Expert Systems With Applications 322 (2026) 132259

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

SGD-SST 2.0: Seamless global daily sea surface temperature products cross-sensors generating from 2003 to 2025

Qi Wang [a], Qiang Zhang [a,*], Tongde Yang [a], Weizhen Sun [a], Qiangqiang Yuan [b]

[a] Information Science and Technology College, Dalian Maritime University, Dalian, 116026, China
[b] School of Geodesy and Geomatics, Wuhan University, Wuhan, 430079, China

ARTICLE INFO

Keywords:
Sea surface temperature
Multi-source products reconstruction
Bayesian optimization
Adaptive ConvLSTM
Validation

ABSTRACT

Sea Surface Temperature, a key parameter of the Earth system, suffers from severe data gaps due to cloud cover and sensor limitations. To address this problem, this study develops an adaptive Convolutional Long Short-Term Memory spatio-temporal reconstruction model via Bayesian optimization. This model generates the Seamless Global Daily SST products (SGD-SST 2.0). The products, which have a spatial resolution of 9 km, support cross-sensor coverage and cover the period from 2003 to 2025. The model develops a dynamic time-step selection mechanism driven by data missing rates. It also introduces an adaptive post-processing strategy that considers latitudinal effects. In particular, addressing the high missing rate characteristic of Moderate-Resolution Imaging Spectroradiometer data, a Visible Infrared Imaging Radiometer Suite pre-trained model transfer learning strategy is developed. It enables seamless reconstruction of cross-sensor multi-source products between Moderate-Resolution Imaging Spectroradiometer sea surface temperature (2003–2012) and Visible Infrared Imaging Radiometer Suite sea surface temperature (2013–2025), generating 22-year global daily sea surface temperature products. For validation experiments, four verification methods are employed: in-situ validation, time-series consistency analysis, comparison with other methods, and latitude-stratified comparison. Validation results demonstrate a Correlation Coefficient of 0.994, Root Mean Square Error of 0.916 K, and Mean Absolute Error of 0.602 K. Compared with the previous SGD-SST 1.0 products, SGD-SST 2.0 products show significantly improved reconstructing accuracy especially in mid-low latitude regions. SGD-SST 2.0 products are available for download at https://huggingface.co/datasets/namelesswang/SGD-SST_2.0/tree/main.

1. Introduction

Sea Surface Temperature (SST) is a key physical quantity characterizing energy exchange at the ocean-atmosphere interface. It is also a crucial element for revealing multi-scale processes in the global climate system (Minnett et al., 2019). SST directly regulates the exchange of heat, momentum, and gas fluxes between the ocean and atmosphere. It profoundly influences atmospheric circulation, ocean thermal structure, and ecosystems. From a societal perspective, spatio-temporal variations of SST drive extreme weather events, long-term climate trends, and regional environmental phenomena. Therefore, accurate monitoring of SST holds significant importance for weather prediction, ecological assessment, fisheries management, and other fields (O'carroll et al., 2019).

Generally, SST data are primarily obtained through two ways: in-situ measurements, and satellite remote sensing retrieval. Satellite remote sensing technology has been operationally implemented since 1981. Owing to its advantages of high spatial and temporal resolution, it has become a core mode for acquiring global ocean temperature distribution characteristics (Haghbin et al., 2021; Zhang et al., 2024b). It complements in-situ measurements in terms of accuracy and regional coverage. However, both techniques have inherent limitations. In-situ measure-

List of Abbreviations: Abbreviation, Full Name; CC, Correlation Coefficient; CIs, Confidence Intervals; DINEOF, Data Interpolation Empirical Orthogonal Function; GAN, Generative Adversarial Network; KD-tree, K-Dimensional Tree; LSTM, Long Short-Term Memory; LWIR, Long-Wave Infrared; MAE, Mean Absolute Error; ME, Mean Error (Bias); MODIS, Moderate-Resolution Imaging Spectroradiometer; MWIR, Mid-Wave Infrared; NOAA, National Oceanic and Atmospheric Administration; R^2, Coefficient of Determination; RMSE, Root Mean Square Error; SGD-SST 1.0, Seamless Global Daily SST Products 1.0; SGD-SST 2.0, Seamless Global Daily SST Products 2.0; SNPP-VIIRS, Visible Infrared Imaging Radiometer Suite (VIIRS) aboard the Suomi National Polar-Orbiting Partnership (SNPP); SSMA, Spatial-Spectral Multi-Head Attention; SST, Sea Surface Temperature; TIR, Thermal Infrared; VIIRS, Visible Infrared Imaging Radiometer Suite; VE-DINEOF, Variable Optimal EOF Improved Algorithm; iQuam, In-Situ SST Quality Monitor.

* Corresponding author.
E-mail address: qzhang95@dlmu.edu.cn (Q. Zhang).

https://doi.org/10.1016/j.eswa.2026.132259
Received 5 January 2026; Received in revised form 22 March 2026; Accepted 25 March 2026
Available online 6 April 2026

ments are constrained by the density of observation platforms, resulting in sparse spatio-temporal coverage. Although satellite retrieval enables large-scale monitoring, it suffers from significant data gaps caused by cloud obstruction, aerosol interference, and solar glint effects. These issues have become the bottlenecks restricting its applications in multiple fields (Kolbe et al., 2025; Luo et al., 2021; Nielsen-Englyst et al., 2023; Zhang et al., 2020).

SST serves as a core parameter for ocean-atmosphere interactions. Its high-precision reconstruction and prediction form a critical foundation for climate monitoring, ecological assessment, and disaster early warning. In recent years, relevant works have developed two main categories of methods: physics-driven approaches and data-driven approaches. These approaches are described as follows:

In terms of physics-driven approaches, HYCOM represents an advanced data assimilation system based on physical mechanisms. Its core advantage lies in combining hybrid coordinate modeling with multi-source data fusion. This provides a physically constrained framework for dynamic repair of missing satellite SST data (Gao, 2008). Although HYCOM has achieved breakthroughs in data assimilation accuracy, it relies on traditional interpolation methods. These methods still result in systematic biases in extreme data missing scenarios. Additionally, computational costs grow exponentially with increasing resolution. Statistical models are represented by Data Interpolation Empirical Orthogonal Function (DINEOF). DINEOF achieves missing data repair through spatio-temporal feature decomposition. It operates without prior assumptions, becoming one of the mainstream methods for spatio-temporal missing data reconstruction. Ping et al. (2016) addressed the efficiency and accuracy bottlenecks of the traditional DINEOF algorithm in high missing rate scenarios. A Variable Optimal EOF Improved Algorithm (VE-DINEOF) was proposed. By dynamically selecting spatio-temporal feature modes, it significantly improves the performance of missing data reconstruction in complex marine environments. However, although VE-DINEOF shows excellent performance in sea data, it relies on pre-setting maximum EOF numbers and sub-region divisions. This results in parameter sensitivity issues (Shi et al., 2024; Sukresno, 2010; Zhang et al., 2024a).

In terms of data-driven approaches, Long Short-Term Memory (LSTM) neural network has become a research hotspot in recent years due to its efficient modeling capability for time-series data. For instance, Jia et al. (2022) constructed a 5-day short-term SST prediction model based on LSTM for the complex dynamic environment of the East China Sea. Through sensitivity experiments and regional transfer analysis, this study provides a new reference method for SST prediction in small and medium-scale sea areas. LSTM demonstrates efficient capture capability for seasonal cycles and short-term trends in single-point time-series modeling. However, its main limitation lies in focusing solely on temporal dimension feature extraction. It insufficiently utilizes the spatial correlation of SST and the spatial propagation mechanisms of dynamic processes (Koner et al., 2016; Xiao et al., 2019; Zhang et al., 2022).

To address the issues of LSTM, Hirahara et al. (2021) proposed a generative adversarial network (GAN) incorporating adversarial physical model loss, to address the cloud obstruction problem in satellite SST data. This method fuses physical constraints from data assimilation with efficient reconstruction capabilities of deep learning. It achieves high-precision for SST restoration, while avoiding the high computational cost of traditional data assimilation methods. However, this method is limited by a fixed 3-day time window, capturing only short-term temporal features. Additionally, this study focuses on local region reconstruction at 64×64 pixels, targeting only small-scale scenarios. It exhibits limitations in handling spatio-temporal continuous features of global-scale SST fields (Gavahi et al., 2023; Marullo et al., 2014).

To provide a structured synthesis of the SST reconstruction methods discussed above, their core characteristics, advantages, and limitations are summarized in Table 1.

Although the aforementioned methods demonstrate advantages in specific scenarios, high-precision global-scale daily SST reconstruction still faces three main challenges:

1) Difficulty in eliminating interpolation biases in extreme missing scenarios;
2) Existing spatio-temporal modeling fragmentation issues (e.g., LSTM focuses solely on single-point time-series, while GAN is limited to local region restoration);
3) Long-term products suffer from sequence discontinuity caused by sensor replacement.

Therefore, three key objectives must be achieved in complex marine dynamic environments: spatio-temporal modeling adaptability, long-term sequence continuity, and multi-dimensional accuracy verification. How to address these issues has become a key breakthrough point in current research. To solve the above problems, this study proposes the following innovative solutions:

- An adaptive ConvLSTM spatio-temporal reconstruction model based on Bayesian optimization is developed. This model incorporates a dynamic time-step selection mechanism driven by data missing rates, and introduces an adaptive post-processing strategy that accounts for latitudinal effects. These innovations not only significantly improve the reconstruction accuracy in regions with high data missing rates, but also effectively preserve key oceanographic features (e.g., ocean fronts).
- Cross-sensor multi-source data reconstruction is realized to build seamless global daily SST products (named SGD-SST 2.0), which cover 2003-2025 with a 9 km spatial resolution by integrating MODIS SST (2003-2012) and VIIRS SST (2013-2025) datasets. To address the high missing rate of MODIS SST data, a cross-sensor strategy adaptation strategy using a VIIRS pre-trained model is developed, extending the reconstruction period to 22 years.
- A multi-dimensional validation strategy is adopted to evaluate the performance of SGD-SST 2.0 products, including four aspects: (a) in-situ validation, (b) time-series consistency analysis, (c) comparison with the DINEOF method, and (d) latitude-stratified comparison. Validation results show that, compared with the previous SGD-SST 1.0 products, SGD-SST 2.0 exhibits significantly improved reconstruction accuracy in mid-low latitude regions. Additionally, it demonstrates better spatio-temporal continuity than traditional interpolation methods.

The remainder of this paper is organized as follows. Section 2 introduces the used data. Section 3 describes the reconstruction model and validation approaches. Section 4 presents the experimental results and validation comparison outcomes. Section 5 provides the conclusions and future prospects.

2. Data

2.1. MODIS Sea surface temperature products

The Moderate-Resolution Imaging Spectroradiometer (MODIS), a core payload of NASA's Earth Observing System, is onboard the Terra (launched December 18, 1999) and Aqua (launched May 4, 2002) satellites. It achieves high-frequency global sea surface coverage through complementary morning (10:30 local time) and afternoon (1:30 local time) orbits. The sensor has 36 bands (0.4–14.4μm). Primary bands for SST retrieval include mid-wave infrared (MWIR) band-20 (3.75μm), band-22 (3.959μm), band-23 (4.040μm), and long-wave infrared (LWIR) band-31 (11.03μm), band-32 (12.02μm). MWIR bands excel in night-time SST retrieval without solar interference, benefiting

Table 1
Summary of mainstream SST reconstruction methods.

Category	Method	Core Idea	Advantages	Limitations
Physics-driven	**HYCOM**	Hybrid coordinate modeling with multi-source data assimilation	Physically constrained framework; breakthrough in assimilation accuracy	Relies on traditional interpolation; systematic bias under extreme missing data; computational cost grows exponentially with resolution
	DINEOF	Spatio-temporal missing data repair via EOF decomposition	No prior assumptions; mainstream method for spatio-temporal reconstruction	Poor performance in high missing-rate scenarios
	VE-DINEOF	Dynamic selection of spatio-temporal feature modes	Significantly improves reconstruction in complex marine environments	Requires pre-setting maximum EOF number and sub-region divisions; parameter sensitivity
Data-driven	**LSTM**	Time-series modeling for SST prediction	Efficient capture of seasonal cycles and short-term trends	Focuses only on temporal dimension; insufficient utilization of spatial correlation and spatial propagation mechanisms
	GAN	Generative adversarial network with adversarial physical model loss	Fuses physical constraints with deep learning; high-precision restoration; avoids high computational cost	Limited by fixed 3-day time window; targets only small-scale (64 × 64 pixels) scenarios; limitations in global-scale spatio-temporal continuous features

from high atmospheric transparency and temperature sensitivity. However, they are susceptible to sensor noise. LWIR bands support day-night observations via mature split-window algorithms, while their accuracy is significantly affected by atmospheric water vapor absorption and scan angle variations (Kilpatrick et al., 2019; Liu & Minnett, 2016).

The missing data issue of MODIS SST products primarily stems from the combined effects of cloud obstruction and sensor performance limitations. Infrared bands' reliance on clear-sky conditions results in substantial observation gaps. After quality control, the global annual average of valid data approximately accounts for 50%–70%, as shown in Fig. 1. In complex scenarios such as equatorial convective regions and polar winters, the proportion of unprocessed pixels due to cloud obstruction significantly increases (Jia & Minnett, 2020; Kang et al., 2014).

Furthermore, cloud detection algorithms (e.g., threshold methods based on brightness temperature differences and texture analysis) exhibit a misclassification rate of approximately 5%–10%. This may incorrectly categorize some clear-sky pixels as cloudy, leading to data gaps or low-quality flags. In complex scenarios including polar winters and thin cloud cover, misclassification rates may locally increase with scene complexity due to reduced spectral contrast and sensor limitations.

Regarding sensor performance, the MWIR bands' insufficient penetration capability for low clouds at night exacerbates observation blind spots in high-latitude winters. Detector noise and early calibration biases have undergone on-orbit corrections. However, complex atmospheric conditions such as high-water vapor content and strong temperature gradients may still introduce systematic errors of certain degrees. These factors all affect the data continuity and reliability of SST products (Frey et al., 2020; Kilpatrick et al., 2015), as shown in Fig. 1c.

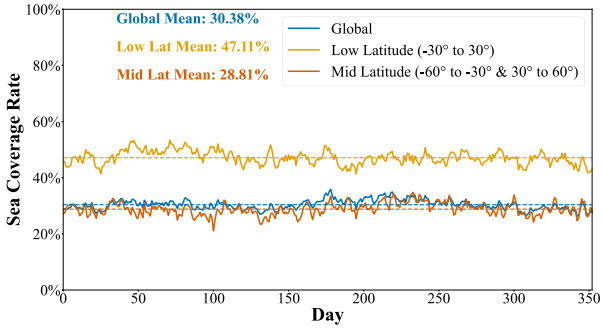
This study reconstructs missing SST data from MODIS Aqua covering the period 2003–2012. The data missing rate for the 2003 dataset is illustrated in Fig. 1a. The MODIS SST products can be downloaded from: https://oceandata.sci.gsfc.nasa.gov/l3/.

2.2. VIIRS Sea surface temperature products

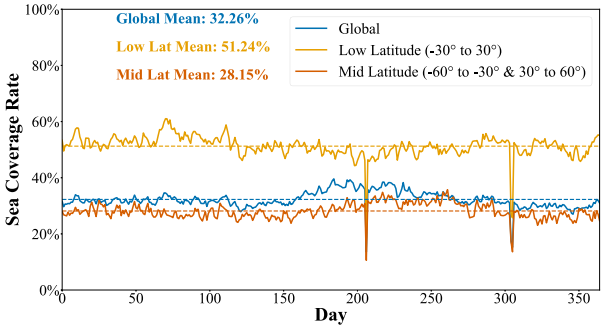
The Suomi NPP satellite, launched on October 28, 2011, serves as the core satellite of the U.S. National Polar-orbiting Partnership program. It operates in a sun-synchronous near-polar orbit with an altitude of 824 km. The satellite daily crosses the equator at 13:30 local time, featuring a 16-day revisit cycle and a 3000 km swath width to achieve gap-free global coverage. The Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the satellite is a key payload for SST retrieval. This sensor inherits technical advantages from MODIS and AVHRR, configured with 22 bands covering the 0.4–12.5μm range. Primary bands for SST retrieval include mid-wave infrared (MWIR) bands M12 (3.61–3.79μm), M13 (3.973–4.128μm), and thermal infrared (TIR) bands M15 (10.263–11.263μm), M16 (11.538–12.488μm). The sensor provides a 750 × 750 m spatial resolution and effectively reduces pixel distortion at scan edges through pixel aggregation technology (Choo et al., 2025; Jiang et al., 2024; Liu & Wang, 2018; Zhang et al., 2025).

