International Journal of Applied Earth Observations and Geoinformation 117 (2023) 103195

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

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

Cloud-covered MODIS LST reconstruction by combining assimilation data and remote sensing data through a nonlocality-reinforced network

Yuting Gong [a], Huifang Li [a,*], Huanfeng Shen [a,b], Chunlei Meng [c], Penghai Wu [d]

[a] School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China
[b] The Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan 430079, China
[c] Institute of Urban Meteorology, China Meteorological Administration, Beijing 100089, China
[d] School of Resources and Environmental Engineering, Anhui University, Hefei 230039, China

ARTICLE INFO

Keywords:
Assimilation data
Land surface temperature reconstruction
Nonlocality-reinforced network
Remote sensing data

ABSTRACT

Reconstruction of cloud-covered thermal infrared land surface temperature (LST) is vital for the measurement of physical properties in land surface at regional and global scales. In this paper, a novel reconstruction method for Moderate Resolution Imaging Spectroradiometer (MODIS) LST data with a 1-km spatial resolution is proposed by combining assimilation data and remote sensing data through a nonlocality-reinforced network (NRN) model. Firstly, a data grading criterion is introduced to evaluate the importance of the various datasets, forming four combinations of multi-modal datasets for the training and testing of the NRN model. Secondly, the NRN model with a multiscale encoding–decoding structure considering the nonlocality-reinforced module is proposed for LST reconstruction. The results suggest that the proposed method can precisely reconstruct cloud-covered LST, with a mean absolute error (MAE) less than 0.8 K, even when no auxiliary remote sensing LST are used (Combination 1). The best result is the full combination (Combination 4), in which the coefficient of determination is 0.8956, the MAE is 0.5219 K, and the root-mean-square error is 0.7622 K. Compared with the traditional harmonic analysis of time series method, the improved enhanced spatial and temporal adaptive reflectance fusion method and the multiscale feature connected convolutional neural network method for LST reconstruction, the proposed method can achieve superior results. The proposed method with Combination 1 has been implemented to reconstruct the daily LST in the study area for 2019. Referring to the meteorological station observations, the reconstructed bias absolute value is less than 1 K, indicating that the proposed model is very effective and valid for regional cloud-covered LST reconstruction.

1. Introduction

Land surface temperature (LST) is an important indicator of the physical characteristics of the land surface at both regional and global scales (Li et al., 2013), and is a key research factor widely used in climatology, ecology, and military reconnaissance (Kustas and Anderson, 2009; Hansen et al., 2010; Zhang et al., 2023). Thus, LST and its spatio-temporal dynamics have always been important research directions in the international academic community (Hong et al., 2021). As a result of the small number of ground stations and their discrete spatial distribution, ground-based LST data cannot represent the continuous spatial distribution characteristics of a complex land surface (Yu et al., 2015). Remote sensing techniques have improved significantly in recent decades, and many thermal infrared (TIR) remote sensing LST products with high accuracies have been produced. Retrieving LST based on remote sensing technology has been used in an extensive manner to obtain spatially contiguous LST (Aires et al., 2004; Wang and Upreti, 2019). However, it is often difficult for satellites to obtain surface thermal information, due to the large cloud coverage with high frequency (Crosson et al., 2012). The large numbers of invalid pixels caused by cloud cover seriously restrict the subsequent application of LST data. Therefore, reconstructing the cloud-covered TIR LST data to obtain gapless LST is crucial for obtaining a continuous heat distribution of the Earth's surface (Lu et al., 2011).

In recent years, a variety of cloud removal methods for TIR LST have been proposed by utilizing the complementary LST information from other regions, times or sources to reconstruct the cloud-covered LST (Mo et al., 2021). Meanwhile, those data reflecting the surface physical/

---

* Corresponding author.
E-mail address: huifangli@whu.edu.cn (H. Li).

https://doi.org/10.1016/j.jag.2023.103195
Received 13 October 2022; Received in revised form 10 January 2023; Accepted 13 January 2023

geometric properties, such as albedo, normalized difference vegetation index (NDVI), and digital elevation model (DEM), have also been utilized to supply important auxiliary information. According to the differences of the data sources, LST reconstruction methods can be divided into two main categories: 1) homologous data supported methods; and 2) heterologous data supported methods. The homologous data supported methods are the early mainstream methods, in which the TIR LST is taken as the primary reference to reconstruct the cloud-covered LST. The complementary relationship is constructed in the spatial (Lyon et al., 2010; Neteler, 2010; Shuai et al., 2014), temporal (Di et al., 2019; Scharlemann et al., 2008a; Xu and Shen, 2013) or spatio-temporal domain (Weiss et al., 2014; Liu et al., 2017; Duan et al., 2017; Pham et al., 2019). Physics-based models, statistical models or deep learning networks (Arslan and Sekertekin, 2019; Chen et al., 2019; Wu et al., 2019; Yao et al., 2021) are built to determine this complementary relationship, so as to reconstruct the cloud-covered LST. Among the different models, deep learning models have shown advantages over the traditional methods for LST reconstruction due to their outstanding abilities to mine complex nonlinear relationships among multiple datasets (Youssef et al., 2020). However, it can be found that the cloudy-sky LST is reconstructed by using the clear-sky auxiliary information in the homologous data supported methods, which is a compromise solution in contrast with the data missing. In fact, cloudy-sky LST differs from in that of clear-sky LST (Hong et al., 2021). How to take this difference into account to obtain a more realistic cloudy-sky LST has become a challenge in the LST reconstruction. The second category of methods—the heterologous data supported methods, is established by considering the LST difference between clear-sky and cloudy-sky conditions (Cheng et al., 2020; Zhang et al., 2020; Zhao and Duan, 2020). Such methods typically use gapless LST data from other sources under the same meteorological conditions as the complementary information, such as passive microwave (PMW) LST (Tang et al., 2022), assimilation LST (Zhang et al., 2021), and indirectly using radiation data to correct LST (Zhao and Duan, 2020). Among the different data sources, PMW LST reflects the temperature of the soil surface with a coarse resolution, so that the heterologous data normalization and spatial downscaling are required when PMW LST data are used for the TIR LST reconstruction. Some complicated data-driven models have been used to solve these two problems, such as random forest model (Zhang et al., 2022) and deep learning models (Wu et al., 2022), from which high reconstruction accuracies have been reported and the strict data requirements have been met. Assimilation data realizes high-precision simulation of land surface parameters based on physical models, providing more real-time and effective information for the cloudy-sky LST reconstruction. In recent years, a few studies have generated all-weather LST using the improved enhanced spatial and temporal adaptive reflectance fusion method (improved-ESTARFM) (Long et al., 2020), the improved LST time series decomposition model (Zhang et al., 2021), and the annual/diurnal temperature cycle-based framework (Hong et al., 2021; Ma et al., 2022) and so on, combining downscaling and bias correction approaches to resolve the spatial incompleteness and tradeoff between the heterologous LST (Jia et al., 2021; Wang et al., 2021). In addition to directly referring to LST products, there is another solution that uses the difference of the radiation in cloudy and clear-sky conditions, radiation data was used to correct the LST results under clear-sky conditions, which has achieved realistic results (Zeng et al., 2018; Yang et al., 2019).

In summary, it can be found that both homologous and heterologous data are valuable for the cloud-covered TIR LST reconstruction (Wu et al., 2022). However, the auxiliary homologous data at other times are also likely to be covered by cloud, and thus fail to provide sufficient complementary information for the cloud-covered LST reconstruction. Therefore, using heterologous data, such as PMW data and assimilation data, to realize cloud-covered LST reconstruction has been a research hotspot over the last few years. PMW data do not suffer from the cloud cover problem, but do suffer from missing data due to orbital gaps (Wu et al., 2022). Beyond this, LST retrieval with PMW data is an ill-posed problem which is difficult to solve (Gao et al., 2013), so that PMW LST data are not easily obtained. Assimilation data can provide spatially continuous auxiliary information reflecting the real thermal state of the land surface, which is very significant for the TIR LST reconstruction. There are now some public assimilation datasets available, providing convenience for its common use. Examples are the Global Land Data Assimilation System (GLDAS) and the China Land Data Assimilation System (CLDAS), for which the correlation coefficient/root mean square error (RMSE) when compared with ground sites are 0.82/4.2 K and 0.98/1.8 K, respectively (Li et al., 2016; Zhang et al., 2021). This is also the reason why we used the assimilation data in this study. Besides the LST data, the basic physical and geometric properties of the Earth's surface, such as the vegetation coverage, the elevation, and etc., are also very significant for the LST reconstruction and should not be neglected. There are differences between clear and cloudy-sky LST under the same physical/geometric attributes. Combining them with data that can reflect the real thermal state of the surface is considered in this paper. Hence, it can be concluded that even though those multiple sources of data can provide complementary information for the cloud-covered LST reconstruction, they differ in importance, spatial coverage and access. How to select the most reasonable and efficient data combinations from multi-source data, and establish the relationship between the multi-source data and the real LST, are important problems that need to be solved for practical LST reconstruction.

Therefore, in this study, we developed a reconstruction method for Moderate Resolution Imaging Spectroradiometer LST data through combining assimilation data and remote sensing data through the NRN model. Four different data combinations were investigated in depth for precise and practical LST reconstruction. The main contributions of this paper are as follows: 1) The NRN model is proposed to combine multiple sources of data, including remote sensing data and assimilation data, to realize LST reconstruction of cloud-covered regions for MOD11A1 LST data. 2) A grading criterion is introduced to quantitatively weight each auxiliary remote sensing LST data source by comprehensively evaluating its acquisition complexity and its significance to the pending data. 3) The performance of four different data combinations is compared, and the best combination can be selected according to the data conditions, which represents a practical way of gapless LST generation in a large region.

