ARTICLE IN PRESS

Scientific Data

https://doi.org/10.1038/s41597-026-06959-0

Article in Press

Gridded millennial summer temperature dataset over the Yangtze River Basin

Received: 3 October 2025

Accepted: 23 February 2026

Cite this article as: Dilawar, A., Duan, J., Liu, Y. *et al.* Gridded millennial summer temperature dataset over the Yangtze River Basin. *Sci Data* (2026). https://doi.org/10.1038/s41597-026-06959-0

Adil Dilawar, Jianping Duan, Yawen Liu, Iman Babaeian & Seyed Asaad Hosseini

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Gridded millennial summer temperature dataset over the Yangtze River Basin

Adil Dilawar[1], Jianping Duan[1*], Yawen Liu[1], Iman Babaeian[2], Seyed Asaad Hosseini[3]

[1]State Key Laboratory of Earth Surface Processes and Disaster Risk Reduction, Faculty of Geographical Science, Beijing Normal University, Beijing, China.
[2]Climate Research Institute, Atmospheric Science and Meteorological Research Center, Mashhad, Iran.
[3]Department of Physical Geography, University of Mohaghegh Ardabili, Ardabil, Iran.

Corresponding author (*): Jianping Duan (duanjp@bnu.edu.cn).

Abstract

High-resolution near-surface air temperature (SAT) datasets are essential for evaluating long-term SAT changes and developing management strategies in the Yangtze River Basin (YRB). However, existing SAT datasets have limited temporal coverage or coarser spatial resolution. In this study, a summer SAT dataset for 850–2005 at a spatial resolution of 1°×1° was developed by integrating the Millennium Global Climate Model (GCM) simulations from the Coupled Model Intercomparison Project Phases 5 and 6 (CMIP5 and CMIP6) with four proxy datasets. A weighted ensemble of multi- GCM simulations was created, and bias-correction was performed using an updated cumulative distribution function (CDF) method to improve gridded summer SAT dataset. The bias-corrected dataset was subsequently integrated with four gridded paleo summer SAT datasets covering the past millennium using a grid-weighted averaging technique based on an entropy method to improve low-frequency signal robustness. Spatiotemporal evaluation metrics showed substantial improvements in corrected summer SAT compared with raw GCMs. The integrated dataset demonstrates robust performance in revealing the spatiotemporal changes in SAT in the YRB during the past millennium.

Background & Summary

The Yangtze River is one of the longest freshwater channels in Eurasia, originating from the

Tibetan Plateau, China, and extending into the East China Sea. The Yangtze River Basin (YRB) is home to approximately 400 million people and produces approximately 50% of China's food, primarily rice[1]. It is experiencing amplified and consistent warming, which has exacerbated the frequency and intensity of heatwaves in the region[2-5]. This heightens the likelihood of extreme climate events and poses significant threats to people's livelihoods. Therefore, it is paramount to investigate in depth the potential consequences of climate change across critical sectors. These endeavors tend to lead to robust adaptation and mitigation policies to combat the detrimental effects of climate change. In this context, the SAT is among the key variables in a broad research field, including climate change, the hydrological cycle, ecology, and agricultural production[6,7]. A high-resolution gridded SAT dataset is needed to explore the long-term spatiotemporal characteristics of the increased frequency and intensity of warm extremes across the YRB.

For SAT data collection, ground stations and gridded datasets are the major sources[8,9]. Previous studies have used ground station data and have shown an accelerating trend of mean annual warming of 0.46 °C/10 yr from 1979–2018[10] and 0.17 °C/10 yr from 1960–2009[11] across the upper reach of the YRB. Similarly, Cui et al. (2017) reported the increasing trend of mean annual SAT at a rate of 0.19 °C/10 yr from 1960–2015 across the YRB[12]. Additionally, other studies used station data and confirmed that the mean annual SAT in the YRB has substantially increased in recent decades[13-15]. Some studies have employed gridded temperature datasets to examine the spatiotemporal characteristics and recurrence of extreme temperature events in the YRB[16,17,18,4]. For instance, Wu and Gao (2013) developed a high-resolution gridded temperature dataset (CN05.1) based on > 2400 ground stations across China using the angular distance weighted technique for the period of 1961–2024, which provides relatively high spatial accuracy[19]. These station SAT records and gridded SAT datasets are widely recognized for their accuracy and reliability[20-23] but their limited temporal span constrains the ability to capture long-term temperature changes over past centuries, necessitating investigations of the potential implications of SAT changes (e.g., changes compared with the preindustrial era) in the YRB. Therefore, given the spatial and temporal limitations of available stations and gridded data across the YRB, the demand for high-resolution SAT datasets remains understudied.

In this context, proxy data can enhance the understanding of the past climate in the region, particularly before instrumental period datasets[24], and aid in future climate projection. Several studies have reconstructed long-term SAT changes using multiproxy data approaches, such as tree rings and historical documents, including the YRB or parts of it. Overall, past climate variations have been studied mainly in the upper YRB, which covers the eastern part of the Tibetan Plateau[25-27]. These reconstructed datasets involve unevenly distributed proxy sites because of the complex elevation of the YRB and cover only a limited spatial and temporal scale; they do not reflect accurate spatial variability, specifically in the lower reach of the YRB. In addition, a few gridded long-term high-resolution SAT datasets have been developed that focus on China and East Asia, including the YRB[28-31]. However, these paleo-reconstruction datasets have coarse spatial resolution (2° × 2° and 5° × 5°). These datasets have used fewer proxy sites because of their unavailability in the lower and middle reaches of the YRB[32]; thus, reflecting the accurate SAT variability across the YRB with complex terrain is challenging.

GCMs are available datasets for investigating climate change[33-39]. GCMs provide continuous spatiotemporal data to examine past climate variability under modeling experiments for the last millennium (850–1850 CE), incorporated within the Paleoclimate Modeling Intercomparison Project Phases 3 and 4 (PMIP3 and PMIP4)[40]. Selected models from the CMIP5[41] and CMIP6[42] are run under these experiments for the past 1000 simulation data[43-45]. However, large uncertainties in the data derived from GCMs are usually reported[46], especially at the regional scale[47]. Therefore, bias- correction of the raw GCM outputs is needed to produce reliable simulations for understanding paleoclimate variability and strengthening mitigation and adaptation policies in the future.

To address these challenges and limitations, this study combines GCM simulations and four available relatively low spatial resolution gridded paleo reanalysis and reconstruction datasets produced for China and East Asia to produce a new and high spatiotemporal resolution summer SAT dataset covering the last millennium across the YRB. First, we create a weighted multiensemble of nine GCMs (seven from CMIP5[48-54] and two from CMIP6[55,56]) instead of a traditional multi-mean ensemble. We subsequently employed an advanced gridded bias-correction using a CDF-based procedure under the normal distribution and utilized the CRU temperature data as reference data. Finally, bias-corrected GCM simulations from the current study and four proxy gridded datasets were integrated through an advanced grid-weighted average technique to improve the decadal and multidecadal signals of past SAT data. This newly developed dataset enabled us to assess past summer temperature changes in the YRB and better constrain the implications of future climate change.

Methods

Study area

The Yangtze River starts from the Tanggula Mountain glacier located in the Qinghai-Tibetan Plateau at an elevation of > 5000 a.s.l.m, while the elevation in the middle and lower reaches of the YRB ranges from 1 a.s.l.m to 2267 a.s.l.m. The complex terrain of the YRB spans an elevation range of 1–7217 m and decreases from west to east. The geographical location and topography of the study area are illustrated in Fig. 1.

Fig. 1 Location of study domain (a), physical characteristics of the YRB (b)[57], and topography with spatial characteristics (c). Topographic and urbanization information is based on the Shuttle Radar Topography Mission (SRTM) and the MODIS satellite.

Data sources

Observation data

The CRU TS monthly SAT dataset was developed by the University of East Anglia, England, at an original spatial resolution of 0.5° × 0.5° from 1901 to 2023; and is accessible at https://data.ceda.ac.uk/badc/cru/data/cru_ts/ [58]. In the current study, the CRU TS monthly SAT dataset was interpolated at a spatial resolution of 1°×1° and split into two periods: calibration

(1951–2005) and validation (1901–1950). In addition, the longest available period (1901–2005) was used to calculate the gridded weights and validate the final generated SAT dataset.

Global climate model simulations

Monthly SAT data were collected from nine GCMs (7 GCMs from CMIP5 and 2 GCMs from CMIP6) covering the period of 850–2005, accessed from https://esgf-node.llnl.gov/search/cmip6/ . Moreover, monthly SAT data were collected for the last millennium period (850–1849, past 1000 simulations), and the historical period (1850–2005, historical simulations) (Table 1). Further details regarding the CMIP5 and CMIP6 experiments are described by Huo et al. (2021)[59]. The CMIP5 simulations were terminated in 2005, and the other paleo reanalysis and reconstruction datasets also end early. Therefore, the historical period was extended to 2023 using observations (CRU TS). Finally, the summer SAT dataset was prepared for the last millennium and covers the period of 850–2023 over the YRB.

Table 1. Detailed description of the GCMs used in this study. The italic font indicates the GCMs from CMIP6, while the normal font represents CMIP5.

Millennium paleoclimate datasets

Four summer SAT paleo reanalysis and reconstruction datasets covering the past millennium were collected for this study (Table 2). The first paleo reanalysis dataset was the Modern Era Reanalysis (ModE), a global monthly dataset covering 1421–2008 with a spatial resolution of 1.8° × 1.8° [60]. The second paleo reanalysis dataset was the Paleo Hydrodynamic Data Assimilation (PHYDA) product, a global hydroclimate reconstruction for the period of 0001–2023 at 2.5° × 1.88° resolution, using a data assimilation approach that integrates a time series of 2,978 proxy-data sites along constraints of climate model[61]. Third, the summer temperature proxy dataset was reconstructed in eastern and south-central Asia covering 900–1999 at a 5° × 5° spatial resolution[30]. It employs a point-to-point regression method using multiproxy data in Asia and is accessible at https://www.ncei.noaa.gov/access/paleo-search/study/18635 .

Fourth, a summer temperature reconstruction was developed for East Asia covering the 800–2005 period at a 2° × 2° resolution using a network of annual tree-ring chronologies[28]. It uses an ensemble point-to-point regression technique to reconstruct summer temperature of each grid point in East Asia, and is accessible at https://www.ncei.noaa.gov/access/paleo-search/study/19523 . However, observational, GCM, and paleo reanalysis and reconstruction datasets have varying spatial resolutions. Therefore, all the datasets were interpolated using the bilinear technique to a uniform spatial resolution of 1º × 1º to ensure fair comparison and analysis. The GCM simulations were bias corrected using the observations (CRU TS) at a spatial resolution of 1º × 1º. Finally, corrected GCM simulations and four paleo reanalysis and reconstruction datasets were integrated to produce the final summer SAT dataset with a spatial resolution of 1º × 1º for the last millennium (850–2005) over the YRB. The detailed procedure for data generation is illustrated in Fig. 2.

