

Research Article

Missing Information Reconstruction of Land Surface Temperature Data Based on LPRN

Chen Xue ¹, Tao Wu ¹, Xiaomeng Huang,^{1,2} and Amir Homayoon Ashrafzadeh³

¹School of Computer Science, Chengdu University of Information Technology, Chengdu, China

²Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China

³RMIT University, Melbourne, Australia

Correspondence should be addressed to Tao Wu; wut@cuit.edu.cn

Received 19 August 2021; Accepted 6 September 2021; Published 21 September 2021

Academic Editor: Xianyong Li

Copyright © 2021 Chen Xue et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Temperature is the main driving force of most ecological processes on Earth, with temperature data often used as a key environmental indicator to guide various applications and research fields. However, collected temperature data are limited by the hardware conditions of the sensors and atmospheric conditions such as clouds, resulting in temperature data that are often incomplete. This affects the accuracy of results using the data. Machine learning methods have been applied to the task of completing missing data, with mixed results. We propose a new data reconstruction framework to improve this performance. Using the MODIS LST map over a span of 9 years (2000–2008), we reconstruct the land surface temperature (LST) data. The experimental results show that, compared with the traditional reconstruction method of LST data, the proportion of effective pixels of the LST data reconstructed by the new framework is increased by 3%–7%, and the optimization effect of the method is close to 20%. The experiment also discussed the influence of different altitudes on the data reconstruction effect and the influence of different loss functions on the experimental results.

1. Introduction

As a key element of biological survival, temperature is an important subject in the field of climate research. Researchers can determine the water-heat balance between the Earth's surface and the atmosphere based on surface temperature data. These data can also be used as a basis for understanding various terrestrial activities, determining fire and seismic zones, exploring geothermal resources, and studying urban heat island effects [1, 2]. Temperature also directly affects evapotranspiration and soil content and is key to evaluating terrestrial water volume, vegetation, and soil biochemical characteristics; atmospheric precipitation; and regional CO₂ content [3–9]. Temperature data are a common input variable into models of atmospheric, ecological, hydrological, and biogeochemical processes. The accuracy of the data directly affects the output accuracy and overall model accuracy [10, 11]. Therefore, it is important to obtain high-precision, continuous surface temperature data.

The collection of surface temperature data was initially carried out on mobile phones through discrete observation stations, obtaining high-precision, all-weather data through a large number of ground observation stations. However, the ground stations do not completely cover the Earth's surface, leading to a loss of data in many areas. Technological advances have led to Earth observation satellites as the mainstream data collection method, with remote sensing technology the only surface temperature observation method able to guarantee thorough coverage at all times in all areas on a global scale [12]. The collection of surface temperature using remote sensing technology is now an effective method for obtaining surface temperatures at a large scale. The theory and method of land surface temperature inversion from thermal infrared remote sensors are now relatively mature and produce high-precision clear-sky surface temperatures [13, 14]. However, thermal infrared remote sensing cannot penetrate clouds, leaving areas under cloud cover inaccessible and affecting data collection.

Solving the problem of missing temperature data due to cloud cover is currently a major topic in the field of climate research.

So far, researchers have applied many mature data reconstruction methods to solve the problem of temperature information loss. These methods fall into three broad categories [15]: space-based methods, spectrum-based methods, and time-based methods. Details of these methods are provided in the discussion in Section 2.2. In the research on the reconstruction of land surface temperature data, Shuo Xu et al. reconstructed the land surface temperature in the Tibetan Plateau (TP) and the Heihe River Basin (HRB) area based on the Bayesian maximum entropy (BME) method and verified the effectiveness of the method by using the soil temperature measured in the field [16]. Zhang et al. merged TIR and MW observations from a perspective of decomposition of LST in temporal dimension and overcame shortcomings of single-source remote sensing [17]. Martins et al. proposed an all-weather LST product based on visible light and infrared observations, which combines clear-sky LST retrieved from the Spinning Enhanced Visible and Infrared Imager on Meteosat Second Generation (MSG/SEVIRI) infrared (IR) measurements with LST estimated with a land surface energy balance (EB) model to fill gaps caused by clouds [18]. Although these different methods can acquire satisfactory recovery results, most of them are employed independently, and they can only be applied to a single specific reconstruction task in limited conditions [19]. On the one hand, most traditional data reconstruction methods such as Bayesian and KNN interpolation are based on linear regression, which requires a linear relationship between the data. However, the actual data missing scene is more complicated, and the data are often nonlinearly related. On the other hand, the ability of traditional data reconstruction methods to process data is limited. When the data reach a certain amount, the time to process the data will increase rapidly, and the data reconstruction experiment will be greatly affected. Therefore, it is necessary to find a better data reconstruction method that can integrate diversified information to overcome the limitations and deficiencies of traditional methods.

In recent years, benefiting from the improvement of computer computing power, LeCun et al. made a huge breakthrough in the field of image recognition using machine learning methods [20–23]. Considering that the satellite remote sensing field also needs to complete the recognition and classification of remote sensing satellite images and other similar tasks, related scholars have tried to introduce machine learning methods into the field of Earth sciences and have achieved certain results. Chen et al. used the Apriori method to find potential correlations in geological data [24]. Xie and Li used deep learning methods to denoise hyperspectral images [25]. Wei et al. improved the accuracy of sharpened images through the residual network [26]. Shah et al. proved the correlation between LST anomalies and earthquakes (EQs) in Pakistan with the ANN method based on MODIS LST data [27]. However, in terms of the reconstruction task of LST data based on machine learning methods, it is still difficult to find related research.

Therefore, we propose the LST palindrome reconstruction network (LPRN) method, which uses a deep convolutional neural network to reconstruct remote sensing images contaminated by dense clouds in MODIS LST data. At the same time, in order to provide high-quality training data and label data for the deep learning framework, the LPRN method also incorporates a variety of data pre-processing schemes, such as histogram equalization, inverse distance weighted interpolation, and so on. Compared with traditional data reconstruction methods, this method can effectively improve the number of “good quality” pixels (also represents effective pixels, explained in Section 2.1) in cloud-contaminated images while also greatly improving computing efficiency.

The rest of this article is organized as follows. In Section 2, we introduce the available datasets for reconstruction, current mainstream methods for reconstructing remote sensing data, and the network structure and specific details of our LPRN data reconstruction model. Section 3 presents our experimental results from using the model to reconstruct the surface temperature dataset. Finally, Section 4 presents our summary, concluding remarks, and directions for future research.