2. Datasets

2.1. Study area

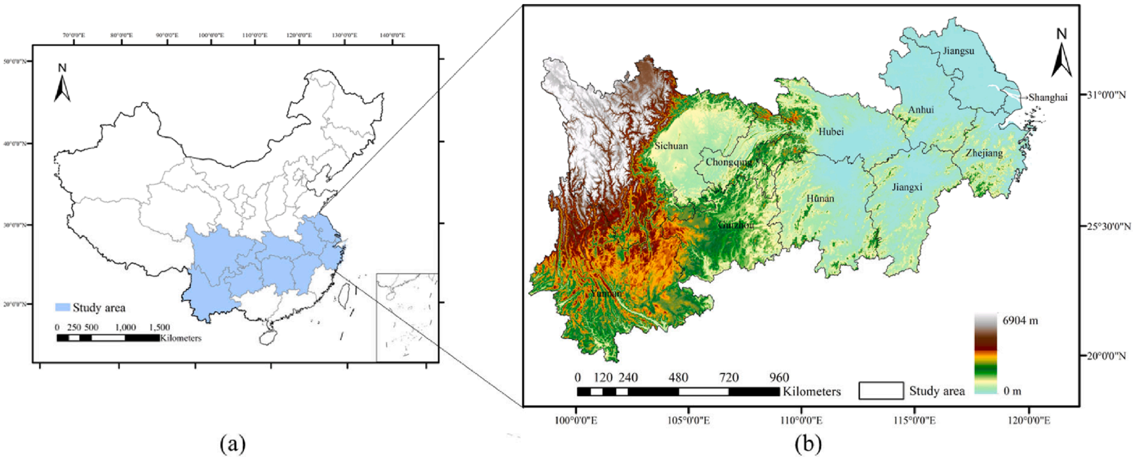
As shown in Fig. 1, the study area is located in the south of China, covering the Yangtze River Economic Belt, including the 11 provinces and cities of Shanghai, Jiangsu, Zhejiang, Anhui, Jiangxi, Hubei, Hunan, Chongqing, Sichuan, Yunnan, and Guizhou. The top-left corner of the study region is 35.13°N, 97.35°E and the bottom-right corner is 21.14°N, 122.84°E. The topography of the study area is complex, including mountains, plateaus, basins, and plains, and the elevation range is widely distributed from 0 to 6904 m. The complex surface conditions can ensure the generality of the proposed method.

2.2. Data collection

All the data utilized in this study are listed in Table 1. There are three types of data: assimilation data, remote sensing data and ground station data, which are described in details below.

2.2.1. Assimilation data

The CLDAS datasets combine observational information obtained by meteorological stations in China and multiple data products (Shi et al., 2011; Zhang et al., 2013). The CLDAS datasets provide high-resolution, high-quality, and near real-time grid data of the surface state and flux parameters. The CLDAS version 2.0 datasets were downloaded from the



(a) (b)

Fig. 1. Location of the Yangtze River Economic Belt in China (a) and its DEM (b).

Table 1
Data collection.

Data type	Data	Spatial/ temporal resolution	Description
Assimilation data	CLDAS DGT	0.0625°/1 day	Daily ground temperature
Remote sensing data	MOD13A2 NDVI	1 km/16 days	16-day normalized difference vegetation index (Aqua)
	MOD11A2 LST	1 km/8 days	8-day land surface temperature (Aqua)
	MYD11A1 LST	1 km/1 day	Daily land surface temperature (Terra)
	MOD11A1 LST	1 km/1 day	Daily land surface temperature (Aqua)
	MOD11B3 LST	6 km/1 month	Monthly land surface temperature (Aqua)
	SRTM DEM	250 m/ -	Elevation
Ground station data	CMA 0-cm LST	—/1 h	Ground 0-cm land surface temperature

China Meteorological Data Service Centre (https://data.cma.cn/). The daily ground temperature (DGT) was used for the reconstruction datasets. Spatial interpolation was used to process the DGT data into a 1-km resolution.

2.2.2. *Remote sensing data*

The remote sensing data utilized in this study included DEM products, NDVI products and four kinds of LST products.

DEM: DEM data are an important factor when attempting to obtain spatio-temporal gapless LST of a coarse-to-fine scale (Zhan et al., 2012; Duan and Li, 2016). The DEM data were downloaded from the Consortium of International Agricultural Research Centers Consortium for Spatial Information (CGIAR-CSI, https://srtm.csi.cgiar.org/index.asp). The elevation variable of the DEM products with a spatial resolution of 250 m was resampled with bicubic interpolation to a 1-km resolution.

NDVI: Many previous studies have demonstrated that NDVI data are effective for reconstructing LST (Zeng et al., 2018). The spatial resolution of the MOD13A2 NDVI product is 1 km. The reason for using a 16-day NDVI product is that there is likely to be no obvious change in the NDVI values within 16 days.

LST: MOD11A1, MYD11A1, MOD11A2, and MOD11B3 LST data were utilized in this study. The first three kinds of data were utilized for the MOD11A1 LST reconstruction. The MOD11B3 LST data were used for the verification of the regional experiment. For these products, the equatorial crossing time of the Terra satellite is about 10:30 a.m., and that of the Aqua satellite is about 13:30p.m.

In this study, the MODIS data were obtained from the Level-1 and Atmosphere Archive and Distribution System Distributed Active Archive Center (https://ladsweb. modaps.eosdis.nasa.gov/search/).

2.2.3. *Ground station data*

The 0-cm LST data from the China Meteorological Administration, were downloaded from the China Meteorological Data Service Centre (http://data.cma.cn/). According to the specifications for surface meteorological observations, the 0-cm LST is measured by a platinum resistance sensor at the automatic observation stations half buried in the soil and half exposed to the air (WMO, 2014). 0-cm LST data from four ground stations were chosen to validated the dependability of the application in regional LST reconstruction. Details of the ground site measurements are provided in Table 2.

3. Methodology

3.1. *Data combinations*

3.1.1. *Datasets preparation*

The aim of this study was to remove the cloud-covered pixels in the MOD11A1 LST product. The missing rate of data (MRD) of the MOD11A1 product in 2019 was calculated, which mainly ranges from 50 % to 90 % and the average MRD is 69.42 %, as shown in Fig. 2. Different masks with different MRDs (50 %, 70 %, and 90 %) were generated to simulate the cloud cover status, and then the dataset comprising of cloudy and clear LST image pairs can be prepared for the deep learning model, as shown in Fig. 3. The white area retains the original MOD11A1 LST, and the black areas simulate missing pixels.

In the process of data clipping, all the images of the study area were included as much as possible, so that the reconstruction datasets covered abundant complex terrain areas. Thus, 3698 pairs of cloudy-sky MOD11A1 images and auxiliary images for the corresponding position with the size of 64 × 64 were prepared, among which 3330 pairs were used for the training and 368 pairs for the testing. Among, them, spatial interpolation was used to process the DGT data into 1 km resolution. The

Table 2
Details of the ground site measurements.

Site	City	Location (° E, °N)	Elevation (m)
No. 56571	Xichang	102.27°, 27.90°	1592.4
No. 56748	Baoshan	99.11°, 25.07°	1652.2
No. 56778	Kunming	102.39°, 25.00°	1888.1
No. 57932	Rongjiang	108.53°, 25.97°	287.4

The detailed information about the ground sites data from the China Meteorological Data Service Centre (http://data.cma.cn).

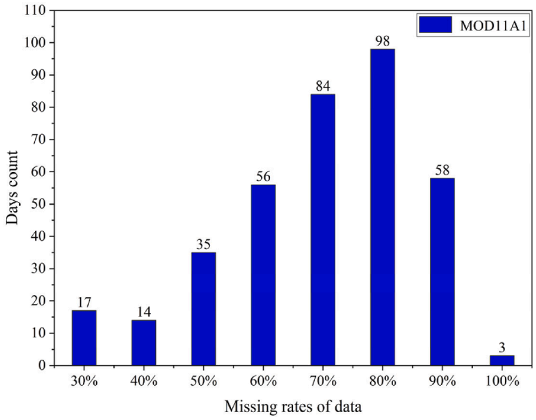


Fig. 2. The missing rate of data (MRD) for the MOD11A1 product in the study region in 2019.

moment matching method is used to normalize CLDAS GST to MOD11A1.

3.1.2. Data grading criterion

A grading criterion was developed to evaluate the importance of the various input auxiliary data by comprehensively considering the missing rate of the auxiliary data and its correlation to the pending data (determination coefficient, R^2). According to the importance of the auxiliary data, the input data were composed of three parts: 1) data that reflect the physical/geometric properties of the ground surface; 2) gapless temperature data that reflect the real-time thermal state of the ground surface; and 3) LST products with complementary information at adjacent times. Among these inputs, the first two parts are necessary inputs, and the third part is selective inputs ranked according to the importance of the data rating.

The surface physical and geometric attribute can be characterized by the NDVI and DEM data, respectively. Many studies have shown that LST and NDVI have a strong correlation (e.g., Sun and Kafatos, 2007; Yue et al., 2007; Tan et al., 2012). Therefore, MOD13A2 NDVI data were used as the necessary data for the reconstruction of the NRN model. DEM data have a powerful influence on the spatial distribution of LST, as LST is affected by elevation and topographic relief. Many studies have found that DEM data have a significant negative correlation with LST (Duan et al., 2017). In addition, SRTM DEM is easy to download, and contains elevation data covering 80 % of the Earth's surface. DEM data are necessary data for an LST reconstruction model.

The gapless temperature used to reflect the real-time thermal state of the surface was the CLDAS DGT, which builds a bridge between the remote sensing data and ground observation data. It has been proved that the relative correlation coefficient between the CLDAS DGT and ground LST is as high as 0.98 in China, and the deviation is less than 2 K (Chen, 2010). The CLDAS DGT data are also gapless, and were thus used as the necessary data in this study.

