Estimating subpixel turbulent heat flux over leads from MODIS thermal infrared imagery with deep learning

Zhixiang Yin [1,2,3], Xiaodong Li [1], Yong Ge [4], Cheng Shang [1,2], Xinyan Li [1,2], Yun Du [1], Feng Ling [1]

[1]Key Laboratory for Environment and Disaster Monitoring and Evaluation, Hubei, Innovation Academy for Precision
Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, 430077, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Anhui Province Key Laboratory of Wetland Ecosystem Protection and Restoration, Anhui University, Hefei, 230601, China
[4]State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences & Natural
Resources Research, Chinese Academy of Sciences, Beijing 100101, China

Correspondence to: F. Ling (lingf@whigg.ac.cn)

Abstract. Turbulent heat flux (THF) over leads is an important variable used for monitoring climate change in the Arctic. Presently, THF over leads is often calculated from satellite imagery. The accuracy of the estimated THF is low for mixed pixels that consist of ice and leads, because the mixed pixels along lead boundaries will lower the accuracy of the surface temperature measured over leads and the corresponding lead map. To address this problem, a deep residual convolutional neural network (CNN)-based framework is proposed to estimate THF over leads at the subpixel scale (DeepSTHF) with remotely sensed imagery. The DeepSTHF allows the production of a sea surface temperature (SST) image and a corresponding lead map with a finer spatial resolution than the input SST image using two CNNs, so that the subpixel scale THF can be estimated from them. The proposed approach is assessed using simulated and real MODIS imagery and compared against the conventional bicubic interpolation and pixel-based methods. The results demonstrate that the proposed CNN-based method can effectively estimate subpixel-scale information from the coarse data and performs well in producing fine spatial resolution SST images and lead maps, thereby allowing researchers to obtain more accurate and reliable THF over leads.

1. Introduction

Leads form as a linear area of open water and thin floating ice within a closed pack ice (Willmes and Heinemann, 2015). They develop as the result of various forces, such as thermal stress and wave action (Tschudi et al., 2002). Through leads, the sea surface contacts the atmosphere allowing a direct exchange of sensible and latent heat flux (Marcq and Weiss, 2012). Even though leads only cover a relatively small part of the total sea ice area in the polar regions, they serve as the primary window for turbulent heat flux (THF) because the sea ice itself significantly reduces any air–sea interaction (Maykut, 1978). In the central Arctic, leads comprise no more than 1% of sea area during winter, but provide a channel for more than 70% of the upward heat flux (Marcq and Weiss, 2012). Additionally, it has been revealed that small changes in leads would cause a considerable change in temperature near the surface (Lüpkes et al., 2008). Consequently, the estimation of THF over leads is a crucial part of climate studies (Maykut, 1978;Ebert and Curry, 1993).

Remotely sensed satellite images have become a promising data source that is often used to estimate the THF occurring over leads (Qu et al., 2019), because on-site measurement is always difficult given the harsh weather conditions typical of polar regions. To calculate the THF over leads with remotely sensed data, both a lead map and associated sea surface

35 temperature (SST) are required. A lead map could be produced from various satellite data, including visible, thermal, and microwave imagery (Lindsay and Rothrock, 1995;Röhrs and Kaleschke, 2012;Willmes and Heinemann, 2015). Once a lead map has been generated, the corresponding SST image that is needed to estimate the THF that is often derived from Thermal Infrared (TIR) bands of satellite data.

Based on the foregoing discussion, satellite TIR imagery is essential for the estimation of the THF occurring over leads

40 because it can be used to generate a lead map and the associated SST imagery simultaneously. Thermal infrared images can be obtained from Landsat-8 and MODIS products. Landsat-8 TIR imagery has a spatial resolution of 100 m. However, the Landsat-8 satellite has an approximately 16-day revisit cycle, making it a challenge to estimate THF with appropriate timing. In contrast, MODIS has a daily repeat frequency, which is attractive for research that focuses on the rapid variation of THF over leads. However, MODIS imagery has a spatial resolution of 1 km, and unusually includes mixed pixels. The MODIS

45 mixed pixel problem not only strongly affects the extraction of a lead map, but also affects the estimated surface temperature of leads. Therefore, the mixed pixel problem of MODIS imagery will result in a large error when calculating THF.