2. Materials and Methods

2.1. Datasets. The dataset used in the experiment came from the Terra-MODIS and the MODIS-Aqua datasets (<https://modis.gsfc.nasa.gov/data/dataproduct/mod11.php>), and the MODIS data product IDs used are MOD11A1, MYD11A1, MOD11A2, and MYD11A2. The Terra-MODIS dataset was from March 2000, and MODIS-Aqua from August 2002, both with daily resolution. The time difference between the Aqua satellite data and the Terra satellite data in the same area can be as short as a few hours, so there is great similarity between the two data. And Terra satellite data can be added to the dataset as supplementary information to the Aqua satellite data. We selected and processed nine years of MODIS Daily data and 8-Day data from 2000 to 2008 as experimental data. MODIS Land Surface Temperature and Emissivity products map land surface temperatures and emissivity values ideally under clear-sky conditions [28]. The underlying algorithms use other MODIS data and further auxiliary maps for input, including geolocation, radiance, cloud masking, atmospheric temperature, water vapor, snow, and land cover [29]. Temperatures are provided in Kelvin. The MODIS LST algorithm is aimed at reaching a better accuracy than 1 Kelvin (± 0.7 K stddev.) for areas with known emissivities in the range -10°C to 50°C [13, 29]. LST is observed by the two MODIS sensors four times per day (01 : 30, 10 : 30, 13 : 30, and 22 : 30, local solar time) originally at 1000 m pixel resolution. Clouds and other atmospheric disturbances, which may obscure parts of or even the entire satellite image, constitute a significant obstacle for continuous LST monitoring; the low-quality pixels of each LST map are marked in an accompanying quality assessment (QA) layer. For the LST maps, we used the open-source software GRASS to remap the MODIS LST data, filter out invalid pixels, and reject the pixels with the following labels

indicated in the QA bitmap: “other error,” “missing pixel,” “poor quality,” “average emissivity error >0.04,” “LST error 2 K–3 K,” and “LST error >3 K,” and the rest are “good quality” pixels. Based on the tags in the bitmap, it is possible to accurately filter the elements in the original LST image and improve the efficiency of LST data preprocessing. It should be noted that the data will only be stored in the database after being processed by the GRASS software. The dataset in the database is still in an unclassified state and there are abnormal values, and data cleaning and data classification need to be completed through the data preprocessing process.

2.2. Existing Remote Sensing Data Reconstruction Methods

Space-Based Methods. Space-based methods were originally used for image reconstruction in the field of computer vision. Among them, the most commonly used method is linear interpolation, which is weighted based on the surrounding information around missing data to estimate the missing information. This method often relies on the correlation and connection between missing information and its surrounding information. However, when the missing information changes greatly compared to the surrounding information, such as boundary values, it often results in inaccurate interpolation results. This method is reasonable, but it is limited by the amount of missing data and regional characteristics.

Spectrum-Based Methods. In the case of multispectral or hyperspectral images, specific spatial correlations between them have led some scholars to propose an innovative scheme for reconstructing missing information using these correlations. Rakwatin et al. used a polynomial linear fitting (LF) method to fit the missing data in the Aqua MODIS data [30]. Chen et al. used histogram matching and local least squares fitting methods to reconstruct Aqua MODIS data [31]. However, this method provides no help for satellite data with cloud cover.

Temporal Methods. Satellite remote sensing data are obtained through continuous satellite scanning of a defined area, with the collected data including the time of collected data. Some methods use these time elements to restore and reconstruct missing information. For example, Chen et al. proposed the use of a space-time weighted regression model (STWR) to obtain continuous cloud-free LST images [31]. Scaramuzza and Barsi proposed a local linear histogram matching (LLHM) method, but in most cases, it did not achieve good results [32]. However, it is undeniable that time differences are quite influential and useful when reconstructing missing information.

2.3. Reference Deep Learning Model

Convolutional Neural Network. CNN [23] is a multi-layer neural network originally used for image

recognition and classification. CNNs are generally composed of four parts: convolutional layer, nonlinear layer, pooling (also called downsampling) layer, and fully connected layer. The convolution operation primarily extracts features in the convolutional layer. The pooling layer reduces the number of calculations in the entire network and prevents overfitting. At the end of the network is the fully connected layer. The size of the convolution kernel is very important, as it determines the size of the neuron’s receptive field. If the kernel is too small, it is difficult to extract effective regional features. If the kernel is too large, the network complexity increases and may exceed the representation ability of the kernel. By convolving the input feature block x with the kernel and then passing the activation function $f(x)$ plus a bias term b , the output y is obtained. The entire convolution operation is given in equation (1). The function of the pooling layer is to reduce the amount of output data, obtaining lower dimensional features, reducing the number of calculations, and preventing feature overfitting. More specifically, neurons in the network are temporarily discarded according to a certain probability to prevent overfitting and improve the generalization ability of the model. Then, added regularization can make the parameters sparser, which makes one part of the optimized parameters become 0 and other part a nonzero real value. The nonzero real value plays a role in selecting important parameters or feature dimensions and at the same time plays a role in removing noise, which can further improve the performance of the model.

$$y = f(wx + b). \quad (1)$$

Fully Convolutional Network. FCNs [33] were first used for semantic segmentation, which differs from ordinary classification tasks that output a specific category only. FCNs require the output and the input to be the same. The structure of FCNs differs from CNNs as well. On the basis of CNNs, FCNs adjusted the last layer of the network and added a “deconvolution layer.” The fully connected layer of a CNN loses some of the position information in the input data. FCNs work to preserve this information by replacing the fully connected layer with a similar “deconvolutional layer.” After convolution and pooling, the feature matrix is restored to the original data dimensions after the “deconvolution layer” and upsampling method. Upsampling is the inverse process of the pooling layer introduced earlier which restores data features extracted through convolution and pooling to the appropriate dimensions to protect the original structure of the data at least in part. The entire network performs a total of five pooling operations, with feature fusion performed in the last step of the network. Feature fusion requires upsampling the feature map with the smallest receptive field obtained after convolution and pooling and then fusing the result with the feature map of the previous layer. The fused feature map is again upsampled with the

feature map of the previous layer, and feature fusion is performed until the feature map is restored to the original number of data dimensions.

