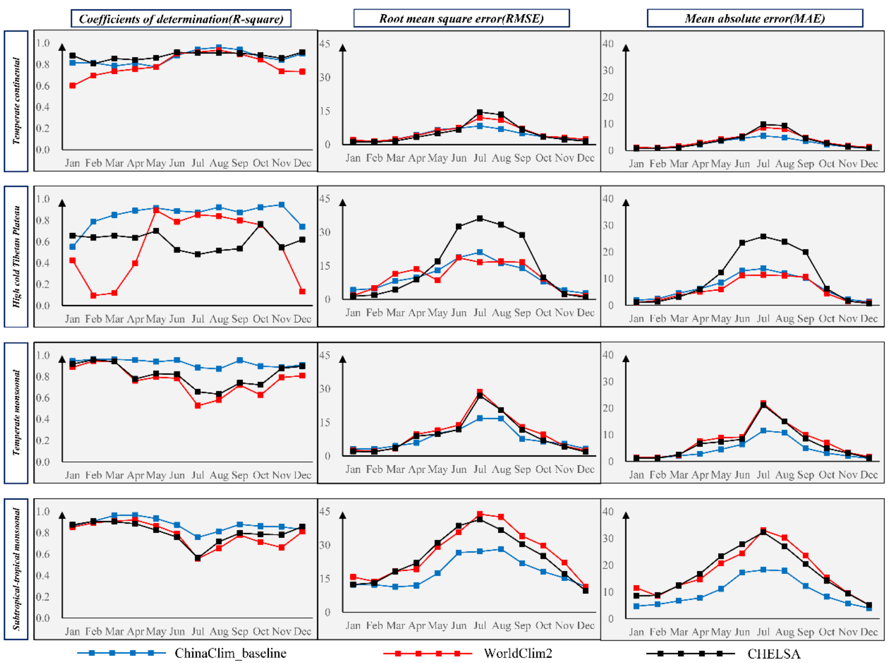

Figure 4. The accuracy of ChinaClim_baseline and WorldClim2 and CHELSA for precipitation in the temperate continental, high cold Tibetan Plateau, temperate monsoonal, and subtropical-tropical monsoonal climate zones.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

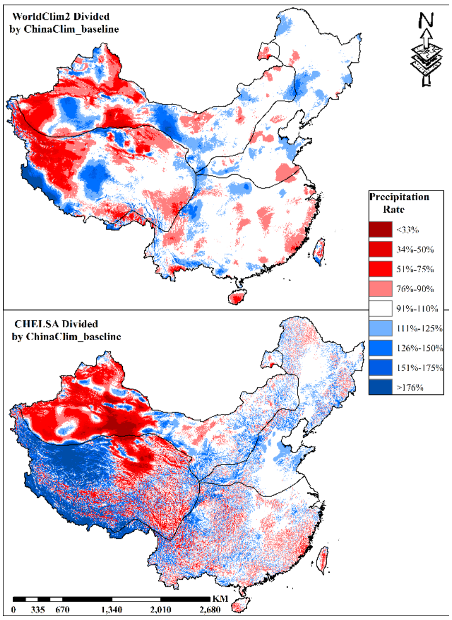

Figure 5. WorldClim2/ ChinaClim_baseline and CHELSA/ ChinaClim_baseline ratio maps (expressed as percentage) of annual precipitation for China.

For temperature elements included average, maximum and minimum temperature, the performance of ChinaClim_baseline was the most excellent, followed by WorldClim2 and CHELSA (Figs 6, 8, and 10). Although there is no obvious discrepancy in the accuracy of temperature estimation in Oct-Mar among ChinaClim_baseline, WorldClim2, and CHELSA, ChinaClim_baseline improved R^2 and greatly reduced *RMSE* and *MAE* in Apr-Sep which were during the growing season of most plants. The temperature estimation of ChinaClim_baseline, WorldClim2 and CHELSA in temperate continental and high cold Tibetan Plateau climate zones performed worse (higher *RMSE* and *MAE* and lower R^2) relative to subtropical-tropical monsoonal and temperate monsoonal climate zones. However, due to more weather stations and powerful algorithms, ChinaClim_baseline tangibly improved the temperature estimation of WorldClim2 and CHELSA in the above two regions. Moreover, the spatial discrepancy between ChinaClim_baseline and WorldClim2 and CHELSA for temperatures were smaller than precipitation as temperature generally follows relatively simple gradients of latitude and elevation (Fig.5). For annual average temperature, most areas showed small differences within -1~1 °C, and worldClim2 and CHELSA were slightly hotter than ChinaClim_baseline. However, for July maximum temperature (Fig.7), they were colder than

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

ChinaClim_baseline except for temperate continental climate zone. In particular, CHELSA was colder than ChinaClim_baseline in vast high-altitude areas. For January minimum temperature (Fig.9), WorldClim2 was generally cooler than ChinaClim_baseline and CHELSA was seriously warmer than ChinaClim_baseline in most high-altitude areas. Overall, there were some spatial discrepancies between ChinaClim_baseline and WorldClim2 and CHELSA in many areas of China, especially in low-density weather station regions such as high cold Tibetan Plateau and temperate continental.