Table 2. Detailed descriptions of the summer SAT paleoclimate datasets used in this study.

Grid-based weight ensembles for the multiple models

Correction_1 employed the grid-weighted average technique to create an ensemble of nine GCMs using monthly SAT data from GCMs and observations (CRU TS) during 1901–2005 at a 1° × 1° resolution. It evaluates the positive and negative relationships between observations (CRU TS) and GCM simulations using the normalized root-mean-square-errors (RMSEs), correlation coefficient (CC), mean bias errors (MBEs), and mean absolute errors (MAEs) at each grid. However, different methods are used to calculate the weights of these indicators[62], and compared with an equal weight strategy, the current study employed an advanced weight calculation method based on the entropy technique to calculate the gridded weights and improve the results[63]. The detailed methodology for the calculation of weight was described by Dilawar et al. (2025)[64]. Finally, the gridded weights of each GCM were rated to create a rational weighted ensemble of nine GCMs to improve the GCM simulations compared with the simple mean of the raw GCMs (correction_1) at a spatial resolution of 1° × 1° for the period of 850–2005.

Spatial bias adjustment procedure

Subsequently, correction_2 (bias-correction) was employed to further improve the ensemble GCM simulations using an updated quantile mapping (QM) technique, which is computationally efficient. Many studies have used it to eliminate the biases from the outputs of GCMs [65-67]. In this study, an updated CDF matching-based correction (CDFM), which considers nonlinear changes rather than maintaining equal distance between the observed and GCM quantile distributions, was adopted. Moreover, it provides more accurate estimates of the distribution of extremes[68,69]. It uses a normal distribution because SAT values vary between negative and positive values.

The following equation was used to perform bias-correction[70].

$$\tilde{t}_{ms.adjust} = t_{ms} + F_{oh}^{-1}\left(F_{ms}(t_{ms})\right) - F_{mh}^{-1}\left(F_{ms}(t_{ms})\right) \ldots\ldots\ldots\ldots\ldots(1).$$

where, $\tilde{t}_{ms.\mathrm{adjust}}$ is the corrected SAT data from the GCMs, and F is the CDF of either the observation (*o*), or the GCM (*m*) during the historical (*h*) or forecast period (*s*).

Fig. 2 Flow diagram of the dataset generation.

The full observation period (1901–2005) was split into two subperiods: 1951–2005 (calibration period) and 1901–1950 (validation period). Furthermore, to evaluate the performance of the bias-correction procedure, we chose an independent validation period of 1901–1950. This maximizes the overlap with the available observations and allows for comprehensive evaluation. Finally, the full common period (1901–2005) of observations (CRU TS) was used for the bias-correction of the GCM simulations to obtain a bias-corrected SAT dataset for the last millennium (850–2005)[71].

Integration of multi-source millennial gridded datasets

While correction_2 employs the CDF correction method, which significantly improves accuracy and inter-annual variability, it is weak at capturing the decadal and multidecadal signals of SAT during the last millennium. Further improvements are needed to better capture these decadal and multidecadal signals. To improve this weakness, correction_3 was employed, which integrates four paleo reanalysis and reconstruction gridded datasets, including the PHYDA summer SAT, ModE summer SAT, Shi summer SAT, and Cook summer SAT, with the corrected GCM dataset from correction_2 (Table 2). These datasets have different temporal spans and spatial resolutions. To preserve data originality and integrity, while maintaining consistency, all the paleo reanalysis and reconstruction datasets were conservatively interpolated at a spatial resolution of 1° × 1°. The gridded anomaly was subsequently calculated with respect to the period of 1961–1990 and spanning 850–2005. Correction_3 integrates these five datasets using a grid-weighted average technique based on the entropy method to enhance the data quality by optimizing the performance at each grid in the YRB. To facilitate the data integration procedure of multiple datasets (Fig. S1), the entire period (850–2005) was divided into four subperiods— 850–1099, 1100–1420, 1421–1999, and 2000–2005— based on the available period of each dataset. Some paleo reanalysis and reconstruction datasets had missing values in the initial years; these years were excluded to maintain originality during the data integration procedure. Gridded weights were calculated on the basis of the longest available period (1901–2005) of observations (CRU TS). Detailed information on integrating paleo reanalysis and reconstruction datasets with corrected GCM data for calculating the gridded weights is presented in Fig. S1.

Evaluation procedure for the correction

Two-sample Kolmogorov–Smirnov (KS) test

The KS test is a widely used nonparametric method for assessing the equivalence of fundamental population distributions of two samples[72]. Assume that the population cumulative distribution functions (CDFs) of the two samples are $F_X(t)$ and $F_Y\,(t)$. Moreover, the first sample size (n_X) for X = (X_1, X_2,…………, X_{n_X}) and the second sample size (n_Y) for Y = $(Y_1,\ Y_2,\ \dots,\ Y_{n_Y})$ are randomly selected. Consider the alternative (H1: $F_X(\mathrm{t}) \neq F_Y(\mathrm{t})$) and null ($H_0 : F_X(\mathrm{t}) = F_Y(\mathrm{t})$) hypotheses for at least one t and for every t. $\hat{\mathrm{F}}_X(\mathrm{t})$ and $\hat{\mathrm{F}}_Y(\mathrm{t})$ are the empirical cdfs using the X and Y samples, respectively. The bootstrap approach employs a consistent resampling strategy by replacing ‘m’ times the dataset to create m samples[73]. As a result, some samples appear multiple times in the sample data. Specifically, for the validation period 10,000 time series are synthesized, and quantiles are calculated in each synthesis. Thus, the median value of the bootstrapped data is calculated for each quantile. The “D” statistics of the KS test reveal the reliability of the bias-corrected SAT for a significance level of 0.05. A smaller value of “D” indicates that the corrected SAT is closer to the observation[71].

Statistical error indicators

The efficiency of the corrected summer SAT from correction_1, correction_2, and correction_3 was evaluated using multiple error metrics from the temporal and spatial perspectives**.** The current study validated the results from correction_1 and correction_2 during the validation period (1901–1950) on the basis of observations (CRU TS) at 1° × 1° resolution over the YRB. However, correction_3 integrates different paleo reanalysis and reconstruction datasets and corrected GCMs to improve the multidecadal signal for the period of 850−2005 and was validated during the longest available observational period of 1901−2005. The following metrics were employed to gauge the performance of the correction procedure:

$$\ldots\ldots\ldots.\,\mathrm{RMSE} = \sqrt{\frac{\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{O_i} - \mathrm{M_i})^2}{\mathrm{n}}}\ \ldots\ldots\ldots\ldots..\ldots..\ldots\ldots.(2),$$

$$\ldots\ldots\ldots.\,\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}|\mathrm{O_i} - \mathrm{M_i}|\ \ldots.\ldots.\ldots\ldots\ldots..\ldots\ldots\ldots.(3),$$

$$\ldots\ldots\ldots.\,\mathrm{MBE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{O_i} - \mathrm{M_i})\ \ldots\ldots\ldots\ldots\ldots\ldots.\ldots..\ldots.(4),$$

$$.\ldots\ldots\ldots.\,\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(O_i - M_i)^2\ \ldots\ldots\ldots\ldots..\ldots\ldots\ldots..\ldots.(5),$$

$$\ldots\ldots\ldots.\,\mathrm{Bias}\ (\%) = \frac{\left|\frac{1}{n}\sum_{i=1}^{n} O_i - \frac{1}{n}\sum_{i=1}^{n} M_i\ \right|}{\frac{1}{n}\sum_{i=1}^{n}|O_i|} \times 100.\ \ldots\ldots\ldots\ldots(6),\ \text{and}$$

$$\ldots\ldots\ldots\mathrm{CC} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{M_i}-\bar{\mathrm{M}})(\mathrm{O_i}-\bar{\mathrm{O}})}{\sqrt{\sum_{i=1}^{n}(M_i-\bar{\mathrm{M}})^2}\sqrt{\sum_{i=1}^{n}(O_i-\bar{\mathrm{O}})^2}}\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots(7).$$

where, O_i is the observational SAT, the M_i is the corrected SAT, and ‘n’ is the number of samples. In addition, $\bar{O}$ and $\bar{M}$ represent the mean value of the observational and corrected SAT. The lower values of error metrics indicate better performance except for CC, while the higher values indicate poorer performance except for CC. These error metrics are annually calculated.

Relative added value evaluation

The spatial performance of the correction_3 SAT data was further assessed using relative added value (RAV). It measures the enhanced performance of the correction_3 SAT data compared with the correction_2 SAT data, which can be described as follows.

$$\ldots\ldots\ldots\ldots\mathrm{RAV}=\frac{\sum_{i=j} Metric_{cor\_2_i}-\sum_{i=j} Metric_{cor\_3_i}}{\sum_{i=j} Metric_{cor\_2_i}}\ldots\ldots\ldots\ldots\ldots(8).$$

Here, “*Metric*” refers to one of the five error metrics, such as the RMSE, MAE, MBE, MSE, and bias (%), where j=1, 2, 3, 4, and 5. The subscript “cor_2” denotes the corrected SAT data from correction_2, whereas “cor_3” refers to the integrated paleo reanalysis and reconstruction SAT data from correction_3. Compared with the correction_2 SAT data, the RAV data quantified the degree of improvement in the correction_3 SAT data using observation data. The performance of the final SAT was evaluated in the past millennium using the most reliable independent global paleo reanalysis SAT dataset (observation) at a 1° × 1° resolution over the YRB. Moreover, a positive RAV indicates improvement in the data, whereas a negative RAV denotes diminished performance.

Data Records

The dataset is available at the public data repository Zenodo and can be accessed at https://doi.org/10.5281/zenodo.18724777 [78]. The current dataset is generated for the summer SAT during the period of 850–2023 at a spatial resolution of 1° × 1° across the YRB. It is an integration of bias-corrected ensemble GCM simulations with four paleo reanalysis and reconstruction datasets for the summer SAT. The dataset is stored in netcdf format (.nc), and contains 1174 timesteps with a longitudinal range of 91.25 °E–119.25 °E and latitudinal range of 25.25 °N–34.25 °N.

Technical Validation

Temporal evaluation of the performance of corrected GCM simulations

For correction_1, simulations from nine GCMs, including seven from CMIP5 and two from CMIP6, were used in the current study to construct an ensemble using the grid-weighted averaging technique for summer SAT in the past millennium (850–2005) at a spatial resolution of 1° × 1°

over the YRB. The summer SAT from the correction_1 was subsequently bias-corrected by employing the correction_2 procedure. Thus, the performance of the correction_1 SAT and correction_2 SAT datasets was comprehensively evaluated using observations (CRU TS) during the validation period (1901–1950) over the YRB. This evaluation was performed across the grids for each year to compare the time series of the raw GCM simulations, correction_1, and correction_2 to assess the improvement in the SAT data through each correction step. In this context, different error metrics such as the RMSE, MAE, MBE, MSE, bias (%), and CC were used to evaluate the performance of the raw GCM simulations, for both corrected datasets across the entire YRB (Fig. 3). According to these error metrics, the simulation values that are closer to the observational values demonstrate better performance, with relatively high correlation and low error magnitude for the error metrics (RMSE, MBE, MSE, and bias (%)).

