Long Time-Series Glacier Outlines in the Three-Rivers Headwater Region From 1986 to 2021 Based on Deep Learning

Longfei Chen , Wanchang Zhang , *Member, IEEE*, Yaning Yi, Zhijie Zhang , and Shijun Chao

Abstract—The deep-learning-based approach has drawn significant attention in glacier extraction due to its advantages over traditional techniques. In this study, to verify the feasibility and effectiveness of LandsNet architecture for glacier extraction, we applied a modified LandsNet (M-LandsNet) to extract the glacier outlines in the Three-Rivers Headwater Region. The band ratio method, U-Net, U-Net++, GlacierNet, SaU-Net, U-Net+cSE, and LandsNet, and two scenes were used for comparison. Analysis of the two scenes indicated that the M-LandsNet had the best performance and generalization ability among the eight methods. Weather conditions had the greatest negative impact on the eight methods, followed by geographic environment and geographic location. We further extracted the glacier outlines in the Three-Rivers Headwater Region in 1986−2021 in a total of 12 periods using the M-LandsNet and through manual adjustments. The glacier area in the Three-Rivers Headwater Region has decreased by 416.40 ± 102.71 km^2 $(16.53 \pm 4.08\%)$ in 1986−2021. The reduction rate $(16.13 \pm 5.63$ km^2 a$^{-1})$ in 2003−2021 was almost twice that $(7.42 \pm 5.97$ km^2 a$^{-1})$ in 1986−2003. The reduction rate of the glacier area varied among different periods and areas. Comparison with previous results indicated that the obtained glacier outline dataset in this study is reliable, and can effectively reflect the glacier area and spatio-temporal glacier changes in the Three-Rivers Headwater Region. A long time-series dataset of glacier outlines in the Three-Rivers Headwater Region in 1986−2021 is available at https://doi.org/10.5281/zenodo.5512064. This study can provide data support for the estimation of regional water resources storage.

Index Terms—Deep learning, long time-series glacier outlines, M-LandsNet, three-river headwater region, U-Net.

Longfei Chen is with the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and with the International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: chenlongfei19@mails.ucas.ac.cn).

Wanchang Zhang is with the Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China (e-mail: zhangwc@radi.ac.cn).

Yaning Yi is with the National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China (e-mail: yaningyi@ninhm.ac.cn).

Zhijie Zhang is with the Department of Geography, University of Connecticut, Storrs, CT 06269 USA (e-mail: zhijie.zhang@uconn.edu).

Shijun Chao is with the Qinghai Research and Design Institute of Environmental Sciences, Xining 810007, China (e-mail: chaoshijun@163.com).

Digital Object Identifier 10.1109/JSTARS.2022.3189277

I. INTRODUCTION

MOUNTAIN glaciers are very sensitive to climate change, and widely recognized as an amplifier of climate change on regional and even global scales [1]. In the context of global warming, most glaciers have retreated and have exhibited negative mass balances since the 1990s [2], which affects the regional runoff and global sea level [3], [4]. Glacier mass loss has been confirmed to contribute to 30% of global sea level rise [5].

The Tibetan Plateau contains the largest volume of mountain glaciers in the mid-latitude regions due to its very high elevations [6]. As the headstream of the Yangtze, Yellow, Lancang, Brahmaputra, Ganges, and Indus rivers, Tibetan Plateau is known as the water tower of Asia [7]. The contributions of meltwater originating from snow and glaciers to the Yangtze, Yellow, Brahmaputra, and Ganges rivers are 8%, 8%, 27%, and 10%, respectively [8]. The food security of more than 63 million people living in these drainage basins will be threatened as a result of reduced water availability [8]. Therefore, it is significant to observe the dynamics of glaciers and explore their responses to climatic change.

Glacier outline extraction is the basis of studying glacier volume and mass changes [3]. Traditional methods of obtaining glacier information rely on field surveys, which are very time-consuming and not feasible because glaciers are usually located in remote and inaccessible high-mountain terrain [9]. The development of remote sensing techniques provides a more feasible method of rapid glacier extraction [10]. In general, the methods of glacier outline extraction from remote sensing images can be summarized into two categories: 1) visual interpretation and 2) automated and semi-automated methods, including the band ratio, normalized difference snow index, supervised classification, and decision tree methods [11]. Fully visual interpretation is widely considered to be the most accurate approach and is usually used to evaluate the accuracy of other automatic extraction approaches. However, visual interpretation is very time consuming and expensive, and the results have a large time lag [12]. For example, the First Chinese Glacier Inventory was performed by fully visual interpretation in 24 years (1978–2002). The Second Chinese Glacier Inventory (SCGI), obtained by band radio method and visual interpretation, integrated into Randolph Glacier Inventory (RGI 6.0) released in 2017, is currently the latest and most accurate glacier outlines in China. But it was still performed in 7 years (2006–2013) and only

obtained the glacier outlines for 2008 [13]. The current automated and semi-automated methods can effectively extract clean glacier areas, but fail to accurately extract glaciers contaminated by snow cover, clouds, and shadows [14]. Moreover, glacier systems are complex, as conditions such as weather, surrounding environment, and glacier dynamic behaviors greatly vary among different places and seasons [15]. The optimal threshold for traditional automatic segmentation of glaciers also varies among different regions, environments, and even at different times, which limits large-scale and long-term glacier extraction. Therefore, it is necessary to explore a more accurate and general method for large-scale and long-term glacier extraction that can adapt to different environments and times.

The vigorous development of deep learning methods has provided a new method for glacier extraction [16], [17]. The application of deep learning in glacier extraction can be summarized in two categories: applying or modifying existing networks, and developing novel networks. Baumhoer *et al.* [18] applied a modified U-Net to extract Antarctic glacier and ice shelf front from Sentinel-1 images. Mohajerani *et al.* [19] used a convolutional neural network (CNN) with a U-Net architecture to automatically detect glacier calving fronts in Greenland from Landsat images. Zhang *et al.* [20] also applied the U-Net architecture to extract ice sheet in Greenland. Zhang *et al.* [21] evaluated the performance of DeepLabv3+ and U-Net combined with three histogram modification strategies, aimed to find the best combination of deep learning networks and histogram modification methods. These research focused on improving the original network by modifying the network parameters or the strategies for classification. Xie *et al.* [22] developed a GlacierNet based on SegNet model for mapping complex glacier boundaries. Robson *et al.* [23] developed a novel CNN method in conjunction with object-based image analysis (OBIA), named CNN–OBIA method, for identifying rock glaciers from satellite data. He *et al.* [24] proposed a self-attention U-Net (SaU-Net) network to extract glacial lakes in the Alatau mountains. Tian *et al.* [25] developed an improved U-Net with channel squeeze and excitation (cSE) model, named U-Net+cSE, to extract glacier in the Pamir Plateau. These research developed novel networks based on existing machine learning methods to extract glacier outlines. The above studies showed that the modified network or the proposed novel network can accurately and efficiently extract the glaciers in the study area. However, they only tested the feasibility of the method but did not further extract long time-series glacier outlines which are important to study glacier mass balance and climatic change.

Yi and Zhang [26] developed a novel deep learning network named LandsNet, based on U-Net architecture. LandsNet has been shown to have a higher accuracy than ResUNet and DeepUNet when used to distinguish between landslides and nonlandslides. But this network has not yet been applied to glacier extraction. In fact, similar to landslide extraction, glacier outline extraction can also be regarded as an image segmentation task that only contains two categories (glacier and nonglacier). Therefore, to verify the feasibility and effectiveness of LandsNet architecture for glacier outline extraction, this study applied a modified LandsNet (M-LandsNet) to extract the
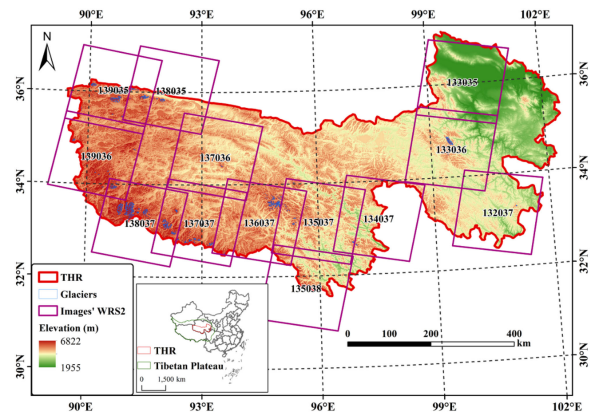

Fig. 1. Location of the study area. The six digit numbers in purple boxes represent the Landsat images' numbers following the Worldwide Reference System 2 (WRS2). The first and last three numbers indicate the path and row numbers in the WRS2, respectively.

glacier outlines in the Three-Rivers Headwater Region (THR) of the Tibetan Plateau. Meanwhile, the band ratio method and six deep learning approaches (U-Net, U-Net++, GlacierNet, SaU-Net, U-Net+cSE, and LandsNet) were also used for the same extraction task. The optimal model for glacier outline extraction was finally obtained through comparison using two scenes. We further extracted the glacier outlines in the THR during 1986–2021 in a total of 12 periods using M-LandsNet, and analyzed the temporal and spatial changes in the glacier area in this region over the past 35 years, which is also the innovation of this study. This article can provide a long time-series dataset of glacier outlines for the estimation of regional water resources storage in the THR.

II. MATERIALS AND METHODS

A. Study Area

The THR (89°24'–102°27'E, 31°32'–37°8'N), situated in the eastern part of the Tibetan Plateau (see Fig. 1), is the origin of the Yangtze River, the Yellow River, and the Lancang River. It is also an important supply of fresh water resources in China. The THR spans about 1183 km from west to east and 634 km from south to north, with an area of about 3.63×10^5 km^2. Its elevation ranges from 1955 to 6822 m and decreases from west to east. According to the SCGI data for 2008, the THR contains 1555 glaciers and covers an area of 2297.93 km^2, and these glaciers are distributed in the northwest, southwest, central, and eastern parts of this region. The annual mean temperature and annual precipitation are about 2 °C and 423 mm [27], respectively. The climate of the THR is dry and cold in the west but warm and humid in the east, which causes the rates of glacier ablation to greatly vary among the western, middle and eastern areas, so glaciers in these regions exhibit very different surface and geometric properties [28].

B. Data Sources and Preprocessing

Obtaining a reliable glacier inventory is the first step to training a deep learning network. In this study, the SCGI was

TABLE I
LANDSAT-5 IMAGES' NUMBERS, TIMES, AND CLOUD COVER PERCENTAGES
USED FOR DEEP LEARNING NETWORKS, AND THE GLACIER AREA IN THE
CORRESPONDING IMAGES

WRS2	Time	Cloud cover (%)	Glacier area (km^2)
132 037	20070714	3.00	3.11
133 035	20070923	4.00	0.67
133 036	20090811	1.00	102.97
134 037	20070930	1.00	9.24
135 037	20070921	2.00	38.07
135 038	20070820	13.00	16.36
136 037	20100429	32.00	197.83
137 036	20070919	0.00	2.20
137 037	20070919	1.00	340.95
138 035	20090830/20100902	0.00/1.00	261.02
138 037	20070505	3.00	1164.88
139 035	20060914/20070731	3.00/2.00	778.93
139 036	20070731	1.00	152.34

Note: The total glacier areas in the 13 images is much larger than the area of the glaciers in the THR due to the overlap between adjacent images.

used as the true glacier outlines, and it was treated as the ground truth. The SCGI investigated by Liu *et al.* [13] was derived from the National Tibetan Plateau Data Center.[1] The SCGI in the THR was obtained using the band ratio method followed by manual adjustment based on Landsat-5 Thematic Mapper (TM) images [29]. Thirteen Landsat-5 images were used to identify the glacier outlines in the THR (see Fig. 1). However, affected by clouds and seasonal snow cover, the glacier outlines were actually extracted based on satellite images from 2006 to 2010, not images taken in one year (see Table I). Field global positioning system measurements and verification through high-resolution remote sensing images revealed that the glacier area of SCGI has a high accuracy, with a relative error of 3.20% [13]. The SCGI can meet the accuracy requirements of scientific research, and has been widely used for research on glaciers [29], [30].

The Landsat-5 (TM) and -8 (Operational Land Imager) images were acquired from the Global Visualization Viewer website of the United States Geological Survey.[2] The Landsat Collection 2 Level-1 product was used in this study. This product consists of quantized and calibrated scaled Digital Numbers (DN) representing the multispectral image data [31]. This product has been registered by the Landsat Ground Control Points from the Global Land Survey database [32]. The geo-registration is consistent and within prescribed image-to-image tolerances of ≤12 m (about half a pixel) radial root mean square errors [33]. In this study, to ensure consistency with the SCGI, we used the same Landsat-5 images as the images used for SCGI in the THR when training the deep learning networks (see Table I). When performing long time-series glacier outline extraction, the Landsat-5 and Landsat-8 images taken during 1986–2012 and 2013–2021 were used, respectively.

[1][Online]. Available: http://data.tpdc.ac.cn
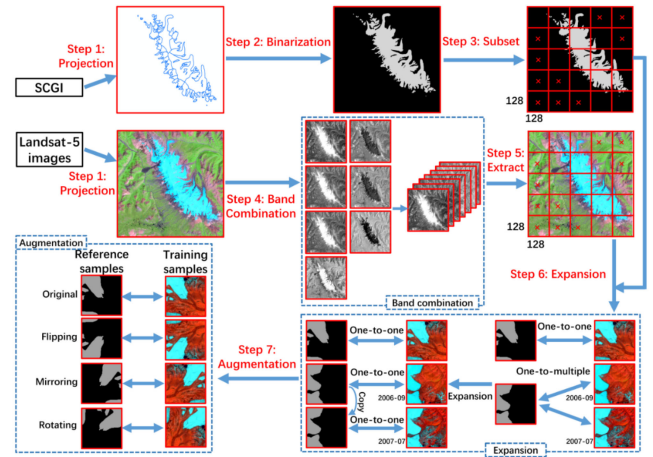[2][Online]. Available: https://glovis.usgs.gov/

Fig. 2. Data preprocessing workflow. The red "×" indicates that the subimages contain no glacier (useless samples) and were not used in the subsequent steps. The contents of the blue dashed boxes indicate the detailed explanations of this step.

The data needed to be preprocessed using the following steps before it was used to train the deep learning models (see Fig. 2).

Step 1: The SCGI for the THR and the Landsat-5 images were projected onto the same coordinate system, i.e., the 1984 World Geodetic System (WGS 84) with Universal Transverse Mercator Zone 46 North.