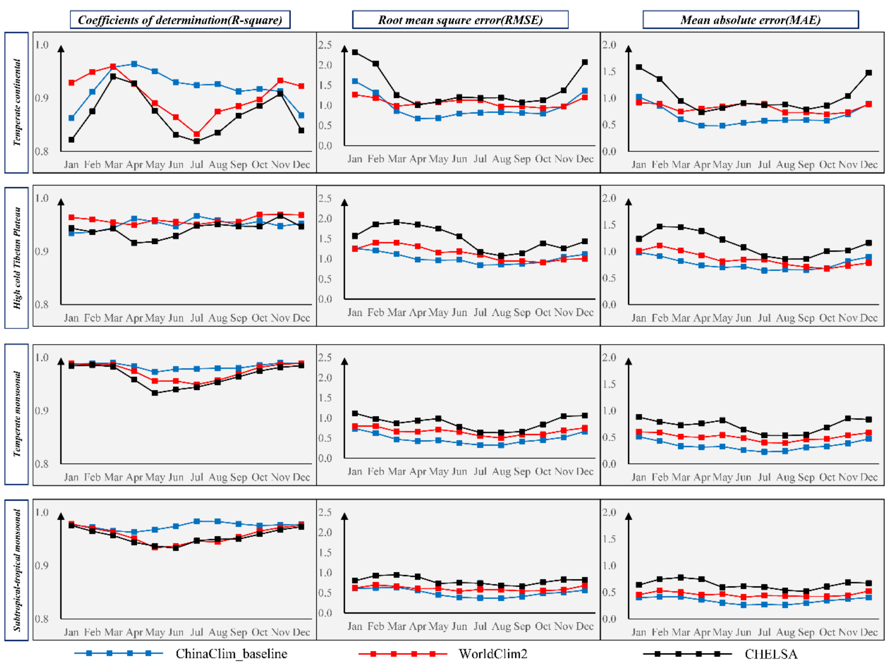

Figure 6. The accuracy of ChinaClim_baseline and WorldClim2 and CHELSA for average temperature in the temperate continental, high cold Tibetan Plateau, temperate monsoonal, and subtropical-tropical monsoonal climate zones.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

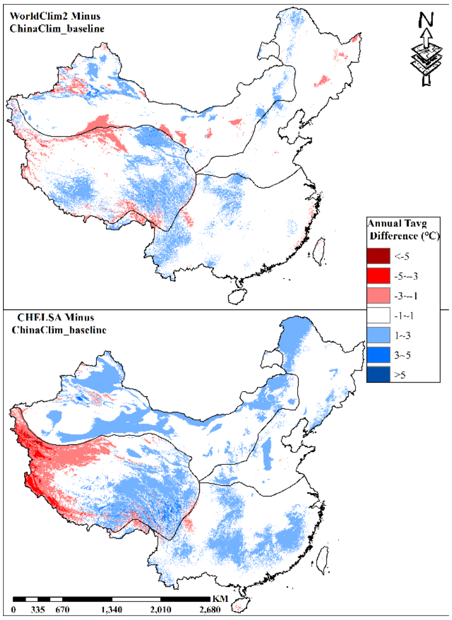

Figure 7. WorldClim2 - ChinaClim_baseline and CHELSA - ChinaClim_baseline difference maps of annual average temperature for China.

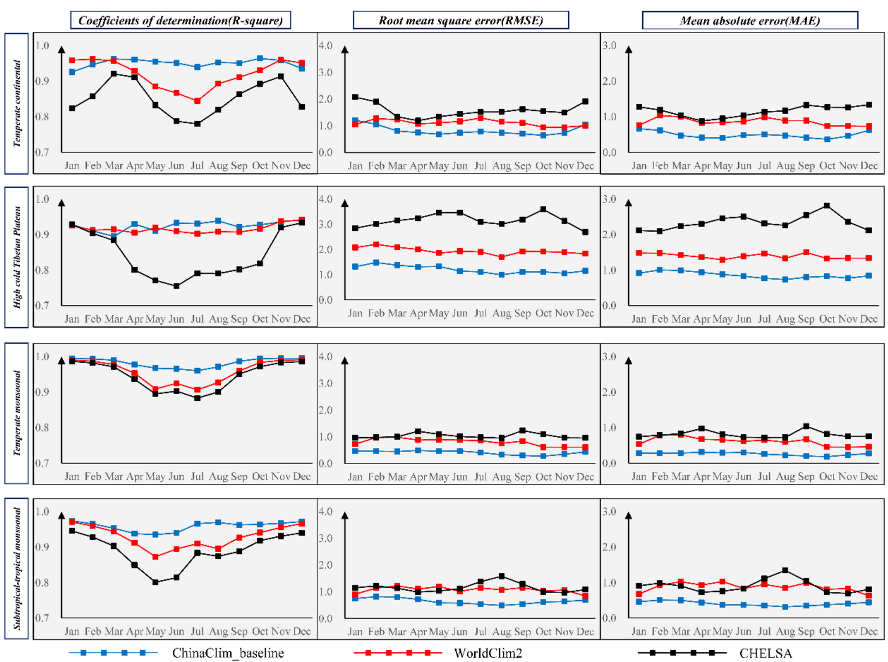

Figure 8. The accuracy of ChinaClim_baseline and WorldClim2 and CHELSA for maximum temperature in the temperate continental, high cold Tibetan Plateau, temperate monsoonal, and subtropical-tropical monsoonal climate zones.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

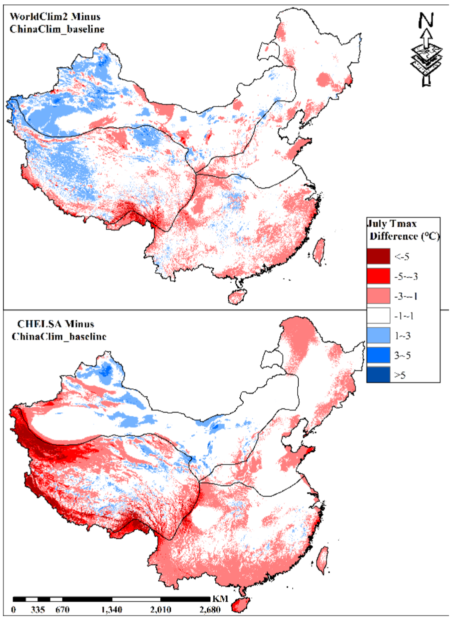

Figure 9. WorldClim2 - ChinaClim_baseline and CHELSA - ChinaClim_baseline difference maps of July maximum temperature for China.