However, the data integrity of VIIRS is constrained by multiple factors. As an infrared sensor, it is highly sensitive to cloud coverage. It relies on the Alternating Decision Tree cloud screening algorithm to exclude contaminated pixels. This algorithm improves effective data coverage by 3%–6% compared to traditional binary decision trees. Nevertheless, in cloud-intensive regions such as the Intertropical Convergence Zone and mid-high latitude storm areas, data gaps still occur due to extensive cloud pixel removal. For instance, certain regions in the equatorial Pacific and western Indian Ocean exhibit extremely sparse buoy data matches. This is mainly caused by persistent heavy cloud coverage. During daytime retrieval, mid-wave infrared bands cannot be used due to solar reflection interference. The retrieval process relies solely on thermal infrared dual bands (M15, M16), which results in slightly lower accuracy compared to nighttime three-band algorithms (the nighttime algorithms combine mid-wave infrared bands for improved performance). Atmospheric complexity further exacerbates data quality challenges: regions affected by Saharan dust include the West African coast and Arabian Sea, where atmospheric aerosols significantly absorb and scatter infrared radiation, causing cold biases in SST retrieval. Quality flag systems screen contaminated pixels; however, high-concentration dust regions may still be misclassified as clear sky, introducing systematic errors (Tu et al., 2015; Zhang et al., 2026).

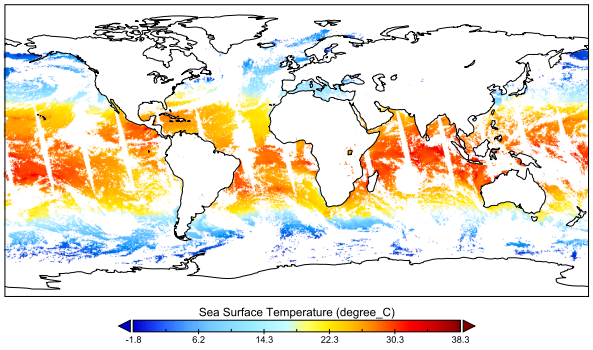
Overall, SNPP-VIIRS provides important data support for global SST retrieval through multi-band infrared radiation measurements. However, its data quality exhibits significant gaps or accuracy degradation in regions such as tropical oceans, polar seas, and shallow coastal zones,



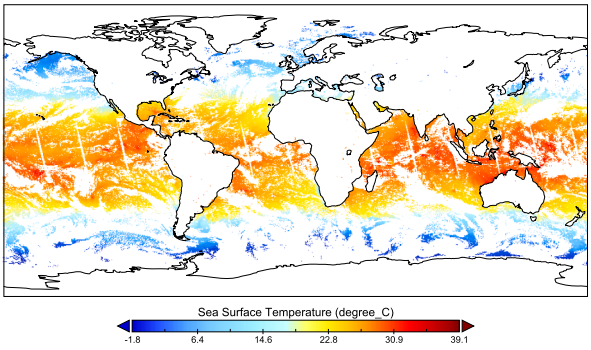
(a) Global SST coverage rate in 2003 (MODIS)



(b) Global SST coverage rate in 2023 (VIIRS)



(c) Global SST in 2003.3.11 (MODIS)



(d) Global SST in 2023.3.11 (VIIRS)

Fig. 1. Global SST coverage rate and data gaps in MODIS and VIIRS Products.

as shown in Fig. 1d. These issues result from the combined effects of cloud coverage, aerosols, complex surface conditions, and instrument noise. This study reconstructs missing SST data from SNPP-VIIRS covering the period 2013–2025. The data missing rate for the 2023 dataset is illustrated in Fig. 1b. The VIIRS SST products can be downloaded from: https://oceandata.sci.gsfc.nasa.gov/l3/.

2.3. Global sea surface temperature station data

The In-situ SST Quality Monitor (iQuam) system, developed by the U.S. National Oceanic and Atmospheric Administration (NOAA), serves as an in-situ measurement data platform supporting satellite SST products calibration. It integrates multi-source data including Global Telecommunication System, Argo buoys, and ship observations. This integration constructs a high-quality dataset covering multiple platforms such as drifting buoys, moored buoys, and ships. The system implements strict quality control through binary checks and statistical methods. It significantly removes outliers while retaining high-precision observations. Data quality levels are indicated through the quality-level field. Specifically, 0–2 indicate data of excessively poor quality for use, while 3–5 respectively represent low quality, acceptable quality, and optimal quality data (Ahmed et al., 2025; Good et al., 2020; Xu & Ignatov, 2014).

This study utilizes iQuam data with quality level 5 as the benchmark. It validates the accuracy of reconstructed SST products through spatio-temporal matching and statistical analysis. It also evaluates error characteristics in complex environments, providing a basis for algorithm optimization and model improvement. The standardized quality control and reliable observations of iQuam offer critical support for quantifying the reliability of reconstructed SST products, ensuring effective applications in climate research and numerical simulations.

3. Methodology

3.1. Reconstruction method

The overall flowchart of the reconstruction method is shown in Fig. 2. This method includes the following steps: First, perform data preprocessing-screen pure ocean regions from the original data, select samples with data missing rates meeting requirements within the time window, simulate missing data using masks, and generate data pairs to construct a sample dataset. Second, define the parameter space: generate parameters iteratively through Bayesian optimization and configure ConvLSTM with adaptive time steps. Third, input the training set into the configured model, train through a joint loss function fused with monthly average constraints, generate reconstructed SST products combined with latitude-adaptive median filtering post-processing, and perform validation. Finally, generate seamless global daily SST products from 2003–2025 based on the optimized model.

3.1.1. Sample dataset construction

To address the spatio-temporal characteristics and reconstruction requirements of SST data, this study constructs a sample dataset with dynamic time windows. This adapts to the adaptive adjustment mechanism of 9–13 day time steps in the model. The original data adopts VIIRS SST data with 9 km resolution. Considering that MODIS SST data has few complete samples due to high missing rates, this experiment prioritizes VIIRS data for sample dataset construction. Preliminary pre-experiments show that pre-trained models based on MODIS high-missing rate data poorly perform in SST reconstruction. In contrast, the high integrity of VIIRS data provides a more reliable spatio-temporal feature learning foundation for the model.

To ensure seasonal coverage and spatio-temporal continuity of samples, the screening process incorporates both spatial and temporal cri-

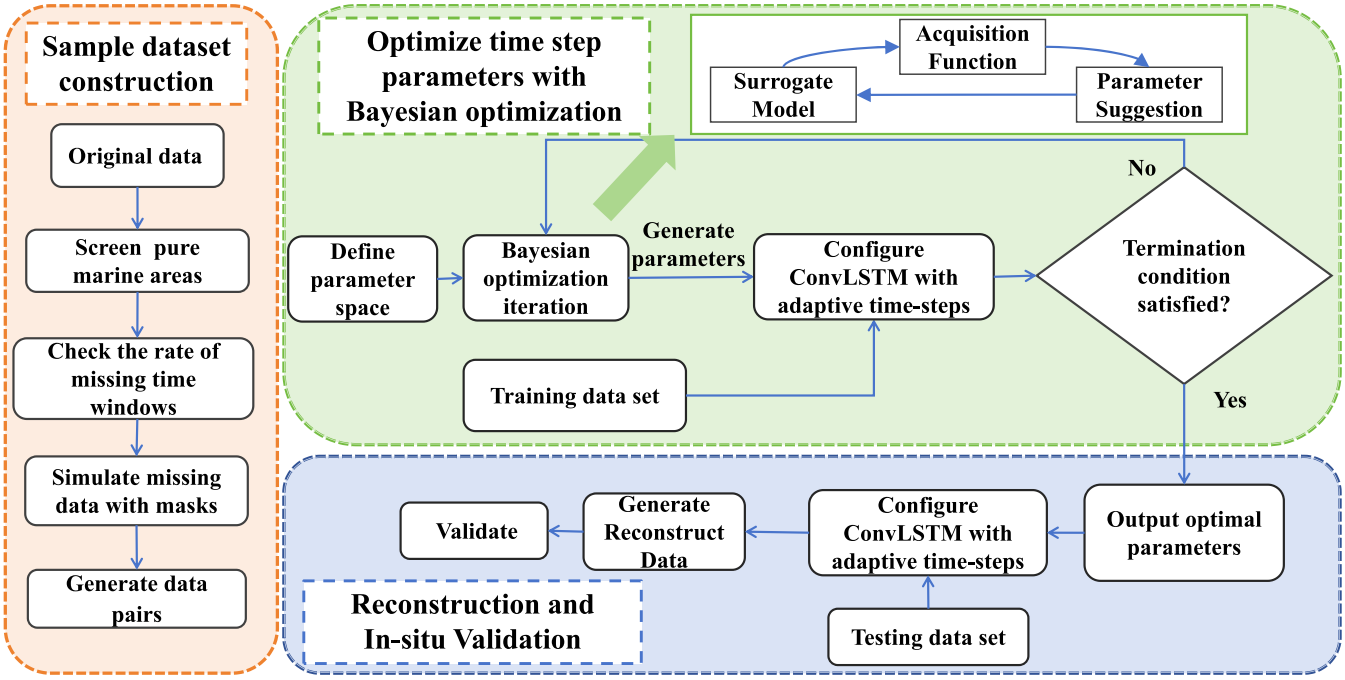


Fig. 2. Flowchart of the SGD-SST 2.0 reconstruction method.

teria: Spatially, only pure ocean pixel blocks (with 0 land pixels) are retained to avoid land interference. Temporally, centered on target date T, the missing rate of ocean pixels must be less than 10% within 13 days from $T-12$ to T. This forms long-term continuous valid label samples.

In the data pair generation stage, the basic time range for each sample is set as $T-12$ to T. Continuous sub-windows of 9–13 days are dynamically extracted according to the model's adaptive time steps. Examples include 13 days from $T-12$ to T and 9 days from $T-8$ to T. The specific process is given as follows: First, extract SST data and corresponding ocean masks from $T-12$ to T, and construct a 3D input group containing time, height, and width dimensions; the mask is used to mark valid observation areas for each time step. Subsequently, select 8250 simulated masks of 40×40 from the VIIRS mask library (2013–2024) with missing rates ranging from 0.3–0.8. Randomly apply these masks to the complete block of date T to generate incomplete input data, and simultaneously update the mask group for the corresponding time window to ensure spatio-temporal consistency between input data and masks (Wang et al., 2025; Young et al., 2024; Zhang et al., 2021).

The finally constructed dataset 9272 sample blocks. Each sample consists of dynamically extracted 9–13 days incomplete SST blocks (corresponding to continuous sub-windows within the $T-12$ to T base time range) and their labels (complete blocks of date T). Each sample constitutes a spatio-temporal cube of 40×40 pixels × 9–13 days, which carries high spatial-temporal information density to support model learning. Moreover, the 8250 simulated missing masks (missing rates: 0.3–0.8) applied to each complete sample form a large-scale targeted data augmentation strategy, enabling the model to capture generalizable features for various missing patterns and ensuring the sufficiency of the 9272 samples for the global 22-year reconstruction task. Through long-time window screening and dynamic sub-window interception, this dataset directly adapts to the time step adjustment logic of the model. Meanwhile, it introduces diverse missing patterns through simulated masks, providing training materials highly matching real scenarios for the deep spatio-temporal reconstruction model. This effectively improves the reconstruction performance of the model under complex missing conditions.

Although all 9272 training samples are derived from 2021, this dataset covers most typical spatiotemporal patterns of SST in the study region, including both normal SST variation areas and regions with local anomalous changes. The core goal of SST missing value reconstruction is to learn the universal spatiotemporal dependencies and variation laws of SST, rather than memorizing year-specific climate signals. The 2021 samples adequately cover the statistical characteristics of SST in both normal and anomalous states (e.g., mean, variance, spatial autocorrelation). The model's reconstruction performance on out-of-sample test data with an independent climate background (e.g., 2023 SST data) is consistent with that on the training dataset, indicating that the model has learned generalizable reconstruction capabilities rather than year-specific features. Thus, the single-year training data does not introduce significant temporal bias. Furthermore, while 2021 does not exhibit extreme global climate anomalies such as strong El Niño/La Niña events, the local and moderate-intensity SST anomalies included in the training set ensure that the model can adapt to most climate scenarios in other years. For extreme climate years, the fundamental spatiotemporal variation mechanism of SST remains unchanged, and the model's ability to capture temperature gradients, spatial continuity, and contextual dependencies can still be effectively applied to guarantee reliable reconstruction performance.

Regarding the dataset partition, this study follows a "training-application-external validation" paradigm. The training set is composed of all 9272 sample blocks derived from the 2021 VIIRS SST data, which serve as high-quality labels paired with simulated missing masks. After training, the model is directly applied to two completely independent sequences for testing: the 2003–2012 MODIS SST data (cross-sensor test) and the 2013–2025 VIIRS SST data (cross-time test). The final validation relies on the performance of the full 22-year reconstructed SST products, which are rigorously evaluated against the independent iQuam in-situ observation database. All quantitative metrics reported in this manuscript are obtained from this strict external validation, which verifies the reliability and generalization of the proposed model and training strategy.

3.1.2. Bayesian optimization-based adaptive ConvLSTM spatio-temporal reconstruction model

1) **Adaptive Time Step Adjustment Mechanism**

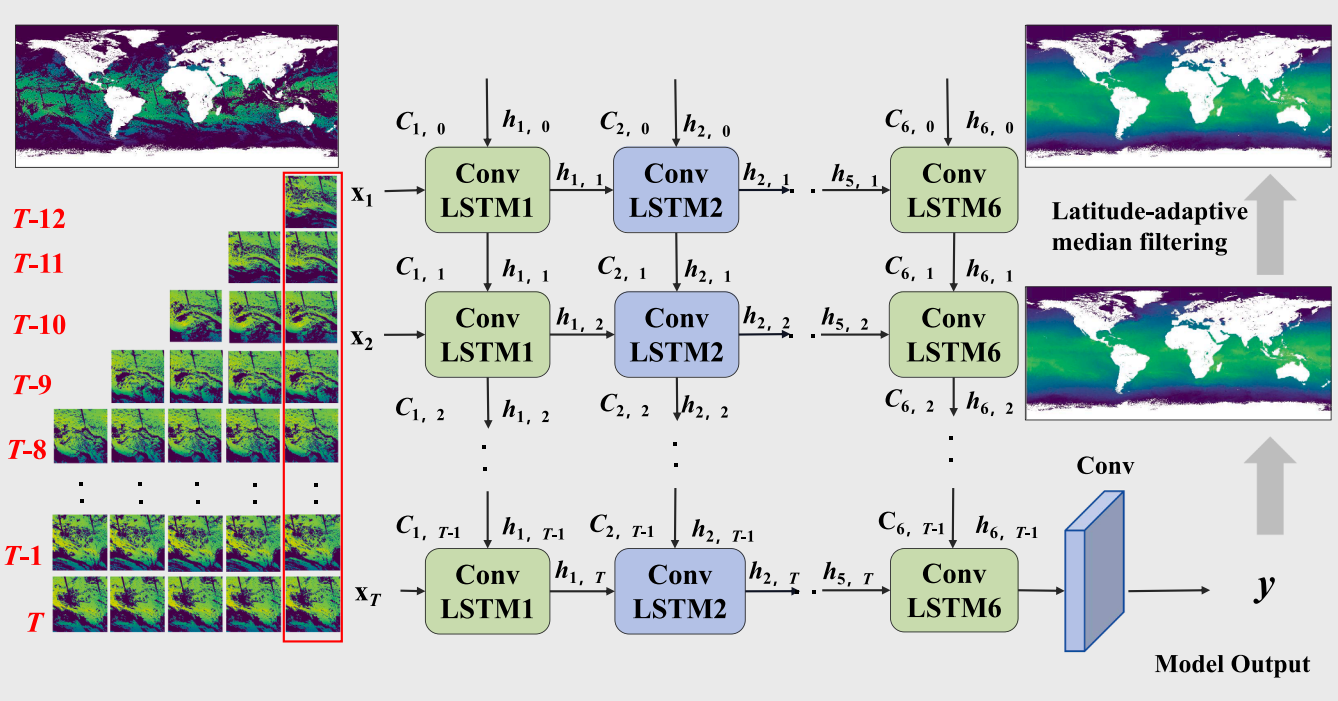


Fig. 3. Bayesian optimization-based adaptive ConvLSTM spatio-temporal reconstruction model.