U-Net. U-Net [34] is a variant of the FCN network, with a changed encoder-decoder structure. There is no fully connected layer in the network, and the structure is divided into two parts: the encoder path and the decoder path. The structure of the encoder path is similar to the traditional CNN structure, and the decoder path is the encoder path's inverse process. The whole structure resembles the letter U, incorporated into the name. The encoder path mainly performs feature extraction and dimensionality reduction. The decoder path is composed of a convolutional layer, an activation layer, and an upsampling layer, which fuse features and restore the data dimensionality. The biggest difference between U-Net and FCN is in the deep and shallow feature information fusion of the data. FCN fuses feature information by adding the values at the corresponding positions, while U-Net fuses using same dimensional splicing. The latter works by placing feature information at different levels in different channels, splicing it together, and then convolving to fuse features. The FCN network fuses data features through addition, which will obscure the details of the data features, while U-Net's splicing structure can retain more position information and reduce the loss of data details.

2.4. Proposed LST Palindrome Reconstruction Network (LPRN). Inspired by the encoder-decoder framework in U-Net, we try to improve the original network structure of U-Net and propose the LPRN framework. This framework applies deep convolutional neural networks to LST data reconstruction research, and the overall architecture is shown in Figure 1. The whole framework includes three processes, namely, the reprojection of MODIS LST (explained in section *Datasets*), the preprocessing of the dataset used by the model, and the structure design of the LPRN model.

2.4.1. Data Preprocessing. The LST data after reprojection can be divided into two categories according to whether they are occluded by clouds: cloud-containing images and cloud-free images. After the cloud-containing image is filtered, the "good quality" pixels will be greatly reduced and cannot be directly used as the training dataset of the deep learning network (when there are too many missing values in the training dataset, the deep learning method cannot capture the potential correlation between the data, nor can it reconstruct the missing real data). Therefore, it is necessary to use data processing methods to interpolate the missing data to a certain extent. In the process of data preprocessing, we first use the histogram equalization (HE) method to determine whether the image is occluded by clouds, then divide the dataset, and finally use the inverse distance weighting (IDW) method (see Figure 2) to interpolate part of the missing data.

The data preprocessing process includes the following steps:

Step 1: according to the experimental needs, download the LST dataset within a specified time scale in batches, classify and store the data according to the satellite, date, and other factors, and establish the corresponding association between the data.

Step 2: divide the cloud-containing image according to the spectral characteristics—take out the image data from the database and convert it into the form of gray frequency histogram; analyze the gray frequency histogram; if two peak frequencies appear and increase with the gray value and the gray value between the peaks has a buffer zone, the gray value of the area corresponding to the second peak is set as an abnormal value and the image is divided into cloud-containing image; if the above requirements are not met, the image is classified as pending image data; the above process is repeated until all the images in the dataset are completely divided.

Step 3: analyze the spectral characteristics of the cloud-containing image through the gray frequency histogram. According to the analysis of the existing data, it is found that the frequency histogram of the nonthin cloud-containing image will show two obvious gray scale peaks. The peak frequency with the higher gray scale will be significantly higher than the one with the lower gray scale and there will be a certain buffer in the frequency between two gray values. This feature is an important basis for determining whether an image is a cloud-containing image. If the image is a cloud-free image, the gray frequency histogram of the image will show an approximate standard normal distribution. According to the above feature, the remote sensing data of LST are preliminarily classified.

Step 4: perform histogram equalization processing on the filtered pending image data and filter the cloud-containing image for the second time according to the processing result—take out the image data in sequence from the pending image dataset, perform the histogram equalization processing on the image data, and calculate the gray average value and second-order moment of the image before and after processing; classify the image data according to equations (2) and (3); if two inequalities are satisfied at the same time, the image will be classified as a cloud-containing image from the pending image dataset; otherwise, the image will be classified as a cloud-free image (where $Mean_b$ represents the average value of the candidate image after equalization processing, $Mean_f$ represents the average value of the candidate image before equalization processing, ASM_b represents the second moment of the candidate image after equalization processing, and ASM_f indicates that the candidate image is equalizing second moment before transformation).

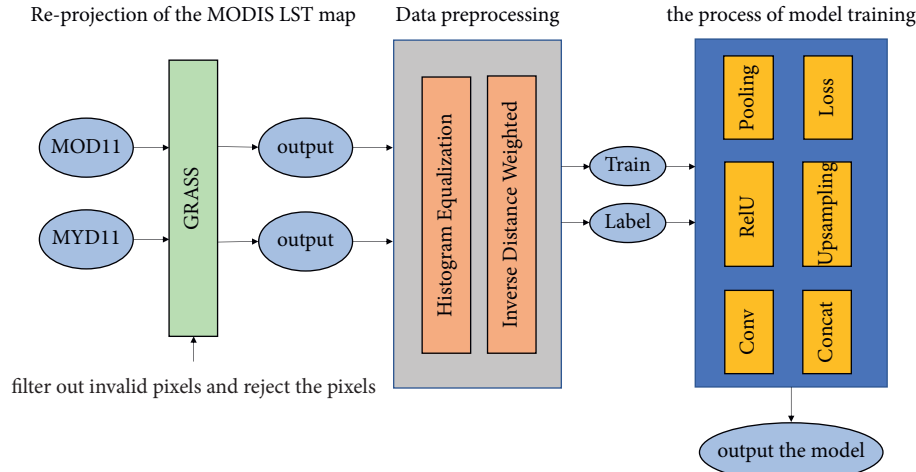


FIGURE 1: Flowchart of the LPRN framework for the missing information reconstruction of LST.