There are three types of LST products at adjacent times: MYD11A1, MOD11A1, and MOD11A2. The instantaneous observation of LST is not independent, but has a certain correlation in the time series, which can provide rich information for reconstruction. If we suppose that the acquisition time of the missing MOD11A1 product is D_1T_1 (day 1, time 1), then the available LST auxiliary data are $D_{1\sim8}T_1$ MOD11A2, D_1T_2 MYD11A1, and D_2T_1 MOD11A1. As shown in Fig. 4, the MRD for the three auxiliary LST products was calculated for the study area in 2019. Clearly, the MRD of the MOD11A2 product is the lowest, and that of MOD11A1 is close to that of MYD11A1. The R values for these three LST products with the D_1T_1 MOD11A1 for each day in 2019 are presented in Fig. 5. On the whole, the R values are, in descending order, $D_{1\sim8}T_1$

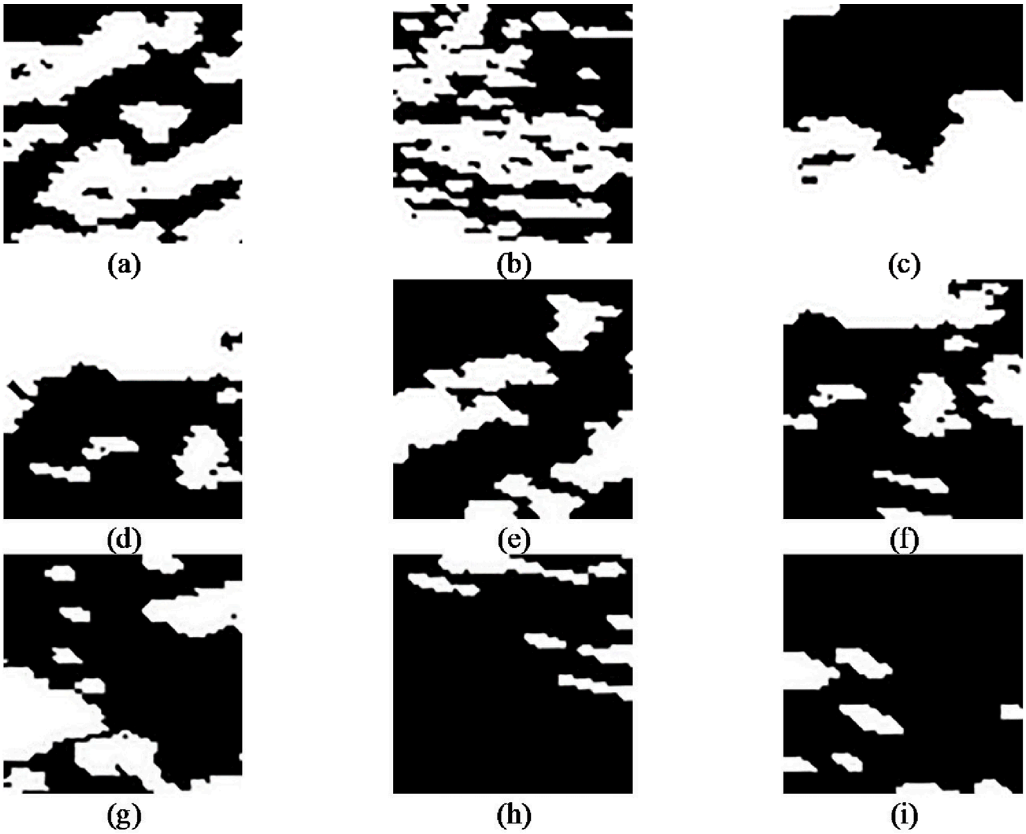


Fig. 3. Masks (size: 64 × 64) with different MRDs for the missing observations. The white and the black pixels represent the clear-sky areas and the missing areas, respectively. (a)–(d) 50 % MRD. (e)–(g) 70 % MRD. (h) and (i) 90 % MRD.

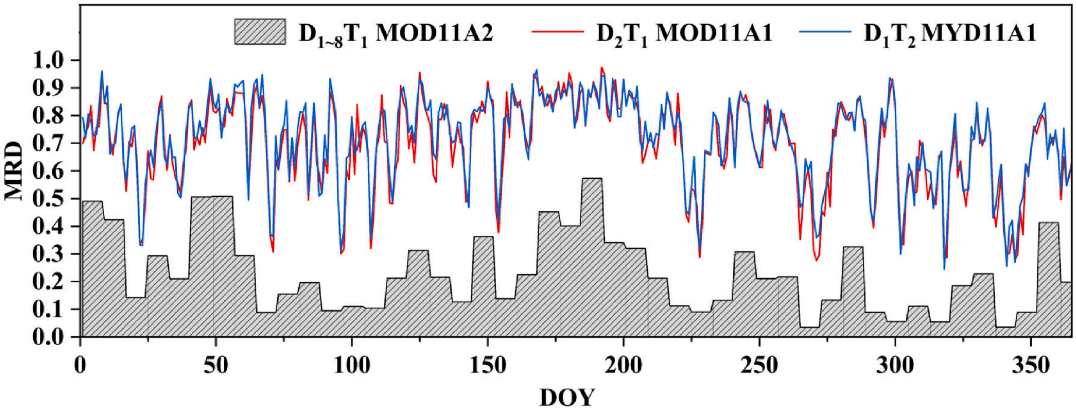


Fig. 4. The MRD for the various auxiliary LST products.

MOD11A2, D_1T_2 MYD11A1, D_2T_1 MOD11A1.

The average statistical MRD and R values for 2019 are listed in Table 3. That the MRD from low to high is $D_{1\sim8}T_1$ MOD11A2 $<$ D_1T_2 MOD11A1 $<$ D_2T_1 MYD11A1, and the R from strong to weak is $D_{1\sim8}T_1$ MOD11A2 $>$ D_1T_2 MYD11A1 $>$ D_2T_1 MOD11A1.

The grading criterion results for the three kinds of auxiliary data are listed in Table 4. Clearly, the total score of MOD11A2, MYD11A1, and MOD11A1 rank from high to low by the importance score.

Based on the above analysis, we scored four input combinations according to their importance grading criterion. Among these combinations, Combination (Com) 1 is composed of the necessary data, and Coms 2, 3, and 4 add the relevant LST variables in sequence according to the scoring standard on the basis of Com 1. The following auxiliary data combinations were considered in the experiments:

Com 1: MOD13A2 NDVI + SRTM DEM + CLDAS DGT.

Com 2: MOD13A2 NDVI + SRTM DEM + CLDAS DGT + MOD11A2 LST.

Com 3: MOD13A2 NDVI + SRTM DEM + CLDAS DGT + MOD11A2 LST + MYD11A1 LST.

Com 4: MOD13A2 NDVI + SRTM DEM + CLDAS DGT + MOD11A2 LST + MYD11A1 LST + MOD11A1 LST (MOD11A1 LST is the adjacent temporal LST of the pending MOD11A1 LST).

Table 3
Average statistical MRD and R for each day in 2019.

	$D_{1\sim8}T_1$ MOD11A2	D_1T_2 MYD11A1	D_2T_1 MOD11A1
MRD	22.90 %	71.21 %	69.42 %
R	0.9241	0.8769	0.7757

Table 4
Importance scores for the various LST data sources.

Data	Correlation to pending data (0–1)	Probability of acquiring gapless data (0–1)	Total score (0–2)
MOD11A2 LST	0.92	0.77	**1.69**
MYD11A1 LST	0.88	0.29	**1.17**
MOD11A1 LST	0.78	0.31	**1.09**

3.2. *The nonlocality-reinforced network model for cloud removal of LST data*

The proposed nonlocality-reinforced network (NRN) model is a U-shaped structure of multiscale encoding–decoding, as shown in Fig. 6. In the encoding structure, a nonlocality-reinforced module is used to enhance the spatial cognition and learning ability of the network, which can effectively learn the longer distance spatial correlation in multi-source data, reconstructing more accurate cloudy-sky LST. The multiscale encoding–decoding structure can increase the receptive field and reduce the amount of computation (Ronneberger et al., 2015), and the skip connection structure can effectively reduce the gradient disappearance and degradation problems (Szegedy et al., 2016). The nonlocality-reinforced module and loss function of the NRN model are described in detail below.

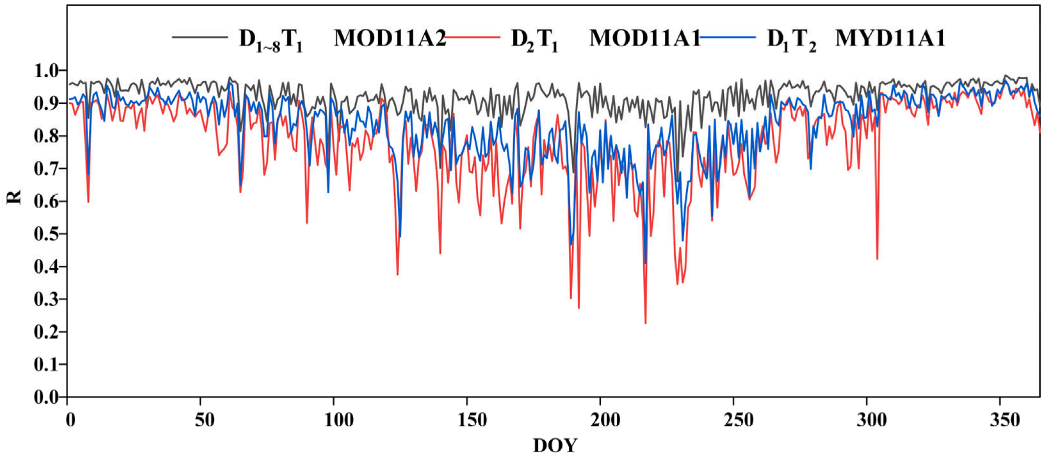


Fig. 5. R between the MOD11A1 product to be reconstructed and the various auxiliary LST products.

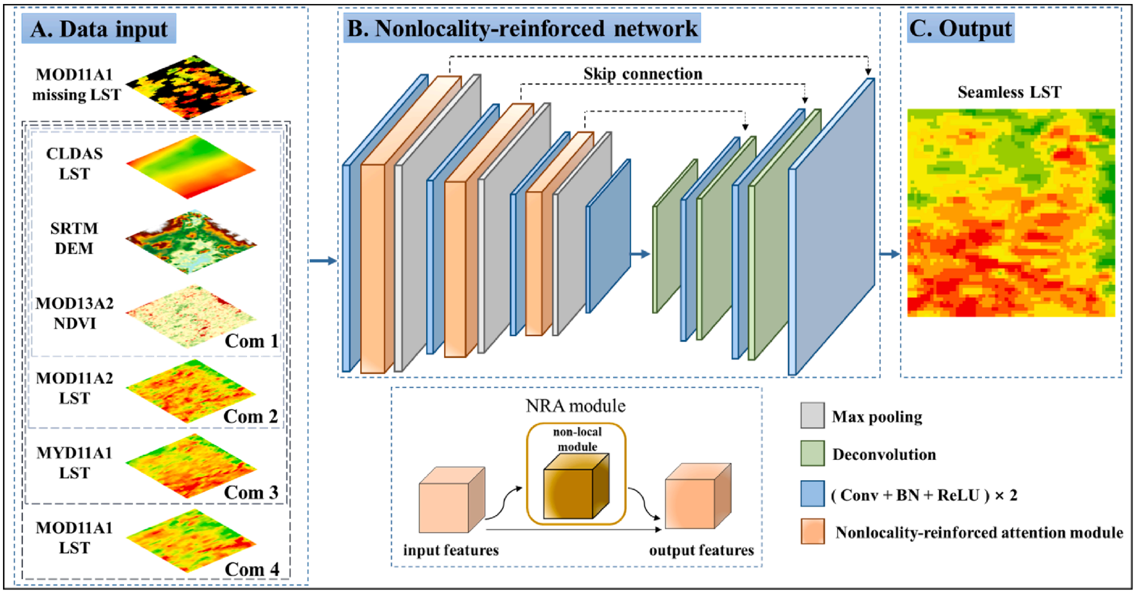


Fig. 6. Flowchart of the proposed method: A. Data input, B. Nonlocality-reinforced network, and C. Output.