The mixed pixel problem of MODIS TIR imagery that affects THF estimation can be solved through conducting a subpixel analysis (Ge et al., 2009;Atkinson, 2013;Ge et al., 2019;Foody and Doan, 2007;Wang et al., 2014;Zhong and Zhang, 2013). Image super-resolution (SR), which aims to enhance the spatial resolution of images, is a representative subpixel scale analysis

50 technique that has been widely adopted in a variety of applications (Zhang, 2006;Ling et al., 2010;Leach and Sherlock, 2013;Foody et al., 2005). A wide range of approaches have been proposed for images with SR (Wang et al., 2020;Glasner et al., 2009). Among them, convolutional neural network (CNN)-based methods have provided significantly improved performance in producing SR imagery, because they have a powerful ability to model the latent nonlinear relationship between fine spatial resolution image and the corresponding coarse spatial resolution one through a large amount of training data (Dong

55 et al., 2014;Ledig et al., 2017;Ling et al., 2019;Jia et al., 2019;Ling and Foody, 2019). A CNN-based SR approach has a great potential in the downscaling of imagery such as fine spatial resolution SST imagery and lead maps that are useful in THF estimation. To the best of our knowledge, studies using CNN in the estimation of the THF occurring over leads have not yet been conducted.

This study proposes a CNN-based framework used to estimate the THF at a subpixel scale from MODIS TIR imagery

60 (hereinafter DeepSTHF). Specifically, a unified framework comprising two CNNs that simultaneously produce Landsat-like SST imagery and a corresponding binary map of leads, which are used in the estimation of the THF over leads, from the MODIS SST imagery, is proposed. The generated Landsat-like SST imagery and lead map are then employed to estimate the THF with an aerodynamic bulk formula (Marcq and Weiss, 2012;Qu et al., 2019). This study will provide a new perspective for solving the mixed pixel problem in the estimation of THF from remotely sensed imagery by extracting subpixel level

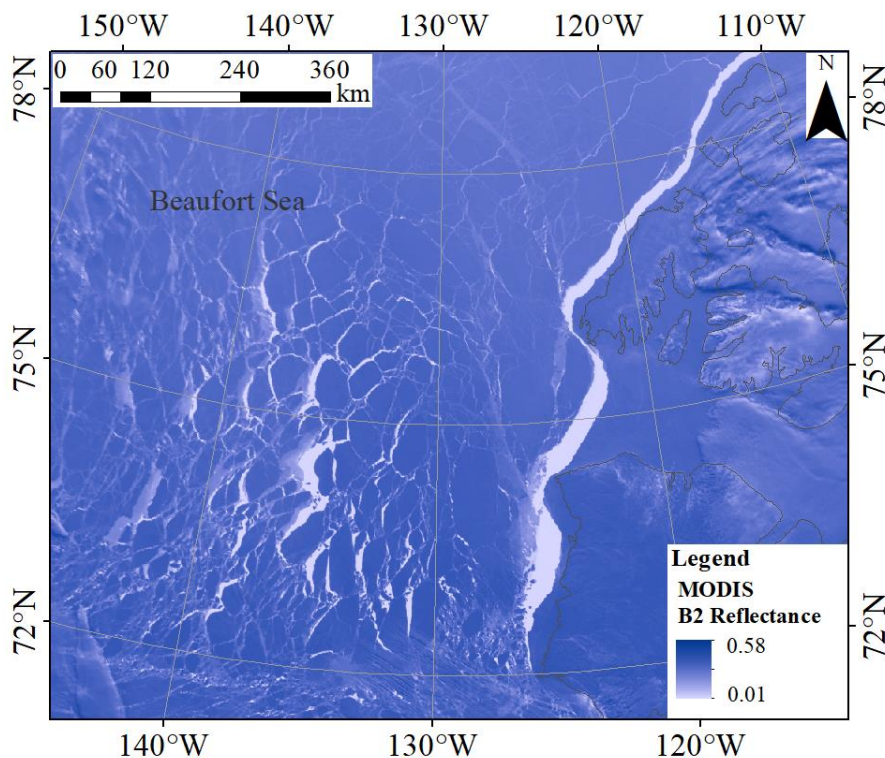
65 information.