The results show that (1) compared with the raw GCMs and correction_1, the correction_2 had lower RMSE, MAE, MBE, MSE, and bias (%) and higher CC (Fig. 3); (2) the RMSEs values of raw GCMs data were higher than both of corrected data, likewise for others error indicators (Fig. 3a); (3) the average RMSEs, MAEs, MBEs, MSEs, and bias (%) for correction_1 were 1.81 °C, 1.43 °C, 1.15 °C, 3.33 °C, 5.4% respectively, while for correction_2 these values were 0.62 °C, 0.54 °C, 0.41 °C, 0.45 °C, 2.5%, respectively; and (4) the average CC during validation period (1901–1950) was relatively higher for correction_2 (0.99) compared to correction_1(0.98).

Fig. 3 Temporal evaluation of summer SAT (°C) correction procedures using multiple error metrics based on observations (CRU TS) during the validation period (1901–1950) over the YRB.

The annual RMSE of correction_2 SAT was lower, highlighting the superiority of the corrected data. Each error metric followed the same pattern for each year in the time series (Fig. 3a). On the other hand, the CC is higher for each year during 1901–1950 (Fig. 3f). Overall, compared with the raw GCMs, both the corrections improved the SAT data. The raw GCMs exhibited larger errors, which were reduced by applying stepwise corrections such as correction_1 and correction_2. These five-error metrics indicated that correction_2 has higher accuracy and the lowest error, reflecting the robustness of the correction method.

Furthermore, compared with the correction_1 data, correction_2 improved SAT data by reducing the biases and uncertainties. However, for the raw GCMs, the probability density function (PDF) showed a wide range for RMSEs (5.4 °C–7.3 °C) and a narrow range for CC (0.965–0.975), whereas it substantially decreased to RMSEs of 1.45 °C–2.25 °C during correction_1 and RMSEs of 0.35 °C–1.4 °C in correction_2. Conversely, the CC increased range to 0.978–0.989 in correction_1 and to 0.99–1.0 in correction_2 during 1901–1950 over the YRB (Fig. S2).

Similarly, the MAE, MBE, MSE, and bias (%) metrices followed the same order with ranges for correction_1 of 1.0 °C–1.85 °C, 0.45 °C–1.95 °C, 2.20 °C–5.20 °C, and 3.5% –9.5%, respectively, and those for correction_2 of 0.30 °C–1.4 °C, –0.25 °C–1.4 °C, 0.15 °C–2.0 °C, and 1.0%–6.0%, respectively (Figs. S2a–S2e). This signifies that after the correction_2, compared with the raw GCMs and correction_1, the summer SAT data were promisingly improved (Fig. S2). Furthermore, the time series of variations in regional mean SAT also indicated that compared with raw GCMs, corrections 1 and 2 substantially reduced the uncertainty with RMSE values to 1.18 °C and 0.49 °C, respectively (Fig. S3).

Evaluation of spatial error distribution over the YRB

The correction procedures were further evaluated in terms of the spatial pattern of multiple error metrics across the YRB for the same validation period. The error metrics were computed across each grid to compare the time series of each correction procedure during the validation period (1901–1950). It was demonstrated that compared with the raw GCMs, the corrections 1 and 2 significantly improved the SAT simulations by reducing the uncertainty across the YRB (Fig. 4). The mean RMSE for the raw GCMs was 6.06 °C, which was considerably reduced to 1.50 °C (0.66 °C) after correction_1(correction_2) (Figs. 4a–4c). The bias (raw GCMs) was greater in the upper reach of the basin, compared with correction_1 and correction_2, which remarkably improved the data in the western part of the YRB by reducing the error values. Similarly, the spatial distribution of the remaining error metrics showed evidence of improved accuracy of the SAT by reducing the MAEs, MBEs, MSEs, and bias (%) over the validation period. These findings demonstrate that in each correction procedure, compared with raw GCMs, the SAT data is greatly improved. The difference in spatial error distribution between the raw GCMs and corrected SAT data from correction_1 and 2 is clear in Fig. 4. In general, correction_1 and correction_2 reduced the error throughout the study region with average values as follows: MAEs from 1.54 °C to 0.54 °C; MBEs from 1.15 °C to 0.41 °C; MSEs from 3.32 °C to 0.45 °C, and bias (%) from 8.34% to 3.07%, while those of the raw GCMs (without correction) were 6.06 °C, 6.01 °C, 39.73 °C, and 32.63%, respectively. Overall, spatial differentiation remarkably improved the SAT data in the upper reach of the YRB.

Fig. 4 Detailed evaluation of spatial distribution of error metrics for the correction procedures based on observations (CRU TS) during 1901–1950.

The data in Fig. S4 further confirm the superiority of correction_2 using cumulative distributions and D statistics of the KS test and relative decrements (%) of biases during the validation period (1901–1950). This finding indicates that the distribution of the SAT data from correction_2 was considerably closer to the observations than the correction_1 SAT data, despite some remaining differences between corrected and observed data. This improvement is further demonstrated by the D statistics of the KS test. The result of the two-sample KS test revealed that compared with the correction_1, the corrected data effectively reduced the D-values at a significance level of 0.05. These results show that correction_1 had a D-value = 1.0 (Fig. S4a), and after correction_2, it decreased to D-value = 0.52 at $p \leq 0.05$ during 1901–1950 (Fig. S4b). Overall, the SAT data greatly improved after the correction_2 procedure across the YRB.

In addition, we investigated the spatial pattern of the decrement of biases of correction_1 and correction_2 in terms of the relative decrements (%) of the RMSE and MAE (Figs. S4c–S4d). As shown in Fig. S4c, the relative decrease (%) in RMSE was greater, specifically in the high-elevation region, because the raw data had a greater bias. This indicates an improvement in the data in the upper reach, with minimum and maximum of 3.12% and 92.8%, respectively, in the RMSE during the validation period (1901–1950). In addition, in some of the grids, the data did not significantly improve in terms of reduced error. Similarly, the spatial distribution of the relative decrement (%) of MAE followed the same pattern as that of the RMSE (Fig. S4d). This shows that correction_2 has significantly reduced uncertainty and errors in the western part of the YRB, including the high-elevation area, with average decreases of 34.4% and 38.3%, respectively. Overall, correction_2 significantly improved the data by reducing the cooling biases across the basin.

Assessment of SAT with correction_3 compared with observational SAT

The temporal evaluation of the SAT after different corrections (2 & 3) was assessed using the different error metrics compared with the longest available observational data during 1901–2005. A comparison among the mean SAT corrections demonstrated the superiority of the correction_3 method in reducing the biases and uncertainty. The correction_3 outperformed correction_2, with higher correlations and lower average biases (Fig. 5a). It is revealed that compared with observed SAT time series, the correction_3 (CC = 0.68) followed a better temporal pattern than correction_2 did (CC = 0.24). In contrast, correction_3 had lower RMSE (0.29 °C) than correction_2 did (0.42 °C). This finding indicates that each of the SAT correction steps improved the SAT quality. Finally, correction_3 greatly improved the decadal and multidecadal signals.

Fig. 5 Temporal comparison of the regional mean summer SAT anomalies (°C) after correction_2 and correction_3 with the observations (CRU TS) during the longest common period 1901–2005 over the YRB. CC means correlation coefficient, and RMSE means root mean square error (a). Temporal variation in regional mean summer SAT anomalies (°C) obtained from the current study, along with the uncertainty (RMSE) range during the past millennium (850–2023) over the YRB. The shaded area shows the uncertainty calculated based on observations (CRUTS) SAT anomaly (°C) for the period of 1901–2005 over the YRB (b).

In addition, a comprehensive assessment was carried out to investigate the improvement in the SAT after correction_3 by comparing the SAT after correction_2 during the longest common period. The temporal variations in the RMSE, MAE, MBE, and MSE anomalies (°C) and bias (%) of the correction_2 SAT and correction_3 SAT are shown in Fig. S5. All the error metrics showed notable improvement in the correction_3 SAT values during each year over the longest common validation period (1901–2005). The RMSE and MAE are shown in Figs. S5a–S5b, respectively, with lower values approaching 0.3 °C, indicating that the accuracy of the correction_3 SAT was better than that of the correction_2 SAT. Fig. S5c, where the MBE values suggest reduction in systematic bias, and the values were closer to zero for the correction_3 SAT than for the correction_2 SAT. Moreover, Fig. S5d presents the MSE values, suggesting lower error variance for the correction_3 SAT during 1901–2005.

Moreover, the annual bias (%) and CC were derived from the correction_3 SAT with observational data during 1901–2005, which greatly improved compared with those obtained from the correction_2 SAT (Figs. S5e–S5f). Overall, the temporal patterns of all error metrics signify the improvement of the final SAT data developed from correction_3 (integrating multiple paleo reanalysis and reconstruction data). For instance, the average RMSE, MAE, MBE, and MSE anomalies (°C) of the correction_2 SAT were 0.51 °C, 0.43 °C, 0.008 °C, and 0.30 °C, respectively. For the correction_3 SAT, these values decreased to 0.36 °C, 0.30 °C, 0.005 °C, and 0.15 °C during 1901–2005 (Fig. 6). In addition, the bias followed the same order: lower for the correction_3 SAT and higher for correction_2 SAT. In contrast, correction_3 SAT had a stronger correlation, whereas correction_2 SAT had a weaker correlation with the observations. The reduction in all the error gauging indicators and an increase in the correlation favor the correction_3 SAT data, which were closer to the observed data than the correction_2 SAT was. These evaluations demonstrate the robustness and effectiveness of the summer SAT, ultimately developed from correction_3 over the past millennium for the YRB.

Fig. 6 Whisker bar plot for spatial improvement of the correction_3 summer SAT anomaly (°C) compared with correction_2 summer SAT anomaly (°C) over the YRB during 1901–2005. The middle dotted black lines in the box represent the median values, and the dots represent pixels. The boxes indicate the range of 25% – 75%.