3.2.1. *Nonlocality-reinforced module*

LST is a spatially continuous and autocorrelated variable. Three nonlocality-reinforced modules are embedded into the cloud removal network to better learn the global spatial features and those of the multi-source auxiliary data (Wang et al., 2018; Woo et al., 2018). In a nonlocality-reinforced module, the response in one location is computed as the weighted sum of the features in all the locations, thus overcoming the shortcoming of convolution processing paying attention to local information. The nonlocality-reinforced module structure is shown in Fig. 7.

In the nonlocality-reinforced module, global feature points are considered for the weighting calculation, which expands the receptive field. The expression is as follows:

$$X_{LST_i}^{NRA} = \frac{1}{C(X_{LST}^{C})} \sum_{\forall j} f(X_{LST_i}^{C} - X_{LST_j}^{C}) g(X_{LST_j}) \tag{1}$$

where X_{LST}^{C} is the output LST signal of the previous step; i represents any position of the output LST feature map; j is the index of all the locations; X_{LST}^{NRA} is the output LST feature map; CX_{LST}^{C} is a normalized function; and $f(x)$ is a function to calculate the feature's similarity between position i and position j. In the proposed method, an embedded Gaussian function, which is a simple variant of the Gaussian function, is used as $f(x)$ to calculate the similarity in embedded space.

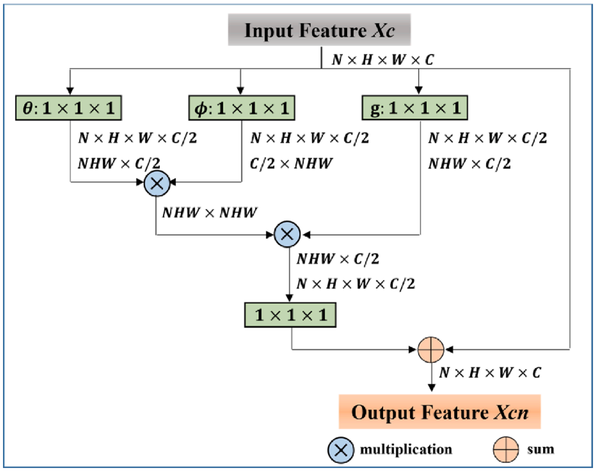


Fig. 7. The structure of the nonlocality-reinforced module.

3.2.2. *Loss function*

The L2-norm and the Structural SIMilarity (SSIM) index are combined as the loss function of the NRN model. The L2-norm loss function is not only derivable everywhere, but is also easy to optimize, and has a relatively stable solution. However, the L2-norm is based on pixel-by-pixel comparison that particularly sensitive to outliers. The SSIM index considers the similarity of two LST images, which is counteract the influence of individual outliers on the overall training of the network (Wang et al., 2004).

The L2-norm loss function measures the Euclidean distance between the true LST and the reconstructed LST of the deep learning networks (Zhang et al., 2017). The mathematical formula is as follows:

$$Loss_{L2-norm} = \frac{1}{N} \sum_{i}^{N} \left(LST_{re(i)} - LST_{es(i)}\right)^2 \tag{2}$$

where LST_{es} and LST_{re} represent the estimated LST value and real LST value, respectively. i represents a point of the LST images. N represents the number of points of the LST images.

The larger the SSIM index value is, the more similar the two LST images are. When the two LST images are identical, the SSIM index value is 1. The SSIM index is calculated as follows:

$$\mathrm{SSIM}(LST_{re} - LST_{es}) = \frac{(2\mu_{LST_{re}}\mu_{LST_{es}} + c_1)(2\sigma_{LST_{re}LST_{es}} + c_2)}{(\mu_{LST_{re}}^2 + \mu_{LST_{es}}^2 + c_1)(\sigma_{LST_{re}}^2 + \sigma_{LST_{es}}^2 + c_2)} \tag{3}$$

$$Loss_{SSIM} = 1 - SSIM \tag{4}$$

where LST_{es} and LST_{re} represent the estimated LST value and real LST value, respectively. $\mu_{LST_{re}}$ and $\mu_{LST_{es}}$ represent the average values of the estimated LST and the real LST, respectively. $\sigma_{LST_{es}}$ and $\sigma_{LST_{re}}$ are the standard deviations of the pixel values of the estimated LST and the real LST, respectively. $\sigma_{LST_{re}LST_{es}}$ is the covariance between the estimated LST and the real LST. c_1 and c_2 are constants.

4. Results

4.1. Examination of the different data combinations for LST reconstruction

4.1.1. Visual assessment

In all four input combinations, the proposed reconstruction method can fully fill the invalid LST pixels caused by cloud coverage, as shown in Fig. 8. Three test data were randomly selected for different terrain in different areas of the study area over different days. The first column shows the missing LST simulated by the reference LST, and the middle four columns are different combinations reconstructed LST. More specifically, in the three samples, Com 4 works best among all the combinations, and its reconstruction result is the closest to the real LST, with the best overall details. The reconstruction results of Com 1 have rough spatial details, and are the worst one among the four combinations, showing some deviation between the reconstructed LST and original real LST. Com 1 has the lowest difficulty of data acquisition, promoting the broad prospect of large-scale application, and is a cost-effective reconstruction method. The reconstruction results of Com 2 and Com 3 are between the best and the worst, and the results of Com 3 have richer detail consistency than those of Com 2. In addition, in the four reconstruction results, with the increase of the MRD, the reconstruction visual effect shows fewer and fewer details. With the exception of the three examples shown in Fig. 8, all of the reconstructed visual effects in the dataset behave similarly.

4.1.2. Quantitative assessment

To evaluate the accuracy of the four combinations for LST reconstruction, the quantitative statistical accuracies with different MRDs for the NRN model were compared. The three evaluation indices considered in this paper are: mean absolute error (MAE), RMSE, and R^2. Among the indices, a smaller MAE, a smaller RMSE, and a larger R^2 indicate better reconstruction results.

Fig. 9 shows the quantitative statistics of the NRN model for the four combinations of MOD11A1 images, where red represents MAE, gray represents RMSE, and blue represents R^2. It can be seen that the overall accuracy evaluation results are consistent with the visual effects. Com 4 obtains the best reconstruction results, with an MAE of 0.5219 K, an RMSE of 0.7622 K, and an R^2 value of 0.8956. The deviation and absolute error are also the smallest between the reconstruction LST and real LST. Com 3 obtains an effect that is second only to that of Com 4, and is slightly better than that of Com 2. Although Com 1 acquires the worst reconstruction result, the R^2 is higher than 0.78 and the MAE is less than 0.78 K, maintaining reasonable consistency with the real LST.

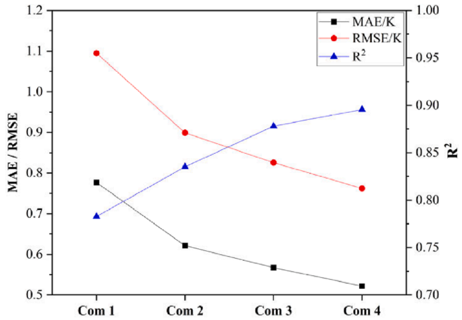


Fig. 9. Quantitative statistics of the NRN model for the four combinations of MOD11A1 images.

To analyze the influence of different MRDs and combinations in more detail, Fig. 10 shows the quantitative statistics of the NRN model for the four combinations of MOD11A1 images with different MRDs. With the increase of the MRD, the reconstruction results all become worse. When only the necessary auxiliary data are used as inputs (Com 1), NRN achieves a relatively poor accuracy, with an RMSE of less than 1.5 K at 90 % MRD. With the addition of more remote sensing multi-temporal LST data (Coms 2–4), the accuracy of the reconstruction results is significantly improved, with a closer consistency with the true LST. This is due to the space–time complementary correlation between the multi-temporal LST products. The quantitative assessment results are provided in Table 5.

Fig. 11 shows the error distribution of the four combinations. All the combinations are all basically normally distributed, and the errors are all distributed from −5 K to 5 K. The errors are mainly concentrated in the −2 K to 2 K range, with peaks close to 0 K, and a few errors with larger values. The mean value of error distribution is very close to 0 K. With the increase of auxiliary data variables, the standard deviation also

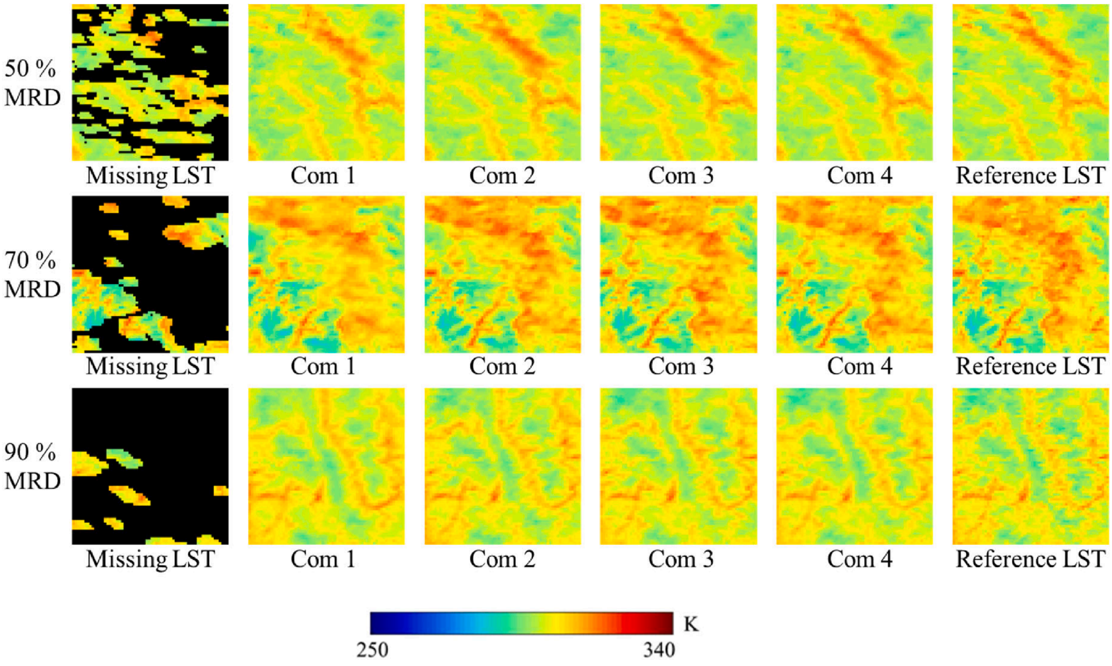


Fig. 8. LST reconstruction results obtained using the NRN model for MOD11A1 during different situations.