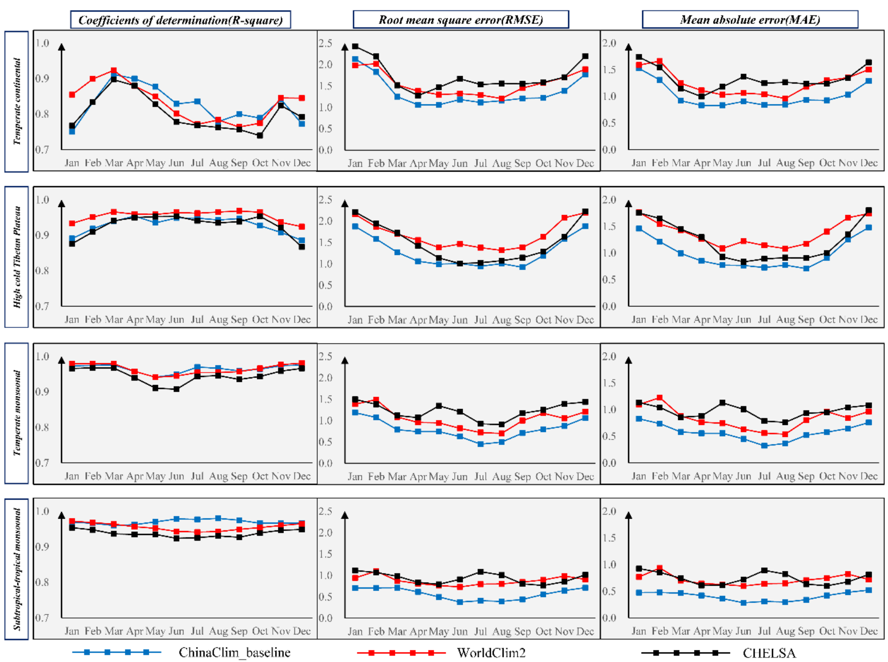

Figure 10. The accuracy of ChinaClim_baseline and WorldClim2 and CHELSA for minimum temperature in the temperate continental, high cold Tibetan Plateau, temperate monsoonal, and subtropical-tropical monsoonal climate zones.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

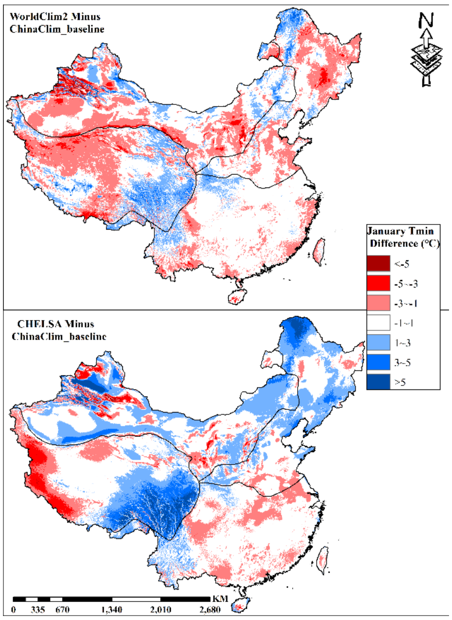

Figure 11. WorldClim2 - ChinaClim_baseline and CHELSA - ChinaClim_baseline difference maps of January minimum temperature for China.

4.2 1km monthly precipitation and temperatures surfaces during1952-2019 (ChinaClim_timeseries)

4.2.1 The optimal models and accuracy of ChinaClim_timeseries with seasonal variation

Models (Model4, Model5, and Model7) considering baseline climatology surface, showed better performance in all months for precipitation anomaly and temperatures anomaly during 1952-1997 and 1952-2000, respectively (Table S12). For precipitation anomaly during 1998-2019, models (Model 14) with the highest average R^2 value for each month all include TRMM3B43 anomaly (Table S17). However, for temperature anomaly during 2000-2019 (Table S17), those models (Model4, Model5 and Model7) that did not consider LST anomaly also exhibited excellent performance.

Our results demonstrated that ChinaClim_timeseries showed excellent performance during 1952-2019 (Table 4). Precipitation had an average R^2 of 0.699~0.923, an average *RMSE* between 7.449

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

mm and 56.756 mm, and an average of *MAE* of 4.263~40.271 mm for all months. Similarly, in terms of seasonal changes, compared with other months, the accuracy of precipitation was slightly worse from Jun to Sep. Average temperature had an average R^2 of 0.966~0.985, an average *RMSE* between 0.807 °C and 1.394 °C, and an average *MAE* of 0.548~0.930 °C for all months. Maximum temperature had an average R^2 of 0.939~0.981, an average *RMSE* between 0.935 °C and 1.391 °C, and an average *MAE* of 0.608°C ~0.877 °C for all months. Minimum temperature had an average R^2 of 0.968~0.977, an average *RMSE* between 0.924 °C and 1.766 °C, and an average *MAE* of 0.641~1.236 °C for all months. The performance of the average temperature was the best, followed by the maximum temperature and the minimum temperature.

Table 4. Tenfold cross-validation statistics for ChinaClim_timeseries during 1952-2019.

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Precipitation	R^2	0.898	0.917	0.921	0.887	0.855	0.799	0.710	0.699	0.746	0.809	0.838	0.845
	RMSE	7.512	9.714	15.326	24.374	36.983	49.620	56.806	54.126	37.955	22.966	13.751	7.592
	MAE	4.370	5.643	9.134	15.111	23.608	33.319	40.403	37.727	24.938	13.857	7.955	4.320
Average temperature	R^2	0.981	0.980	0.978	0.974	0.967	0.966	0.972	0.974	0.976	0.980	0.985	0.983
	RMSE	1.394	1.286	1.091	0.953	0.933	0.888	0.820	0.807	0.846	0.909	1.027	1.268
	MAE	0.930	0.859	0.736	0.653	0.627	0.584	0.554	0.548	0.589	0.636	0.722	0.866
Maximum temperature	R^2	0.977	0.973	0.966	0.953	0.939	0.941	0.951	0.958	0.961	0.967	0.980	0.981
	RMSE	1.391	1.350	1.242	1.144	1.116	1.065	0.987	0.940	0.935	1.010	1.075	1.249
	MAE	0.877	0.870	0.825	0.769	0.743	0.704	0.685	0.647	0.608	0.647	0.697	0.801
Minimum temperature	R^2	0.975	0.977	0.976	0.975	0.971	0.968	0.972	0.973	0.973	0.976	0.979	0.976
	RMSE	1.766	1.600	1.332	1.110	1.067	1.008	0.924	0.949	1.045	1.167	1.357	1.637
	MAE	1.236	1.125	0.902	0.803	0.770	0.703	0.641	0.666	0.753	0.852	0.979	1.160

4.2.2 Comparison of ChinaClim_timeseries to other datasets

Here, we compared the accuracy of ChinaClim_timeseries with Peng's climate surface (Peng et al., 2019) and CHELSAcruts (Karger et al., 2017) by R^2, *RMSE* and *MAE* in China and four climate zones (temperate continental, high cold Tibetan Plateau, temperate monsoonal and subtropical-tropical monsoonal climate zones). The independent weather stations extracted from a tenfold cross-

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

validation approach were used to diagnose the performance of ChinaClim_timeseries, while only these independent weather stations from CMD with small deviations (< 200 m) between the recorded and actual elevation (1 km DEM) were used to assess the accuracy of CHELSAcruts and Peng's climate surface.

