A Conditional Random Fields-Based Identification for Small Lakes Over Multiple Melt Seasons Using Sentinel-2 Imagery in the Larsemann Hills, East Antarctica

Tingting Zhu, Xiangbin Cui, Yu Zhang, Kai Lu, and Yuande Yang

***Abstract*—This study provides the first long-time series of spatial and temporal distributions for small lakes in the Larsemann Hills (69°23′S, 76°20′E) in the East Antarctic. In the Larsemann oasis, there is a significant number of over 150 small lakes, which can be observed with high spatial resolution in remote sensing imagery. However, accurately identifying and analyzing these small water bodies and elongated rivers has been challenging due to the mixed pixels effect and limitations in available middle spatial resolution imagery. In our study, we propose a data-driven approach within the conditional random fields framework, which considers three scales: superpixel, pixel, and subpixel, to refine the boundaries of small water bodies efficiently. The superpixel level quickly identifies the main water body and normalized difference water index provides a buffer region, while the pixel level employs support vector machine (SVM) to obtain a more precise boundary. Subpixel mapping technology within the pixel level further reduces mixed pixel effects for improved accuracy. The waterbodies were extracted from Sentinel-2 images with a spatial resolution of 10 m. The lake boundaries derived from the proposed algorithm in this study showed good agreement with in situ measurements of the lake shoreline delineated by aerial images from the 39th Chinese Antarctic Scientific Expedition. The analysis revealed distinct seasonal patterns across the Larsemann Hills, while the lake areas achieved their peak extents earlier, specifically in February before 2020 and in January after 2020. The water body mapping based on the proposed algorithm can contribute to Antarctic remote sensing hydrological observations, particularly in the monitoring of outburst events. These findings demonstrate the potential of extending this method to other Antarctic oases to enhance intra-annual lake observations. Moreover, Sentinel-2 images provide valuable remote sensing data for studying the seasonal cycles of water bodies, including those of varying sizes in the Larsemann Hills, based on long-term time series imagery.**

***Index Terms*—Conditional random fields (CRF), small lakes, spatial and temporal analysis, superpixel-pixel-subpixel.**

I. Introduction

THE Antarctic water bodies primarily consist of supraglacial and subglacial lakes, playing a crucial role in ice sheet hydrology and the stability of ice sheets [2], [5], [11]. Most supraglacial lakes tend to be quite small, particularly when situated within glacier fractures and ice shelf crevasses, potentially serving as significant precursors to ice shelf disintegration. Given the inherent tradeoff between spatial and temporal resolution in remote sensing, developing a subpixel mapping approach becomes essential to accurately assess the interannual variability of lake distribution. In the Larsemann Hills oasis, there are 150 lakes distinguished by their compact sizes, with a maximum area of 0.13 km^2. Monitoring the spatial and temporal distribution of lakes in the Larsemann Hills is important for understanding the dynamics of the local hydrological system and its response to climate change, scientists gain insights into the water cycle in the region, its impact on the local ecosystem, and the potential implications for broader Antarctic ice sheet dynamics. The Larsemann Hills oasis is prone to regular and unpredictable occurrences of lake drainage, including outbursts, which have been observed annually. Of notable concern is the Boulder outburst flood that took place on 30 January, 2017, located in the vicinity of the Russian expedition station [3]. These instabilities in the oasis demonstrate the dynamic nature of the region and its susceptibility to abrupt changes in lake behavior that holds immense importance for monitoring the temporal and spatial variations of water bodies within the Larsemann Hills oasis even in the wider Antarctic region.

Long-time series monitoring using remote sensing technology is significant in understanding the lake dynamic changes of lakes, offering valuable insights into local climate variability and

Manuscript received 4 December 2023; revised 23 February 2024 and 18 March 2024; accepted 16 April 2024. Date of publication 22 April 2024; date of current version 8 May 2024. This work was supported in part by the "China High Resolution Earth Observation System Program" under Grant GFZX04032502, in part by the National Natural Science Foundation of China under Grant 42376253, Grant 41901275, and Grant 41801266, in part by the Wuhan Knowledge Innovation Special Project under Grant 2023010201020267, in part by Key Laboratory of Polar Environment Monitoring and Public Governance Wuhan University under Grant 202404, and in part by the Open Fund of Hubei Luojia Laboratory under Grant 230100011. *(Corresponding authors: Xiangbin Cui; Yu Zhang.)*

Tingting Zhu is with the College of Geomatics Science and Technology, Nanjing Tech University, Nanjing 210037, China (e-mail: zhutingting62008@163.com).

Xiangbin Cui and Kai Lu are with the Polar Research Institute of China, Shanghai 200136, China (e-mail: cuixiangbin@pric.org.cn; lukai@pric.org.cn).

Yu Zhang and Yuande Yang are with the Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China, and also with the Key Laboratory of Polar Environment Monitoring and Public Governance, Ministry of Education and Luojia Laboratory, Wuhan 430079, China (e-mail: yuzhang_spl@whu.edu.cn; yuandeyang@whu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2024.3391881

Antarctic ice sheet dynamics [30]. Satellite remote sensing provides continuous coverage, allowing for day and night monitoring of lakes with high spatial–temporal–spectral resolution. The remote and harsh environment of Antarctica presents significant challenges in establishing hydrological stations for widespread lake monitoring. Only a limited number of large lakes near the Indian, China Zhongshan Station, and Russian stations have hydrological stations dedicated to lake monitoring. This limited coverage poses difficulties in meeting the continuous and extensive monitoring requirements in the region. Satellite remote sensing imagery can provide extensive data support enabling continuous monitoring of temporal and spatial changes in lakes. Optical remote sensing sensors are classified into three categories based on their spatial resolution: coarse resolution (>200 m), medium resolution (5–200 m), and high resolution (<5 m). High-resolution optical remote sensing images have distinct advantages in observing and classifying land-use categories, particularly in capturing detailed information and accurately distinguishing small-sized land cover classifications, such as small water bodies [13], [28]. In polar regions, the availability of remote sensing imagery is mainly concentrated at medium resolution. Enhancing the spatial resolution of optical satellites has always been one of the most critical tasks in the field of polar remote sensing. The subpixel unmixing technique is important in achieving improved spatial resolution by extracting subpixel information to enhance the classification of land cover [12], [23]. Specifically, the medium-resolution Sentinel-2 data, which is freely accessible and offers a resolution of 10 m, is well-suited for the monitoring of small-scale lakes within the Larsemann Hills. In addition to deep learning-based approaches for water body extraction, conventional methods can be categorized into three types. The first method is based on water body indices, wherein discriminative water body indices are formulated by relying on the spectral responses in different bands, and efforts are made to effectively combine them for water body extraction [26]. The second method involves exploring the spatial features of water bodies on imagery, such as edges and shapes, and performing pixel clustering to separate water from nonwater regions [16], [20]. The third method combines the first and second approaches [31], wherein a water body extraction method considering both spectral and spatial features is constructed within the framework of conditional random fields, enabling effective extraction of water bodies.

In the context of accurately detecting and delineating small lakes within the Larsemann Hills, the utilization of high-resolution imagery, such as WorldView-2 data, has been employed by Jawak et al. [34] to generate a series of normalized difference water index (NDWI) features. The selection of different water indices for various types of water bodies determines the accuracy of identification and analysis. Nevertheless, there still exists a notable disparity in achieving precise extraction of small-sized lakes when employing imagery with lower spatial resolutions. The inaccuracies in water body extraction results at lower spatial resolution primarily are attributed to the presence of mixed pixels containing rock or shadow elements, which commonly occur at the water body boundaries. Subpixel mapping techniques can be employed to provide a spatial distribution of each class within these mixed pixels [21], [22]. To ensure the efficiency of the unmixing and subpixel mapping algorithm, the conditional random fields (CRF) can be employed to describe the interior, exterior, and boundary components of the water body at pixel, superpixel, and subpixel, respectively. Subpixel mapping aims to obtain more detailed information by dividing low-resolution pixels into higher resolution subpixels. This process proves to be particularly effective in identifying boundaries, especially for small lakes in middle-resolution remote sensing imagery [24], [27]. By performing subpixel mapping under the constraints of least squares and non-negativity, each coarse pixel can be divided into subpixels, resulting in more precise boundary information. It is worth mentioning that the CRF framework takes spatial relations into account when modeling subpixel labels. By incorporating spatial adaptive attraction values and a spatially smoothed prior, the CRF framework ensures that the subpixel labels among neighboring pixels are accurately represented, while also considering the abundance constraints [4]. At the superpixel scale, a segmentation technique is utilized to group similar pixels into superpixels, which greatly enhances the efficiency and time-saving aspect of the process. When it comes to superpixels, such as those used in the CRF framework, they are constructed based on pairwise potential and unary terms [33]. Unary terms provided by support vector machine (SVM) do not take into account contextual information, but we can enhance the classification results by incorporating multiple textural information, such as gray level co-occurrence matrix (GLCM).

Continuous observations of ice-free lakes during the austral summer in the oasis are primarily based on field experiments involving meteorological and hydrological measurements. However, a comprehensive investigation of the spatial and temporal changes in small lakes, which is crucial for understanding the seasonal water cycle and monitoring lake outburst susceptibility, is lacking. This study aims to address this gap by providing long-term spatial and temporal distributions of lake areas in Larsemann Hills. Our focus is on developing algorithms for detecting small lake types using mid-high spatial resolution remote sensing technology and multiple machine learning algorithms.

