Fine Spatial and Temporal Ice/Snow Surface Temperature Generation: Evaluation Spatiotemporal Fusion Methods in Greenland Ice Sheet

Qing Cheng, Zejun Zhang, Dong Liang, and Fan Ye

***Abstract*—Monitoring ice/snow surface temperature (IST) variations with high spatial and temporal resolution data from satellites are essential for research on the mass balance of the Greenland ice sheet (GrIS). However, the tradeoff between satellite sensors' bandwidth and re-entry cycle, coupled with the influence of cloudy weather, limits their ability to fine-monitor IST. Spatiotemporal data fusion is a way of producing high spatiotemporal datasets. This article uses four spatiotemporal fusion algorithms to fuse the Landsat 8 IST data and the Moderate Resolution Imaging Spectrometer IST to generate fine spatial-temporal IST in the GrIS regions. The quantitative evaluation of the different fusion data shows that the R^2 are all above 0.9. The spatial and temporal nonlocal filter based fusion model (STNLFFM) dual-temporal algorithm provided the highest accuracy with a root mean square error of 2.427 K, followed by the STNLFFM mono-temporal algorithm, the spatial and temporal adaptive reflectance fusion model (STARFM), the flexible spatiotemporal data fusion model, and enhanced STARFM. From the results, the fusion data are accurate and detailed in different regions. That is, the spatiotemporal fusion technique has the potential to generate IST datasets that possess high spatial and temporal resolutions for Greenland.**

***Index Terms*—Greenland, ice and snow, ice/snow surface temperature (IST), spatiotemporal fusion, surface temperature.**

I. Introduction

MONITORING the ice/snow surface temperature (IST) is an essential element in observing the mass balance of the Greenland ice sheet (GrIS). It helps keep track of alterations in the surface energy balance and the exchange of energy between the atmosphere and the surface [1], [2]. According to energy balance models created by several academics [3], [4], if the annual average surface temperature of the GrIS increases by 1 K, the total ablation will rise by approximately 20% to 50%, whereas an increase in summertime temperatures would cause faster melting [5] and result in more material loss [6]. And research on the correlation between the freeze-thaw state of the polar regions and temperature proved a strong correlation between melt area and temperature [7], [8]. Thus, estimating the IST and analyzing its variations is critical to observing variations in the GrIS's mass and tracking environmental processes.

Surface temperature variability in the GrIS can be analyzed from automatic weather station (AWS) measurements and satellite thermal infrared (TIR) data [9]. In general, the high temporal resolution and long-term records of AWS measurements make them an effective means of capturing the temporal variability of IST. However, the measured values do not reflect IST changes across the GrIS. Furthermore, given technical and budgetary constraints, satellite sensors necessitate a compromise between bandwidth and temporal resolution [10], [11]. Therefore, a single satellite sensor cannot deliver high spatial and temporal resolution TIR data for GrIS [12]. The commonly satellite-derived surface temperature datasets include AVHRR polar pathfinder extended dataset [13], MODIS products (MOD11, MOD21) [14], Landsat TM/ETM+/TIRS retrieval data [15], ASTER LST product (AST08) [16], SEVIRI LST product [17], etc. Different sensors have various features. For instance, the TIR bands of TM/ETM+ and ASTER exhibit a spatial resolution between 60 to 120 m. However, their re-entry cycles are longer than half a month. In contrast, sensors, such as MODIS and AMSR-E can obtain daily surface temperature data with extended coverage, but the spatial resolution is 1–10 km [18]. Spatial details and temporal variations are challenging to capture over GrIS.

Spatiotemporal fusion is the process of fusing different remote sensing data from various sensors to produce data with high spatiotemporal resolution [19]. These fusion algorithms can support the fine monitoring of spatial and temporal dynamics of the ground surface. In previous work, spatiotemporal data fusion algorithms were classified into four groups: weight function-based; unmixing-based; learning-based; and hybrid approaches

Manuscript received 7 June 2023; revised 19 June 2023 and 19 August 2023; accepted 3 October 2023. Date of publication 11 October 2023; date of current version 17 November 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 42171383, in part by the International Research Centre of Big Data for Sustainable Development Goals under Grant CBASYX0906, in part by the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) under Grant CUG2106212, and in part by the Open Research Project of The Hubei Key Laboratory of Intelligent Geo-Information Processing under Grant KLIGIP-2023-B04. *(Corresponding author: Dong Liang.)*

Qing Cheng, Zejun Zhang, and Fan Ye are with the School of Computer Science, China University of Geoscience, Wuhan 430079, China (e-mail: qingcheng@whu.edu.cn; zhangzejun@cug.edu.cn; yefan@cug.edu.cn).

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

Digital Object Identifier 10.1109/JSTARS.2023.3323742

[10], [20]. The typical weight function-based fusion method is the spatial and temporal adaptive reflectance fusion model (STARFM). It generates daily reflectance data by weighting the spectral, temporal, and spatial differences between MODIS and Landsat data [21]. Zhu et al. [22] suggested an enhanced STARFM called ESTARFM, which incorporates two conversion factors to improve the precision of predictions in heterogeneous landscapes. Using a Bayesian approach, the spatial and temporal reflectance unmixing model separates coarse pixels to estimate the variations of fine-resolution endmembers. This process involves the use of moving windows to produce a fused image [23]. The spatial and temporal nonlocal filter-based fusion model (STNLFFM), which efficiently avails of the redundant data in remote sensing data with high accuracy and robustness, was proposed by Cheng et al. [24]. Wang and Atkinson [25] proposed a three-step method consisting of regression model fitting, spatial filtering, and residual compensation, called fit-FC, for the case of a small correlation between coarse data and a large time span. Other complex learning-based approaches include the sparse-representation-based spatiotemporal reflectance fusion model [26] and the wavelet-artificial intelligence fusion approach [27]. Hybrid approaches combine the advantages of different methods to improve the performance of spatiotemporal data fusion. Flexible spatiotemporal data fusion (FSDAF) [28] is a typical representative of this kind of fusion method. It based on the spectral unmixing analysis, weighted function, and thin plate spline interpolation to obtain better spatiotemporal fusion data, even with rapid changes in ground cover. In addition, numerous algorithms based on deep learning networks capture the potential relationships among multiple environmental variables, demonstrating remarkable performance. For instance, Lin et al. [29] introduced a cross-attention-based adaptive weighting fusion network based on the residual learning framework. This network finds application in the spatiotemporal fusion of MODIS-Landsat data, producing dense medium-resolution images. Furthermore, Chen et al. [30] proposed the ROBust OpTimization-based fusion model for global data processing. This method necessitates no parameter tuning, boasts computational efficiency, and lends itself to diverse data structures. In addition, Wu et al.'s DSTFN model leverages residual dense blocks and attention mechanism modules to enhance feature representation and extraction [31].

Spatiotemporal fusion methods have extensive applications in various fields [32], [33], [34], [35], [36], including agriculture, ecology, and surface parameter estimation. For instance, they are utilized for crop yield and gross primary productivity estimation, monitoring the cultivation and administration of crops [37], [38], [39], [40], estimating ecological variables, monitoring ecosystem dynamics [41], [42], and estimating surface parameters such as reflectance and land surface temperature (LST) [21], [43], [44], [45], [46]. More studies have been conducted on LST for surface urban heat islands (SUHI) [47], [48], [49], [50]. Shen et al. [48] employed multitemporal and multisensor fusion methods to generate LST for Wuhan from 1988 to 2013. Chang et al. [49] explored the daily variation of SUHI and thermal differences between different land covers in the Boston metropolitan area. Huang et al. [50] used Landsat and MODIS for Beijing in 2002 to generate dense time-series LST data. The above studies showed the feasibility of fusion methods for LST. Nevertheless, few studies of spatiotemporal fusion methods have applied to ice/snow temperature surfaces (IST). In recent years, monitoring the spatiotemporal dynamics of the IST in the poles has become increasingly important owing to the effects of climate change on polar regions [51].

TABLE I
VARIABLE DEFINITIONS

t_0, t_1	reference date
t_k	target date
F	high spatial and low temporal resolution data (i.e., fine-resolution data)
C	low spatial and high temporal resolution data (i.e., coarse-resolution data)
(x_i, y_i)	given pixel coordinate index
W_i	weight of i th similar pixel
w	size of the window
FP	predicted results
$F(x_i, y_i, t)$	high pixel value at t for a given coordinate position (x_i, y_i)
$C(x_i, y_i, t)$	low pixel value at t for a given coordinate position (x_i, y_i)

This article intends to evaluate and apply the spatiotemporal fusion algorithms to produce high spatiotemporal resolution IST data for the GrIS. Four classical and efficient algorithms, STARFM, ESTARFM, FSDAF, and STNLFFM, are selected to fuse MOD21 and Landsat 8 surface temperature products to generate 30 m IST fusion data daily. We opt for these four methods based on: they have already demonstrated accurate results in previous spatiotemporal fusion studies [52], [53], [54]; the deep learning algorithms require substantial high-resolution data and are sensitive to noise, which may introduce potential biases in fusion results; and the GrIS surface cover types are relatively uniform with limited variability.

Leveraging spatiotemporal fusion techniques to produce high-resolution IST data can enhance our analysis of melting ice cap boundaries in fine spatial and temporal dynamics. This article can deepen the application of spatiotemporal fusion to IST and support fine monitoring and research in GrIS.

II. MATERIALS AND METHODS

Before presenting the procedure of each spatiotemporal fusion method, the following explanations of important variable definitions is provided for clarity of presentation (see Table I). The mono- and dual-temporal spatiotemporal fusion schematic diagram is shown (see Fig. 1).

