https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

Open Access Earth System Science Data Discussions

An AI-Driven Reconstruction of Global Surface Temperature with Emphasis on Refining the Antarctic Record

Chenxi Ouyang[1,2,3], Qingxiang Li[1,2,3], Zichen Li[1,2,3], Sihao Wei[1,2,3]

[1] School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, China
[2] Key Laboratory of Tropical Atmosphere–Ocean System, Ministry of Education, Zhuhai, China
[3] Southern Laboratory of Ocean Science and Engineering (Guangdong Zhuhai), Zhuhai, China

Correspondence to: Qingxiang Li (liqingx5@mail.sysu.edu.cn)

Abstract. Accurate estimates of long-term surface temperature (ST) changes are fundamental not only for assessing observed warming, but also for improving the reliability of future climate projections. However, substantial missing information in global ST datasets, remains a major source of uncertainty in estimating global or regional temperature changes. Recent advances in artificial intelligence (AI) have promoted the effective application of deep learning approaches, such as image inpainting and transfer learning, in reconstructing incomplete geophysical datasets. In this study, partial convolutional neural network (PConv) models were trained using the 20CR reanalysis data and CMIP6 climate model outputs as training samples, with the aim of achieving a proper reconstruction of the global surface temperature dataset. To address differences among existing sea surface temperature (SST) datasets, we reconstruct global monthly ST fields since 1850 by merging the China global Land Surface Air Temperature (C-LSAT2.1) dataset with Extended Reconstructed Sea Surface Temperature (ERSSTv6) dataset and Met Office Hadley Centre's sea surface temperature (HadSST4) dataset, respectively. Although both reconstructions reliably reproduce large-scale spatial patterns and long-term variations, the merge of C-LSAT2.1 with HadSST4 exhibits greater physical consistency and is therefore adopted as our preferred reconstruction. In particular, validation against station observations indicates that the reconstructions perform well over the Antarctica after 1961, where observational coverage is extremely sparse. Based on this framework, we developed the China global Artificial Intelligence Reconstructed Surface Temperature$_{\mathrm{20CR/CMIP6}}$ (C-AIRST$_{\mathrm{R/M}}$) datasets, providing spatially complete global monthly ST anomaly reconstructions since 1850 with a spatial resolution of 5°×2.5°. These datasets offer improved support for extending long-term climate records and for applications in polar climate assessment, as well as in climate monitoring, detection, and attribution studies. The C-AIRST$_{\mathrm{R/M}}$ datasets can be downloaded at https://doi.org/10.6084/m9.figshare.30663797.v1 (Ouyang et al., 2025). They are also available from http://www.gwpu.net/en/h-col-103.html (last access: 21 November 2025).

1 Introduction

Global surface temperature (ST) is one of the most fundamental variables in the climate system, directly reflecting the state of the Earth's energy balance, it plays a central role in the monitoring and assessment of climate change (IPCC, 2013, 2021).

Although sporadic surface temperature (ST) observations have been available since the late seventeenth century, continuous observational datasets capable of representing global-scale ST variability did not emerge until the mid-nineteenth century (Cowtan and Way, 2014; Jones, 2016). The sparse distribution of observation sites introduces uncertainties in estimates of both global and regional climate changes (Katz et al., 2013; Karl et al., 2015; Huang et al., 2017; Li et al., 2021, 2022; Sun et al., 2021, 2022). Because the number of meteorological sensors and stations is limited, it is difficult to derive globally representative conclusions directly from observational records. Therefore, developing effective approaches to reconstruct global climate information has become particularly necessary (Vose et al., 2021; Morice et al., 2021).The evolution of the atmosphere and ocean follows the fundamental physical laws of mass, momentum, and energy conservation, these constraints imply that the climate field exhibits a certain degree of spatial and temporal continuity and predictability, and based on these properties, it is possible to infer missing information through statistical or dynamical relationships even in regions with sparse observations (Lorenz, 1963; Trenberth et al., 2003). Consequently, researchers have employed various methods to reconstruct missing climate information, including smoothing and interpolation techniques (Rayner et al., 2003; Vose et al., 2012; Lenssen et al., 2019; Li et al., 2021), principal component analysis (PCA) and its variants such as empirical orthogonal teleconnection (EOT) (Huang et al., 2017; Sun et al., 2021, 2022), and data-interpolating empirical orthogonal functions (DINEOF) (Beckers et al., 2003; Huang et al., 2017). These methods have played an important role in filling missing values and extracting climate signals from noisy data (Beckers et al., 2003; Wang and Clow, 2020). However, missing information inevitably introduces uncertainties and structural biases. The propagation of observational errors, inconsistencies among reanalysis products, and the inherent limitations of interpolation assumptions can all affect the reliability of reconstruction results (Huang et al., 2017; Morice et al., 2021). These issues are particularly pronounced in regions such as Africa, South America, and Antarctica, where sparse observations and harsh environmental conditions pose greater challenges for traditional reconstruction methods.

The Antarctica holds an irreplaceable position in the global climate system, and its enormous glacier masses store a substantial portion of the world's freshwater and play a major role in determining future sea level change (IPCC, 2013). At the same time, this region's radiation budget and the complex interactions among ice, atmosphere, and ocean exert profound influences on the atmospheric circulation of the Southern Hemisphere (SH) and the global energy balance (Kennicutt et al., 2019). Consequently, accurately characterizing the spatiotemporal variations of ST over Antarctica is essential for assessing the polar amplification effect, diagnosing changes in ice sheet mass balance, and improving the performance of global climate models. However, due to its extreme geographical conditions, harsh climate, and logistical and communication constraints, observational data from Antarctica remain extremely scarce and temporally uneven. Since the International Geophysical Year (1957/1958), national research programs have gradually established an automatic weather station (AWS) network across the Antarctica, providing valuable data for long-term climate monitoring (Jones et al., 2019; Wang et al., 2023). Nevertheless, most of these stations are concentrated along the coastal regions and near research bases, while observations over the interior plateau remain sparse. As a result, observation-based regional temperature fields often hard to capture the overall spatial structure of ST variability (Bromwich et al. 2025a). Against this background, Antarctic temperature reconstruction has become a major focus of polar climate research. Researchers have commonly adopted multi-source data fusion strategies that integrate limited in situ observations with reanalysis and satellite remote sensing products, using statistical and machine learning methods for spatial and temporal interpolation or extrapolation. For instance, Nicolas et al. (2014) and Bromwich et al. (2025a) applied ordinary kriging with different datasets as weights to spatially extrapolate Antarctic ST, effectively mitigating the problem of sparse observational coverage. Ma et al. (2025) reconstructed the Antarctic temperature field using a deep learning model. Nielsen et al. (2023) improved the spatiotemporal consistency of polar temperature estimates by calibrating MODIS land/ice ST against AWS observations through linear regression. Moreover, Xie et al. (under review) combined limited in situ station data with MODIS ST retrievals within a Bayesian framework, reproducing the evolution of Antarctic climate variability since the beginning of this century. Despite these advances, Antarctic temperature reconstruction still faces considerable challenges. Systematic biases exist among different data sources, and reanalysis products often show substantial uncertainties at high latitudes. Satellite observations are significantly affected by cloud cover and complex topography, while traditional statistical interpolation methods struggle to fully capture nonlinear spatial structures (Wang et al., 2023; Wang et al., 2025; Bromwich et al., 2025; Ma et al., 2025). Therefore, developing reliable methods to reconstruct Antarctic surface air temperature under limited observational constraints remains one of the central scientific challenges in polar climate reconstruction research.

At present, a relatively mature technical framework for global ST reconstruction has been developed. The Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2021) includes five observational ST products: HadCRUT5, NOAAGlobalTemp-Interim, GISTEMPv4, Berkeley Earth, and China-MST-Interim. These datasets apply various reconstruction or interpolation methods to generate homogenized global ST records with as complete a spatial coverage as possible (Morice et al., 2021; Vose et al., 2021; Lenssen et al., 2019; Rohde, 2020; Sun et al., 2021). With the rapid development of artificial intelligence (AI), deep learning has created new opportunities in atmospheric science (Liu et al., 2018; Ham et al., 2019; Kadow et al., 2020; Irrgang et al., 2021). For example, NOAAGlobalTempv6 incorporates an artificial neural network (ANN) to extend data coverage and update the dataset to a globally complete product (Yin et al., 2024). Bochow et al. (2025) applied fast Fourier convolution to fill missing values in HadCRUT4, while Plésiat et al. (2024) examined the ability of partial convolution-based networks to reproduce historical spatial patterns of climate extremes. Partial convolutional networks (PConv), originally proposed by Liu et al. (2018) for image inpainting, perform convolution operations using only valid (non-missing) pixels and dynamically update the validity mask during training. This design greatly enhances reconstruction accuracy, particularly for fields with extensive missing regions. The underlying concept has since been extended to climate reconstruction tasks, in which PConv-based models learn spatial structures and nonlinear dependencies from large climate datasets, thereby enabling “intelligent” completion of incomplete climate fields (Kadow et al., 2020; Zhou et al., 2022; Jiao et al., 2023; Bochow et al., 2025; Ma et al., 2025). Building on these previous efforts, the present study applies PConv reconstruction framework for global ST anomaly fields. We construct training samples from long-term reanalysis products and climate model simulations, and use in situ Antarctic station observations as the primary reference. By combining statistical approaches with convolutional neural networks, the methodology seeks to balance model

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

interpretability with nonlinear representation capability. The reconstructed temperature anomaly fields are then systematically compared with existing observational and reconstructed products to assess the strengths and limitations of AI-based methods for this application.

2 Data and methods

2.1 Data resources

The China global Merged Surface Temperature (China-MST/C-MST) dataset is an established global ST product (Yun et al., 2019; Sun et al., 2021, 2022; Li et al., 2020, 2021). The latest version, C-MST3.0, is classified into three variants (C-MST3.0-Nrec, C-MST3.0-Imin and C-MST3.0-Imax) based on the spatial coverage of reconstructed Arctic sea-ice ST (Li et al., under review). C-MST3.0 is constructed by combining ST from the China land surface air temperature (LSAT) dataset C-LSAT2.1 (Wei et al., 2025) with the Extended Reconstructed Sea Surface Temperature version 6 (ERSSTv6), released by NOAA/NCEI (Huang et al., 2025a, 2025b). In this study, LSAT from C-LSAT2.1 are separately merged with the sea surface temperature (SST) products ERSSTv6 and the Met Office Hadley Centre's sea surface temperature (HadSST4.1.1.0) dataset (Kennedy et al., 2019), and the resulting merged datasets, referred to as “Merge-E” and “Merge-H”, are used as the original inputs for the AI-based reconstruction. Both datasets provide monthly mean ST from 1850 to 2024.

In addition, two historical monthly ST datasets are employed to construct training sets for the AI reconstructions. The Twentieth Century Reanalysis version 3 (20CR; Slivinski et al., 2019), which provides monthly ST fields for 1850–2015, is used to train the “20CR-AI” model, whereas the “Historical” simulations from the Coupled Model Intercomparison Project Phase 6 (CMIP6), offering monthly ST fields for 1850–2014, are used to train the “CMIP6-AI” model. The 20CR dataset contains 80 ensemble members, whereas the CMIP6 dataset contains 105 ensemble members (Table S1).

To better reconstruct realistic climate conditions over Antarctica, we further employ monthly Antarctic station observations. Most stations originate from SCAR READER (Turner et al., 2004), GHCNm v4 (Menne et al., 2018), AntAWS (Wang et al., 2023), and GSOD (NCEI, 1999). Additional stations are obtained from the OSU Polar Meteorology Group (Bromwich et al., 2025b), Météo-France (Météo-France, 2025), the University of Wisconsin-Madison (South Pole Meteorology Office, 2025), and the National Institute of Water and Atmospheric Research Ltd (NIWA, 2025). Stations with more than 25 valid years during 1961–1990 are selected, yielding a temporal span of 1957–2024. Following Bromwich et al. (2025a), necessary gap-filling procedures are applied to the Antarctic station data. Subsequently, homogenization tests are performed to eliminate discontinuities caused by station relocations, sensor changes, or other non-climatic shifts. Station observations are mapped to the corresponding 5°×2.5° grid cells (based on station longitude and latitude) and treated as valid data constraints for the reconstruction. Details of the station metadata and temperature anomaly time series are shown in Table S2 and Figure S1 in the Supplement.