To enhance the model's adaptability to dynamic features of spatio-temporal data, the Smooth Time Step Selector module is designed in Fig. 3. It dynamically adjusts the time step length input to ConvLSTM based on data missing rates. This module achieves smooth transitions using the Sigmoid function, avoiding discrete jumping issues of traditional fixed threshold methods. Specifically, the module receives data missing rate r as input. It maps r to the real number domain through linear transformation:

$$x = (r - \theta) \cdot s \tag{1}$$

where θ is the threshold and s is the slope parameter. After generating weight w via the Sigmoid function $w = \frac{1}{1+e^{-x}}$, the time step length is calculated by Eq. 2:

$$t = t_{\min} + (t_{\max} - t_{\min}) \cdot w \tag{2}$$

This enables continuous adjustment within the interval $[t_{\min}, t_{\max}]$ (during training, $t_{\min} = 9$ and $t_{\max} = 13$; parameters remain consistent during testing). This mechanism allows the model to dynamically adapt to scenarios with different data integrity levels: When the missing rate is high, expand the time window to capture long-term temporal dependencies. When data integrity is high, shrink the window to focus on local spatio-temporal features, balancing computational efficiency and feature extraction capability.

2) **Bayesian Optimization Strategy**

Considering the data processing differences between training and testing phases, a Bayesian optimization strategy is adopted to independently optimize the model's key parameters (threshold θ, slope s) as below (Frazier, 2018; Pelikan, 2005; Snoek et al., 2012; Wang et al., 2023).

Data block division and missing rate characteristics: In the training phase, the global 2160×4320 data is divided into 40×40 small blocks. Missing values are artificially introduced through simulated masks with a relatively low missing rate range focusing on the $0.1 \sim 0.2$ interval. This controls the complexity of training data. In the testing phase, 540×540 data blocks under real scenarios are directly processed. The missing rates are affected by actual observation conditions, showing a wide distribution range mainly concentrated in the $0.5 \sim 0.9$ interval.

Training Phase: Threshold $\theta_{\text{train}} \in [0.1, 0.2]$ and slope $s_{\text{train}} \in [10, 100]$. For low missing rate scenarios, a larger slope enhances function response sensitivity. This ensures the model accurately captures local spatio-temporal features when data is relatively complete.

Testing Phase: Threshold $\theta_{\text{test}} \in [0.5, 0.9]$ and slope $s_{\text{test}} \in [1, 20]$. For high missing rate scenarios, a smaller slope slows the function change rate. This allows the model to stably adjust the time window during large-scale missing, avoiding overfitting caused by sparse data.

Optimization Implementation Details: The objective function is guided by minimizing a custom loss function (see Section 3.1.3 for specific formulation), integrating the Adam optimizer and MultiStepLR learning rate scheduler. An early stopping mechanism is embedded in the training process. Training terminates when the training loss does not decrease significantly for 3 consecutive epochs to avoid overfitting. Experimental pruning is implemented through Optuna (an open-source hyperparameter optimization framework Akiba et al. (2019)) to improve search efficiency in high-dimensional parameter spaces. This strategy deeply couples the time step adjustment mechanism with data scenarios. It ensures the model efficiently extracts spatio-temporal features in both simulated training and testing scenarios.

Numerical Example: To concretely illustrate the mechanism of the adaptive time step adjustment, consider a set of parameters optimized for a typical high-missing-rate testing scenario: threshold $\theta = 0.6$, slope $s = 5$, with bounds $t_{\min} = 9$ and $t_{\max} = 13$.

Case 1 – High missing rate ($r = 0.8$): Input missing rate r is above the threshold θ.

Compute: $x = (r - \theta) \times s = (0.8 - 0.6) \times 5 = 1.0$

Then: $w = \text{Sigmoid}(x) = \text{Sigmoid}(1.0) \approx 0.73$

Finally: $t = \text{round}\big(t_{\min} + (t_{\max} - t_{\min}) \times w\big) = \text{round}\big(9 + (13 - 9) \times 0.73\big) \approx \text{round}(11.9) = 12$

Interpretation: For a data block with 80% missing pixels, the mechanism selects a longer time step (12 days), expanding the temporal window to integrate information from more distant observations. **Case 2 – Low missing rate ($r = 0.3$):** Input missing rate r is below the threshold θ.

Compute: $x = (r - \theta) \times s = (0.3 - 0.6) \times 5 = -1.5$

Then: $w = \text{Sigmoid}(x) = \text{Sigmoid}(-1.5) \approx 0.18$

Finally: $t = \text{round}\big(t_{\min} + (t_{\max} - t_{\min}) \times w\big) = \text{round}\big(9 + (13 - 9) \times$

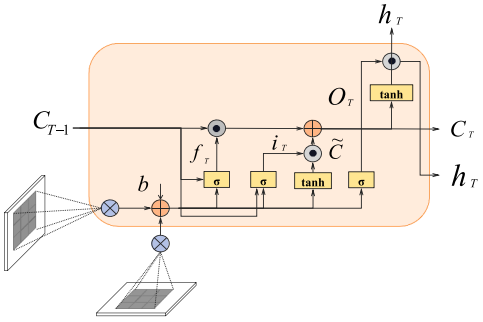


Fig. 4. Basic unit structure of ConvLSTM.

$0.18) \approx \text{round}(9.7) = 10$

Interpretation: For a relatively complete data block (30% missing), the mechanism opts for a shorter time step (10 days), focusing on the nearest frames to capture fine-grained local dynamics.

3) **ConvLSTM spatio-temporal reconstruction model**

In the spatio-temporal feature extraction stage, the proposed framework employs a ConvLSTM model to process incomplete SST data over consecutive periods. The key is to enhance the spatial feature capture capability of traditional LSTM through convolution operations (Graves, 2012). As shown in Fig. 3, the model is composed of multiple stacked ConvLSTM Cell units. Each unit realizes joint modeling of spatio-temporal information through a convolutional gating mechanism (Chai et al., 2019; Hu et al., 2020; Lin et al., 2020; Moishin et al., 2021; Mustaqeem & Kwon, 2020).

Each ConvLSTM Cell inherits the gating mechanism of traditional LSTM, while it replaces fully connected operations with 3×3 convolutional layers to preserve the spatial structure of the input data. As depicted in Fig. 4, the specific gating process is given as follows:

Forget Gate: The forget gate determines the retention degree of the previous cell state c_{t-1}. It calculates the spatial correlation between the current input x_t and the previous hidden state h_{t-1} through a 3×3 convolution layer:

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + b_f) \tag{3}$$

where σ is the Sigmoid function (output range $[0, 1]$). W_{xf} and W_{hf} are 3×3 convolution kernel weights. b_f is the bias term. $*$ denotes the convolution operation. The forget gate output f_t is used to selectively retain valid information in c_{t-1}.

Input Gate: The input gate is responsible for screening new information from the current input and consists of two parts:

a) Update Signal: Determine the positions requiring update through the Sigmoid layer:

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + b_i) \tag{4}$$

b) Candidate State: Generate information-containing candidate cell states through the Tanh layer:

$$\tilde{c}_t = \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \tag{5}$$

After element-wise multiplication of these two components, they jointly update the cell state together with the forget gate output.

Cell State Update: Combining the outputs of the forget gate and input gate, the current cell state c_t is updated as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{6}$$

where $\odot$ denotes element-wise multiplication, enabling selective forgetting of historical states and fusion of new information.

Output Gate: The output gate determines the output proportion of cell state c_t. First, the output weight is determined through the Sigmoid layer, then the hidden state h_t is generated after Tanh processing:

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + b_o) \tag{7}$$

This process ensures the output contains only valid information related to current spatio-temporal features.

In the proposed framework, the ConvLSTM model is configured with 6 layers, with hidden dimensions of $[16, 32, 64, 64, 64, 64]$ sequentially and 3×3 convolution kernels. After dimension adjustment, input data is fed into each ConvLSTM Cell layer sequentially by time steps. The first layer receives single-channel input and extracts basic spatio-temporal features through 3×3 convolution. Subsequent layers first capture multi-scale spatio-temporal dependencies with incrementally increasing hidden dimensions ($16 \rightarrow 32 \rightarrow 64$). The last three layers maintain 64-dimensional hidden states to deep feature extraction. Each layer performs cyclic calculations by time steps to explicitly capture dynamic changes in the temporal dimension of input data. Outputs are passed to the next layer after ReLU activation. The hidden state output from the final layer takes the result of the last time step as spatio-temporal feature representation, which is processed by the final convolution layer for reconstruction. This design incorporates cyclic processing of time steps in each layer. It also combines spatial feature extraction by 3×3 convolution kernels. This enables the model to capture dynamic evolution of time-series while mining spatial local correlations. It achieves deep fusion of spatio-temporal features, providing refined feature support for missing SST data reconstruction in complex regions.

Building on this foundation, this paper designs a latitude-adaptive median filtering post-processing method for regions with high missing rates such as high-latitude sea ice-covered areas. It achieves differential smoothing for different latitude regions by dynamically adjusting the filtering window. The method calculates the filtering window size based on latitude values: For the i-th latitude row, the window width w is determined as follows:

$$w = 2 \cdot \left\lfloor \frac{90 - |\text{Lat}[i]|}{90} \cdot 7 \right\rfloor + 1 \tag{8}$$

where $\text{Lat}[i]$ represents the latitude value corresponding to the current row. This design allows lower-latitude areas to obtain larger windows to capture broader spatial correlations. While higher-latitude areas use smaller windows to avoid detail loss caused by over-smoothing.

In practical implementation, one-dimensional median filtering is applied individually to each latitude row (as illustrated in algorithm 1). The SST matrix $T_{m \times n}$ is processed row by row to generate the filtered matrix T_{filtered}. During processing, land pixels are set to NaN using a land mask $\text{Mask}_{\text{land}}$ to prevent interference from land features on ocean region filtering. This method effectively suppresses outliers in high-missing-rate areas while preserving local detailed features of SST. It enhances the spatial continuity of reconstruction results along latitude gradients. It is particularly suitable for noise smoothing in complex high-latitude sea ice environments such as the Arctic (Moishin et al., 2021).

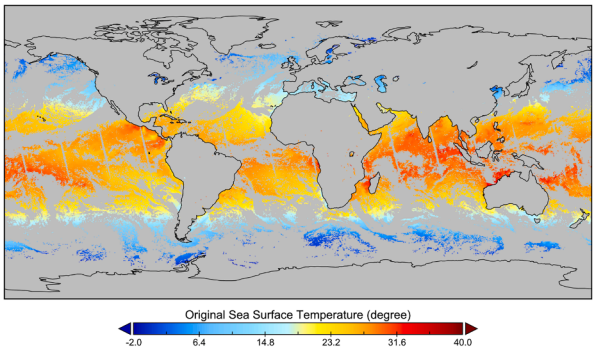
Algorithm 1 Latitude-adaptive Median Filter.

Input:
$T_{m \times n}$: Sea surface temperature matrix
Lat: Latitude vector (degrees)
Output: T_{filtered}: Filtered SST matrix with land masked

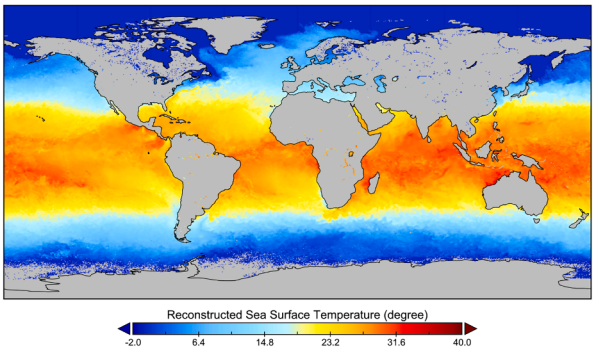
for $i \leftarrow 1$ to m **do**
  $w \leftarrow 2 \cdot \left\lfloor \frac{90 - |\text{Lat}[i]|}{90} \cdot 7 \right\rfloor + 1$
  $T_{\text{filtered}}[i, :] \leftarrow \text{MedianFilter1D}(T[i, :], w)$
end for
return T_{filtered}

3.1.3. Model training and optimization

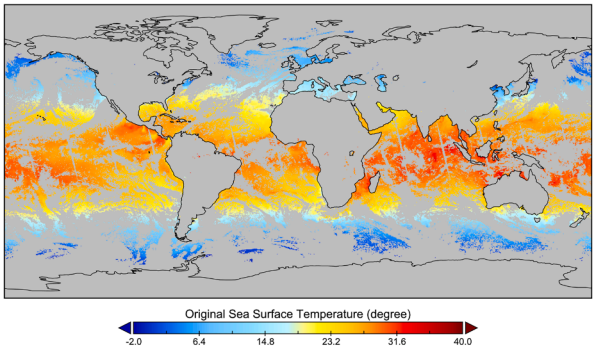
In the model training and optimization process, this framework designs a loss function based on the spatio-temporal distribution characteristics of SST data. It also introduces a post-processing mechanism to improve reconstruction quality. Considering that monthly VIIRS SST data has higher effective data coverage and spatial consistency than daily