The rest of this article is organized as follows. Section II presents the study area and data used for detecting small lakes in middle-spatial resolution remote sensing images. Section III provides details on the proposed algorithms, while Section IV showcases the major experimental results and an assessment of these experiments. A comparison of different machine learning algorithms and strategies is presented. Finally, Section V concludes this article.

II. Study Area and Materials

A. Study Area, Climate, Lakes

The Larsemann Hills (69°23′S, 76°20′E) are located on the shoreline of Princess Elizabeth Land, in the southeastern region of Prydz Bay, East Antarctica. Situated between the Amery Ice Shelf and Dålk Ice Shelf, the Larsemann Hills predominantly consists of outcrop rocks, making it a coastal ice-free oasis. Fig. 1(a) displays the map provided by the Sentinel-2

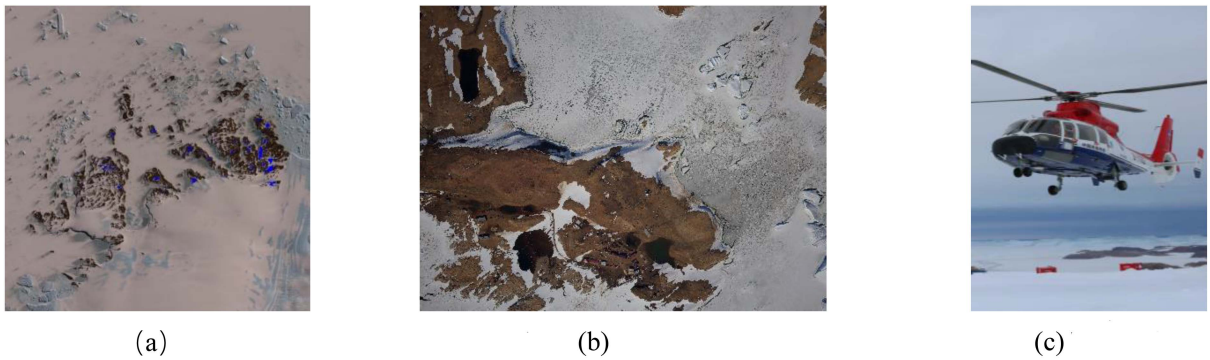


Fig. 1. Location of the Larsemann Hills. (a) Major inland water bodies in blue. Water map generated from Sentinel-2 and layers derived through the proposed algorithm. (b) The accuracy of the lake areas was verified through in situ measurements taken by the 39th Chinese Antarctic Scientific Expedition and (c) aerial images were collected from the Dolphin helicopter which is equipped with an ice survey camera payload.

optical imagery depicting the Larsemann Hills, which exhibit three primary land cover classes: water bodies, glacier ice, and exposed rocks [14]. Lakes can be found in the vicinity of rock islands which can be observed from the aerial images in Fig. 1(b) collected from the 39th Chinese Antarctic Scientific Expedition, see Fig. 1(c). The weather of the Larsemann Hills is under the influence of the Katabatic winds that blow from the northeast continuously throughout the year. It rarely rains in the Antarctic ice-free zone (Larsemann Hills also known as an oasis). The region's average annual precipitation (in the form of snow) is only 100–150 mm of water equivalent. According to the reference elevation model for Antarctica, the terrain around the Larsemann Hills consists of hills ranging in elevation from 30 to 120 m [8]. The average air temperature in summer is about 3°C, whereas the average air temperature in winter is around –16°C. Consequently, the Larsemann Hills have been selected as expedition sites by China, Russia, and India. With over 150 lakes, varying in water temperature from 0 to 10°C and water depth ranging from 5 to 50 m, these lakes cover an area ranging from 0.001 to 0.13 km^2. Most of these lakes are landlocked and occupy depressions around the rock and ice-free areas. They can be classified as supraglacial lakes, epiglacial lakes, and englacial lakes. During the austral summer, these lakes may be free of ice. Supraglacial lakes are typically located at the edges of the continental ice sheet, often in blue ice regions like Boulder Lake, which is characterized by a thick ice cover. Epiglacial lakes occur at the boundary between rocks and the continental ice sheet, such as Lake Ledyanoe. Notably, on 30 January, 2017, the drainage of Dålk lakes occurred due to outburst floods. This study primarily focuses on two representative lake types in the Larsemann Hills: landlocked lakes and supraglacial lakes.

B. Sentinel-2 Data

The Sentinel satellite system is a key component of the Copernicus Program, which was announced by the European Space Agency. The Sentinel-2A (S2A) platform was launched on 23 June, 2015, and operates in a near-polar orbit with a repeat cycle of 10 days. Since the launch of the Sentinel-2B (S2B) satellite on 7 March, 2017, the Sentinel-2 constellation system has been able to provide high-quality optical images with a repeat time of up to 5 days. With 13 spectral channels covering the visible/near-infrared and short-wave infrared (SWIR) spectral ranges, the S2A and S2B optical instruments are designed to support the Global Monitoring for Environment and Security service, as well as land and ocean environment observation. This article utilizes a total of 67 Sentinel-2 imagery scenes, including both L1C and L2A data, spanning from September 2016 to March 2023, for lake extraction in the Larsemann Hills of Antarctica. The L1C data provided include level-1 top-of-atmosphere reflectances in UTM projection, with tiles measuring 100 by 100 km. The aforementioned data have been subject to both radiometric and geometric corrections. The L2A data, on the other hand, consist of surface reflectance products. To produce the monthly lake area product, we have selected 49 scenes from each year and month, spanning from September to March. The lake extraction experiments utilized Sentinel-2 data with a spatial resolution of 10 m, including the red, green, blue, and near-infrared (NIR) bands. Detailed information regarding the Sentinel-2 data can be found in Table I. Sentinel-2 Level-1A and C data covering the Larsemann Hills are collected and then, these Sentinel-2 images are preprocessed to eliminate the invalid pixels covered by the cloud.

C. In Situ Data

We use the aerial survey dataset for evaluating the results of the lake areas and boundaries. For the validation, the in situ measurements were collected during the 39th Chinese Antarctic Scientific Expedition. An aerial survey experiment and application demonstration of sea ice conditions in key Antarctic regions was conducted using the polar research helicopter and the developed high-resolution airborne polar remote sensing software and hardware system. The study focused on the Zhongshan Station in Antarctica and its surrounding sea ice areas. The aim was to provide data and technological support for the assessment of resource potential in polar regions and the planning of Antarctic expedition icebreaker routes. The aerial observation missions for two survey lines were conducted on the morning and evening of 8th January during the 39th Chinese Antarctic Scientific

TABLE I
SENTINEL-2 SATELLITE DATA

Band name	Central wavelength/nm		Bandwidth/nm		Spatial resolution/m
	S2A	S2B	S2A	S2B	
1-Coastal aerosol	443.9	442.3	27	45	60
2-Blue	496.6	492.1	98	98	10
3-Green	560.0	559.0	45	46	10
4-Red	664.5	665.0	38	39	10
5-Vegetation Red Edge	703.9	703.8	19	20	20
6-Vegetation Red Edge	740.2	739.1	18	18	20
7-Vegetation Red Edge	782.5	779.7	28	28	20
8-NIR	835.1	833.0	145	133	10
8A-Vegetation Red Edge	864.8	864.0	33	32	20
9-Water Vapor	945.0	943.2	26	27	60
10-SWIR-Cirrus	1373.5	1376.9	75	76	60
11-SWIR	1613.7	1610.4	143	141	20
12-SWIR	2202.4	2185.7	242	238	20

Expedition. The observation areas included the offshore sea ice area of Zhongshan Station and the Larsemann Hills region. The maximum straight-line distance of each survey line was 50 km, with a total flight distance of 300 km. The helicopter maintained a constant altitude of 2 km and a speed of 150 kmph throughout the flight, resulting in a flight time of over 2 h for each survey line. A total of more than 1600 high-resolution photographs were successfully captured during the entire survey, meeting the predetermined quality standards and fulfilling the planned objectives.

III. METHODOLOGY

The proposed algorithm can extract water boundaries at three different scales. Algorithms for extracting water bodies at pixel, superpixel, and subpixel scales are introduced in the following, respectively.

A. Pixel-Level Lake Extraction By Normalized Difference Water Index (NDWI)

To infer water pixels of Larsemann Hills, the multispectral water indices are utilized to automatically distinguish water pixels from nonwater pixels based on reflectance differences in the Green ($\rho_{\text{Green}} = 0.56\ \mu\text{m}$) and NIR spectra ($\rho_{\text{NIR}} = 0.842\ \mu\text{m}$), respectively. The NDWI water index proposed by McFeeters [35] is based on the strong absorption of water in NIR and reflectance peaks in the Green band. It can be expressed as follows:

$$\text{NDWI} = \frac{\rho_{\text{Green}} - \rho_{\text{NIR}}}{\rho_{\text{Green}} + \rho_{\text{NIR}}}. \tag{1}$$

In Sentinel-2 imagery, the green spectra refer to band 3, and NIR is band 8. In large-scale water mapping, spectral indices are significant such as NDWI, MNDWI, and AWEI [6]. Then, NDWI is used to extract water features and then Otsu's global threshold [15] is used to divide NDWI calculation results into water pixels and nonwater pixels. Otsu's method is a nonparametric and unsupervised image segmentation algorithm based on the threshold to segment NDWI into water and nonwater pixels. The threshold is selected from the maximum interclass variance globally and calculated from the corresponding histogram of the NDWI of the image. Otsu's method is the simplest and adaptive algorithm for selecting thresholds since it is independent of image brightness and contrast. Nevertheless, small-scale water mapping faces some challenges. For example, the accurate water boundary can be omitted since the muddy and shallow water bodies widely can be observed in the in situ images. The snow, the muddy and shallow water bodies, and the shadows of the outcrops can contribute to the noise and thus impede the water extraction. To decrease the above-mentioned noise artifact and improve the accuracy of the boundary extraction, the subpixel mapping based on CRF for lake extraction has been proposed and described in Section III-B.