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

2.2 LSAT and SST merging method

In this study, we adopted the data merging approach of Yun et al. (2019) to merge LSAT from C-LSAT2.1 separately with SST from ERSSTv6 and HadSST4. The land–ocean mask used in this process was obtained from the NCAR Command Language (NCL) "landsea.nc" mask file (download at: http://www.ncl.ucar.edu/Applications/Data/cdf/landsea.nc, last access: 23 July 2025). The mask file contains five categories: 0 for ocean, 1 for land, 2 for lake, 3 for small island, and 4 for ice shelf. In this study, land, lake, small island, and ice shelf were treated as land, while the other category was considered ocean. Due to the relatively low resolution of the reconstructed datasets (5°×2.5°), this merging approach results in some grid cells along the land–ocean boundaries, particularly along the Antarctica where observational data are sparse, containing only ocean information. AI-based reconstruction is influenced by features at the edges of missing data (Liu et al., 2018). To minimize the impact of ocean-dominated grid cells along land–ocean boundaries, we reasonably expanded the Antarctic land mask relative to the original land–sea mask at the 5°×2.5° scale (Fig. S2). This expansion has negligible effect on the merging process in regions with extremely sparse Antarctic observations but allows the AI reconstruction to better capture ST patterns in the Antarctic coastal and peripheral areas.

2.3 AI training and reconstruction process

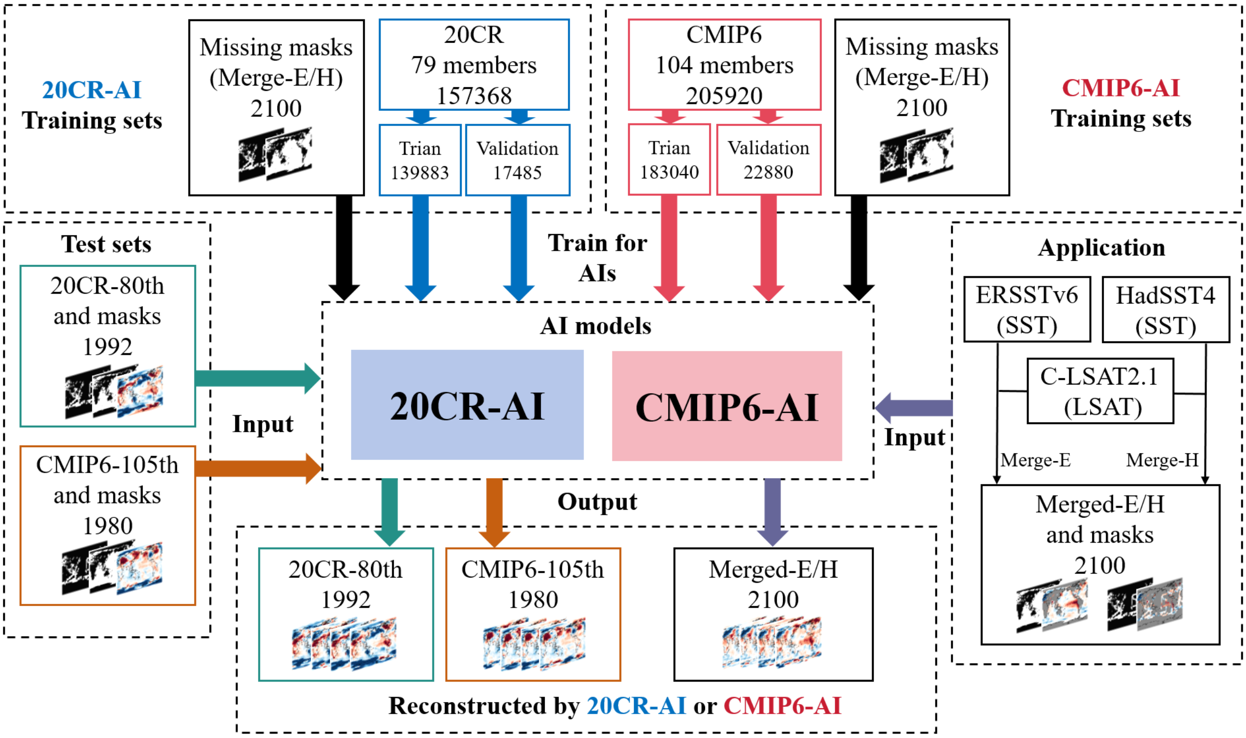


Figure 1: AI training and reconstruction.

The PConv method has demonstrated excellent performance in reconstructing globally missing climate data, including ST anomaly fields, wind fields, and surface solar radiation (Kadow et al., 2020; Zhou et al., 2022; Jiao et al., 2023). The underlying principle is to treat the climate field as a two-dimensional image, with missing regions regarded as irregular holes to be inpainted. PConv is then trained to learn spatial features from a large set of climate data samples, enabling the inference of plausible spatial patterns in the missing regions (Kadow et al., 2020). In this study, the AI reconstruction follows a similar approach. The reconstruction workflow is illustrated in Figure 1, and consists of the following steps:

(1) Resolution standardization: All datasets used for training and reconstruction were regridded to a common spatial resolution of 5°×2.5°, resulting in a 72×72 grid.

(2) Sample partitioning: Monthly temperature anomaly fields from 20CR (80 ensemble members) and CMIP6 (105 ensemble members) for 1850–2015 (2014) relative to the 1961–1990 climatology were divided into training and testing sets. One ensemble member was randomly selected as the test set, while the remaining 79 members from 20CR were used to train the 20CR-AI model, and 104 CMIP6 members were used to train the CMIP6-AI model.

(3) Training sample construction: The training samples were randomly partitioned, with 8/9 used for model training and 1/9 for validation. This resulted in 139,883 training samples and 17,485 validation samples for the 20CR-AI model, and 183,040 training samples and 22,880 validation samples for the CMIP6-AI model.

(4) Mask configuration: Observational data from Merge-E and Merge-H, covering monthly values from 1850 to 2024, were used as masks for missing values in the respective reconstruction schemes. Observed grid points were marked as "1" and missing grid points as "0", serving as inputs for model training and validation.

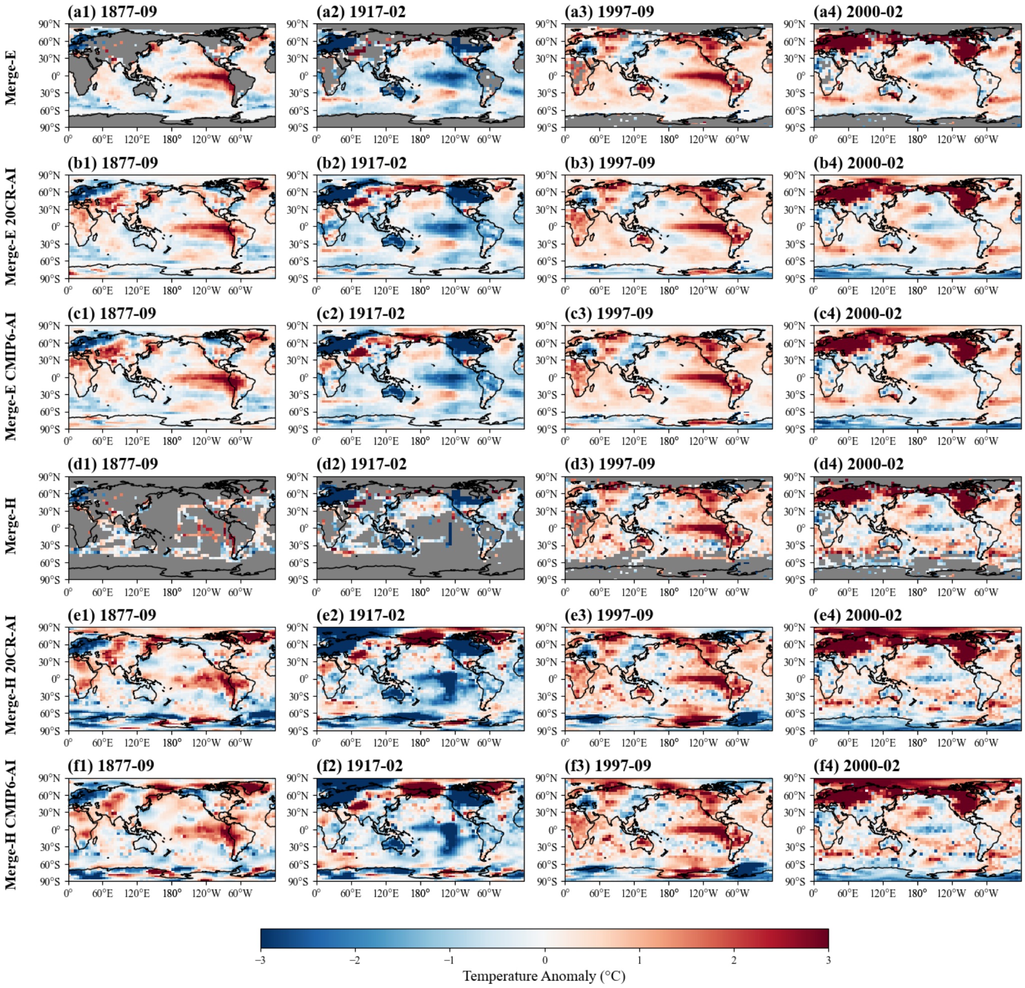
(5) Model training and fine-tuning: Both AI models were trained for 500,000 iterations with a learning rate of 0.0002, followed by an additional 500,000 iterations with a reduced learning rate of 0.00006. The batch size was set to 16, and computations were performed on an NVIDIA GeForce RTX 4060 GPU for approximately 8 hours.

2.4 Model Post-Processing

To ensure comparability across different data sources, all external datasets, including observations, reanalysis products, and reconstructed data, were first regridded to a common 72×72 regular grid consistent with the AI reconstruction output, with a uniform spatial resolution of 5°×2.5°. All data were converted to ST anomalies relative to the 1961–1990 climatology to eliminate differences in the climate reference among datasets. Subsequently, the performance of the AI reconstruction was systematically evaluated. The primary evaluation metrics included the spatial correlation coefficient and root mean square error (RMSE), which were used to assess the consistency and bias of the reconstructed fields relative to reference datasets in the spatial domain. Temporal stability and reconstruction accuracy were further examined using annual mean correlation coefficients and RMSE time series (Fig. S4, S5 and S6). To comprehensively assess the global applicability and reliability of the AI reconstruction, it was compared against multiple representative global temperature datasets, including Berkeley Earth, HadCRUT5, NOAAGlobalTempv6, C-MST3.0-Imax, and the four reconstruction scenarios in this study (Merge-E 20CR-AI, Merge-E CMIP6-AI, Merge-H 20CR-AI, Merge-H CMIP6-AI). The comparison included global annual mean temperature anomaly time series, linear trend estimates, and their statistical significance. In the Antarctica, where observational coverage is sparse, targeted evaluation of the AI reconstruction was performed. Reference datasets included the ERA5 reanalysis, 180 Berkeley Earth, HadCRUT5, GISTEMPv4, the reconstruction by Bromwich et al. (2025a), and Antarctic ground-based station observations used in this study. Cross-validation among these multiple data sources was conducted to assess the reliability and consistency of the AI reconstruction under conditions of sparse polar observations.

3 Global Reconstruction Results

3.1 Characteristics of the Global Reconstruction Results



185

Figure 2: Global temperature anomaly fields before and after reconstruction for four typical months. Merge-E original (a1–a4), Merge-E reconstructed with 20CR-AI (b1–b4), Merge-E reconstructed with CMIP6-AI (c1–c4), Merge-H original (d1–d4), Merge-H reconstructed with 20CR-AI (e1–e4), Merge-H reconstructed with CMIP6-AI (f1–f4).