A. Spatial and Temporal Adaptive Reflectance Fusion Model

The STARFM model is based on a weight function that computes the predicted reflectance values of the fine pixels while accounting for spectrum variances, temporal differences,

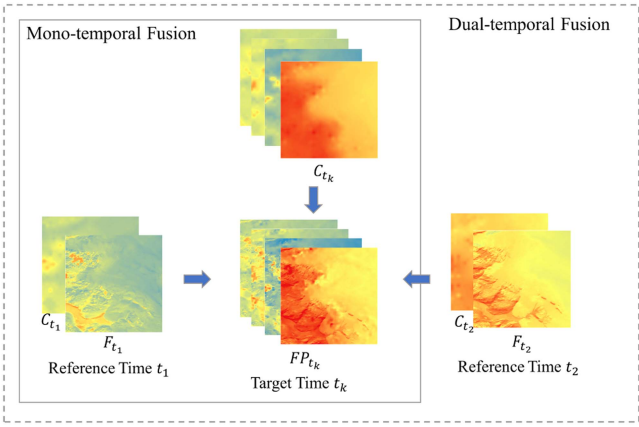

Fig. 1. Schematic diagram of mono- and dual-temporal spatiotemporal fusion.

and spatial distances. Assuming that coarse-resolution pixels are homogeneous, the data preprocessing, acquisition time, sensor bandwidth, and geographic location are responsible for the observed errors in coarse-fine resolution pixel reflectance. Therefore, if the ground cover type and observation errors do not change between t_0 and t_k, disregarding geometric and atmospheric correction errors, the fine-resolution image pixels at t_k can be expressed as follows:

$$F\left(x_i, y_i, t_k\right) = C\left(x_i, y_i, t_k\right) + F\left(x_i, y_i, t_0\right) - C\left(x_i, y_i, t_0\right). \tag{1}$$

When the ground cover type is complicated or changing, mixed pixels can appear in the coarse-resolution images. STARFM places the target pixel in the center of a moving window and uses the surface reflectance of the red and NIR spectra to identify spectrally similar pixels. Similar pixels are weighted according to spectral, temporal, and spatial distances. Ultimately, the target pixel is calculated as follows:

$$FP\left(x_{w/2}, y_{w/2}, t_k\right) = \sum_{i=1}^{w}\sum_{j=1}^{w}\sum_{m=1}^{n} W_{ijm} \times \left(C\left(x_i, y_i, t_k\right) + F\left(x_i, y_i, t_0\right) - C\left(x_i, y_i, t_0\right)\right) \tag{2}$$

$$W_{ijm} = \frac{1/\left(S_{ijm}T_{ijm}D_{ijm}\right)}{\sum_{i=1}^{w}\sum_{j=1}^{w}\sum_{m=1}^{n}\left[1/\left(S_{ijm}T_{ijm}D_{ijm}\right)\right]}. \tag{3}$$

In (2), w refers to the moving window's size, $(x_{w/2}, y_{w/2})$ indicates the target image pixel, n signifies the count of reference dates, and W_{ijm} denotes the weight of the similar pixel, calculated by (3). S_{ijm} represents the spectral difference (i.e., the absolute value of the spectral difference between the image pixels at the same coordinate position in the coarse-fine resolution image in the same time phase). T_{ijm} represents the temporal difference (i.e., the absolute value of the reflectance difference between two coarse-resolution image pixels in different time dates at the same position). D_{ijm} represents the spatial distance (i.e., the relative distance between the target pixel and a similar pixel).

B. Flexible Spatiotemporal Data fusion

The FSDAF algorithm employs unmixing-based techniques, spatial interpolation, and weighting functions. To produce high-resolution data of the target date, FSDAF requires input data of the reference date at both coarse and fine-resolutions and coarse-resolution data of the target date. It generates the fusion result in six main steps.

1) Unsupervised classification of fine-resolution images of t_0. Calculate the specific gravity f_c of each class in the pixel of coarse resolution image.
2) Analysis of the temporal variation in reflectance $\Delta F(c)$ across object classes from coarse-resolution images at t_0 and t_k. l denotes the total number of land cover types within a coarse image pixel, and $\Delta C\left(x_i, y_i\right)$denotes the change in reflectance of coarse pixels from t_0 to t_k

$$\Delta C\left(x_i, y_i\right) = C_k\left(x_i, y_i\right) - C_0\left(x_i, y_i\right) = \sum_{c=1}^{l} f_c\left(x_i, y_i\right) \times \Delta F\left(c\right). \tag{4}$$

3) Calculate the temporal prediction image $F_{t_k}^{\mathrm{TP}}$ and residual R by using $\Delta F(c)$. $F_{t_0}(x_{ij}, y_{ij})$ denotes the reflectance of the jth fine pixel of (x_i, y_i) at t_0, m is the number of fine pixels within one coarse pixel

$$F_{t_k}^{TP}\left(x_{ij}, y_{ij}\right) = F_{t_0}\left(x_{ij}, y_{ij}\right) + \Delta F_k\left(c\right) \tag{5}$$

$$R\left(x_i, y_i\right) = \Delta C\left(x_i, y_i\right) - \frac{1}{m}\left[\sum_{j=1}^{m} F_{t_k}^{\mathrm{TP}}\left(x_{ij}, y_{ij}\right) - \sum_{j=1}^{m} F_{t_0}\left(x_{ij}, y_{ij}\right)\right]. \tag{6}$$

4) Adopt TPS to downscale the coarse-resolution image at t_k and obtain the spatial prediction image $F_{t_k}^{SP}$. N is the total number of pixels in the coarse resolution image. The constraints are $\sum_{i=1}^{N} b_i = \sum_{i=1}^{N} b_i x_i = \sum_{i=1}^{N} b_i y_i = 0$. Then, a_0, a_1, a_2, and b_i are optimized by minimizing $\sum_{i=1}^{N}\left|C_{t_k}(x_i, y_i) - \sum_{j=1}^{m} F_{t_k}^{SP}(x_{ij}, y_{ij})^2\right|$

$$F_{t_k}^{SP}\left(x_{ij}, y_{ij}\right) = a_0 + a_1 x_{ij} + a_2 y_{ij} + \frac{1}{2}\sum_{i=1}^{N} b_i r_i^2 \log r_i^2 \tag{7}$$

where $r^2 = (x - x_i)^2 + (y - y_i)^2$.
5) The weight for combining the $F_{t_k}^{\mathrm{TP}}$ and $F_{t_k}^{\mathrm{SP}}$ through homogeneity index HI is derived. Then, the residual is distributed $r(x_{ij}, y_{ij})$ by the normalized weighting coefficient $W(x_{ij}, y_{ij})$

$$\mathrm{CW}\left(x_{ij}, y_{ij}\right) = \left(F_{t_k}^{\mathrm{SP}}\left(x_{ij}, y_{ij}\right) - F_{t_k}^{\mathrm{TP}}\left(x_{ij}, y_{ij}\right)\right) \times \mathrm{HI}\left(x_{ij}, y_{ij}\right) + R\left(x_i, y_i\right) \times \left(1 - \mathrm{HI}\left(x_{ij}, y_{ij}\right)\right) \tag{8}$$

$$W\left(x_{ij}, y_{ij}, b\right) = \mathrm{CW}\left(x_{ij}, y_{ij}, b\right) \Big/ \sum_{j=1}^{m} \mathrm{CW}\left(x_{ij}, y_{ij}, b\right) \tag{9}$$

$$r\left(x_{ij}, y_{ij}\right) = m \times R\left(x_i, y_i\right) \times W\left(x_{ij}, y_{ij}\right) \tag{10}$$

where $\mathrm{HI}(x_{ij}, y_{ij}) = (\sum_{k=1}^{m} I_k)/m$, If the kth fine pixel within a moving window, whose size is one coarse pixel, has the same land cover type as the target pixel (x_{ij}, y_{ij}), then $I_k = 1$; otherwise, $I_k = 0$.

6) The final fine-resolution prediction image FP_{t_k} is formed using the domain information. The weights W_c of all similar image pixels are calculated based on relative distance D_c, and n is the total number of similar pixels

$$D_c = \frac{1 + \sqrt{(x_c - x_i)^2 - (y_c - y_i)^2}}{w/2} \tag{11}$$

$$W_c = (1 + D_c) \Big/ \sum_{c=1}^{n} (1 + D_c) \tag{12}$$

$$\mathrm{FP}_{t_k}(x_{ij}, y_{ij}) = F_{t_0}(x_{ij}, y_{ij}) + \sum_{c=1}^{n} W_c \times (r(x_{ij}, y_{ij}) + \Delta F(c, B)). \tag{13}$$

C. Enhanced STARFM Algorithm

The enhanced STARFM algorithm (ESTARFM) focuses on the trend of reflectance of the ground cover type, which is more conducive to the spatiotemporal integration of heterogeneous areas. The ESTARFM inputs two sets of coarse- and fine- resolution images at known dates and a coarse-resolution image at the target dates to output a fine-resolution image. The relationship between coarse- and fine-resolution pixels can be represented linearly for homogeneous pixels, presuming that coarse-resolution images have been resampled to the same spatial resolution as the fine-resolution images, and (x_i, y_i) is taken as fine-resolution pixels within a mixed coarse-resolution pixel

$$F(x_i, y_i, t_0) = a \times C(x_i, y_i, t_0) + b. \tag{14}$$

The conversion factor $v(x_i, y_i)$ is introduced for heterogeneous pixels. Where $h(x_i, y_i)$ signifies the change rate of pixel (x_i, y_i), h_c is the change rate of type c, l denotes the total number of land cover types within a coarse image pixel, and f_c represents the specific gravity of each type in the pixel of coarse resolution image

$$v(x_i, y_i) = \frac{h(x_i, y_i)}{\sum_{c=1}^{l} \frac{f_c h_c}{a}} \tag{15}$$

$$F(x_i, y_i, t_k) = F(x_i, y_i, t_0) + v(x_i, y_i) \times (C(x_i, y_i, t_k) - C(x_i, y_i, t_0)). \tag{16}$$

The ESTARFM inherits the moving window mechanism from STARFM to filter similar pixels. In (17), $F(x_{w/2}, y_{w/2}, t)$ determines the reflectance of the fine pixel at the center of the window, W_i represents the weight of the i th similar image pixel, is calculated with the same formula as STARFM

$$F_0(x_{w/2}, y_{w/2}, t_k) = F(x_{w/2}, y_{w/2}, t_0) + \sum_{i=1}^{n} (W_i \times v(x_i, y_i) \times (C(x_i, y_i, t_k) - C(x_i, y_i, t_0))). \tag{17}$$

The temporal weights T_0 and T_1 are calculated from the magnitude of the change in reflectance of the coarse resolution image at t_0 and t_1, and the final fusion equation for predicting the temporal t_k is as follows:

$$T_p = \frac{1/|\sum_{i=1}^{w} C(x_i, y_i, t_p) - \sum_{i=1}^{w} C(x_i, y_i, t_k)|}{\sum_{p=0,1} (1/|\sum_{i=1}^{w} C(x_i, y_i, t_p) - \sum_{i=1}^{w} C(x_i, y_i, t_k)|)}, \quad p = 0, 1 \tag{18}$$

$$F(x_{w/2}, y_{w/2}, t_k) = T_0 \times F_0(x_{w/2}, y_{w/2}, t_k) + T_1 \times F_1(x_{w/2}, y_{w/2}, t_k). \tag{19}$$