B. Superpixel-Level Lake Boundary From Conditional Random Fields CRF

CRF model includes unary potential and pairwise potential. The unary potential individually deals with the pixel-level spectral information. The pairwise potential models the contextual information of neighborhood pixels. CRF models considering the superpixel strategy have successfully integrated multiscale contextual information into remote sensing identification. As a discriminative probability model, CRF directly builds the posterior distribution of the labels X conditioned on the observations Y. The pairwise potential is constructed on superpixels, which can be obtained by the mean-shift segmentation method. The unary and pairwise potentials in the CRF model are described as follows:

$$P(x|y) = \frac{1}{Z_{\text{CRF}}} \exp\left[\sum_{i\in S} f_i^{\text{unary}}(x_i) + \sum_{i\in S}\sum_{j\in N_i} f_{ij}^{\text{pairwise}}(x_i, x_j|y)\right]. \tag{2}$$

The unary potential f_i^{unary} contributes to the probability of a single superpixel taking the label of water or nonwater. It

can be constructed using the SVM classifier, which takes the multispectral bands (bands 3 and 8 in Sentinel-2) and NDWI features as the input. In this study, 24-dimensions data have been used to construct the unary potential. The pairwise potential f_{ij}^{pairwise} models the contextual information by considering the interactions of the neighborhood superpixels. The pairwise potential is achieved by the graph cut method, which divides two nodes into water and nonwater. In the graph cut method, the edge is defined as the neighbor superpixels. The relationships between the nodes (central superpixel) and neighbor superpixels are calculated. Under the minimum criteria of least energy loss object term, each superpixel takes the label of water or nonwater. Since the melting ice of the lakes may lead to some misclassification, we design a further step to remove these types of errors. It can also improve the accuracy along the lake boundary. The spatial relation between a superpixel and its neighboring superpixels is taken into account for comparing the superpixel and pixel level information that will be provided in a detailed description as the following. Within the water body, some pixels are classified as nonwater labels. If the comparison results indicate inconsistency, we utilized the contextual information from the pairwise potential to update the wrong superpixels, which get the new label as most of the neighborhood superpixels. Especially during the freezing period, some lakes are buried by snow, some misclassified superpixels can be corrected by this strategy.

C. Subpixel-Level Lake Boundary Based Subpixel Mapping

Since lakes in Larsemann Hills are very small lakes ($<$0.13 km^2), pixel and subpixel level lake identification are collaborative on Sentinel-2 imagery with a spatial resolution of 10 m. Pixel-level classification provides both the criteria for the pairwise potential and refinement of the unary potential which is constructed on the superpixels. Subpixel mapping can provide the spatial distribution of water within each mixed boundary pixel, leading to a more accurate lake boundary. Along the boundary, the mixed pixels usually degrade the water area calculation which can be resolved by using the spectral unmixing technique. Since the lakes in the Larsemann Hills are usually small ($<$0.13 km^2), the CRF algorithm is insufficient to describe the characteristics of small lakes ($<$0.01 km^2), which often leads to missed detection. On the other hand, the highest spatial resolution of Sentinel-2 is 10 m, and the lake boundary area often exists in the form of mixed pixels.

1) Features for the SPM Based on CRFs: To improve the ability to identify small lakes and improve the accuracy of lake boundary extraction, this article proposes a subpixel mapping method based on the CRF model, which can provide the probability belonging to the lake pixel in the unary potential. In this study, the water relative features are calculated, including the pixel-level feature NDWI, and the subpixel-level feature, including the abundance feature map, spatial attraction value, and the adaptive attraction value for water. In addition, in this study, GLCM features including mean, variance, homogeneity, second moment, and entropy have been used for water mapping. The features constructed for the unary potential in the CRF model can be expressed as the following:

$$DN = \sum_{i=1}^{k=1} f_t^k T_k, \text{ where } f_k > 0 \ \& \ \sum_{k=1}^{K} f_k = 1 \tag{3}$$

$$\text{at}_{i,j}^k = \sum_{t=1}^{N} \frac{f_t^k}{D_{t\ i,j}}, \text{ where } D_{t\ i,j} = \sqrt{d^2 + (0.5d + 0.5)^2} \tag{4}$$

$$\text{Homogeneity} = \frac{\sum_{i=1}^{k=1} \sum_{j=1}^{K} (i - \mu_x)(j - \mu_y) S_d(i,j)}{\sigma_x \sigma_y} \tag{5}$$

$$\text{Entropy} = \sum_{i=1}^{k=1} \sum_{j=1}^{K} S_t(i,j) \log_{10} S_t(i,j) \tag{6}$$

$$\text{Angular Second Moment (ASM)} = \sum_{i=1}^{K} \sum_{j=1}^{K} S_{d,\theta}(i,j)^2. \tag{7}$$

Equation (3) is the abundance calculation method using fully constrained least squares unmixing techniques, where f_t^k is the abundance under the constraint that each abundance should be non-negative and sum-to-one [9].

In (4), $\text{at}_{i,j}^k$ is the spatial attraction value at the subpixel position i, j for the category k, N is the number of subpixels within each single pixel, and $D_{t\ i,j}$ represents the Euclidean distance between the subpixel located at coordinates i, j and the neighboring pixel at position t. The parameter d denotes the scale of the subpixel, and in this study, we have set d equal to 5.

In (5)–(7), μ_x, μ_y represents the mean value and σ_x, σ_y is the corresponding standard deviation of different cells. S_t is a descriptor that computes the similarity of different pixels over each cell. For GLCM features, the window size is 4, the separated distance t is 4, the moving steps are set to 4, and the gray level is 64.

2) Lake Boundary Mapping Based on SPM With CRFs: Because the lake boundary area is usually composed of mixed pixels, to improve the positioning accuracy of the lake boundary, this study employs a subpixel mapping (SPM) method based on CRFs, taking into account the influences of spatial position and observation sequence. In addition, adaptive attraction features are introduced to enhance the accuracy of lake boundaries.

In the extraction of adaptive attraction features, we initially leverage the spectral reflectance information from the original Sentinel-2 imagery to obtain pixel-level abundance information. Building upon this, a subpixel-level attraction feature is constructed through a resampling strategy, where the spatial attraction feature calculation formula is depicted in (4). The scaling factor is set to 5, indicating that each pixel in the original image corresponds to a $5 * 5$ region in the subpixel-level image. The spatial attraction primarily determines the category of each subpixel by calculating the attraction value between neighboring subpixels and within the central subpixel, which does not require any prior information. Subpixel mapping essentially involves mapping pixel-level features to the subpixel

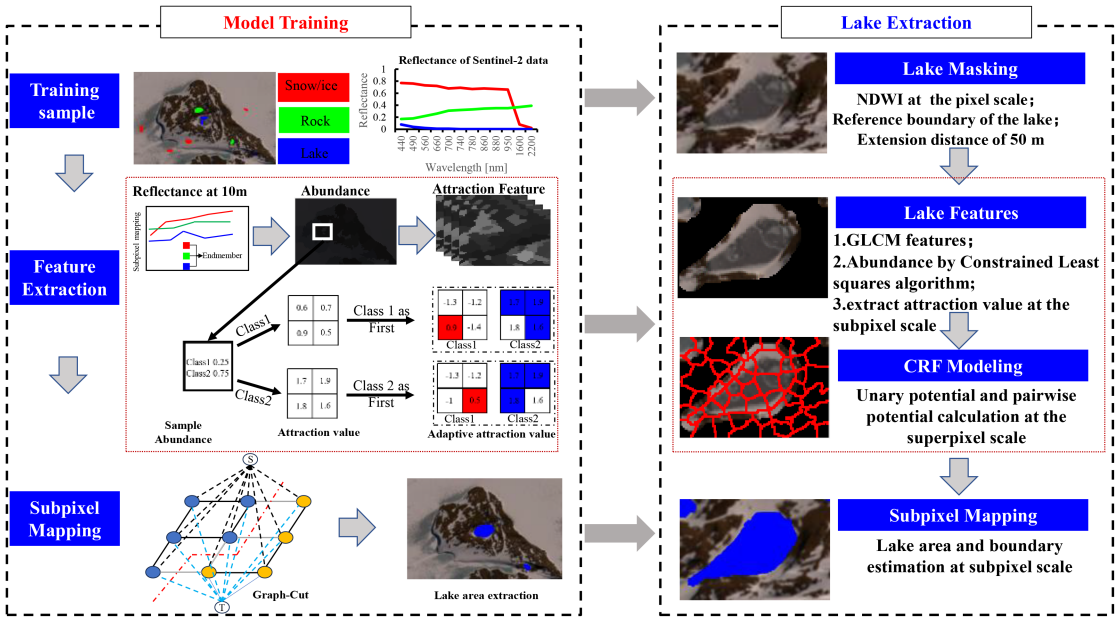


Fig. 2. Flowchart of the process of extracting small lake areas from Sentinel-2 data.

space. The accurate mapping of these features determines the assignment of subpixel categories. In this study, we introduced three endmembers, namely lakes, rocks, and snow/ice, and their allocation in the subpixel space is determined based on the spatial attraction values. The spatial attraction value considers the spatial and contextual information between pixels, while the observation order takes into account the relationships within subpixels. Without prior information, assuming lakes are known observed objects, the attraction value of lakes to subpixels of the lake category should be greater than that for subpixels of other categories. By iterating through the defined categories using this approach, we obtain three-dimensional features for each endmember. Combining these with the 24-dimensional features, such as the GLCM features constructed in Section I, results in a total of 33 dimensions of features. The primary focus of this study is on lake extraction, and thus the Moran index is employed to determine the observation order for water bodies. All these features are used to construct the CRF models.