Prior to the 19th century, the C-LSAT2.1 dataset contained substantial missing data in Asia, South America, Africa, and Antarctica due to the sparse distribution of land-based stations, and this situation was improved considerably after the 19th century (Wei et al., 2025). In addition, Antarctic observational stations began recording data progressively after the 1957/1958 International Geophysical Year; however, observational coverage in Antarctica remains extremely sparse (Wang et al., 2023). As shown in Figure S3, the original ERSSTv6 dataset exhibits higher ocean data coverage compared to HadSST4. This is because ERSSTv6 employs an ANN reconstruction method, which allows for more accurate and stable reconstruction of SST in sparsely observed regions, reduces excessive smoothing, and better preserves spatial variability (Huang et al., 2025a, 2025b). In contrast, HadSST4 primarily relies on statistical interpolation and ensemble-based methods, which aim to correct biases and estimate uncertainties within grid cells, without performing large-scale spatial reconstruction (Kennedy et al., 2019).

Considering the differences between the two SST datasets, this study performed AI-based reconstruction using two distinct merged datasets to investigate potential discrepancies under different conditions. Figures 2 (b1–b4, c1–c4, e1–e4, f1–f4) show the complete temperature anomaly fields reconstructed from Merge-E and Merge-H using the 20CR-AI and CMIP6-AI models. The global spatial patterns of the four reconstructed fields are largely consistent across most regions; however, in high-latitude areas with extremely sparse early observations, the spatial distributions of Merge-E and Merge-H reconstructions show some differences. The reconstruction performance of the model improves as the coverage of original valid data in the model validations (Fig. S4, S5 and S6). Compared to Merge-H, Merge-E exhibits higher spatial coverage and fewer missing gaps in polar regions, particularly over Antarctica and its surrounding seas. This leads to better AI reconstruction performance under the Merge-E mask than under Merge-H (Fig. S4 and S6). Visually, the reconstructed fields in Antarctica and adjacent regions are smoother in Merge-E than in Merge-H (Fig. 2), indicating that during image inpainting, the AI reconstruction is influenced to some extent by the colour, texture, and style features at the edges of missing regions in the two different datasets (Liu et al., 2018; Nazeri et al., 2019), thereby leading to distinct reconstructed features in the early periods when large areas of missing data occur in Merge-E and Merge-H.

Moreover, the AI reconstruction effectively reproduces characteristic climate events in key years. In particular, the results shown in Figures 2 (e1, f1) clearly capture the strong El Niño event of 1877, characterized by significantly positive SST anomalies in the central and eastern tropical Pacific and a distinct east–west dipole pattern of warm and cold anomalies along the equator, reflecting the typical signal of the El Niño–Southern Oscillation (ENSO). This demonstrates that the model is capable of reconstructing large-scale spatial patterns and temporal evolution of the temperature field even under extremely sparse observational coverage, indicating robust performance and spatial consistency.

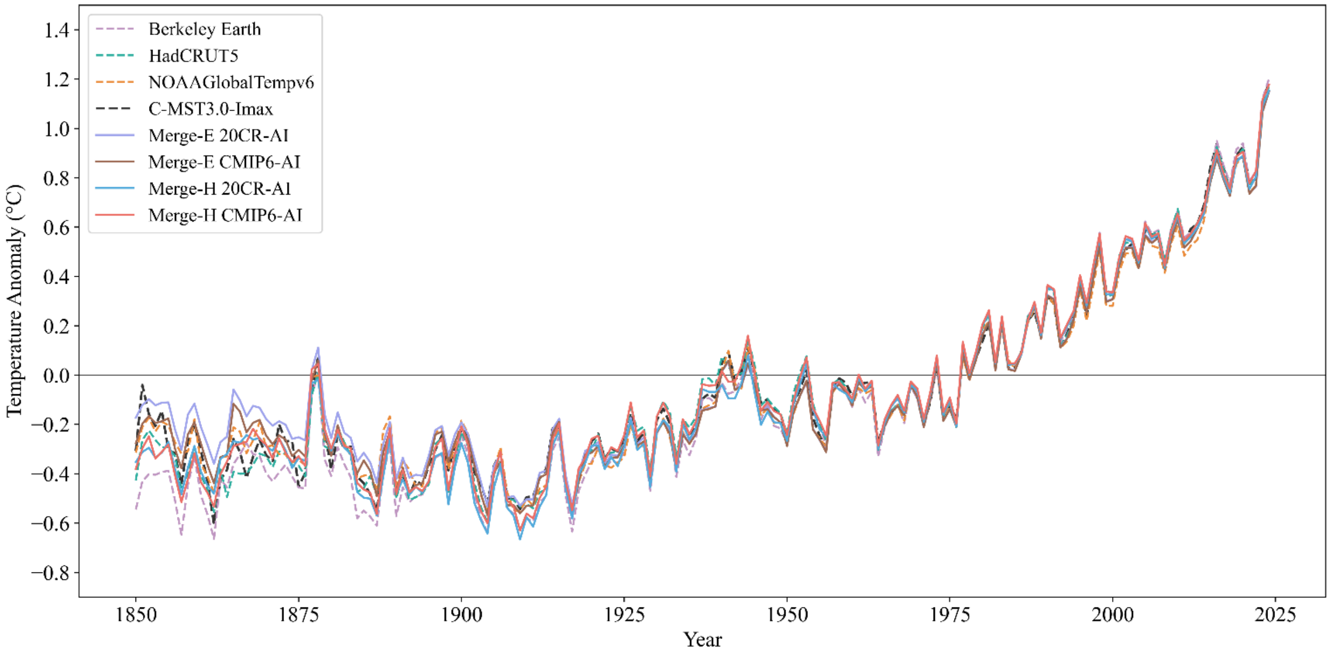


Figure 3: Global mean temperature anomaly time series from 1850 to 2024 (relative to the 1961–1990 climatology).

Table 1: Trends of global annual mean surface temperature over different periods and 95% confidence intervals (°C per decade; The global warming level (GWL) denotes the increase in global mean temperature (°C) in 2024 relative to the 1850–1900 reference period).

Dataset/Period	1850–2024	1900–2024	1950–2024	1979–2024	GWL
Berkeley Earth	0.070±0.006	0.107±0.008	0.163±0.104	0.206±0.024	1.62
HadCRUT5	0.065±0.006	0.100±0.008	0.156±0.015	0.203±0.023	1.53
NOAAGlobalTempv6	0.058±0.006	0.098±0.008	0.155±0.013	0.196±0.024	1.45
C-MST3.0-Imax	0.061±0.006	0.098±0.008	0.159±0.014	0.210±0.022	1.50
Merge-E 20CR-AI	0.052±0.007	0.097±0.008	0.157±0.013	0.199±0.023	1.37
Merge-E CMIP6-AI	0.057±0.007	0.099±0.008	0.157±0.013	0.199±0.023	1.43
Merge-H 20CR-AI	0.064±0.006	0.105±0.008	0.155±0.014	0.196±0.023	1.45
Merge-H CMIP6-AI	0.064±0.006	0.101±0.008	0.156±0.014	0.199±0.023	1.52

The Global mean temperature anomaly time series reconstructed by AI is shown in Figure 3, reflecting the long-term variations of global mean temperature from 1850 to 2024. The AI reconstruction results from Merge-E and Merge-H exhibit overall consistent trends; however, certain differences exist during periods with sparse early observations. Prior to the 1890s, the global ST anomalies in the Merge-E reconstruction are generally higher than those in the Merge-H reconstruction. Given the significant contribution of SST to global annual mean temperature variability, this difference is mainly attributable to the differing SST datasets used in the two reconstructions. PConv rely heavily on the spatial structures present in regions with valid observations when reconstructing climate fields containing extensive missing data (Liu et al., 2018; Reichstein et al., 2019; Toms et al., 2020). In the Merge-E scenario, when early land observations are extremely sparse but oceanic grid points exhibit relatively complete spatial coverage (Fig. S3), the dominant spatial gradients, covariance structures, and anomaly patterns learned by the model during training are inherently governed by the ocean. According to the general properties of deep learning, convolutional neural networks preferentially learn the most frequent and statistically stable features in the input data (LeCun et al., 2015). Consequently, during 1850–1890, when land observations are limited, the PConv model is inevitably dominated by oceanic signals during feature extraction. This causes the model to "fill" missing regions with spatial structures resembling those of the ocean, which are typically smoother and warmer, thereby producing systematic biases in the land driven by ocean-dominated features. Around the 1890s, however, the substantial increase in land-based observation stations (Wei et al., 2025) enriches the spatial information over land. As the spatial patterns and variability of land anomalies begin to occupy sufficient weight in the training data, the PConv model becomes capable of simultaneously learning both land and ocean features. This reduces the early-period dominance of oceanic coverage and allows the reconstruction to gradually converge toward a more realistic combined land–ocean spatial structure. In the experiments described above, the SST product ERSSTv6 in Merge-E already incorporates ANN-based spatial infilling, resulting in smoother and more homogeneous oceanic features (Huang et al., 2025a, 2025b). This further amplifies the dominance of oceanic characteristics during feature learning. In contrast, in the Merge-H scenario, both land and ocean fields contain extensive missing data in the early period (Wei et al., 2025; Kennedy et al., 2019), meaning that no single domain provides a strong dominant signal. As a result, the PConv model is more likely to learn spatial structures jointly constrained by both the land and the ocean domains, thereby reducing the risk of domain-specific biases, particularly biases associated with the ocean. Consequently, the Merge-E reconstruction exhibits pronounced overfitting to oceanic structures before the late nineteenth century, producing warmer biases and ultimately yielding a lower long-term trend and global warming level (GWL) than that of Merge-H (Table 1). Based on the above results, we adopt the more physically consistent merging scheme, the Merge-H product generated by combining C-LSAT2.1 with HadSST4, as the primary focus of our analysis.

It is noteworthy that the two AI models show small differences in their long-term reconstructions and trends under the Merge-H scenario (Fig. 3, Table 1). For the Merge-H case, the 1850–2024 warming trend produced by both AI reconstructions is 0.064 ± 0.006 °C per decade, which is broadly consistent with the trend in HadCRUT5. The GWL reconstructed by CMIP6-AI reaches 1.52 °C, which is slightly higher than the 1.5 °C estimated from 20CR-AI. This indicates that both AI reconstructions from Merge-H suggest that global warming in 2024 has reached 1.5 °C above the baseline of 1850–1900.

At the decadal scale, all reconstruction sequences clearly reproduce the pronounced global warming associated with the extreme 1876–1877 El Niño event, which led to the warmest years prior to the 1940s (Huang et al., 2020). From 1850 to the early 20th century, global mean temperatures remained relatively stable, occasionally interrupted by short-term cooling events associated with enhanced volcanic activity and slight decreases in solar radiation (Fang et al., 2022). Between the 1910s and 1940s, global mean temperature experienced the first rapid and sustained warming phase of the 20th century, with a marked increase in magnitude. From the mid-20th century to the mid-1970s, the warming trend slowed, followed by a renewed and significant warming phase starting in the late 1970s, which continues to the present day, reflecting the dominant influence of anthropogenic greenhouse gas emissions (IPCC, 2013, 2021). It is noteworthy that none of the AI reconstruction sequences show a significant "warming slow-down" during 1998–2012, consistent with the findings of Li et al. (2021). Overall, the AI reconstruction sequences shown in Figure 3 reasonably reproduce the phased characteristics of global climate changes over the past 175 years at the annual scale, further demonstrating the robustness and reliability of the AI framework for long-term climate reconstruction.