D. Spatial and Temporal Nonlocal Filter-Based Fusion Model

The STNLFFM algorithm combines the STARFM algorithm with nonlocal filtering. Its selection of similar pixels is not limited to the moving window, but selects more similar image pixels of the same feature type from the global perspective, effectively using the redundant information in remote sensing images. STNLFFM assumes a linear relationship between the reference and target dates. The regression coefficients a and b are calculated from the coarse resolution reflectance images of multiple dates, Ω denotes the set of similar image pixels, n denotes the number of reference dates, and the target pixels are calculated as follows:

$$C(x_i, y_j, t_p) = a(x_i, y_j, \Delta t) \times C(x_i, y_j, t_k) + b(x_i, y_j, \Delta t) \tag{20}$$

$$F(x_i, y_j, t_p) = \sum_{i,j}^{\Omega} \sum_{k=1}^{n} W(x_i, y_j, t_k) \times [a(x_i, y_j, \Delta t) \times F(x_i, y_j, t_k) + b(x_i, y_j, \Delta t)]. \tag{21}$$

STNLFFM divides the weights of similar pixels into individual weights W_{in} and overall weights W_{wh}. The former is based on the spatial correlation of the image pixels and assigned in a way borrowed from the nonlocal mean filtering algorithm, whereas the latter is based on the degree of difference between images in the auxiliary and target dates. The smaller the difference, the larger the weight

$$W_{\text{wh}}(x_i, y_j, t_k) = \frac{1 \Big/ \sum_{i=1}^{w} \sum_{j=1}^{w} (|C(x_i, y_j, t_k) - C(x_i, y_j, t_0)|)}{\sum_k \left(1 \Big/ \sum_{i=1}^{w} \sum_{j=1}^{w} (|C(x_i, y_j, t_k) - C(x_i, y_j, t_0)|)\right)} \tag{22}$$

$$W_{\text{in}}(x_i, y_j, t_k) = \exp\left(-\frac{G \times \|C(P(x_i, y_j, t_k)) - C(P(x_i, y_j, t_0))\|}{h^2}\right) \tag{23}$$

$$W(x_i, y_j, t_k) = W_{\text{in}}(x_i, y_j, t_k) \times W_{\text{wh}}(x_i, y_j, t_k) \tag{24}$$

where $C(P(x_i, y_j, t))$ is the neighborhood matrix of similar pixels in the coarse resolution image, $P(x_i, y_j, t)$ is the neighborhood block centered at (x_i, y_j). G is the Gaussian kernel, h is the filtering parameter, and W is the final weight value.

TABLE II
RELEVANT PARAMETERS OF LANDSAT, MODIS, AND SITE DATA USED IN THE EXPERIMENT

AWS	UPE_L	UPE_U	KAN_B	KAN_L	KAN_M	KAN_U
Latitude (°N)	72.893	72.888	67.125	67.096	67.067	67.000
Longitude (°W)	54.296	53.578	50.183	49.951	48.836	47.025
Elevation (m)	220	940	350	670	1270	1840
AWS Timing	16:00		15:00			
Landsat Timing	15:35–15:45		14:40–14:50			
MODIS Timing	14:20–16:35		14:00–15:50			
Number of Landsat	17/4	17/4	19/6	41/16	41/24	22/15
Number of MODIS	127/75	127/99	165/55	165/72	261/171	256/210
Bias (Landsat-AWS)	−2.889	−2.795	−2.527	−3.622	−4.121	−3.531
Bias (MODIS-AWS)	−2.320	0.226	−0.451	−1.300	−0.012	−1.508
R^2(Landsat-AWS)	0.890					
R^2(MODIS-AWS)	0.973					

Note: For A/B, A is the number of remote data. B is the number of valid values in site data.

III. EXPERIMENTS

A. Study Sites and Data

The high-resolution products studied in this article are Landsat 8 surface temperature [15]. These data have a spatial resolution of 30 m and are collected every 16 days.[1] A single-channel algorithm is used to generate these products. Low-resolution data in this article refers to the MOD21 [55] surface temperature product obtained from MODIS. These data have a spatial resolution of 1 km and are collected daily, providing a broader temporal coverage compared with the high-resolution datasets.[2] The data preprocessing date includes the following.

1) The MOD21 was reprojected to the same coordinate system as Landsat data and resampled to 30 m.
2) The vacant data of the study area is filled using spatial neighbor interpolation.
3) The MOD21 is geographically aligned based on Landsat 8 data to eliminate errors between sensors.

We use AWS data and Landsat data to verify the accuracy of different algorithms. Since 2007, the programme for monitoring of the Greenland ice sheet (PROMICE) has been actively collecting data on climate and ice sheet characteristics, recording hourly data, including wind speed, air temperature, and surface temperature.[3] Considering the site location and data time continuity, we select six AWSs of PROMICE for the verification: UPE_L, UPE_U, KAN_B, KAN_L, KAN_M, and KAN_U. Some parameters of remote sensing data and AWS data are given in Table II.

Some researchers have evaluated the accuracy of surface temperature remote sensing products. The single-channel approach employed for Landsat yields root-mean-square errors (RMSEs), ranging from 2.2 to 3.5 K [56], [57]. The temperature emissivity separation algorithm used in MOD21 results in an RMSE of less than 1.5 K [58]. We also calculate the pertinent indexes for the two datasets in the Greenland regions. Landsat IST has a relatively smaller sample size, and the presence of a few extreme deviation values can exert a notable influence on the final outcomes. Thus, results indicate that the Landsat 8 IST data exhibit varied deviations at different AWS, with a maximum bias of −4.121 K and a minimum of −2.527 K, and an overall correlation R^2 of 0.890. The MOD21 IST data have a bias of approximately −1 K and display a correlation R^2 of 0.973, as given in Table II.

B. Experiment Design

To investigate the applicability of different spatiotemporal fusion methods to predict the Greenland IST, we design two comparison experiments by controlling the number of reference dates. First experiment: monotemporal fusion by evaluating STARFM; FSDAF; and STNLFFM. We input a Landsat 8 IST and MOD21 pair for the reference time t_1 and MOD21 at the target time t_k. The output is a fusion image for the target time t_k. Second experiment: dual-temporal fusion; adopting ESTARFM; and STNLFFM. Fig. 1 shows that two Landsat 8 IST and MOD21 pairs for reference times t_1 and t_2 were input and MOD21 for the target time t_k and fusion image for the target time t_k was output. Table III gives experimental reference dates for each AWS.

We validate the fusion image accuracy through quantitative evaluation (PROMICE AWSs data) and Landsat verification. In the quantitative evaluation part, the IST values of fused data are compared with PROMICE AWSs data for long-time series. For the Landsat validation part, we compare the fused data with Landsat 8 IST data. The validation results include a quantitative evaluation with the indices of the RMSE, mean absolute error (MAE), coefficient of determination (R^2), and qualitative analysis.

IV. RESULTS AND DISCUSSION

A. Verification With AWS Measurements

This article investigates the performance of four spatiotemporal fusion methods (STARFM, FSDAF, ESTARFM, and STNLFFM) for IST fusion in GrIS. Fig. 2 presents the scatter plots of 670 sets of fused results and AWSs data relevance and fit functions for all methods. The blue line in the scatter plots represents the $Y = X$ line. We observed that the fusion results of different methods are generally consistent. The slope of the fitted function for all methods ranged from a maximum of 0.991

[1](Data source: https://earthexplorer.usgs.gov/)
[2](Data source: https://search.earthdata.nasa.gov/)
[3] (Data source: http://www.promice.dk)

TABLE III
REFERENCE DATES FOR AWSS

	UPE_L	UPE_U	KAN_B	KAN_L	KAN_M	KAN_U
Reference Dates	04/29/2019 08/14/2020	04/29/2019 08/14/2020	10/07/2019 10/09/2020	10/07/2019 10/09/2020	08/29/2019 10/09/2018	08/29/2019 08/15/2020

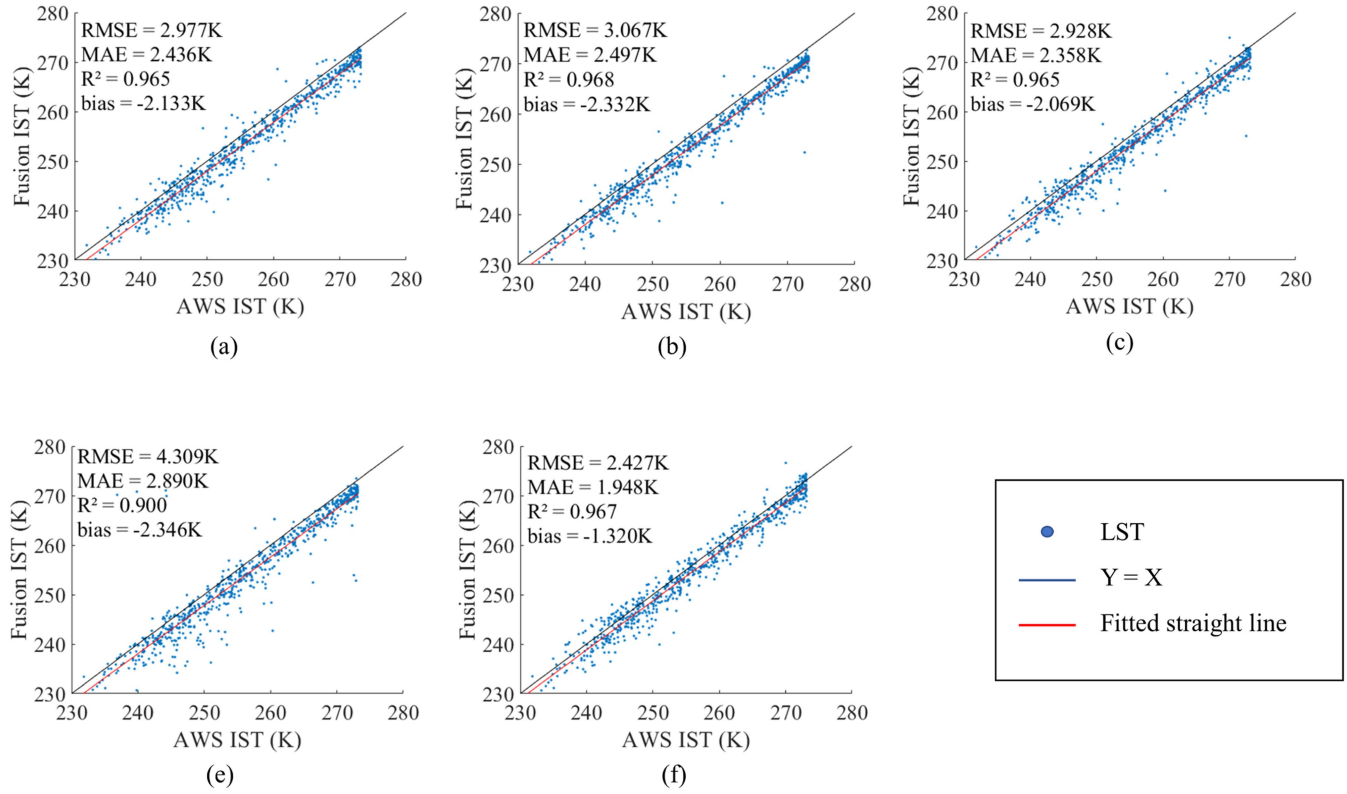

Fig. 2. Quantitative evaluation results of the spatiotemporal fusion methods. (a) STARFM $y = 0.982x + 2.405$. (b) FSDAF $y = 0.982x + 2.206$. (c) STNLFFM (Mono-temporal Fusion) $y = 0.988x + 1.091$. (e) ESTARFM $y = 0.975x + 4.043$. (f) STNLFFM (Dual-temporal Fusion) $y = 0.991x + 1.084$.