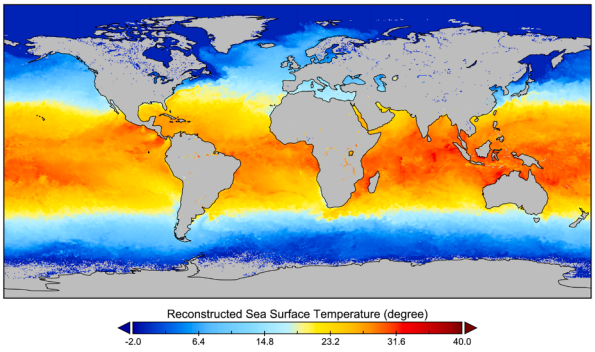
(a) Original SST in 2020.3.10



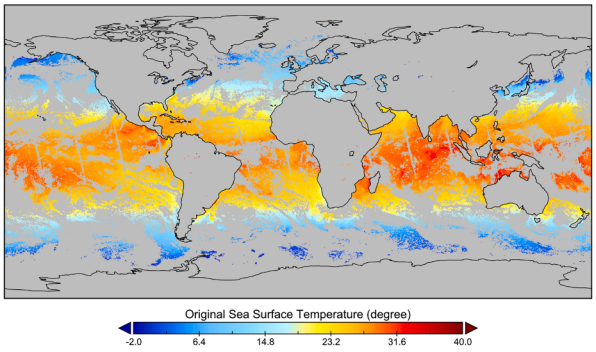
(b) SGD-SST 2.0 in 2020.3.10



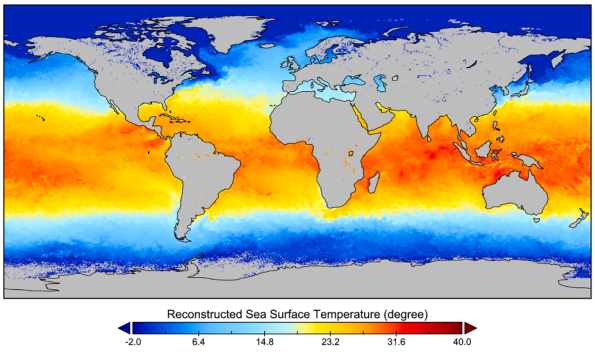
(c) Original SST in 2020.3.11



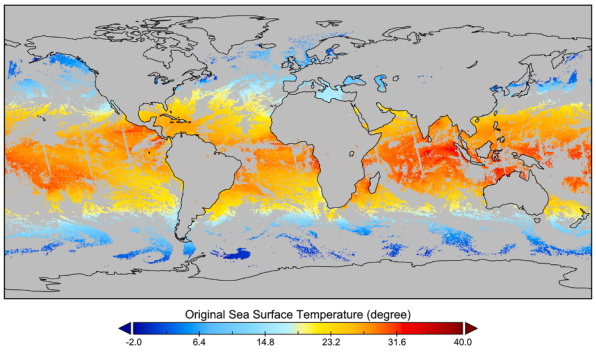
(d) SGD-SST 2.0 in 2020.3.11



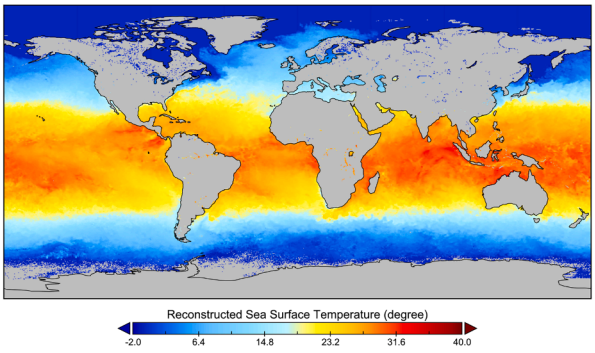
(e) Original SST in 2020.3.12



(f) SGD-SST 2.0 in 2020.3.12

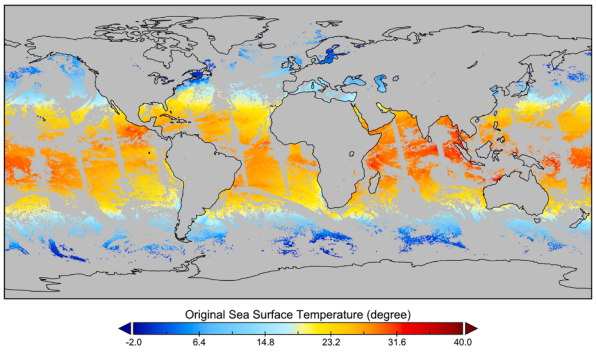


(g) Original SST in 2020.3.13

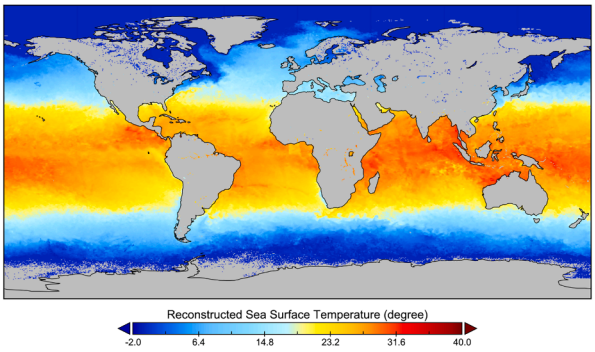


(h) SGD-SST 2.0 in 2020.3.13

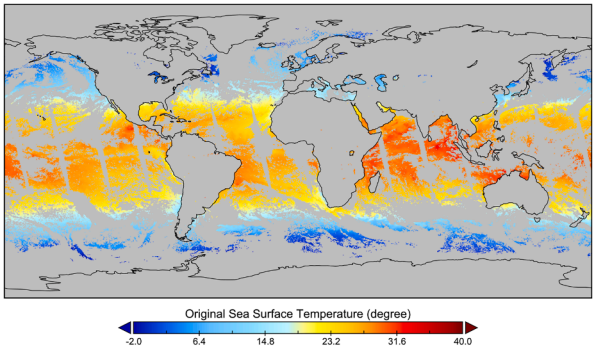
Fig. 5. Original and reconstructed SGD-SST 2.0 results from March 10, to 13, 2020.



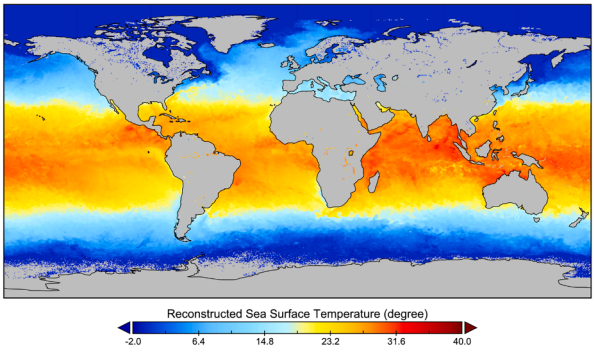
(a) Original SST in 2005.4.6



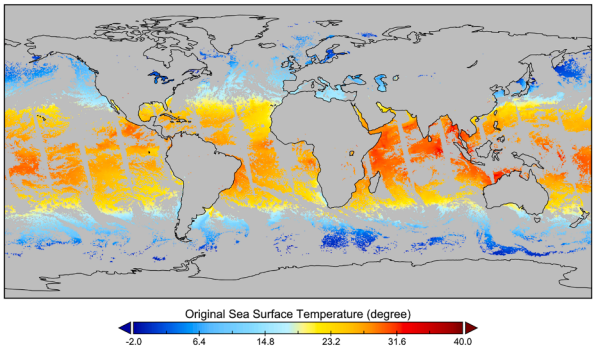
(b) SGD-SST 2.0 in 2005.4.6



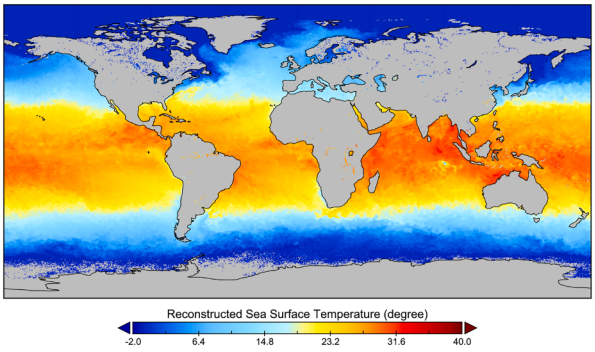
(c) Original SST in 2005.4.7



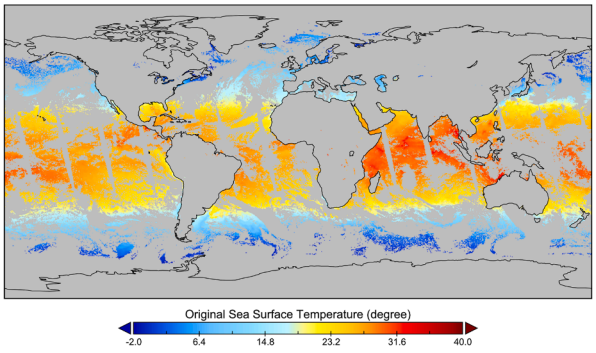
(d) SGD-SST 2.0 in 2005.4.7



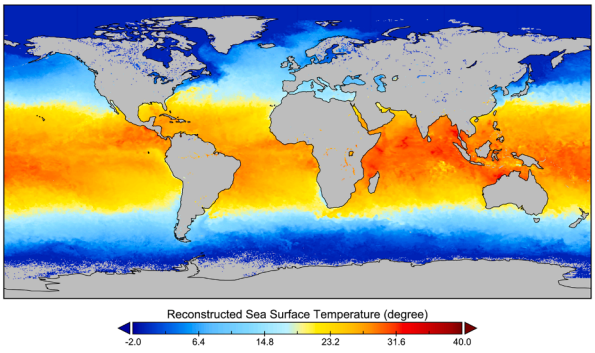
(e) Original SST in 2005.4.8



(f) SGD-SST 2.0 in 2005.4.8



(g) Original SST in 2005.4.9



(h) Reconstructed SST in 2005.4.9

Fig. 6. Original and reconstructed SGD-SST 2.0 results from April 6, to 8, 2005.

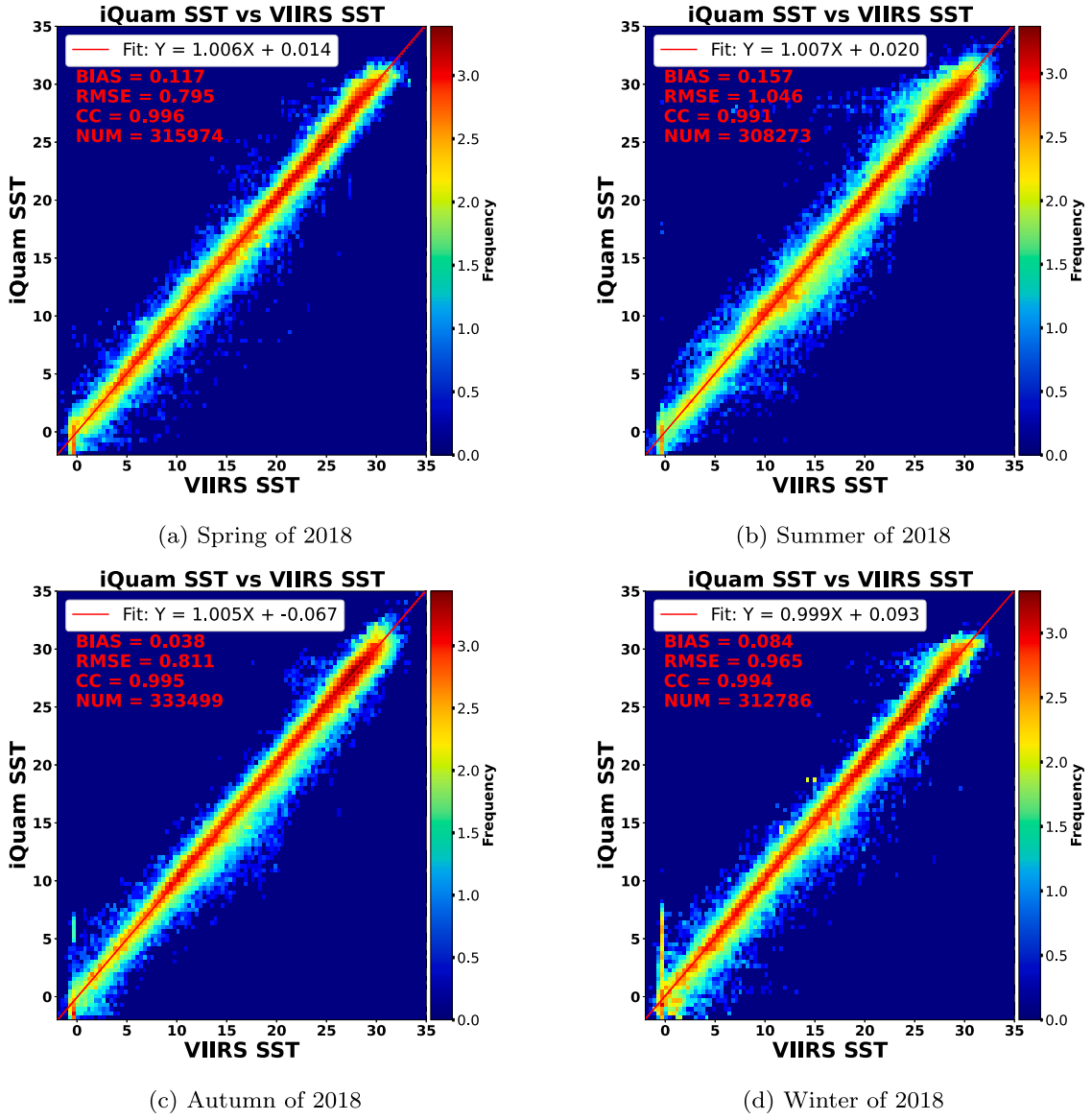


Fig. 7. Spatio-temporal matching scatter plots of the original VIIRS/SGD-SST 2.0 products and in-situ data for the four seasons in 2018.

data, the total loss function combines reconstruction loss and monthly smoothing loss. This combination focuses on constraining the reconstruction accuracy of missing regions. The reconstruction loss measures the pixel-wise difference between the reconstruction result and the real label. It drives the model to learn the mapping from spatio-temporal features to true values. The monthly smoothing loss suppresses outliers using the spatial consistency of monthly average SST.

The reconstruction loss is defined as:

$$L_{\text{rec}} = \frac{1}{2N} \left\| (1 - M_T) \odot (SST_{\text{rec}} - SST_{\text{ori}}) \right\|_2^2 \tag{9}$$

The smoothing loss is designed as:

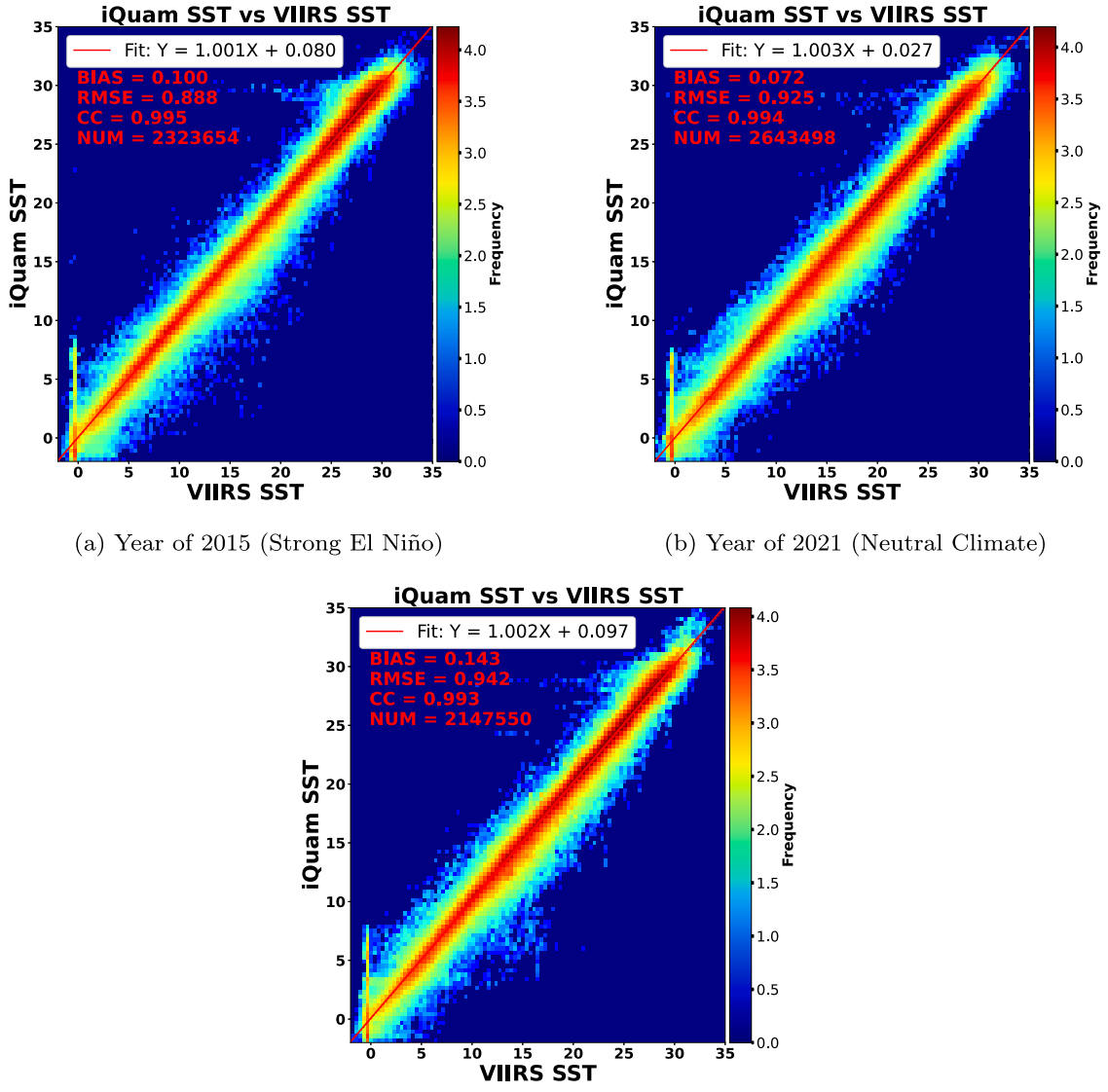
$$L_{\text{smo}} = \frac{1}{2N} \left\| (1 - M_T) \odot (SST_{\text{rec}} - SST_{\text{month}}) \right\|_2^2 \tag{10}$$

where N represents the number of samples. M_T denotes the mask for the current SST data at date T. SST_{rec} and SST_{ori} respectively represent the reconstructed SST block and the original seamless SST block. SST_{month} denotes the monthly average SST block. The total loss function of the deep spatio-temporal fusion model combines the reconstruction loss L_{rec} and the smoothing loss L_{smo} as follows:

$$L_{\text{total}} = L_{\text{rec}} + \omega L_{\text{smo}} \tag{11}$$

In the reconstruction loss function, the model gradually approximates real data by evaluating the difference between reconstructed SST data and true labels. This enables global SST reconstruction. In the smoothing loss function, outliers in SST reconstruction are suppressed using monthly average values. Both loss functions only calculate errors for missing regions. This focuses the deep fusion model on reconstructing missing areas. In the total loss function, the weight factor ω that balances L_{rec} and L_{smo} is fixed at 0.1.