The accuracy of ChinaClim_timeseries for precipitation estimation showed better performance than Peng's climate surface and CHELSAcruts in China with R^2 increased by 6.15 % and 5.68 %, *RMSE* decreased by 14.71 % and 15.36 % and *MAE* decreased by 15.15 % and 16.22 %, respectively (Table 5). Specifically, ChinaClim_timeseries and CHELSAcruts showed similar accuracy in Temperate continental; although ChinaClim_timeseries had a slightly lower R^2 than CHELSAcruts, CHELSAcruts had higher *RMSE* and *MAE*. Moreover, ChinaClim_timeseries in high cold Tibetan Plateau, R^2 increased by 13.08 % and 15.87 %, *RMSE* decreased by 27.46 % and 32.97 % and *MAE* decreased by 23.13 % and 30.81 %, respectively. Compared with CHELSA_cruts, except the R^2 of our product was slightly lower in the temperate continental, all indicators were obviously better in different zones, especially in high cold Tibetan Plateau and subtropical-tropical monsoonal. Remarkably, in terms of interannual variations (Fig.12), ChinaClim_timeseries performed slightly better than other datasets before 1998, while its accuracy was greatly improved during 1998-2019.

Table 5. The overall accuracy of precipitation for ChinaClim_timeseries, Peng's climate surface and CHELSAcruts in China and four climate zones

	Precipitation	*R2*	*RMSE*	*MAE*
China	ChinaClim_timeseries	**0.855**	**33.868**	**18.063**
	Peng's climate surface	0.805	39.707	21.290
	CHELSAcruts	0.809	40.015	21.560
temperate continental climate	ChinaClim_timeseries	0.822	**14.805**	**7.729**
	Peng's climate surface	0.791	16.575	8.881
	CHELSAcruts	**0.832**	15.043	7.892
high cold Tibetan Plateau	ChinaClim_timeseries	**0.807**	**22.942**	**12.454**
	Peng's climate surface	0.714	31.625	16.201
	CHELSAcruts	0.696	34.228	18.000
temperate monsoonal	ChinaClim_timeseries	**0.851**	**26.222**	**13.588**
	Peng's climate surface	0.817	29.151	15.496
	CHELSAcruts	0.831	28.819	15.375
subtropical-tropical monsoonal	ChinaClim_timeseries	**0.820**	**45.501**	**27.502**
	Peng's climate surface	0.758	52.426	31.612
	CHELSAcruts	0.760	52.950	32.364

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

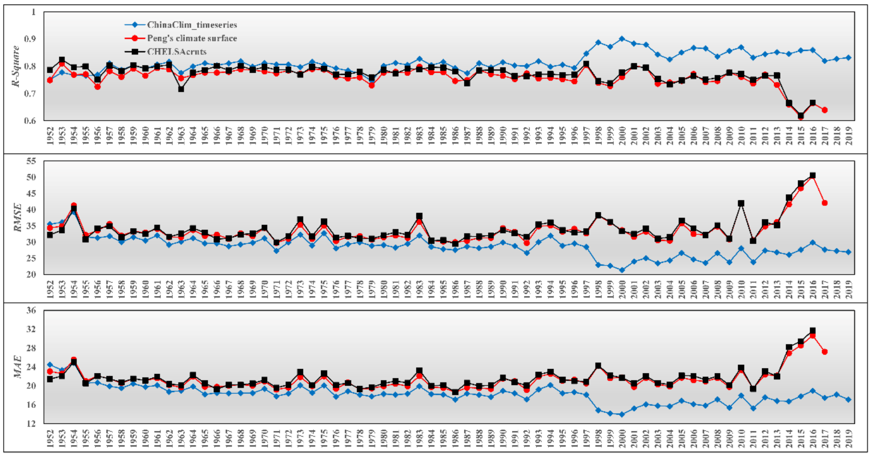

Figure 12. The accuracy of interannual variations in ChinaClim_timeseries, Peng's climate surface and CHELSAcruts for precipitation.

The accuracy of ChinaClim_timeseries for temperatures estimation also showed better performance than Peng's climate surface and CHELSAcruts in China and different climate zones (Tables 6-7). In China, the R^2, *RMSE* and *MAE* of maximum temperature were 0.989, 1.167 °C and 0.724 °C, respectively, and those of minimum temperature were 0.989, 1.303 °C and 0.892°C, respectively. Our results showed that all R^2 were very high among three datasets, but compared to Peng's and CHELSAcruts, our *RMSE* decreased by 10.17 % and 19.14 % for maximum temperature, and by 8.37 % and 14.42 % for minimum temperature; *MAE* decreased by 25.73 % and 34.02 % for maximum temperature and by 16.92 % and 20.74 % for minimum temperature. In different climate zones, the accuracy of ChinaClim_timeseries was much better than Peng's climate surface and CHELSAcruts. Especially in the High cold Tibetan Plateau, our *RMSE* decreased by 23.31 % and 36.52 % for maximum temperature and by 21.61 % and 9.65 % for minimum temperature, respectively; our *MAE* decreased by 39.61 % and 50.00 % for maximum temperature and by 29.35 % and 16.81 % for minimum temperature, respectively. Moreover, in terms of interannual variation (Figs 13-14), our results demonstrated that despite the accuracy of Peng's climate surface and CHELSAcruts also very well with high R^2, ChinaClim_timeseries undoubtedly showed more powerful performance in almost all years with lower *RMSE* and *MAE*.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

Table 6. The overall accuracy of maximum temperature for ChinaClim_timeseries, Peng's climate surface and CHELSAcruts in China and four climate zones.