Step 2: The SCGI was binarized to binary labels, with 1 and 0 representing glacier and nonglacier areas, respectively.

Step 3: The size of original satellite images was large. To make it compatible with the network input requirements, the original SCGI images were subset into subimages with the size of 128×128. One subimage only contained part of the glacier for some large glaciers. The nonglacial area (3.63×10^5 km^2) is far greater than the glacial area (2297.93 km^2) in the THR, so the positive and negative samples are extremely unbalanced, which has negative impacts on model performance [34]. To maintain the balance of positive and negative samples as much as possible, the subimages without glacier were removed from the samples [35], [36] and were not used in the subsequent steps. In this study, a total of 692 SCGI subimages were retained and used for the networks and were regarded as the ground truth.

Step 4: The original Landsat-5 images contained seven single images corresponding to seven bands. To facilitate the correspondence between the image and the label, the original seven single images were combined into one 7-band image.

Step 5: The combined Landsat-5 images were extracted using the retained SCGI subimages to obtain the Landsat-5 subimages. In theory, we should have obtained 692 Landsat-5 subimages. However, due to the overlap between adjacent images and the use of images with multiple phases, we actually obtained 999 Landsat-5 subimages. These subimages were used as the training samples of the networks.

Step 6: The number of Landsat-5 subimages was far greater than the number of SCGI subimages as the Landsat-5 images for multiple phases were used for the interpretation
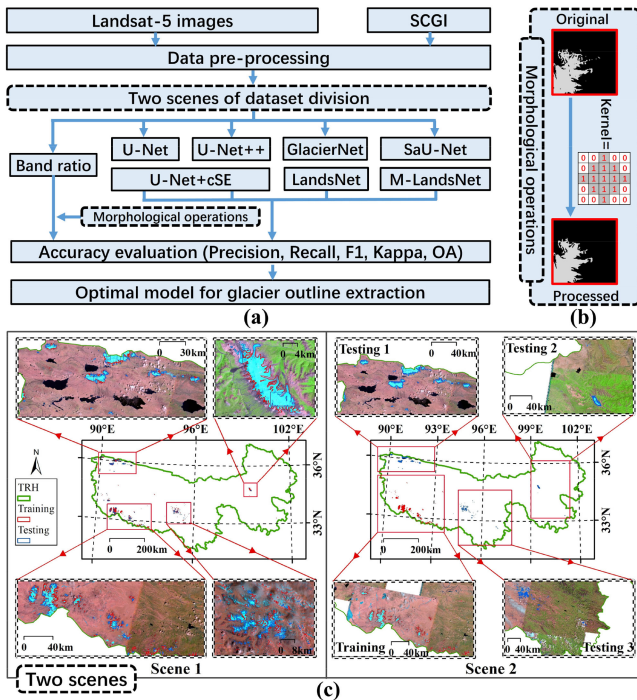
Fig. 3. Overview of the general workflow (a), the description of morphological operations (b) and two scenes (c).

and the adjacent images usually partially overlapped. This phenomenon means that one SCGI subimage may correspond to multiple Landsat-5 subimages. We expanded the number of SCGI subimages to 999 to achieve a one-to-one correspondence between the SCGI and Landsat-5 sub-images. More specifically, if an SCGI subimage named "A" corresponds to two Landsat-5 subimages named "a2006-09" and "a2007-07," respectively, then we copy this SCGI subimage, and rename them to "A2006-09" and "A2007-07," respectively, so that each Landsat-5 subimage can correspond to an SCGI subimage. The expansion method is shown in Fig. 2.

Step 7: To teach the network the desired invariance and robustness properties, the ground truth and training samples were processed using the same data augmentation method (including flipping, mirroring, and rotating). Finally, 3996 ground truth and training samples with a size of 128×128 were obtained.

C. Methodology

In this article, the glacier outlines in the THR were extracted using eight methods: the band ratio method, U-Net, U-Net++, GlacierNet, SaU-Net, U-Net+cSE, LandsNet, and M-LandsNet. The extraction accuracies of the eight methods were assessed using five evaluation metrics: the precision, recall, F1 score (F1), Kappa coefficient (Kappa), and overall accuracy (OA) [37]. An overview of the general workflow of this study is shown in Fig. 3(a).

The original Landsat-5 and SCGI datasets required preprocessing, the workflow of the data preprocessing is shown in Fig. 2. The extraction results of the band ratio method were

processed using morphological operations to remove the noise and improve the extraction accuracy. A filter matrix with a size of 5×5 was applied in the morphological operations in this study [see Fig. 3(b)]. To comprehensively compare the applicability of the eight methods in different regions and under different sample conditions, we selected two scenes to compare the extraction performances of the eight methods [see Fig. 3(c)]. The two scenes are described in detail below.

Scene 1: The dataset was randomly divided into 70% (2797) training samples and 30% (1199) testing samples. The training and testing samples were mutually exclusive. The parameter optimization processes of models were performed by randomly dividing the training samples into 80% training set and 20% validation set, and did not use the testing samples. This scene was used to evaluate the overall performances of the networks in the THR. However, it failed to reflect the differences in the generalization abilities of the networks since the testing and training samples were randomly generated in the same area and had a high similarity. To verify and compare the generalization abilities of the networks, we set scene 2.

Scene 2: In this scene, the training and testing areas were mutually exclusive. The dataset from the images with Worldwide Reference System 2 (WRS2s) of 139 036, 137 036, 138 037, and 137 037 was selected for training, and then, the trained model was applied to three testing areas. Testing area 1 was the images with WRS2s of 139 035 and 138 035. In testing area 1, the testing and training samples had similar geographic environments but different geographical locations compared to those of the training area, and the glaciers were only slightly covered by clouds and fog. This testing area was used to compare the generalization abilities of the models when the testing samples were similar to the training samples. Testing area 2 included the images with WRS2s of 133 035, 133 036, and 132 037. The glaciers in testing area 2 were also very clear. However, compared with the training area, the ground of testing area 2 has more vegetation cover, and the glaciers in this area have larger glacier ablation rate due to lower elevation, warmer and wetter climate. So the geographic environment surrounding the glaciers was quite different from that in the training area. Testing area 3 consisted of four images with WRS2s of 136 037, 135 037, 134 037, and 133 038. In testing area 3, the geographic environment surrounding the glaciers was quite different from that in the training area. Moreover, most of the glaciers in testing area 3 were covered by clouds and fog. Testing areas 2 and 3 were used to compare the generalization abilities of the models when the testing samples were quite different from the training samples. Therefore, the comparison of the two scenes was used to comprehensively evaluate the performances of the eight methods in different regions and under different sample conditions.

D. Band Ratio Method

The band ratio method is considered to be the fastest, simplest, most accurate, and most robust method of glacier extraction, compared with other automated and semi-automated methods [38]. This method is based on the low reflectance of ice and snow

in the shortwave infrared (SWIR) and the high reflectance of ice and snow in the visible spectrum . When a threshold is applied in the band ratio method with equivalent bands TM3 (red) and TM5 (SWIR), the small differences in the spectral features of the objects are greatly enhanced and glacier extraction can be accomplished. The formula is as follows:

$$\text{Ratio} = \frac{\text{Band}_3}{\text{Band}_5} \tag{1}$$

where Band_3 and Band_5 indicate the DN of bands 3 and 5 in the Landsat-5 images, respectively. Ratio is the result of the band ratio method. The determination of the threshold is the key to accurate glacier outline extraction. Typical values of the threshold are 2.0 ± 0.5 [38]. That is, the areas with radio $>$ threshold and radio $\leq$ threshold are treated as glacier and nonglacier areas, respectively.

E. U-Net and U-Net++

The U-Net was proposed by Ronneberger et al. [39] and is a two-dimensional image segmentation network based on fully convolutional network. Due to its ability to accurately segment images based on very little training data [39], the network attracted increasing attention in the field of remote sensing [16]. The network architecture consists of a contracting path and an expansive path, and it contains a total of 23 convolutional layers. The contracting path can be seen as the repeated applications of two successive 3×3 convolutions, each followed by a rectified linear unit and a 2×2 max pooling operation. The expansive path consists of the repeated application of an up-sampling layer and two 3×3 convolutions. In addition, the expansive path is concatenated with the correspondingly cropped feature map from the contracting path. Finally, each 64-component feature vector is mapped to the desired number of classes through a 1×1 convolution layer, and the image segmentation is performed.

One of the key ideas of U-Net is to skip connections. Skip connections have shown to help recover the full spatial resolution at the network output, making fully convolutional methods suitable for semantic segmentation [40]. On the basis of U-Net, Zhou et al. [40] proposed a new segmentation architecture based on nested and dense skip connections, named U-Net++. It improves U-Net by adding dense convolution blocks between encoder and decoder, and it reduces the semantic gap between feature maps generated by the encoder module and those generated by the decoder module. In this study, U-Net and U-Net++ were used to compare the performance for glacier extraction with M-LandsNet.

F. GlacierNet, SaU-Net, and U-Net+cSE

GlacierNet was proposed by Xie et al. [22] in 2020. It was designed using SegNet as a reference model by appropriately choosing the type, number and size of layers and filters, and encoder depth. It was used to automatically delineate the debris-covered glacier. The result indicated that the GlacierNet can effectively extract the debris-covered glacier in Karakoram and Nepal Himalayan, with F1 values ranging from 82% to 94%.
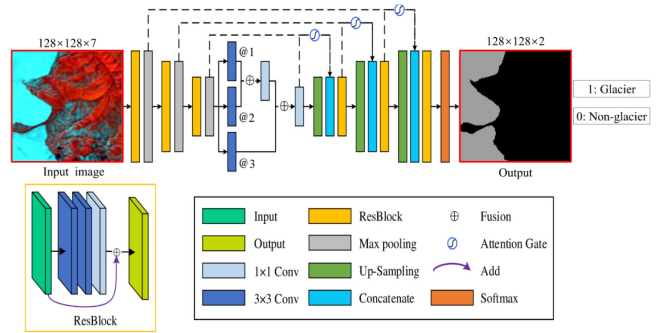

Fig. 4. Architecture of M-LandsNet. For 3×3 convolutional layer, the default dilation rate is 1. @2 denotes that the dilation rate is 2. In LandsNet, the dilation rates are 2, 4, and 6, while in M-LandsNet, the dilation rates are 1, 2, and 3, respectively.

SaU-Net was proposed by He et al. [24] in 2021 through integrating a self-attention mechanism module into U-Net network. The SaU-Net was used to extract glacial lakes from multispectral image. The experimental results in the Alatau mountains indicated that the proposed SaU-Net can effectively segment glacial lakes. Compared with the standard U-Net network, the true positive of SaU-Net increased by 15.95% for the combination of 3, 5, and 6 bands and by 5.79% for all bands.

U-Net+cSE, developed by Tian et al. [25] in 2022, was used to extract mountain glaciers. This network improved the original U-Net by integrating a channel squeeze and excitation model. Compared with U-Net and GlacierNet, the U-Net+cSE showed higher accuracy when tested on glaciers in the Pamir Plateau.

These three networks were specially developed for the extraction of glaciers or glacial lakes. In this study, GlacierNet, SaU-Net, and U-Net+cSE were also used to compare the performance for glacier extraction with M-LandsNet.

G. M-LandsNet

LandsNet was proposed by Yi and Zhang [26]. LandsNet is a cascaded end-to-end deep learning network based on the U-Net architecture. It has been used to intelligently detect and map earthquake-triggered landslides on RapidEye satellite images. Compared with the ResUNet and DeepUNet networks, which have been commonly used for image segmentation, LandsNet was demonstrated to have a higher accuracy for landslide detection in two test sites (the Ya'an-Lushan and Jiuzhaigou regions in Sichuan Province, China).

LandsNet mainly consists of a residual block (ResBlock), an attention module, and multiscale fusion operations (see Fig. 4). The ResBlock was designed by Yi et al. [41], and it is mainly used to alleviate the vanishing gradient. The ResBlock contains two successive 3×3 convolutional layers and a 1×1 convolutional layer, and both the stride and dilation rate of the three convolutional layers are 1. The operation processes of LandsNet are as follows.

Step 1. The input image is processed through three repeated applications of the ResBlock, and each application is followed by max pooling. The max pooling uses a 2×2 convolutional layer with a stride of 2.

Step 2. Three 3×3 dilated convolutional layers with dilation rates of 2, 4, and 6 are adopted to obtain three images. The first two images are fused and then processed by a 1×1 convolutional layer. The fusion result is fused with the third image and then processed by a 1×1 convolutional layer. It should be noted that the different dilated convolution results are not fused directly but are sequentially aggregated from local to global instead, which maintains the hierarchical dependencies. [42].

Step 3. The result of step 2 is processed through three repeated applications of the up-sampling layer, and each application is followed by a concatenation layer and a ResBlock layer. The up-sampling uses a 2×2 convolutional layer. Three concatenation layers are obtained by concatenating the third result of the max pooling with the result of step 2 through an attention gate, concatenating the second result of the max pooling with the fourth ResBlock result through an attention gate, and concatenating the first result of the max pooling with the fifth ResBlock result through an attention gate.

Step 4. The sixth ResBlock result is mapped to between 0 and 1 through a softmax layer, and then, the final output binary image is obtained. The values of 1 and 0 indicate glacier and nonglacier areas, respectively. Refer to paper [26] for more details about the LandsNet architecture and the parameters.

The original LandsNet used three 3×3 dilated convolutional layers with dilation rates of 2, 4, and 6 when performing multiscale fusion. Using dilated convolution can enlarge the receptive field without reducing the resolution [43], [44]. A larger dilation rate is more effective for the segmentation of large objects, but it may not be able to extract some of the small objects [45]. Compared with earthquake-triggered landslides, the sizes of glaciers in different regions can vary significantly. There are many small glaciers in the THR, and these glaciers are fragmented. Most glaciers only occupied a small proportion in the subimages with the size of 128×128. A large dilation rate may not be able to extract these small and fragmented glaciers. Therefore, in this study, we modified the dilation rates from 2, 4, and 6 to 1, 2, and 3, named M-LandsNet, when performing the multiscale fusion based on the characteristics of the glaciers in the THR.