The Adam optimizer was used during the training phase. The initial learning rate was set to 0.001, decaying by a factor of 0.5 every 100 epochs. The batch size was fixed at 128, with a total of 500 training epochs to fully capture data distribution characteristics. Experiments were implemented based on the PyTorch framework, relying on hardware configuration including NVIDIA RTX 4090 GPU, i9-12900K CPU, and 64GB RAM. Based on the above hardware and training configuration, the total training time of the proposed model is approximately 36 hours. For the inference (reconstruction) task, processing a single daily global SST field with a resolution of 2160×4320 pixels takes about 45 to 60 seconds on the same GPU. The main computational cost is concentrated in the training phase, and the model exhibits favorable efficiency for batch processing long-term time series data after train-



(a) Year of 2015 (Strong El Niño)

(b) Year of 2021 (Neutral Climate)

(c) Year of 2022 (Moderate Anomaly)

Fig. 8. Spatio-temporal matching scatter plots of iQuam SST vs. VIIRS SST for years with distinct climate backgrounds (2015, 2021, 2022).

ing. This ensured high efficiency for large-scale data training. By combining monthly scale constraints with adaptive filtering, this solution effectively balances global reconstruction consistency and local detail accuracy. It provides optimized support for reliable reconstruction of complex global SST products.

3.2. Validation method

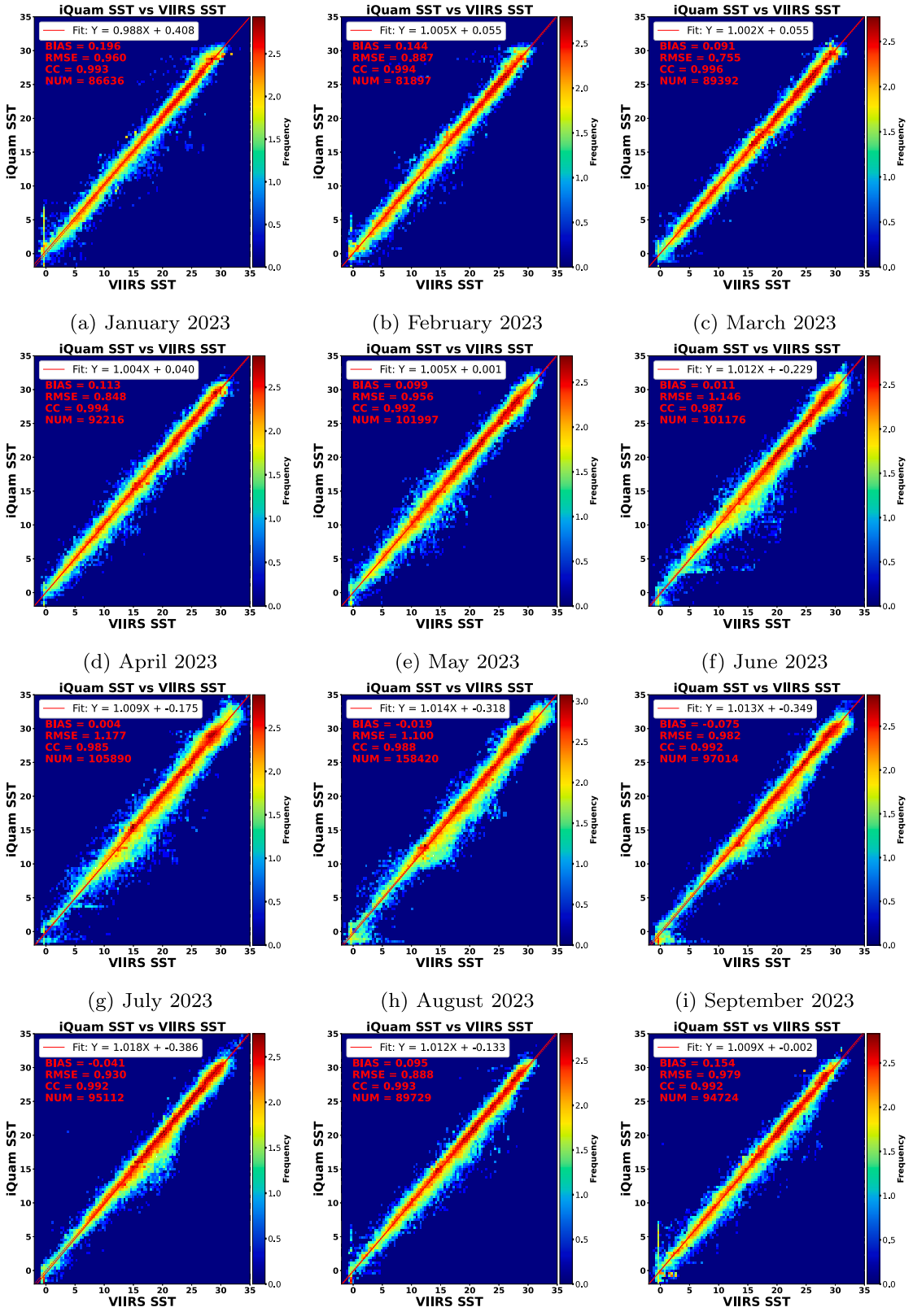
To further verify the accuracy and effectiveness of the SGD-SST 2.0 products in this work, four validation methods are adopted:

a) In-situ validation: Using in-situ observation data from the iQuam system as the true values of SST, to validate the reconstructed SST data.
b) Time-series consistency analysis: From the temporal dimension, evaluating the temporal continuity between reconstructed SST data and original valid SST data, and analyzing the consistency of long-term variation trends.
c) Comparison with DINEOF method: Comparing the reconstruction results of SGD-SST 2.0 products with DINEOF, a classical reconstruction method for SST. Through quantitative indicators and visual analysis, verifying the reconstruction accuracy and spatio-temporal feature preservation capability of the proposed method in complex missing scenarios.
d) Latitude-stratified comparison: Latitude-stratified comparison focuses on the mid-low latitude region (60°S – 60°N). SGD-SST 2.0 and 1.0 products are respectively matched and compared with in-situ observation data.

For evaluating the accuracy of reconstructed SST data, the site-product spatio-temporal matching method plays a critical role. This study uses in-situ SST data with quality level-5 provided by the iQuam system, covering all types of measured data.

In the temporal dimension, considering that both S-NPP and Aqua satellites pass by at 13:30 local time, data from the 12:30-14:30 time window in the iQuam station data is selected as the matching dataset for validation. This selection ensures the tight temporal matching between SST station data and reconstructed products, effectively reflecting SST conditions around the satellite overpass time.

In the spatial dimension, taking the iQuam in-situ data as the center, KD-tree strategy is used for fast search. KD-tree is a tree-like data structure that stores data points in k-dimensional space for ef-



(a) January 2023 (b) February 2023 (c) March 2023

(d) April 2023 (e) May 2023 (f) June 2023

(g) July 2023 (h) August 2023 (i) September 2023

(j) October 2023 (k) November 2023 (l) December 2023

Fig. 9. Spatio-temporal matching scatter plots of the original VIIRS/SGD-SST 2.0 products and in-situ data for 12 months in 2023.

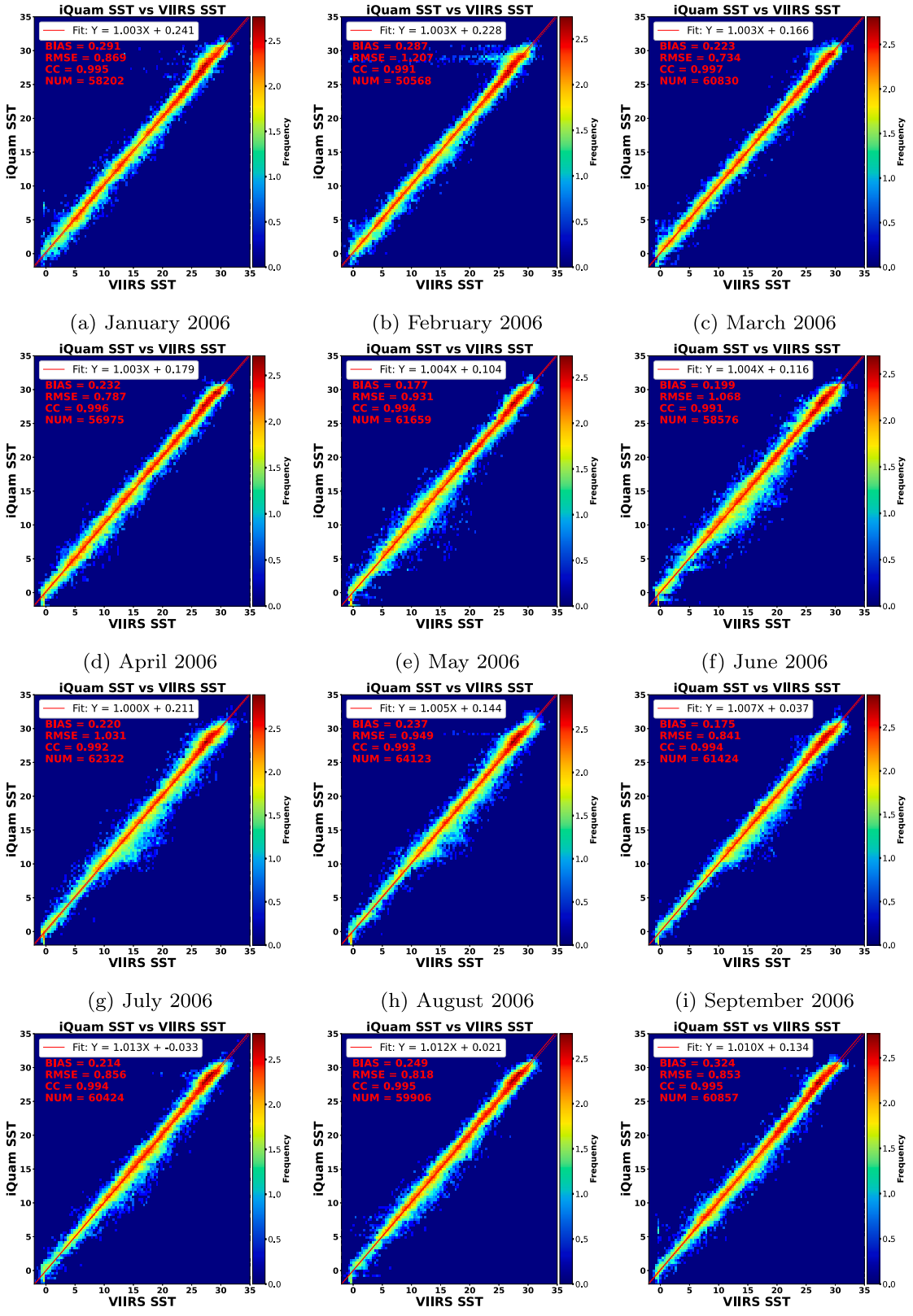
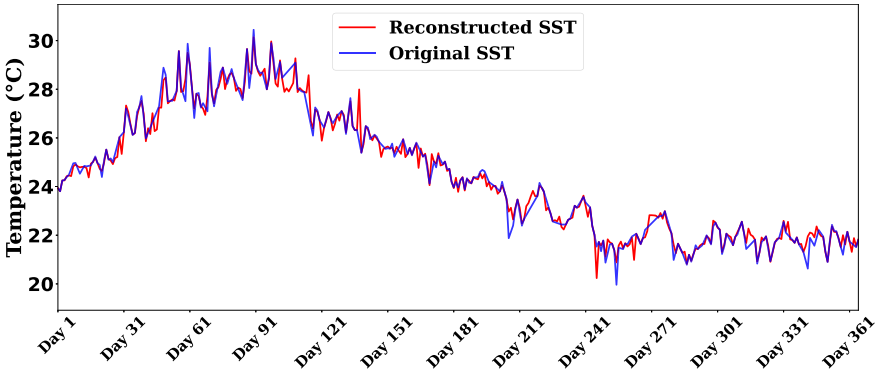
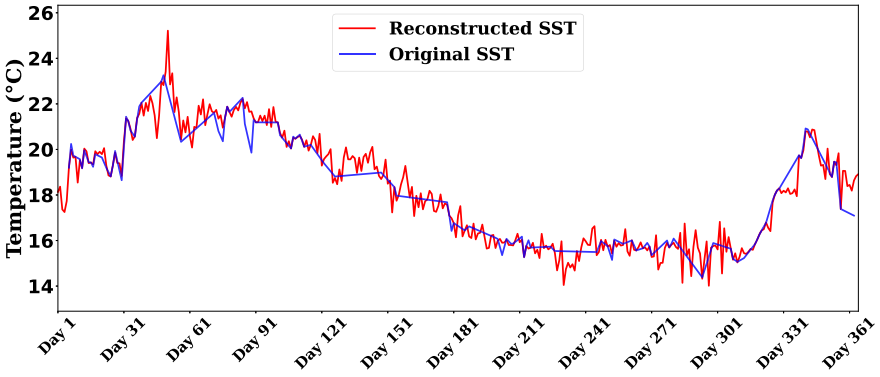


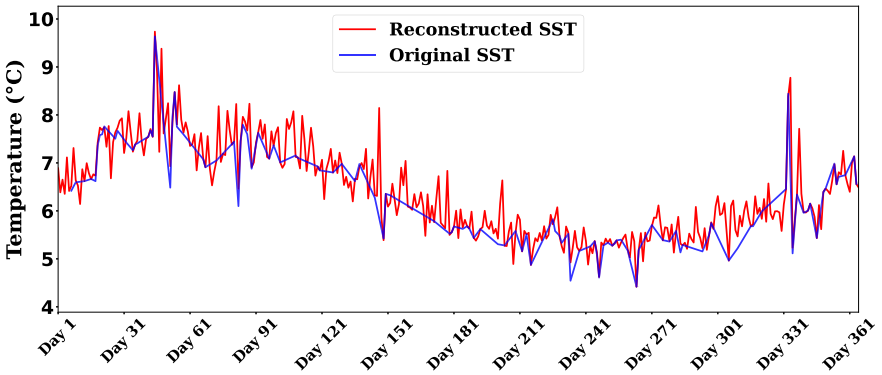
Fig. 10. Spatio-temporal matching scatter plots of the original VIIRS/SGD-SST 2.0 products and in-situ data for 12 months in 2006.