In CRFs modeling, the discriminant model can be described in the following corresponding energy function:

$$E\ (\mathrm{L}) = -\log P\ (x|y) = \log Z_{\mathrm{CRF}} - \sum_{i\in s} f_i^{\mathrm{unary}}\ (x_i) - \sum_{i\in S}\sum_{j\in N_i} f_{ij}^{\mathrm{pairwise}}\ (x_i, x_j|y)\,. \tag{8}$$

Minimizing energy is equivalent to maximizing the posterior probability of the label field in the CRF framework. Therefore, based on the maximum posterior probability criterion, we can derive the new expressions for unary and pairwise potentials

$$f_i^{\mathrm{unary}}\ (x_i) = \sum_{i\in s} \omega_i \delta\ (y_i, l)\ \varphi_i\ (x_i)\,. \tag{9}$$

The unary is constructed by the normalized features $\varphi_i(x_i)$, we have described in Section III-A. ω_i is the weight coefficient of the features, and $\delta(y_i, l)$ is the Dirac function indicating a value of 1 when the current superpixel label y_i for water otherwise with 0. In the subpixel scale, similar pixels are grouped into superpixels, which can be obtained by the mean-shift algorithm

$$f_{ij}^{\mathrm{pairwise}}\ (x_i, x_j|y) = \sum_{i\in S}\sum_{j\in N_i} \varphi_{i,j}\ (x_i, x_j|y)\,. \tag{10}$$

The pairwise potential corresponds to the correlation between the superpixel S and neighbor superpixels N_i. When the labels of both are the same, the value is 1 divided by their Euclidean distance, whereas when the labels are different, the value is 0.

3) Overall Workflow: The flowchart of the proposed Tri-SF algorithm and the validation of the obtained lake boundary is illustrated in Fig. 2.

The algorithm proposed in this article comprises two main components: model training and lake boundary extraction. The model training involves three steps. The first step is the selection of training samples, where the NDWI is employed to compute potential water regions. A lake buffer zone is established, within which endmembers (pure pixels) are chosen as training samples (1000 pixels per class). Then, feature extraction, including abundance, spatial attraction value, and GLCM statistics (mean, variance, homogeneity, second moment, and entropy), is performed using four original bands including red, green, blue, and NIR. Abundance features are computed based on the reflectance of the four original Sentinel-2 bands using a fully constrained least squares algorithm. Then, adaptive spatial attraction features are computed at the subpixel scale, considering the position relationships between subpixels and original pixels and the impact of different class observation sequences. The final subpixel scale CRF model is constructed using the unary and pairwise potentials. To separate the water body and background, a minimum

energy function loss constraint is applied using a criterion based on the maximum a posteriori principle. A graph-cut method is employed to classify the buffer zone into water and background classes.

The second part concerning the lake boundary extraction, involves four sequential steps. Initially, the algorithm computes the NDWI based on the original image to identify potential water distribution areas. A buffer of 50 m with the water boundary identified by NDWI is then applied to set the potential lake boundary buffer zone. Within this buffer zone, independent testing samples are selected, ensuring mutual independence from the training samples. There are 8203 testing pixels, including 4063 water pixels, 2302 snow/ice pixels, and 1838 rock pixels, 19.57% of the total buffer zone pixels. The second step involves upsampling the features obtained during model training to the subpixel scale. This includes GLCM, abundance features, and spatial attraction features, which are computed at the subpixel scale. Consequently, mean-shift oversegmentation is applied at the subpixel scale, and a superpixel segmentation algorithm is constructed at the subpixel scale, incorporating both unary and pairwise potentials of the CRF model.

IV. Results and Discussion

A. Experimental Setup and Accuracy Evaluation

The calculation of the lake area in the Larsemann Hills is conducted utilizing the proposed algorithm. Sentinel-2 imagery with a spatial resolution of 10 m was selected for a long-time series analysis spanning from 2016 to 2023. Since the SVM and CRF require training samples, we manually delineated the ground truth by manual interpretation according to Antarctic aerial imagery, the Antarctic Digital Database (ADD) shape file[1] and the Sentinel-2 true color imagery. The samples were divided into training and testing datasets. Certain lakes were present in the ADD but were not visible in the remote sensing imagery, while others could be observed in the imagery but were not recorded in the ADD. Therefore, lakes of these types were not selected for analysis. For example, on 30 January, 2017, the outburst of the englacial Lake Dålk flooded to Prydz Bay and finally became a moulin. In this study, we focus on three types of lakes including supraglacial, epiglacial, and englacial lakes. Accuracy assessment and area estimation of the lake area calculation results were performed. In addition to the overall accuracy (OA), the standard accuracy variance and kappa coefficients were calculated to assess the performance of each algorithm including the comparison algorithm and the proposed algorithm.

B. Overall Evaluation of the Extraction Results

1) Comparison With Other Algorithms: Different lake mapping algorithms are compared in both qualitative and quantitative terms in this section. We design three algorithms based on the minimum processing unit into groups superpixel, pixel, and subpixel. We deploy three algorithms on these datasets including NDWI, CRF, and SVM in addition to the proposed algorithm, which concurrently considers the three above-mentioned scales, denoted as the Tri-SF algorithm. The SVM comparative algorithm utilizes the combination of RGB and NIR bands as input data. Four bands are considered, and for each band, five GLCM features are computed, including mean, variance, homogeneity, second moment, and entropy. Consequently, a total of 24-dimensional features are extracted and employed as input for the SVM algorithm. As SVM is a supervised classifier, training samples are necessary. For each class (lake, snow, rock), 1000 pixels were selected as training samples. The supervised algorithms such as the CRF and SVM utilize the buffer regions based on the lake shape file, which can be obtained by the NDWI or the manual way. We set the buffer regions with a distance of 50 m. The testing dataset consists of over 8203 pixels, including 4063 water pixels, 2302 snow/ice pixels, and 1838 rock pixels. The training samples and testing samples are mutually independent in this study. Ultimately, during lake extraction, a buffer of approximately 15 326 pixels was generated, representing 19.57% of the total. Similarly, both the CRF algorithm and the proposed algorithm utilize the same set of training samples and testing samples. Since the unary construction of the CRF algorithm is based on the SVM classifier, the same ensemble of features is chosen as input. This ensures consistency in feature representation across algorithms, facilitating a meaningful comparative analysis of their performance. From 67 scenes (L1C with 67 scenes, L2A with 55 scenes), we select two scenes as the example shown in Fig. 2 to visually show the performance among these comparison algorithms.

To explore the water body during different periods that the proposed algorithm can extract, a detailed comparison of Progress Lake and Nella Lake. We presented some visual results for periods including freezing onset (end of February to early March), freezing (April to September), early melting onset (October and November), and melting season (December to January) in Fig. 3. Compared with machine learning algorithms only without considering the different scales including the superpixel, pixel, and subpixels, the proposed algorithm displays excellent potential for investigating the fine boundaries of small water bodies. Fig. 3 also indicates that the proposed algorithm can perform well during different seasonal changes such as during the freezing onset, the comparison method provided rougher boundaries with the same performance during the early melting onset. The main reason for the unsatisfactory results is the mixed pixels effect of small lakes. The second reason is associated with errors introduced by the presence of floating ice on the water surface. The algorithm proposed in this study integrates three different scales to address the influences of both the edges and the floating ice. However, the subpixel mapping could be provided for the small water body extraction.

The scene in the first row during the transition from the freezing period to melting onset shows the different performance on the Nella Lake, where lake ice was misclassified by the NDWI algorithm. It can be attributed to the ice on the lake starting to melt and the signal is not easy to be detected. In the second scene during the melting season, we can find the index from NDWI performs worth on some large lakes. From the visual inspection,