TABLE IV
QUANTITATIVE EVALUATION VALUES OF MONO-TEMPORAL FUSION

AWS	STARFM			FSDAF			STNLFFM (Mono-temporal Fusion)		
	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
UPE_L	4.440	3.629	0.893	4.204	3.294	0.911	3.842	2.911	0.957
UPE_U	2.499	1.814	0.965	3.561	2.265	0.942	3.011	1.715	0.971
KAN_B	2.780	2.197	0.926	2.876	2.321	0.925	3.387	2.720	0.914
KAN_L	4.269	4.026	0.978	3.719	3.447	0.980	2.889	2.481	0.975
KAN_M	1.660	1.381	0.990	1.972	1.708	0.989	1.372	1.051	0.990
KAN_U	2.884	2.683	0.992	2.876	2.689	0.993	2.708	2.500	0.992
All	2.977	2.436	0.965	3.067	2.497	**0.968**	**2.928**	**2.358**	0.965

Unit: K
The bold entities represent the optimal performance among the quantitative evaluation of the same indicator.

to a minimum of 0.975, and the difference between them was insignificant. The results of all the methods are distributed mostly below the $Y = X$ line. The reason is that the original Landsat data are generally lower than the AWSs data, with the negative values of the bias (see Table II). However, Fig. 2 shows agreements between the fusion IST and the situ measurements. All fusion methods' correlation coefficient R^2 values range from 0.900 to 0.968. The RMSE values are from 4.309 K (ESTARFM) to 2.427 K (dual-temporal STNLFFM). These results suggest a satisfactory accuracy of the fusion methods for generating IST data.

Tables IV and V list the quantitative evaluation results of monotemporal and dual-temporal fusion methods at each PROMICE site. The quantitative metrics of the three monotemporal fusion methods, STARFM, FSDAF, and STNLFFM (see Table IV), differed very little for all sites. Table V gives that, for the two dual-temporal fusion methods ESTARFM and

TABLE V
QUANTITATIVE EVALUATION VALUES OF DUAL-TEMPORAL FUSION

AWS	ESTARFM			STNLFFM (Dual-temporal Fusion)		
	RMSE	MAE	R^2	RMSE	MAE	R^2
UPE_L	7.856	6.420	0.673	2.981	2.153	0.920
UPE_U	6.137	2.972	0.760	1.643	1.374	0.990
KAN_B	4.482	3.893	0.927	2.986	2.356	0.923
KAN_L	3.951	3.502	0.944	3.642	3.275	0.973
KAN_M	1.743	1.167	0.980	1.996	1.724	0.990
KAN_U	2.734	2.526	0.992	2.862	2.668	0.993
All	4.309	2.890	0.900	**2.427**	**1.948**	**0.967**

Unit: K
The bold entities represent the optimal performance among the quantitative evaluation of the same indicator.

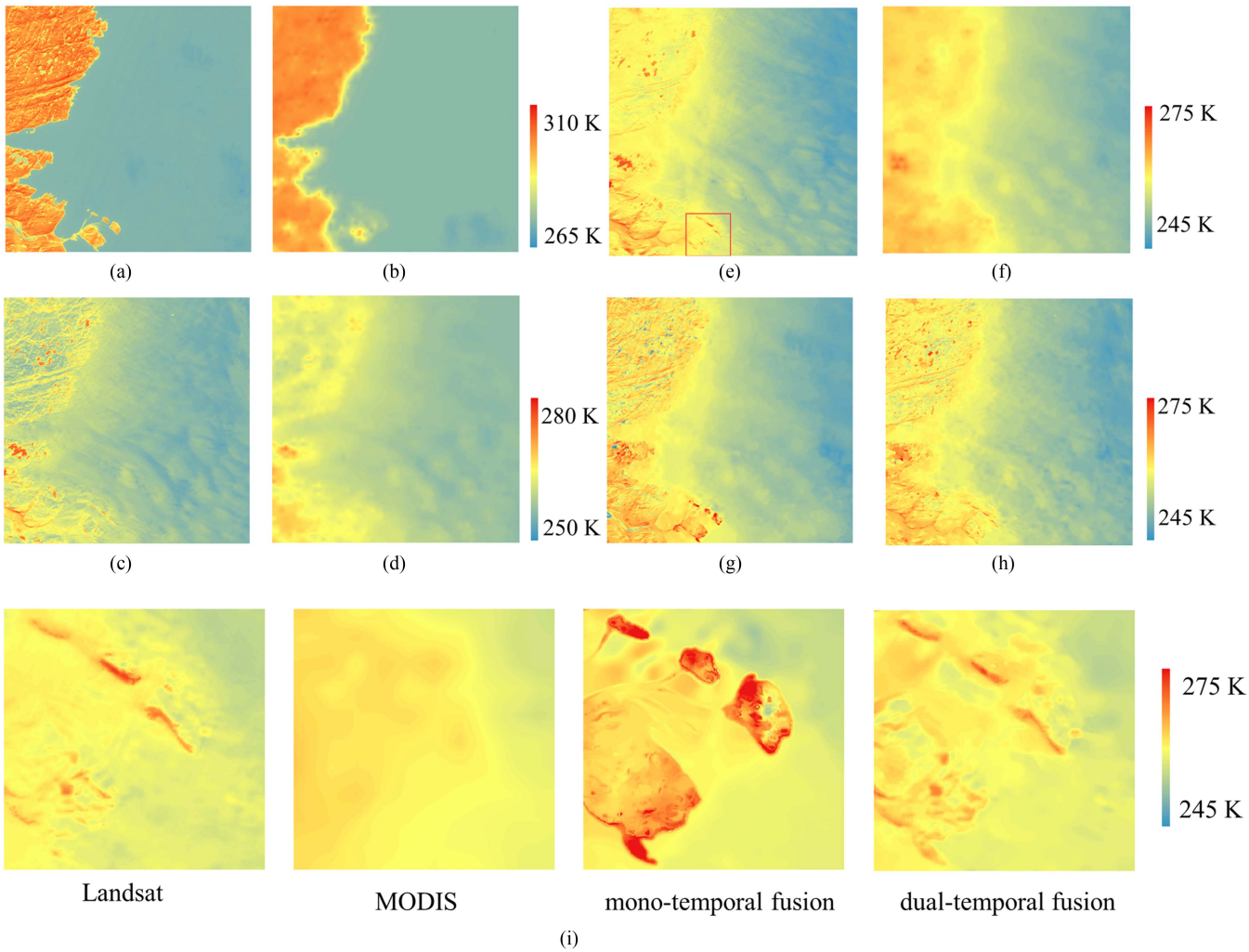

Fig. 3. (a)–(d) Reference Landsat-MODIS pairs on June 8, 2019 and September 30, 2020. (e)–(h) Observed Landsat-MODIS pairs and prediction results on October 14, 2019. (g) Mono-temporal fusion result. (h) Dual-temporal fusion result. (i) Zoomed-in detail figures of (e)–(h).

STNLFFM, the RMSE values were 4.309 K and 2.427 K, and the R^2 values were 0.900 and 0.967, respectively. The fusion IST of STNLFFM has a better correlation and is more accurate. In terms of each site results, at UPE_L and UPE_U, the R^2 values of the dual-temporal STNLFFM are 0.920 and 0.990, respectively, and the RMSE values are 2.981 and 1.643 K, which performed better than other methods. At the four AWSs of KAN_N, KAN_L, KAN_M, and KAN_U, the R^2 values of all algorithms are higher than 0.910. The STNLFFM method has better performance in RMSE and MAE. Overall, the fusion results of STNLFFM have better agreement with the site IST data, and this method is more robust. Moreover, the dual-temporal STNLFFM is more

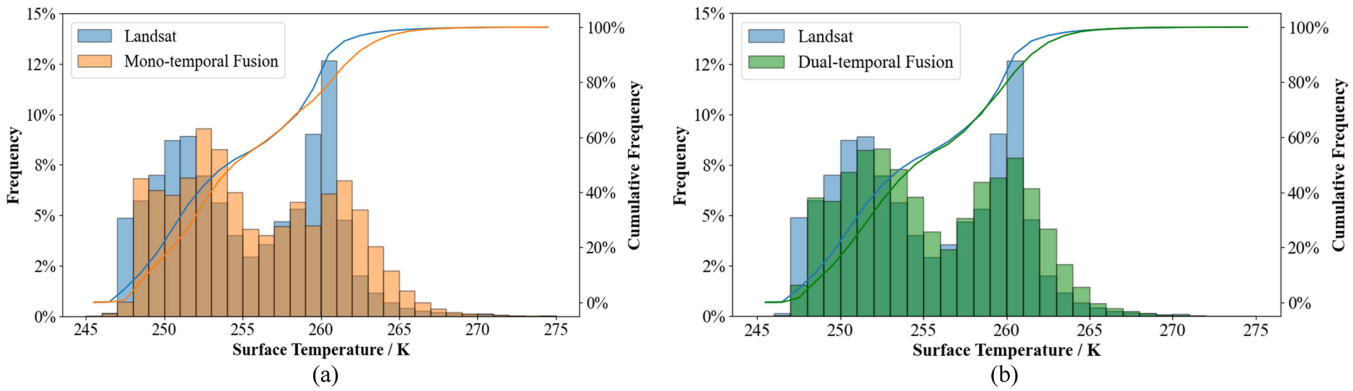

Fig. 4. Frequency distribution of Landsat surface temperature and Fusion data in case 1. (a) Mono-temporal fusion data. (b) Dual-temporal fusion data.

accurate than the mono-temporal STNLFFM for generating IST data.