(a) VIIRS low-latitude region (1.81°S, 107.96°W) in 2023



(b) VIIRS mid-latitude region (35.88°S, 116.88°W) in 2023



(c) VIIRS high-latitude region (55.71°S, 61.13°W) in 2023

Fig. 11. (a) Continuous daily time-series variation curves of original and reconstructed SGD-SST 2.0 in Different Regions for VIIRS and MODIS SST data (Part 1: VIIRS).

ficient retrieval. In this study, it can efficiently find points close to the in-situ data in large datasets. Meanwhile, the Haversine formula is used to calculate geographical distances. Both S-NPP VIIRS and Aqua MODIS SST data have a spatial resolution of 9 km. Based on this factor, a radius of 4.5 km is selected as the matching spatial range. This setting could effectively ensure that matching points precisely fall within individual pixels, greatly improving the accuracy and representativeness of mutual matching between in-situ data and reconstructed products. It makes the in-situ validation results more reliable and persuasive.

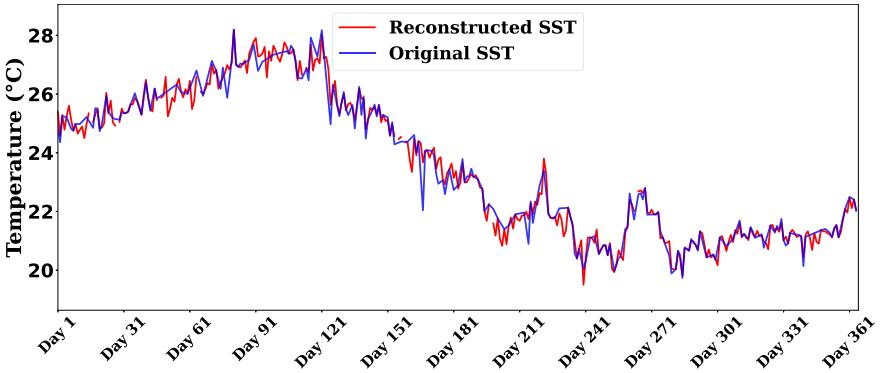
4. Experimental results and validation

In this section, we present experimental results and related validations, to demonstrate the effectiveness of the proposed reconstruction framework. Through this framework, this work successfully generates seamless global daily SST long-term products from Aqua MODIS and S-NPP VIIRS for the period from January 1, 2003 to December 12, 2025. The SGD-SST 2.0 products are stored in NetCDF4 format and can be downloaded from the provided link.

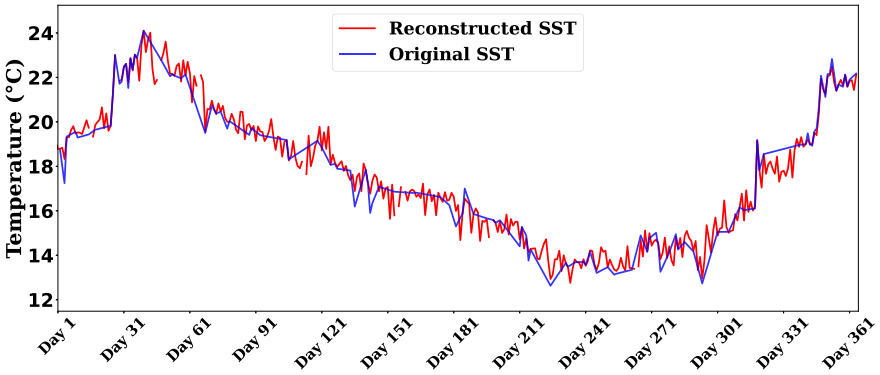
In the experiments of this section, Standard Deviation (Std_Dev, Eq. 12), Root Mean Square Error (RMSE, Eq. 13), Mean Absolute Error (MAE, Eq. 14), and Correlation Coefficient (CC, Eq. 15) are used to evaluate the performance of the reconstruction method, as depicted below:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{y})^2} \tag{12}$$

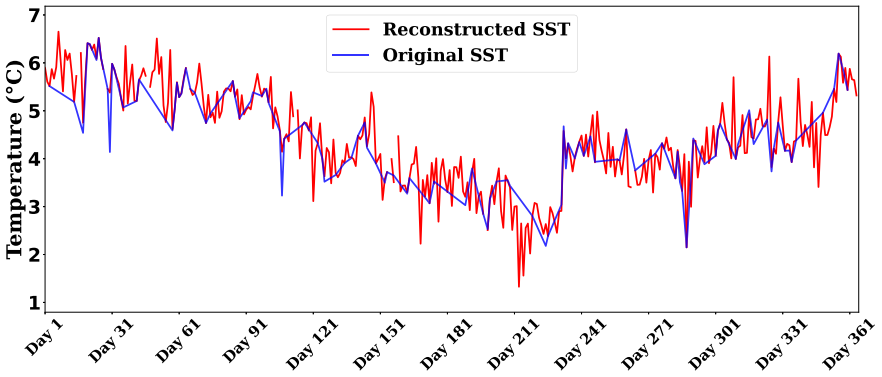
$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2} \tag{13}$$



(d) MODIS low-latitude region (1.81°S, 107.96°W) in 2003



(e) MODIS mid-latitude region (35.88°S, 116.88°W) in 2003



(f) MODIS high-latitude region (55.71°S, 61.13°W) in 2003

Fig. 11. (b) Continuous daily time-series variation curves of original and reconstructed SGD-SST 2.0 in Different Regions for VIIRS and MODIS SST data (Part 2: MODIS, continued).

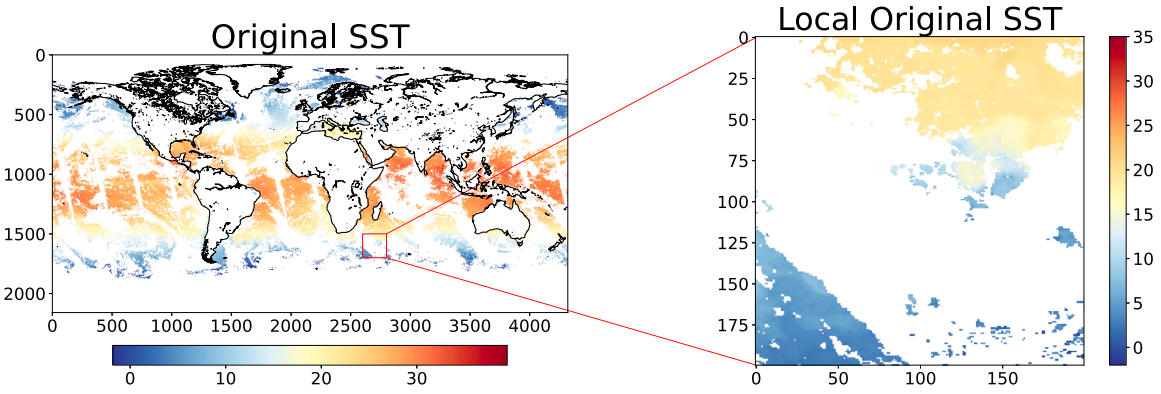
$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i| \tag{14}$$

$$CC = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{15}$$

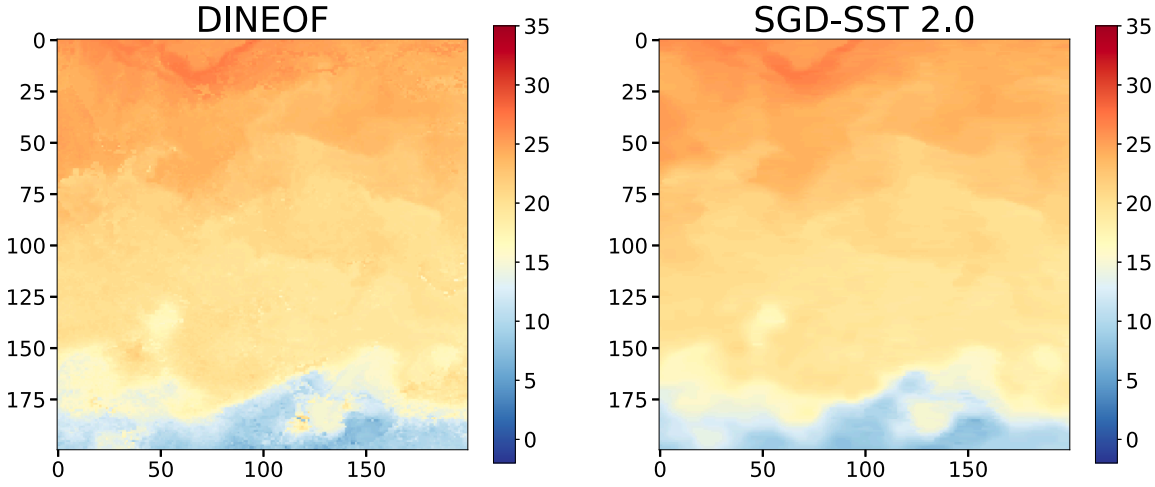
To improve the statistical rigor of the quantitative evaluation, 95% confidence intervals (CIs) for all above metrics were further calculated using the bootstrap method with 1000 resamples. The statistical analysis was conducted based on 86,947 spatio-temporally matched point pairs from January 2023. Narrow confidence intervals across all metrics indicate high statistical precision and reliability of the reconstruction results. The RMSE for January 2023 is 0.9599 ˚C, which is slightly higher than the global average RMSE of 0.916 K over the period 2013–2025. This difference is reasonable and consistent with statistical and physical principles: the single-month dataset only covers short-term spatiotemporal samples affected by local climate fluctuations and data variability, leading to a relatively higher error. In contrast, the long-term multi-year global average smooths out short-term anomalies and random fluctuations, resulting in a lower overall error. The detailed metrics with 95% CIs are summarized in Table 2.

Through the site-product spatio-temporal matching method described above, the accuracy comparison between the original SST products and the reconstructed SGD-SST 2.0 products is conducted, as listed in Table 3.

To further demonstrate the temporal fidelity and statistical rigor of the reconstruction, we conducted an in-depth validation focusing on May 2023-a period with abundant in-situ observations-incorporating



(a) Original SST in May 4, 2003



(b) Reconstructed results of DINEOF (c) Reconstructed results of SGD-SST 2.0

Fig. 12. Comparison of SST Reconstructed Results between DINEOF and SGD-SST 2.0 in May 4, 2003.

Table 2
Key metrics with 95% CI (Bootstrap: 1,000 resamples; 86k spatiotemporal pairs, Jan 2023).

Metric	Value (95% Confidence Interval)
Bias (Mean Error, ME)	-0.1959 °C [-0.2023, -0.1893] °C
Correlation Coefficient (CC)	0.9931 [0.9929, 0.9933]
Root Mean Square Error (RMSE)	0.9599 °C [0.9432, 0.9762] °C
Mean Absolute Error (MAE)	0.6097 °C [0.6048, 0.6143] °C

Table 3
In-situ validation evaluation metrics of original SST products and the reconstructed SGD-SST 2.0 products from 2013 to 2025.

SST Products	Evaluation index			
	Number	CC	RMSE	MAE
Original SST	502,091	0.993	0.779	0.526
SGD-SST 2.0	1,204,901	0.994	0.916	0.602

additional diagnostic metrics as suggested by the reviewer. Key results include:

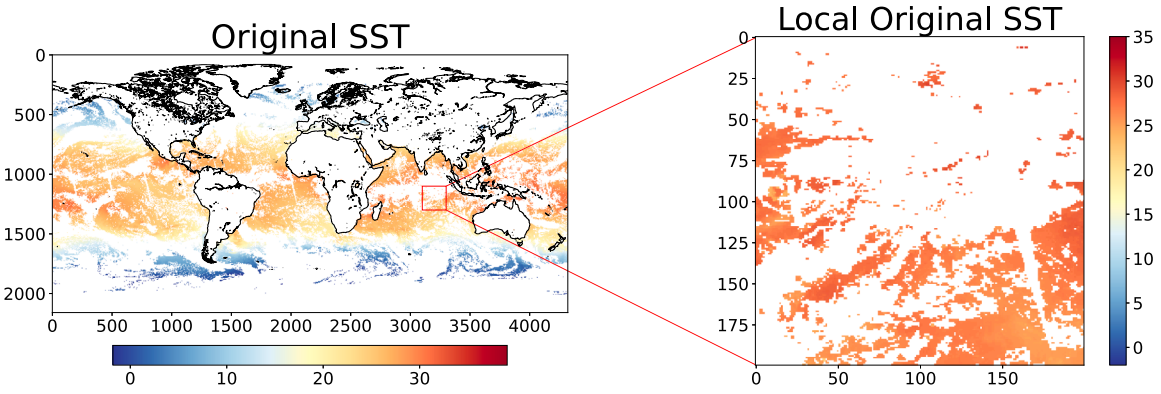
- **Bias (Mean Error, ME)**: 0.177°C, indicating a slight systematic overestimation.
- **Coefficient of Determination (R^2)**: 0.980, showing that the model explains 98.0
- **Time-Series Correlation**: Based on 566 stations with ≥15 consecutive days of observations, the average correlation is 0.617 (median: 0.705), with over 50
- **Root Mean Square Error (RMSE)**: 1.067°C.
- **Mean Absolute Error (MAE)**: 0.716°C.
- **Overall Correlation Coefficient (CC)**: 0.990.

These results confirm that the proposed model not only maintains reliable overall accuracy ($R^2 = 0.980$, overall CC = 0.990) but also exhibits stable temporal consistency at well-observed stations, with over half of the stations achieving time-series correlations greater than 0.7. The model's inherent capability to preserve spatio-temporal structures, as highlighted by other diagnostic metrics, is further discussed in the Discussion section.

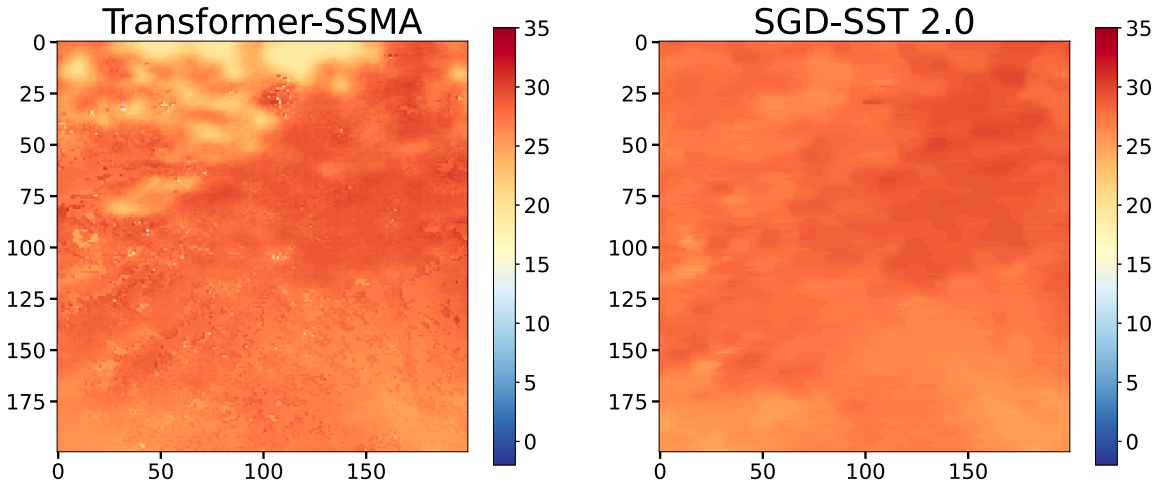
4.1. Reconstruction results of SGD-SST 2.0 products

As shown in Figs. 5 and Fig. 6, this section presents SGD-SST 2.0 products from S-NPP VIIRS during March 10–13, 2020, and from Aqua MODIS during April 6–9, 2005. The left columns of Figs. 5 and Fig. 6 show the original incomplete SST data, while the right columns display the reconstructed results of SGD-SST 2.0 products.