1.Online: [Available]: https://www.add.scar.org/

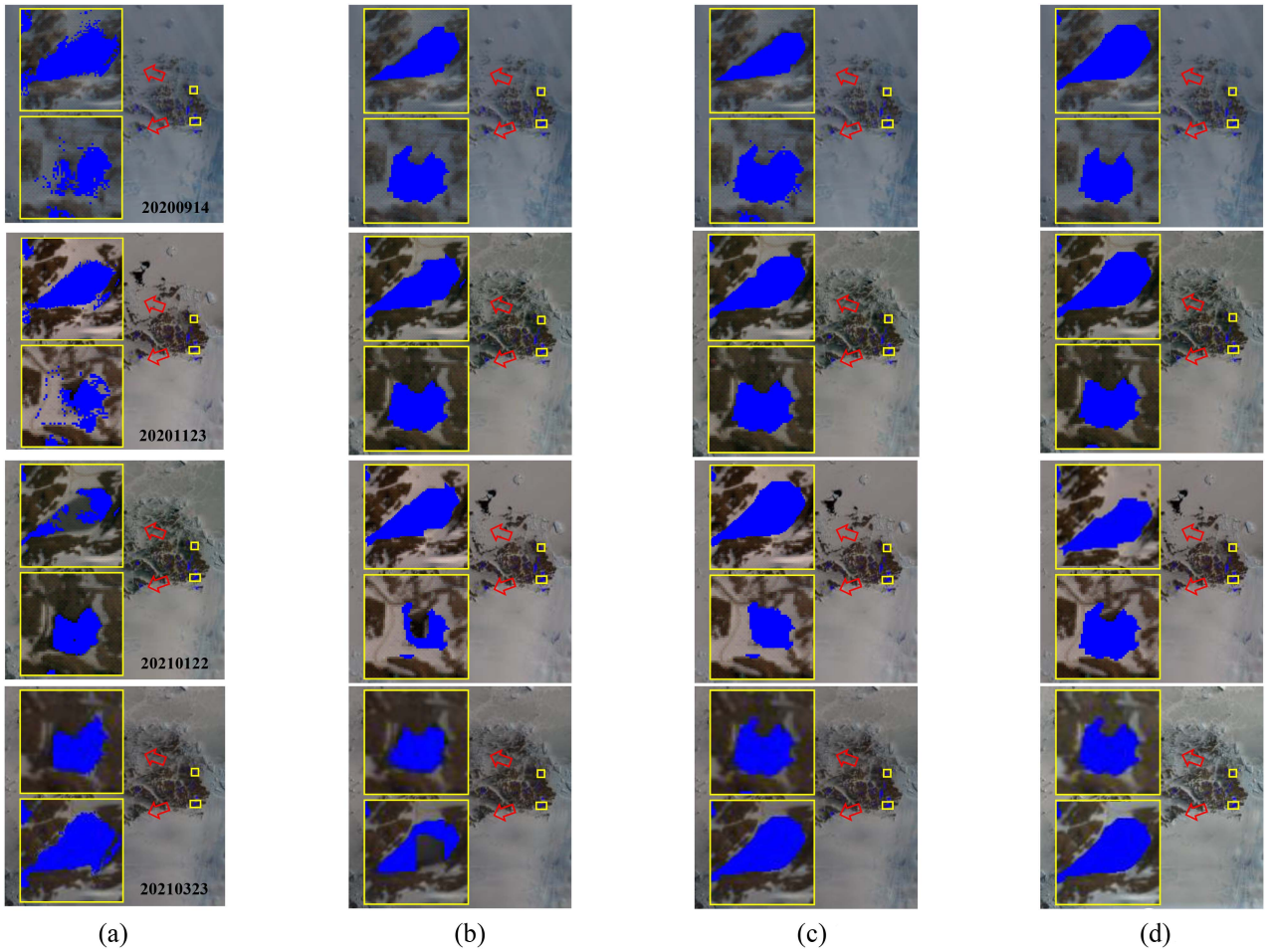


Fig. 3. Results from different algorithms based on different processing units. Examples show two representative lake extraction results using different algorithms during different periods. Lake extraction results from the (a) pixel level, (b) superpixel level, (c) pixel-level SVM, and (d) subpixel-pixel-superpixel level results from the respective algorithm are shown by column.

Nella Lake has been missed by the CRF methods in the third row. However, it is still obvious CRF is not suitable for this small lake type. Although the SVM algorithm provides similar visual results, its computational speed is relatively slow due to pixel-level processing. Subpixel mapping can provide more accurate lake boundaries for different-sized lakes during different scenes. The details can be found in the following discussion. The proposed algorithm obtains a good result in terms of edge precision and visual classification accuracy compared to the reference map and the visual inspection.

The monthly area for lakes in Larsemann Hills from 2016 to 2023 was mapped. The accuracies were calculated based on the confusion matrix, from which the average precision, Kappa coefficients, and standard deviation were obtained shown in Table II. The average estimated OA for the traditional NDWI is $94.88\% \pm 5.24\%$. Meanwhile, the CRF achieves an estimated accuracy of $95.23\% \pm 4.21\%$. The SVM performs superior to NDWI which can be attributed to the lake ice being misclassified as ice. As NDWI with the largest standard deviations, the NDWI is susceptible in Cryosphere lakes mapping. Although the CRF algorithm can obtain higher OA compared to NDWI, the Kappa coefficient is the smallest. This is because CRF utilizes the superpixel strategies that cause some lake pixels to be missed. Due to the NDWI deals pixel by pixel, NDWI provides better results for the small lake type than the superpixels algorithm such as the CRF. By benefiting from the subpixel mapping technique, the proposed algorithm demonstrates the best performance, as depicted in Table II. Based on the Tri-SF algorithm, a time series of the monthly lake area for Progress Lake and Nella Lake is shown in Fig. 4. The interannual and intra-annual variations in the water area of Progress Lake and Nella Lake can be observed from the 7-year dataset. Significant changes occurred primarily after 2020. We speculate that these changes may be related to

TABLE II
ACCURACIES OF DIFFERENT ALGORITHMS IN THE STUDY AREA

Methods	Average OA/%	Kappa	Standard deviation/%
NDWI	94.88	0.9316	5.24
CRF	95.23	0.9231	4.21
SVM	97.26	0.9464	2.07
Tri-SF algorithm	97.84	0.9541	1.87

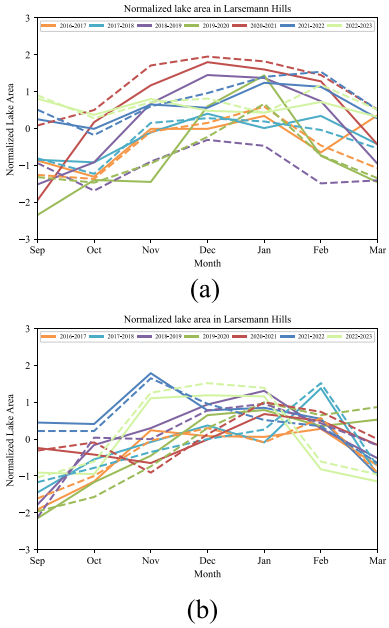


Fig. 4. Time series of the monthly lakes for two typical lakes including (a) Progress Lake and (b) Nella Lake. The solid line shows the lake area extracted by the proposed algorithm and the dashed line shows the lake area extracted by the algorithm that does not consider the subpixel scale.

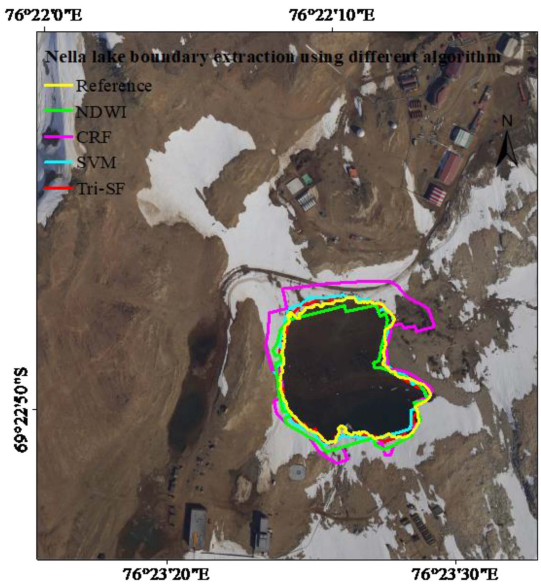


Fig. 5 Lake boundary extraction from different algorithms. The aerial data were acquired from the 39th Antarctic scientific expedition's aerial survey conducted on 8 January, 2023. The Sentinel-2 image as background was acquired on 19 January, 2023, with a time gap of 11 days.

the overflow events that took place in 2020. We will discuss this further in the subsequent sections.

An analysis was conducted comparing the lake boundaries extracted from satellite remote sensing imagery with the reference results obtained from aerial data, shown as yellow in Fig. 5. Taking Nella Lake as an example, it can be observed that the CRF algorithm does not exhibit an advantage in edge

TABLE III
LAKE AREA AND PERIMETER CALCULATION FROM DIFFERENT ALGORITHMS

Methods	Lake Area (km^2)	Lake perimeter (m)
Reference	0.06503	869.03
NDWI	0.06748	930.74
CRF	0.07124	1188.04
SVM	0.06677	884.26
Tri-SF algorithm	0.06437	854.26

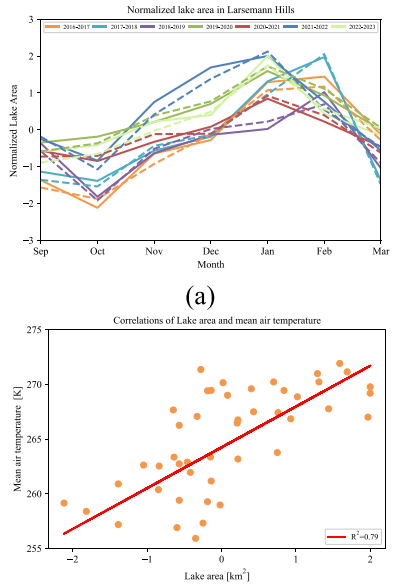


Fig. 6. Time series of lake areas in Larsemann Hills. (a) Time series of the monthly lake area. (b) Scatter plot showing the correlation between surface temperature and the lake area.

extraction, while NDWI tends to underestimate the lake extent. The SVM algorithm lacks accuracy in delineating snow-water boundaries. However, our proposed algorithm demonstrates the resemblance to the in situ lake boundaries reference.

Except for the CRF algorithm, the comparative algorithms exhibit relatively smaller deviations from the lake boundaries derived from the aerial reference imagery. As shown in Table III, the proposed algorithm shows a reduction of 1.7% in the lake perimeter compared to the aerial observations. Similarly, the lake area demonstrates a similar trend, where the CRF algorithm overestimates the lake area by 9.55% relative to the aerial reference imagery. The deviation between the lake area computed in this study and the aerial imagery is only 1.02%.