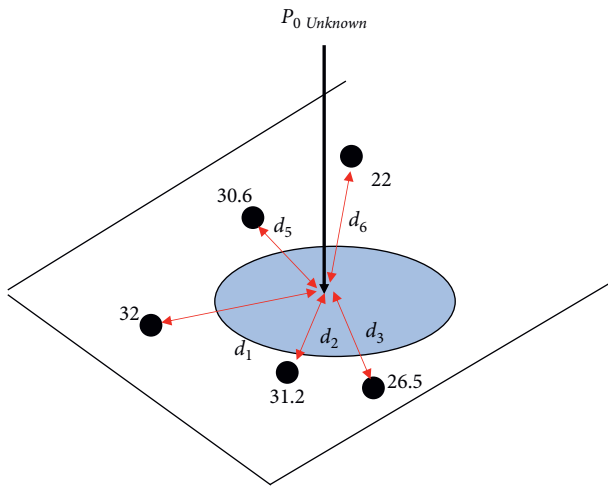


FIGURE 2: The inverse distance weighting.

$$2.3 < \frac{\text{Mean}_b}{\text{Mean}_f} < 3 \quad (2)$$

$$-0.03 \leq \text{ASM}_f - \text{ASM}_b \leq 0.03.$$

Step 5: record the cloud occlusion range of the cloud-containing image, divide the cloud occlusion area, and count the cloud occlusion ratio—for the data classified as cloud-containing images in step 2, the cloud occlusion range is divided according to the highest gray value peak of the cloud-containing image data in the frequency histogram, and the pixels in the abnormal gray scale peak interval are all cloud occlusion areas, and count the number of cloud pixels; for the data classified as cloud-containing images in step 4, the cloud occlusion range is divided according to the average gray value of the buffer area in the distribution of the frequency histogram. If there is a gray peak value that is much larger than the average value, the area is divided into a cloud occlusion area, and count the amount of cloud pixels.

Step 6: filter data according to the cloud occlusion ratio of each sample in the cloud-containing images—set the threshold according to the proportion of cloud-occluded pixels calculated in step 5 and determine whether to accept the land surface temperature data according to the threshold (the threshold needs to be flexibly determined according to the actual amount of data, and if the threshold is set too high, too much data will be deleted).

Step 7: take out the cloud-containing image in turn and use the inverse distance weighting method to reconstruct the pixels in the cloud occlusion area according to equations (4), (5), and (3); if the range of the missing pixels is large, the pixel value of the missing area of the picture is reconstructed according to the ratio of the gray values between the pixels in the image at different moments in the same area (where d_i represents the distance from the target pixel to the surrounding pixels, λ_i represents the calculated weight based on the reciprocal of the distance, and $\hat{P}(x, y)$ represents the sum of the products of the surrounding pixels and the weight).

$$d_i = \sqrt{(x - x_i)^2 + (y - y_i)^2},$$

$$\lambda_i = \frac{1/d_i}{\left(\sum_{i=1}^n 1/d_i\right)}, \quad (3)$$

$$\hat{P}(x, y) = \sum_{i=1}^n \lambda_i P(x_i, y_i).$$

The main purpose of the above process is to eliminate outliers and obtain the standard dataset for the LPRN model, and the brief process is shown in Figure 3.

The above method can remove the data missing problem caused by cloud occlusion to a certain extent and reduce the numerical error caused by cloud occlusion. However, in most cases, because the range of missing values is relatively large, the effect of interpolation methods based on space and time is limited.

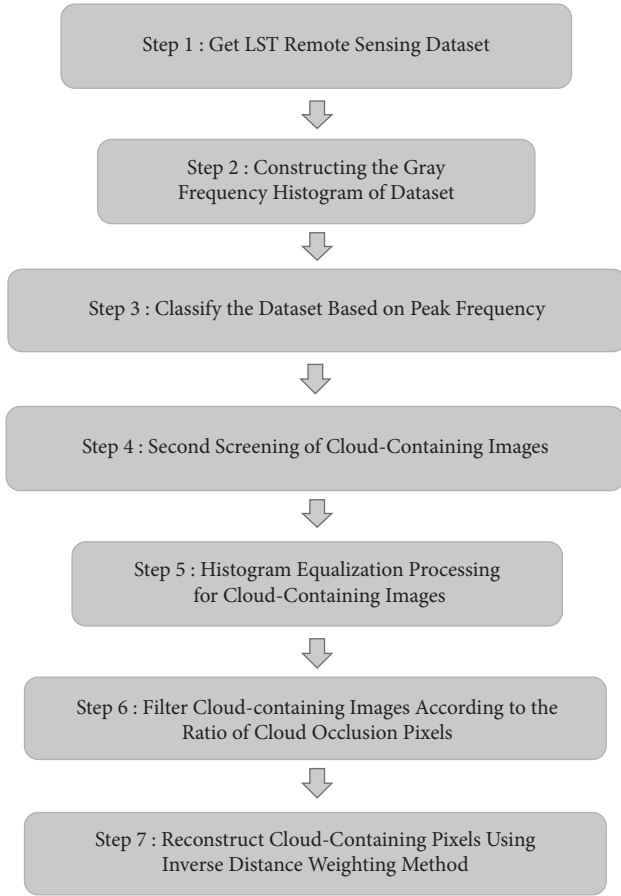


FIGURE 3: The flowchart of data preprocessing.

2.4.2. LPRN Framework. With reference to the structure of U-Net, the LPRN model proposed in the paper is based on the encoder-decoder architecture to reconstruct the missing data of LST. The LPRN model consists of 14 convolutional layers, 3 pooling layers, and 3 upsampling layers, and since the data characteristic specifications obtained by each layer are different, each layer has constraints on the specifications of the input data and the output data. Considering the overall calculation amount of the model, the characteristics that the model needs to obtain, and the LST data at different times and dates, the model uses convolution kernels of different sizes to obtain the characteristics of the LST data as comprehensively as possible. The overall architecture of the LPRN framework is shown in Figure 4. The label data in the proposed model are the original 8-Day LST image, and the training data in the proposed model are the preprocessed Daily LST image.

As shown in the framework, the slice size of the input Daily LST data sample is $200 * 248$. In the framework, five channels of different sizes are set, which are 1, 8, 16, 32, and 64. Three convolution kernels of different sizes are set to realize feature extraction of different scales, which are $[3, 3]$, $[2, 2]$, and $[1, 1]$. In order to avoid the loss of too much detailed information in the data, the downsampling layer only uses a convolution kernel with a size of $[2, 2]$ to compress the data volume features. The framework adopts a

decoder-encoder structure to effectively reduce the loss of original data information due to pooling and convolution. This structure retains the intermediate information of the data during the decoding process and then uses convolution and data splicing methods to fuse the original information and the extracted features during the encoding process, which can reconstruct the missing data according to the characteristic correlation information in the original data.

In the deep learning network, the loss function is used to estimate the degree of inconsistency between the predicted value $f(x)$ of the model and the true value Y . The smaller the loss function, the better the stability of the model. Relying on the guidance of the loss function, the deep learning model can realize the learning process.