Different from conventional cross-sensor strategy adaptation based on network parameter fine-tuning and layer freezing, the cross-sensor transfer from VIIRS to MODIS in this work is implemented via the dynamic adaptation of reconstruction strategies, rather than modifying



(a) Original SST in January 15, 2023



(b) Reconstructed results of Transformer-SSMA (c) Reconstructed results of SGD-SST 2.0

Fig. 13. Comparison of SST Reconstructed Results between Transformer-based Transformer-SSMA and SGD-SST 2.0 (high missing rate scenario).

network weights or freezing specific layers. The primary discrepancy between VIIRS and MODIS SST data lies in the distribution of data-missing patterns, instead of low-level spatial features. To enable the VIIRS pre-trained model to adapt to MODIS data, a Bayesian optimization strategy is designed as a lightweight task adaptation layer. During the inference phase, it analyzes the actual missing rate of MODIS SST and dynamically optimizes key operational parameters (e.g., time-window threshold) to match the missing characteristics of the target dataset. This "pre-trained model plus dynamic strategy adaptation" scheme effectively transfers the high-quality reconstruction capability learned from VIIRS to MODIS, ensuring stable and reliable reconstruction performance on the MODIS dataset.

From the spatial dimension, the reconstructed SGD-SST 2.0 results show reasonable spatial continuity in connecting missing regions with adjacent valid regions. For instance, in Fig. 5d, the southern Indian Ocean experienced cloud cover, creating missing regions in the original SST data. Through spatio-temporal context correlation, the reconstructed SGD-SST 2.0 result generates the smooth area. This area could match the surrounding SST gradients, reducing the edge discontinuities common in traditional interpolation methods.

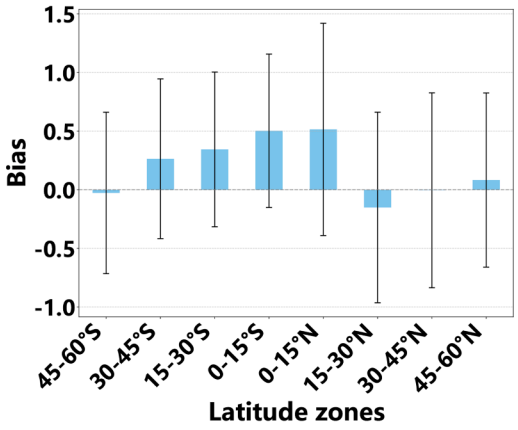
Temperature extreme distributions appear in two critical regions: the Western Pacific warm pool and the Eastern Atlantic cold pool. In the reconstruction results, these distributions maintain consistent spatial patterns with the original SST data. No abnormal jumps are observed in Figs. 5 and Fig. 6, and unreasonable temperature gradient reversals are also absent in SGD-SST 2.0 products.

From the temporal dimension, although the original daily SST results show high similarity and correlation, there are still differences and variations to varying degrees between them. The reconstructed SGD-SST 2.0 results are given in the right columns of Figs. 5 and Fig. 6. They demonstrate the performance of the proposed reconstruction method in two key aspects: firstly, it maintains consistent information across consecutive time-series SST data; secondly, it effectively fills specific information in different time-series SST data. In summary, the proposed method demonstrates superior consistency and reliability in spatio-temporal feature fusion and detailed information recovery.

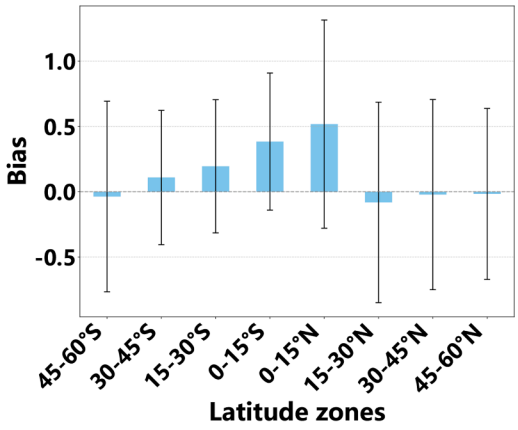
4.2. *In-situ validation*

In-situ SST data provided by iQuam could be served as the true SST values. After applying the site-product spatio-temporal matching method, it could be used to validate remote sensing satellite SST retrieval and reconstruction results. Details of the site-product spatio-temporal matching method are provided in Section 3.2.

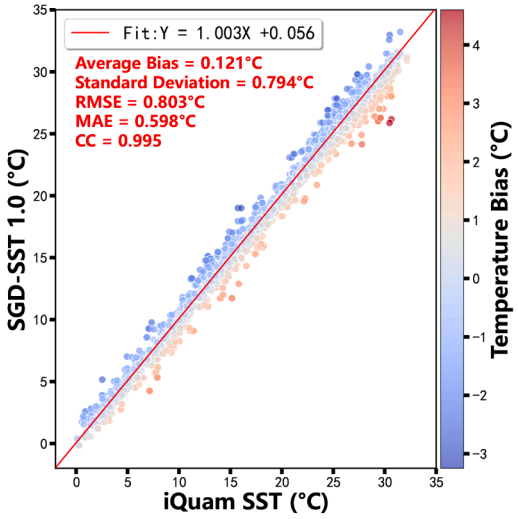
Fig. 7 presents spatio-temporal matching scatter plots between SGD-SST 2.0 products and in-situ data for the four seasons of 2018. The horizontal axis represents original VIIRS and gap-filled SGD-SST 2.0 data, while the vertical axis represents in-situ SST data obtained from the iQuam system. In Fig. 7, the red dashed line indicates the 1:1 line. Ideally, data points where the two datasets are completely consistent should lie on this line. The closer the data points are to



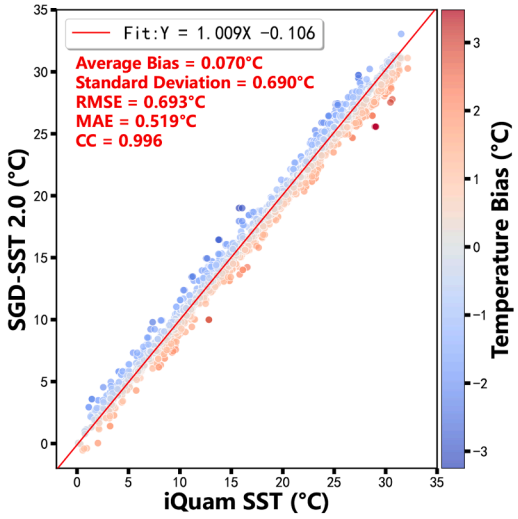
(a) Temperature bias box plots of SGD-SST 1.0 in different latitudinal regions on April 8, 2020



(b) Temperature bias box plots of SGD-SST 2.0 in different latitudinal regions on April 8, 2020



(c) Scatter plots of SGD-SST 1.0 matched with in-situ observations on April 8, 2020



(d) Scatter plots of SGD-SST 2.0 matched with in-situ observations on April 8, 2020

Fig. 14. Comparison results among SGD-SST 2.0, SGD-SST 1.0 and other related products.

this line, the higher the consistency between VIIRS SST and iQuam SST.

Statistical results from the figures show an average of approximately 317,600 matching points (with 333,499, 315,974, 308,273, and 312,786 points in each figure, respectively). CC ranges from 0.991 to 0.996, indicating high correlation between the two datasets and confirming the reliability of the reconstructed SGD-SST 2.0 data. RMSE ranges from 0.795 to 1.046, and bias ranges from 0.038 to 0.157. These results demonstrate that the reconstructed SGD-SST 2.0 products have high accuracy and credibility.

This subsection presents spatio-temporal matching analyses between reconstructed data and in-situ data for VIIRS SST (2023, Fig. 9) and MODIS SST (2006, Fig. 10), respectively. For VIIRS SST, the monthly average number of matching points is approximately 107,933. CC ranges from 0.985 to 0.996. RMSE (0.755–1.177) and BIAS (-0.075–0.196) show that RMSE reaches 1.177 in July due to increased summer weather systems (monsoons, typhoons) and cloud cover, resulting in slightly prominent errors.

To further validate the cross-year generalization ability of the proposed model, we conducted supplementary quantitative experiments on years with distinct climate backgrounds. The model was trained exclusively on 2021 neutral climate data, and evaluated on three datasets: 2015 (strong El Niño), 2022 (moderate anomaly), and 2021 (the training/neutral year), with corresponding spatio-temporal matching visualizations and statistical metrics (see Table 4 and Fig. 8).

Table 4
Cross-year generalization metrics of the proposed model.

Year	BIAS	RMSE	CC
2021 (neutral/training)	0.072	0.925	0.994
2022 (moderate anomaly)	0.143	0.942	0.993
2015 (strong El Niño)	0.100	0.888	0.995

Statistical results in Table 4 show that core metrics remain stable across diverse climate conditions: the variation of RMSE is less than 0.06, and CC ranges from 0.993 to 0.995. This quantitatively confirms the strong cross-year generalization ability of the proposed model, indicating that it learns general spatiotemporal dependencies rather than year-specific climate signals. Furthermore, these results empirically sup-

port our assertion that the 2021 dataset covers most typical spatiotemporal patterns.

Both types of satellite data show high matching with in-situ data (strong correlation reflected by the CC values of VIIRS and MODIS). Despite error fluctuations in certain months (e.g., VIIRS July, MODIS February and July) due to environmental factors, SGD-SST 2.0 data demonstrates high overall reliability and could provide solid data support for global daily SST analysis.

4.3. Time-series consistency analysis

To further verify the temporal consistency of the SGD-SST 2.0 products, this subsection selects three different regions (low, mid, and high latitudes) from VIIRS and MODIS SST data as analysis objects: low-latitude region (1.806°S, 107.958°W), mid-latitude region (35.875°S, 116.875°W), and high-latitude region (55.708°S, 61.125°W). Continuous daily time-series variation curves for these three regions are presented, as shown in Fig. 11.

The horizontal axis represents the continuous daily time-series from day 1 to day 365 in 2023 for VIIRS SST data and in 2003 for MODIS SST data. The vertical axis represents SST values. In Fig. 11, the blue curve represents original valid SST values, and the red curve represents reconstructed SGD-SST 2.0 values for missing SST values.

As depicted in Fig. 11, the reconstructed SGD-SST 2.0 results in different sea areas exhibit fine temporal consistency with the original SST data. In high-latitude regions, due to large-area missing original SST data caused by cloud and sea ice coverage, the satellite-observed original SST data are only available for 117 days in a year. The SGD-SST 2.0 products still demonstrate reliable temporal consistency, as shown in Fig. 11c.

Overall, in regions with less missing of original SST data, the SGD-SST 2.0 products could stably reflect the temporal consistency and specific characteristics. This time-series consistency analysis demonstrates the stability of the proposed method and the usability of the SGD-SST 2.0 products.

4.4. Comparison with DINEOF and deep learning-based methods

To verify the superiority of the SGD-SST 2.0 products in spatial consistency, this section selects the Data Interpolating Empirical Orthogonal Functions (DINEOF) method as the benchmark for comparison with the SGD-SST 2.0 products. DINEOF's core principle involves performing feature decomposition on the original spatio-temporal field, extracting the first few principal modes that explain the maximum data variance, and then using these modes' spatio-temporal characteristics to reconstruct missing values.

As shown in Fig. 12, although the reconstruction results of the DINEOF could reflect the overall distribution trend of SST, obvious noise is visually observed in the results. In contrast, the reconstruction results of SGD-SST 2.0 exhibit better visual smoothness, preserving key spatial features of SST while effectively suppressing interference from irrelevant noise. This difference is more pronounced in scenarios requiring capture of subtle fluctuations such as SST diurnal variations: SGD-SST 2.0 could maintain detailed features while avoiding false fluctuations caused by noise introduction.

From the perspective of methodological principles, DINEOF relies on modal truncation strategies for its robustness to data noise. When noise energy overlaps with effective signal energy in the modal space, perfect separation becomes difficult to achieve. In contrast, SGD-SST 2.0 leverages the nonlinear modeling capabilities of deep learning models. By constructing ConvLSTM with spatio-temporal feature interactions, it could automatically learn complex patterns in SST data. This characteristic enables SGD-SST 2.0 to more accurately characterize the real variation trends of SST while maintaining spatial continuity.

To further comprehensively evaluate the performance of the proposed method, we additionally compare it with a recent deep learning approach: the Transformer-based SSMA (Spatial-Spectral Multi-head Attention) method. This method adopts a spatial-spectral multi-head attention mechanism to capture long-range dependencies in SST data, which is representative of state-of-the-art deep learning reconstruction frameworks. The comparative experimental results are presented in Fig. 13, with quantitative metrics supplemented in the corresponding table.

Visually, as shown in Fig. 13, the Transformer-based SSMA method achieves relatively smooth reconstruction results but struggles to adapt to the complex missing patterns of satellite SST data (e.g., large-area continuous missing regions caused by cloud cover). In contrast, SGD-SST 2.0 maintains higher spatial consistency in large missing regions and better preserves subtle SST gradient features. Quantitatively, the SSMA method also lags behind SGD-SST 2.0 in key metrics such as RMSE and CC, especially under high missing rates (≥0.6).

To further quantify and compare the performance differences of the two methods under different missing rates, this paper presents detailed quantitative comparison results in the table. Specifically, the table reports the RMSE, MAE, and CC values of SGD-SST 2.0 and Transformer-SSMA at three representative missing rates (69.2855%, 67.4914%, and 66.0507%). All metrics are derived from original experimental data, ensuring a fair and reliable performance comparison under identical missing data scenarios.

The core reason for this performance gap lies in the design focus of the methods: the Transformer-based SSMA emphasizes general long-range feature capture through complex attention mechanisms, but lacks targeted adaptation to the specific missing characteristics of satellite SST data. In contrast, SGD-SST2.0's core innovation is the dynamic adaptive reconstruction mechanism tailored to missing patterns-via Bayesian optimization, it dynamically adjusts key operational parameters (e.g., time-window threshold) according to the actual missing rate and distribution of the input data. This design directly addresses the key challenge of satellite SST reconstruction (i.e., diverse and variable missing patterns) and avoids the inefficiency of merely increasing network complexity. Thus, SGD-SST 2.0 achieves more stable and reliable performance in this specific task Table 5.

Table 5
Comparison of RMSE, MAE and CC under different missing rates.

Missing Rate (%)	Method	RMSE	MAE	CC
69.2855	SGD-SST 2.0	0.8294	0.5151	0.9924
	Transformer-SSMA	1.7693	1.0130	0.9661
67.4914	SGD-SST 2.0	0.7137	0.5226	0.9942
	Transformer-SSMA	1.7133	0.9205	0.9704
66.0507	SGD-SST 2.0	0.7234	0.5294	0.9951
	Transformer-SSMA	1.3968	0.8204	0.9776

Table 6
Comparison between SGD-SST 1.0 and 2.0 products.

SST Products	SGD-SST 1.0	SGD-SST 2.0
Period	2013-2024	**2003-2025**
Sensor	VIIRS	**VIIRS and MODIS**
R	0.991	**0.993**
RMSE	0.818	**0.696**
MAE	0.612	**0.522**
Average Bias	0.136	**0.089**
Standard Deviation	0.806	**0.690**

4.5. Comparison between SGD-SST 2.0 and SGD-SST 1.0

Based on the situ-product spatio-temporal matching method, this section conducts in-situ validation for the SGD-SST 1.0 and SGD-SST 2.0 products. The comparison between SGD-SST 1.0 and 2.0 products is listed in Table 6.