2. Study area and data

2.1. Study area

The Beaufort Sea, a marginal sea of the Arctic Ocean situated north of Canada and Alaska, was selected as the study area (Fig. 1). The typically severe climate of the Beaufort Sea keeps the surface frozen most of the year. Multiple forces such as the anti-
70 cyclonic motion of the Beaufort Gyre cause the ice pack to fracture (Lewis and Hutchings, 2019) and linear cracks (leads) to form. Recently, global warming is causing the multi-year ice pack to shrink rapidly (Barber et al., 2014); as a consequence, the size and spatial extent of leads are increasing in the Beaufort Sea. Various floes along with leads of varying widths and lengths now occur in this region.

2.2. Datasets and preprocessing

75 The proposed method, DeepSTHF, used MODIS SST images as the input data to calculate the THF at a subpixel scale. Remotely sensed images derived from the Landsat-8 Operational Land Imager (OLI) were used as fine-resolution data to train the CNN models and assess the accuracy of the result in the experiments. Additionally, associated meteorological data (i.e., wind speed, air temperature, and dew point temperature) were obtained as well to estimate the THF occurring over leads.

80 **Figure 1. Location of the study area. Background image is the band 2 (B2) reflectance of a Moderate Resolution Imaging Spectroradiometer (MODIS) image acquired on 25 April 2015. The black area generally represents leads.**

2.2.1. MODIS data

The MODIS Level-1B product MOD021KM, acquired by the sensor aboard Terra satellite, was used in this research, and it
85 was obtained from the US National Aeronautics and Space Administration's Level 1 and Atmosphere Archive and Distribution
System Distributed Active Archive Center (https://ladsweb.modaps.eosdis.nasa.gov/). The MOD021KM data set was stored
in the Hierarchical Data Format-Earth Observing System swath structure, which was designed to support the data archiving
and storage needs of the Earth Observing System. It is mainly comprised of data and geolocation fields. The data file contained
36 calibrated and geolocated spectral bands from the optical to TIR wavelength regions and have a spatial resolution of 1 km
90 comprised of 1354 by 2030 pixels. Cloud pixels in the MOD021KM were identified with MODIS cloud mask product
(MOD35_L2) and pixels with a zenith angle > 25 ° were not used to reduce the panoramic bowtie effect (Eythorsson et al.,
2019). The TIR bands 31 and 32 which are respectively centered on 11.03 and 12.02 μm, were used to retrieve the SST using
a split-window algorithm (Hall and Riggs, 2001) that is adapted for use with MODIS data, whose accuracy was reported within
2 ℃. Additionally, longitude and latitude coordinates, which were provided in the geolocation field and given at a 5 km
95 resolution, were used to tie the swath to points on the Earth's surface. We collected MODIS images with cloud cover of < 10%
during March–May from 2013 to 2020, because leads in these three months were abundant with a variety of sizes and shapes
from visual inspection.

2.2.2. Landsat-8 data

This study used the Landsat-8 Level 1 terrain-corrected (L1T) data product, which was acquired from the United States
100 Geological Survey Earth Explorer website (http://earthexplorer.usgs.gov/). Because the Landsat-8 data product includes
images acquired from May 2013 to present, here we selected scenes with < 10% cloud cover acquired during 2013–2020.
Likewise, only images acquired during March–May in each year were obtained.