	Maximum temperature	*R2*	*RMSE*	*MAE*
China	ChinaClim_timeseries	**0.989**	**1.167**	**0.724**
	Peng's climate surface	0.988	1.299	0.974
	CHELSAcruts	0.987	1.443	1.097
temperate continental	ChinaClim_timeseries	**0.989**	**1.346**	**0.799**
	Peng's climate surface	0.985	1.591	1.202
	CHELSAcruts	0.981	1.835	1.358
high cold Tibetan Plateau	ChinaClim_timeseries	**0.958**	**1.705**	**1.115**
	Peng's climate surface	0.951	2.224	1.847
	CHELSAcruts	0.947	2.686	2.231
temperate monsoonal	ChinaClim_timeseries	**0.996**	**0.766**	**0.519**
	Peng's climate surface	0.993	1.090	0.847
	CHELSAcruts	0.993	1.225	0.962
subtropical-tropical monsoonal	ChinaClim_timeseries	**0.980**	**1.107**	**0.712**
	Peng's climate surface	0.978	1.252	0.935
	CHELSAcruts	0.980	1.314	1.035

Table 7. The overall accuracy of minimum temperature for ChinaClim_timeseries, Peng's climate surface and CHELSAcruts in China and four climate zones.

	Minimum temperature	*R2*	*RMSE*	*MAE*
China	ChinaClim_timeseries	**0.989**	**1.303**	**0.892**
	Peng's climate surface	0.988	1.422	1.074
	CHELSAcruts	0.987	1.523	1.125
temperate continental	ChinaClim_timeseries	0.979	1.770	**1.270**
	Peng's climate surface	**0.982**	**1.765**	1.351
	CHELSAcruts	0.976	2.004	1.461
high cold Tibetan Plateau	ChinaClim_timeseries	**0.966**	**1.784**	**1.271**
	Peng's climate surface	0.944	2.276	1.800
	CHELSAcruts	0.958	1.975	1.528
temperate monsoonal	ChinaClim_timeseries	**0.992**	**1.202**	**0.871**
	Peng's climate surface	0.991	1.324	1.032
	CHELSAcruts	0.989	1.585	1.196
subtropical-tropical monsoonal	ChinaClim_timeseries	**0.987**	**0.885**	**0.621**
	Peng's climate surface	0.977	1.254	0.938
	CHELSAcruts	0.984	1.119	0.878

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

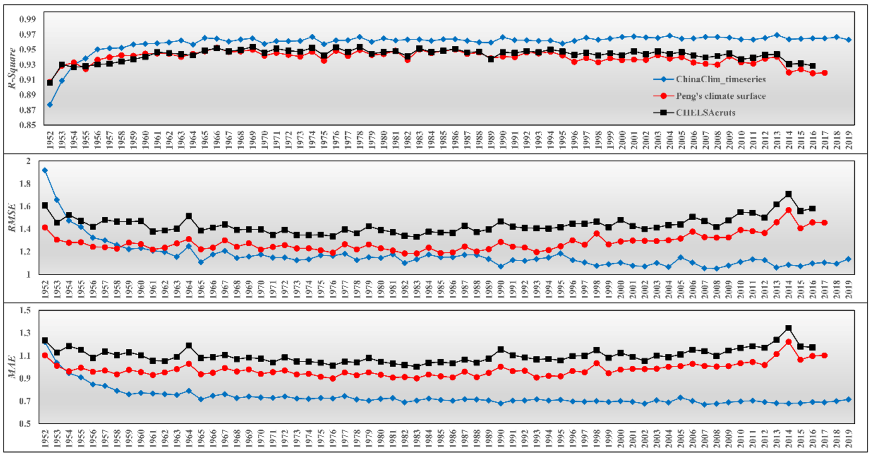

Figure 13. The accuracy of interannual variations in ChinaClim_timeseries, Peng's climate surfaces and CHELSAcruts for maximum temperature.

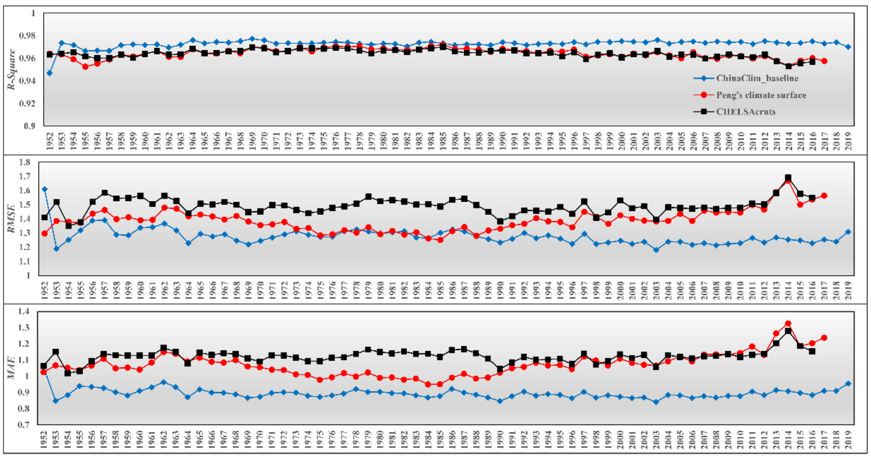

Figure 14. The accuracy of interannual variations in ChinaClim_timeseries, Peng's climate surface and CHELSAcruts for minimum temperature.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

5 Data availability

ChinaClim_baseline is a brand-new and high-quality baseline climatology surface for China at spatial resolution of 1km. The data now is freely available through Zenodo at https://doi.org/10.5281/zenodo.4287824 (Gong, 2020a), which can be downloaded in TIFF format. The scale factor of the data is 0.01.

ChinaClim_timeseries is a monthly temperatures and precipitation dataset in China for the period of 1952-2019 of 1km spatial resolution. The data now are freely available through Zenodo at https://doi.org/10.5281/zenodo.4288388 (Gong, 2020b) , https://doi.org/10.5281/zenodo.4288390 (Gong, 2020c) and https://doi.org/10.5281/zenodo.4288392 (Gong, 2020d), which can be downloaded in TIFF format. The scale factor of the data is 0.1.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

6 Discussion

There are a number of baseline climatology surface products for global land surface (Hijmans et al., 2005; Karger et al., 2017; New et al., 1999; New et al., 2002; Fick et al., 2017), while few weather stations are employed to generate these surfaces in China, which might result in insufficient accuracy of these surfaces, and further affect the availability of long-term climate datasets with these surfaces as input. Application TPS algorithm, we considered much more weather stations and satellite-driven variables in climate interpolation to create baseline climatology surface.