B. Verification With Landsat Data

In the quantitative assessment in Section IV-A, the STNLFFM fusion method performs more satisfactorily. Thus, in this part, we chose the mono- and dual-temporal fusion methods of STNLFFM to generate IST data for the Greenland regions. The fused IST data and Landsat 8 IST are compared in detail. In this test, we use the Landsat 8 IST and MOD21 data on the same day, with less than two hours of intervals. Each fusion image covers an area of 60 km × 60 km and contains 2000 × 2000 image pixels with a spatial resolution of 30 m.

The Greenland IST volatility follows a particular regional pattern, with significant spatial variation as a seasonal change. The IST variation trend varies enormously from north to south in spring and fall, but from east to west in winter. Meanwhile, the north is experiencing the most significant warming, whereas the southwest and southeast experience the highest temperatures, but are the most prevalent cooling regions [59], [60]. During the summers of 2012 and 2019, the GrIS experienced the most severe ice cap melting [61], [62]; summer surface temperature increases were most pronounced in the southwest and north. This part chooses marginal lands and inland ice caps in the southwest as the study areas [63].

1) Case 1 for Marginal Ice Zone: The study region is located on the edge land of southwest Greenland. Fig. 3(a)–(d) are the input reference data. Fig. 3(a) and (b) show the Landsat IST and MOD21 on 8 June 2019. Fig. 3(c) and (d) shows the Landsat IST and MOD21 on September 30, 2020. The target date is October 14, 2019. Fig. 3(e) is the real Landsat IST used to assess the performance of the fusion results. Fig. 3(f) is the input MOD21 on the target date. Fig. 3(g) and (h) show the mono- and dual-temporal STNLFFM fusion results, respectively. The three sets of data span up to 15 months.

The GrIS experienced an extremely warm melt in 2019, with positive temperature anomalies in the summer (June–August) concentrated in the southwest [64]. Fig. 3 shows that STNLFFM can estimate the IST accurately, although the surface temperature at the reference date differs from that at the target date. The fusion results are with more details compared with MOD21. The quantitative assessment results are presented in Table VI. The R^2 values of the mono- and dual-temporal fusion IST are 0.865 and 0.964, the RMSE values are 2.002 and 1.077 K, respectively. Fusion results can accurately reproduce the temperature variation. However, the mono-temporal fusion result introduces some noise from the reference date in the highly heterogeneous regions, whereas the dual-temporal fusion data [see Fig. 3(h)] is in better agreement with the actual Landsat data [see Fig. 3(e)]. For the red-boxed area in Fig. 3(e), Fig. 3(i) presents an enlarged view of this region, illustrating that the fusion data appears clearer than the MOD21 IST, and the result of dual-temporal STNLFFM is closer to the Landsat data.

Fig. 4 shows that the IST values are between 247 and 263 K, and the spatiotemporal fusion results are generally consistent with the IST distribution of the Landsat. As illustrated in the Fig. 5, the difference between Landsat IST and the fusion IST

TABLE VI
QUANTITATIVE EVALUATION VALUES

Case 1 (unit: K)		
	STNLFFM (Mono)	STNLFFM (Dual)
R^2	0.865	**0.964**
RMSR	2.002	**1.077**
bias	-0.733	**-0.536**

The bold entities represent the optimal performance among the quantitative evaluation of the same indicator.

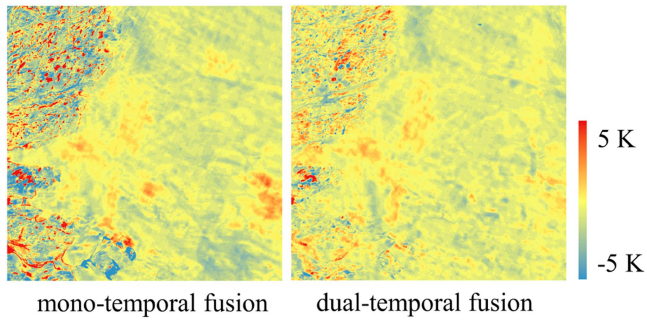

Fig. 5. Difference of IST (Landsat IST minus the fusion data) for the case 1.

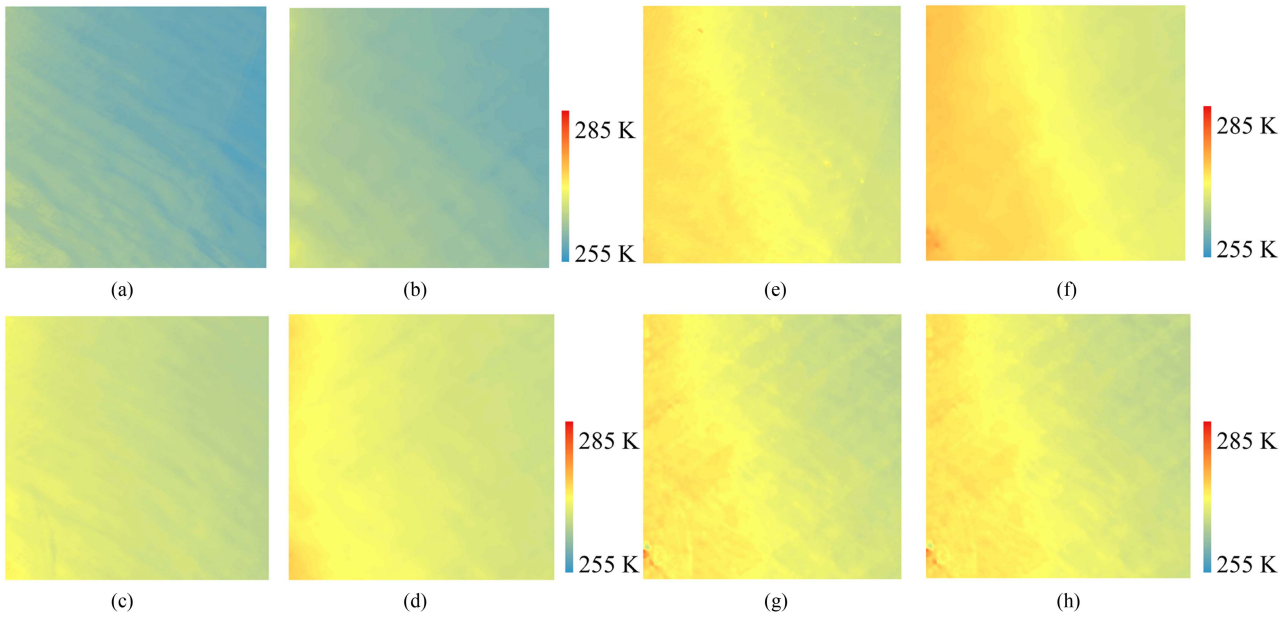

Fig. 6. (a)–(d) Reference Landsat-MODIS pairs on April 9, 2020 and October 9, 2020. (e)–(h) Observed Landsat-MODIS pairs and prediction results on August 15, 2020. (g) Mono-temporal fusion result. (h) Dual-temporal fusion result.

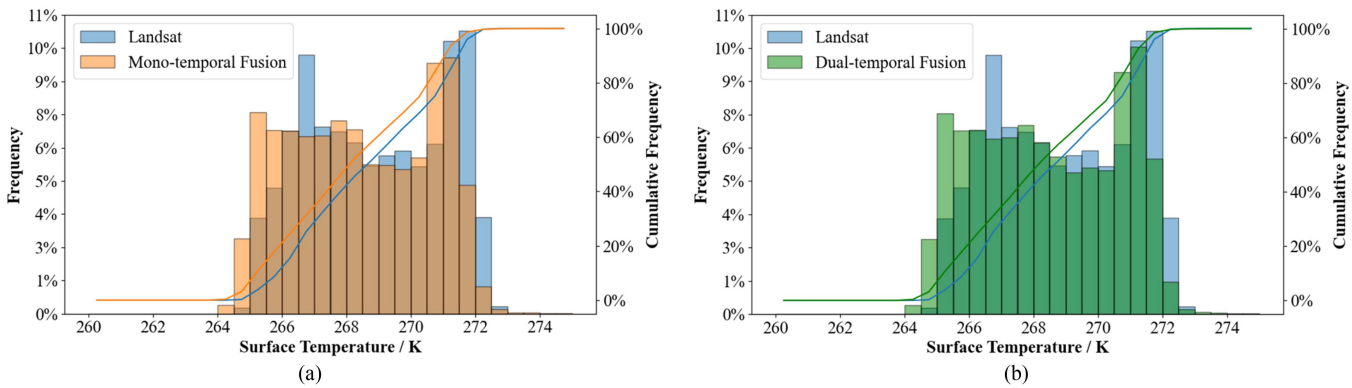

Fig. 7. Frequency distribution of Landsat surface temperature and fusion data. (a) Mono-temporal fusion data. (b) Dual-temporal fusion data.

exhibits minor variations across the entire region. The temperature difference in the heterogeneous edge region displays apparent deviations, revealing intricate details in the reference data. Specifically, the reference images include features not fully captured by MODIS IST at the target date, resulting in fusion differences in temperature.

2) Case 2 for Island Zone: This part used a region in the GrIS as the study region. Fig. 6(a)–(d) shows the Landsat IST and MOD21 image pairs, acquired on April 9, 2020 and October 9, 2020, respectively. Fig. 6(f) shows the MOD21 of the August 15, 2020 target date. Fig. 6(g) and (h) are the fused data of the target date. The Landsat 8 IST in Fig. 6(e) was used to verify the accuracy of the fusion results.

In case 2, the study area is snow and ice covered all year round. Thus, the IST does not vary much between different dates. However, the surface temperature is higher on the predicted date (summer) than the reference dates' spring and autumn phases. The IST is still below 273 K in most regions, with no major ice melting. The mono- and dual-temporal STNLFFM fusion results

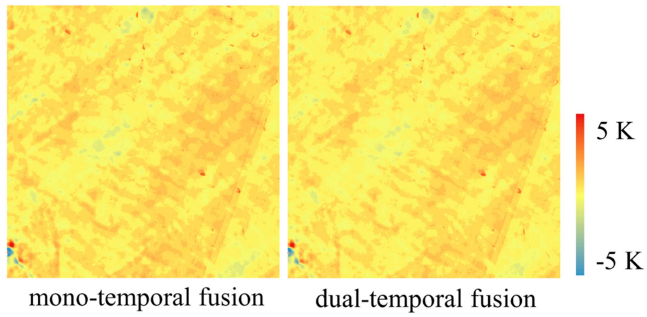

Fig. 8. Difference of IST (Landsat IST minus the fusion data) for the case 2.

and Landsat IST are consistent in the overall spatial distribution. Table VII gives that the R^2 values of the mono- and dual-temporal fusion results are 0.958 and 0.965, respectively, and the RMSE values are 0.658 and 0.606 K, which suggest that the fusion results are satisfactory.