These Landsat-8 data were used to produce fine-resolution SST reference images and lead maps. The split-window
algorithm that was developed for the Landsat-8 data (Du et al., 2015), which is suitable for various surface types including ice
105 and water, was employed to retrieve SST data. The algorithm estimates temperature from two thermal infrared bands of
Landsat-8 data and has an accuracy of better than 1.0 ℃ (Du et al., 2015). Note that the Landsat-8 TIR Sensor had an issue
with stray light, which refers to unwanted radiance from outside the field-of-view entering the optical system (Montanaro et
al., 2015). Nevertheless, corrections have been applied in the current Landsat-8 LIT data product and the stray light artifact
has been reduced (Gerace and Montanaro, 2017). However, this artifact was amplified and obvious in the generated SST image
110 (Fig. 2a), which could impact the estimation of THF. A median filtering method (Eppler and Full, 1992;Qu et al., 2019) was
then used to further remove the noise in the Landsat-8 SST image caused by this type of artifact (Fig. 2b). The reference lead
maps were manually drawn by referring to Landsat-8 OLI spectral bands; an example is shown in Fig. 2c.

For the obtained Landsat-8 data product, both the OLI spectral bands and the TIRS bands have a pixel size of 30 m. Considering the fact that TIRS bands at 30 m resolution were up-sampled from 100 m raw data by cubic interpolation to match
115 the OLI spectral bands, the TIRS bands were resampled to 100 m by average strategy to retain the original spatial resolution.

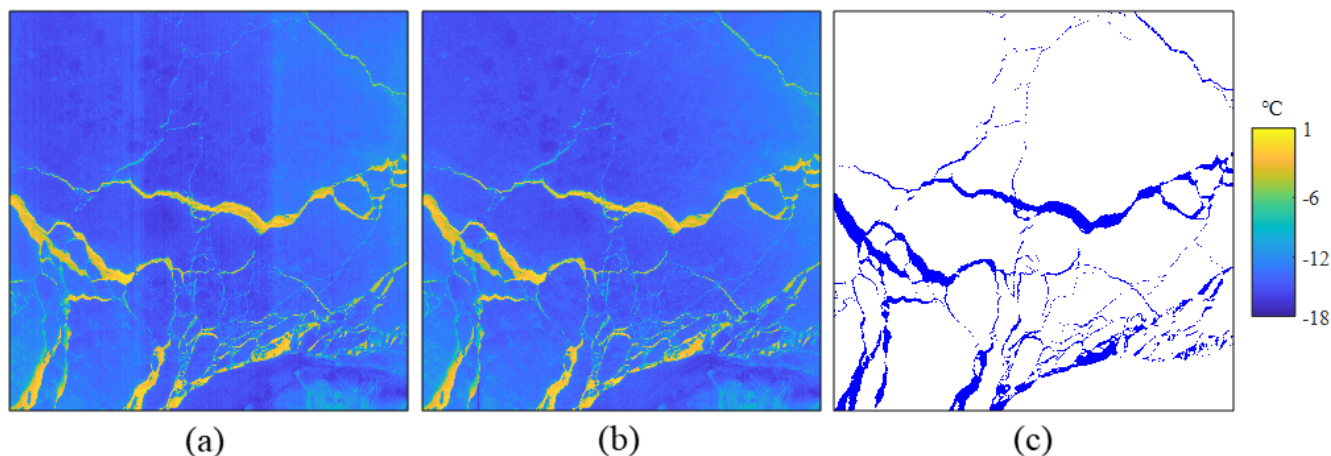



Figure 2. Sea surface temperature (SST) images derived from Landsat-8 imagery and corresponding corrected SST image on 31 March 2020: (a) original SST image, (b) corrected SST image by the median filtering method; and (c) the manually produced lead map. Lead and ice-covered areas are marked as blue and white, respectively.