The deep learning methods were implemented based on Keras using the Tensorflow framework as the backend in the Windows system. The networks were trained in the Kaggle platform,[3] which provides high-performance graphic processing unit cloud servers for free. In this study, the glorot normal initializer [46] was adopted to initialize weights and parameters of networks, the cross-entropy loss was employed in the training process, and the Adam optimizer [47] was employed to optimize the training loss. When training the networks, a batch size of 4 was adopted in the experiment. We set the initial learning rate to 0.001, but it was divided by a factor of 10 after every ten epochs. This variable learning rate was adopted because the step size should decrease as the result approached an optimal value [48]. To prevent overfitting, all deep learning models were trained

using early stopping with a patience parameter of ten epochs [49], which means that if the validation accuracy is not improved after a period of ten epochs, the training will stop. In this study, the models were approximately converged within 80 epochs. The model with the highest validation accuracy was saved, and used to identify the glacier from the testing samples. It should be noted that the networks used the same data, and the training and testing samples were mutually exclusive.

H. Evaluation Metrics and Glacier Area Uncertainty Estimation

In this article, five evaluation metrics, including the precision, recall, F1, Kappa, and OA, were used to compare the accuracies of the eight methods of glacier outline extraction. First, we specified that the glacier and nonglacier areas were positive and negative samples, respectively. The formulas for the precision, recall, and F1 are given in (2)–(5), respectively. The precision is the proportion of correctly predicted positive samples among the total predicted samples, and the recall is the proportion of correctly predicted positive samples among the total actual positive samples. It is one-sided to merely compare the precision and recall. F1, Kappa, and OA are usually used to comprehensively evaluate the goodness of the model. F1 is actually the geometric mean between the precision and recall [50]. Kappa is a robust metric used to measure the agreement between the predicted and actual classification results. OA is the proportion of correctly predicted samples to the total samples

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

$$\text{OA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \tag{5}$$

where TP, FP, FN, and TN denote true positive, false positive, false negative, and true negative, respectively.

Glacier area uncertainty estimation is necessary to assess the significance of the results and to avoid misinterpretation of the mapping of the glacier area [51]. The potential error of glacier mapping mainly arises from positional error [52]. The error can be calculated based on a buffer drawn around the glacier outlines as suggested by Bolch *et al.* [52] and Granshaw and Fountain [53]. Spatial resolution of Landsat images is 30 m and the registration error of these images is about half a pixel [33], so a 15 m buffer size was used to calculate the mapping uncertainty of the glacier area. The uncertainty of the change in the glacier area (ε) can be estimated using the following equation proposed by Hall *et al.* [54]:

$$\varepsilon = \sqrt{(e_1)^2 + (e_2)^2} \tag{6}$$

where e_1 and e_2 denote the estimated errors associated with the glacier area of two different time periods.

[3][Online]. Available: https://www.kaggle.com/

TABLE II
Precision, Recall, f1, Kappa, and OA of the Band Ratio Method for Glacier Outline Extraction in the THR for 12 Thresholds

Thresholds	Precision (%)	Recall (%)	F1 (%)	Kappa (%)	OA (%)
1.5	84.98	**98.01**	91.03	88.07	95.53
1.7	86.79	97.69	91.92	89.29	96.02
1.9	88.22	97.36	92.57	90.18	96.38
2.1	89.36	96.98	93.01	90.79	96.62
2.3	90.35	96.58	93.36	91.27	96.82
2.5	91.15	96.12	93.57	91.56	96.94
2.7	91.77	95.68	93.68	91.73	97.01
2.9	92.34	95.21	93.75	91.83	97.06
3.1*	92.82	94.75	**93.77**	**91.87**	**97.08**
3.3	93.24	94.21	93.72	91.82	97.07
3.5	93.64	93.49	93.57	91.63	97.02
3.7	**93.96**	92.6	93.27	91.27	96.91

Note: The numbers in bold indicate the maximum values for the 12 thresholds, and ∗ indicates the optimal threshold.

III. Results

A. Determination of the Optimal Threshold for the Band Ratio Method

In this study, according to the typical threshold of 2.0 ± 0.5 suggested by Paul *et al.* [38], the accuracy of the band ratio method for glacier outline extraction was evaluated for 12 different thresholds (1.5 to 3.7 with an interval of 0.2). The optimal extraction result was used for the subsequent comparison.

The precision of the extraction results increased from 84.98% to 93.96% and the recall decreased from 98.01% to 92.60% as the threshold increased from 1.5 to 3.7 (see Table II). F1, Kappa, and OA initially increased and then decreased. When the threshold was 3.1, F1, Kappa, and OA reached their maximum values of 93.77%, 91.87%, and 97.08%, respectively. The F1, Kappa, and OA of the extraction results ranged from 91.03% to 93.77%, from 88.07% to 91.87%, and from 95.53% to 97.08%, respectively. This indicates that the selection of the optimal threshold can contribute about 2.74% to the F1, 3.8% to the Kappa, and 1.55% to the OA. The optimal threshold for the band ratio method was determined to be 3.1 in the THR. Therefore, the extraction results of the band ratio method for a threshold of 3.1 were used for the comparison with the seven deep learning methods.

B. Visual Comparison Between Ground Truth and the Extraction Results of the Eight Methods

In this article, we selected two scenes to compare the performance of the eight methods, i.e., the band ratio method, U-Net, U-Net++, GlacierNet, SaU-Net, U-Net+cSE, LandsNet, and M-LandsNet, for glacier outline extraction in the THR. The results of the glacier outline extraction obtained by applying the eight methods to the two scenes are shown in Fig. 5–8.

When the extraction results of the eight methods for scene 1 were visually compared, it was found that except for GlacierNet, the false positive (red) and false negative (yellow) areas of the other six deep learning methods were significantly smaller than those of the band ratio method [see Fig. 5(a)–(i)]. This result indicated that except for GlacierNet, the other six deep learning methods performed better than the band ratio method even though the optimal threshold was selected for the band ratio method. Fig. 5(j)–(t) show the results of the eight methods of glacier outline extraction for scene 1 in more detail. The band ratio method tended to identify snow cover and water as glaciers, and the identified glaciers were usually larger than the actual glaciers. U-Net, U-Net++, and U-Net+cSE had the same problem as the band ratio method when the snow cover was thick [see Fig. 5(n)–(p)]. However, GlacierNet, SaU-Net, and LandsNet tended to omit some glaciers on the edge of the glaciers [see Fig. 5(l) and (r)]. M-LandsNet distinguished between snow, water, and glaciers better, and the obtained glacier outlines were more consistent with the ground truth.

In testing area 1, the band ratio method, U-Net, U-Net++, and U-Net+cSE tended to identify snow cover as glaciers, while LandsNet tended to omit some glaciers. GlacierNet and SaU-Net had these two problems [see Fig. 6(j)−(t)]. When the glaciers were partly covered by thin clouds [see Fig. 6(n)], the eight methods were affected by the clouds to some extent. The band ratio method, U-Net and SaU-Net were nearly completely unable to extract the glaciers, while U-Net++, GlacierNet, U-Net+cSE, LandsNet, and M-LandsNet could extract part of the glaciers. When there are water and thick snow cover near the glaciers [see Fig. 6(p), (s), (t)], the band ratio method, U-Net, U-Net++, and LandsNet identified a large area of water as glaciers, and the GlacierNet and SaU-Net identified a large area of snow as glaciers. U-Net+cSE and M-LandsNet were able to better distinguish between the water, snow cover, and glaciers.

In testing area 2, the performances of the eight methods decreased to a certain extent compared to their performances in testing area 1, that is, the proportions of false positive and false negative areas increased [see Fig. 7(a)–(i)]. However, the comparison of the eight methods showed that M-LandsNet still had a higher accuracy based on a visual comparison. More specifically, for areas covered by snow, the performances of the seven deep learning methods were better than that of the band ratio method [see Fig. 7(j)−(t)]. GlacierNet, SaU-Net, and LandsNet tended to seriously omit the ends of some glaciers and areas where the color of the glacier's surface was different from that of the surrounding area, and U-Net also had this problem but not as seriously as LandsNet [see Fig. 7(k), (p), (s), and (t)]. The glaciers identified by U-Net++, U-Net+cSE, and M-LandsNet were closer to ground truth, but U-Net++ and U-Net+cSE tended to identify more snow cover as glaciers [see Fig. 7(k) and (s)]. However, in the debris-covered glacier areas, the eight methods failed to effectively extract the glacier outlines [see Fig. 7(m) and (q)]. Compared with the seven deep learning methods, the band ratio method identified more glaciers in these debris-covered glacier areas. Nevertheless, for the entire testing area, M-LandsNet still performed better.

In testing area 3, the performances of the eight methods decreased significantly compared to their performances in testing area 1 [see Fig. 8(a)−(i)]. The band ratio method tended to identify snow cover as glaciers when the areas were covered
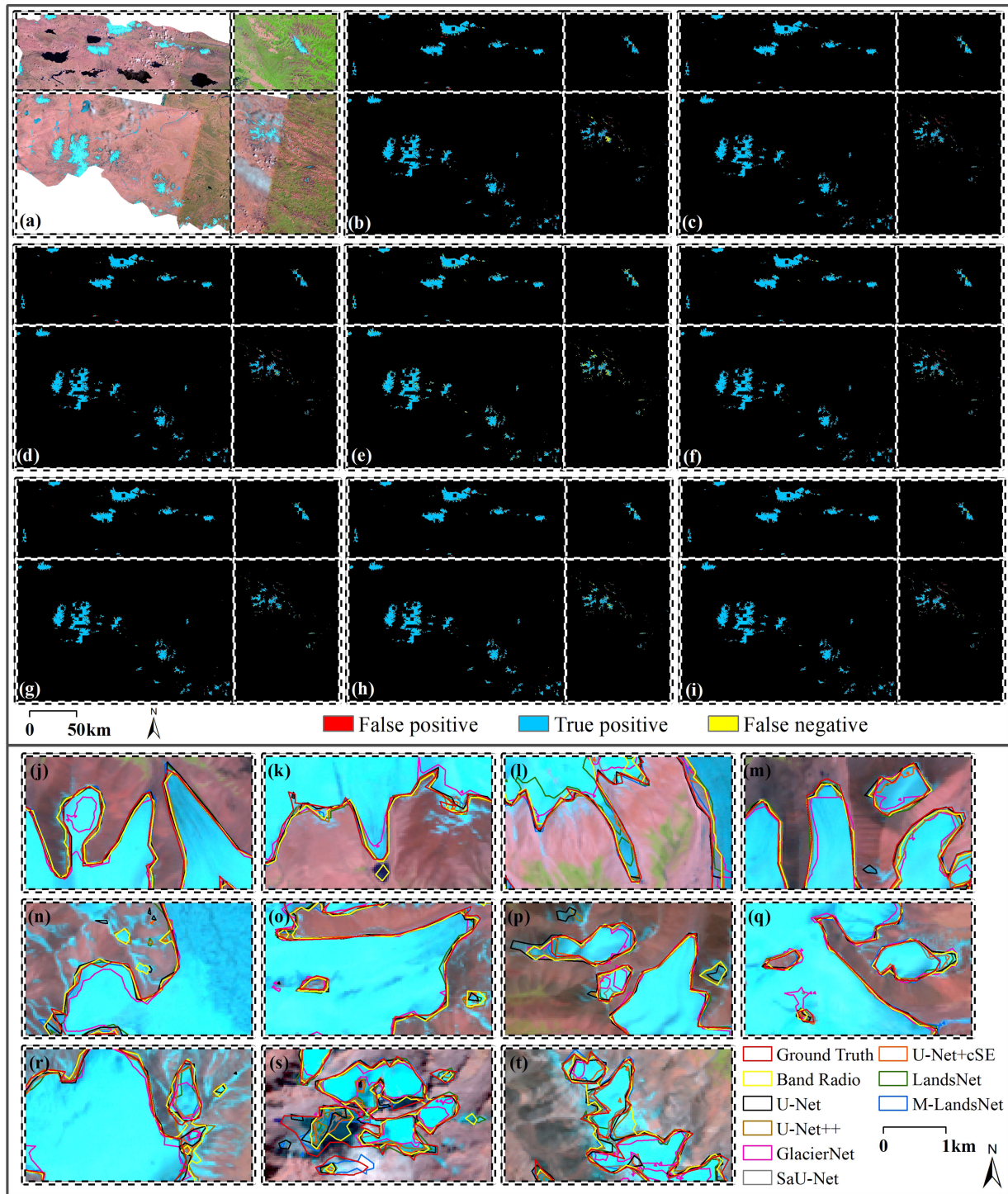
Fig. 5. Comparison of the ground truth and the results of the eight methods of glacier outline extraction for scene 1. (a) A false color composite Landsat-5 image (band7-SWIR, band4-near infrared (NIR), and band3-Red for R/G/B); (b) Band ratio method; (c) U-Net; (d) U-Net++; (e) GlacierNet; (f) SaU-Net; (g) U-Net+cSE; (h) LandsNet; (i) M-LandsNet; (j–t) More detailed comparison of the ground truth and the results of the eight methods.