The precipitation estimation of ChinaClim_baseline performed well in all months with R^2 greater than 0.86. The *RMSEs* and *MAEs* in summer (Jun-Aug) were much higher than other months, which is also reported by the previous studies that the estimation of summer precipitation is pretty difficult than that of winter precipitation, especially in the monsoon zones (Chen et al., 2018; Fick et al., 2017). It is because summer precipitation is deeply affected by summer monsoon. However, compared to WorldClim2 and CHELSA, ChinaClim_baseline deeply improved the accuracy of precipitation with higher R^2, lower *RMSE*, and *MAE*, especially in temperate continental and high cold Tibetan Plateau zones and summer months. Because ChinaClim_baseline used much more weather stations, and the spatially continuous satellite-driven TRMM3B43 which can distinguish the rain shadow effect of mountains and provide enough information in sparse areas of weather station, while WorldClim2 and CHELSA cannot (Deblauwe et al., 2016). Moreover, our models constructed by each month could well reveal the seasonal variation of precipitation. So, our work could provide a good reference for accurately estimating precipitation (especially for precipitation in rainy season), which allows us to better understand the hydrological processes and execute more meaningful ecological modeling based on ChinaClim_baseline.

The accuracy of the average temperature was the best, followed by the maximum temperature, and the minimum temperature. In the meantime, the accuracy of summer temperatures was better than winter temperatures. It is not difficult to understand that summer temperature and maximum temperature often simply changes with elevation, while winter temperature, and minimum temperature, have a more complex relationship with elevation (Daly et al., 2008; Gustavsson et al., 1998). CHELSA simply used temperature lapse rates to estimate temperatures, which might make mistakes in temperature estimations without sufficient weather stations for corrections in high-altitude regions. Although WorldClim2 considered LST, it did not consider the effects of LST on the model accuracy for each month (Hijmans et al., 2005; Fick et al., 2017), which ignored the improvements helped by LST during key months such as vegetation growth season; In comparison with WorldClim2 and CHELSA, our model deeply improved the accuracy of temperature element during growing season facilitated by select optimal multiple model formulations for each month, which is very important for revealing the vegetation-climate relationship. Previous findings claimed that models only using latitude, longitude, elevation can be consistently superior for temperatures estimation (Parmentier et al., 2014), while our results showed that LST can greatly improve the accuracy of temperatures estimation, especially in summer months. Some studies showed that LST just improved the estimates for maximum temperature (Kilibarda et al., 2014), while our results found that LST improved not only the estimate of maximum temperature, but also the estimates of average and minimum temperature was greatly improved.

There was a large amount of evidence to suggested that the CAI method can better generate long-term monthly climate surface (Abatzoglou et al., 2018; Becker et al., 2013; C. Vega et al., 2017; Karger et al., 2017; Mosier et al., 2014; Peng et al., 2019; Willmott and Robeson, 2010). Our results proved Peng's climate surface and CHELSAcruts datasets, relying on coarse CRU anomaly and high-quality baseline climatology surfaces with CAI method, had relatively high accuracy (high R^2) with a few weather stations in China at 1km spatial resolution. (Karger et al., 2017; Peng et al., 2019). However, those studies rarely incorporated satellites-driven products into climate interpolation, and the performance of baseline climatology surface covering China (WorldClim2 and CHELSA) using the CAI method was also troubling, especially in high cold Tibetan Plateau. ChinaClim_timeseries used a higher precision baseline climatology surface (ChinaClim_baseline) as input in CAI method. We also implemented TPS interpolation by selecting optimal multiple model formulations for each month. Not only can we make full use of the time-series weather stations, but also consider the satellites-driven anomaly as either independent spline variables or linear covariates to further improve the accuracy of the final monthly climate surface. Our results showed that ChinaClim_timeseries was indeed a better climate dataset than Peng's climate surface and CHELSAcruts in China with higher R^2, and lower *RMSE* and *MAE*, especially in high cold

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Open Access Earth System Science Data Discussions

Tibetan Plateau for precipitation estimation, R^2 increased by 13.08 % and 15.87 %, *RMSE* decreased by 27.46 % and 32.97 % and *MAE* decreased by 23.13 % and 30.81 %, respectively.

Previous studies indicated that baseline climatology surface, considering detailed topographic information, the effects of distance to the nearest coast and satellite-driven variables, is physically representative and has a fine-scale distribution of meteorological variables (Marchi et al., 2019; Mosier et al., 2014; Peng et al., 2017; Platts et al., 2015). Thus, a superior baseline climatology surface is helpful to improve the accuracy of the long-term monthly climate surface. In this study, the baseline climatology surface was not only used as one of the inputs of CAI method, but also as one of the variables of TPS models to calculate the monthly anomaly surface. Our results showed that TPS models considering baseline climatology surface, showed better performance for precipitation and temperatures anomaly in all months during 1952-2019 (Table S12 and S17), which was helpful to improve the estimates of ChinaClim_timeseries. However, in terms of interannual variation, compared with other datasets, the estimates of ChinaClim_timeseries for precipitation performed slightly better during 1952-1997, while the performance was much better during 1998-2019 (Fig.12). Owing to the utilization of satellite-driven TRMM3B43 anomaly in climate interpolation after 1997, we believed that satellite-driven anomaly can greatly improve the estimates of precipitation and baseline climatology surface only slightly improve the estimates of precipitation. Remarkably, we only considered satellites-driven TRMM3B43 anomaly as either independent spline variables or linear covariates to generate the final monthly precipitation surface and this process was not implemented in the temperature estimation because the optimal TPS model did not reveal that LST anomaly can effectively improve the temperature estimation (Table S17). The temperatures estimation of ChinaClim_timeseries still performed well during 1952-2019 (Figs 13-14), which might be attributed to a better baseline climatology surface as input in CAI method. Although previous studies illustrated that satellites-driven products can greatly improve the accuracy of climate elements estimation (Kilibarda et al., 2014; Kolios and Kalimeris, 2020; Yao et al., 2020), our results showed that satellite-driven anomaly cannot substantially improve the estimates of temperatures and baseline climatology surface can play a key role in long-term temperature estimation.