120 **2.2.3. Meteorological data**

An aerodynamic bulk formula was adopted to calculated the THF occurring over leads in this research (Goosse et al., 2001). Aside from the leads' surface temperature, the aerodynamic bulk formula requires the corresponding 10 m wind speed, 2 m air temperature, and dew point temperature. These corresponding meteorological data were collected from the European Center for Medium-Range Weather Forecasts ERA5 reanalysis hourly dataset (https://cds.climate.copernicus.eu/). All of the data
125 were resampled to 100 m resolution.

2.2.4. Co-registration of MODIS and Landsat-8 imagery

Note that MODIS and Landsat-8 products employ different spatial reference systems, because the MOD021KM data set uses a geolocation subset comprised of longitude and latitude coordinates to provide a geographic location, while Landsat-8 imagery uses a projected coordinate system. These datasets need to be converted to the same spatial reference system when used in the
130 same experiment. To avoid deviations of footprints in the different projected coordinate systems and achieve an accurate registration, we transformed Landsat-8 data into the geolocation data grid of MOD021KM using latitude and longitude data. Specifically, the geolocation data of MOD021KM given at a 5 km resolution was interpolated to form a 100 m resolution geolocation grid by a subpixel interpolation strategy. This strategy comprised three processes (Fig. 3). First, for a 100 m resolution pixel P to be interpolated, four bounded 5 km resolution pixels P1 to P4 were searched. Second, we obtained two
135 pixels P' and P'' with 100 m resolution on the along-track line, which lies the same along-scan line with P; then, the positions

(longitude, latitude) of P' and P" were interpolated using P1/P3 and P2/P4. One reasonable approximation done here is that two successive scan lines are parallel to each other. Third, the position of P was interpolated with P'/P". For each pixel in the interpolated geolocation grid, corresponding Landsat-8 pixels were searched using the latitude and longitude of the grid center. Since points on the earth surface with the same longitude and latitude are identical to each other in difference spatial reference

140   systems, the proposed approach can register the different datasets accurately.

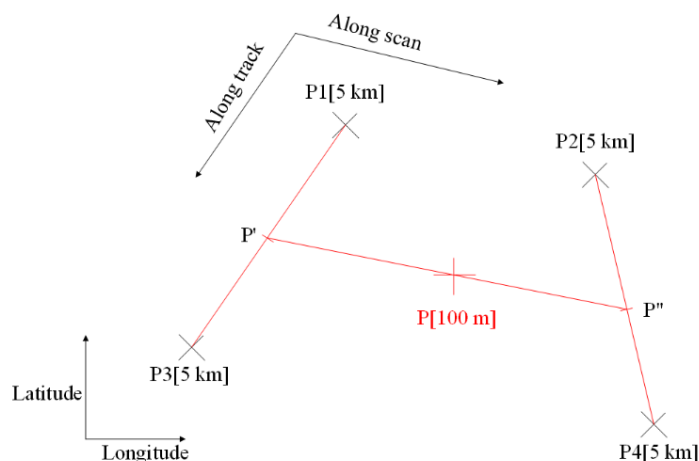
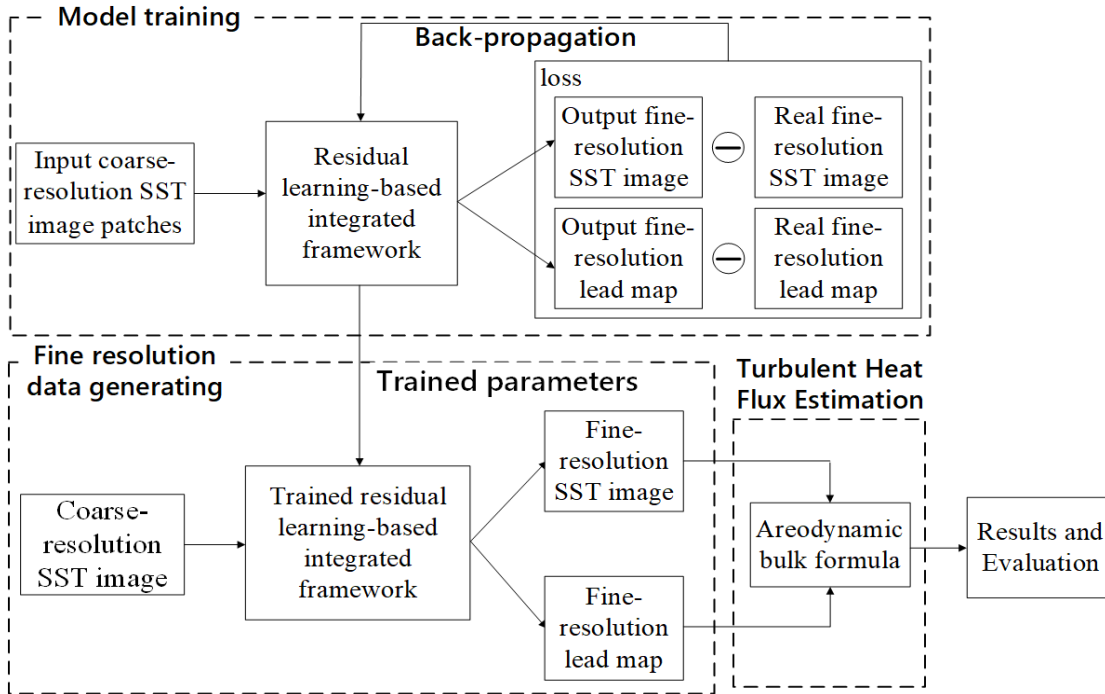