In the experiment, we tried a variety of different loss functions. Among them, the original loss function of the U-Net network is the cross entropy function, and the commonly used loss functions of deep learning networks include mean square error (MSE) and root mean square error (RMSE). However, the experimental results show that aforementioned loss function has very limited guiding effect on LPRN, so we tried the relatively unpopular normalized cross correlation (NCC) loss function and achieved relatively good results (the experimental results of different loss functions are provided in the discussion in Section 3).

RMSE: RMSE is often used as a standard for measuring the prediction results of machine learning models, which is expressed as follows:

$$\text{RMSE}(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x_i) - y_i)^2}, \quad (4)$$

where X represents the training sample, m represents the number of training samples, $h(x)$ represents the predicted value of the training sample, and y represents the true value of the training sample, which is also called the label value.

NCC: NCC is primarily used with image matching, and it is the process of finding the subimage having greatest similarity to a real time image for the purpose of identifying a target image, which is expressed as follows:

$$\rho(x, y) = \frac{\sigma(S_{x,y}, g)}{\sqrt{D_{x,y}D}}, \quad (5)$$

where $S_{x,y}$ and g , respectively, represent the corresponding subblocks of the two compared images, $\rho(x, y)$ represents the correlation coefficient, which is used to determine whether the two subblocks are related, $\sigma(S_{x,y}, g)$ represents the covariance of $S_{x,y}$ and g , $D_{x,y}$ represents the variance of datum $S_{x,y}$, and D represents the variance of datum g .

The inequality $|\rho(x, y)| \leq 1$ is used to measure the correlation between two subblocks, and the value of similarity is within $[-1, 1]$. The correlation coefficient characterizes the description of the degree of approximation between the two

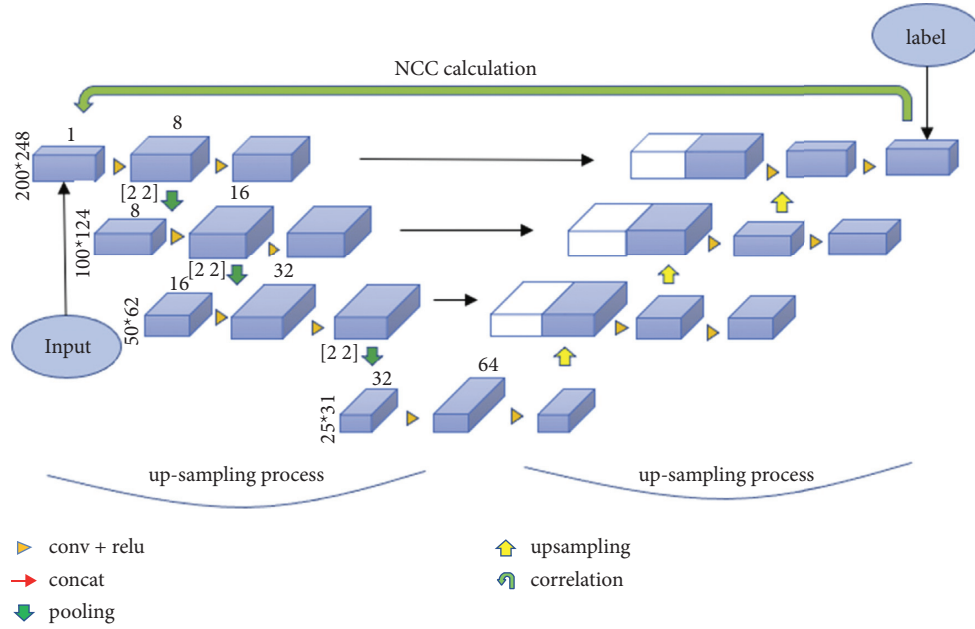


FIGURE 4: Architecture of the proposed LPRN framework.

data. Generally speaking, the closer it is to 1, the closer the linear relationship between the two data is.

3. Results and Discussion

Uncertainty in the processed MODIS LST data as proposed in this paper is of two types: (1) intrinsic uncertainty in the downloaded MODIS data and (2) error introduced by processing. The former includes error in the radiometric and geometric precision of the MODIS instrument and uncertainties in the known emissivity values of land surfaces, cirrus, or other atmospheric phenomena. In the experiment, the reprojection of MODIS LST and the preprocessing of the dataset are used to minimize the impact of related errors. The latter is mainly the potential error caused by the calculation of the average value in the preprocessing process and the data reconstruction process. This error may be reduced in the future by using additional data sources.

In order to comprehensively evaluate the impact of uncertainty on the accuracy of LPRN reconstruction data, this study carried out two different LST reconstruction data accuracy evaluation experiments.

In the first experiment, we mainly examine the change of a single LST value before and after reconstruction. The experiment randomly selected 2000 LST original samples from the cloud-free image, so as to avoid selecting too many similar regions. Then, we randomly select 5% of the “good quality” pixels from each original sample and modify them as missing data and record the location of the modified pixels. Finally, we put the “modified” LST original sample into the LPRN model for reconstruction and find the reconstructed LST data according to the position of the “modified” pixel and compare it with the actual value of the LST in the original sample (see Table 1)

(because the images are randomly selected, the distribution of the number of images in each year is not the same).

Analyzing the results of the first experiment and comparing the reconstructed data from 2000 to 2002 with the original image, the average temperature error between the corresponding pixels is larger, and the average error is above 1 K. Comparing the reconstructed data from 2003 to 2008 with the original image, the average temperature error between the corresponding pixels is small, the average error is below 0.4 K, and the smallest error is close to 0.1 K. Analysis of the reasons for this phenomenon is mainly due to the different reconstruction effects caused by the amount of training data. The data from 2003 to 2008 are compared with the data from 2000 to 2002. The latter lacks MODIS-Aqua data, which leads to the reduction of data associations that can be referred to by the model, so the reconstruction effect of the model also decreases accordingly.

The first experiment mainly evaluates the influence of the LPRN model on the reconstruction of the original data, and the result proves that the model will not cause a large change in the normal data. However, for the newly created information in the missing area, since they are data that do not exist in the original dataset, their accuracy cannot be verified, so we introduced a data correlation coefficient $\rho(T_{re}, T_{tr})$ in the second experiment, as shown in equation (6). In the second experiment, based on the correlation between the reconstructed data and the label data, we further verify the accuracy of the reconstructed data of the LPRN model (experiment 2 used all the extracted cloud-free data into the LPRN model for experimentation, and because there are few cloud-free data available from 2000 to 2002, they were not used in the experiment).

TABLE 1: Average error statistics of the single pixel.