by snow [see Fig. 8(k), (q), and (r)]. When the clouds were thin and the glaciers could still be easily interpreted visually, the band ratio method could only identify part of the glaciers, while U-Net, GlacierNet, U-Net+cSE, and LandsNet significantly overestimated the glacier areas [see Fig. 8(n)]. The glaciers identified by U-Net++, SaU-Net, and M-LandsNet were almost identical to the ground truth in this situation. When the cloud cover reached a certain thickness and it was difficult to interpret the glaciers visually, the band ratio method completely failed to extract the glacier outlines [see Fig. 8(o)]. The seven deep learning methods could identify parts of the glaciers, but their performances were unsatisfactory. Among them, the glaciers identified by M-LandsNet were closer to ground truth. Therefore, compared with the other seven methods, M-LandsNet still
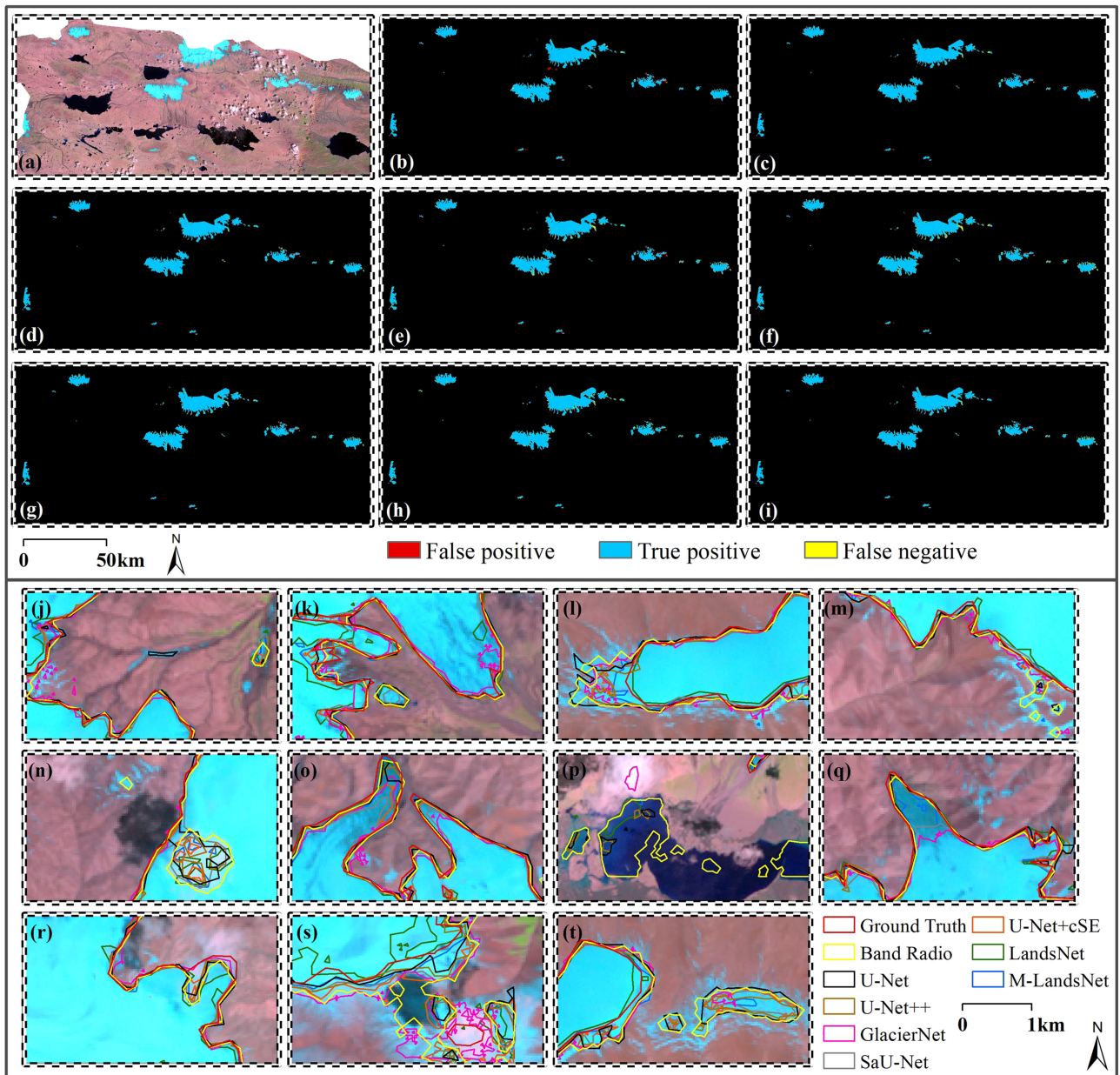
Fig. 6.   Same as in Fig. 5 but for testing area 1 in scene 2.

had better performance in this situation. As can be seen from Fig. 8(p)–(t), when the clouds were very thin and had almost no effect on the glacier outline extraction, the pros and cons of the eight methods were the same as aforementioned.

C. Accuracy Assessment of the Eight Methods

To quantitatively evaluate the performances of the eight methods, we further used five evaluation metrics to compare the accuracies of the eight methods for glacier extraction. The results of the accuracy evaluation are presented in Table III.

For scene 1, except for GlacierNet, the performances of the other six deep learning methods were better than that of the band ratio method (see Table III). M-LandsNet (F1 = 96.01%, Kappa = 94.81%, OA = 98.16%) had the best performance,

followed by U-Net+cSE, U-Net, U-Net++, SaU-Net, LandsNet, the band ratio method, and GlacierNet. Although GlacierNet had the highest precision (97.03%) and U-Net had the highest recall (96.60%), the performance of M-LandsNet was still better than that of GlacierNet and U-Net when comparing the three comprehensive evaluation metrics (i.e., F1, Kappa, and OA).

For scene 2, the F1, Kappa, and OA values of the U-Net+cSE (F1 = 95.46%, Kappa = 93.43%, OA = 97.72%) were slightly higher than that of the M-LandsNet (F1 = 95.36%, Kappa = 93.28%, OA = 97.12%) in testing area 1. However, in testing areas 2 and 3, the F1 values of the M-LandsNet were 93.63% and 80.48%, respectively; the Kappa values were 91.63% and 78.82%, respectively; the OA values were 96.95% and 96.94%, respectively; and they were the maximum values among the eight
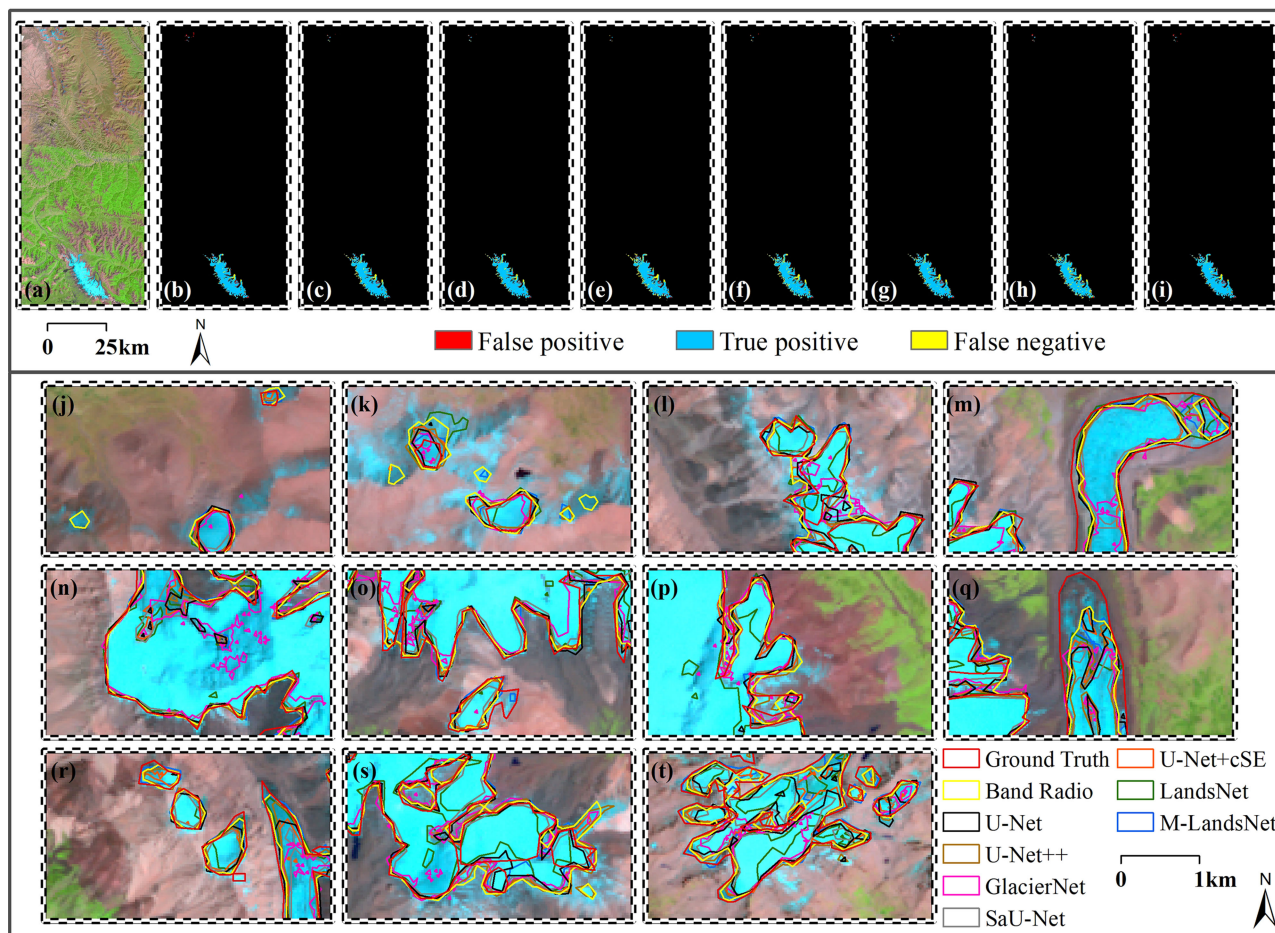
Fig. 7. Same as in Fig. 5 but for testing area 2 in scene 2.

methods. Therefore, M-LandsNet still had the best performance in scene 2. However, compared with scene 1, the performances of the eight methods decreased in scene 2. Particularly in testing area 3, the F1 values of the eight methods decreased by >15% compared to those for scene 1. Among them, the performance of U-Net exhibited the largest decrease, with F1, Kappa, and OA decreasing by 24.84%, 26.43%, and 3.44%, respectively, followed by GlacierNet (21.50%, 21.74%, and 1.09%), LandsNet (19.88%, 20.62%, and 1.79%), U-Net+cSE (17.76%, 18.33%, and 1.31%), SaU-Net (17.54%, 17.98%, and 1.23%), U-Net++ (17.51%, 17.91%, and 1.13%), the band ratio method (16.68%, 16.58%, and 0.46%), and M-LandsNet (15.53%, 15.99%, and 1.22%). Therefore, when the testing samples were different from the training samples to a certain extent, the performances of the seven deep learning methods decreased. M-LandsNet was the most accurate and stable model, and it had the best generalization ability among the eight methods.

In testing area 1, the F1, Kappa, and OA values of the seven deep learning methods decreased slightly, while the F1 of the band ratio method increased compared to that in scene 1. This result implies that changes in the geographic location have little impact on the performances of the eight methods. LandsNet had the highest precision (96.79%), while U-Net had the highest recall (95.91%). However, neither of them

was the best model. U-Net+cSE (F1 = 95.46%, Kappa = 93.43%, OA = 97.72%) slightly outperformed the M-LandsNet (F1 = 95.36%, Kappa = 93.28%, OA = 97.12%), and was the best model.

In testing area 2, the F1, Kappa, and OA values of the eight methods decreased compared to that in testing area 1, except for the Kappa and OA values of the band ratio method. The band ratio method achieved good performance, only worse than the M-LandsNet. LandsNet had the highest precision (97.76%), but the recall (81.20%) was the lowest among the eight methods, which indicates that LandsNet tends to seriously omit glaciers. Compared to testing area 1, the F1, Kappa, and OA values of M-LandsNet decreased by 1.73%, 1.65%, and 0.17%, respectively; while the F1 and Kappa values of the band ratio method only decreased by 0.67% and 0%, respectively, and the OA value even increased by 0.59%. However, the performance of M-LandsNet (F1 = 93.63%, Kappa = 91.63%, OA = 96.95%) was still slightly better than that of the band ratio method (F1 = 93.39%, Kappa = 91.29%, OA = 96.81%). These results suggest that changes in the geographical environment have an impact on the performances of the eight methods, and the deep learning methods are more affected than the band ratio method. Nevertheless, M-LandsNet was still the best method in testing area 2.
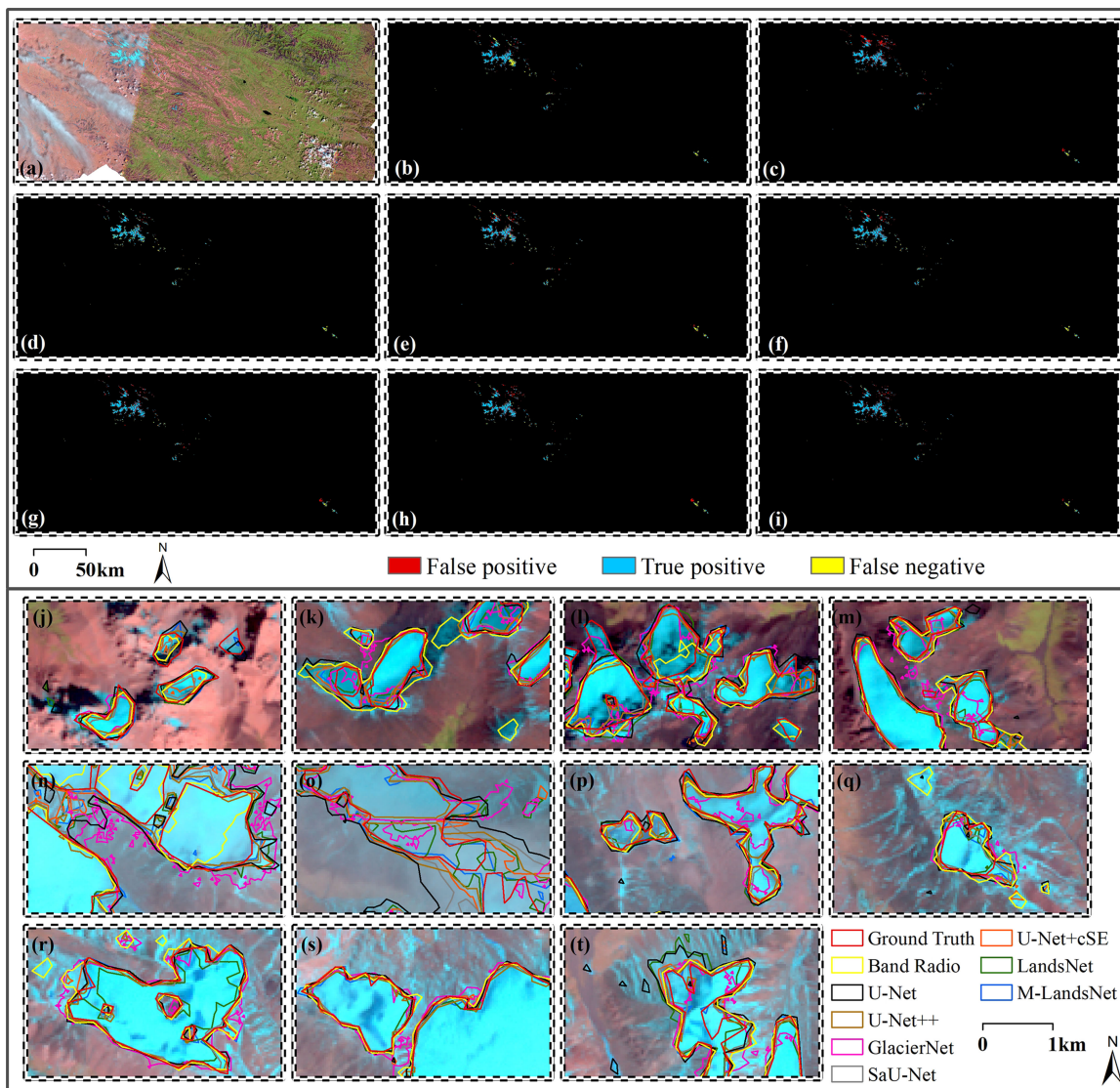
Fig. 8.   Same as in Fig. 5 but for testing area 3 in scene 2.