Figure 3. The process used to generate a fine resolution geolocation grid

3. Methods

The estimation of the THF at a subpixel scale involved two main processing steps (Fig. 4). First, a fine-resolution SST image

145   and the corresponding fine-resolution lead map are produced from a coarse-resolution SST image with a CNN-based integrated framework. Second, the THF is estimated from the fine-resolution SST image and a lead map using an aerodynamic bulk formula, and finally the accuracy of the results is assessed.

150 **Figure 4. The flowchart of the estimation of turbulent heat flux at the subpixel scale with the proposed deep residual convolutional neural network-based framework. Note: SST, sea surface temperature.**

3.1. Generation of a fine-resolution SST image and lead map

With the coarse-resolution SST image as input, the proposed integrated framework (Fig. 4) aims to produce an SST image and a lead map both with fine resolution. Specifically, provided that the original coarse-resolution SST image I_{CR}^{SST} has $H \times W$ pixels,

155 the objective of the integrated framework is to generate an SST image I_{FR}^{SST} and a lead map M_{FR}^{Leads}, both of which contain $(z \times H) \times (z \times W)$ pixels, where z is the scaling factor that equals ten in this study.

Although the key factors for SR SST image reconstruction and lead mapping both involve modeling the nonlinear relationship between coarse- and fine-resolution data, the objectives of each one are different. The process of generating fine-resolution SST images focuses on recovering a fine spatial pattern, while in producing a fine-resolution lead map, it is crucial

160 to classify each sub-pixel in addition to recovering the fine spatial pattern. Therefore, we used two CNNs, a very deep residual CNN and a multi-level feature fusion residual CNN, with different structures in the integrated framework to achieve generation of a fine-resolution SST image and a lead map. For fine-resolution SST image reconstruction, a very deep residual CNN which has been widely used in image SR (Zhang et al., 2018;Ledig et al., 2017) was used (Fig. 5). The SR lead mapping method essentially is a type of image segmentation. Considering the good performance of an encoder-decoder structure in image

165 segmentation (Ronneberger et al., 2015;Badrinarayanan et al., 2017), we combined the very deep CNN and encoder-decoder structure, and applied a multi-level feature fusion residual CNN (Fig. 5) to SR lead mapping.