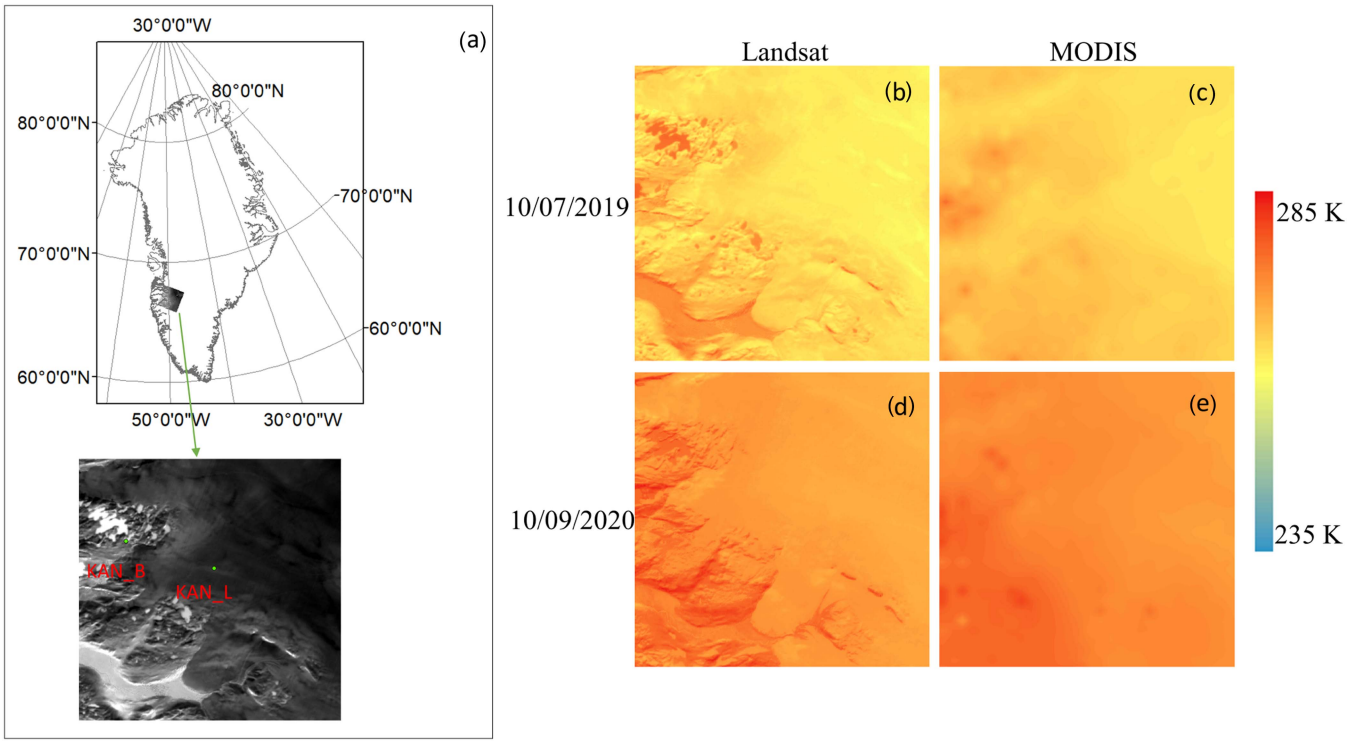

Fig. 9. (a) Study area, which includes two sites (KAN_B, KAN_L). (b)–(e) Reference Landsat-MODIS pairs on October 9, 2019 and October 9, 2020.

TABLE VII
QUANTITATIVE EVALUATION VALUES

Case 2 (unit: K)		
	STNLFFM (Mono)	STNLFFM (Dual)
R^2	0.958	**0.965**
RMSR	0.658	**0.606**
bias	0.482	**0.439**

The bold entities represent the optimal performance among the quantitative evaluation of the same indicator.

Fig. 7 shows the histograms of Landsat and fused IST from the areas of interest. The temperatures range between 264 and 273 K, and the fusion results are consistent with the Landsat. This scene has a slight banding effect, as shown in Fig. 8. However, the temperature difference between Landsat IST and fusion IST is relatively subtle, floating around 1 K. Comparing the two cases showed that the fusion results are better for the homogeneous surface cover type scenario.

C. Application for Long-Term Fusion

The STNLFFM dual-temporal fusion algorithm has high accuracy in fusing IST data in Greenland, as analyzed in Section IV-A and IV-B. To validate the applicability of this spatiotemporal fusion method further, a long-term series of experiments is designed for the southwestern region of Greenland [see Fig. 9(a)]. It is delineated at one year to cycle the IST change. The STNLFFM algorithm fuses data from two reference dates, October 7, 2019 and October 9, 2020 [see Fig. 9(b)–(e)], with the time interval for the output data being about 30 days. All input and output remote sensing data cover an area of 30 km × 30 km, with the number of image pixels being 1000 × 1000.

Fig. 10 shows that the fusion data produced by the STNLFFM algorithm shows crisper features than MOD21, with the edge visible. From August 2019 to August 2020, the fusion data in the region underwent a process of decreasing, leveling off, and finally increasing. This temperature change pattern is consistent with the temperature change pattern of the GrIS.

For the validation of long-term fusion data, we conducted a comparison between the fusion IST and the measurements from two PROMICE sites, as shown in Fig. 11 and Table VIII. This process generated a dataset consisting of 18 different dates. Our analysis reveals that KAN_B, located in the southwestern region of the ice sheet, exhibited visible rock features. From October 2019 to March 2020, the surface temperature is below 0°C. In summer, it could be above freezing. The RMSE between KAN_B measurements and MOD21 IST is 2.210 K, and the bias is 1.340 K. The spatiotemporal fusion data indicate improved results, with an RMSE of 1.383 K and a bias of 0.875 K. KAN_L is located inland in Greenland, with an average altitude of 670 m. The accuracy of the fusion data is slightly lower than that of MODIS in KAN_L site, but they show consistent change characteristics.

The results obtained from this long-term series experiment have further validated the applicability of the spatiotemporal fusion algorithm in producing highly spatial and temporal IST in Greenland. The high resolution and well-defined features have facilitated the identification of melting boundaries and the accurate capture of the spatiotemporal variations of the surface temperature for the GrIS.

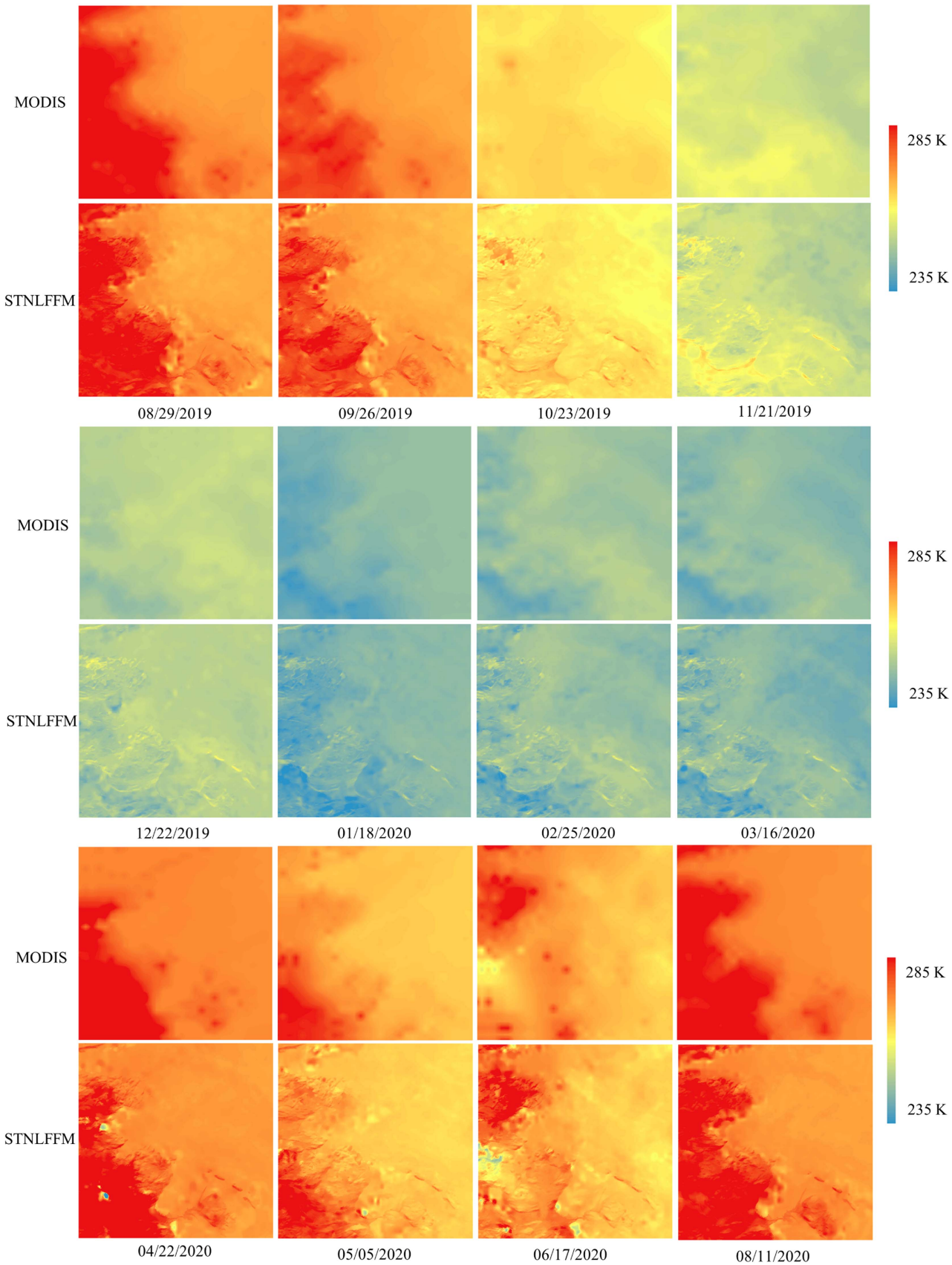

Fig. 10. Fusion data from 2019 to 2020 fused using the dual-temporal fusion algorithm of STNLFFM, compared with MOD21 data.