As shown above, ChinaClim_baseline is a brand-new and high-quality baseline climatology surface

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

in China currently released. Baseline climatology surface, not only could be applied in history and paleo climate models, but also can be combined with GCM products to generate future climate change scenarios with high resolution (Peng et al., 2019; Platts et al., 2015). Besides, the quality of baseline climatology surface has a fundamental role in predictions of the potential impact of climate change on organisms and natural ecosystems (Marchi et al., 2019, Vega et al 2017). ChinaClim_timeseries is a very high-quality monthly climate surface and can successfully reveal the spatial-temporal change patterns of precipitation and temperatures for China. At the same time, it can be used as a good data source for long-term modeling of hydrology, ecology, and other related fields. In particular, ChinaClim_timeseries also could help to reduce the uncertainty of the input of climate parameters in high cold Tibetan Plateau zones, and better quantify the region's ecosystem variation in the context of global changes.

The TRMM3B43 improves the estimate of precipitation, while the 0.25 degree resolution of TRMM might be fail to represent many important finer-scale climatic features due to the uncertainties caused by the simply resampling process from 0.25 degree to 1km (Deblauwe et al., 2016). In the meantime, MODIS-LST was effective to improve the algorithms for estimating air temperatures (Jin and Dickinson, 2010; Mildrexler et al., 2011; Parmentier et al., 2014; Yao et al., 2020). However, land surface temperature is tightly related to land cover, which is itself highly affected by human activities. Therefore, incorporating TRMM3B43 and LST into the generation of ChinaClim_baseline and ChinaClim_timeseries maybe present challenges (Deblauwe et al., 2016). Besides, it should also be noted that there is a temporal mismatch between the datasets from weather stations (1980–2010) and from average TRMM3B43 (1998-2019) and LST (2001–2019) in estimating ChinaClim_baseline. With the emergence of high-resolution and long-term climate remote sensing products in the future, and the improvement of multiple remote sensing data fusion technology, we could greatly reduce the uncertainty of climate interpolation and improve the accuracy of product estimation, particularly in places with very few weather stations or strong gradients change or complex terrain (Immerzeel et al., 2009; Li and Shao, 2010; Fick et al., 2017; Vega et al 2017). Although our research showed that TPS method could be used well in climate interpolation, many studies have pointed out that this method accounted for direct elevation effects only, and had difficulty in considering the sharp changes in the relationship between climate and elevation (Daly et al., 2008; Daly et al., 2007; Marchi et al., 2019). Therefore, it is essential to comprehensively quantify the non-linear relationship between environmental variables and climate elements. Thus, future work ought to couple these nonlinear relationships with TPS or new algorithm for the better estimates of climate elements.

Author Contributions. Haibo Gong formed the original idea and wrote the original manuscript; Huiyu Liu offered valuable comments and was responsible for the manuscript revisions; Xueqiao Xiang and Xiaojuan Xu participated in the data collection and analysis; FuSheng Jiao and Zhenshan Lin created figures and tables.

Competing interests. The author declare that they have no conflict of interest

Acknowledgements. We thank the all people and institutions who contributed to the establishment of this dataset

Financial support. This research had been funded by the National Natural Science Foundation of China (No. 41971382, 31870454) and the Priority Academic Program Development of Jiangsu Higher Education Institutions (164320H116).

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., and Hegewisch, K.C.: TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015, Scientific Data., 5,170191, 2018.

Becker, A., et al.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, Earth System Science Data., 5,1(2013-02-21), 5(1), 921-998, 2013.

Belda, M., Holtanova, E., Kalvova, J., and Halenka, T.: Global warming-induced changes in climate zones based on CMIP5 projections, Climate Research., 71(1), 17-31, 2017.

Biasutti, M., Yuter, S.E., Burleyson, C.D., and Sobel, A.H.: Very high resolution rainfall patterns measured by TRMM precipitation radar: seasonal and diurnal cycles, Climate Dynamics., 39(1), 239-258, 2012.

Boer, E.P.J., Beurs, K.M.de., and Hartkamp, A.D.: Kriging and thin plate splines for mapping climate variables, International Journal of Applied Earth Observation and Geoinformation., 3(2), 146-154, 2001.

Vega, G.C., Pertierra, L.R., and Olalla-Tárraga, M.Á.: MERRAclim, a high-resolution global dataset of remotely sensed bioclimatic variables for ecological modelling, Scientific Data., 4(1), 170078, 2017.

Chaney, N.W., Sheffield, J., Villarini, G., and Wood, E.F.: Development of a High-Resolution Gridded Daily Meteorological Dataset over Sub-Saharan Africa: Spatial Analysis of Trends in Climate Extremes, Journal of Climate., 27(15), 5815-5835, 2014.

Chen, Y., et al.: A new downscaling-integration framework for high-resolution monthly precipitation estimates: Combining rain gauge observations, satellite-derived precipitation data and geographical ancillary data, Remote Sensing of Environment., 214, 154-172, 2018.

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L., and Pasteris, P.: A knowledge-based approach to the statistical mapping of climate, Climate Research., 22(2), 99-113, 2002.

Daly, C., et al.: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, International Journal of Climatology., 28(15) , 2008.

Daly, C., Smith, J.W., Smith, J.I., and Mckane, R.B.: High-Resolution Spatial Modeling of Daily Weather Elements for a Catchment in the Oregon Cascade Mountains, United States, Journal of Applied Meteorology & Climatology., 46(10): 1565-1586, 2007.

Deblauwe, V., et al.: Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics, Global Ecology & Biogeography; 25(4): 443-454, 2016.

Gao, L., et al.: A high-resolution air temperature data set for the Chinese Tian Shan in 1979–2016, Earth Syst. Sci. Data., 10(4): 2097-2114, 2018.

Gong, H.: A Brand-New and High-Quality Baseline Climatology Surface for China (ChinaClim_baseline), Zenodo, https://doi.org/10.5281/zenodo.4287824, 2020a

Gong, H.: 1 km Monthly Precipitation Dataset for China from 1952 to 2019 (ChinaClim_timeseries), Zenodo, https://doi.org/10.5281/zenodo.4288388, 2020b

Gong, H.: 1 km Monthly Maximum Temperature Dataset for China from 1952 to 2019 (ChinaClim_timeseries), Zenodo, https://doi.org/10.5281/zenodo.4288390, 2020c

Gong, H.: 1 km Monthly Minimum Temperature Dataset for China from 1952 to 2019 (ChinaClim_timeseries), Zenodo, https://doi.org/10.5281/zenodo.4288392, 2020d

Gustavsson, T.R., Karlsson, M., Bogren, J.R., and Lindqvist, S.: Development of Temperature Patterns during Clear Nights, J. Appl. Meteorol., 37(6): 559-571, 1998.