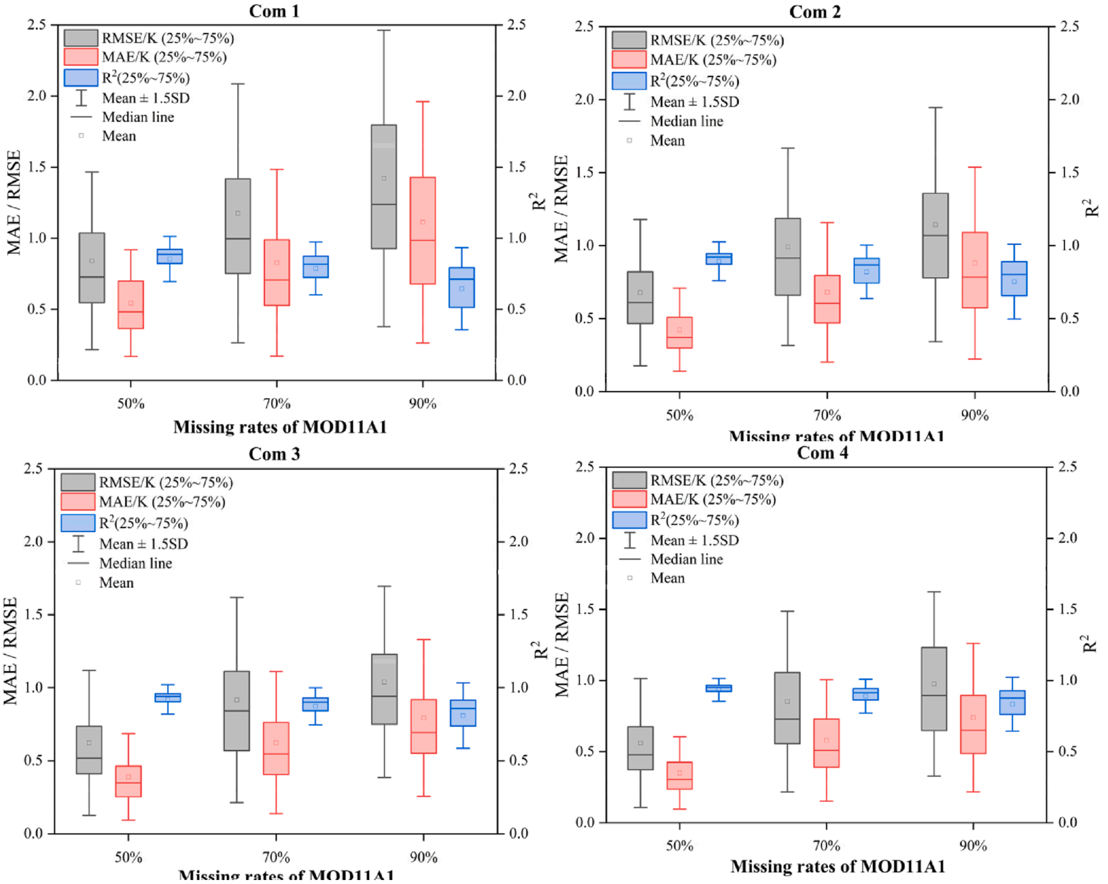


Fig. 10. Quantitative statistics of the NRN model for the four combinations of MOD11A1 images with different missing rates.

Table 5
Quantitative statistics for the reconstruction results of the four combinations with different missing rates.

Indicator	MRD	Com 1	Com 2	Com 3	Com 4
MAE/K	50 %	0.5442	0.4248	0.3895	0.3501
	70 %	0.8285	0.6805	0.6249	0.5784
	90 %	1.1124	0.8809	0.7936	0.7395
	Average	**0.7767**	**0.6216**	**0.5672**	**0.5219**
RMSE/K	50 %	0.8414	0.6780	0.6231	0.5611
	70 %	1.1756	0.9924	0.9168	0.8524
	90 %	1.4210	1.1447	1.0406	0.9767
	Average	**1.0950**	**0.8990**	**0.8255**	**0.7622**
R^2	50 %	0.8541	0.8930	0.9206	0.9345
	70 %	0.7878	0.8209	0.8727	0.8901
	90 %	0.6451	0.7534	0.8094	0.8342
	Average	**0.7828**	**0.8351**	**0.8780**	**0.8956**

decreases, and the peak of the curve becomes higher and the distribution is more concentrated, which means that the errors are smaller. This shows that the multiple dataset combination method proposed in this paper provides rich auxiliary information and can improve the accuracy of LST reconstruction.

4.2. Comparison with other methods

To investigate the NRN model performance, the traditional method of the harmonic analysis of time series (HANTS) (Xu and Shen, 2013), improved ESTARFM method (Long et al., 2020) and the deep learning method of MFCTR-CNN (Wu et al., 2019) were compared with the proposed method. The spatial and temporal fusions and bias-corrected LST are used in the improved ESTARFM reconstruction method, and MFCTR-CNN uses a multiscale CNN combined with spatial attention to learn high-level features of the adjacent-time LST for reconstruction.

The quantitative statistical results of these four methods are listed in Table 6. Except that the R^2 values of improved-ESTARFM is lower than 0.5, the other three methods are higher than 0.7. There are significant differences in MAE and RMSE. The MAE of the improved-ESTARFM is greater than 3.5 K, while that of HANTS, MFCTR-CNN and the NRN model are less than 1 K. 1) The improved-ESTARFM, although systematically bias corrected for the results of the spatio-temporal fusion of CLDAS GST and MOD11A1, still has a large bias with original MOD11A1. The improved-ESTARFM method only considers simple linear relationships in spatio-temporal, making it difficult to achieve better results. The HANTS methods only used time series LST information, achieving an acceptable result, without taking into account the LST difference between clear and cloudy-sky. The other two methods take into account the spatio-temporal complementarity with good results, while deeply mining its characteristics. 2) The NRN model (Com 1) obtains the best results in the four evaluation indicators. This is probably due to the HANTS, the improved ESTARFM and the MFCTR-CNN methods only considering LST as auxiliary data, while the NRN model also considers the data reflecting physical/geometric attributes (NDVI/DEM) and the real-time thermal state (CLDAS GST) of the surface as necessary auxiliary data.

To verify the capabilities of the NRN structure, Unet (Ronneberger et al., 2015) and MFCTR-CNN were compared in this study. Unet is a three-times down-up sampling structure, and MFCTR-CNN adds the spatial attention structure to upsampling (named UnetSA in this paper), while the NRN model adds a nonlocality-reinforced module in the downsampling.

As shown in Fig. 12, with the four input combinations, the test results of UnetSA and the NRN model are improved, compared with the Unet structure, but the NRN model shows a more significant improvement.

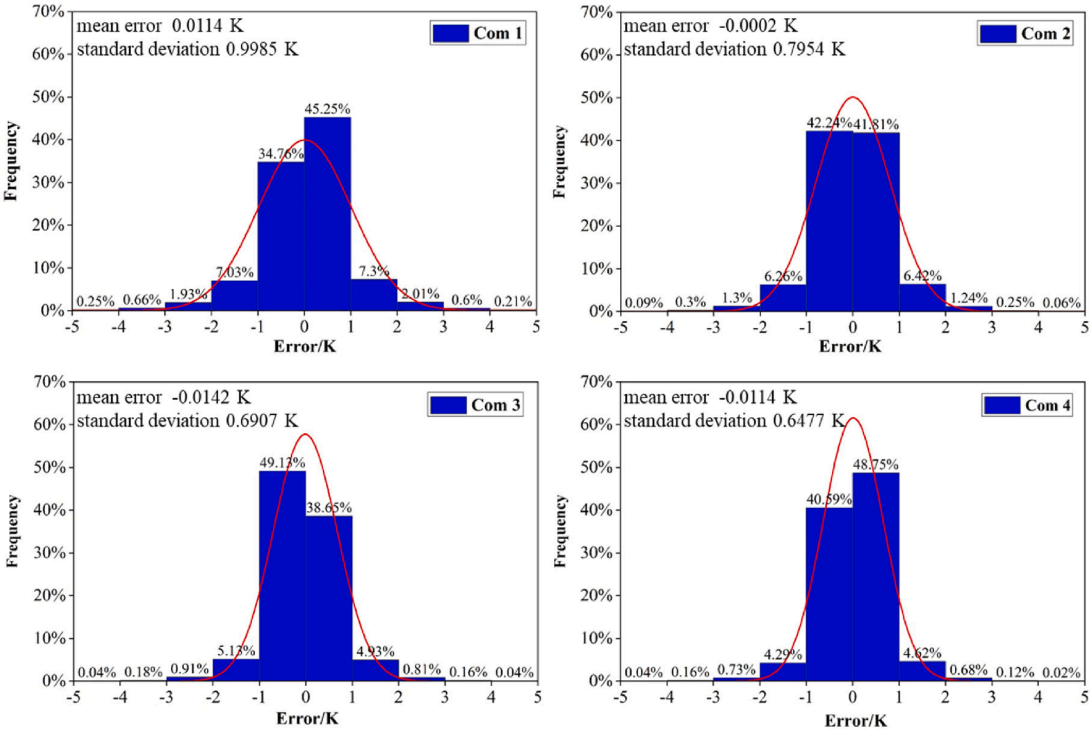


Fig. 11. The error distributions of the four combinations.