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

3.2 Characteristics of Regional Surface Temperature Changes

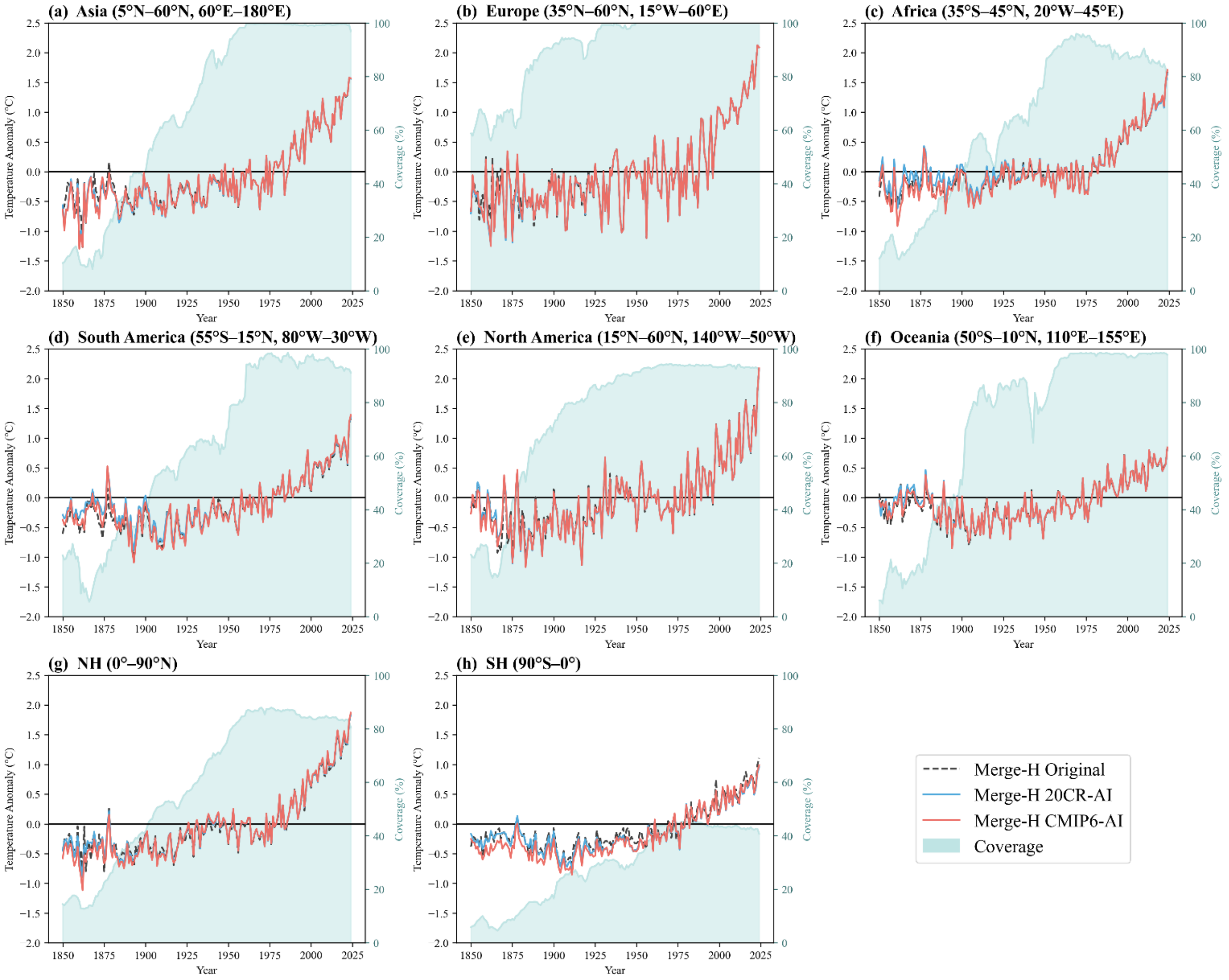


Figure 4: Annual mean surface temperature anomaly time series from 1850 to 2024 in different regions under the Merge-H Scenario. Asia, Europe, Africa, South America, North America, Oceania, the Northern Hemisphere, and the Southern Hemisphere (a–h). Coverage indicates the original data availability in each region.

Global reconstruction of land regions is one of the main focuses of this study. We first analyse the long-term variations of ST in different land regions before and after reconstruction (Fig. 4). In this analysis, the original Merge-E dataset is compared with its two AI-based reconstructions. Considering the extremely sparse observational coverage over Antarctica, this subsection does not discuss the Antarctic region, which will be addressed in detail in Section 4 (Antarctic Reconstruction Results). As a reference for reconstruction performance, the effective data coverage of the original datasets in each region is also presented. Table 2 provides the linear trends of annual mean ST and their 95% CI for different regions from 1850 to 2024.

Table 2: Temperature trends and 95% confidence intervals (CI) in different regions from 1850 to 2024 (°C per decade).

Region/Dataset	Merge-H Original	Merge-H 20CR-AI	Merge-H CMIP6-AI
Asia	0.083 ± 0.010	0.090 ± 0.010	0.091 ± 0.010
Europe	0.094 ± 0.014	0.099 ± 0.014	0.097 ± 0.014
Africa	0.069 ± 0.009	0.060 ± 0.009	0.074 ± 0.009
South America	0.068 ± 0.009	0.055 ± 0.010	0.064 ± 0.010
North America	0.080 ± 0.012	0.080 ± 0.013	0.083 ± 0.012
Oceania	0.041 ± 0.008	0.035 ± 0.009	0.038 ± 0.008
NH	0.087 ± 0.010	0.089 ± 0.010	0.097 ± 0.010
SH	0.054 ± 0.007	0.050 ± 0.008	0.063 ± 0.007

Figure 4 shows that as observational coverage increases, the differences between different reconstruction results decrease markedly. This feature is particularly evident in regions with increasingly abundant observational data, where the AI-reconstructed annual variations become more stable. In contrast, in periods or regions with sparse early observations and low data coverage, discrepancies among reconstructions are more pronounced. The differences between the two AI reconstructions are mainly concentrated in regions such as Africa and South America, where early observations were scarce. Due to the large extent of missing data in these areas, which corresponds to large holes in the input images that need to be filled, the accuracy of the AI reconstructions is significantly affected (Liu et al., 2018). Considering Table 2, the differences between the two AI reconstructions are unavoidable in some regions during periods of low early data coverage. For the Merge-H reconstruction, the two AI models exhibit slight differences in long-term trends, with 20CR-AI showing a lower temperature trend in all regions except Europe compared to CMIP6-AI. However, the trends remain within reasonable ranges relative to the pre-reconstruction data. Except for regions with substantial early data gaps, such as Africa, South America, and Oceania, most AI-reconstructed temperature trends exhibit stronger warming compared with pre-reconstruction trends.

On long timescales, all regions show significant warming trends, particularly since the 1970s, when the warming rate accelerated. Nevertheless, warming is uneven across continents due to differences in response scales and persistence (Li et al., 2022). According to Table 2, in the Merge-H reconstruction scenario, Europe experienced the most pronounced warming between 1850 and 2024, with trends of 0.099 ± 0.014 and 0.097 ± 0.014 °C per decade, whereas Oceania experienced the smallest warming, with corresponding trends of 0.035 ± 0.009 and 0.038 ± 0.008 °C per decade. The warming rate in the SH is noticeably lower than in the NH. This north–south asymmetry under global warming results from significant differences in

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

land–ocean distribution, with the Southern Ocean absorbing the majority of heat, leading to greater climate inertia in the SH (Hansen et al., 2010).

A comprehensive analysis of regional ST anomaly time series indicates that NH land areas contribute most significantly to global land warming. Within the NH, Europe, Asia, and North America are the primary contributors. In conditions of extremely low data coverage, the both reconstructions inevitably exhibit biases. However, as coverage increases, the AI reconstruction experiments show good consistency and stability in these regions.

3.3 Characteristics of seasonal surface temperature changes

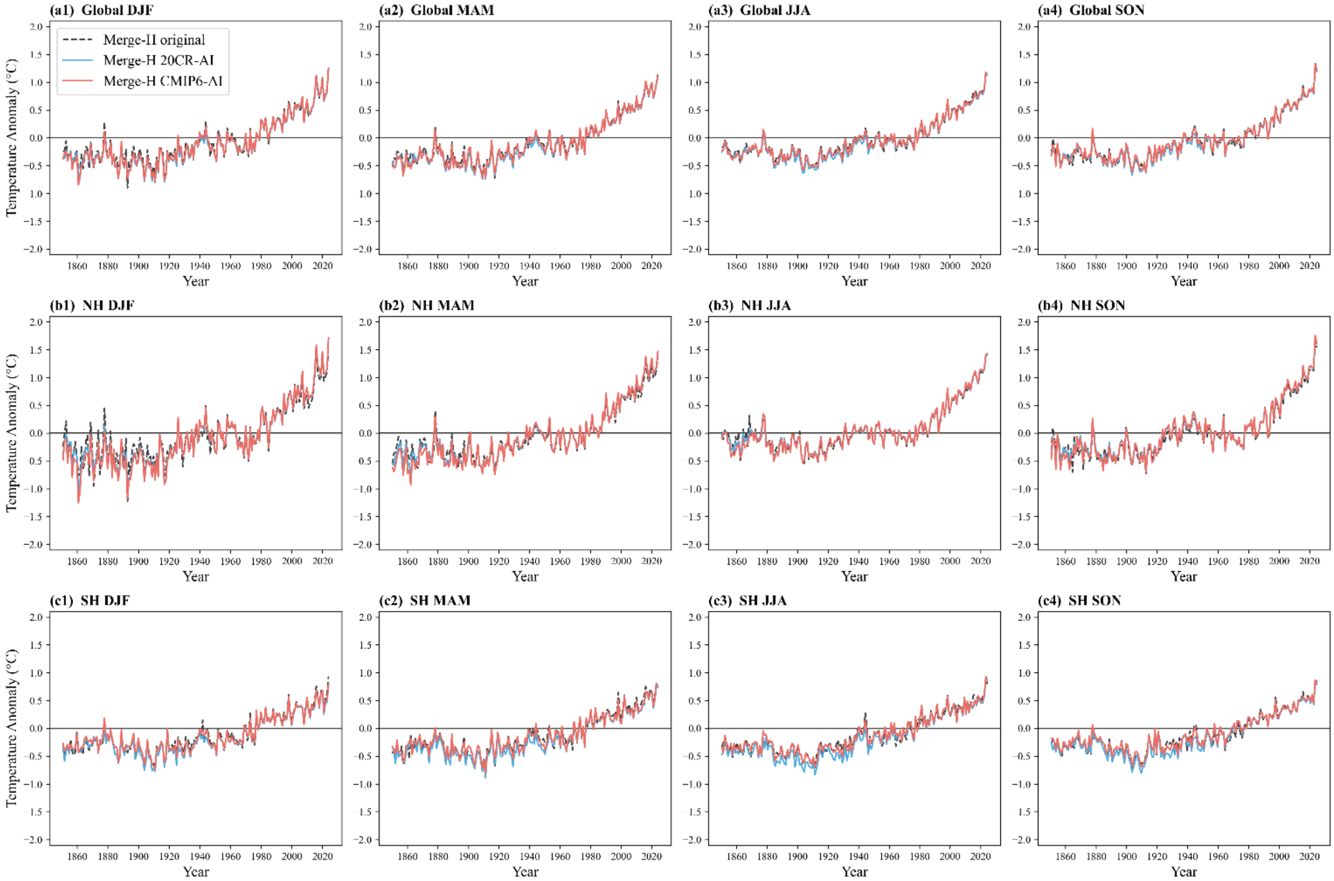


Figure 5: Seasonal surface temperature anomaly time series from 1850 to 2024 under the Merge-H scenario. Global winter, spring, summer and autumn(a1–a4), Northern Hemisphere winter, spring, summer and autumn (b1–b4), Southern Hemisphere winter, spring, summer and autumn (c1–c4).

The seasonal ST anomaly time series from 1850 to 2024 are shown in Figure 5. It can be seen that all seasons exhibited a slight cooling trend prior to the 20th century. The period from the 1910s to 1940s corresponds to the first notable warming phase, followed by a relatively stable trend in global land temperatures during the 1940s–1970s. Since the late 1970s, global

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

land temperatures have shown a pronounced and rapid warming trend, consistent with the overall global temperature changes.

From the perspective of interannual variability, winter (DJF) temperatures in the NH exhibit the largest interannual fluctuations, whereas summer (JJA) shows the smallest variability. This contrast is mainly attributed to the dominance of the NH. The larger interannual variability in winter land temperatures is associated with a relatively unstable atmospheric circulation, enhanced planetary wave activity, and strong snow and sea ice feedbacks, which make land temperatures more sensitive to external perturbations. In contrast, the smaller interannual variability in summer is due to the dominance of solar radiation, a more stable boundary layer, and land processes such as evapotranspiration and soil moisture that provide thermal buffering, thereby reducing the impact of atmospheric disturbances on temperature (Jones et al., 2014). The four seasons in the SH do not exhibit significant differences in interannual variability, reflecting stronger thermal inertia and lower year-to-year fluctuations.