Year	Error (K)
2000	1.3
2001	1.4
2002	1.0
2003	0.4
2004	0.1
2005	0.3
2006	0.4
2007	0.2
2008	0.3
Mean	0.6

$$\rho(T_{re}, T_{tr}) = \frac{Cov(T_{re}, T_{tr})}{\sqrt{D(T_{re})} \sqrt{D(T_{tr})}} \quad (6)$$

where T_{re} and T_{tr} , respectively, indicate reconstructed data and label data and $Cov(\cdot)$ and $D(\cdot)$ indicate covariance and variance, respectively.

The second experiment first finds out the corresponding data in the label dataset according to the date of the cloud-free data and then puts all the cloud-free data into the trained LPRN framework for reconstruction and obtains the reconstructed cloud-free dataset. Finally, the correlation and error between the original cloud-free dataset, the reconstructed cloud-free dataset, and the label dataset are calculated to verify the accuracy of the reconstructed data (see Table 2).

The focus of the introduction of the correlation coefficient is to test the accuracy of the reconstruction results of the LPRN model. Analyzing the data results in Table 2, we first analyze the correlation between the original data and the label data. The maximum correlation coefficient appeared in 2004 and was 0.90, and the error in that year was also the smallest, with an average error of 2.92 K. The correlation coefficient in 2006 was the smallest, 0.87, and the error in that year was also the largest (5.07 K). The average value of the correlation coefficient between the original data and the label data is 0.883, and the average value of the error is 4.265. Then, we analyze the correlation coefficient and error between the LST data reconstructed by the LPRN method and the label data. The maximum value of the correlation coefficient is 0.93, the minimum value is 0.90, the minimum error is 1.21 K, and the maximum value is 2.62 K. After the reconstruction of the LST data, the average value of the correlation coefficient is improved by nearly 0.03, and the error is reduced by 2.4 K on average.

The results of experiment two show that the correlation coefficient and the error are linearly related, and the smaller the error, the greater the correlation. Analyzing the error between the original data before and after reconstruction and the label data, it can be determined that the reconstruction effect of the LPRN model is obvious. After comprehensive analysis of the results of the first experiment and the second experiment, the LPRN model can significantly reduce the error between the original data and the label data. But when the correlation coefficient of the two data reaches a

TABLE 2: Evaluations of the reconstructed data.

Year	$\rho(T_{ori}, T_{tr})$	Error(T_{ori}, T_{tr})	$\rho(T_{re}, T_{tr})$	Error(T_{re}, T_{tr})
2003	0.88	4.90 K	0.92	1.68 K
2004	0.90	2.92 K	0.93	1.21 K
2005	0.88	4.31 K	0.91	2.15 K
2006	0.87	5.07 K	0.92	1.74 K
2007	0.89	3.83 K	0.90	2.62 K
2008	0.88	4.56 K	0.93	1.43 K
Mean	0.883	4.265 K	0.918	1.805 K

certain value, the model reconstruction effect will be limited. Therefore, more diversified label data can be considered in future research, and the reconstruction effect of the LPRN model can be improved by reducing the correlation between the data. At the same time, it should be noted that, compared with the first experiment, as the number of verification data increases, the error will also increase to a certain extent. This may be due to the small amount of data used in the first experiment, resulting in greater randomness in the results.

In addition, in order to test the effect of the LPRN model on the reconstruction of missing data, the experiment selects the method proposed in [35] for comparison. Reference [35] mainly uses traditional data reconstruction methods such as nearest neighbor algorithm and histogram equalization. Considering that the experimental results are all reconstructed data and lack the reference of real data, the experiment mainly compares the number of “good quality” pixels in the image reconstructed by the two methods. The experiment selects the same area as in research [35] (46:45:00N-45:36:50N; 10:06:33E-12:46:30E) and data size (original size after data download). Table 3 shows the statistical analysis results for the spatial distribution of “good quality” pixels. In the comparative experiment, in addition to analyzing the “good quality” pixels in the observation data, the correlation between the altitude factor and the number of “good quality” pixels in the data was also analyzed and studied.

Analyzing the experimental results in Table 2, the LST data reconstructed by using the LPRN method show similar characteristics to the data reconstructed by the traditional method. Compared with high-altitude areas (>1500 m), low-altitude areas (such as flat bottoms and valleys) are more likely to be obscured by coverings, resulting in fewer “good quality” pixels that can be collected. It is reflected in Table 1 that the proportion of “good quality” pixels has dropped significantly. At the same time, compared with traditional reconstruction methods, in the LST data reconstructed by LPRN (loss function using NCC), the proportion of “good quality” pixels will be further increased, and the maximum increase can reach 7%.

Based on the overall analysis of the experimental results, after removing the interference of altitude factors, the percentage of “good quality” pixels that can be collected throughout the year is between 37% and 48%. Compared with the data reconstructed by the traditional method, the data reconstructed by the LPRN method can increase the proportion of “good quality” pixels by 3%–7%.

TABLE 3: The spatial distribution of “good quality” pixels in the daily MODIS LST obtained after dividing by altitude (each value represents the percentage of “good quality” pixels; the data reference comes from [35]).

Year	0–499 m	LPRN	500–1499 m	LPRN	>1500 m	LPRN
2000	30.0	35.3	36.0	40.2	36.1	43.2
2001	32.4	36.5	40.3	46.5	40.5	46.1
2002	27.0	30.9	33.6	40.3	35.9	39.3
2003	35.5	38.2	46.1	49.9	47.6	53.8
2004	31.5	36.7	39.2	46.2	41.6	46.9
2005	35.1	40.0	44.7	50.1	45.1	50.3
2006	34.7	39.2	43.8	44.5	46.1	51.0
2007	38.2	45.1	46.6	53.2	46.5	52.6
2008	28.3	33.1	33.4	40.1	34.0	46.2
Mean	32.5	37.2	40.4	45.6	41.5	47.7

In order to prove that the use of different loss functions affects the effect of data reconstruction, the experiment designed a comparative experiment with the loss function as a variable (only RMSE and NCC are listed here; MAE and MSE have poor results, so they are not listed here). Figure 5 shows the results of the proportion of “good quality” pixels in the LST data reconstructed by the model trained using RMSE and NCC as the loss function in an area with an altitude of less than 500 m.