In the current study, an additional advanced evaluation (RAVs) was adopted to confirm the superior performance of the correction_3 SAT compared with the correction_2 SAT (Fig. S6). Most of the areas of the YRB had positive RAV grids for all the error metrics (RMSE, MAE, MBE, MSE, and bias (%)). This is further evident from the subplots, where the RAVs for all the error metrics were predominantly within the range of 0–1. The spatial pattern of the RAVs for all the error metrics showed a high positive value across the western part (mainly the upper reach) of the YRB, indicating a significant improvement in the correction_3 SAT data compared with the eastern part (mainly the lower reach). The average RAVs for the RMSE, MAE, MBE, MSE, and bias (%) were 0.30, 0.31, 0.76, 0.49, and 0.69, respectively, indicating remarkable added value in the summer SAT dataset (correction_3) over the YRB. The higher RAVs in the western part indicate the better efficacy of correction_3 SAT data in the area because more raw data from proxies were used in the paleo reanalysis and reconstruction datasets.

Spatiotemporal comparison of our dataset with observations, paleo reanalysis datasets, and regional proxy reconstructions

The spatial pattern of the summer SAT dataset developed in this study was compared with observations and two independent global gridded paleo-reanalysis SAT datasets over the observation period 1901–2005. The two paleo-reanalysis SAT datasets were developed by Mann et al. (2009)[74] (accessed at https://www.science.org/doi/10.1126/science.1177303) and LMR data by Meng et al. (2025)[75] (accessed at https://zenodo.org/records/17268597). All used datasets were interpolated at a spatial resolution of 1° × 1° over the YRB to compare their spatial patterns (Fig. S7). The results indicate that the current dataset better captures the observed spatial distribution of mean summer SAT than the other two paleo-reanalysis datasets across the YRB during the observed period 1901–2005.

Further spatial pattern correlations and time series comparisons demonstrated that the summer gridded SAT dataset from this study was much more consistent with observations than the other two paleo reanalysis datasets over the YRB (Figs. 7a–7c). The gridded CC between our dataset and the observations during 1901–2005 was in the range of 0.4–0.8, whereas those of Mann et al. 2009 and LMR were 0–0.53 and –0.1–0.4, respectively. Moreover, the temporal variations in regional mean summer SAT from the current study were closer to the observations than the other two millennial datasets (Fig. 7d). It appeared that the performance of the SAT dataset developed by Mann et al. (2009) is better than that of the LMR paleo reanalysis over the YRB. Therefore, we considered the Mann et al. (2009) dataset to further evaluate the performance of the current summer SAT over the YRB during the past millennium (850–2005).

Fig. 7 Spatial pattern of gridded correlation coefficient (CC) between observed and reconstructed summer SAT anomaly (°C) developed in the current study (a), Mann et al. (2009) (b), and the LMR (c) over the YRB during 1901–2005. Temporal comparisons of the regional mean summer SAT anomalies (°C) among the current study, observations (CRU TS), and the other two paleo temperature datasets over the YRB during their common period 1901–2005 (d).

In this context, RAVs were computed to assess the performance of the current summer SAT dataset developed herein during the full period (850–2005) across the YRB (Fig. S8). The summer SAT dataset by Mann et al. (2009) was used as the comparable dataset to calculate the error metrics for both corrections (2&3) during the period of 850–2005. The spatial pattern of the RAVs for all the metrics indicates positive values (Figs. S8a–S8e); in particular, the highest added values were observed in the upper reach of the YRB. This finding demonstrates that the final SAT dataset (correction_3) obtained from the current study substantially improved the summer SAT performance compared to a comparable dataset across all the grids located in the YRB for the past millennium. The average RAVs for the RMSE, MAE, MBE, MSE, and bias (%) were 0.30, 0.32, 0.40, 0.50, and 0.33, respectively, indicating remarkable added value in the summer SAT dataset (correction_3) over the past millennium. The subplot information indicated that all the RAVs for all the metrics were on the positive axis, which confirms the reliability of the SAT data developed in the current study during the 850–2005.

The temporal variation of regional mean summer SAT during 850–2005 showed clear interannual and multidecadal variabilities, demonstrating distinct cooling and warming phases occurred over the YRB in the past millennium (Fig. 5b). The selected extreme cold and warm years were shown to demonstrate the spatial patterns of the temperature anomalies (with respect to 1961–1990) across the YRB. The cold years (1405 and 1817) exhibited negative anomalies ranging from –0.24 °C to –1.28 °C, whereas the warm years (1247 and 2005) showed positive anomalies ranging from approximately 0 °C to 0.5 °C (Fig. S9). In addition, a few regional-scale summer SAT reconstructions were available to assess the decadal and multidecadal signals of the correction_3 (final integrated) SAT dataset. For the upper reach, we chose an August–September SAT independent reconstruction based on the maximum latewood density (MXD) of tree rings with a high explained variance across the southeastern Tibetan Plateau, covering part of the upper YRB[76]. The comparison revealed consistent warming and cooling intervals. However, prominent cooling phases were observed at approximately the 1600s, 1640s, 1675–1680, early 1710s, 1815–1820, and 1830–1840, and at approximately1860s, early 1990s and 1965, as well as the warming periods were during the early 1600s, 1800s, late 1800s, 1935–1950s, late 1900s, and early 2000s (Figs. S10a–S10b). Notably, some discrepancies were observed between the datasets in the late 1500s and 1650s, which are likely attributed to regional and seasonal differences in the temperature reconstructions.

For the comparisons in the middle and lower reaches, we chose the available millennial-scale decadal-resolution annual temperature reconstruction developed via the multiple climate proxy data, which mainly included tree rings, ice cores, historical documents, sediments, etc., over China[77]. In the comparison, both datasets revealed the decadal warming intervals (e.g., 900s–925s, approximately 1200s–1260s, and after the 1920s). On the other hand, cooling intervals were observed during 1270–1340, and 1640–1720, as well as between 1810–1870 (Figs. S11a–S11b). The comparison showed that recent warming was more pronounced after the 1920s. However, in the early 1100s, the warming phase was identified in the YRB but was not indicated in the Chinese SAT reconstruction. In general, both datasets captured consistent decadal warming and cooling patterns during the period of 865–1995. Discrepancies are observed in the decadal comparison during the early periods between the two datasets. The reconstruction proxy temperatures indicate a warming phase approximately 975s–1050s, whereas the current study’s data fail to capture warming evidently. These differences are likely to be from variations in the spatial scope of the two datasets and the target seasons.

Data Availability

The dataset generated in the current study can be accessed at https://doi.org/10.5281/zenodo.18724777 [78].

Code availability

The software and code associated with the current study's bias correction are publicly available[79]. Additional codes related to the current study are shared on the public data repository Zenodo with the dataset [78].

Acknowledgements

This study is supported by the National Key Research and Development Program of China (2023YFC3206601).

Author information

Authors and Affiliations

State Key Laboratory of Earth Surface Processes and Disaster Risk Reduction, Faculty of Geographical Science, Beijing Normal University, Beijing, China.
Adil Dilawar, Jianping Duan, & Yawen Liu
Climate Research Institute, Atmospheric Science and Meteorological Research Center, Mashhad, Iran.
Iman Babaeian
Department of Physical Geography, University of Mohaghegh Ardabili, Ardabil, Iran. Seyed Asaad Hosseini

Authors' contributions

J.D. conceived and designed the study with input from A.D., I.B., and S.A.H. A.D. performed data analysis and produced figures with input from Y.L. J.D. supervised the study and structured the paper with input from A. D., I.B., and S.A.H. A.D. wrote the original draft with input from input from J.D. Y.L., I.B., and S.A.H. All co-authors contributed to the interpretation of data and review & editing of the manuscript.

Ethics declaration

Not applicable, because this article does not contain any studies with human or animal subjects.

Competing interests

The authors declare no competing interests.

References

1 Birkinshaw, S. J. *et al.* Climate change impacts on Yangtze River discharge at the Three Gorges Dam. *Hydrology and Earth System Sciences* **21**, 1911-1927 (2017).

2 Ma, S., Zhou, T., Stone, D. A., Angélil, O. & Shiogama, H. Attribution of the July–August 2013 heat event in Central and Eastern China to anthropogenic greenhouse gas emissions. *Environmental Research Letters* **12**, 054020 (2017).

3 Yan, M. *et al.* The exceptional heatwaves of 2017 and all-cause mortality: An assessment of nationwide health and economic impacts in China. *Science of the Total Environment* **812**, 152371 (2022).

4 Hua, W., Dai, A., Qin, M., Hu, Y. & Cui, Y. How unexpected was the 2022 summertime heat extremes in the middle reaches of the Yangtze River? *Geophysical Research Letters* **50**, e2023GL104269 (2023).

5 Liao, Z., An, N., Chen, Y. & Zhai, P. On the possibility of the 2022-like spatio-temporally compounding event across the Yangtze River Valley. *Environmental Research Letters* **19**, 014063 (2024).

6 Sippel, S., Meinshausen, N., Fischer, E. M., Székely, E. & Knutti, R. Climate change now detectable from any single day of weather at global scale. *Nature Climate Change* **10**, 35-41 (2020).

7 Chen, R. *et al.* Association between ambient temperature and mortality risk and burden: time series study in 272 main Chinese cities. *British Medical Journal* **363** (2018).

8 Dos Santos, R. S. Estimating spatio-temporal air temperature in London (UK) using machine learning and earth observation satellite data. *International Journal of Applied Earth Observation and Geoinformation* **88**, 102066 (2020).

9 Fu, P. & Weng, Q. Variability in annual temperature cycle in the urban areas of the United States as revealed by MODIS imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* **146**, 65-73 (2018).

10 Yan, Y., You, Q., Wu, F., Pepin, N. & Kang, S. Surface mean temperature from the observational stations and multiple reanalyses over the Tibetan Plateau. *Climate Dynamics* **55**, 2405-2419 (2020).

11 Wang, S., Jiao, S. & Xin, H. Spatio-temporal characteristics of temperature and precipitation in Sichuan Province, Southwestern China, 1960–2009. *Quaternary International* **286**, 103-115 (2013).

12 Cui, L. *et al.* Innovative trend analysis of annual and seasonal air temperature and rainfall in the Yangtze River Basin, China during 1960–2015. *Journal of Atmospheric and Solar-Terrestrial Physics* **164**, 48-59 (2017).

13 Sang, Y.-F. Spatial and temporal variability of daily temperature in the Yangtze River Delta, China. *Atmospheric Research* **112**, 12-24 (2012).

14 Sang, Y.-F., Wang, Z. & Liu, C. Spatial and temporal variability of daily temperature during 1961–2010 in the Yangtze River Basin, China. *Quaternary International* **304**, 33-42 (2013).

15 Zhang, S. *et al.* Changes in the mean and extreme temperature in the Yangtze River Basin over the past 120 years. *Weather and Climate Extremes* **40**, 100557 (2023).

16 Long, Y., Li, J., Zhu, Z. & Zhang, J. Predictability of the anomaly pattern of summer extreme high-temperature days over southern China. *Climate Dynamics* **59**, 1027-1041 (2022).

17 Liu, B., Zhu, C., Ma, S., Yan, Y. & Jiang, N. Subseasonal processes of triple extreme heatwaves over the Yangtze River Valley in 2022. *Weather and Climate Extremes* **40**, 100572 (2023).