Table 3: Seasonal temperature trends and 95% CI from 1850 to 2024 (°C per decade).

Season/Dataset		Merge-H Original	Merge-H 20CR-AI	Merge-H CMIP6-AI
Global	DJF	0.062±0.007	0.067±0.007	0.068±0.007
	MAM	0.064±0.007	0.067±0.007	0.068±0.007
	JJA	0.057±0.006	0.059±0.007	0.058±0.006
	SON	0.062±0.007	0.063±0.007	0.062±0.006
NH	DJF	0.073±0.010	0.084±0.010	0.090±0.010
	MAM	0.071±0.008	0.078±0.008	0.084±0.008
	JJA	0.056±0.008	0.057±0.008	0.062±0.008
	SON	0.070±0.008	0.071±0.009	0.075±0.009
SH	DJF	0.051±0.006	0.050±0.006	0.046±0.006
	MAM	0.056±0.006	0.056±0.006	0.051±0.006
	JJA	0.056±0.005	0.061±0.006	0.055±0.006
	SON	0.052±0.006	0.054±0.006	0.049±0.006

As shown in Table 3, the reconstructed global temperature trends are fastest in DJF, with trends of 0.067 ± 0.007 and 0.068 ± 0.007 °C per decade, followed by MAM (0.067 ± 0.007 and 0.068 ± 0.007 °C per decade), while JJA shows the slowest warming rate (0.059 ± 0.007 and 0.058 ± 0.006 °C per decade), with the NH contributing the most to the long-term warming trend. A further comparison of different AI reconstruction scenarios indicates that CMIP6-AI reconstructions generally exhibit slightly higher seasonal temperature trends than the 20CR-AI reconstructions, whereas in the SH the former shows lower trends than the latter. Overall, the reconstructed seasonal temperature series capture the major climate

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

change features revealed by both observations and models. The slight differences between CMIP6-AI and 20CR-AI highlight the importance of the coverage, spatial characteristics, and training samples of the original datasets for AI reconstruction performance, providing a reference for further improving multi-source climate data fusion and AI-based reconstruction methods.

340 4 Antarctic Reconstruction Results

Table 4: Validation between Antarctic station observations and monthly temperature anomalies from AI-reconstructed grid cells under two reconstruction schemes (r: correlation coefficient; RMSE: root mean square error, °C)

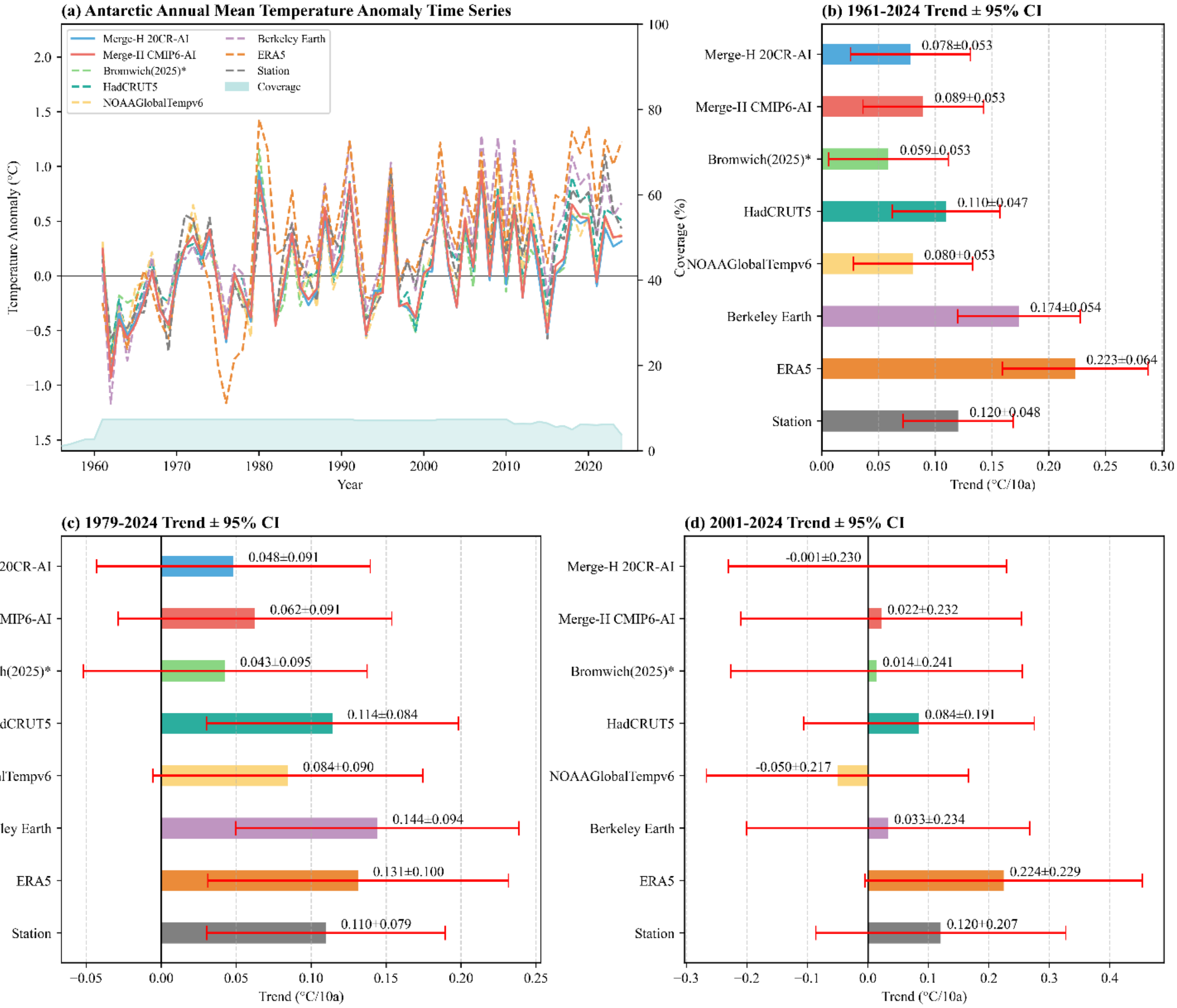
Station	Latitude	Longitude	Elevation(m)	Coverage	Merge-H 20CR-AI		Merge-H CMIP6-AI	
					r	RMSE	r	RMSE
Dome A	−80.4	77.4	4084	2006–2019	0.52	1.28	0.64	1.17
Drescher	−79.2	−19.0	27	1993–2003	0.60	0.72	0.56	0.77
Ferrel	−77.9	170.8	45	2007–2023	0.84	0.65	0.88	0.58
G3	−70.9	69.9	84	2002–2020	0.49	0.79	0.44	0.81
GC41	−71.6	111.3	2763	1984–2005	0.80	0.57	0.96	0.43
General Belgrano	−78.0	−38.8	32	1956–1978	0.70	0.76	0.69	0.77
GF08	−68.5	102.1	2125	1987–2007	0.87	0.46	0.87	0.45
LG10	−71.3	59.2	2619	1993–2005	0.73	0.43	0.71	0.43
LG35	−76.0	65.0	2345	1994–2007	0.70	0.51	0.64	0.53
LG59	−73.5	76.9	2565	1994–2003	0.63	0.45	0.54	0.48
Mount Siple	−73.2	−127.1	230	1993–2006	0.84	0.34	0.84	0.34
Nordenskiold	−73.1	−13.4	497	1995–2019	0.82	0.31	0.81	0.32
Russkaya	−74.7	−136.9	100	1981–1989	0.73	0.66	0.69	0.70
Thurston Island	−72.5	−97.6	225	2011–2021	0.74	0.56	0.85	0.43
Average (14 stations)					0.72	0.61	0.72	0.59

Among the 81 Antarctic stations used in this study (Table S2 and Fig. S1), the vast majority began recording Antarctic 345 climate data after the 1957/1958 International Geophysical Year. Even so, the effective data coverage in Antarctica remained only around 10% after 1961 (Figure 6a). During model validation for Antarctic data (Fig. S6), the reconstructed results from both AI models showed high performance, with correlation coefficients with the test dataset reaching approximately 0.9 and RMSE below 1.3 °C after 1961. Due to the scarcity of observational data prior to 1961, validation and assessment of

https://doi.org/10.5194/essd-2025-717
Preprint. Discussion started: 14 January 2026

Antarctic data for earlier years are not feasible. Therefore, this study focuses on the systematic validation and analysis of Antarctic reconstructed data from 1961 onward.

To better evaluate the performance of the AI models in reconstructing ST over the Antarctic continent since 1961 and to ensure the validity of the verification, we selected 14 independent Antarctic stations that were not used in the reconstruction and were located outside the effective grid cells. The monthly observed anomalies at these stations were compared with the corresponding reconstructed anomalies to assess the consistency and statistically test whether the reconstructions reliably reproduced temperature variability (Monaghan et al., 2008; Nicolas et al., 2014; Bromwich et al., 2025a). The results indicate that most stations exhibit high correlation with the reconstructed grid cell temperature series, with RMSE below 0.9°C except for Dome A (Table 4). Dome A, located at the highest point of the East Antarctic Plateau, is influenced by surface winds controlled by synoptic-scale circulation, and its temperature records can reach extreme lows, making it more challenging for the reconstruction method to capture its variability (Scambos et al., 2018). Overall, the average correlation between the 14 stations and the two reconstruction schemes is 0.72, with corresponding mean RMSE values of 0.61°C and 0.59°C. Although there are slight differences among the reconstruction schemes, the CMIP6-AI reconstruction performs slightly better than the 20CR-AI reconstruction, and both schemes show good agreement with the observational data.



(a) Antarctic Annual Mean Temperature Anomaly Time Series

(b) 1961-2024 Trend ± 95% CI

(c) 1979-2024 Trend ± 95% CI

(d) 2001-2024 Trend ± 95% CI

Figure 6: Annual mean surface temperature anomaly time series over Antarctica from 1961 to 2024 (a). Linear trend of annual mean temperature and 95% confidence interval (°C per decade) during 1961–2024 (b), 1979–2024 (c) and 2001–2024 (d). "*" indicates data ending in 2022.

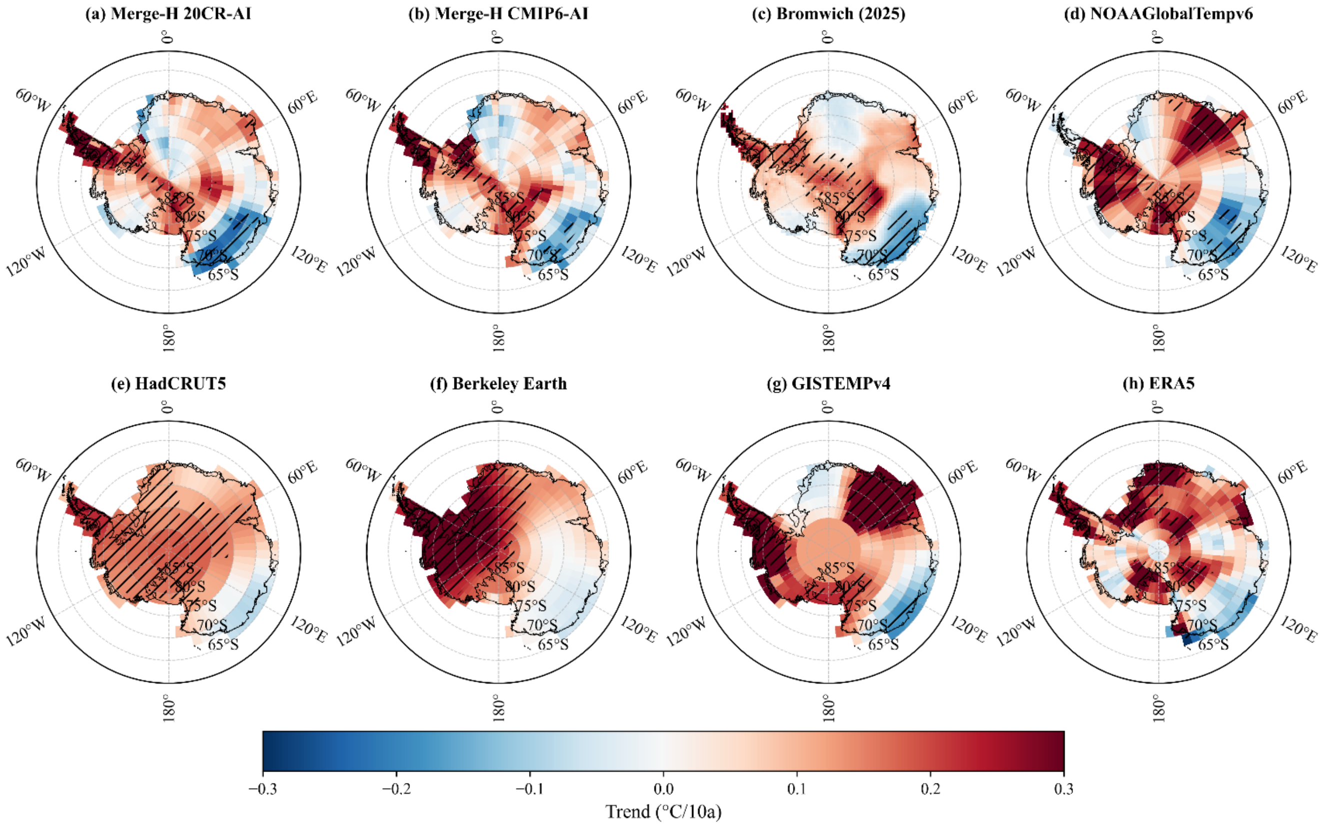
Figure 6 presents the annual mean ST anomalies over the Antarctic continent for 1961–2024, along with the linear temperature trends and their 95% CI for three representative periods. To comprehensively evaluate the reliability and performance of the AI reconstructions, the results were systematically compared with the arithmetic mean of 81 Antarctic observational stations used in this study (Station in Figure 6), the reconstruction from Bromwich et al. (2025a), HadCRUT5, NOAAGlobalTempv6, Berkeley Earth, and the ERA5 reanalysis product.