In Figure 5, “ori” represents the proportion of “good quality” pixels in the output result using traditional reconstruction methods, “RMSE” represents the proportion of “good quality” pixels in the output result of using RMSE as the loss function of the LPRN model, and “NCC” represents the proportion of “good quality” pixels in the output result of using NCC as the loss function of the LPRN model. In the area below the altitude of 500 m, the proportion of “good quality” pixels in the output result of the model using the loss function NCC can be increased by about 4.7% on average. The reconstruction effect of the model based on the loss function RMSE is improved to a certain extent compared with the traditional method, but it is lower than the LST data reconstruction effect of the model based on the loss function NCC.

Figure 6 shows the results of the proportion of “good quality” pixels in the LST data reconstructed by the model trained using RMSE and NCC as the loss function in an area with an altitude ranging from 500 m to 1500 m. The proportion of “good quality” pixels in the output result of the model using the loss function NCC can be increased by about 5.2%.

Figure 7 shows the results of the proportion of “good quality” pixels in the LST data reconstructed by the model trained using RMSE and NCC as the loss function in an area with an altitude above 500 m to 1500 m. The proportion of “good quality” pixels in the output result of the model using the loss function NCC can be increased by about 6.2%.

The experimental results proved that the model effect trained by the deep learning network has great correlation with the choice of loss function. In different application scenarios, the use of different loss functions will play different roles. In the LST data reconstruction experiment, we compared the use of RMSE as the loss function with the use of NCC as the loss

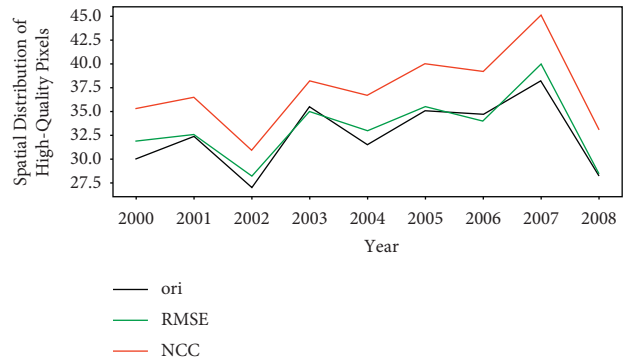


FIGURE 5: Proportion of “good quality” pixels at altitudes <500 m (RMSE and NCC).



FIGURE 6: Proportion of “good quality” pixels at altitudes 500 m–1500 m (RMSE and NCC).

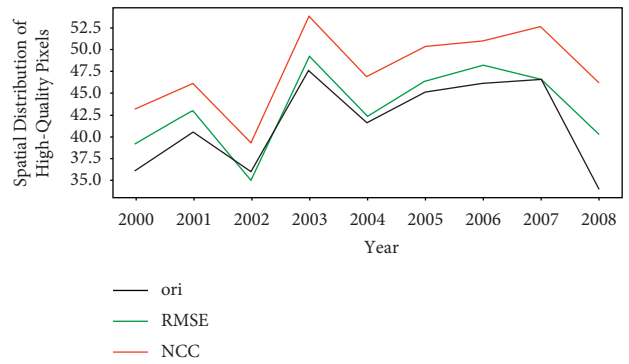


FIGURE 7: Proportion of “good quality” pixels at altitudes >1500 m (RMSE and NCC).

function, indicating that in the application scenario of LST data reconstruction, using the loss function NCC will significantly increase the proportion of “good quality” pixels, which proves that NCC is more suitable for this field.

4. Conclusions

High-resolution, high-precision satellite remote sensing data have tremendous value for many fields including the Earth sciences. However, the satellite remote sensing data collected

are often contaminated, particularly by cloud cover. LST data processed using existing methods lose significant information due to occlusions, necessitating one or more passes through reconstruction methods to restore the missing information.

In this article, we use the histogram equalization and inverse distance weighting method to preprocess the LST raw data, remove the outliers in the data and construct the standard dataset usable by the deep learning network, and finally use the LPRN framework to reconstruct LST data. In the experiment to verify the accuracy of the reconstructed data, the experiment proved that the LPRN method can reduce the error between the training data and the label data. Therefore, in future extended experiments, the reconstruction effect of the LPRN method can be further enhanced by using higher precision or higher precision label data. Compared with traditional reconstruction methods, in the LST data reconstructed by LPRN, the proportion of “good quality” pixels will be further increased, and the maximum increase can reach 7%. Compared with traditional reconstruction methods, in the LST data reconstructed by LPRN, the proportion of “good quality” pixels will be further increased, and the maximum increase can reach 7%. In addition, the experimental results show that compared with high-altitude areas (>1500 m), low-altitude areas (such as flat bottoms and valleys) are more likely to be obscured by coverings, resulting in fewer “good quality” pixels that can be collected, and the proportion of “good quality” pixels will drop significantly. Finally, the experiment also found that choosing a suitable loss function has a significant effect on the experimental results. The experimental results potentially showed that increasing amounts of training data will change the experimental results. This is related to the inherent nature of the deep learning method. Therefore, it can be expected that when the amount of data reaches a certain level in the future, deep learning methods can explore more hidden information in the data.

Existing deep learning framework research has shown that training data with time continuity contain potential time and space attributes, with the ability to predict the future to some degree based on them. In the field of Earth sciences, explorations of these phenomena are rare. In the future, we plan to incorporate more diversified relevant data into the model, such as wind speed, sunshine duration, and light intensity, by integrating multiple data fusion methods into the data reconstruction methods. In this way, we hope to reduce the effects of data information loss caused by external interference and to maximize the value of data.

Data Availability

Publicly available datasets were analyzed in this study. These data can be found in the following website: <https://modis.gsfc.nasa.gov/data/dataproduct/mod11.php>.

Disclosure

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the National Key Research and Development Program of China (grant nos. 2020YFA0607900 and 2018YFC1507005), the National Natural Science Foundation of China (grant no. 42075137), and the Sichuan Science and Technology Program (grant no. 2020YFG0189).