18 Huang, H., Zhu, Z. & Li, J. Disentangling the unprecedented Yangtze River basin extreme high temperatures in summer 2022: Combined impacts of the reintensified La Niña and strong positive NAO. *Journal of Climate* **37**, 927-942 (2024).

19 Wu, J. & Gao, X.-J. A gridded daily observation dataset over China region and comparison with the other datasets. *Chinese Journal of Geophysics-Chinese Edition* **56**, 1102-1111, doi:10.6038/cjg20130406 (2013).

20 Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* **7**, 109 (2020).

21 C3S. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate copernicus climate change Service climate data store (CDS). (2017).

22 Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A. & Hegewisch, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data* **5**, 1-12 (2018).

23 Schneider, U. *et al.* GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology* **115**, 15-40 (2014).

24 Jones, P., Briffa, K., Barnett, T. & Tett, S. High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures. *The Holocene* **8**, 455-471 (1998).

25 Yang, B. *et al.* A 622-year regional temperature history of southeast Tibet derived from tree rings. *The Holocene* **20**, 181-190 (2010).

26 Duan, J., Wang, L., Li, L. & Chen, K. Temperature variability since AD 1837 inferred from tree-ring maximum density of Abies fabri on Gongga Mountain, China. *Chinese Science Bulletin* **55**, 3015-3022 (2010).

27 Gou, X. *et al.* Rapid tree growth with respect to the last 400 years in response to climate warming, northeastern Tibetan Plateau. *International Journal of Climatology: A Journal of the Royal Meteorological Society* **27**, 1497-1503 (2007).

28 Cook, E. R. *et al.* Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 CE. *Climate Dynamics* **41**, 2957-2972 (2013).

29 PAGES 2k Consortium. Continental-scale temperature variability during the past two millennia. *Nature Geoscience* **6**, 339–346 (2013).

30 Shi, F. *et al.* A multi-proxy reconstruction of spatial and temporal variations in Asian summer temperatures over the last millennium. *Climatic Change* **131**, 663-676 (2015).

31 Shi, F. *et al.* Northern Hemisphere temperature reconstruction during the last millennium using multiple annual proxies. *Climate Research* **56**, 231-244 (2013).

32 Yang, F. *et al.* Comparison of the dryness/wetness index in China with the Monsoon Asia Drought Atlas. *Theoretical and Applied Climatology* **114**, 553-566 (2013).

33 Meehl, Gerald A., Thomas F. Stocker, William D. Collins, Pierre Friedlingstein, Amadou T. Gaye, Jonathan M. Gregory, Akio Kitoh et al. "Global climate projections. Chapter 10." (2007).

34 Bronstert, A., Kolokotronis, V., Schwandt, D. & Straub, H. Comparison and evaluation of regional climate scenarios for hydrological impact analysis: General scheme and application example. *International Journal of Climatology: A Journal of the Royal Meteorological Society* **27**, 1579-1594 (2007).

35 Phipps, S. J. *et al.* Paleoclimate data–model comparison and the role of climate forcings over the past 1500 years. *Journal of Climate* **26**, 6915-6936 (2013).

36 Neukom, R. *et al.* Inter-hemispheric temperature variability over the past millennium. *Nature Cimate Change* **4**, 362-367 (2014).

37 Stocker, T.F. *et al.* The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. *Change, IPCC Climate* **10** (2013).

38 Solomon, S. et al. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change 2007. (2007).

39 Arias, P. *et al.* Climate Change 2021: the physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; technical summary. (2021).

40 Kageyama, M. et al. The PMIP4 contribution to CMIP6-Part 4: Scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum experiments and PMIP4 sensitivity experiments. Geoscientific Model Development 10, 4035-4055 (2017). https://doi.org/10.5194/gmd-10-4035-2017.

41 Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society* **93**, 485-498 (2012).

42 Eyring, V. *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9**, 1937-1958 (2016).

43 Goosse, H. *et al.* Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. *Journal of Geophysical Research: Atmospheres* **115** (2010).

44 Goosse, H., Renssen, H., Timmermann, A. & Bradley, R. S. Internal and forced climate variability during the last millennium: a model-data comparison using ensemble

simulations. *Quaternary Science Reviews* **24**, 1345-1360 (2005).

45 Schmidt, G. *et al.* Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1. 1). *Geoscientific Model Development* **5**, 185-191 (2012).

46 Maraun, D. Bias correcting climate change simulations-a critical review. *Current Climate Change Reports* **2**, 211-220 (2016).

47 Di Luca, A., Pitman, A. J. & de Elía, R. Decomposing temperature extremes errors in CMIP5 and CMIP6 models. *Geophysical Research Letters* **47**, e2020GL088031 (2020).

48 Wu, T. *et al.* An overview of BCC climate system model development and application for climate change studies. *Journal of Meteorological Research* **28**, 34-56 (2014).

49 Gent, P. R. *et al.* The community climate system model version 4. *Journal of Climate* **24**, 4973-4991 (2011).

50 Schmidt, G. A. *et al.* Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems* **6**, 141-184 (2014).

51 Dufresne, J.-L. *et al.* Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics* **40**, 2123-2165 (2013).

52 Watanabe, S. *et al.* MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development* **4**, 845-872 (2011).

53 Yukimoto, S. et al. Present-day climate and climate sensitivity in the Meteorological Research Institute coupled GCM version 2.3 (MRI-CGCM2.3). *Journal of the Meteorological Society of Japan* 84, 333-363 (2006). https://doi.org/10.2151/jmsj.84.333.

54 Giorgetta, M. A. *et al.* Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems* **5**, 572-597 (2013).

55 Yukimoto, S. *et al.* The Meteorological Research Institute Earth System Model version 2.0, MRI-ESM2. 0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan. Ser. II* **97**, 931-965 (2019).

56 Lim Kam Sian, K. T. C., Wang, J., Ayugi, B. O., Nooni, I. K. & Ongoma, V. Multi-decadal variability and future changes in precipitation over Southern Africa. *Atmosphere* **12**, 742 (2021).

57 Commission, M. R. Ministry of Water Resources of China. *Joint observation and evaluation of the emergency water supple***-** *ment from China to the Mekong River* (2016).

58 Harris, I., Osborn, T. J., Jones, P. & Lister, D. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data* **7**, doi:10.1038/s41597-020-0453-3 (2020).

59 Huo, R. *et al.* Extreme precipitation changes in Europe from the last millennium to the end of the twenty-first century. *Journal of Climate* **34**, 567-588 (2021).

60 Valler, V. *et al.* ModE-RA: a global monthly paleo-reanalysis of the modern era 1421 to 2008. *Scientific Data* **11**, 36 (2024).

61 Steiger, N. J., Smerdon, J. E., Cook, E. R. & Cook, B. I. A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Scientific Data* **5**, 1-15 (2018).

62 Srinivasa Raju, K., Kumar, D.N. (2018). Selection of Global Climate Models. In: Impact of Climate Change on Water Resources. Springer Climate. Springer, Singapore. https://doi.org/10.1007/978-981-10-6110-3_2

63 Raju, B. K. & Nandagiri, L. Analysis of historical trends in hydrometeorological variables in the upper Cauvery Basin, Karnataka, India. *Current Science*, 577-587 doi: https://doi.org/10.18520/cs/v112/i03/577-587 (2017).

64 Dilawar, A. *et al.* Approaching the 1.5° C warming threshold and escalating seasonal warm extremes in the Yangtze River Basin. *Journal of Hydrology: Regional Studies* **61**, 102744 (2025).

65 Cannon, A. J., Sobie, S. R. & Murdock, T. Q. Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *Journal of Climate* **28**, 6938-6959 (2015).

66 Thrasher, B., Maurer, E. P., McKellar, C. & Duffy, P. B. Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences* **16**, 3309-3314 (2012).

67 Shastri, H., Ghosh, S. & Karmakar, S. Improving Global Forecast System of extreme precipitation events with regional statistical model: Application of quantile-based probabilistic forecasts. *Journal of Geophysical Research: Atmospheres* **122**, 1617-1634 (2017).

68 Moraes, A. G. d. L. & Motlagh, S. K. The Climate Data for Adaptation and Vulnerability Assessments and the Spatial Interactions Downscaling Method. *Scientific Data* **11**, doi:10.1038/s41597-024-03995-6 (2024).

69 Gupta, R., Bhattarai, R. & Mishra, A. Development of Climate Data Bias Corrector (CDBC) Tool and Its Application over the Agro-Ecological Zones of India. *Water* **11**, doi:10.3390/w11051102 (2019).

70 Li, H., Sheffield, J. & Wood, E. F. Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research-Atmospheres* **115**, doi:10.1029/2009jd012882 (2010).

71 Huo, R. *et al.* Flood variability in the upper Yangtze River over the last millennium—Insights from a comparison of climate-hydrological model simulated and reconstruction. *Science China Earth Sciences* **66**, 547-567 (2023).

72 Hollander, M. & Wolfe, D. A. Nonparametric Statistical Methods." Wiley, New York. (1999).

73 Davison, A. C. & Hinkley, D. V. *Bootstrap methods and their application*. (Cambridge university press, 1997).

74 Mann, M. E. *et al.* Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly. *Science* **326**, 1256-1260, doi:10.1126/science.1177303 (2009).

75 Meng, Z., Hakim, G. J. & Steig, E. J. Coupled Seasonal Data Assimilation of Sea Ice, Ocean, and Atmospheric Dynamics over the Last Millennium. *Journal of Climate* **38**, 7229-7247, doi:10.1175/jcli-d-25-0048.1 (2025).

76 Duan, J., Ma, Z., Li, L. & Zheng, Z. August-September temperature variability on the Tibetan Plateau: Past, present, and future. *Journal of Geophysical Research: Atmospheres* **124**, 6057-6068 (2019).

77 Ge, Q., Hao, Z., Zheng, J. & Shao, X. Temperature changes over the past 2000 yr in China and comparison with the Northern Hemisphere. *Climate of the Past* **9**, 1153-1160 (2013).

78 Dilawar, A., Duan, J., Liu, Y., Babaeian, I., & Hosseini, S. A. (2026). Millennial summer temperature dataset over the Yangtze River Basin [Data set]. Zenodo. https://doi.org/10.5281/zenodo.18724777

79 Agro-Climatic Tools. (2019). AgroClimaticTools/CDBC: Climate Data Bias Corrector [First Release] (v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.2745243

Main Tables: Scientific Data

Gridded millennial summer temperature dataset over the Yangtze River Basin

Adil Dilawar[1], Jianping Duan[1*], Yawen Liu[1], Iman Babaeian[2], Seyed Asaad Hosseini[3]

[1]State Key Laboratory of Earth Surface Processes and Disaster Risk Reduction, Faculty of Geographical Science, Beijing Normal University, Beijing, China.
[2]Climate Research Institute, Atmospheric Science and Meteorological Research Center, Mashhad, Iran.
[3]Department of Physical Geography, University of Mohaghegh Ardabili, Ardabil, Iran.