Table 6
Quantitative statistics for the reconstruction results of the four methods.

Indicator	Improved-ESTARFM	HANTS	MFCTR-CNN	Proposed (Com 1)
MAE/K	3.3695	1.8134	0.8013	**0.7767**
RMSE/K	4.9385	2.3825	1.1455	**1.0950**
R^2	0.4683	0.7238	0.7630	**0.7828**

Firstly, the spatial attention module enhances the weight of the important features, but is still a local calculation, with a small receptive field. The nonlocality-reinforced module not only increases the weight of the important features, but also solves the limitation of the CNN paying too much attention to the local features. Secondly, UnetSA pays attention to the high-level features and achieves certain results. The NRN model uses the nonlocality-reinforced module for the downsampling, paying more attention to the features obtained from the input combinations of data rather than the features obtained after multiple convolution transformations. These points fully prove the superiority of the nonlocality-reinforced structure and its insertion position. In addition, with the increase of the optional data, the accuracy of all three networks is significantly improved, which verifies the generality and expansibility of the various combinations proposed in this paper.

4.3. Application to regional MODIS LST reconstruction

In Section 3.1, we described how the four input combinations provided the possibility of reconstructing LST for the whole study area in 2019. In Com 1, the input data only include the pending MOD11A1 LST, MOD13A2 NDVI, SRTM DEM and CLDAS DGT data, for which gapless data are easily acquired. To validate the performance of the NRN model in a regional experiment, the time-series results were compared with the ground-based 0-cm LST data from the China Meteorological Administration. Data from four ground stations were selected for the verification in homogeneous areas.

4.3.1. Regional LST mapping

The spatial distribution pattern of the daily missing LST and the reconstructed LST products of the study area in different seasons in 2019 are shown in Fig. 13. The left side shows the daily LST missing images in winter, spring, summer and autumn, respectively, and the right side shows the reconstruction results of Combination 1. Obviously, the Combination 1 can be reconstructed to seamless LST of the proposed method. If there were more optional auxiliary data, a more accurate reconstructed LST is available. A more accurate evaluation of the regional reconstruction results was done later using ground 0-cm LST data.

4.3.2. Accuracy validation

Fig. 14 shows the scatter plot distribution of the 0-cm LST and MODIS LST at four ground stations for the time-series data in 2019, where the orange and blue dots represent the original and the reconstructed LST, respectively. The correlation between the 0-cm LST and MOD11A1 before and after reconstruction is very close, and the absolute bias value is less than 0.9976 K. Excluding their systematic error, the data distribution before and after reconstruction maintains a good consistency using the proposed method, with no abnormal values, proving that the proposed NRN model has good stability and reliability.

5. Discussions

5.1. Advantages of the proposed method

To address the problem of the missing pixels caused by cloud coverage of MODIS LST data, based on the analysis of multi-source data correlation and missing rate, four data combinations are established to provide complementary information in this paper. A nonlocality-reinforced network is constructed to mine the complex nonlinear complementary mapping of multi-source data to achieve high-precision reconstruction of LST. The advantages of the proposed method compared with existing methods are mainly reflected in the following three aspects: 1) Compared with the traditional time-series (Xu et al.,

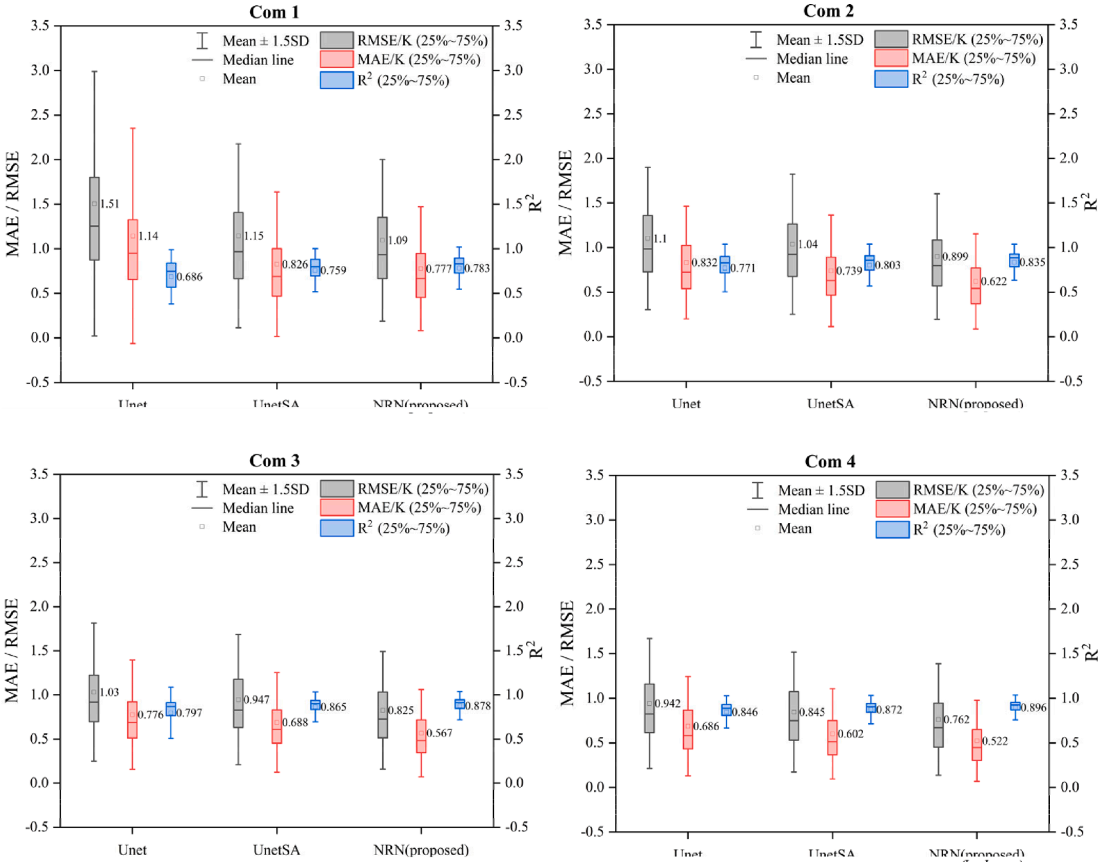


Fig. 12. Quantitative statistics for the LST reconstruction of the three networks during four combinations.

2013) and spatio-temporal fusion reconstruction methods (Long et al., 2020), the reconstruction accuracy of the proposed method in this paper is significantly improved due to the introduction of multi-source data and deep learning methods, as shown in Table 6. And the analysis of different data combinations also makes the proposed method more practical. 2) Compared with the LST reconstruction methods using the assimilated data, the accuracy of our method can reach the RMSE of 0.7622 K, which is more accurate compared with the annual/diurnal temperature cycle-based framework (Hong et al., 2021; Ma et al., 2022). Because the convolutional neural network can deeply mine the spatial features of the auxiliary data and learn the nonlinear relationships among multi-source auxiliary data. 3) Compared with existing deep learning methods, the proposed method introduces nonlocality-reinforced module, which enables the network to better take into account the spatial distribution characteristics of LST, and the reconstruction accuracy is higher. As shown in Fig. 12, the proposed method in this paper is more accurate in RMSE, that the MAE and R^2 are the highest with the same data combination as input.

5.2. Extensions of the proposed method

It is known that deep learning networks are highly dependent on datasets, resulting in limited generalization and extension. To testify the generalization of the proposed method, the trained model with 2019 data is used for the LST reconstruction in the other spatial areas and captured in the other years. In this case, spatial adaptation analysis was tested using data from other regions (36.08°N, 125.11°E–43.58°N, 96.81°E) in 2019, and temporal adaptation analysis was tested using data from the study area in spring 2018, both using 500 pairs of test data.

The results of the comparative experimental test accuracy of the four combinations are shown in Table 7. It can be seen that the NRN model has the highest test accuracy on the original data set. The test accuracy decreases in other times and spaces, but acceptable results can be obtained. In the construction of the training dataset, we have tried to make the data pairs cover various land cover types and various seasons as much as possible, so that NRN can be applied to more regions and times.

5.3. Limitations of the proposed method

Datasets are of great significance to deep learning methods. Only one year’s data of the study area was used to establish a dataset in this paper. Although data from other regions and time can also be processed, the accuracy still needs to be further improved. In order to improve the spatio-temporal generalization ability of the NRN method, future work should expand the temporal domain of the dataset and make the dataset balanced across land cover types. To make full use of the auxiliary data with gaps and introduce more assimilation data at different time may be an effective method, which is what we will discuss further in the future.

Moreover, the temporal difference of the auxiliary data used in this study might bring uncertainty to the final result. The merging of TIR- and passive microwave (PMW)-based LST is regarded as the most promising option to solve this problem and develop an all-weather LST product (Li et al, 2023), even though the thermal sampling depths of these two types of data are not equal. In addition to that, there exist large scale differences between TIR- and PMW-based LST. Deep learning models have the ability to learn complex nonlinear correlations, which may benefit the exploration of the physical mechanisms and spatial patterns, and support the merging of these two different sourced temperature data. It is a challenging and very valuable direction which worth further studying in the future.