2) Seasonal Changes of Lakes in Larsemann Hills: The lakes in the Larsemann Hills show continuous area expansion during 2016–2023. The monthly lake areas from September to March display seasonal cycles in the Larsemann Hills Fig. 6(a). The lake areas increase from October to December and then decrease after peaking in January or February. The intra-annual lake extents in Larsemann Hills present different seasonal features before and after 2020. The area reaches its peak usually in February before 2020 while it arrived in January ahead of schedule in later years. This can be ascribed to the melting ice

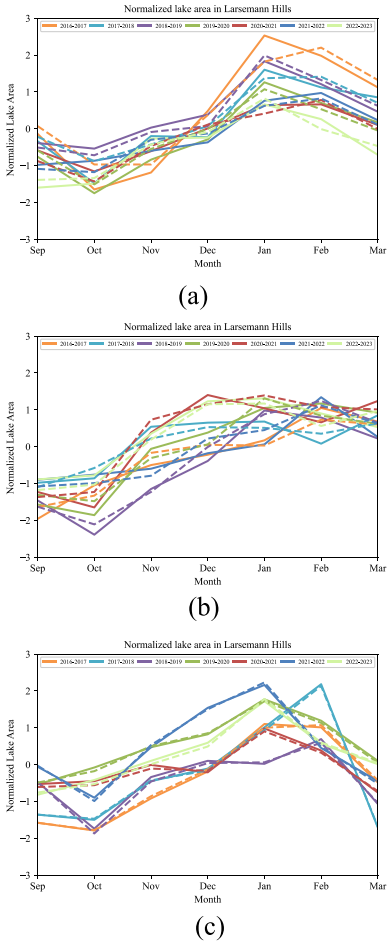


Fig. 7. Time series of different lakes in different divided regions. (a) west (14 lakes) lon $< 70°10'$. (b) middle area (28 lakes) $70°10' <$ lon $< 70°20'$. (c) east (23 lakes) lon $< 70°20'$.

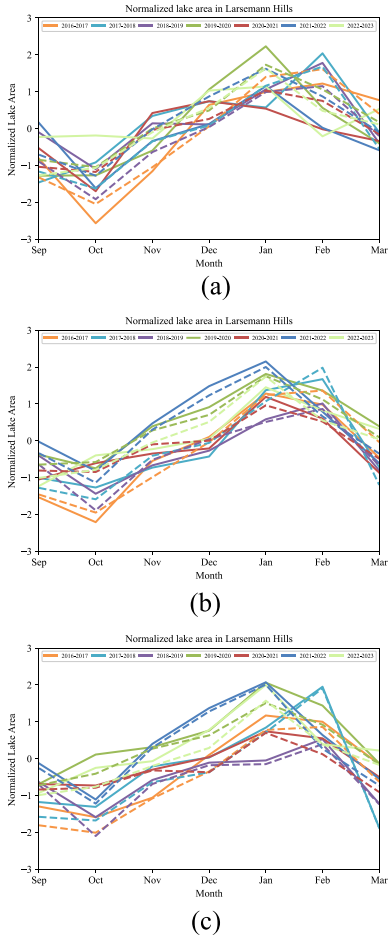


Fig. 8. Time series of different lakes with different sizes. (a) $>0.05\,\text{km}^2$ (6 lakes). (b) $0.01\sim0.05\,\text{km}^2$ (23 lakes). (c) $<0.01\,\text{km}^2$ (36 lakes).

and snow that provide additional water supply for lakes in the polar region. The lake expansion is dominantly driven by the melting snow and ice due to the increasing surface temperature Fig. 6(b). However, the lake expansion may lead to lake overflow or drainage. The most probable occurrence of outburst events is during the austral summer, especially in January or February.

C. Spatial and Temporal Analysis

To investigate the regions prone to significant changes in water body extent leading to outburst events, an analysis of their spatial distribution characteristics is essential. We divided Larsemann Hills into three parts including the west region, the middle region, and the east region for analyzing these lakes as shown in Fig. 7. Lake areas reach their valley value in October, while they reach their peaks in January or February. In the west, lakes reached their maximum area in January during 2016–2019, while the middle and east lakes continued their expansion to February. During 2020–2023, their lake's maximum area reached its peak in January. In addition, the overall seasonal cycles of lakes in the Larsemann Hills also display their expansion accelerating due to the increasing surface temperature.

There are 150 lakes in Larsemann Hills, in the west of the Larsemann Hills, lake areas are larger than in the middle and east. Larger lakes reach their peaks earlier compared with smaller lakes. The rapid accumulation and subsequent expansion and contraction of water bodies in the east indicate a higher likelihood of outburst events in that region.

As shown in Fig. 8, the lake area larger than $0.05\,\text{km}^2$ is mainly located in the eastern region (89.35%), which is consistent with the change in lake area in the eastern region, with the largest seasonal fluctuation in area, and the overall area remains stable. The spatial distribution of the 0.01–$0.05\,\text{km}^2$ lakes is relatively uniform, and the overall change trend is consistent with the changing trend of the total lake area in the Larsemann Hills. Lakes smaller than $0.01\,\text{km}^2$ are mainly located in the central and western regions (55% in the central region and 30% in the western region), with the smallest interannual fluctuations. The accelerating rate of water body expansion in larger lakes, possibly attributed to topography advantages, is counterbalanced by their faster seasonal reduction in area compared to smaller lakes. We can also conclude that the smaller the area

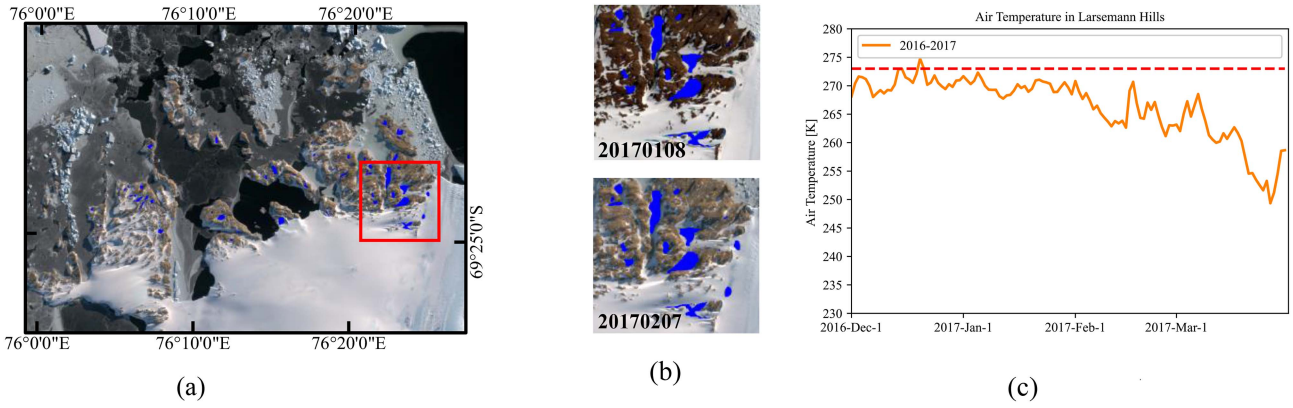


Fig. 9. (a) Location of the Boulder Lake outburst caused the drainage of Lake Dålk on 30 January, 2017. (b) Before and after Sentinel-2 imagery is collected and shows the lake expansion. During the outburst period, temperatures above 273 K have been shown in (c).

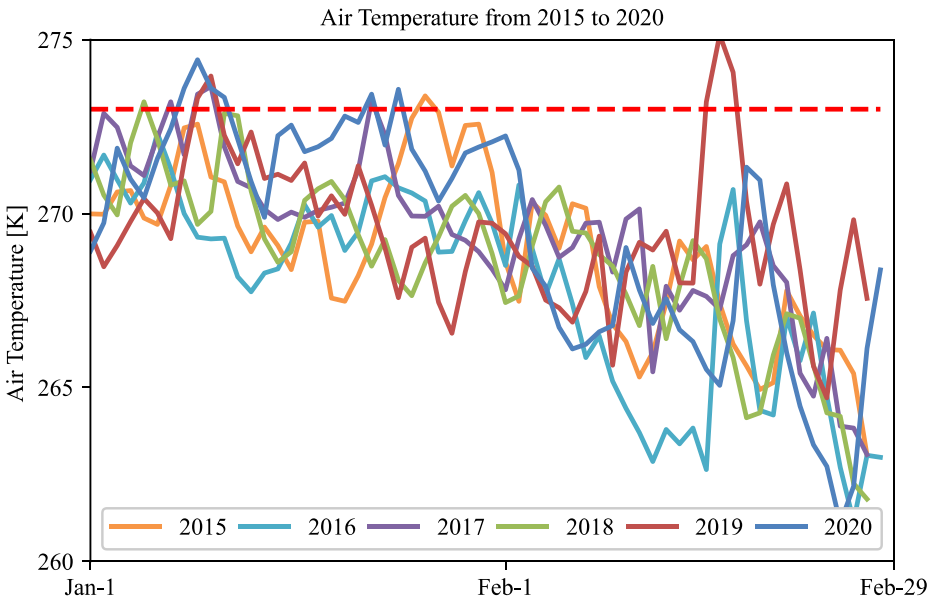


Fig. 10. Daily air temperature from ERA5 reanalysis data for the Larsemann Hills.

of the lake, the greater the deviation of the estimated lake area before and after reducing the mixed pixel effect, and the larger the area, the more stable it is. It proves that the lake boundary does exist in the form of mixed pixels, which highlights the significance of the proposed algorithm. It also shows that the main reason affecting the accuracy of lake area estimation is the determination of the lake boundary. In addition, the proportion of the area $<0.01\,km^2$ to the total lake area ($\sim$12%) is small, so the impact on the overall extraction accuracy is limited.