Overall, the AI reconstructions exhibit high consistency with multiple observational and reconstruction products in terms of interannual variability. In particular, the reconstructed anomalies closely track the phases of strong cold and warm years, indicating that the AI method reliably captures interannual climate signals in the Antarctic region. During 1961–2024, the linear warming trends derived from the two AI reconstruction schemes are 0.078 ± 0.053 and 0.089 ± 0.053°C per decade, respectively. The trends are similar in magnitude, show minor differences, and are both statistically significant at the 0.05 level, demonstrating a pronounced warming trend over the Antarctic continent since the 1960s (Fig. 6b).

Notably, since 1979, the warming trends in the AI reconstructions remain non-significant, consistent with Bromwich and NOAAGlobalTempv6, whereas HadCRUT5, Berkeley Earth, and ERA5 still show significant warming trends, in agreement with the arithmetic mean of the station data (Station) (Fig. 6c). The higher trend estimates in these products may stem from biases introduced by statistical and reanalysis methods under sparse observational conditions in polar regions. Previous studies indicate that the ERA5 reanalysis system exhibits substantial seasonal biases in near-surface Antarctic temperatures, likely due to weak constraints on surface turbulent processes and increased assimilation errors under strong inversions and low-friction conditions (Garza-Girón et al., 2024; Yang et al., 2025). Additionally, Berkeley Earth's spatial statistical extrapolation may produce systematic overestimation in high-latitude Antarctica (Rohde et al., 2020), although the exact contribution of statistical and assimilation methods to these biases remains to be quantified. These results highlight that systematic differences among data sources must be carefully considered when assessing long-term climate trends in polar regions.

Since 2001, both the AI reconstructions and all datasets except ERA5 indicate that Antarctic warming is not statistically significant (Fig. 6d). During this period of accelerated global mean temperature increase (Fig. 3), no significant warming is detected over Antarctica, suggesting that internal climate variability continues to exert a stronger influence on Antarctic climate than external forcing, warranting further attention.

Overall, the AI reconstruction results demonstrate reasonable temporal consistency, trend significance, and magnitude of variability, validating the feasibility and potential utility of AI-based climate reconstruction for the Antarctic region.


(a) Merge-H 20CR-AI

(b) Merge-H CMIP6-AI

(c) Bromwich (2025)

(d) NOAAGlobalTempv6

(e) HadCRUT5

(f) Berkeley Earth

(g) GISTEMPv4

(h) ERA5

Figure 7: Spatial distribution of Antarctic annual mean surface temperature trends for 1979–2024 ("///" indicates regions where the temperature trend is statistically significant at the 0.05 level).

The spatial pattern of Antarctic temperature change is of critical scientific importance for understanding regional climate variability and its potential driving mechanisms. Due to substantial differences in topography, sea ice coverage, circulation features, and ocean–atmosphere interactions across the Antarctic continent, temperature changes are not uniformly distributed but exhibit complex regional responses (Turner et al., 2005, 2019; Marshall, 2003, 2006). Figure 7 presents the spatial distribution of ST trends and their statistical significance for 1979–2024 derived from the two AI reconstruction schemes, together with the results from other representative datasets. Overall, both AI reconstruction schemes reveal significant warming over the Antarctic Peninsula, near the Ronne Ice Shelf, and the northeastern Ross Ice Shelf, while the Wilkes Land coast shows significant cooling. These spatial patterns are broadly consistent with the reconstructions of Bromwich, the NOAAGlobalTempv6 dataset, and reported observations or model simulations in these regions (Vaughan et al., 2003; Wang et al., 2025; Darelius et al., 2016; Clem et al., 2020; Sheehan et al., 2024), demonstrating the reliability of AI reconstructions in capturing key climate signals in Antarctica.

In contrast, the ERA5 product exhibits some spatial discrepancies: it shows significant warming near Queen Maud Land and west of the Ronne Ice Shelf, whereas cooling along the Wilkes Land coast is not statistically significant. The

HadCRUT5, Berkeley Earth, and GISTEMPv4 global temperature reconstructions rely on limited ground-based observations in Antarctica and extend spatially using different statistical methods (optimal interpolation, EOF, or spatially weighted averaging) (Morice et al., 2021; Rohde et al., 2020; Lenssen et al., 2019). As a result, their temperature trend fields are smoother and have weaker spatial gradients. All three datasets indicate significant warming over the Antarctic Peninsula, with HadCRUT5 and Berkeley Earth also showing extensive warming near the South Pole, while GISTEMPv4 indicates significant warming east of the Ross Ice Shelf and in eastern Queen Maud Land. Notably, NOAAGlobalTempv6 also shows warming in eastern Queen Maud Land, whereas the two AI reconstruction schemes, reconstruction from Bromwich, and ERA5 reveal positive but statistically insignificant trends in this region, which lies in the transitional zone of Antarctic continental temperature variability.

In summary, the AI reconstructions successfully capture the main regional temperature trends and their statistical significance across Antarctica, reflecting climate signals consistent with observations and reanalysis data while exhibiting plausible spatial variability in data-sparse areas. This demonstrates that AI-based climate reconstruction not only provides high temporal consistency but also possesses strong potential for spatial applications, offering a novel tool for understanding Antarctic climate change processes and their underlying drivers.

5 Limitations and future perspectives

AI reconstruction results exhibit high consistency with existing observational and reconstructed datasets in terms of spatial structure, long-term trends, and interannual variability, demonstrating the feasibility and reliability of this AI approach for climate reconstruction. However, compared with traditional statistical interpolation or reanalysis models, AI methods still present challenges regarding model dependency, physical interpretability, and long-term consistency that warrant further investigation.

The AI reconstruction approach employed in this study is based on image inpainting techniques, which treat climate fields as two-dimensional images and use PConv to infer missing data from learned spatial features in local neighborhoods. This method performs well in mid-latitude and low-latitude regions where spatial continuity is relatively strong, but because the computation relies on planar convolution operations, it cannot fully account for the spherical geometry of the Earth. Consequently, systematic limitations remain in high-latitude regions. Similarly, Kadow et al. (2020) and Bochow et al. (2025) highlighted that whether using partial convolution or Fourier convolution, while historical climate fields can be reasonably reconstructed, projecting the spherical Earth onto a two-dimensional equidistant grid introduces geometric distortion at high latitudes, affecting continuity on the sphere and potentially generating “edge discontinuities” or “artifacts” at the poles or at the image seams (at the 0°/360° boundary of the global map). The “artifact” features observed in the edge reconstruction of our images support this observation (Fig. 2 and S8). Esteves et al. (2023) reported that scaling spherical convolutional neural networks (Scaling Spherical CNNs) can directly model data on the sphere, achieving competitive results in weather forecasting tasks, demonstrating the potential of spherical convolutional frameworks in atmospheric sciences. However, high-resolution global climate reconstruction with such methods remains significantly limited by current GPU memory capacity and the high cost of high-performance computing. Overall, balancing spherical geometric accuracy with computational feasibility while maintaining global spatial continuity is a key direction for future AI climate reconstruction. Integrating spherical deep learning structures with physical constraints may enable more realistic and accurate reconstructions of polar climate variability.

The stability of AI reconstruction results largely depends on the spatial patterns of the training samples and the availability of original valid data. In this study, models trained on CMIP6 outputs generally captured higher long-term trends than those trained on 20CR, particularly, systematic biases also exist between the Merge-E and Merge-H reconstruction schemes. This indicates that while deep learning methods possess strong nonlinear fitting capabilities, they are constrained by the spatial structure and statistical characteristics of the input data. Moreover, we found that when missing regions are excessively large and contain no valid data, the AI model's prior correlation and RMSE deteriorate substantially, as observed in the pre-1961 polar regions. After introducing station observations post-1961 to increase effective coverage, the prior correlation improved and RMSE decreased, particularly for Merge-H, which had larger missing regions than Merge-E (Fig. S6). This finding emphasizes that while AI reconstruction can generate plausible climate fields under sparse observational conditions, sufficient image information or spatial constraints are required. Because the AI model does not explicitly include energy conservation or atmospheric dynamics constraints, extrapolations to data-sparse regions, such as early periods and polar regions, heavily rely on surrounding temperature spatial patterns, that is, the edge features of the missing regions (Liu et al., 2018). Thus, land–ocean boundaries, complex terrain, and extensive missing-data areas are major sources of AI reconstruction errors. Future studies may improve reconstructions by incorporating terrain corrections, adjusting land–ocean weighting, introducing local dynamical parameterizations, or using physics-informed loss functions. Furthermore, combining AI's ability to capture nonlinear features and local spatial patterns with the interpretability, computational transparency, and uncertainty quantification advantages of statistical methods, forming "AI plus physics-driven" or "AI plus statistical fusion" frameworks, may represent a promising direction for climate reconstruction.

AI method demonstrates strong potential and scalability in climate reconstruction, yet its application in modeling and reconstructing climate systems remains an evolving area. The physical consistency and long-term stability of current AI reconstruction outputs still require further validation under more stringent constraints. Future research should incorporate climate dynamics and energy balance constraints while preserving the nonlinear fitting advantages of deep learning to ensure interpretability and generalizability in complex climate contexts. With continued advancement in high-performance computing and climate big data, AI reconstruction methods may achieve higher resolution and stronger constraint global climate field reconstructions. By integrating spherical neural networks, physical information, and multi-source data fusion frameworks, future AI reconstruction systems could more accurately reproduce climate evolution in data-sparse regions, supporting studies on polar amplification, local climate change, and data monitoring, detection, and evaluation. Overall, AI is expected to play an increasingly important role in climate science, providing a robust technological foundation for understanding past and projecting future global and regional climate change.

6 Data availability

The C-AIRST$_{R/M}$ datasets are publicly available on the website at https://doi.org/10.6084/m9.figshare.30663797.v1 (Ouyang et al., 2025). They can also be accessed at http://www.gwpu.net/en/h-col-103.html (last access: 21 November 2025) for free.

7 Code availability

The code utilized in this project can be downloaded here or cloned here at https://github.com/FREVA-CLINT/climatereconstructionAI.