The procedure of CNN-based SR SST reconstruction and lead mapping comprises three parts: (1) training data preparation, (2) CNN model training, and (3) fine-resolution SST image and lead map prediction with the trained CNN models. The following subsection explains the two CNNs more fully.

170 ### 3.1.1 Architecture of the integrated framework

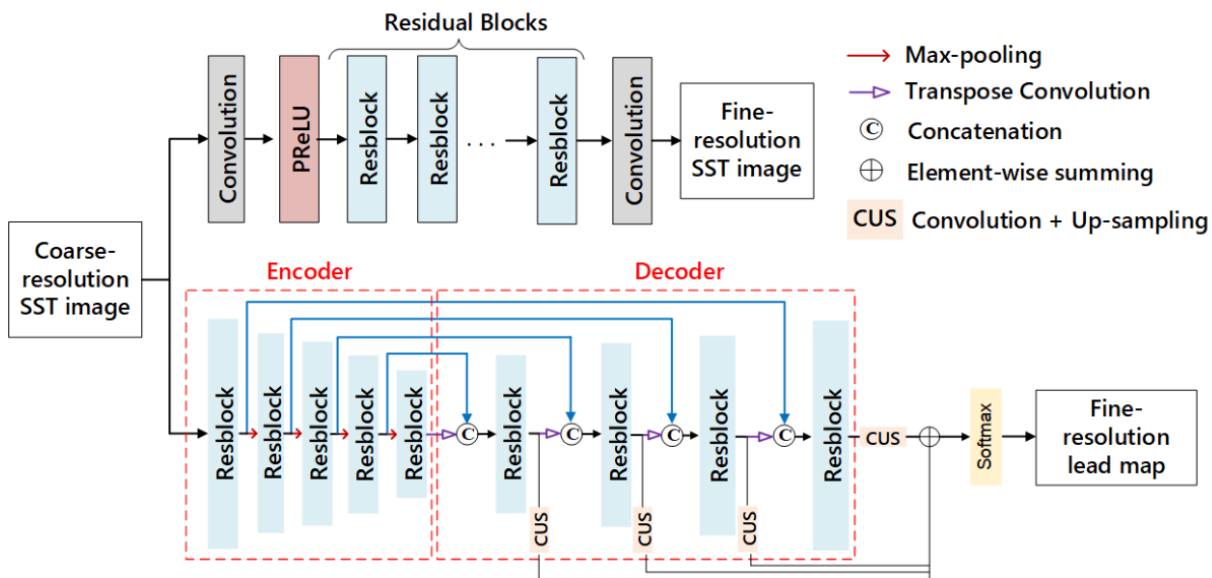



Figure 5. Architecture of the two convolutional neural networks for sea surface temperature (SST) image super-resolution reconstruction and super-resolution lead mapping. Note: PReLU, parametric rectified linear unit.

A very deep residual CNN model with 48 layers was used for SR SST reconstruction. The input coarse SST image was
175 followed by a convolution layer with 64 filters of size 3 ×3 and a parametric rectified linear unit which was used to ensure the output is a nonlinear expression of the input data. At the core of the very deep residual CNN are nine residual blocks, each of which contains two convolution layers with 64 filters of size 3 ×3 ×64 followed by batch norm layers and parametric rectified linear unit functions (termed here as Resblock). The last layer of the very deep residual CNN consists of a single filter of size 3 ×3 ×64.