In testing area 3, the eight methods had unsatisfactory performances. The best model was still M-LandsNet, with F1, Kappa, and OA values of 80.48%, 78.82%, and 96.94%, respectively, followed by U-Net++, SaU-Net, U-Net+cSE, the band ratio method, LandsNet, U-Net, and GlacierNet. The SaU-Net and U-Net had the highest precision (80.12%) and recall (87.88%), respectively. However, neither of these methods was the best. This indicates that the SaU-Net tends to omit glaciers, while U-Net tends to identify other objects as glaciers. Therefore, clouds and fog had a greater impact on the performances of the eight methods, and M-LandsNet still performed best in this situation.

In the three testing areas, the precision of LandsNet in testing areas 1 and 2 and the precision of the SaU-Net in testing area 3 were the maximum values among the eight methods, but neither of them was the best method, which suggests that they tend to identify glaciers as other objects and underestimate the glacier area. The recall values of U-Net in testing areas 1 and 3 were the maximum values, but U-Net was not the best model,

and the same situation occurred for the band ratio method in testing area 2. This phenomenon suggests that U-Net and the band ratio method tend to identify more objects as glaciers and overestimated the glacier area. Expect that, the F1, Kappa, and OA values of the U-Net+cSE were slightly higher than that of the M-LandsNet in the testing area 1, the F1, Kappa, and OA values of M-LandsNet in the testing areas 2 and 3 were the maximum values among the eight methods. Therefore, M-LandsNet had the best performance and generalization ability in scene 2.

D. Long Time-Series of Glacier Outlines From 1986 to 2021

In the previous section, it was demonstrated that M-LandsNet had the highest accuracy and the best generalization ability. In this study, we used M-LandsNet to extract the glacier outlines in the THR from 1986 to 2021 in a total of 12 periods. Affected by clouds and seasonal snow cover, satellite images with less cloud and snow cover within three years were combined and used to extract the glacier outlines for one period. Thus, we obtained

TABLE III
ACCURACY COMPARISON OF THE EIGHT METHODS FOR THE GLACIER OUTLINE EXTRACTION IN THE TWO SCENES

		Precision (%)	Recall (%)	F1 (%)	Kappa (%)	OA (%)
Scene 1	Band Radio	92.55	94.91	93.71	91.79	97.08
	U-Net	95.27	**96.60**	95.93	94.69	98.10
	U-Net++	95.55	95.96	95.76	94.47	98.02
	GlacierNet	**97.03**	87.07	91.78	89.48	96.39
	SaU-Net	96.61	94.88	95.74	94.46	98.00
	U-Net+cSE	96.81	95.10	95.95	94.78	98.12
	LandsNet	95.93	94.97	95.45	94.09	97.91
	M-LandsNet	96.03	95.99	**96.01**	**94.81**	**98.16**
Scene 2	Testing area 1					
	Band Radio	93.08	95.07	94.06	91.29	96.22
	U-Net	94.11	**95.91**	95.00	92.67	96.82
	U-Net++	94.75	95.47	95.11	92.93	96.98
	GlacierNet	93.89	93.36	93.62	90.70	95.99
	SaU-Net	96.34	91.13	93.66	90.86	96.11
	U-Net+cSE	96.63	94.32	**95.46**	**93.43**	**97.72**
	LandsNet	**96.79**	93.98	95.36	93.27	97.12
	M-LandsNet	96.74	94.02	95.36	93.28	97.12
	Testing area 2					
	Band Radio	94.21	**92.59**	93.39	91.29	96.81
	U-Net	96.44	88.59	92.35	90.02	96.43
	U-Net++	95.77	89.92	92.75	90.51	96.75
	GlacierNet	94.58	84.04	89.00	85.74	94.95
	SaU-Net	97.25	83.13	89.64	86.64	95.32
	U-Net+cSE	96.50	88.83	92.51	90.23	96.50
	LandsNet	**97.76**	81.20	88.71	85.51	94.90
	M-LandsNet	95.05	92.26	**93.63**	**91.63**	**96.95**
	Testing area 3					
	Band Radio	78.26	75.84	77.03	75.21	96.62
	U-Net	59.68	**87.88**	71.09	68.26	94.66
	U-Net++	74.07	82.92	78.25	76.56	96.89
	GlacierNet	66.55	74.46	70.28	67.74	94.10
	SaU-Net	**80.12**	76.37	78.20	76.49	96.82
	U-Net+cSE	72.91	84.29	78.19	76.45	96.78
	LandsNet	71.25	80.45	75.57	73.47	96.12
	M-LandsNet	76.92	84.38	**80.48**	**78.82**	**96.94**

Note: The values in bold indicate the maximum values of the corresponding metrics

the glacier outlines for 12 different periods (see Table IV). To improve the accuracy, the dataset was further processed through manual adjustments as the extraction of some glaciers may require the use of multiple images and may be affected by clouds, seasonal snow cover, and shadows.

The glacier areas in the THR and its four subareas during 12 periods indicate that the total glacier area in the entire study area during 1986–1998 was 2518.65 ± 72.69 km^2 [see Table V(a)], and by 2021, the glacier area had reduced to 2102.25 $\pm$ 72.56 km^2, with a reduction of 416.40 ± 102.71 km^2 ($16.53 \pm 4.08\%$) over the past 35 years [see Table V(b)]. The total glacier area in the THR exhibited a decreasing trend (see Fig. 9), with a reduction rate of about 11.90 ± 2.93 km^2 a^{-1} ($0.47 \pm 0.12\%$ a^{-1}) from 1986 to 2021 [see Table V(b)]. The reduction rate of glacier area varied during the different periods. From 1986 to 2003, the total glacier area reduced by 126.11 ± 101.47 km^2 ($5.01 \pm 4.03\%$), with a reduction rate of about 7.42 ± 5.97 km^2 a^{-1} ($0.29 \pm 0.24\%$ a^{-1}). However, the total glacier area reduced from 2392.54 ± 70.80 km^2 in 2001–2003 to 2102.25 $\pm$ 72.56 km^2 in 2019–2021, with a reduction of 290.29 ± 101.38 km^2

($12.13 \pm 4.24\%$). The reduction rate from 2003 to 2021 was about 16.13 ± 5.63 km^2 a^{-1} ($0.67 \pm 0.24\%$ a^{-1}), which is almost twice that from 1986 to 2003. In the four subareas, the reduction rate in 2003−2021 was also higher than those in 1986−2003.

The glacier area exhibited a decreasing trend in the four subareas (see Fig. 9). However, the reduction rate varied among the different areas. The glacier area in the training area reduced from 1405.86 ± 37.77 km^2 in 1986−2003 to 1173.83 ± 38.70 km^2 in 2003−2021, making it the largest reduction (232.03 ± 54.08 km^2) and reduction rate ($6.63 \pm 1.55\%$ a^{-1}), followed by testing area 3 (103.83 ± 22.63 km^2 and $2.97 \pm 0.65\%$ a^{-1}), testing area 1 (52.28 ± 19.94 km^2 and $1.49 \pm 0.57\%$ a^{-1}), and testing area 2 (28.28 ± 61.15 km^2 and $0.81 \pm 1.75\%$ a^{-1}). Although the reduced area in testing area 3 was less than that in the training area, it had the largest proportional reduction ($37.08 \pm 8.08\%$) and reduction rate ($1.06 \pm 0.23\%$ a^{-1}), followed by testing area 2 ($23.70 \pm 51.25\%$ and $0.68 \pm 1.46\%$ a^{-1}), the training area ($16.50 \pm 3.85\%$ and $0.47 \pm 0.11\%$ a^{-1}) and testing area 1 ($7.33 \pm 2.80\%$ and $0.21 \pm 0.08\%$ a^{-1}) [see Table V(b)]. Similar results were also found during 1986−2003 and 2003−2021.

TABLE IV
LANDSAT IMAGES USED TO EXTRACT THE GLACIER OUTLINES

Periods	132037	133035	133036	134037	135037	135038	136037	137036	137037	138035	138037	139035	139036
1986 −1988	19870621 (0.00%)	19870815 (3.00%)	19870815 (3.00%)	19860615 (7.00%)	**19880815 (2.00%)**/ 19881002 (4.00%)	19870626 (62.00%)/ **19880815 (2.00%)**	**19870820 (23.00%)**/ 19881025 (1.00%)	**19870624 (30.00%)**/ 19880813 (52.00%)	19870624 (46%)/ **19880728 (14.00%)**/ 19881016 (2.00%)	19860730 (0.00%)	19880719 (22.00%)/ **19881023 (1.00%)**	19871012 (10.00%)	19881014 (3.00%)
1989 −1991	19910830 (59.00%)	19890921 (0.00%)	**19890921 (0.00%)**/ 19910911 (19.00%)	19890827 (6.00%)/ **19900830 (5.00%)**	19890802 (21.00%)	**19890802 (37.00%)**/ 19910824 (33.00%)	**19900828 (27.00%)**/ 19910714 (26.00%)	19910822 (4.00%)	19910822 (7.00%)/ **19911009 (0.00%)**	19900810 (15.00%)	19910914 (0.00%)	19890729 (0.00%)/ **19900817 (2.00%)**	19890729 (0.00%)
1992 −1994	19920906 (41.00%)	19940802 (4.00%)	19930831 (6.00%)	19920803 (0.00%)	19940731 (3.00%)	19920810 (20.00%)/ **19940715 (7.00%)**	19940924 (0.00%)	19940814 (0.00%)	19940814 (2.00%)	19940821 (0.00%)	19940805 (0.00%)	19930910 (0.00%)	19940929 (0.00%)
1995 −1997	**19960816 (7.00%)**/ 19970819 (26.00%)	19950821 (7.00%)	19970810 (1.00%)	19950812 (2.00%)	19950904 (32.00%)/ **19960821 (4.00%)**/ 19960821 (18.00%) (3.00%)	19950725 (3.00%)/ 19950826 (6.00%)	**19950725 (3.00%)**/ 19950826 (6.00%)	19950902 (2.00%)	**19950817 (1.00%)**/ 19960803 (12.00%)	19970728 (2.00%)	19940922 (0.00%)	19960716 (9.00%)/ **19970719 (0.00%)**	19960716 (3.00%)
1998 −2000	19990825 (40.00%)	19990731 (1.00%)	19980914 (17.00%)/ **19990731 (3.00%)**/ 20000701 (29.00%)	19990924 (4.00%)	**19990729 (13.00%)**/ 20000715 (4.00%)	19980928 (2.00%)	19980717 (0.00%)	19980910 (8.00%)	19980809 (0.00%)	19980816 (0.00%)	19980917 (42.00%)/ **20000821 (18.00%)**	19980823 (42.00%)/ **19980908 (9.00%)**/ 19990826 (42.00%)	**19990709 (5.00%)**/ 19990725 (33.00%)/ 19990826 (42.00%)
2001 −2003	20010814 (1%)	**20010720 (15.00%)**/ 20030827 (55.00%)	**20010720 (14.00%)**/ 20020824 (15.00%)/ 20030912 (10.00%)	**20020815 (6.00%)**/ 20020831 (8.00%)	20010904 (10.00%)/ **20020907 (5.00%)**	**20010904 (28.00%)**/ 20020907 (18.00%)	**20010709 (7.00%)**/ 20020829 (57.00%)	20020804 (10.00%)	20020804 (24.00%)/ 20030722 (24.00%)/ **20030807 (25.00%)**	20010723 (14.00%)	**20010723 (11.00%)**/ 20010808 (12.00%)/ 20031017 (1.00%)	20030922 (0.00%)	20011103 (2.00%)/ **20030922 (0.00%)**
2004 −2006	20040806 (0%)	20060904 (41.00%)	**20040914 (1.00%)**/ 20050715 (21.00%)	20060725 (4.00%)	20060801 (8.00%)	20040811 (45.00%)/ **20060801 (29.00%)**	**20040903 (10.00%)**/ 20060723 (41.00%)/ 20060808 (14.00%)	20060714 (11.00%)	**20040926 (22.00%)**/ 20060628 (31.00%)/ 20060714 (20.00%)	20060907 (12.00%)	20040901 (24.00%)/ **20060907 (20.00%)**	20060728 (30.00%)/ **20060914 (3.00%)**	20040519 (1.00%)/ **20040908 (16.00%)**
2007 −2009	20070714 (3%)	20070822 (5.00%)/ **20070923 (4.00%)**	20090811 (1.00%)	20070914 (5.00%)	20070921 (2.00%)	20070820 (13.00%)	20090816 (27.00%)	20070802 (9.00%)	20070802 (0.00%)/ **20070919 (1.00%)**	20090830 (0.00%)	20070505 (3.00%)	20070731 (2.00%)	**20070731 (1.00%)**/ 20090704 (13.00%)
2010 −2012	20110725 (1%)	20100814 (4.00%)	20100814 (1.00%)	20110919 (2.00%)	20110831 (2.00%)	20100828 (26.00%)/ **20110831 (11.00%)**	**20110806 (8.00%)**/ 20110822 (23.00%)	20100725 (1.00%)	20100725 (11.00%)/ **20110829 (7.00%)**	20100902 (1.00%)	20100630 (39.00%)/ **20100801 (8.00%)**	20100504 (7.00%)	**20100707 (18.00%)**/ 20110811 (29.00%)
2013 −2015	20140919 (4.72%)	20140809 (18.20%)	20150812 (1.31%)	20130813 (0.04%)	20150725 (2.06%)	20150927 (2.67%)	20130811 (5.06%)	20130802 (0.13%)	20130802 (0.51%)	**20130622 (13.99%)**/ 20150730 (13.97%)	20130809 (20.92%)/ **20151002 (0.23%)**	20130731 (0.81%)	20130731 (1.66%)
2016 −2018	20160807 (15.42%)	20160729 (0.81%)	20160627 (6.43%)/ **20160729 (3.29%)**	20160821 (2.70%)	20160727 (5.20%)	**20160727 (8.80%)**/ 20170916 (22.82%)	**20160819 (7.14%)**/ 20180825 (54.15%)	20180715 (1.58%)	20160725 (7.78%)/ **20160810 (8.35%)**	20160801 (2.54%)	**20160716 (23.85%)**/ 20171007 (0.25%)	20160909 (2.03%)	20170811 (1.51%)
2019 −2021	20200903 (5.35%)	20200825 (0.53%)	20200825 (0.12%)	20190729 (10.33%)	**20190821 (55.79%)**/ 20200823 (46.30%)	20200823 (69.13%)/ 20210725 (44.45%)/ **20210810 (65.96%)**	20200830 (6.75%)	20200922 (1.20%)	**20190819 (16.62%)**/ 20200922 (7.51%)	20200929 (0.12%)	20190725 (0.44%)	20200819 (4.03%)	20200819 (2.77%)

Note: The numbers in parentheses indicate cloud cover percentages. The bold numbers indicate the main reference images among the multiple images, and the underlined number indicates that the image taken on September 22, 1994 was used because no image was available in scene 138037 during 1995–1997.