Corresponding author (*): Jianping Duan (duanjp@bnu.edu.cn).

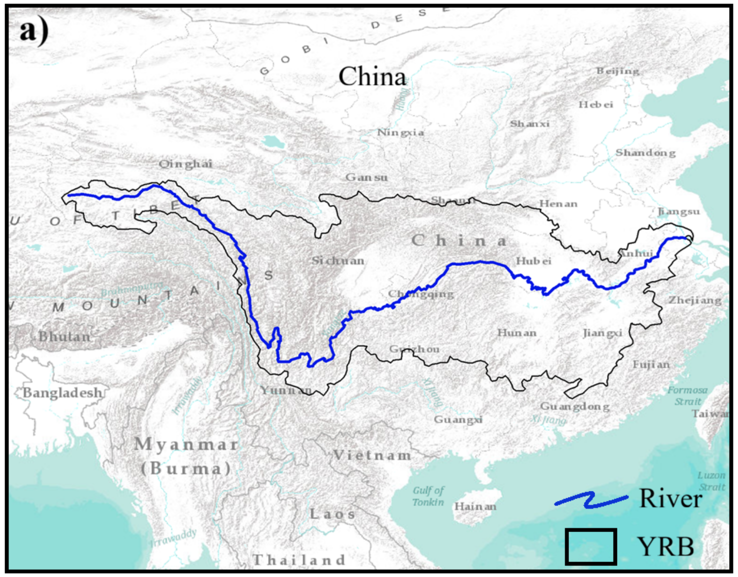
File description:
This file contains two tables (Table 1 and Table 2) are cited in the main text.

Table 1. Detailed description of the GCMs used in this study. The italic font indicates the GCMs from CMIP6, while the normal font represents CMIP5.

No	Model Name	Developer Name	Spatial Resolution	Time Period	Variable Name	Reference
1	**BCC_CSM 1**	Beijing Climate Center (BCC) and China Meteorological Administration (CMA), China	2.8° × 2.8°	850–2005	Tas	48
2	**CCSM4**	National Center for Atmospheric Research (NCAR), USA	0.9° × 1.25°	850–2005	Tas	49
3	***MRIESM2***	*Japan Meteorological Agency, Japan*	*1.13° × 1.13°*	*850–2005*	*Tas*	*55*
4	**GISS-E2**	NASA's Goddard Institute for Space Studies, USA	2.0° × 2.5°	850–2005	Tas	50
5	***MIROC_ES2L***	*Japan Agency for Marine–Earth Science and Technology, Atmosphere and Ocean Research Institute, The University of Tokyo, National Institute for Environmental Studies, and RIKEN Center for Computational Science/Japan*	*2.8° × 2.8°*	*850–2005*	*Tas*	*56*
6	**IPSL_CM5A_LR**	Institute Pierre-Simon Laplace in France	2.5° × 3.75°	850–2005	Tas	51
7	**MIROC_ESM**	Center for Climate System Research, Japan	2.8° × 2.8°	850–2005	Tas	53
8	**MRICGCM3**	Meteorological Research Institute, Japan	1.12° × 1.12°	850–2005	Tas	54
9	**MPI_ESM_P**	Max Planck Institute for Meteorology, Germany	1.84° × 1.84°	850–2005	Tas	55

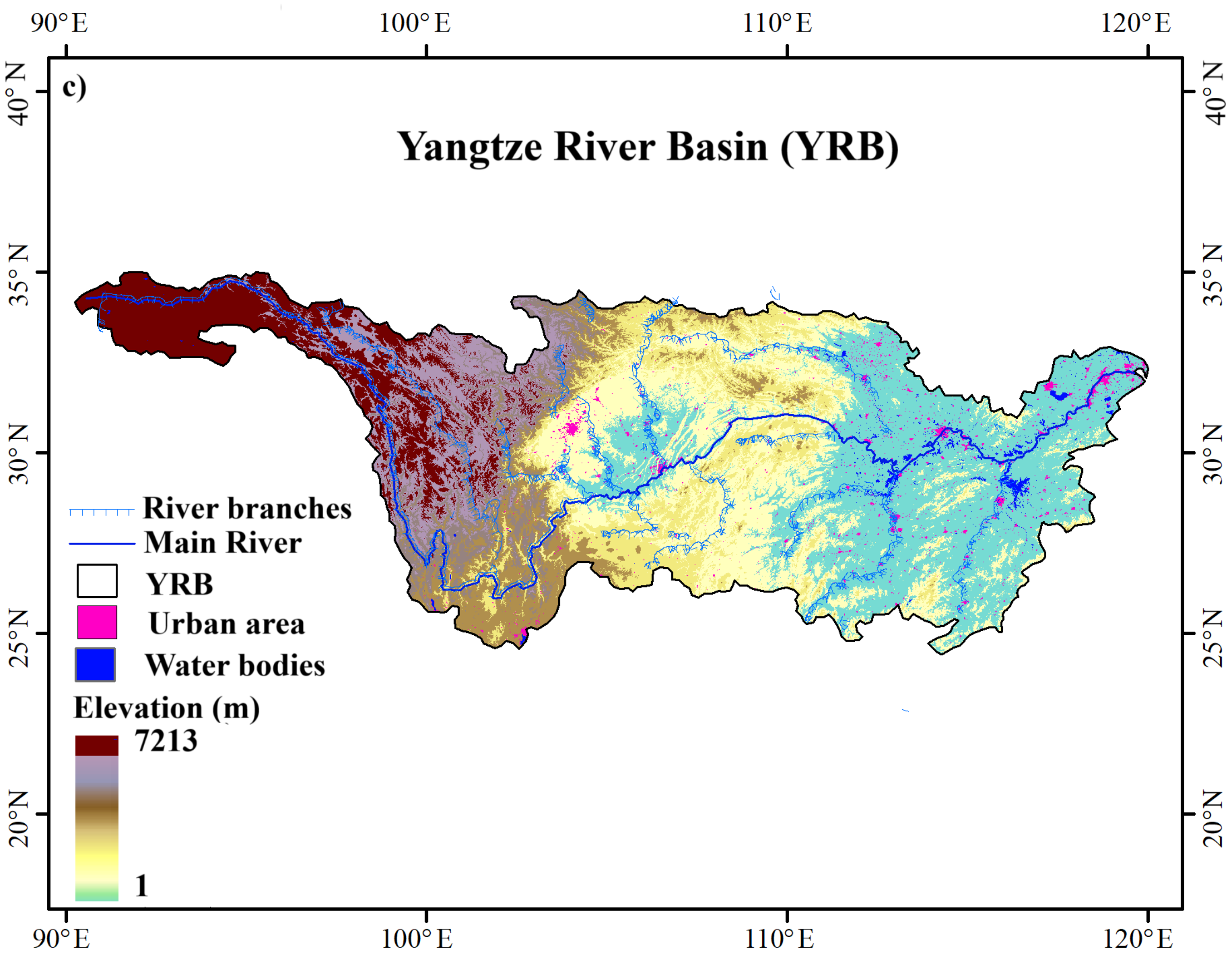
Table 2. Detailed descriptions of the summer SAT paleoclimate datasets used in this study.

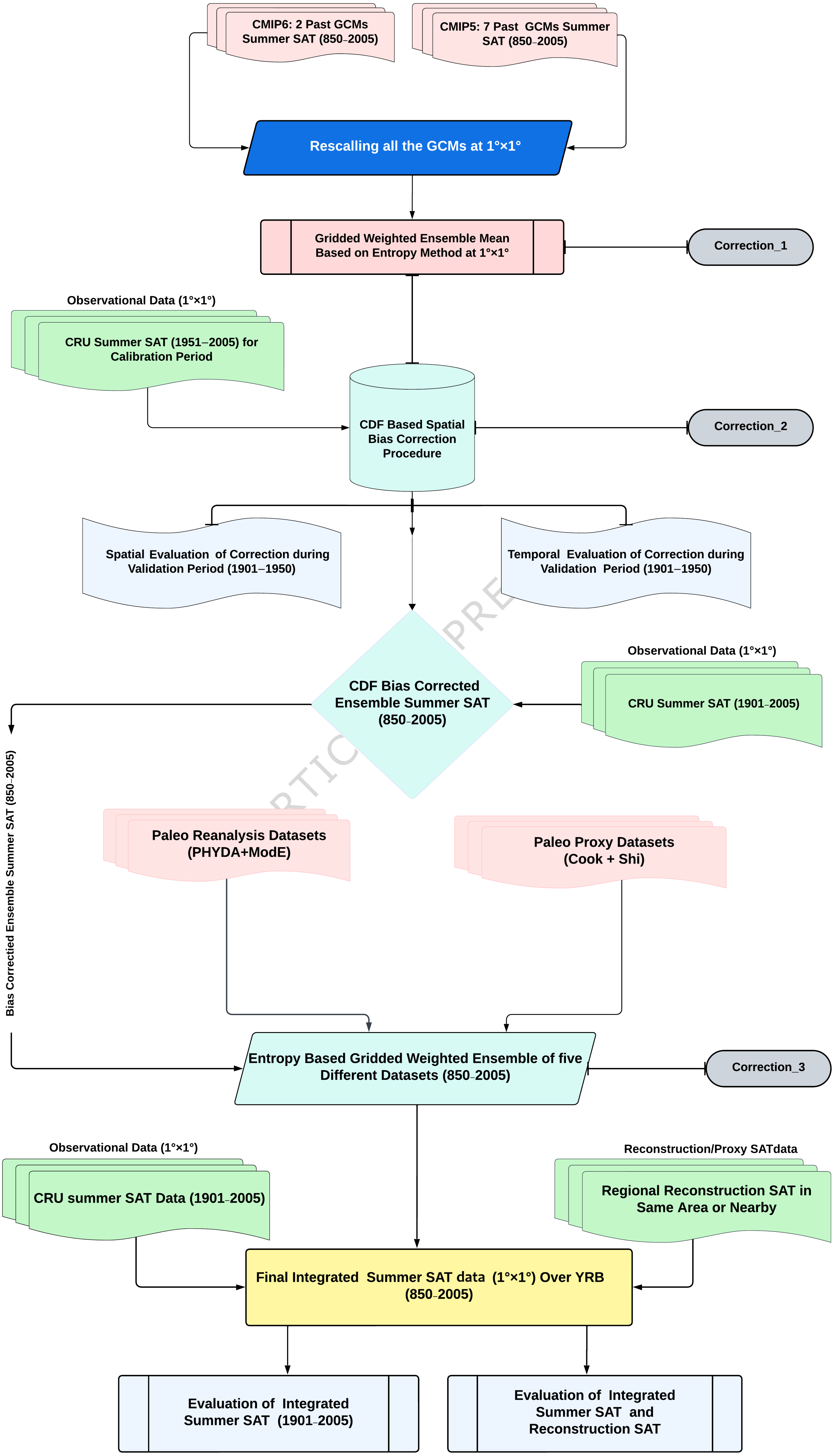
No	Dataset	Season	Time period	Spatial resolution	Data type	Region
1	**ModE**	Summer	1421–2005	$1.8^o \times 1.8^o$	Paleo reanalysis	Global
2	**PHYDA**	Summer	850–2005	$2.5^o \times 1.88^o$	Paleo reanalysis	Global
3	**Shi et al**	Summer	900–1999	$5^o \times 5^o$	Paleo reconstruction	Asia
4	**Cook et al**	Summer	850–2005	$2^o \times 2^o$	Paleo reconstruction	East Asia