As shown in Fig. 14, a comparison of the matching results with in-situ data reveals that the SGD-SST 2.0 products demonstrates optimized bias performance in most latitudinal regions relative to the SGD-SST

1.0 products. For instance, in regions such as 0–15°S, the SGD-SST 2.0 products exhibit smaller bias fluctuations, indicating more stable performance. This enhancement stems from fundamental mechanistic improvements in the core model of SGD-SST 2.0 products.

As displayed in Fig. 14, in the 45°S-30°N domain, SGD-SST 2.0 products show superior bias performance over 1.0 in all regions except 0–15°N, with absolute bias <0.2°C indicating enhanced reconstruction accuracy. These indicators collectively show that SGD-SST 2.0 products have better consistency with in-situ data.

Besides, the reconstruction accuracy of SGD-SST 2.0 products has significantly improved compared to the SGD-SST 1.0 products in the mid-low latitude regions (60°S-60°N).

The SGD-SST 1.0 products employ a deep spatio-temporal fusion model, which represents a static architecture. It extracts and fuses features through fixed downsampling-upsampling paths and gated convolutions. This design inherently assumes a uniform expectation for the spatio-temporal completeness of input data. When confronted with high-proportion, non-uniform data gaps prevalent in real-world scenarios (especially in high-latitude sea-ice zones or areas with high cloud coverage), its fixed receptive field and temporal window struggle to capture sufficient long-term dependencies. This limitation may lead to discontinuities in the reconstructed regions.

In contrast, SGD-SST 2.0 products are built upon a data-driven adaptive ConvLSTM model. Its core innovation is the introduction of a "data-missing-rate-driven dynamic time-step selection mechanism." This mechanism dynamically and smoothly adjusts the model's observational time window length based on the real-time missing data status. Specifically, it automatically expands the window to mine long-term climatological signals when data is sparse, and contracts it to focus on local fine-scale evolution when data is complete. This "adaptive adjustment based on data status" capability endows the model with inherent robustness against complex and variable observational conditions.

Furthermore, the post-processing in SGD-SST 1.0 products did not explicitly account for the geographical zonation of oceanic phenomena. SGD-SST 2.0 products address this by introducing a "latitude-effect-adaptive post-processing method." The core of this method is a median filter with a dynamically adjusted window size, where the window size decreases as the absolute latitude increases. This design incorporates a clear physical rationale: the SST field in low latitudes has a large spatial correlation scale, making it suitable for larger windows to smooth noise; conversely, key features like oceanic fronts and eddies in mid-to-high latitudes are smaller in scale and exhibit strong gradients, requiring smaller windows to preserve their morphological structure and sharpness. Consequently, this processing significantly improves the fidelity of key oceanic dynamical features in mid-to-high latitudes, while also effectively suppressing random noise globally. This allows the SGD-SST 2.0 products to enhance accuracy without causing feature blurring.

5. Conclusions

To address the issue of missing SST data, this study constructs an adaptive ConvLSTM spatio-temporal reconstruction model via Bayesian optimization, generating seamless global daily products from 2003 to 2025 (SGD-SST 2.0). In-situ validation, time-series consistency analysis, comparison with DINEOF method, and latitude-stratified comparison are conducted. These analyses confirmed the SGD-SST 2.0's advantages in accuracy, temporal consistency, and spatial continuity. Notably, the detail preservation capability of SGD-SST 2.0 products is significantly improved in mid-low latitude regions, comparing with previous SGD-SST 1.0 products.

Despite these achievements, this study still has two notable limitations that need to be addressed in future research. First, the current model relies solely on satellite SST time-series data for reconstruction, without integrating other auxiliary physical variables (e.g., sea surface height, sea surface wind, and salinity). The incorporation of such multi-source data could provide stronger physical constraints, which may further enhance the physical rationality and reconstruction accuracy, especially in regions with persistent cloud cover and long-term data gaps. Second, the reconstruction performance is inherently constrained by the coverage of input observations. For areas with nearly complete lack of satellite observations (e.g., certain high-latitude ocean regions in winter), the model cannot generate reliable SST estimations-a common challenge for global remote sensing product reconstruction. Nevertheless, current research is limited to satellite SST data. Future plans include incorporating auxiliary data such as sea wind and multi-source heterogeneous information fusion. Besides, subsequent works will promote the integration with physical model.

CRediT authorship contribution statement

Qi Wang: Methodology, Writing – Original Draft; **Qiang Zhang:** Supervision, Methodology, Writing – Review & Editing; **Tongde Yang:** Software, Writing – Review & Editing. ; **Weizhen Sun:** Conceptualization, Writing – Review & Editing; **Qiangqiang Yuan:** Validation, Writing – Review & Editing.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Qiang Zhang reports financial support was provided by National Natural Science Foundation of China. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study is supported in part by the National Natural Science Foundation of China under Grant 62401095; And in part by the Dalian Science and Technology Talent Innovation Supporting Project under Grant 2024RQ028; And in part by the China Postdoctoral Science Foundation under Grant 2023M740460 and 2025T180065; And in part by the Natural Science Foundation of Liaoning Province under Grant Natural Science Foundation of Liaoning Province2025-BS-0236; And in part by the Fundamental Research Funds for the Central Universities under Grant 3132026267.

References

Ahmed, A., Fox-Kemper, B., Watkins, D. M., Wexler, D., & Wilhelmus, M. M. (2025). Estuarine temperature variability: Integrating four decades of remote sensing observations and in-situ sea surface measurements. *Remote Sensing of Environment*, *322*, 114643.

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 2623–2631).

Chai, D., Chai, D., Newsam, S., Zhang, H. K., Qiu, Y., & Huang, J. (2019). Cloud and cloud shadow detection in landsat imagery based on deep convolutional neural networks. *Remote Sensing of Environment*, *225*, 307–316.

Choo, M., Jung, S. C., Im, J.-H., & Han, D. (2025). CARE-SST: Context-Aware reconstruction diffusion model for sea surface temperature. *ISPRS Journal of Photogrammetry and Remote Sensing*, *220*, 454–472.

Frazier, P. (2018). A tutorial on bayesian optimization. arXiv:1807.02811, 1–13.

Frey, R. A., Ackerman, S. A., Holz, R. E., Dutcher, S., & Griffith, Z. (2020). The continuity MODIS-VIIRS cloud mask. *Remote Sensing*, *12*(20), 3334.

Gao, S. (2008). Sea surface temperature simulation of tropical and north pacific basins using a hybrid coordinate ocean model (HYCOM). *Marine Science Bulletin*, *8*(2), 1024–1032.

Gavahi, K., Foroumandi, E., & Moradkhani, H. (2023). A deep learning-based framework for multi-source precipitation fusion. *Remote Sensing of Environment*, *295*, 113723.

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., Roberts-Jones, J., Searle, T., Waters, J., While, J. et al. (2020). The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sensing*, *12*(4), 720.

Graves, A. (2012). Long short-term memory. *Supervised Sequence Labelling with Recurrent Neural Networks*, (pp. 37–45).

Haghbin, M., Sharafati, A., Motta, D., Al-Ansari, N., & Noghani, M. H. M. (2021). Applications of soft computing models for predicting sea surface temperature: A comprehensive review and assessment. *Progress in Earth and Planetary Science*, *8*(1), 4.

Hirahara, N., Sonogashira, M., & Iiyama, M. (2021). Cloud free sea surface temperature image reconstruction from anomaly inpainting network. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–11.

Hu, W.-S., Li, H.-C., Pan, L., Li, W., Tao, R., & Du, Q. (2020). Spatial–spectral feature extraction via deep convLSTM neural networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *58*(6), 4237–4250.

Jia, C., & Minnett, P. J. (2020). High latitude sea surface temperatures derived from MODIS infrared measurements. *Remote Sensing of Environment*, *251*, 112094.

Jia, X., Ji, Q., Han, L., Liu, Y., Han, G., & Lin, X. (2022). Prediction of sea surface temperature in the east china sea based on LSTM neural network. *Remote Sensing*, *14*(14), 3300.

Jiang, J., Yao, Z., & Liu, Y. (2024). Nighttime fog and low stratus detection under multi-scene and all lunar phase conditions using S-NPP/VIIRS visible and infrared channels. *ISPRS Journal of Photogrammetry and Remote Sensing*, *218*, 102–113.

Kang, D., Im, J., Lee, M.-I., & Quackenbush, L. J. (2014). The MODIS ice surface temperature product as an indicator of sea ice minimum over the arctic ocean. *Remote Sensing of Environment*, *152*, 99–108.

Kilpatrick, K. A., Podestá, G., Williams, E., Walsh, S., & Minnett, P. J. (2019). Alternating decision trees for cloud masking in MODIS and VIIRS NASA sea surface temperature products. *Journal of Atmospheric and Oceanic Technology*, *36*(3), 387–407.

Kilpatrick, K. A., Podestá, G., Walsh, S., Williams, E., Halliwell, V., Szczodrak, M., Brown, O. B., Minnett, P. J., & Evans, R. (2015). A decade of sea surface temperature from MODIS. *Remote Sensing of Environment*, *165*, 27–41.

Kolbe, W. M., Dybkjær, G., Tonboe, R. T., Eastwood, S., Nielsen-Englyst, P., Høyer, J., Jensen, A. T., & Suhr, M. B. (2025). Arctic and antarctic surface temperatures from AVHRR thermal infrared satellite sensors 1982–2023. *Remote Sensing of Environment*, *328*, 114816.

Koner, P. K., Harris, A., & Maturi, E. (2016). Hybrid cloud and error masking to improve the quality of deterministic satellite sea surface temperature retrieval and data coverage. *Remote Sensing of Environment*, *174*, 266–278.

Lin, Z., Li, M., Zheng, Z., Cheng, Y., & Yuan, C. (2020). Self-attention convlstm for spatiotemporal prediction. In *Proceedings of the AAAI conference on artificial intelligence* (pp. 11531–11538). (*vol. 34*).

Liu, X., & Wang, M. (2018). Gap filling of missing data for VIIRS global ocean color products using the DINEOF method. *IEEE Transactions on Geoscience and Remote Sensing*, *56*(8), 4464–4476.

Liu, Y., & Minnett, P. J. (2016). Sampling errors in satellite-derived infrared sea-surface temperatures. Part I: Global and regional MODIS fields. *Remote Sensing of Environment*, *177*, 48–64.

Luo, B., Minnett, P. J., & Nalli, N. R. (2021). Infrared satellite-derived sea surface skin temperature sensitivity to aerosol vertical distribution field data analysis and model simulations. *Remote Sensing of Environment*, *252*, 112151.

Marullo, S., Santoleri, R., Ciani, D., Le Borgne, P., Péré, S., Pinardi, N., Tonani, M., & Nardone, G. (2014). Combining model and geostationary satellite data to reconstruct hourly SST field over the mediterranean sea. *Remote Sensing of Environment*, *146*, 11–23.

Minnett, P. J., Alvera-Azcárate, A., Chin, T. M., Corlett, G. K., Gentemann, C. L., Karagali, I., Li, X., Marsouin, A., Marullo, S., Maturi, E. et al. (2019). Half a century of satellite remote sensing of sea-surface temperature. *Remote Sensing of Environment*, *233*, 111366.

Moishin, M., Deo, R. C., Prasad, R., Raj, N., & Abdulla, S. (2021). Designing deep-based learning flood forecast model with convLSTM hybrid algorithm. *IEEE Access*, *9*, 50982–50993.

Mustaqeem, Kwon, S., (2020). CLSTM: Deep feature-based speech emotion recognition using the hierarchical ConvLSTM network. *Mathematics*, *8*(12), 2133.

Nielsen-Englyst, P., Høyer, J. L., Kolbe, W. M., Dybkjær, G., Lavergne, T., Tonboe, R. T., Skarpalezos, S., & Karagali, I. (2023). A combined sea and sea-ice surface temperature climate dataset of the arctic, 1982–2021. *Remote Sensing of Environment*, *284*, 113331.

O'carroll, A. G., Armstrong, E. M., Beggs, H. M., Bouali, M., Casey, K. S., Corlett, G. K., Dash, P., Donlon, C. J., Gentemann, C. L., Høyer, J. L. et al. (2019). Observational needs of sea surface temperature. *Frontiers in Marine Science*, *6*, 420.

Pelikan, M. (2005). Bayesian optimization algorithm. In *Hierarchical bayesian optimization algorithm: Toward a new generation of evolutionary algorithms* (pp. 31–48). Springer.

Ping, B., Su, F., & Meng, Y. (2016). An improved DINEOF algorithm for filling missing values in spatio-temporal sea surface temperature data. *PloS One*, *11*(5), e0155928.

Shi, C., Wang, T., Wang, S., Jia, A., Zheng, X., Leng, W., & Du, Y. (2024). Mdineof: A scheme to recover land surface temperatures under cloudy-sky conditions by incorporating radiation fluxes. *Remote Sensing of Environment*, *309*, 114208.

Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in Neural Information Processing Systems*, *25*.

Sukresno, B. (2010). Empirical orthogonal functions (EOF) analysis of SST variability in indonesian water concerning with ENSO and IOD. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science Journal*, *38*(8), 116–121.

Tu, Q., Pan, D., & Hao, Z. (2015). Validation of S-NPP VIIRS sea surface temperature retrieved from NAVO. *Remote Sensing*, *7*(12), 17234–17245.

Wang, Q., Zhang, Q., Xie, H., Liu, Z., & Dong, Y. (2025). SGD-SST: Seamless global daily sea surface temperature products reconstruction and validation via deep spatio-Temporal fusion model. *Expert Systems with Applications*, *293*, 128703.

Wang, X., Jin, Y., Schmitt, S., & Olhofer, M. (2023). Recent advances in bayesian optimization. *ACM Computing Surveys*, *55*(13s), 1–36.

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., & Chen, Z. (2019). Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-Adaboost combination approach. *Remote Sensing of Environment*, *233*, 111358.

Xu, F., & Ignatov, A. (2014). In situ SST quality monitor (i quam). *Journal of Atmospheric and Oceanic Technology*, *31*(1), 164–180.

Young, C.-C., Cheng, Y.-C., Lee, M.-A., & Wu, J.-H. (2024). Accurate reconstruction of satellite-derived SST under cloud and cloud-free areas using a physically-informed machine learning approach. *Remote Sensing of Environment*, *313*, 114339.

Zhang, Q., Dong, Y., Zheng, Y., Yu, H., Song, M., Zhang, L., & Yuan, Q. (2024a). Three-Dimension spatial-spectral attention transformer for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–13.

Zhang, Q., Yuan, Q., Li, J., Li, Z., Shen, H., & Zhang, L. (2020). Thick cloud and cloud shadow removal in multitemporal imagery using progressively spatio-temporal patch group deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 148–160.

Zhang, Q., Yuan, Q., Li, J., Wang, Y., Sun, F., & Zhang, L. (2021). Generating seamless global daily AMSR2 soil moisture (SGD-SM) long-term products for the years 2013–2019. *Earth System Science Data*, *13*(3), 1385–1401.

Zhang, Q., Yuan, Q., Song, M., Yu, H., & Zhang, L. (2022). Cooperated spectral low-Rankness prior and deep spatial prior for HSI unsupervised denoising. *IEEE Transactions on Image Processing*, *31*, 6356–6368.

Zhang, Q., Zhang, X., Xiao, Y., & Xie, H. (2026). Deep low-rank tensor embedded network for hyperspectral image super-resolution. *Expert Systems with Applications*, *299*, 129864.

Zhang, Q., Zheng, Y., Yuan, Q., Song, M., Yu, H., & Xiao, Y. (2024b). Hyperspectral image denoising: From model-Driven, data-Driven, to model-Data-Driven. *IEEE Transactions on Neural Networks and Learning Systems*, *35*(10), 13143–13163.

Zhang, Q., Zhu, J., Dong, Y., Zhao, E., Song, M., & Yuan, Q. (2025). 10-Minute forest early wildfire detection: Fusing multi-type and multi-source information via recursive transformer. *Neurocomputing*, *616*, 128963.