TABLE VIII
COMPARISONS BETWEEN AWSS IST AND MOD21, AWSS IST AND FUSION IST THE BIAS REPRESENTS MODIS/FUSION IST MINUS AWSS ST

	KAN_B				KAN_L			
	RMSE	MAE	R^2	bias	RMSE	MAE	R^2	bias
MODIS	2.210	1.707	0.998	1.340	2.015	1.518	0.983	−1.438
Dual-temporal Fusion	1.383	1.203	0.983	0.857	2.869	2.627	0.989	−2.627

Unit: K

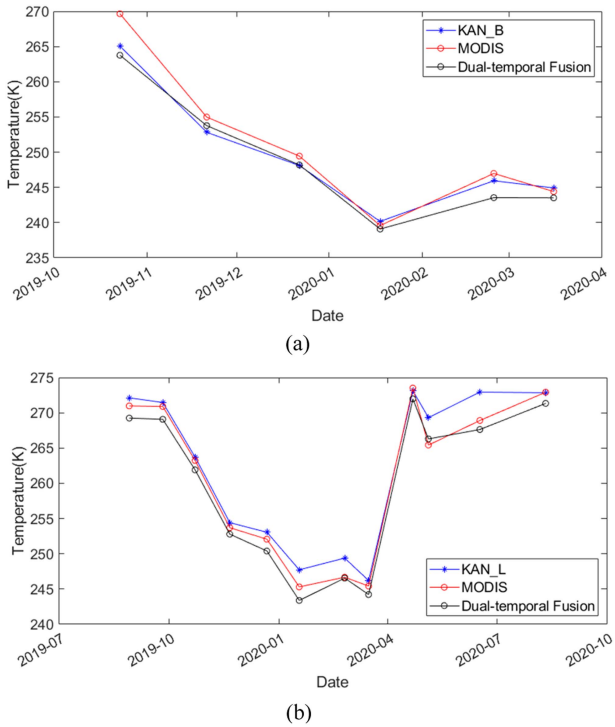

Fig. 11. Surface temperature comparison from AWSs, MOD21, and dual-temporal fusion data. (a) KAN_B. (b) KAN_L.

V. CONCLUSION

Greenland ice sheet remote sensing data showed prolonged cloud cover and low spatial or temporal resolution data. To solve the problem, this article evaluated the applicability of several spatiotemporal fusion methods on the ice/snow surface temperature data of GrIS. In the first part, we use the surface temperature data from AWS as a reference for quantitatively evaluating the fusion results of different methods. In the second part, based on the quantitative evaluation results, we selected mono- and dual-temporal STNLFFM fusion methods to test two cases of different land types. In the third part, to investigate further the applicability of the STNLFFM algorithm to IST data in GrIS, we conducted a group of fusion experiments for one year.

The results of the above experiments are as follows: first, five spatiotemporal fusion methods obtained good performances with all R^2 values above 0.9, whereas the results of the ESTARFM method differ larger from the AWS measurements. The dual-temporal STNLFFM method is more accurate than the other methods for the fusion of IST data. Second, in the heterogeneous marginal ice zone, the dual-temporal STNLFFM method can accurately predict the changes in the spatial distribution of IST. The mono-temporal fusion results introduce some noise compared with the original Landsat 8 IST data. Third, in the Greenland ice cap regions, the mono- and dual-temporal STNLFFM fusion methods have accurate results, with RMSE values of 0.658 and 0.606 K. The two dual-temporal methods exhibit remarkable superiority, attributed to their reliance on more comprehensive reference data, enabling accurate capture of temperature variations in long temporal studies. Furthermore, IST spatiotemporal fusion accuracy generally surpasses the existing LST fusion results. That might be attributed to the relatively stable and homogeneous feature types of snow and ice.

This article explores the potential of spatiotemporal fusion methods for fusing ice and snow surface temperatures, solving the low frequency problem for Landsat 8 IST data in GrIS. The spatiotemporal fusion technique can generate IST datasets for GrIS with high spatial and temporal resolutions. It provides an effective way for the detailed observation and study of Greenland ice sheet. There is still room for improvements of our study. Compared to other surface types, ice/snow surfaces exhibit less pronounced variations and higher homogeneity, resulting in superior spatiotemporal fusion accuracy. Additionally, ice/snow surfaces are highly reflective, leading to reduced absorption and surface temperature. In future work, the fusion algorithms based on the characteristics of ice and snow surface temperature could be developed.

REFERENCES

[1] J. E. J. Reeves Eyre and X. Zeng, "Evaluation of Greenland near surface air temperature datasets," *Cryosphere*, vol. 11, no. 4, pp. 1591–1605, Jul. 2017.
[2] A. C. Adolph, M. R. Albert, and D. K. Hall, "Near-surface temperature inversion during summer at Summit, Greenland, and its relation to MODIS-derived surface temperatures," *Cryosphere*, vol. 12, no. 3, pp. 907–920, Mar. 2018.
[3] J. Oerlemans, "The mass balance of the Greenland ice sheet: Sensitivity to climate change as revealed by energy-balance modelling," *Holocene*, vol. 1, no. 1, pp. 40–48, Mar. 1991.
[4] I. Janssens and P. Huybrechts, "The treatment of meltwater retention in mass-balance parameterizations of the Greenland ice sheet," *Ann. Glaciol.*, vol. 31, pp. 133–140, Jan. 2000.
[5] T. J. Young et al., "Rapid basal melting of the Greenland ice sheet from surface meltwater drainage," *Proc. Nat. Acad. Sci. United States Amer.*, vol. 119, no. 10, Mar. 2022, Art. no. e2116036119.
[6] J. E. Walsh et al., "Extreme weather and climate events in northern areas: A review," *Earth-Sci. Rev.*, vol. 209, Oct. 2020, Art. no. 103324.
[7] D. Liang, H. Guo, Q. Cheng, L. Zhang, and L. Kong, "Correlation and interaction between temperature and freeze-thaw in representative regions of Antarctica," *Int. J. Digit. Earth*, vol. 15, no. 1, pp. 2296–2318, Dec. 2022, doi: 10.1080/17538947.2022.2158242.
[8] D. Liang et al., "Analyzing Antarctic ice sheet snowmelt with dynamic big earth data," *Int. J. Digit. Earth*, vol. 14, no. 1, pp. 88–105, Jan. 2021.
[9] Y. Qian et al., "Retrieval of surface temperature and emissivity from ground-based time-series thermal infrared data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 4, pp. 284–292, Jan. 2020, doi: 10.1109/JSTARS.2019.2959794.
[10] B. Chen, B. Huang, and B. Xu, "Comparison of spatiotemporal fusion models: A review," *Remote Sens.*, vol. 7, no. 2, pp. 1798–1835, Feb. 2015.

[11] H. K. Zhang, B. Huang, M. Zhang, K. Cao, and L. Yu, "A generalization of spatial and temporal fusion methods for remotely sensed surface parameters," *Int. J. Remote Sens.*, vol. 36, no. 17, pp. 4411–4445, Sep. 2015.

[12] H. Xia, Y. Chen, A. Gong, K. Li, L. Liang, and Z. Guo, "Modeling daily temperatures via a phenology-based annual temperature cycle model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 4, pp. 6219–6229, Jun. 2021, doi: 10.1109/JSTARS.2021.3085342.

[13] J. Key, X. Wang, Y. Liu, R. Dworak, and A. Letterly, "The AVHRR polar pathfinder climate data records," *Remote Sens.*, vol. 8, no. 3, Feb. 2016, Art. no. 167.

[14] Z. Wan, Y. Zhang, Q. Zhang, and Z. - L. Li, "Quality assessment and validation of the MODIS global land surface temperature," *Int. J. Remote Sens.*, vol. 25, no. 1, pp. 261–274, Jan. 2004.

[15] N. K. Malakar, G. C. Hulley, S. J. Hook, K. Laraby, M. Cook, and J. R. Schott, "An operational land surface temperature product for Landsat thermal data: Methodology and validation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5717–5735, Oct. 2018.

[16] A. Gillespie, S. Rokugawa, T. Matsunaga, J. S. Cothern, S. Hook, and A. B. Kahle, "A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1113–1126, Jul. 1998, doi: 10.1109/36.700995.

[17] I. F. Trigo, I. T. Monteiro, F. Olesen, and E. Kabsch, "An assessment of remotely sensed land surface temperature," *J. Geophys. Res., Atmos.*, vol. 113, Sep. 2008, Art. no. D17108.

[18] D. Liang, H. Guo, L. Zhang, H. Li, and X. Wang, "Sentinel-1 EW mode dataset for Antarctica from 2014–2020 produced by the CASEarth cloud service platform," *Big Earth Data*, vol. 6, no. 4, pp. 385–400, Oct. 2022.

[19] X. Zhu, F. Cai, J. Tian, and T. K. - A. Williams, "Spatiotemporal fusion of multisource remote sensing data: Literature survey, taxonomy, principles, applications, and future directions," *Remote Sens.*, vol. 10, no. 4, Mar. 2018, Art. no. 527.

[20] M. Belgiu and A. Stein, "Spatiotemporal image fusion in remote sensing," *Remote Sens.*, vol. 11, no. 7, Apr. 2019, Art. no. 818.

[21] F. Gao, J. Masek, M. Schwaller, and F. Hall, "On the blending of the Landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2207–2218, Aug. 2006, doi: 10.1109/TGRS.2006.872081.

[22] X. Zhu, J. Chen, F. Gao, X. Chen, and J. G. Masek, "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions," *Remote Sens. Environ.*, vol. 114, no. 11, pp. 2610–2623, Nov. 2010.

[23] C. M. Gevaert and F. J. García-Haro, "A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion," *Remote Sens. Environ.*, vol. 156, pp. 34–44, Jan. 2015.

[24] Q. Cheng, H. Liu, H. Shen, P. Wu, and L. Zhang, "A spatial and temporal nonlocal filter-based data fusion method," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4476–4488, Aug. 2017, doi: 10.1109/TGRS.2017.2692802.

[25] Q. Wang and P. M. Atkinson, "Spatio-temporal fusion for daily Sentinel-2 images," *Remote Sens. Environ.*, vol. 204, pp. 31–42, Jan. 2018.

[26] B. Huang and H. Song, "Spatiotemporal reflectance fusion via sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3707–3716, Oct. 2012, doi: 10.1109/TGRS.2012.2186638.

[27] V. Moosavi, A. Talebi, M. H. Mokhtari, S. R. F. Shamsi, and Y. Niazi, "A wavelet-artificial intelligence fusion approach (WAIFA) for blending Landsat and MODIS surface temperature," *Remote Sens. Environ.*, vol. 169, pp. 243–254, Nov. 2015.

[28] X. Zhu, E. H. Helmer, F. Gao, D. Liu, J. Chen, and M. A. Lefsky, "A flexible spatiotemporal method for fusing satellite images with different resolutions," *Remote Sens. Environ.*, vol. 172, pp. 165–177, Jan. 2016.