8 Conclusions

This study employed an AI model (PConv) to reconstruct global ST fields from two fused observational datasets. The results demonstrate that the AI reconstruction shows high consistency with multiple representative climate datasets in terms of interannual variability, long-term temperature trends, and spatial patterns, validating the effectiveness and reliability of deep learning-based image inpainting approaches for climate reconstruction tasks. A comprehensive comparison of the global mean ST time series derived from different reconstruction schemes in this study indicates that the datasets exhibit largely consistent long-term trends. In particular, for Antarctica after 1961, the AI reconstruction aligns well with both observational and reanalysis data, indicating that the AI-based approach provides reliable support for reconstructing continuous global ST fields since the mid-19th century. The main conclusions of this study are as follows:

(1) Convolutional neural network-based image inpainting can effectively fill global temperature data gaps. Both the 20CR-AI and CMIP6-AI models, which are trained on large-sample datasets, produced satisfactory reconstruction results. The differences between the two models in reconstruction quality and trend estimation are minor, demonstrating good generalization and stability. Reconstruction performance improves significantly with increased data coverage, highlighting the importance of data completeness for model performance.

(2) AI reconstructions provide a novel tool for climate system research. Compared with traditional statistical interpolation and reanalysis products, the AI method has advantages in capturing complex spatial structures and nonlinear changes. It can restore large-scale temperature anomaly patterns such as those associated with ENSO, and generate physically reasonable, fully covered climate fields even in data-sparse regions.

(3) High reliability of Antarctic AI reconstructions after 1961. Although the effective data coverage in Antarctica is only around 10% after 1961, the AI models achieve correlation coefficients of approximately 0.9 and RMSE values generally below 1.3°C in this region. The reconstructed sequences show strong consistency with previous studies in both temporal evolution and spatial distribution, indicating that deep learning models can effectively recover ST variations in high-latitude regions with sparse observations when sufficient climate information is present.

(4) Consistent long-term trends across different AI models in global and Antarctic regions. The 20CR-AI reconstructions show slightly lower global temperature trends than CMIP6-AI, but both reveal that 19th-century temperature anomalies were generally lower, with a slow cooling trend from the mid-19th century to the early 20th century. Since the mid-20th century, particularly during the satellite era, global temperatures have increased significantly, and no pronounced "warming hiatus" has been detected since 1998. In Antarctica, surface temperatures have shown a gradual warming trend since 1961, statistically significant at the 0.05 level, with strong warming particularly evident over the Antarctic Peninsula, Ronne Ice Shelf, Ross Ice Shelf, and surrounding regions.

Based on the Merge-H AI reconstruction schemes, this study developed spatially complete global monthly ST anomaly datasets for 1850–2024 with a spatial resolution of 5°×2.5°, termed the China global Artificial Intelligence Reconstructed Surface Temperature$_{\mathrm{20CR/CMIP6}}$ (C-AIRST$_{\mathrm{R/M}}$) datasets, which are reconstructed independently using the 20CR-AI and CMIP6-AI schemes based on the merged C-LSAT2.1 and HadSST4. Both datasets exhibit high temporal and spatial continuity, providing a solid foundation for extending long-term climate records, assessing polar climate change, and supporting climate monitoring, detection, and attribution. Overall, this study demonstrates the potential and application value of AI in climate data reconstruction. With further advancements in deep learning, physics-informed learning, and high-performance computing, future AI-based climate reconstruction frameworks are expected to achieve breakthroughs in global continuity and high resolution, offering more robust scientific support for understanding the evolution of the Earth's climate system.

Author contributions

CO: conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing (original draft preparation; review and editing). QL: conceptualization, funding acquisition, methodology, project administration, resources, supervision, writing (review and editing). ZL: resources, validation. SW: resources, validation.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Earth System Science Data.

Financial support

This research has been supported by the National Natural Science Foundation of China (grant no. 42375022) and the National Key Research and Development Program of China (grant no. 2023YFC3008002).

References

Beckers, J. M. and Rixen, M.: EOF calculations and data filling from incomplete oceanographic datasets, J. Atmos. Oceanic Technol., 20, 1839–1856, https://doi.org/10.1175/1520-0426(2003)020<1839:ECADEF>2.0.CO:2, 2003.

Bochow, N., Poltronieri, A., Rypdal, M., and Boers, N.: Reconstructing historical climate fields with deep learning, Sci. Adv., 11, eadp0558, https://doi.org/10.1126/sciadv.adp0558, 2025.

Bromwich, D., Wang, S. H., Zou, X., and Ensign, A.: An updated reconstruction of Antarctic near-surface air temperatures at monthly intervals since 1958, Earth Syst. Sci. Data, 17, 2953–2962, https://doi.org/10.5194/essd-17-2953-2025, 2025a.

Bromwich, D., Wang, S. H., and Nicolas, J. P.: Reconstructed Byrd temperature record (1957–2022), Polar Meteorology Group [data set], https://polarmet.osu.edu/datasets/Byrd_recon/, last access: 23 July 2025, 2025b.

Clem, K. R., Fogt, R. L., Turner, J., et al.: Record warming at the South Pole during the past three decades, Nat. Clim. Change, 10, 762–770, https://doi.org/10.1038/s41558-020-0815-z, 2020.

Cowtan, K. and Way, R. G.: Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends, Q. J. Roy. Meteor. Soc., 140, 1935–1944, https://doi.org/10.1002/qj.2297, 2014.

Cowtan, K., Hausfather, Z., Hawkins, E., Jacobs, P., Mann, M. E., Miller, S. K., Steinman, B. A., Stolpe, M. B., and Way, R. G.: Robust comparison of climate models with observations using blended land air and ocean sea surface temperatures, Geophys. Res. Lett., 42, 6526–6534, https://doi.org/10.1002/2015GL064888, 2015.

Darelius, E., Fer, I., and Nicholls, K.: Observed vulnerability of the Filchner–Ronne Ice Shelf to wind-driven inflow of warm deep water, Nat. Commun., 7, 12300, https://doi.org/10.1038/ncomms12300, 2016.

Esteves, C., Slotine, J. J., and Makadia, A.: Scaling spherical CNNs, Proc. 40th Int. Conf. Mach. Learn., 202, 9396–9411, https://doi.org/10.48550/arXiv.2306.05420, 2023.

Fang, S. W., Timmreck, C., Jungclaus, J., and Krüger, K.: On the additivity of climate responses to the volcanic and solar forcing in the early 19th century, Earth Syst. Dynam., 13, 1535–1555, https://doi.org/10.5194/esd-13-1535-2022, 2022.

Garza-Girón, R. and Tulaczyk, S. M.: Brief communication: Significant biases in ERA5 output for the McMurdo Dry Valleys region, Antarctica, The Cryosphere, 18, 1207–1213, https://doi.org/10.5194/tc-18-1207-2024, 2024.

Ham, Y., Kim, J., and Luo, J.: Deep learning for multi-year ENSO forecasts, Nature, 573, 568–572, https://doi.org/10.1038/s41586-019-1559-7, 2019.

Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, Rev. Geophys., 48, 4, https://doi.org/10.1029/2010RG000345, 2010.

Huang, B., Thorne, P. W., Smith, T. M., Liu W., Lawrimore J., Banzon, V. F., Zhang, H. M., Peterson, T. C., and Menne M.: Further exploring and quantifying uncertainties for Extended Reconstructed Sea Surface Temperature (ERSST) version 4 (v4), J. Clim., 29, 3119–3142, https://doi.org/10.1175/JCLI-D-15-0430.1, 2015.

Huang, B., Menne, M. J., Boyer, T., Freeman, E., Gleason, B. E., Lawrimore, J. H., Liu, C., Rennie, J. J., Schreck, C. J., Sun, F., Vose, R., Williams, C. N., Yin, X., and Zhang, H. M.: Uncertainty estimates for sea surface temperature and land surface air temperature in NOAA GlobalTemp version 5, J. Clim., 33, 1351–1379, https://doi.org/10.1175/JCLI-D-19-0395.1, 2020.

Huang, B., Yin, X., Boyer, T., Liu, C., Menne, M., Rao, Y. D., Smith, T., Vose, R., and Zhang, H.: Extended Reconstructed Sea Surface Temperature, Version 6 (ERSSTv6). Part I: An artificial neural network approach, J. Clim., 38, 1105–1121, https://doi.org/10.1175/JCLI-D-23-0707.1, 2025a.

Huang, B., Yin, X., Boyer, T., Liu, C., Menne, M., Rao, Y. D., Smith, T., Vose, R., and Zhang, H.: Extended Reconstructed Sea Surface Temperature, Version 6 (ERSSTv6). Part II: Upgrades on quality control and large-scale filter, J. Clim., 38, 1123–1136, https://doi.org/10.1175/JCLI-D-24-0185.1, 2025b.

Huang, J., Zhang, X., Zhang, Q., Lin, Y., Hao, M., Luo, Y., Zhao, Z., Yao, Y., Chen, X., Wang, L., Nie, S., Yin, Y., Xu, Y., and Zhang, J.: Recently amplified Arctic warming has contributed to a continual global warming trend, Nat. Clim. Change, 7, 875–879, https://doi.org/10.1038/s41558-017-0009-5, 2017.

IPCC: Climate Change 2013: The Physical Science Basis, Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., https://doi.org/10.1017/CBO9781107415324, 2013.

IPCC: Climate Change 2021: The Physical Science Basis, Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp., https://doi.org/10.1017/9781009157896, 2021.

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., and Saynisch-Wagner, J.: Towards neural Earth system modelling by integrating artificial intelligence in Earth system science, Nat. Mach. Intell., 3, 667–674, https://doi.org/10.1038/s42256-021-00374-3, 2021.

Jiao, B., Su, Y., Li, Q., Manara, V., and Wild, M.: An integrated and homogenized global surface solar radiation dataset and its reconstruction based on a convolutional neural network approach, Earth Syst. Sci. Data, 15, 4519–4535, https://doi.org/10.5194/essd-15-4519-2023, 2023.

Jones, P. D.: The reliability of global and hemispheric surface temperature records, Adv. Atmos. Sci., 33, 269–282, https://doi.org/10.1007/s00376-015-5194-4, 2016.

Jones, P. D., Harpham, C., and Vinther, B. M.: Winter-responding proxy temperature reconstructions and the North Atlantic Oscillation, J. Geophys. Res. Atmos., 119, 6279–7013, https://doi.org/10.1002/2014JD021561, 2014.

Kadow, C., Hall, D. M., and Ulbrich, U.: Artificial intelligence reconstructs missing climate information, Nat. Geosci., 13, 408–413, https://doi.org/10.1038/s41561-020-0582-5, 2020.

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., Peterson, T. C., Vose R. S., and Zhang H.: Possible artifacts of data biases in the recent global surface warming hiatus, Science, 348, 1469–1472, https://doi.org/10.1126/science.aaa5632, 2015.

Katz, R. W., Craigmile, P. F., Guttorp, P., Haran, M., Sansó, B., and Stein, M. L.: Uncertainty analysis in climate change assessments, Nat. Clim. Chang., 3, 769–771, https://doi.org/10.1038/nclimate1980, 2013.

Kennedy, J. J., Rayner, N. A., Atkinson, C. P., and Killick, R. E.: An ensemble data set of sea surface temperature change from 1850: the Met Office Hadley Centre HadSST.4.0.0.0 data set, J. Geophys. Res. Atmos., 124, 7719–7763, https://doi.org/10.1029/2018JD029867, 2019.

Kennicutt II, M. C., Bromwich, D., Liggett, D., et al.: Sustained Antarctic research: a 21st century imperative, One Earth, 1, 95–113, https://doi.org/10.1016/j.oneear.2019.08.014, 2019.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521, 436–444, https://doi.org/10.1038/nature14539, 2015

Lenssen, N. J. L., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., and Zyss, D.: Improvements in the GISTEMP uncertainty model, J. Geophys. Res. Atmos., 124, 6307–6326, https://doi.org/10.1029/2018JD029522, 2019.

Li, Q., Sheng, B., Huang, J., Chao, L., Song, Z., Chao, L., Sun, W., Yang, Y., Jiao, B., Guo, Z., Liao, L., Li, X., Sun, C., Li, W., Huang, B., Dong, W., and Jones, P.: Different climate response persistence causes warming trend unevenness at continental scales, Nat. Clim. Chang., 12, 343–349, https://doi.org/10.1038/s41558-022-01313-9, 2022.