More intuitively, we mapped the glacier outlines in the THR from 1986 to 2021 in a total of 12 periods (see Fig. 10). It was also showed that the glacier area has been shrinking since 1986, and the rate of shrinking increased over time. The comparison of the four subareas indicated that the glacier area in the training area had the largest reduction. However, the glacier area in testing area 3 had the largest proportional reduction, as the glacier was small in this subarea. Some small glacier almost completely disappeared.

IV. DISCUSSION

A. Comparison of the Eight Methods for Glacier Outline Extraction

In this article, we compared the performances of eight methods, i.e., the band ratio method, U-Net, U-Net++, GlacierNet, SaU-Net, U-Net+cSE, LandsNet, and M-LandsNet, for glacier outline extraction in the THR using two scenes. The results indicate that M-LandsNet performs best in the two scenes.

TABLE V
GLACIER AREA (A) AND AREA CHANGES (B) IN THE THR AND ITS FOUR SUBAREAS DURING 1986−2021

(a):

Periods	Total area (km²)	Training area (km²)	Testing area 1 (km²)	Testing area 2 (km²)	Testing area 3 (km²)
1986−1988	2518.65 ± 72.69	1405.86 ± 37.77	713.48 ± 13.92	119.33 ± 41.05	279.98 ± 16.90
1989−1991	2482.42 ± 72.72	1384.03 ± 37.91	710.62 ± 13.74	118.17 ± 41.43	269.59 ± 16.93
1992−1994	2468.93 ± 72.79	1373.97 ± 37.95	710.49 ± 13.70	118.18 ± 40.36	266.30 ± 17.10
1995−1997	2441.26 ± 72.62	1361.61 ± 37.85	706.20 ± 13.67	114.55 ± 43.05	258.90 ± 16.81
1998−2000	2404.11 ± 71.88	1343.79 ± 37.53	696.01 ± 13.59	114.40 ± 39.53	249.90 ± 16.80
2001−2003	2392.54 ± 70.80	1348.26 ± 36.92	693.87 ± 13.46	111.63 ± 43.06	238.78 ± 16.11
2004−2006	2344.12 ± 71.27	1319.56 ± 37.26	684.22 ± 13.43	110.12 ± 43.36	230.21 ± 16.24
2007−2009	2303.75 ± 70.51	1297.14 ± 37.18	682.25 ± 13.53	106.65 ± 43.01	217.72 ± 15.50
2010−2012	2256.21 ± 72.15	1262.80 ± 38.14	687.32 ± 13.45	100.25 ± 48.06	205.84 ± 15.75
2013−2015	2209.73 ± 73.17	1241.87 ± 38.62	672.84 ± 14.20	98.61 ± 45.60	196.41 ± 15.78
2016−2018	2163.58 ± 72.93	1219.42 ± 38.79	666.35 ± 14.05	95.43 ± 45.97	182.37 ± 15.49
2019−2021	2102.25 ± 72.56	1173.83 ± 38.70	661.20 ± 14.28	91.05 ± 45.33	176.15 ± 15.05

(b):

			1986−2003	2003−2021	1986−2021
THR	Area change	km²	−126.11 ± 101.47	−290.29 ± 101.38	−416.4 ± 102.71
		%	−5.01 ± 4.03	−12.13 ± 4.24	−16.53 ± 4.08
	Area change rate	km² a⁻¹	−7.42 ± 5.97	−16.13 ± 5.63	−11.90 ± 2.93
		% a⁻¹	−0.29 ± 0.24	−0.67 ± 0.24	−0.47 ± 0.12
Training area	Area change	km²	−57.6 ± 52.82	−174.43 ± 53.49	−232.03 ± 54.08
		%	−4.10 ± 3.76	−12.94 ± 3.97	−16.50 ± 3.85
	Area change rate	km² a⁻¹	−3.39 ± 3.11	−9.69 ± 2.97	−6.63 ± 1.55
		% a⁻¹	−0.24 ± 0.22	−0.72 ± 0.22	−0.47 ± 0.11
Testing area 1	Area change	km²	−19.61 ± 19.36	−32.67 ± 19.62	−52.28 ± 19.94
		%	−2.75 ± 2.71	−4.71 ± 2.83	−7.33 ± 2.80
	Area change rate	km² a⁻¹	−1.15 ± 1.14	−1.82 ± 1.09	−1.49 ± 0.57
		% a⁻¹	−0.16 ± 0.16	−0.26 ± 0.16	−0.21 ± 0.08
Testing area 2	Area change	km²	−7.70 ± 59.49	−20.58 ± 62.52	−28.28 ± 61.15
		%	−6.45 ± 49.85	−18.44 ± 56.01	−23.70 ± 51.25
	Area change rate	km² a⁻¹	−0.45 ± 3.50	−1.14 ± 3.47	−0.81 ± 1.75
		% a⁻¹	−0.38 ± 2.93	−1.02 ± 3.11	−0.68 ± 1.46
Testing area 3	Area change	km²	−41.20 ± 23.35	−62.63 ± 22.05	−103.83 ± 22.63
		%	−14.72 ± 8.34	−26.23 ± 9.23	−37.08 ± 8.08
	Area change rate	km² a⁻¹	−2.42 ± 1.37	−3.48 ± 1.22	−2.97 ± 0.65
		% a⁻¹	−0.87 ± 0.49	−1.46 ± 0.51	−1.06 ± 0.23

However, some issues deserve to be discussed further when comparing the eight methods.

First, for the band ratio method, we found that the F1, Kappa, and OA values initially increased and then decreased with increasing threshold. This phenomenon implies that when the threshold is too small, the band ratio method tends to identify more objects as glaciers. As the threshold increases, this problem is improved. However, when the threshold is too large, the band ratio method tends to identify glaciers as other objects. Moreover, many studies have used the typical values (2.0 ± 0.5) of the threshold [38], [55]. However, in this study, it was found that the optimal threshold was 3.1 in the THR, which is far greater than the typical threshold range. When we simply used 1.5 as the threshold for the band ratio method, the F1, Kappa, and OA values decreased by 2.74, 3.8, and 1.55%, respectively. This indicates that the optimal threshold varies greatly in different regions, and the determination of the optimal threshold is difficult.

Second, the comparison of the eight methods revealed that the band ratio method tends to identify snow cover and water as glaciers, and it fails to extract some glaciers when covered by clouds or fog. This phenomenon explains why the band ratio method has the highest recall (92.59%) in testing area 2 and the second highest precision (78.26%) in testing area 3, but it is not the best of the eight methods. Many studies have confirmed that the band ratio method will obtain confusing results when snow, water, shadows, clouds, and fog are present in conjunction with glaciers [56], [57]. U-Net tends to identify snow cover, water, and even some relatively thick clouds and fog as glaciers, which explains why its recall values in scene 1 (96.60%), testing area 1 (95.91%), and testing area 3 (87.88%) are the highest, but it is not the best of the eight methods. This result indicates that U-Net tends to identify other objects with very small spectral differences from those of glaciers as glaciers, so as to overestimate target objects, which is consistent with several previous studies [41], [58], [59]. The performance of
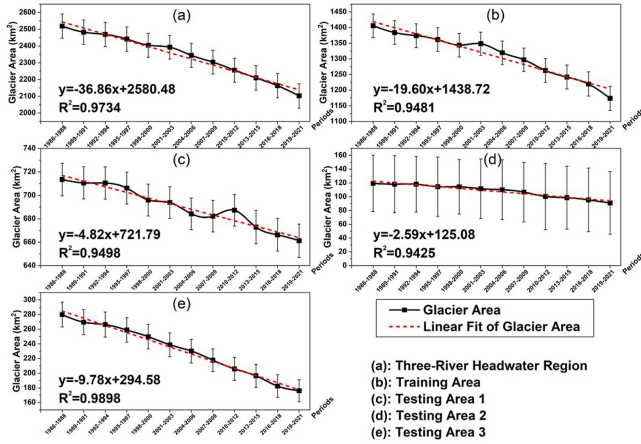
Fig. 9. Change in the glacier area over time in the THR (a) and the four subareas [training area (b), testing area 1 (c), testing area 2 (d) and testing area 3 (e)] from 1986–2021. The error bars denote the uncertainty in the respective years, and the dashed red lines are the linear fitting lines of the glacier area.
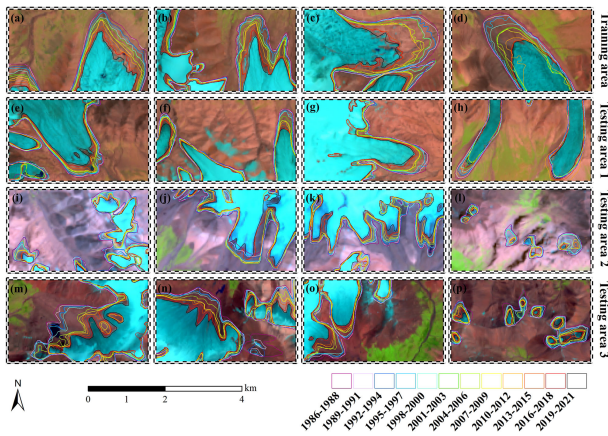

Fig. 10. Changes in glacier outlines in the THR from 1986 to 2021. The four rows of figures represent the training area, the testing area 1, the testing area 2 and the testing area 3, respectively. The base maps are the Landsat images in the period of 2019−2021.

GlacierNet and SaU-Net was worse than the U-Net, as they were developed for debris-covered glacier and glacial lake extraction, respectively, but not for bare glacier extraction. The U-Net+cSE developed for extracting bare glaciers performed better than the U-Net. Similar results were obtained in the study conducted by Tian *et al.* [25], and their results showed that GlacierNet performed worse than U-Net, while U-Net+cSE outperformed U-Net for bare glacier extraction. In contrast to U-Net, LandsNet is too strict in identifying glaciers, so it tends to omit the ends of some glaciers and areas where the color of the glacier's surface is different from that of the surrounding area. This explains why LandsNet has the highest precision but is not the best model in testing areas 1 and 2. U-Net++ had the same problems as LandsNet and U-Net, but was not as seriously as them. Therefore, U-Net++ has better performance than LandsNet and U-Net in scene 2. However, when we modified the dilation rates of LandsNet, these problem was greatly improved. When we decrease the dilation rates, M-LandsNet can capture the features and improve its ability to recognize smaller objects,

so the extraction results are not as fragmented as those of LandsNet. Therefore, it has the best performance among the eight methods for scenes 1 and 2. These results indicate that M-LandsNet, which is based on the LandsNet architecture but with a reasonable modification, is able to extract the glacier outlines in the THR well.

Third, compared to scene 1, the extraction accuracies of the eight methods in the three testing areas in scene 2 decreased to some extent, except for that of the band ratio method in testing area 1. For scene 1, the dataset was randomly divided into training and testing samples, so the testing and training samples had almost exactly the same distribution. For scene 2, the testing areas were different from the training area in terms of geographic location, geographic environment around the glacier, and weather conditions to some extent. The phenomena of different objects with the same spectral reflectance and the same objects with different spectral reflectance are often observed, especially when the environment of the object is quite different [41], which increases the difficulty of identification. For the band ratio method, when affected by changes in the geographic environment and weather conditions, the optimal threshold for this area may no longer be 3.1 since the optimal threshold of 3.1 was obtained for the entire THR. For the deep learning models, the changes in the geographic environment and the weather conditions directly led to large spectral differences in the images, so some of the features learned in the training areas may not have been applicable in the testing areas. These phenomena are primarily responsible for the decreases in the accuracies of the methods in scene 2.

Fourth, the performances of the eight methods decreased from testing area 1 to testing area 3. Among them, the eight methods performed best in testing area 1, almost the same as in scene 1, which indicates that the eight methods have good generalization abilities when the testing area is highly similar to the training area. In fact, most previous studies verified or compared the performances of deep learning models using the dataset division method used in scene 1 and in testing area 1 in scene 2 [26], [60], [61]. In testing area 2, when the glaciers were very clear but the surrounding geographic environment of the glaciers was quite different from that in the training area, the performances of the eight methods decreased to a certain extent. The changes in the geographic environment should be partly responsible for these decreases. Moreover, several debris-covered glacier areas were distributed in testing area 2. The extraction of the debris-covered glacier areas was much more difficult than the extraction of the bare glacier areas because of the spectral similarities between the debris-covered glaciers and the adjacent areas [62]. More data types [i.e., digital elevation model, slope angle, profile curvature, tangential curvature [22], and/or images with a high spatial resolution (≤1 m)] and a high radiometric resolution (i.e., Quickbird-2, GeoEye-1, WorldView-1, and WorldView-2) are usually required to train a model [56]. However, the debris-covered glaciers were not considered in this study since we only used optical images. Additionally, the area of the debris-covered glaciers in the THR is very small, and they are mainly distributed in the image with a WRS2 of 133 036, so we did not have enough debris-covered glacier samples to train the model. When the

TABLE VI
COMPARISON BETWEEN OUR GLACIER OUTLINES DATASET WITH PREVIOUS GLACIER RESEARCH WITHIN THE STUDY REGION

Regions	Years	Previous study (km^2)	Present study (km^2)	Relative error (%)	References
(a)	1992	109.21	112.45	2.97	
	2001	102.80	107.00	4.09	[63]
	2010	99.34	96.59	2.77	
	2009	103.52	102.09	1.38	SCGI
(b)	1994	196.14	195.72	0.21	
	2002	192.81	192.36	0.23	[64]
	2010	191.44	191.27	0.09	
	2007	188.80	189.90	0.58	SCGI
(c)	1992	869.17	864.12	0.58	
	2002	846.81	856.04	1.09	[65]
	2007	845.17	845.44	0.03	SCGI
(d)	2018	613.90	608.42	0.89	[66]
	2007	612.48	613.15	0.11	SCGI

geographic environment surrounding the glaciers was different from that in the training area and the glaciers in the testing area 3 were covered by clouds and fog, the performances of the eight methods decreased significantly. The large area of cloud and snow cover was primarily responsible for this phenomenon. Though M-LandsNet still performed the best in this scene, the extraction results were almost meaningless due to the low accuracy. In fact, there is currently no method to completely eliminate the influences of thick clouds and fog during glacier outline extraction. Therefore, in glacier extraction work, images with less cloud cover and fog should be selected.