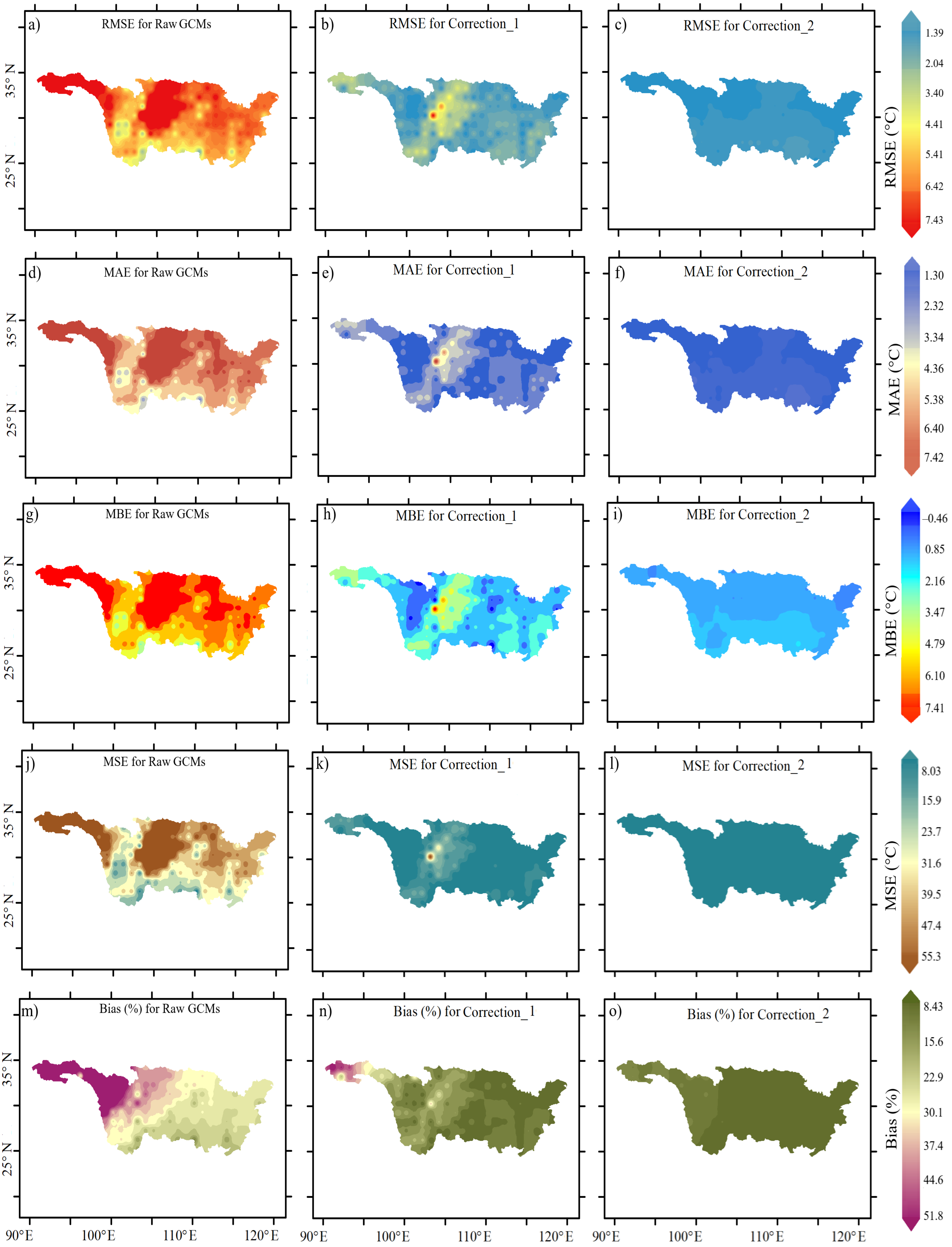
b)

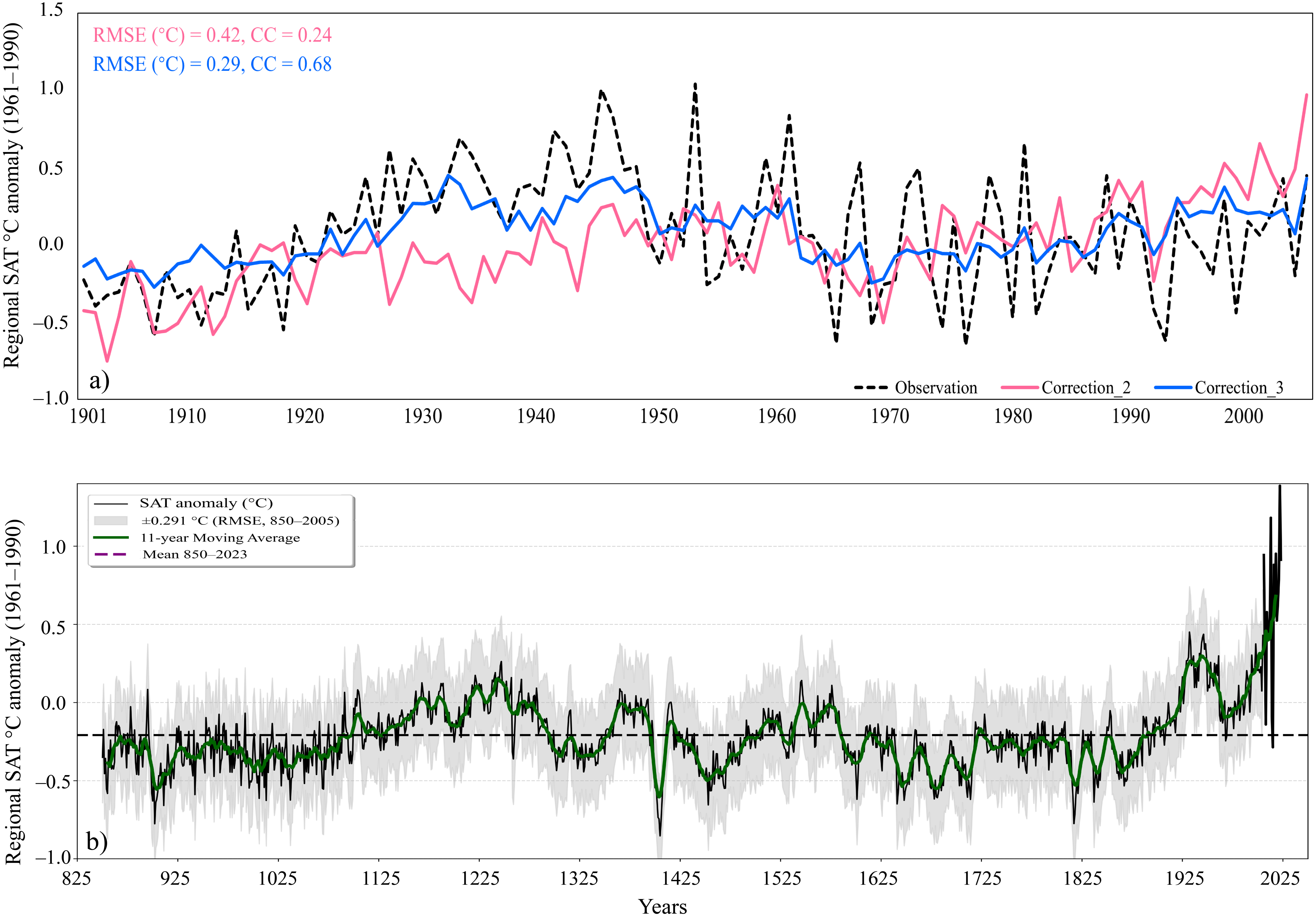
No	Parameters	YRB
1	Basin area (10^3Km^2)	1800
2	River length (10^3Km^2)	6.30
3	Water discharge (Km^3/yr)	896
4	Population (million)	420
5	Latitude range	24 ° 30 " , 35 ° 45 "