Li, Q., Sun, W., Huang, B., Dong, W., Wang, X., Zhai, P., and Jones, P.: Consistency of global warming trends strengthened since 1880s, Sci. Bull., 65, 1709–1712, https://doi.org/10.1016/j.scib.2020.06.009, 2020.

Li, Q., Sun, W., Yun, X., Huang, B., Dong, W., Wang, X., Zhai, P., and Jones, P.: An updated evaluation of the global mean land surface air temperature and surface temperature trends based on CLSAT and CMST, Clim. Dynam., 56, 635–650, https://doi.org/10.1007/s00382-020-05502-0, 2021.

Li, Z., et al.: Arctic warming trends and their uncertainties based on surface temperature reconstruction under different sea ice extent scenarios, Adv. Clim. Chang. Res., 14, 335–346, https://doi.org/10.1016/j.accre.2023.06.003, 2022.

Liu, G., Reda, F. A., Shih, K. J., et al.: Image inpainting for irregular holes using partial convolutions, Springer International Publishing, https://doi.org/10.48550/arXiv.1804.07723, 2018.

Lorenz, E. N.: Deterministic nonperiodic flow, in: The Theory of Chaotic Attractors, edited by: Hunt, B. R., Li, T. Y., Kennedy, J. A., and Nusse, H. E., Springer, New York, https://doi.org/10.1007/978-0-387-21830-4_2, 1963.

Ma, Z., Huang, J., Zhang, X., Luo, Y., Dou, T., and Ding, M.: Deep learning-based reconstruction of monthly Antarctic surface air temperatures from 1979 to 2023, Sci. Data, 12, 847, https://doi.org/10.1038/s41597-025-05175-6, 2025.

Marshall, G. J.: Trends in the Southern Annular Mode from observations and reanalyses, J. Climate, 16, 4134–4143, https://doi.org/10.1175/1520-0442(2003)016<4134:TITSAM>2.0.CO;2, 2003.

Marshall, G. J.: Half-century seasonal relationships between the Southern Annular Mode and Antarctic temperatures, Int. J. Climatol., 27, 373–383, https://doi.org/10.1002/joc.1407, 2006.

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., and Lawrimore, J. H.: The Global Historical Climatology Network Monthly Temperature Dataset, Version 4, J. Climate, 31, 9835–9854, https://doi.org/10.1175/JCLI-D-18-0094.1, 2018.

Météo-France: Dumont d'Urville station from Météo-France [data set], https://meteofrance.com/, last access: 23 July 2025.

Monaghan, A. J., Bromwich, D. H., Chapman, W., and Comiso, J. C.: Recent variability and trends of Antarctic near-surface temperature, J. Geophys. Res.-Atmos., 113, D04105, https://doi.org/10.1029/2007JD009094, 2008.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An updated assessment of near-surface temperature change from 1850: the HadCRUT5 dataset, J. Geophys. Res. Atmos., 126, e2019JD032361, https://doi.org/10.1029/2019JD032361, 2021.

Nazeri, K., Ng, E., Joseph, T., Qureshi, F. Z., and Ebrahimi, M.: EdgeConnect: Structure guided image inpainting using edge prediction, IEEE/CVF Int. Conf. Comput. Vis. Workshop, https://doi.org/10.1109/ICCVW.2019.00408, 2019.

Nicolas, J. P. and Bromwich, D. H.: New reconstruction of Antarctic near-surface temperatures: multidecadal trends and reliability of global reanalyses, J. Climate, 27, 8070–8093, https://doi.org/10.1175/JCLI-D-13-00733.1, 2014.

Nielsen, E. B., Katurji, M., Zawar-Reza, P., and Meyer, H.: Antarctic daily mesoscale air temperature dataset derived from MODIS land and ice surface temperature, Sci. Data, 10, 833, https://doi.org/10.1038/s41597-023-02720-z, 2023.

NIWA (National Institute for Water and Atmospheric Research): Scott Base station, NIWA [data set], https://cliflo.niwa.co.nz, last access: 23 July 2025.

Ouyang, C., Li, Q., Li, Z., and Wei, S: China global Artificial Intelligence Reconstructed Surface Temperature$_{20CR/CMIP6}$ (C-AIRST$_{R/M}$), figshare [data set], https://doi.org/10.6084/m9.figshare.30663797.v1, 2025.

Plésiat, É., Dunn, R. J. H., Donat, M. G., and Kadow, C.: Artificial intelligence reveals past climate extremes by reconstructing historical records, Nat. Commun., 15, 9191, https://doi.org/10.1038/s41467-024-53464-2, 2024.

Rayner, N. A., Parker, D. E., Horton, E. B., et al.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, J. Geophys. Res. Atmos., 108, D14, https://doi.org/10.1029/2002JD002670, 2003.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzier, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Rohde, R. A. and Hausfather, Z.: The Berkeley Earth Land/Ocean Temperature Record, Earth Syst. Sci. Data, 12, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020, 2020.

Rohde, R., Muller, R. A., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D., and Wickham, C.: A new estimate of the average Earth surface land temperature spanning 1753 to 2011, Geoinfor. Geostat.: An Overview, 1, 1–7, https://doi.org/10.4172/2327-4581.1000101, 2013.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention, MICCAI 2015, Lecture Notes in Computer Science, 9351, Springer Cham, https://doi.org/10.1007/978-3-319-24574-4_28, 2015.

Rüttgers, M., Lee, S., Jeon, S., and You, D.: Prediction of a typhoon track using a generative adversarial network and satellite images, Sci. Rep., 9, 6057, https://doi.org/10.1038/s41598-019-42339-y, 2019

Scambos, T. A., Campbell, G. G., Pope, A., Haran, T., Muto, A., Lazzara, M., Reijmer, C. H., and van den Broeke, M. R.: Ultralow surface temperatures in East Antarctica from satellite thermal infrared mapping: the coldest places on Earth, Geophys. Res. Lett., 45, 6124–6133, https://doi.org/10.1029/2018GL078133, 2018.

Sheehan, P. M. F., and Heywood, K. J.: Ross Ice Shelf frontal zone subjected to increasing melting by ocean surface waters, Sci. Adv., 10, eado6429, https://doi.org/10.1126/sciadv.ado6429, 2024.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., and et al.: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system, Q. J. R. Meteorol Soc., 145, 2876–2908, https://doi.org/10.1002/qj.3598, 2019.

South Pole Meteorology Office: Amundsen-Scott South Pole Station climatology data, 1957–present (ongoing), AMRDC Data Repository [data set], https://doi.org/10.48567/szgp-6h49, last access: 23 July 2025.

Sun, W., Li, Q., Huang, B., Dong, W., Wang, X., Zhai, P., and Phil Jones: The assessment of global surface temperature change from 1850s: the C-LSAT2.0 ensemble and the CMST-Interim datasets, Adv. Atmos. Sci., 38, 875–888, https://doi.org/10.1007/s00376-021-1012-3, 2021.

Sun, W., Yang, Y., Chao, L., Dong, W., Huang, B., Jones, P., and Li, Q.: Description of the China global Merged Surface Temperature version 2.0, Earth Syst. Sci. Data, 14, 1677–1693, https://doi.org/10.5194/essd-14-1677-2022, 2022.

Toms, B. A., Barnes, E. A., and Ebert-Uphoff, I.: Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability, J. Adv. Model. Earth Syst., 12, e2019MS002002, https://doi.org/10.1029/2019MS002002, 2020.

Trenberth, K. E., and Stepaniak, D. P.: Covariability of components of poleward atmospheric energy transports on seasonal and interannual timescales, J. Climate, 16, 3691–3705, https://doi.org/10.1175/1520-0442(2003)016<3691:COCOPA>2.0.CO;2, 2003.

Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., Lagun, V., Reid, P. A., and Iagovkina, S.: The SCAR READER Project: Toward a high-quality database of mean Antarctic meteorological observations, J. Climate, 17, 2890–2898, https://doi.org/10.1175/1520-0442(2004)017<2890:TSRPTA>2.0.CO;2, 2004.

Turner, J., Colwell, S. R., Marshall, G. J., Lachlan-Cope, T. A., Carleton, A. M., Jones, P. D., Lagun, V., Reid, P. A., and Svetlana, L.: Antarctic climate change during the last 50 years, Int. J. Climatol., 25, 279–294, https://doi.org/10.1002/joc.1130, 2005.

Turner, J., Marshall, G. J., Clem, K., et al.: Antarctic temperature variability and change from station data, Int. J. Climatol., 40, 2986–3007, https://doi.org/10.1002/joc.6378, 2019.

Vaughan, D. G., Marshall, G. J., Connolley, W. M., et al.: Recent rapid regional climate warming on the Antarctic Peninsula, Clim. Change, 60, 243–274, https://doi.org/10.1023/A:1026021217991, 2003.

Vose, R. S., Arndt, D., Banzon, V. F., et al.: NOAA's Merged Land Ocean Surface Temperature Analysis, Bull. Amer. Meteor. Soc., 93, 1677–1685, https://doi.org/10.1175/BAMS-D-11-00241.1, 2012.

Vose, R. S., Huang, B., Yin, X., Arndt, D., Easterling, D. R., Lawrimore, J. H., Menne, M. J., Sanchez Lugo, A., and Zhang, H. M.: Implementing full spatial coverage in NOAA's global temperature analysis, Geophys. Res. Lett., 48, e2020GL090873, https://doi.org/10.1029/2020GL090873, 2021.

Wang, K., and Clow, G. D.: Reconstructed global monthly land air temperature dataset (1880–2017), Geosci. Data J., 7, 4–12, https://doi.org/10.1002/gdj3.84, 2020.

Wang, R., Zheng, F., and Wang, H.: Influence of Winter Tasman Sea SST on the Antarctic Peninsula: A perspective from historical simulations, Adv. Atmos. Sci., 42, 1533–1547, https://doi.org/10.1007/s00376-024-4350-0, 2025.

Wang, S., Li, G., Zhang, Z., Zhang, W., Wang, X., Chen, D., Chen, W., and Ding, M.: Recent warming trends in Antarctica revealed by multiple reanalysis, Adv. Clim. Chang. Res., 16, 447–459, https://doi.org/10.1016/j.accre.2025.03.003, 2025.

Wang, Y., Zhang, X., Ning, W., Lazzara, M. A., Ding, M., Reijmer, C. H., et al.: The AntAWS dataset: A compilation of Antarctic automatic weather station observations, Earth Syst. Sci. Data, 15, 411–429, https://doi.org/10.5194/essd-15-411-2023, 2023.

Wei, S., Li, Q., Xu, Q., Li, Z., Zhang, H., and Lin, J.: Updates to C-LSAT2.1 and the development of high-resolution land surface air temperature and diurnal temperature range datasets, Earth Syst. Sci. Data, 17, 4985–5005, https://doi.org/10.5194/essd-17-4985-2025, 2025.

Xu, W., Li, Q., Jones, P. D., Wang, X. L., et al.: A new integrated and homogenized global monthly land surface air temperature dataset for the period since 1900, Clim. Dyn., 50, 2513–2536, https://doi.org/10.1007/s00382-017-3755-1, 2018.

Yang, Q., Zhao, C., Feng, J., Li, G., Gu, J., Xia, Z., Xu, M., and Yang, Z.: Comprehensive evaluation of iAMAS (v1.0) in simulating Antarctic meteorological fields with observations and reanalysis, Geosci. Model Dev., 18, 5373–5396, https://doi.org/10.5194/gmd-18-5373-2025, 2025.

Yin, X. G., Huang, B. Y., Menne, M., et al.: NOAAGlobalTemp Version 6: An AI-based global surface temperature dataset, Bull. Amer. Meteor. Soc., 105, E2184–E2193, https://doi.org/10.1175/BAMS-D-24-0012.1, 2024.

Yun, X., Huang, B., Cheng, J., Xu, W., Qiao, S., and Li, Q.: A new merge of global surface temperature datasets since the start of the 20th Century, Earth Syst. Sci. Data, 11, 1629–1643, https://doi.org/10.5194/essd-11-1629-2019, 2019.

Zhou, L., Liu, H., Jiang, X., Ziegler, A., Azorin-Molina, C., Liu, J., and Zeng, Z.: An artificial intelligence reconstruction of global gridded surface winds, Sci. Bull., 67, 2060–2063, https://doi.org/10.1016/j.scib.2022.09.022, 2022.