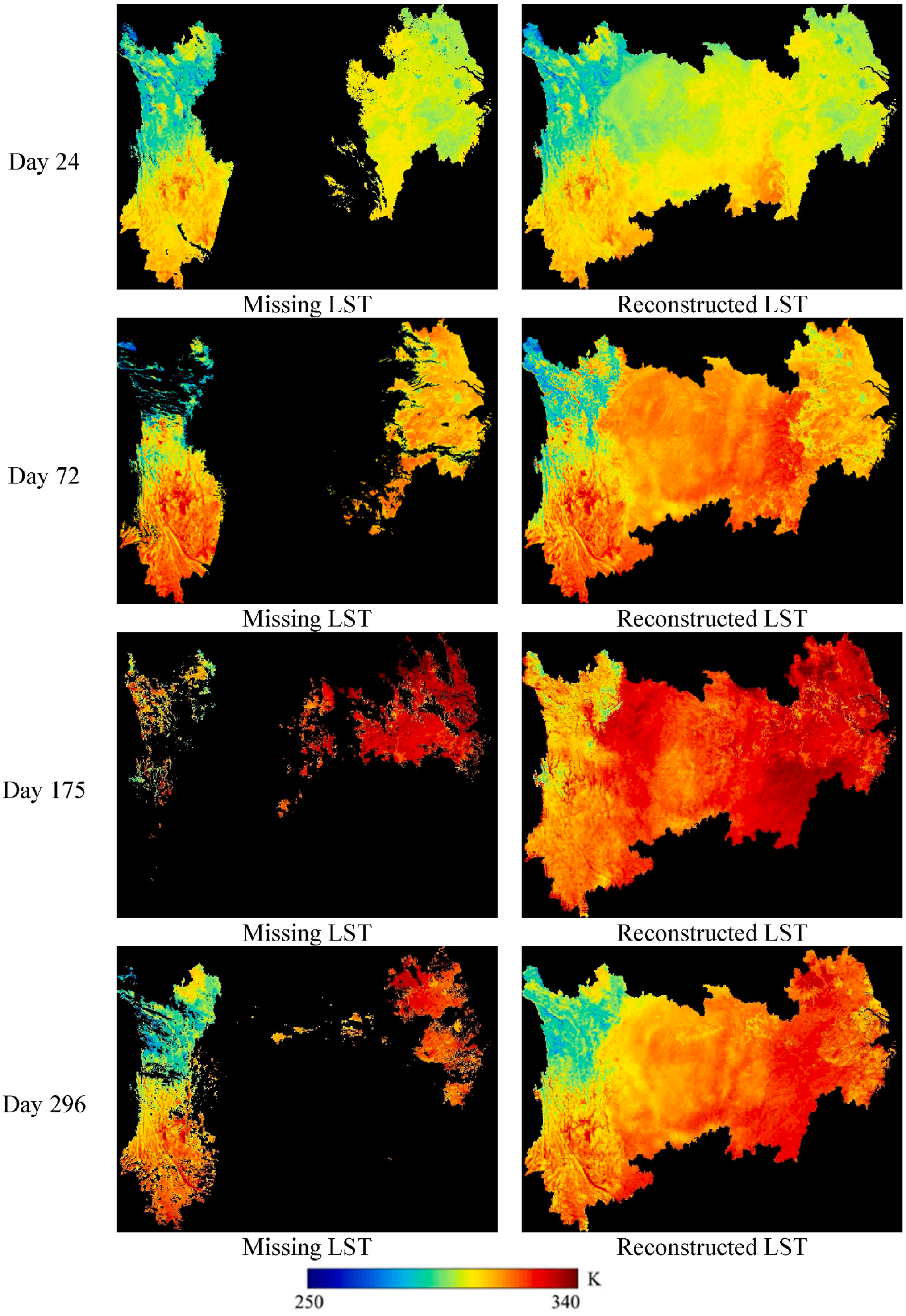


Fig. 13. Spatial distribution pattern of the daily missing LST and the reconstructed LST products of the study area in different seasons in 2019.

6. Conclusions

In this paper, we have described how assimilation data and remote sensing data can be combined through the proposed nonlocality-reinforced deep learning network to achieve cloudy-sky LST reconstruction. Four specific data combinations were proposed according to the data grading criterion, by considering both the significance and acquisition of the data. The results showed that the NRN model can yield good results even using the basic data combination (Com 1), indicating that the combination of assimilation data and remote sensing data is very effective for LST reconstruction. Quantitatively, the reconstruction accuracy was the highest when using the full data combination (Com 4), owing to the injection of more effective information. The proposed method was compared to the traditional the harmonic analysis of time series method, the improved ESTARFM method and some other deep learning methods. A multi-source dataset was shown to be more effective than only using remote sensing data. The comparison with other deep learning networks proved the superiority of the nonlocality-reinforced structure and its insertion position. In the practical regional LST reconstruction, the reconstructed monthly average LST with a 1-km

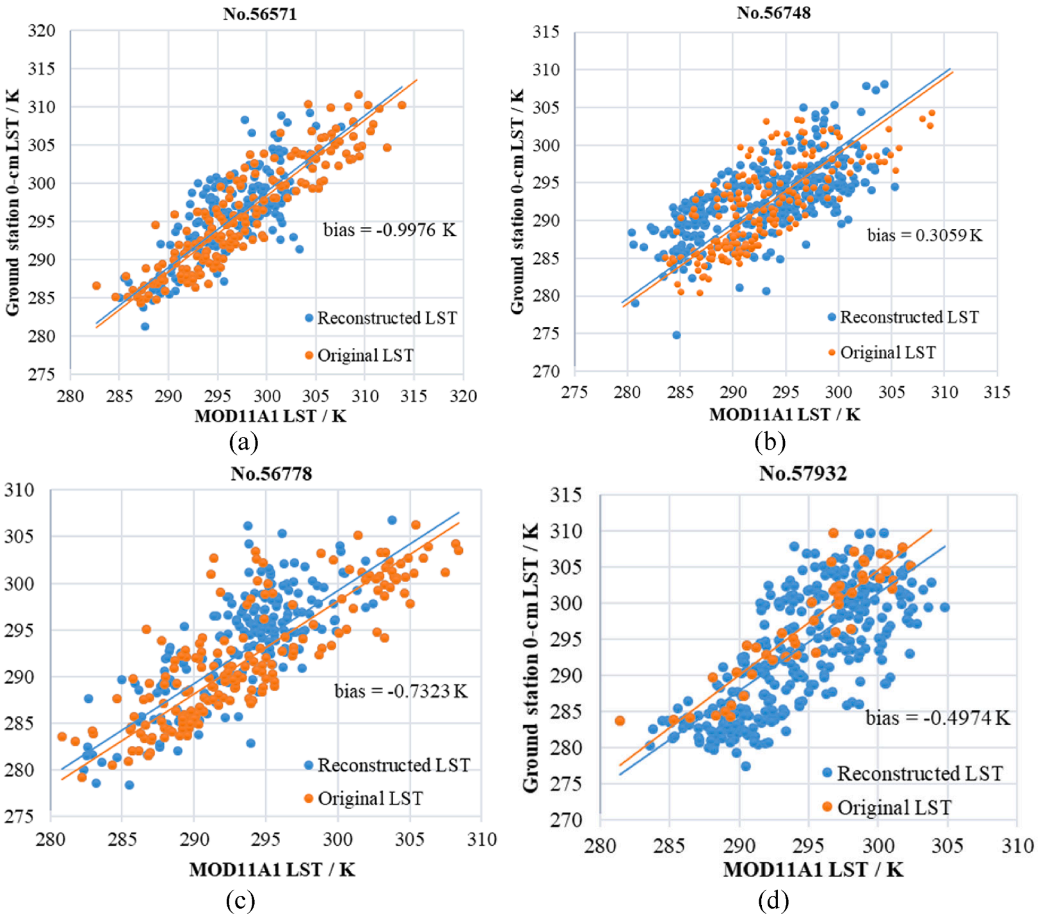


Fig. 14. Scatter plots of the time-series LST against the 0-cm LST for the four selected ground stations: (a) No. 56571, (b) No. 56748, (c) No. 56778, and (d) No. 57932.

Table 7
Quantitative statistics for the LST reconstruction of the original dataset, the data covering other spatial areas and the data captured at other time.

Combinations	Indicator	Original datasets	Data covering other areas	Data captured at other time
Com 1	MAE/K	0.7767	0.9629	1.5219
	RMSE/K	1.0950	1.3433	2.0919
	R^2	0.7828	0.6277	0.6936
Com 2	MAE/K	0.6216	0.7647	1.0364
	RMSE/K	0.8990	1.0846	1.4245
	R^2	0.8351	0.7498	0.8248
Com 3	MAE/K	0.5672	0.5943	0.9047
	RMSE/K	0.8255	0.8747	1.2917
	R^2	0.8780	0.8390	0.8475
Com 4	MAE/K	0.5219	0.5867	0.8801
	RMSE/K	0.7622	0.8503	1.2593
	R^2	0.8956	0.8582	0.8625

resolution was consistent with the official monthly MOD11B3 LST with a 6-km resolution, in both the spatial and temporal domains. This validated the practicability and efficiency of the proposed method. The proposed method actually represents an open framework that can be used to combine more data from other sources, and the subsequent performance in LST reconstruction needs further investigation.

CRediT authorship contribution statement

Yuting Gong: Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Huifang Li:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing. **Huanfeng Shen:** Methodology, Project administration. **Chunlei Meng:** Data curation, Writing – review & editing. **Penghai Wu:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was funded by the National Key Research and Development Program of China (2022YFF1301103) and the National Natural Science Foundation of China (41871246).

References

Aires, F., Prigent, C., Rossow, W.B., 2004. Temporal interpolation of global surface skin temperature diurnal cycle over land under clear and cloudy conditions. J. Geophys. Res.. Atmos. 109.

Arslan, N., Sekertekin, A., 2019. Application of Long Short-Term Memory neural network model for the reconstruction of MODIS Land Surface Temperature images. J. Atmos. Sol.-Terrestrial Phys. 194.

Chen, H., 2010. Simulation of land surface processes over China and its validation part: soil temperature. Sci. Meteorol. Sin.

Chen, M., Newell, B.H., Sun, Z., Corr, C., Gao, W., 2019. Reconstruct missing pixels of Landsat land surface temperature product using a CNN with partial convolution. SPIE-Intl. Soc. Opt. Eng. 11.

Cheng, Y., Li, Y., Wu, H., Li, F., Li, Y., He, L., 2020. Reconstructing Modis Lst Products Over Tibetan Plateau based on Random Forest. In: GARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 6226–6229.

Crosson, W.L., Al-Hamdan, M.Z., Hemmings, S.N.J., Wade, G.M., 2012. A daily merged MODIS Aqua-Terra land surface temperature data set for the conterminous United States. Remote Sens. Environ. 119, 315–324.

Di, W.u., Chen, J., Shi, M., Qin, B., Shengyang, L.I., 2019. Reconstruction of land surface temperature time-series datasets of FY-2F based on Savitzky-Golay filter. Remote Sens. L. Resour.

Duan, S.-B., Li, Z.-L., 2016. Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: case study in Northern China. IEEE Trans. Geosci. Remote Sens. 54, 6458–6469.