B. Comparison With Previous Glacier Research Results Within the Study Area

In this study, we extracted the glacier outlines in the THR from 1986 to 2021 in a total of 12 periods using M-LandsNet. The obtained glacier outlines were then processed through manual adjustments. Except for SGCI in 2008, there is currently no other study of the glacier area in the entire THR, so we compared the glacier areas in four different regions with the results of previous studies (see Table VI). Compared with the SCGI, the relative errors of the glacier area in our study were 1.38%, 0.58%, 0.03%, and 0.11% in regions (a), (b), (c), and (d). We also compared our results with four references. Wang *et al.* [63] obtained the glacier area in region (a) through fully manual digitization, and the glacier area were 109.21, 102.80, and 99.34 km^2 in 1992, 2001, and 2010, respectively. The relative errors between their results and ours were 2.97, 4.09, and 2.77%, respectively. Jiang *et al.* [64] obtained the glacier area in region (b) of 196.14, 192.81, and 191.44 km^2 in 1994, 2002, and 2010, respectively, using supervised classification and manual interpretation. Compared with our results, the relative errors were only 0.21, 0.23, and 0.09%, respectively. Ye *et al.* [65] used band ratio method and manual interpretation to extract the glacier outlines in region

(c), with glacier area of 869.17 and 846.81 km^2 in 1992 and 2002, respectively. The relative errors between their results and ours were only 0.58 and 1.09% respectively. Zhang *et al.* [66] manually digitized the glacier outlines in region (d) from Landsat images, and obtained the glacier area of 613.90 km^2 in 2018, which was also close to our results, with a relative error of only 0.89%. These relative errors range from 0.09 to 4.09%, and are larger than that of the SCGI. The major reasons for these different results are that we used the same Landsat images as the SCGI when extracting the glacier outlines in 2007–2009 in these regions, and the M-LandsNet model was trained using the same images. Therefore, the predicted results are almost the same as the training results. However, for the other four references, the dates of the images used were different from those used in this study, and some of the images were even more than 2 years apart, which is mainly responsible for the larger relative errors. Nevertheless, with the exception of region (a), the relative errors of the other three regions were all less than 1.10%. Therefore, the glacier outline dataset obtained in this article is reliable and can effectively reflect the glacier area in the THR.

We further analyzed the temporal and spatial changes in the glacier area in the study area and found that the rate of decrease from 2003 to 2021 was higher than that from 1986 to 2003 in the entire THR and its four subareas. Ye *et al.* [67] studied the glacier changes on the Tibetan Plateau in the 1970s, 2000, and 2013, and they also found that compared with the period of 1970s–2000, the glacier area in the eastern part of the Tibetan Plateau had a greater rate of decrease from 2001 to 2013. Duan *et al.* [68] and Chen *et al.* [69] also obtained the same result that the glacier area has been shrinking more rapidly in the last 20 years by studying the glaciers in the Tanggula Mountains, which contain the training area and testing area 3. More specifically, Duan *et al.* [68] found that the reduction rate of the glacier area in the Tanggula Mountains was –1.3% a^{-1} during 2007−2015, which is more than twice of that during 1991–2007 (-0.5% a^{-1}). In this study, the reduction rate of the glacier area in the training area and the testing area 3 during 2003−2021 was also about twice of that during 1986−2003. Additionally, Jin *et al.* [70] studied the glaciers in Geladandong region in the training area, and found that the reduction rate of the glacier area was 0.30% a^{-1} during 1986−2004. In our study, it was found to be 0.24±0.22% a^{-1} during 1986−2003. Therefore, the glacier outline dataset produced in this study can also effectively reflect the temporal and spatial changes in the glacier area in the THR.

V. CONCLUSION

In this study, to verify the feasibility and effectiveness of LandsNet architecture for glacier outline extraction, we applied M-LandsNet to extract the glacier outlines in the THR. The band ratio method, U-Net, U-Net++, GlacierNet, SaU-Net, U-Net+cSE, LandsNet, and two scenes were used for comparison. The optimal model was further used to extract the long time-series glacier outlines in the THR.

The results indicate that the optimal threshold of the band ratio method for glacier outline extraction varies greatly in different regions and is difficult to determine. The band ratio

method tends to identify snow cover and water as glaciers, and it fails to extract some glaciers when they are covered by clouds or fog. U-Net tends to identify other objects with very small spectral differences from those of glaciers as glaciers. GlacierNet and SaU-Net performed worse than U-Net, while U-Net+cSE performed better than U-Net. LandsNet tends to omit some glaciers when the color of the glacier's surface is different from that of the surrounding area. U-Net++ has the same problems as U-Net and LandsNet, but is not as seriously as them. M-LandsNet has the best performance among the eight methods. However, when the testing samples are different from the training samples to a certain extent, the performances of the eight methods decrease. Weather conditions have the greatest impact on the eight methods, followed by geographic environment and geographic location. M-LandsNet is still the most accurate and stable model, and it has the best generalization ability among the eight methods. We further extracted the glacier outlines in the THR from 1986 to 2021 in a total of 12 periods using M-LandsNet and manual adjustments, and analyzed the changes in glacier area. We found that the glacier area in the THR reduced by 416.40 ± 102.71 km^2 ($16.53 \pm 4.08\%$) from 1986 to 2021. The reduction rate (16.13 ± 5.63 km^2 a^{-1} ($0.67 \pm 0.24\%$ a^{-1})) from 2003 to 2021 was almost twice of that [7.42 ± 5.97 km^2 a^{-1} ($0.29 \pm 0.24\%$ a^{-1})] from 1986 to 2003. From 1986 to 2021, the glacier area in the training area experienced the largest reduction (232.03 ± 54.08 km^2) and reduction rate ($6.63 \pm 1.55\%$ a^{-1}), followed by testing area 3, testing area 1, and testing area 2, while testing area 3 had the largest proportional reduction ($37.08 \pm 8.08\%$) and reduction rate ($1.06 \pm 0.23\%$ a^{-1}), followed by testing area 2, the training area, and testing area 1. These results indicate that the reduction rate of the glacier area varies among different periods and regions. Compared with the SGCI and previous studies, the relative errors of glacier areas in this study are 0.03–4.09%, which indicates that the glacier outline dataset obtained in this study is reliable and can effectively reflect the glacier area in the THR.

In summary, in this study, a more accurate and generalizable method of glacier outline extraction was proposed and a long time-series dataset of glacier outlines in the THR from 1986 to 2021 was presented. Our study can provide data support for the estimation of regional water resources storage. The dataset is now available.[4]

ACKNOWLEDGMENT

[4][Online]. Available: https://doi.org/10.5281/zenodo.5512064

[5][Online]. Available: http://data.tpdc.ac.cn

[6][Online]. Available: https://glovis.usgs.gov/

[7][Online]. Available: https://www.letpub.com/

REFERENCES

[1] D. Petrakov *et al.*, "Accelerated glacier shrinkage in the Ak-Shyirak massif, inner Tien Shan, during 2003–2013," *Sci. Total Environ.*, vol. 562, pp. 364–378, Aug. 2016, doi: 10.1016/j.scitotenv.2016.03.162.

[2] T. Bolch *et al.*, "The state and fate of Himalayan glaciers," *Science*, vol. 336, no. 6079, pp. 310–314, Apr. 2012, doi: 10.1126/science.1215828.

[3] M. Zemp *et al.*, "Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016," *Nature*, vol. 568, no. 7752, pp. 382–386, Apr. 2019, doi: 10.1038/s41586-019-1071-0.

[4] S. Wang, C. Xu, W. Zhang, H. Chen, and B. Zhang, "Human-Induced water loss from closed inland lakes: Hydrological simulations in China's Daihai lake," *J. Hydrol.*, vol. 607, Apr. 2022, Art. no. 127552, doi: 10.1016/j.jhydrol.2022.127552.

[5] A. S. Gardner *et al.*, "A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009," *Science*, vol. 340, no. 6134, pp. 852–857, May 2013, doi: 10.1126/science.1234532.

[6] D. Farinotti *et al.*, "A consensus estimate for the ice thickness distribution of all glaciers on earth," *Nature Geosci.*, vol. 12, no. 3, pp. 168–173, Feb. 2019, doi: 10.1038/s41561-019-0300-3.

[7] H. D. Pritchard, "Asia's shrinking glaciers protect large populations from drought stress," *Nature*, vol. 569, no. 7758, pp. 649–654, May 2019, doi: 10.1038/s41586-019-1240-1.

[8] W. W. Immerzeel, L. P. Van Beek, and M. F. Bierkens, "Climate change will affect the Asian water towers," *Science*, vol. 328, no. 5984, pp. 1382–1385, Jun. 2010, doi: 10.1126/science.1183188.

[9] X. Wang *et al.*, "Mapping and inventorying active rock glaciers in the northern Tien Shan of China using satellite SAR interferometry," *Cryosphere*, vol. 11, no. 2, pp. 997–1014, Apr. 2017, doi: 10.5194/tc-11-997-2017.

[10] W. W. Immerzeel *et al.*, "High-resolution monitoring of himalayan glacier dynamics using unmanned aerial vehicles," *Remote Sens. Environ.*, vol. 150, pp. 93–103, Jul. 2014, doi: 10.1016/j.rse.2014.04.025.

[11] S. Kaushik, P. K. Joshi, and T. Singh, "Development of glacier mapping in Indian Himalaya: A review of approaches," *Int. J. Remote Sens.*, vol. 40, no. 17, pp. 6607–6634, Sep. 2019, doi: 10.1080/01431161.2019.1582114.

[12] S. Garg *et al.*, "Revisiting the 24 year (1994-2018) record of glacier mass budget in the Suru sub-basin, Western Himalaya: Overall response and controlling factors," *Sci. Total Environ.*, vol. 800, Dec. 2021, Art. no. 149533, doi: 10.1016/j.scitotenv.2021.149533.

[13] S. Liu *et al.*, "The contemporary glaciers in China based on the second Chinese glacier inventory," *Acta Geogr. Sin.*, vol. 70, no. 1, pp. 3–16, Jan. 2015, doi: 10.11821/dlxb201501001.

[14] H. Frey, F. Paul, and T. Strozzi, "Compilation of a glacier inventory for the western himalayas from satellite data: Methods, challenges, and results," *Remote Sens. Environ.*, vol. 124, pp. 832–843, Sep. 2012, doi: 10.1016/j.rse.2012.06.020.

[15] D. B. Jones, S. Harrison, K. Anderson, and W. B. Whalley, "Rock glaciers and mountain hydrology: A review," *Earth-Sci. Rev.*, vol. 193, pp. 66–90, Jun. 2019, doi: 10.1016/j.earscirev.2019.04.001.

[16] L. Ma *et al.*, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm.*, vol. 152, pp. 166–177, Jun. 2019, doi: 10.1016/j.isprsjprs.2019.04.015.

[17] P. Du *et al.*, "Advances of four machine learning methods for spatial data handling: A review," *J. Geovisualization Spatial Anal.*, vol. 4, no. 1, May 2020, Art. no. 13, doi: 10.1007/s41651-020-00048-5.

[18] C. A. Baumhoer, A. J. Dietz, C. Kneisel, and C. Kuenzer, "Automated extraction of antarctic glacier and ice shelf fronts from sentinel-1 imagery using deep learning," *Remote Sens.*, vol. 11, no. 21, Oct. 2019, Art. no. 2529, doi: 10.3390/rs11212529.

[19] Y. Mohajerani, M. Wood, I. Velicogna, and E. Rignot, "Detection of glacier calving margins with convolutional neural networks: A case study," *Remote Sens.*, vol. 11, no. 1, Jan. 2019, Art. no. 74, doi: 10.3390/rs11010074.

[20] E. Zhang, L. Liu, and L. Huang, "Automatically delineating the calving front of Jakobshavn Isbræ from multitemporal TerraSAR-X images: A deep learning approach," *Cryosphere*, vol. 13, no. 6, pp. 1729–1741, Jun. 2019, doi: 10.5194/tc-13-1729-2019.

[21] E. Zhang, L. Liu, L. Huang, and K. S. Ng, "An automated, generalized, deep-learning-based method for delineating the calving fronts of Greenland glaciers from multi-sensor remote sensing imagery," *Remote Sens. Environ,* vol. 254, Mar. 2021, Art. no. 112265, doi: 10.1016/j.rse.2020.112265.

[22] Z. Xie, U. K. Haritashya, V. K. Asari, B. W. Young, M. P. Bishop, and J. S. Kargel, "GlacierNet: A deep-learning approach for debris-covered glacier mapping," *IEEE Access*, vol. 8, pp. 83495–83510, 2020, doi: 10.1109/ACCESS.2020.2991187.

[23] B. A. Robson *et al.*, "Automated detection of rock glaciers using deep learning and object-based image analysis," *Remote Sens. Environ.*, vol. 250, Dec. 2020, Art. no. 112033, doi: 10.1016/j.rse.2020.112033.

[24] Y. He *et al.*, "An extraction method for glacial lakes based on landsat-8 imagery using an improved U-Net network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6544–6558, 2021, doi: 10.1109/JS-TARS.2021.3085397.

[25] S. Tian, Y. Dong, R. Feng, D. Liang, and L. Wang, "Mapping mountain glaciers using an improved U-Net model with cse," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 463–477, Dec. 2022, doi: 10.1080/17538947.2022.2036834.

[26] Y. Yi and W. Zhang, "A new deep-learning-based approach for earthquake-triggered landslide detection from single-temporal rapideye satellite imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 6166–6176, 2020, doi: 10.1109/JSTARS.2020.3028855.

[27] L. Chen, W. Zhang, and H. Gao, "Spatiotemporal dynamic characteristics of snow cover from 1980 to 2019 in the three-river-source region and its response to climate change," *J. Glaciol. Geocryol.*, vol. 44, no. 1, pp. 133–146, Mar. 2022.