References

- [1] L. Chen, W. T. Huang, J. Wu et al., “Ore-forming temperature variation and prospecting target of the Wuxu ore field in Guangxi,” *Geochimica*, vol. 44, no. 6, pp. 546–555, 2015.
- [2] J. A. Voogt and T. R. Oke, “Thermal remote sensing of urban climates,” *Remote Sensing of Environment*, vol. 86, no. 3, pp. 370–384, 2003.
- [3] Z. L. Li, S. B. Duan, B. H. Tang et al., “Review of methods for land surface temperature derived from thermal infrared remotely sensed data,” *Journal of remote sensing*, vol. 20, no. 5, pp. 899–920, 2016.
- [4] X. Guo and G. Cheng, “Advances IN the application OF remote sensing to evapotranspiration research,” *Advances in Earth Science*, vol. 19, no. 1, pp. 107–114, 2004.
- [5] L. H. Zhang, Y. N. Chen, R. F. Zhao, and W. H. Li, “Impact OF temperature and soil water content ON soil res-piration IN temperate deserts, China,” *Chinese Journal of Plant Ecology*, vol. 33, no. 5, p. 936, 2009.
- [6] A. J. Arnfield, “Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island,” *International Journal of Climatology*, vol. 23, no. 1, pp. 1–26, 2003.
- [7] J. D. Kalma, T. R. McVicar, and M. F. McCabe, “Estimating land surface evaporation: a review of methods using remotely sensed surface temperature data,” *Surveys in Geophysics*, vol. 29, no. 4, pp. 421–469, 2008.
- [8] F. N. Kogan, “Operational space technology for global vegetation assessment,” *Bulletin of the American Meteorological Society*, vol. 82, no. 9, pp. 1949–1964, 2001.
- [9] Y. Jiang, D. Yu, M. Zhao, H. Bai, C. Wang, and L. He, “Analysis of semi-supervised text clustering algorithm on marine data,” *Computers, Materials & Continua*, vol. 64, no. 1, pp. 207–216, 2020.
- [10] E. Valor and V. Caselles, “Mapping land surface emissivity from NDVI: application to European, African, and South American areas,” *Remote Sensing of Environment*, vol. 57, no. 3, pp. 167–184, 1996.
- [11] L. Zhang, L. Bai, X. Zhang, Y. Zhang, F. Sun, and C. Chen, “Comparative variance and multiple imputation used for missing values in land price DataSet,” *Computers, Materials & Continua*, vol. 61, no. 3, pp. 1175–1187, 2019.
- [12] Z.-L. Li, B.-H. Tang, H. Wu et al., “Satellite-derived land surface temperature: current status and perspectives,” *Remote Sensing of Environment*, vol. 131, pp. 14–37, 2013.
- [13] Z. Wan, Y. Zhang, Q. Zhang, and Z.-L. Li, “Quality assessment and validation of the MODIS global land surface temperature,” *International Journal of Remote Sensing*, vol. 25, no. 1, pp. 261–274, 2004.

- [14] S.-B. Duan and Z.-L. Li, "Spatial downscaling of MODIS land surface temperatures using geographically weighted regression: case study in northern China," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6458–6469, 2016.
- [15] H. Shen, X. Li, Q. Cheng et al., "Missing information reconstruction of remote sensing data: a technical review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 3, no. 3, pp. 61–85, 2015.
- [16] S. Xu, J. Cheng, and Q. Zhang, "Reconstructing all-weather land surface temperature using the bayesian maximum entropy method over the Tibetan plateau and Heihe river basin," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 9, pp. 3307–3316, 2019.
- [17] X. Zhang, J. Zhou, F.-M. Göttsche, W. Zhan, S. Liu, and R. Cao, "A method based on temporal component decomposition for estimating 1-km all-weather land surface temperature by merging satellite thermal infrared and passive microwave observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 4670–4691, 2019.
- [18] J. P. A. Martins, I. F. Trigo, N. Ghilain et al., "An all-weather land surface temperature product based on MSG/SEVIRI observations," *Remote Sensing*, vol. 11, no. 24, p. 3044, 2019.
- [19] M. K.-P. Ng, Q. Yuan, L. Yan, and J. Sun, "An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3367–3381, 2017.
- [20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [21] B. Zhan, D. Li, X. Wu, J. Zhou, and Y. Wang, "Multi-modal mri image synthesis via gan with multi-scale gate mergece," *IEEE Journal of Biomedical and Health Informatics*, 2021.
- [22] Y. Wang, L. Zhou, B. Yu et al., "3D auto-context-based locality adaptive multi-modality GANs for PET synthesis," *IEEE Transactions on Medical Imaging*, vol. 38, no. 6, pp. 1328–1339, 2018.
- [23] Y. LeCun, B. Boser, J. Denker et al., "Handwritten digit recognition with a back-propagation network," *Advances in Neural Information Processing Systems*, vol. 2, 1989.
- [24] M. Chen, X. Luo, Y. Zhu, Y. Li, W. Zhao, and J. Wu, "An apriori-based learning scheme towards intelligent mining of association rules for geological big data," *Intelligent Automation and Soft Computing*, vol. 26, no. 5, pp. 973–987, 2020.
- [25] W. Xie and Y. Li, "Hyperspectral imagery denoising by deep learning with trainable nonlinearity function," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 11, pp. 1963–1967, 2017.
- [26] Y. Wei, Q. Yuan, H. Shen, and L. Zhang, "Boosting the accuracy of multispectral image pansharpening by learning a deep residual network," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 10, pp. 1795–1799, 2017.
- [27] M. Shah, R. U. Qureshi, N. G. Khan, M. Ehsan, and J. Yan, "Artificial Neural Network based thermal anomalies associated with earthquakes in Pakistan from MODIS LST," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 215, Article ID 105568, 2021.
- [28] Z. Wan, *MODIS Land Surface Temperature Products Users' Guide*, Institute for Computational Earth System Science, University of California, Santa Barbara, CA, USA, 2006.
- [29] Z. Wan, *Collection-5 MODIS Land Surface Temperature Products Users' Guide*, ICESSE, University of California, Santa Barbara, CA, USA, 2007.
- [30] P. Rakwatin, W. Takeuchi, and Y. Yasuoka, "Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 2, pp. 613–627, 2008.
- [31] B. Chen, B. Huang, L. Chen, and B. Xu, "Spatially and temporally weighted regression: a novel method to produce continuous cloud-free Landsat imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 27–37, 2016.
- [32] P. Scaramuzza and J. Barsi, "Landsat 7 scan line corrector-off gap-filled product development," *Proceeding of Pecora*, vol. 16, pp. 23–27, 2005.
- [33] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, 2015.
- [34] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science*, pp. 234–241, Springer, Berlin, Germany, 2015.
- [35] M. Neteler, "Estimating daily land surface temperatures in mountainous environments by reconstructed MODIS LST data," *Remote Sensing*, vol. 2, no. 1, pp. 333–351, 2017.