Duan, S.B., Li, Z.L., Leng, P., 2017. A framework for the retrieval of all-weather land surface temperature at a high spatial resolution from polar-orbiting thermal infrared and passive microwave data. Remote Sens. Environ. 195, 107–117.

Gao, C., Jiang, X., Qian, Y., Qiu, S., Ma, L., Li, Z.L., 2013. A neural network based method for land surface temperature retrieval from AMSR-E passive microwave data. Int. Geosci. Remote Sens. Symp. 469–472.

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. Rev. Geophys. 48.

Hong, F., Zhan, W., Göttsche, F.-M., Lai, J., Liu, Z., Hu, L., Fu, P., Huang, F., Li, J., Li, H., Wu, H., 2021. A simple yet robust framework to estimate accurate daily mean land surface temperature from thermal observations of tandem polar orbiters. Remote Sens. Environ. 264, 112612.

Jia, A., Ma, H., Liang, S., Wang, D., 2021. Cloudy-sky land surface temperature from VIIRS and MODIS satellite data using a surface energy balance-based method. Remote Sens. Environ. 263.

Kustas, W., Anderson, M., 2009. Advances in thermal infrared remote sensing for land surface modeling. Agric. For. Meteorol. 149, 2071–2081.

Li, Z.L., Tang, B.H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F., Sobrino, J.A., 2013. Satellite-derived land surface temperature: current status and perspectives. Remote Sens. Environ. 131, 14–37.

Li, Z., Wu, H., Duan, S., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X., Zhu, J., Sun, Y., Si, M., Liu, M., Li, J., Zhang, X., Shang, G., Tang, B., Yan, G., Zhou, C., 2023. Satellite remote sensing of global land surface temperature: definition, methods, products, and applications. Rev. Geophys. 61, 1–116.

Li, X., Zhou, J., Chen, Y., Huang, J., Wang, L., Liu, W., Yang, K., Chen, D., Qi, J., 2016. Validation of the global land data assimilation system based on measurements of soil temperature profiles. Agric. For. Meteorol.

Liu, Z., Wu, P., Duan, S., Zhan, W., Ma, X., Wu, Y., 2017. Spatiotemporal reconstruction of land surface temperature derived from FengYun geostationary satellite data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4531–4543.

Long, D., Yan, L., Bai, L., Zhang, C., Li, X., Lei, H., Yang, H., Tian, F., Zeng, C., Meng, X., Shi, C., 2020. Generation of MODIS-like land surface temperatures under all-weather conditions based on a data fusion approach. Remote Sens. Environ. 246, 111863.

Lu, L., Venus, V., Skidmore, A., Wang, T., Luo, G., 2011. Estimating land-surface temperature under clouds using MSG/SEVIRI observations. Int. J. Appl. Earth Obs. Geoinf. 13, 265–276.

Lyon, S.W., Sorensen, R., Stendahl, J., Seibert, J., 2010. Using landscape characteristics to define an adjusted distance metric for improving kriging interpolations. Int. J. Geogr. Inf. Sci. 24, 723–740.

Ma, J., Shen, H., Wu, P., Wu, J., Gao, M., Meng, C., 2022. Generating gapless land surface temperature with a high spatio-temporal resolution by fusing multi-source satellite-observed and model-simulated data. Remote Sens. Environ. 278, 1–36.

Mo, Y., Xu, Y., Chen, H., Zhu, S., 2021. A Review of Reconstructing Remotely Sensed Land Surface Temperature under Cloudy Conditions. Remote Sens. 13, 2838.

Neteler, M., 2010. Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data. Remote Sens. 2, 333–351.

Pham, H.T., Kim, S., Marshall, L., Johnson, F., 2019. Using 3D robust smoothing to fill land surface temperature gaps at the continental scale. Int. J. Appl. Earth Obs. Geoinf. 82, 101879.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Springer Int. Publ.

Scharlemann, J.P.W., Benz, D., Hay, S.I., Purse, B.V., Tatem, A.J., Wint, G.R.W., Rogers, D.J., 2008a. Global Data for Ecology and Epidemiology: A Novel Algorithm for Temporal Fourier Processing MODIS Data. PLoS One 3, e1408.

Shi, C.X., Xie, Z.H., Qian, H., Liang, M.L., Yang, X.C., 2011. China land soil moisture EnKF data assimilation based on satellite remote sensing data. Sci. China Earth Sci. 54, 1430–1440.

Shuai, T., Zhang, X., Wang, S., Zhang, L., Shang, K., Chen, X., Wang, J., 2014. A spectral angle distance-weighting reconstruction method for filled pixels of the MODIS land surface temperature product. IEEE Geosci. Remote Sens. Lett. 11, 1514–1518.

Sun, D., Kafatos, M., 2007. Note on the NDVI-LST relationship and the use of temperature-related drought indices over North America. Geophys. Res. Lett. 34.

Tan, K.C., Lim, H.S., MatJafri, M.Z., Abdullah, K., 2012. A comparison of radiometric correction techniques in the evaluation of the relationship between LST and NDVI in Landsat imagery. Environ. Monit. Assess. 184, 3813–3829.

Tang, W., Xue, D., Long, Z., Zhang, X., Zhou, J., 2022. Near-real-time estimation of 1-km all-weather land surface temperature by integrating satellite passive microwave and thermal infrared observations. IEEE Geosci. Remote Sens. Lett. 19.

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process.

Wang, X., Girshick, R., Gupta, A., He, K., 2018. Non-local neural networks. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 7794–7803.

Wang, Z.H., Upreti, R., 2019. A scenario analysis of thermal environmental changes induced by urban growth in Colorado River Basin, USA. Landsc. Urban Plan. 181, 125–138.

Wang, A., Yang, Y., Pan, X., Zhang, Y., Hu, X., 2021. Research on land surface temperature downscaling method based on diurnal temperature cycle model deviation coefficient calculation. Natl. Remote Sens. Bull. 25, 1735–1748.

Weiss, D.J., Atkinson, P.M., Bhatt, S., Mappin, B., Hay, S.I., Gething, P.W., 2014. An effective approach for gap-filling continental scale remotely sensed time-series. ISPRS J. Photogramm. Remote Sens. 98, 106–118.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2016. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. AAAI Conf. Artif. Intell 31 (1).

Woo, S., Park, J., Lee, J.Y., Kweon, I.S., 2018. CBAM: Convolutional block attention module. Lect. Notes Comput. Sci 3–19.

Wu, P., Su, Y., Duan, S. bo, Li, X., Yang, H., Zeng, C., Ma, X., Wu, Y., Shen, H., 2022. A two-step deep learning framework for mapping gapless all-weather land surface temperature using thermal infrared and passive microwave data. Remote Sens. Environ. 277, 113070.

WMO, 2014. Guide to meteorological instruments and methods of observation. (CIMO guide), World Meteorological Organisation, Geneva, Switzerland.

Wu, P., Yin, Z., Yang, H., Wu, Y., Ma, X., 2019. Reconstructing geostationary satellite land surface temperature imagery based on a multiscale feature connected convolutional neural network. Remote Sens. 11.

Xu, Y., Shen, Y., 2013. Reconstruction of the land surface temperature time series using harmonic analysis. Comput. Geosci. 61, 126–132.

Yang, G., Sun, W., Shen, H., Meng, X., Li, J., 2019. An integrated method for reconstructing daily MODIS land surface temperature data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 1026–1040.

Yao, R., Wang, L., Huang, X., Sun, L., Chen, R., Wu, X., Zhang, W., Niu, Z., 2021. A robust method for filling the gaps in MODIS and VIIRS land surface temperature data. IEEE Trans. Geosci. Remote Sens. 59, 10738–10752.

Youssef, R., Aniss, M., Jamal, C., 2020. Machine learning and deep learning in remote sensing and urban application: a systematic review and meta-analysis. Proceedings of the 4th Edition of International Conference on Geo-IT and Water Resources 2020, Geo-IT and Water Resources 2020, 1–5.

Yu, W., Nan, Z., Wang, Z., Chen, H., Wu, T., Zhao, L., 2015. An effective interpolation method for MODIS land surface temperature on the Qinghai-Tibet Plateau. IEEE. J. Sel. Top. Appl. Earth Obs. Remote Sens. 8, 4539–4550.

Yue, W., Xu, J., Tan, W., Xu, L., 2007. The relationship between land surface temperature and NDVI with remote sensing: Application to Shanghai Landsat 7 ETM+ data. Int. J. Remote Sens. 28, 3205–3226.

Zeng, C., Long, D., Shen, H., Wu, P., Cui, Y., Hong, Y., 2018. A two-step framework for reconstructing remotely sensed land surface temperatures contaminated by cloud. ISPRS J. Photogramm. Remote Sens. 141, 30–45.

Zhan, W., Chen, Y., Wang, J., Zhou, J., Quan, J., Liu, W., Li, J., 2012. Downscaling land surface temperatures with multi-spectral and multi-resolution images. Int. J. Appl. Earth Obs. Geoinf. 18, 23–36.

Zhang, Q., Körnich, H., Holmgren, K., 2013. How well do reanalyses represent the southern African precipitation? Clim. Dyn. 40, 951–962.

Zhang, Q., Cheng, J., Wang, N., 2022. Fusion of all-weather land surface temperature from AMSR-E and MODIS data using random forest regression. IEEE Geosci. Remote Sens. Lett. 19.

Zhang, H., Liu, Y., Li, X., Feng, R., Gong, Y., Jiang, Y., Guan, X., Li, S., 2023. Combing remote sensing information entropy and machine learning for ecological environment assessment of Hefei-Nanjing-Hangzhou region, China. J. Environ. Manage. 325, 116533.

Zhang, D., Shao, J., Hu, G., Gao, L., 2017. Sharp and real image super-resolution using generative adversarial network. Lect. Notes Comput. Sci 217–226.

Zhang, Q., Wang, N., Cheng, J., Xu, S., 2020. A stepwise downscaling method for generating high-resolution land surface temperature from AMSR-E data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13, 5669–5681.

Zhang, X., Zhou, J., Liang, S., Wang, D., 2021. A practical reanalysis data and thermal infrared remote sensing data merging (RTM) method for reconstruction of a 1-km all-weather land surface temperature. Remote Sens. Environ. 260.

Zhao, W., Duan, S.B., 2020. Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated MODIS/Terra land products and MSG geostationary satellite data. Remote Sens. Environ. 247.