[28] Y. Zhang, S. Liu, and X. Wang, "A dataset of spatial distribution of degree-day factors for glaciers in high mountain Asia," *China Sci. Data*, vol. 4, no. 3, pp. 141–151, Apr. 2019, doi: 10.11922/csdata.2019.0009.zh.

[29] W. Guo *et al.*, "The second Chinese glacier inventory: Data, methods and results," *J. Glaciol.*, vol. 61, no. 226, pp. 357–372, Feb. 2015, doi: 10.3189/2015JoG14J209.

[30] M. Sun, S. Liu, X. Yao, W. Guo, and J. Xu, "Glacier changes in the Qilian mountains in the past half-century: Based on the revised first and second Chinese glacier inventory," *J. Geographical Sci.*, vol. 28, no. 2, pp. 206–220, Feb. 2018, doi: 10.1007/s11442-018-1468-y.

[31] C. Teixeira Pinto, X. Jing, and L. Leigh, "Evaluation analysis of landsat level-1 and level-2 data products using in situ measurements," *Remote Sens.*, vol. 12, no. 16, Aug. 2020, Art. no. 2597, doi: 10.3390/rs12162597.

[32] J. Storey, M. Choate, and K. Lee, "Landsat 8 operational land imager on-orbit geometric calibration and performance," *Remote Sens.*, vol. 6, no. 11, pp. 11127–11152, Nov. 2014, doi: 10.3390/rs61111127.

[33] J. L. Dwyer *et al.*, "Analysis ready data: Enabling analysis of the landsat archive," *Remote Sens.*, vol. 10, no. 9, Aug. 2018, Art. no. 1363, doi: 10.3390/rs10091363.

[34] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, Mar. 2019, Art. no. 27, doi: 10.1186/s40537-019-0192-5.

[35] M. Marochov, C. R. Stokes, and P. E. Carbonneau, "Image classification of marine-terminating outlet glaciers in Greenland using deep learning methods," *Cryosphere*, vol. 15, no. 11, pp. 5041–5059, Nov. 2021, doi: 10.5194/tc-15-5041-2021.

[36] Q. He, Z. Zhang, G. Ma, and J. Wu, "Glacier identification from landsat 8 OLI imagery using deep U-Net," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 5, no. 3, pp. 381–386, Mar. 2020, doi: 10.5194/isprs-annals-V-3-2020-381-2020.

[37] X. Liu, X. Liu, Z. Wang, G. Huang, and R. Shu, "Classification of laser footprint based on random forest in mountainous area using GLA full-waveform features," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2284–2297, 2022, doi: 10.1109/JSTARS.2022.3151332.

[38] F. Paul *et al.*, "The glaciers climate change initiative: Methods for creating glacier area, elevation change and velocity products," *Remote Sens. Environ.*, vol. 162, pp. 408–426, Jun. 2015, doi: 10.1016/j.rse.2013.07.043.

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, Nov. 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[40] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested U-Net architecture for medical image segmentation," in *Proc. Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support*, Sep. 2018, pp. 3–11, doi: 10.1007/978-3-030-00889-5_1.

[41] Y. Yi *et al.*, "Semantic segmentation of urban buildings from VHR remote sensing imagery using a deep convolutional neural network," *Remote Sens.*, vol. 11, no. 15, Jul. 2019, Art. no. 1774, doi: 10.3390/rs11151774.

[42] Y. Liu *et al.*, "Semantic labeling in very high resolution images via a self-cascaded convolutional neural network," *ISPRS J. Photogramm.*, vol. 145, pp. 78–95, Nov. 2018, doi: 10.1016/j.isprsjprs.2017.12.007.

[43] R. Hussein, S. Lee, R. Ward, and M. J. McKeown, "Semi-dilated convolutional neural networks for epileptic seizure prediction," *Neural. Netw.*, vol. 139, pp. 212–222, Jul. 2021, doi: 10.1016/j.neunet.2021.03.008.

[44] J. Lee, J. Jo, and T. Park, "Segmentation of vehicles and roads by a low-channel LiDAR," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 11, pp. 4251–4256, Nov. 2019.

[45] R. Wang, M. Gong, and D. Tao, "Receptive field size versus model depth for single image super-resolution," *IEEE Trans. Image Process.*, vol. 29, pp. 1669–1682, 2020, doi: 10.1109/TIP.2019.2941327.

[46] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *J. Mach. Learn. Res.*, vol. 9, pp. 249–256, Jan. 2010.

[47] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," Dec. 2014, doi: 10.48550/arXiv.1412.6980.

[48] W. Wang, S. Dou, Z. Jiang, and L. Sun, "A fast dense spectral–spatial convolution network framework for hyperspectral images classification," *Remote Sens.*, vol. 10, no. 7, Jul. 2018, Art. no. 1068, doi: 10.3390/rs10071068.

[49] L. Prechelt, "Early stopping - but when?," in *Neural Networks: Tricks of the Trade*. Berlin, Germany: Springer, 1998, pp. 55–69.

[50] M. Volpi and D. Tuia, "Dense semantic labeling of sub-decimeter resolution images with convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 881–893, Feb. 2017, doi: 10.1109/TGRS.2016.2616585.

[51] A. Chowdhury, M. C. Sharma, S. K. De, and M. Debnath, "Glacier changes in the Chhombo Chhu watershed of the Tista basin between 1975 and 2018, the Sikkim Himalaya, India," *Earth Syst. Sci. Data*, vol. 13, no. 6, pp. 2923–2944, Jun. 2021, doi: 10.5194/essd-13-2923-2021.

[52] T. Bolch *et al.*, "A glacier inventory for the western Nyainqentanglha range and the Nam co basin, Tibet, and glacier changes 1976–2009," *Cryosphere*, vol. 4, no. 3, pp. 419–433, Sep. 2010, doi: 10.5194/tc-4-419-2010.

[53] F. D. Granshaw and A. G. Fountain, "Glacier change (1958–1998) in the North Cascades National Park Complex, Washington, USA," *J. Glaciol.*, vol. 52, no. 177, pp. 251–256, Jan. 2006, doi: 10.3189/172756506781828782.

[54] D. K. Hall, K. J. Bayr, W. Schöner, R. A. Bindschadler, and J. Y. L. Chien, "Consideration of the errors inherent in mapping historical glacier positions in Austria from the ground and space (1893–2001)," *Remote Sens. Environ.*, vol. 86, no. 4, pp. 566–577, Aug. 2003, doi: 10.1016/S0034-4257(03)00134-2.

[55] P. Rastner, T. Strozzi, and F. Paul, "Fusion of multi-source satellite data and dems to create a new glacier inventory for Novaya Zemlya," *Remote Sens.*, vol. 9, no. 11, Nov. 2017, Art. no. 1122, doi: 10.3390/rs9111122.

[56] D. D. Yavaşlı, C. J. Tucker, and K. A. Melocik, "Change in the glacier extent in Turkey during the landsat era," *Remote Sens. Environ.*, vol. 163, pp. 32–41, Jun. 2015, doi: 10.1016/j.rse.2015.03.002.

[57] C. Narama, A. Kääb, M. Duishonakunov, and K. Abdrakhmatov, "Spatial variability of recent glacier area changes in the Tien Shan Mountains, Central Asia, using Corona (∼1970), Landsat (∼2000), and Alos (∼2007) satellite data," *Glob. Planet. Change*, vol. 71, no. 1, pp. 42–54, Mar. 2010, doi: 10.1016/j.gloplacha.2009.08.002.

[58] N. Y. Q. Abderrahim, S. Abderrahim, and A. Rida, "Road segmentation using U-Net architecture," in *Proc. IEEE Int. Conf. Moroccan Geomatics*, 2020, pp. 1–4, doi: 10.1109/Morgeo49220.2020.9121887.

[59] J. H. Kim *et al.*, "Objects segmentation from high-resolution aerial images using U-Net with pyramid pooling layers," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 1, pp. 115–119, Jan. 2019.

[60] M. Rahnemoonfar, M. Yari, J. Paden, L. Koenig, and O. Ibikunle, "Deep multi-scale learning for automatic tracking of internal layers of ice in radar data," *J. Glaciol.*, vol. 67, no. 26, pp. 39–48, Feb. 2021, doi: 10.1017/jog.2020.80.

[61] S. Srivastava, J. E. Vargas-Muñoz, and D. Tuia, "Understanding urban landuse from the above and ground perspectives: A deep learning, multimodal solution," *Remote Sens. Environ.*, vol. 228, pp. 129–143, Jul. 2019, doi: 10.1016/j.rse.2019.04.014.

[62] M. Kneib *et al.*, "Mapping ice cliffs on debris-covered glaciers using multispectral satellite images," *Remote Sens. Environ.*, vol. 253, Feb. 2021, Art. no. 112201, doi: 10.1016/j.rse.2020.112201.

[63] K. Wang, T. Yang, Y. He, and Q. JI, "Research for glaciers and climate change of Anyemaqen mountain nearly 30 years," *Res. Soil Water Conservation*, vol. 22, no. 3, pp. 300–303, Jun. 2015, doi: 10.13869/j.cnki.rswc.2015.03.054.

[64] S. Jiang, T. Yang, and H. Tian, "Glacier shrinkage and its dependence on climate in the Malan mountain in past 40 years based on RS and GIS," *J. Glaciol. Geocryol.*, vol. 34, no. 3, pp. 522–528, Jun. 2012.

[65] Q. Ye, S. Kang, F. Chen, and J. Wang, "Monitoring glacier variations on Geladandong mountain, Central Tibetan Plateau, from 1969 to 2002 using remote-sensing and GIS technologies," *J. Glaciol.*, vol. 52, no. 179, pp. 537–545, Oct. 2006, doi: 10.3189/172756506781828359.

[66] Z. Zhang *et al.*, "Glacier variations at Xinqingfeng and Malan ice caps in the Inner Tibetan Plateau since 1970," *Remote Sens.*, vol. 12, no. 3, Jan. 2020, Art. no. 421, doi: 10.3390/rs12030421.

[67] Q. Ye *et al.*, "Glacier changes on the Tibetan plateau derived from Landsat imagery: Mid-1970s–2000–13," *J. Glaciol.*, vol. 63, no. 238, pp. 273–287, Jan. 2017, doi: 10.1017/jog.2016.137.

[68] H.-Y. Duan *et al.*, "Glacier change in the Tanggula mountains, Tibetan plateau, in 1969–2015," *J. Mountain. Sci.*, vol. 16, no. 11, pp. 2663–2678, Jul. 2019, doi: 10.1007/s11629-018-5011-5.

[69] A. A. Chen *et al.*, "Region-wide glacier mass budgets for the tanggula mountains between ∼1969 and ∼2015 derived from remote sensing data," *Arctic Antarctic Alpine Res.*, vol. 49, no. 4, pp. 551–568, Nov. 2017, doi: 10.1657/AAAR0016-065.

[70] S. Jin, Y. Zhang, and H. Wu, "Study on glacial advancement and retreatment in Geladandong region of Changjiangyuan in recent 40 years," *J. Nat. Resour.*, vol. 28, no. 12, pp. 2095–2104, Aug. 2013.

Longfei Chen received the B.S. degree in physical geography sciences from Shanxi University, Taiyuan, China, in 2019. Since July 2019, he has been working toward the master's degree in the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China.

His research interests include, remote sensing image understanding, glaciers, snow cover, and hydrology modeling.

Wanchang Zhang (Member, IEEE) received the M.S. degree in geography from the Cold and Arid Regions Environmental and Engineering Research Institute (CAREERI), Chinese Academy of Sciences (CAS), Lanzhou, China, in 1992, the second M.S. degree in hydrospheric-atmospheric science from the IHAS Special Program sponsored by UNISCO Nagoya University, Nagoya, Japan, in 1996, and the Ph.D. degree in remote sensing and GIS from the Department of the Earth and Planetary Sciences, Graduate School of Science, Nagoya University, Nagoya, Japan, in 2000.

Since 2000, he returned to China and was employed as a Full Professor/a Ph.D. Supervisor with the International Institute of Earth System Science, Nanjing University, Nanjing, China, and was promoted as a Research Scientist of "100 Talent Researcher Program" sponsored by CAS, in 2005, engaged in the research field of Global Changes with the Institute for Atmospheric Physics, CAS. He works at the Aerospace Information Research Institute, CAS. In total, he has authored or coauthored more than 340 peer-reviewed domestic and international journal articles, in which about 180 articles published in Science Citation Index (SCI), or Engineering Index (EI) indexed journals as first or corresponding author. His research interests include but not limited in application of remote sensing (RS) and GIS in earth system sciences; integration of RS and GIS to construct the distributed ecohydro-climatic models for hydrology, water resource, and water environment studies; and global disaster and environmental studies with RS and GIS techniques associated with modeling schemes.

Dr. Zhang was invited as the Co-Chair for many international conferences as a Committee Member of IEEE from 2005, in recent years, and served as an Editorial Board member and a Reviewer in many domestic and international journals. He has also applied eight domestic invention potent items and owned ten sets of software registered as the first inventor in the Chinese National Copyright Administration. Three website platforms have been constructed under his supervision and were online accessible to public for various applications.

Yaning Yi received the B.S. degree in surveying and mapping engineering from Central South University, Changsha, China, in 2016, and the Ph.D. degree from the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2021.

Since July 2021, he has been a Postdoc with the National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing. His research interests include landslide detection, landslide susceptibility mapping, deep learning, and remote sensing applications.

Zhijie Zhang received the B.S. degree in electrical engineering from the Nanjing University of Post and Telecommunications, Nanjing, China, in 2014, and the M.S. degree in computer engineering from Rutgers University, New Brunswick, NJ, USA, in 2018. He is working toward the Ph.D. degree in geography with the University of Connecticut, Storrs, CT, USA.

From 2016 to 2017, he was a Research Student in WINLAB, Rutgers, NJ. He is involved in two Internet of things (sensor network projects. Since 2018, he has been a Research Assistant with the Geography Department, University of Connecticut. His research interest includes applying deep learning methods in remote sensing, hydrology studies, and landslide studies.

Shijun Chao received the B.S. degree in soil and water conservation and desertification combating from the Northwest A&F University, Xianyang, China, in 2005, and the M.S. degree in soil and water conservation and desertification combating from Shaanxi Normal University, Xi'an, China, in 2008.

Since October 2019, he has been the deputy dean in Qinghai Research and Design Institute of Environmental Science. His research interest includes remote sensing, hydrology, and glaciers.