D. Boulder Outburst Flood

Boulder Lake is a supraglacial type and is located near Progress Lake as shown in Fig. 9(a). The lake is usually frozen it can be observed from the satellite remote sensing imagery during Austral summer. The lake size is a challenge for investigation since it is covered by thick ice. The drainage of Dålk Lake has attracted attention due to the outburst event, which occurred on 30 January, 2017. From the surface temperature data, it can be seen that the Larsemann Hills melting period was earlier compared with the other year, and the temperature above 273 K first appeared in early November 2016 and continued to melt after December until late January 2017. Moreover, the temperature in early February was still fluctuating around 273 K. The lake could not freeze quickly, which led to an increase in the water area expansion in the area, which eventually led to overflow events. From the images of the two scenes before and after the outburst event as shown in Fig. 9(b), it can be seen that the boundary of the lake has a clear tendency to expand, then the Lake Ledyanoe shown at the bottom of Fig. 9(b).

Days of the abnormally high temperature above 273 K in 2017 are far more than other years that showed significant air temperature forcing as shown in Fig. 10. It can be explained that

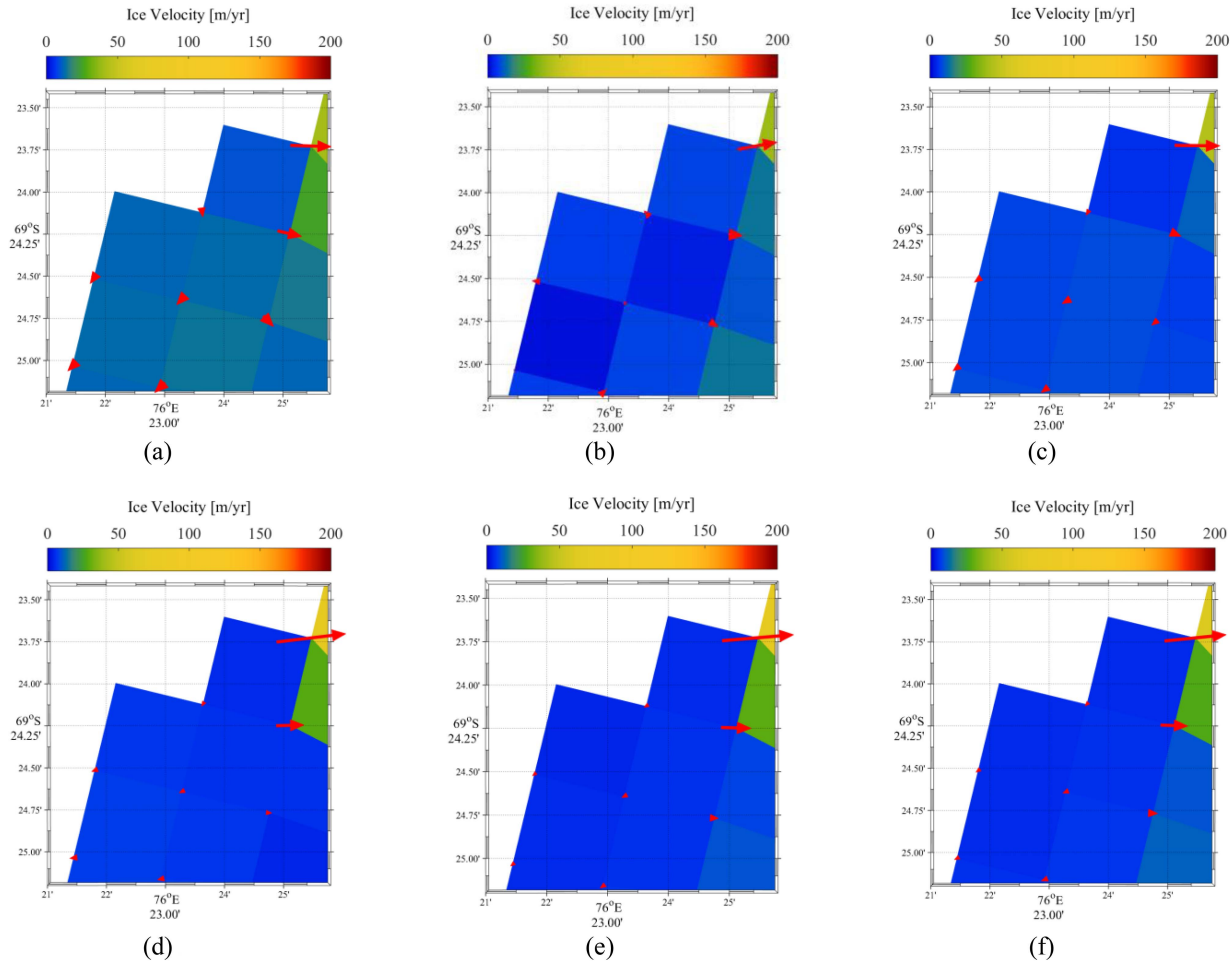


Fig. 11. Time series ice velocity of the region before and after the drainage. Ice velocity of Larsemann Hills data is collected from MeaSUREs. (a) 2014-2015. (b) 2015-2016. (c) 2016-2017. (d) 2017-2018. (e) 2018-2019. (f) 2019-2020.

the lakes melted early in the Larsemann Hills and eventually led to outburst events. Comparing the ice velocity products, it can be seen that the annual ice velocity in the vicinity of Boulder Lake area is slightly higher than that in other areas, and after 2017, the ice velocity, as shown in Fig. 11, has increased dramatically. From the direction of ice velocity, the ice flow direction is consistent with the drainage direction, which eventually leads to the outburst of Boulder Lake to the direction of Lake Dålk. In the early stage of the outburst event on 24 January, 2017, the surface temperature in the Larsemann Hills increased significantly, causing continuous melting, and the area of Boulder Lake expansion.

V. Conclusion

In this article, a novel algorithm considering the superpixel–pixel–subpixel scales was proposed for extracting the small lakes from the satellite remote sensing images. The proposed algorithm for detecting small lakes outperforms those of the compared algorithms and demonstrates good agreement with the lake shoreline delineated by aerial images, achieving an OA of 97.84%, a Kappa coefficient of 0.9541, and a standard deviation of 1.87. Results show that significant seasonal changes in lakes have occurred in the Larsemann Hills. The lake areas present a slowly increasing trend in total, with an accelerated rate after 2020. In terms of the spatial and temporal analysis, the area of lakes in the western region has decreased, while lakes in the central region are stable. Nevertheless, there is a trend of increasing in the northeast region. Affected by surface air temperature, the maximum total lake area has been advanced from February to January each year since 2020.

Under the proposed superpixel–pixel–subpixel multisource data fusion framework, airborne data can quickly extract the accurate boundaries of lakes through superpixel processing strategies. Under the precise registration optimization framework, accurate verification is provided for subpixel mapping achieved in the case of insufficient resolution of onboard data. These comprehensive evaluation results from field data suggest that the proposed algorithm can extract small waterbodies in different seasonal changes, which can be used to understand some of the characteristics of seasonal and incident events. Thus, outburst hydrographs, channel diameters, volume, and duration of floods can be calculated. Future research could be conducted to generate a comprehensive map of ice-free water systems across Antarctica.

REFERENCES

[1] P. M. Atkinson, "Mapping subpixel boundaries from remotely sensed images," in *Innovations in GIS*. London, U.K.: Taylor & Francis, 1997.

[2] R. E. Bell et al., "Antarctic ice shelf potentially stabilized by export of meltwater in surface river," *Nature*, vol. 544, no. 7650, pp. 344–348, 2017, doi: 10.1038/nature22048.

[3] A. Boronina, S. Popov, G. Pryakhina, A. Chetverova, E. Ryzhova, and S. Grigoreva, "Formation of a large ice depression on Dålk Glacier (Larsemann Hills, East Antarctica) caused by the rapid drainage of an englacial cavity," *J. Glaciol.*, vol. 67, no. 266, pp. 1121–1136, 2021, doi: 10.1017/jog.2021.58.

[4] Y. Chen et al., "Subpixel mapping for remote sensing imagery based on spatial adaptive attraction model and conditional random fields," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 1624–1640, 2023, doi: 10.1109/JSTARS.2023.3237745.

[5] D. Corr, A. Leeson, M. McMillan, C. Zhang, and T. Barnes, "An inventory of supraglacial lakes and channels across the West Antarctic Ice Sheet," *Earth Syst. Sci. Data*, vol. 14, no. 1, pp. 209–228, 2022, doi: 10.5194/essd-14-209-2022.

[6] G. L. Feyisa, H. Meilby, R. Fensholt, and S. R. Proud, "Automated water extraction index: A new technique for surface water mapping using Landsat imagery," *Remote Sens. Environ.*, vol. 140, pp. 23–35, 2014, doi: 10.1016/j.rse.2013.08.029.

[7] S. D. Grigoreva, M. R. Kuznetsova, and E. R. Kiniabaeva, "New data on Progress Lake (Larsemann Hills, East Antarctica): Recently discovered subglacial part of the basin," *Polar Sci.*, vol. 38, 2023, Art. no. 100925, doi: 10.1016/j.polar.2023.100925.