[29] L. Lin, Y. Shen, J. Wu, and F. Nan, "CAFE: A cross-attention based adaptive weighting fusion network for MODIS and landsat spatiotemporal fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 5001605, doi: 10.1109/LGRS.2023.3286463.

[30] S. Chen, J. Wang, and P. Gong, "ROBOT: A spatiotemporal fusion model toward seamless data cube for global remote sensing applications," *Remote Sens. Environ.*, vol. 294, Aug. 2023, Art. no. 113616.

[31] J. Wu, L. Lin, T. Li, Q. Cheng, C. Zhang, and H. Shen, "Fusing Landsat 8 and Sentinel-2 data for 10-m dense time-series imagery using a degradation-term constrained deep network," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 108, Apr. 2022, Art. no. 102738.

[32] Y. Yang et al., "Impact of tile drainage on evapotranspiration in South Dakota, USA, based on high spatiotemporal resolution evapotranspiration time series from a multisatellite data fusion system," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, pp. 2550–2564, Jun. 2017, doi: 10.1109/JSTARS.2017.2680411.

[33] B. Wang et al., "Generating spatiotemporally consistent fractional vegetation cover at different scales using spatiotemporal fusion and multiresolution tree methods," *ISPRS J. Photogramm. Remote Sens.*, vol. 167, pp. 214–229, Sep. 2020.

[34] L. Han et al., "Using spatiotemporal fusion algorithms to fill in potentially absent satellite images for calculating soil salinity: A feasibility study," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 111, Jul. 2022, Art. no. 102839.

[35] X. Li et al., "Monitoring high spatiotemporal water dynamics by fusing MODIS, Landsat, water occurrence data and DEM," *Remote Sens. Environ.*, vol. 265, Nov. 2021, Art. no. 112680.

[36] Y. Zhao, B. Huang, and H. Song, "A robust adaptive spatial and temporal image fusion model for complex land surface changes," *Remote Sens. Environ.*, vol. 208, pp. 42–62, Apr. 2018.

[37] Y. Zheng, B. Wu, M. Zhang, and H. Zeng, "Crop phenology detection using high spatio-temporal resolution data fused from SPOT5 and MODIS products," *Sensors*, vol. 16, no. 12, Dec. 2016, Art. no. 2099.

[38] J. Jiang et al., "HISTIF: A new spatiotemporal image fusion method for high-resolution monitoring of crops at the subfield level," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4607–4626, 2020, doi: 10.1109/JSTARS.2020.3016135.

[39] Y. Shen, G. Shen, H. Zhai, C. Yang, and K. Qi, "A Gaussian kernel-based spatiotemporal fusion model for agricultural remote sensing monitoring," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3533–3545, 2021, doi: 10.1109/JSTARS.2021.3066055.

[40] J. Adrian, V. Sagan, and M. Maimaitijiang, "Sentinel SAR-optical fusion for crop type mapping using deep learning and Google earth engine," *ISPRS J. Photogramm. Remote Sens.*, vol. 175, pp. 215–235, May 2021.

[41] H. Nagendra et al., "Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats," *Ecol. Indicators*, vol. 33, pp. 45–59, Oct. 2013.

[42] N. Pettorelli et al., "Satellite remote sensing of ecosystem functions: Opportunities, challenges and way forward," *Remote Sens. Ecol. Conservation*, vol. 4, no. 2, pp. 71–93, Jun. 2018.

[43] Q. Weng, P. Fu, and F. Gao, "Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data," *Remote Sens. Environ.*, vol. 145, pp. 55–67, Apr. 2014.

[44] Y. Li et al., "Fusing Sentinel-2 and Landsat-8 surface reflectance data via pixel-wise local normalization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 7359–7374, 2022, doi: 10.1109/JSTARS.2022.3200713.

[45] Z. Yin et al., "Spatiotemporal fusion of land surface temperature based on a convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1808–1822, Feb. 2021, doi: 10.1109/TGRS.2020.2999943.

[46] P. Wu, H. Shen, L. Zhang, and F. - M. Göttsche, "Integrated fusion of multi-scale polar-orbiting and geostationary satellite observations for the mapping of high spatial and temporal resolution land surface temperature," *Remote Sens. Environ.*, vol. 156, pp. 169–181, Jan. 2015.

[47] P. Wu et al., "Spatially continuous and high-resolution land surface temperature product generation: A review of reconstruction and spatiotemporal fusion techniques," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 3, pp. 112–137, Sep. 2021, doi: 10.1109/MGRS.2021.3050782.

[48] H. Shen, L. Huang, L. Zhang, P. Wu, and C. Zeng, "Long-term and fine-scale satellite monitoring of the urban heat island effect by the fusion of multi-temporal and multi-sensor remote sensed data: A 26-year case study of the city of Wuhan in China," *Remote Sens. Environ.*, vol. 172, pp. 109–125, Jan. 2016.

[49] Y. Chang, J. Xiao, X. Li, D. Zhou, and Y. Wu, "Combining GOES-R and ECOSTRESS land surface temperature data to investigate diurnal variations of surface urban heat island," *Sci. Total Environ.*, vol. 823, Jun. 2022, Art. no. 153652.

[50] B. Huang, J. Wang, H. Song, D. Fu, and K. Wong, "Generating high spatiotemporal resolution land surface temperature for urban heat island monitoring," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1011–1015, Sep. 2013, doi: 10.1109/LGRS.2012.2227930.

[51] J. E. Overland, "Rare events in the arctic," *Climate Change*, vol. 168, no. 3-4, Oct. 2021, Art. no. 27.

[52] M. Wu, H. Li, W. Huang, Z. Niu, and C. Wang, "Generating daily high spatial land surface temperatures by combining ASTER and MODIS land surface temperature products for environmental process monitoring," *Environ. Sci., Processes Impacts*, vol. 17, pp. 1396–1404, 2015.

[53] H. Yang et al., "Measuring the urban land surface temperature variations under Zhengzhou city expansion using Landsat-like data," *Remote Sens.*, vol. 12, Mar. 2020, Art. no. 801.
[54] Y. Shen, H. Shen, Q. Cheng, and L. Zhang, "Generating comparable and fine-scale time series of summer land surface temperature for thermal environment monitoring," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2136–2147, 2021, doi: 10.1109/JSTARS.2020.3046755.
[55] G. Hulley, R. Freepartner, N. Malakar, and S. Sarkar, "Moderate resolution imaging spectroradiometer (MODIS) land surface temperature and emissivity product (MxD21) user guide collection-6∗," Mar, 2019. [Online]. Available: https://dataverse.jpl.nasa.gov/dataset.xhtml?persistentId=hdl:2014/45681&version=1.0
[56] V. García-Santos, J. Cuxart, D. Martínez-Villagrasa, M. A. Jiménez, and G. Simó, "Comparison of three methods for estimating land surface temperature from Landsat 8-TIRS sensor data," *Remote Sens.*, vol. 10, no. 9, Sep. 2018, Art. no. 1450.
[57] J. Guo, H. Ren, Y. Zheng, S. Lu, and J. Dong, "Evaluation of land surface temperature retrieval from Landsat 8/TIRS images before and after stray light correction using the SURFRAD dataset," *Remote Sens.*, vol. 12, no. 6, Mar. 2020, Art. no. 1023.
[58] G. C. Hulley, C. G. Hughes, and S. J. Hook, "Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data: LST&E UNCERTAINTIES," *J. Geophys. Res. Atmospheres*, vol. 117, Dec. 2012, Art. no. D23113.
[59] Z. Fang, N. Wang, Y. Wu, and Y. Zhang, "Greenland-ice-sheet surface temperature and melt extent from 2000 to 2020 and implications for mass balance," *Remote Sens.*, vol. 15, no. 4, Feb. 2023, Art. no. 1149.
[60] S. Liu, J. Jiang, Q. Sun, J. Wan, and H. Sheng, "Assessment of the Greenland ice sheet change (2011–2021) derived from CryoSat-2," *Polar Sci.*, vol. 36, Jun. 2023, Art. no. 100940.
[61] D. Houtz, C. Mätzler, R. Naderpour, M. Schwank, and K. Steffen, "Quantifying surface melt and liquid water on the Greenland ice sheet using l-band radiometry," *Remote Sens. Environ.*, vol. 256, Apr. 2021, Art. no. 112341.
[62] B. Yang, S. Liang, H. Huang, and X. Li, "An elevation change dataset in Greenland ice sheet from 2003 to 2020 using satellite altimetry data," *Big Earth Data*, vol. 9, pp. 1–18, Sep. 2022.
[63] P. Halas, J. Mouginot, B. De Fleurian, and P. M. Langebroek, "Impact of seasonal fluctuations of ice velocity on decadal trends observed in Southwest Greenland," *Remote Sens. Environ.*, vol. 285, Feb. 2023, Art. no. 113419.
[64] K. Yang et al., "Seasonal evolution of supraglacial lakes and rivers on the Southwest Greenland ice sheet," *J. Glaciol.*, vol. 67, no. 264, pp. 592–602, Aug. 2021.

Qing Cheng received the B.S. degree in the geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2010 and 2015, respectively.

She is currently a Professor with the School of Computer Science, China University of Geoscience, Wuhan, China. Her research interests include polar remote sensing, remote sensing data reconstruction, and data fusion.

Zejun Zhang received the B.S. degree in software engineering from the Taiyuan University of Technology, Taiyuan, China, in 2021. She is currently working toward the M.S. degree in computer technology with the School of Computer Science, China University of Geoscience, Wuhan, China.

Her research interests include remote sensing data reconstruction and data fusion.

Dong Liang received the B.Sc. degree in applied mathematics from the University of Hull, Hull, U.K., in 2006, the M.Sc. degree in applied mathematics from Mälardalen University, Västerås, Sweden, in 2009, and the Ph.D. degree in cartography and geographic information system from the University of Chinese Academy of Sciences, Beijing, China, in 2022.

He is currently the Assistant Director-General of the International Research Center of Big Data for Sustainable Development Goals, the Deputy Secretary-General of the Digital Belt and Road Program, the Deputy Secretary-General of the Chinese National Committee of the International Society for Digital Earth, and an Editorial Board Member of Big Earth Data. His research interests include polar remote sensing and Big Earth Data applications.

Fan Ye received the M.S. degree in surveying and mapping engineering from Fuzhou University, Fuzhou, China, in 2021. She is currently working toward the Ph.D. degree in geoscience information engineering with the School of Computer Science, China University of Geoscience, Wuhan, China.

Her research interests include polar remote sensing, and remote sensing data reconstruction and application.