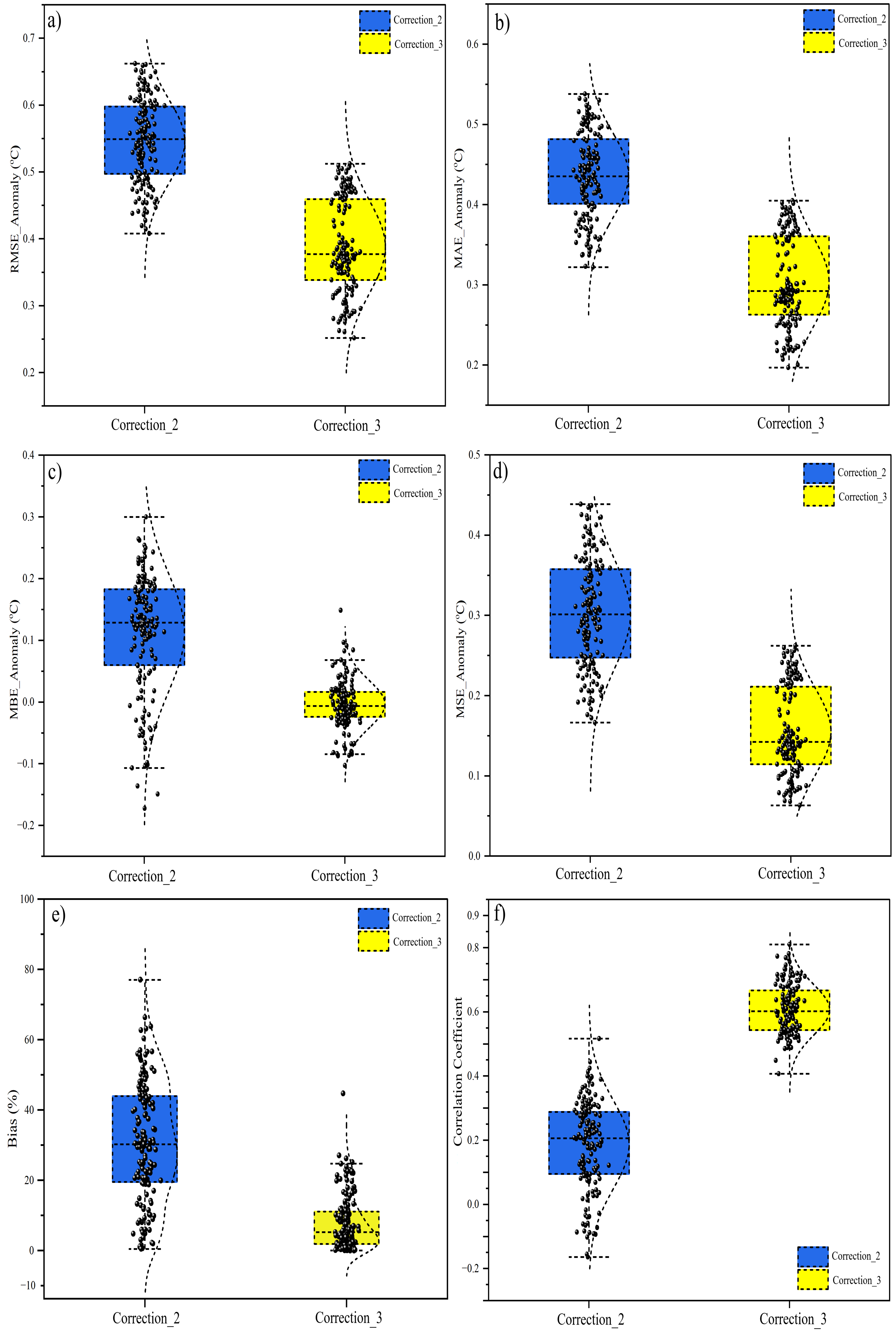

CMIP6: 2 Past GCMs Summer SAT (850–2005)
CMIP5: 7 Past GCMs Summer SAT (850–2005)
Rescalling all the GCMs at 1°×1°
Gridded Weighted Ensemble Mean Based on Entropy Method at 1°×1°
Correction_1
Observational Data (1°×1°)
CRU Summer SAT (1951–2005) for Calibration Period
CDF Based Spatial Bias Correction Procedure
Correction_2
Spatial Evaluation of Correction during Validation Period (1901–1950)
Temporal Evaluation of Correction during Validation Period (1901–1950)
Observational Data (1°×1°)
CDF Bias Corrected Ensemble Summer SAT (850–2005)
CRU Summer SAT (1901–2005)
Bias Corrected Ensemble Summer SAT (850–2005)
Paleo Reanalysis Datasets (PHYDA+ModE)
Paleo Proxy Datasets (Cook + Shi)
Entropy Based Gridded Weighted Ensemble of five Different Datasets (850–2005)
Correction_3
Observational Data (1°×1°)
CRU summer SAT Data (1901–2005)
Reconstruction/Proxy SATdata
Regional Reconstruction SAT in Same Area or Nearby
Final Integrated Summer SAT data (1°×1°) Over YRB (850–2005)
Evaluation of Integrated Summer SAT (1901–2005)
Evaluation of Integrated Summer SAT and Reconstruction SAT

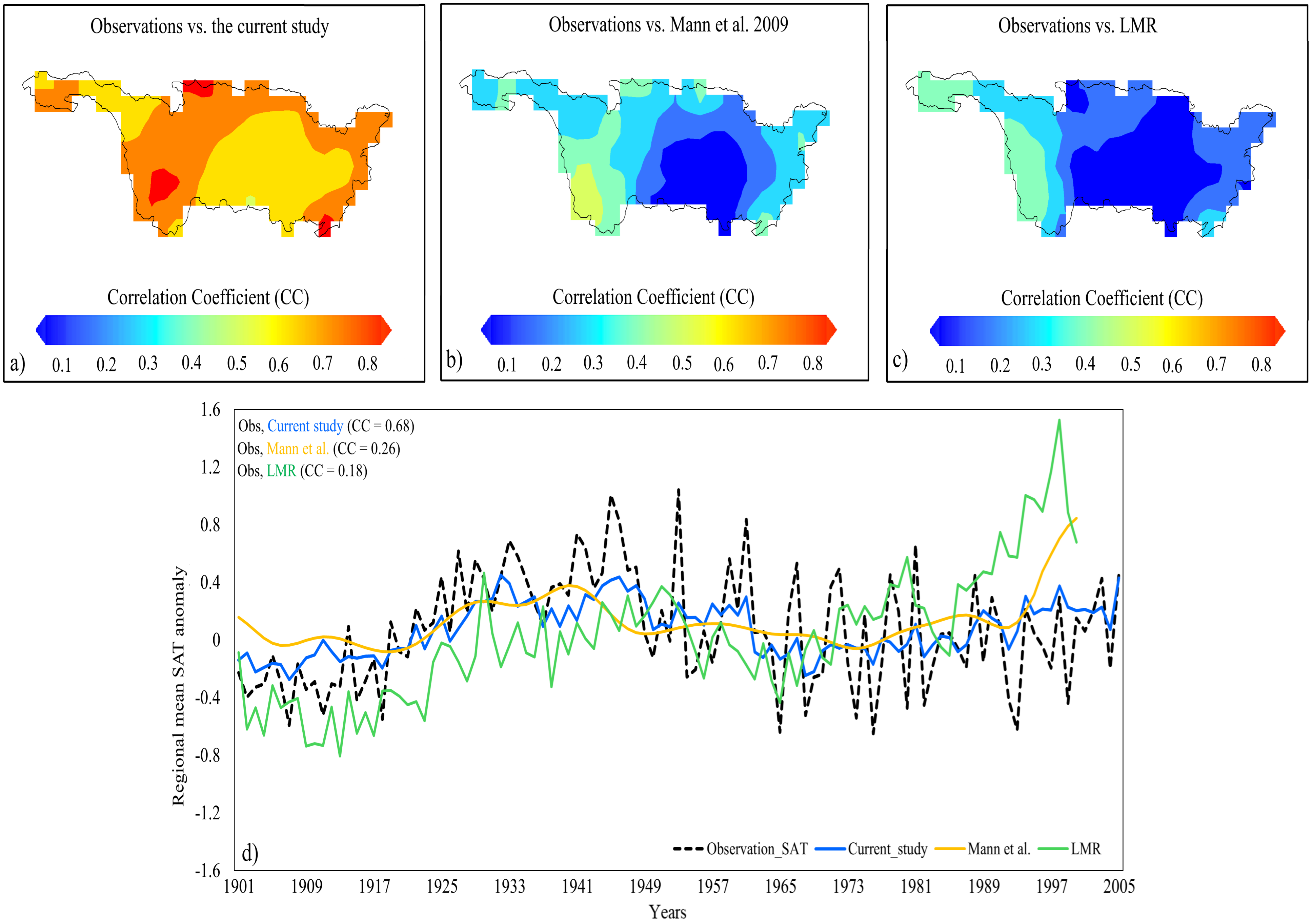

Raw GCMs
Correction_1
Correction_2
a)
RMSE (°C)
0
1
2
3
4
5
6
7
8
1901
1906
1911
1916
1921
1926
1931
1936
1941
1946
1950
Years
b)
MAE (°C)
0
1
2
3
4
5
6
7
8
1901
1906
1911
1916
1921
1926
1931
1936
1941
1946
1950
Years
c)
MBE (°C)
-0.5
0.5
1.5
2.5
3.5
4.5
5.5
6.5
7.5
1901
1906
1911
1916
1921
1926
1931
1936
1941
1946
1950
Years
d)
MSE (°C)
0
2
5
15
25
35
45
1901
1906
1911
1916
1921
1926
1931
1936
1941
1946
1950
Years
e)
Bias (%)
0
5
10
15
20
25
30
1901
1906
1911
1916
1921
1926
1931
1936
1941
1946
1950
Years
f)
Correlation Coefficient
0.95
0.96
0.97
0.98
0.99
1
1901
1906
1911
1916
1921
1926
1931
1936
1941
1946
1950
Years


a)
RMSE for Raw GCMs
b)
RMSE for Correction_1
c)
RMSE for Correction_2
RMSE (°C)
1.39
2.04
3.40
4.41
5.41
6.42
7.43
d)
MAE for Raw GCMs
e)
MAE for Correction_1
f)
MAE for Correction_2
MAE (°C)
1.30
2.32
3.34
4.36
5.38
6.40
7.42
g)
MBE for Raw GCMs
h)
MBE for Correction_1
i)
MBE for Correction_2
MBE (°C)
−0.46
0.85
2.16
3.47
4.79
6.10
7.41
j)
MSE for Raw GCMs
k)
MSE for Correction_1
l)
MSE for Correction_2
MSE (°C)
8.03
15.9
23.7
31.6
39.5
47.4
55.3
m)
Bias (%) for Raw GCMs
n)
Bias (%) for Correction_1
o)
Bias (%) for Correction_2
Bias (%)
8.43
15.6
22.9
30.1
37.4
44.6
51.8
35° N
25° N
90°E
100°E
110°E
120°E



RMSE (°C) = 0.42, CC = 0.24
RMSE (°C) = 0.29, CC = 0.68
Regional SAT °C anomaly (1961–1990)
1.5
1.0
0.5
0.0
−0.5
−1.0
a)
Observation
Correction_2
Correction_3
1901
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
SAT anomaly (°C)
±0.291 °C (RMSE, 850–2005)
11-year Moving Average
Mean 850–2023
Regional SAT °C anomaly (1961–1990)
1.0
0.5
0.0
−0.5
−1.0
b)
825
925
1025
1125
1225
1325
1425
1525
1625
1725
1825
1925
2025
Years


a)
RMSE_Anomaly (°C)
0.2
0.3
0.4
0.5
0.6
0.7
b)
MAE_Anomaly (°C)
0.2
0.3
0.4
0.5
0.6
c)
MBE_Anomaly (°C)
−0.2
−0.1
0.0
0.1
0.2
0.3
0.4
d)
MSE_Anomaly (°C)
0.0
0.1
0.2
0.3
0.4
0.5
e)
Bias (%)
−10
0
20
40
60
80
100
f)
Correlation Coefficient
−0.2
0.0
0.2
0.4
0.6
0.8
0.9
Correction_2
Correction_3


Observations vs. the current study
Correlation Coefficient (CC)
a)
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8
Observations vs. Mann et al. 2009
Correlation Coefficient (CC)
b)
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8
Observations vs. LMR
Correlation Coefficient (CC)
c)
0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8
Obs, Current study (CC = 0.68)
Obs, Mann et al. (CC = 0.26)
Obs, LMR (CC = 0.18)
d)
Observation_SAT
Current_study
Mann et al.
LMR
Regional mean SAT anomaly
1.6
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
-1.6
1901 1909 1917 1925 1933 1941 1949 1957 1965 1973 1981 1989 1997 2005
Years