[8] I. M. Howat, C. Porter, B. E. Smith, M.-J. Noh, and P. Morin, "The reference elevation model of Antarctica," *Cryosphere*, vol. 13, no. 2, pp. 665–674, 2019, doi: 10.5194/tc-13-665-2019.

[9] R. Heylen, D. Burazerovic, and P. Scheunders, "Fully constrained least squares spectral unmixing by simplex projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4112–4122, Nov. 2011.

[10] S. D. Jawak and A. J. Luis, "Very high-resolution satellite data for improved land cover extraction of Larsemann Hills, Eastern Antarctica," *J. Appl. Remote Sens.*, vol. 7, no. 1, 2013, Art. no. 073460, doi: 10.1117/1.jrs.7.073460.

[11] J. T. M. Lenaerts et al., "Meltwater produced by wind–albedo interaction stored in an East Antarctic ice shelf," *Nature Climate Change*, vol. 7, no. 1, pp. 58–62, 2017, doi: 10.1038/nclimate3180.

[12] C. Liu, J. Shi, X. Liu, Z. Shi, and J. Zhu, "Subpixel mapping of surface water in the Tibetan Plateau with MODIS data," *Remote Sens.*, vol. 12, no. 7, 2020, Art. no. 1154, doi: 10.3390/rs12071154.

[13] Y. Li, B. Dang, Y. Zhang, and Z. Du, "Water body classification from high-resolution optical remote sensing imagery: Achievements and perspectives," *ISPRS J. Photogrammetry Remote Sens.*, vol. 187, pp. 306–327, 2022, doi: 10.1016/j.isprsjprs.2022.03.013.

[14] A. Markov et al., "The conditions of the formation and existence of 'Blue Ice Areas' in the ice flow transition region from the Antarctic ice sheet to the Amery Ice Shelf in the Larsemann Hills area," *Polar Sci.*, vol. 22, 2019, Art. no. 100478, doi: 10.1016/j.polar.2019.08.004.

[15] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. TSMC-9, no. 1, pp. 62–66, Jan. 1979.

[16] B. Qi, Y. Zhuang, H. Chen, S. Dong, and L. Li, "Fusion feature multi-scale pooling for water body extraction from optical panchromatic images," *Remote Sens.*, vol. 11, 2019, Art. no. 245, doi: 10.3390/rs11030245.

[17] E. Shevnina, M. Potes, T. Vihma, T. Naakka, P. R. Dhote, and P. K. Thakur, "Evaporation over a glacial lake in Antarctica," *Cryosphere*, vol. 16, no. 8, pp. 3101–3121, 2022, doi: 10.5194/tc-16-3101-2022.

[18] E. Shevnina, E. Kourzeneva, Y. Dvornikov, and I. Fedorova, "Retention time of lakes in the Larsemann Hills oasis," *East Antarctica Cryosphere*, vol. 15, no. 6, pp. 2667–2682, 2021, doi: 10.5194/tc-15-2667-2021.

[19] E. Shevnina and E. Kourzeneva, "Thermal regime and components of water balance of lakes in Antarctica at the Fildes peninsula and the Larsemann Hills," *Tellus, Ser. A, Dynamic Meteorol. Oceanogr.*, vol. 69, no. 1, 2017, Art. no. 1317202, doi: 10.1080/16000870.2017.1317202.

[20] H. Sui, G. Chen, and L. Hua, "An automatic integrated image segmentation, registration and change detection method for water-body extraction using HSR images and GIS data," *Int. Arch. Photogram., Remote Sens. Spatial Inf. Sci.*, vol. 7/W2, pp. 237–242, 2013, doi: 10.5194/isprsarchives-XL-7-W2-237-2013.

[21] P. Wang, L. Zhang, G. Zhang, H. Bi, M. Dalla Mura, and J. Chanussot, "Superresolution land cover mapping based on pixel-, subpixel-, and superpixel-scale spatial dependence with pansharpening technique," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 10, pp. 4082–4098, Oct. 2019.

[22] Q. Wang, X. Ding, X. Tong, and P. M. Atkinson, "Spatio-temporal spectral unmixing of time-series images," *Remote Sens. Environ.*, vol. 259, 2021, Art. no. 112407, doi: 10.1016/j.rse.2021.112407.

[23] Q. Wang, W. Shi, and P. M. Atkinson, "Enhancing spectral unmixing by considering the point spread function effect," *Spatial Statist.*, vol. 28, pp. 271–283, 2018, doi: 10.1016/j.spasta.2018.03.003.

[24] X. Wang, F. Ling, H. Yao, Y. Liu, and S. Xu, "Unsupervised sub-pixel water body mapping with sentinel-3 OLCI image," *Remote Sens.*, vol. 11, no. 3, 2019, Art. no. 327, doi: 10.3390/rs11030327.

[25] Y. Wang et al., "Self-supervised feature learning with CRF embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 2628–2642, May 2019.

[26] S. Wu et al., "An improved subpixel mapping algorithm based on a combination of the spatial attraction and pixel swapping models for multispectral remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 7, pp. 1070–1074, Jul. 2018.

[27] W. Wu, Q. Li, Y. Zhang, X. Du, and H. Wang, "Two-step urban water index (Tsuwi): A new technique for high-resolution mapping of urban surface water," *Remote Sens.*, vol. 10, 2018, Art. no. 1704, doi: 10.3390/rs10111704.

[28] X. Yang et al., "Monthly estimation of the surface water extent in France at a 10-m resolution using Sentinel-2 data," *Remote Sens. Environ.*, vol. 244, 2020, Art. no. 111803, doi: https://doi.org/10.1016/j.rse.2020.111803.

[29] L. Yu et al., "Features of extreme precipitation at Progress Station," *Antarctica J. Climate*, vol. 31, no. 22, pp. 9087–9105, 2018, doi: 10.1175/JCLI-D-18-0128.1.

[30] G. Zhang et al., "Underestimated mass loss from lake-terminating glaciers in the greater Himalaya," *Nature Geosci.*, vol. 16, pp. 333–338, 2023, doi: 10.1038/s41561-023-01150-1.

[31] G. Zhang, T. Yao, H. Xie, W. Wang, and W. Yang, "An inventory of glacial lakes in the third pole region and their changes in response to global warming," *Glob. Planet. Change*, vol. 131, pp. 148–157, 2015, doi: 10.1016/j.gloplacha.2015.05.013.

[32] Y. Zhang, G. Zhang, and T. Zhu, "Seasonal cycles of lakes on the Tibetan Plateau detected by Sentinel-1 SAR data," *Sci. Total Environ.*, vol. 703, 2020, Art. no. 135563, doi: 10.1016/j.scitotenv.2019.135563.

[33] T. Zhu, F. Li, G. Heygster, and S. Zhang, "Antarctic sea-ice classification based on conditional random fields from RADARSAT-2 dual-polarization satellite images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2451–2467, Jun. 2016.

[34] S. D. Jawak, K. Kulkarni,, and A. J. Luis, "A review on extraction of lakes from remotely sensed optical satellite data with a special focus on cryospheric lakes," *Adv. Remote Sens.*, vol. 4, no. 3, pp. 196–213, 2015, doi: 10.4236/ars.2015.43016.

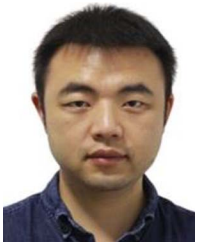
[35] S. K. McFeeters, "The use of the normalized difference water index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.*, vol. 17, no. 7, pp. 1425–1432, 1996, doi: 10.1080/01431169608948714.

Tingting Zhu received the B.S. degree in surveying and mapping from Nanjing Normal University, Nanjing, China, in 2012, and the M.S. and Ph.D. degrees in geodesy and survey engineering from the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China, in 2014 and 2017, respectively.

She is currently with the College of Geomatics Science and Technology, Nanjing Tech University, Nanjing. Her research interests include change detection from high-resolution remotely sensed imagery and multisensor remote sensing in the cryosphere.

Xiangbin Cui received the Ph.D. degree in geological resources and geological engineering from the Department of Earth Sciences, Zhejiang University, Zhoushan, China, in 2010.

He is currently with the Polar Research Institute of China, Shanghai, China. He became a Professor in 2021. He has been to Antarctica six times and has good experience in geophysical and remote sensing observation and research in Antarctica. His research interests include studying the Antarctic ice sheet and its potential change through radio glaciological, geophysical and remote sensing methods.



Yu Zhang received the B.S. degree in communication engineering from the Harbin University of Science and Technology, Harbin, China, in 2011, and the Ph.D. degree in signal and information processing from Wuhan University, Wuhan, China, in 2016.

He is currently working with the Chinese Antarctic Center of Surveying and Mapping, Wuhan University. His research interests include signal processing and microwave remote sensing in the cryosphere.

Kai Lu received the Ph.D. degree in geophysics from the King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, in 2019.

He is currently with the Polar Research Institute of China, Shanghai, China, as an Assistant Professor. He has been to Antarctica twice and has good experience in geophysical and remote sensing observation and research in Antarctica. His research interests include studying the Antarctic ice sheet and its potential change through radio glaciological and seismic methods.

Yuande Yang received the M.S. degree in solid geophysics and the Ph.D. degree in geodesy and geomatics from Wuhan University, Wuhan, China, in 2006 and 2010, respectively.

From 2019 to 2020, he was a Visiting Scholar with Newcastle University, Newcastle, U.K. He is currently a Professor with the Chinese Antarctic Center of Surveying and Mapping, Wuhan University. His work consists of inland water level monitoring, mass balance of ice sheets, and land subsidence monitoring. His research interests include the applications of satellite altimetry, GRACE, and GNSS.