Hamann, A., Roberts, D.R., Barber, Q.E., Carroll, C., and Nielsen, S.E.: Velocity of climate change algorithms for guiding

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

conservation and management, Glob. Change Biol., 21(2): 997-1004, 2015.

Harris, I., Jones, P.D., Osborn, T.J., and Lister, D.H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, International Journal of Climatology., 34(3): 623-642, 2014.

Hartkamp, A.D., De Beurs, K., Stein, A., and White, J.W.: Interpolation Techniques for Climate Variables Interpolation., 1999.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, International Journal of Climatology., 25(15): 1965-1978, 2005.

Huffman, G.J., Adler, R.F., Bolvin, D.T., and Nelkin, E.J.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. j hydrometeor., 2010.

Hutchinson, M.F.: Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 9(4): 385-403, 1995.

Hutchinson, M.F, Xu T.: *ANUSPLIN Version 4.4 User Guide*. Australian National University: Canberra., 2013.

Immerzeel, W.W., Rutten, M.M., and Droogers, P.: Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula, Remote Sensing of Environment., 113(2): 362-370, 2009.

Jin, M., and Dickinson, R.E.: Land surface skin temperature climatology: benefitting from the strengths of satellite observations, Environmental Research Letters., 5(4): 44004, 2010.

Karger, D.N., et al.: Climatologies at high resolution for the earth's land surface areas, Scientific Data., 4(1),170122, 2017.

Kilibarda, M., et al.: Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution, Journal of Geophysical Research: Atmospheres., 119(5), 2294-2313, 2014.

Kolios, S., and Kalimeris, A.: Evaluation of the TRMM rainfall product accuracy over the central Mediterranean during a 20-year period (1998–2017), Theoretical and Applied Climatology., 139(1): 785-799, 2020.

Lawrimore, J.H., et al.: An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3, Journal of Geophysical Research Atmospheres., 116(D19), 2011.

Li, M., and Shao, Q.: An improved statistical approach to merge satellite rainfall estimates and raingauge data, Journal of Hydrology., 385(1-4): 51-64, 2010.

Liu, Q., et al.: The hydrological effects of varying vegetation characteristics in a temperate water-limited basin: Development of the dynamic Budyko-Choudhury-Porporato (dBCP) model, Journal of Hydrology., 595-611, 2016.

Marchi, M., Sinjur, I., Bozzano, M., and Westergren, M.: Evaluating WorldClim Version 1 (1961-1990) as the Baseline for Sustainable Use of Forest and Environmental Resources in a Changing Climate, Sustainability., 11(11): 14, 2019.

Michaelides, S., et al.: Precipitation: Measurement, remote sensing, climatology and modeling, Atmospheric Research., 94(4): 512-533, 2009.

Mildrexler, D.J., Zhao, M., and Running, S.W.: A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. 116(G3) , 2011.

Mosier, T.M., Hill, D.F., and Sharp, K.V.: 30-Arcsecond monthly climate surfaces with global land coverage, International Journal of Climatology., 34(7) , 2014.

Muller, R.A., Rohde, R., Jacobsen, R., Muller, E., and Wickham, C.: A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011, 2013.

New, M., Hulme, M., and Jones, P.: Representing Twentieth-Century Space–Time Climate Variability. Part I: Development of a 1961–90 Mean Monthly Terrestrial Climatology, Journal of Climate., 12(3): 829-856, 1999.

New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, Climate Research., 21(1): 1-25, 2002.

Parmentier, B., et al.: An Assessment of Methods and Remote-Sensing Derived Covariates for Regional Predictions of 1 km Daily Maximum Air Temperature, Remote Sensing., 6(9): 8639-8670, 2014.

Peng, S., Ding, Y., Liu, W., and Li, Z.: 1 km monthly temperature and precipitation dataset for China from 1901 to 2017, Earth Syst. Sci. Data., 11(4): 1931-1946, 2019.

https://doi.org/10.5194/essd-2020-361
Preprint. Discussion started: 8 January 2021

Peng, S., et al.: Spatiotemporal change and trend analysis of potential evapotranspiration over the Loess Plateau of China during 2011–2100, Agricultural and Forest Meteorology., 233: 183-194, 2017.

Pfister, L., et al.: Statistical reconstruction of daily precipitation and temperature fields in Switzerland back to 1864, Clim. Past., 16(2): 663-678, 2020.

Platts, P.J., Omeny, P.A., and Marchant, R.: AFRICLIM: high - resolution climate projections for ecological applications in Africa, African Journal of Ecology., 53(1), 103-108, 2015.

Ray, D., et al.: Comparing the provision of ecosystem services in plantation forests under alternative climate change adaptation management options in Wales, Reg. Envir. Chang., 15(8): 1501-1513, 2015.

Simpson, J., Kummerow, C., Tao, W.K., and Adler, R.F.: The Tropical Rainfall Measuring Mission (TRMM) Sensor Package, Meteorology & Atmospheric Physics., 60(1-3): 19-36, 1996.

Siuki, S.K., Saghafian, B., and Moazami, S.: Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products, International Journal of Remote Sensing., 38(1-2): 558-571, 2017.

Fick, S.E., and Hijmans, R.J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, International Journal of Climatology., 37(12), 4302-15, 2017.

Sterl, A., Komen, G.J., and Cotton, P.D.: Fifteen years of global wave hindcasts using winds from the European Centre for Medium-Range Weather Forecasts reanalysis: Validating the reanalyzed winds and assessing the wave climate, Journal of Geophysical Research Oceans., 103, 5477-5492, 1998.

Thornton, P.E., Running, S.W., and White, M.A.: Generating surfaces of daily meteorological variables over large regions of complex terrain, Journal of Hydrology., 190(3-4): 214-251, 1997.

Willmott, C.J., and Robeson, S.M.: Climatologically aided interpolation (CAI) of terrestrial air temperature, International Journal of Climatology., 15(2), 2010.

Wu, T., and Li, Y.: Spatial interpolation of temperature in the United States using residual kriging, Applied Geography., 44: 112-120, 2013.

Yao, R., Wang, L., Huang, X., Li, L., and Jiang, W.: Developing a temporally accurate air temperature dataset for Mainland China, Science of The Total Environment., 706, 136037, 2020.