180 A multi-level feature fusion residual CNN model was used for SR lead mapping. The model includes a symmetric encoder-decoder module and a feature fusion unit. The backbone of the encoder-decoder also consisted of nine residual blocks. The sizes of filters for the first convolution layer in each Resblock in the decoder part was 3 ×3 ×128, while all other filters in the Resblock had a size of 3 ×3 ×64. Additionally, a max-pooling layer which contained 64 filters of size 2 ×2 with a sliding step of 2 was added behind each Resblock in the encoder procedure to downscale the feature maps and amplify the receptive
185 field. A transpose convolution, which is a reverse process to normal convolution (Noh et al., 2015), and a concatenation operation were used in the decoder procedure to enlarge the size of feature maps and fuse multi-level features, respectively. An attention mechanism module was employed in the concatenation to increase the feature difference at the boundary of a

lead and ice. In the feature fusion part, extracted features in the decoder part were up-sampled and fused with element-wise

summing to combine multi-level features. The scaling factors of the four up-sampling modules (from left to right) were 8, 4,

190   2, and 1, respectively, according to the scale differences between the extracted features and the output image. The last layer

after the feature fusion part comprised a softmax function as the activation function to estimate the class label (lead or not lead)

of every pixel.

3.1.2 Implementation details of the CNNs

According to the structure of the CNN models, the input SST image should match the size of the fine-resolution images.

195   Therefore, the coarse-resolution SST image in training samples must be interpolated. A standard method such as bicubic

interpolation may be used to generate this input data set. The CNN models can be trained using the interpolated coarse-

resolution SST image patches x, the corresponding referenced fine-resolution SST image patches y, and lead map patches l.

Given the different objectives of image SR SST reconstruction and image SR lead mapping, mean square error (MSE) loss

and cross entropy loss were used as the loss functions for the two associated CNNs, respectively. They can be calculated

200   following Eqs. (1) and (2):

$$L_{SR-SST}\left(w_{SR-SST}\right)=\frac{1}{n}\sum_{i=1}^{n}\left\|y_{i}-F_{SR-SST}\left(w_{SR-SST},x_{i}\right)\right\|^{2}, \tag{1}$$

$$L_{SR-LM}\left(w_{SR-LM}\right)=-\frac{1}{n}\sum_{i=1}^{n}\left[l_{i}*\log\left(F_{SR-LM}\left(w_{SR-LM},x_{i}\right)\right)+\left(1-l_{i}\right)*\log\left(1-F_{SR-LM}\left(w_{SR-LM},x_{i}\right)\right)\right], \tag{2}$$

where n is the number of training samples, $F(\cdot)$ denotes the networks, w is the weight parameters of the networks to be updated

in the training process, subscripts represent different networks, SR-SST and SR-LM denote SR of SST and lead mapping,

205   respectively. For optimization, adaptive moment estimation (Adam) (Kingma and Ba, 2014) with standard back propagation

was applied to minimize the loss and update the network weights until convergence; the parameters of Adam were set as: β_1=0,

β_2=0.999. The learning rate α was initialized as 1×10^{-4}.

Once the two networks have been trained, they can be used to generate the fine spatial resolution SST images and

corresponding lead maps. During this procedure, the coarse-resolution SST image was fed into the integrated framework. The

210   fine-resolution SST image was directly produced from the SR SST reconstruction CNN, while the output of the SR lead

mapping CNN was a lead indicator image. An appropriate threshold was used to binarize the lead indicator image into a lead

map according to specific requirements. Here, the threshold value was empirically set to 0.5, meaning that a fine spatial

resolution pixel in the indicator image exceeding 0.5 would be classified as a lead pixel.

3.2. Estimating THF over leads

215   Given a set of data including fine-resolution SST images, the corresponding lead maps and related meteorological data (10 m

wind speed, 2 m air temperature, and dew point temperature), the THF over each lead was estimated using the traditional

aerodynamic bulk formula (Brodeau et al., 2017;Goosse et al., 2001). Note that overall THF is comprised of two parts including sensible (H_s) and latent heat fluxes (H_l). In the bulk formulae, it is assumed that H_s and H_l were mainly determined by the temperature and humidity differences between the leads' surface and atmosphere at a certain height r (2 m used in this study),

220 and they can be calculated following Eqs. (3) and (4):

$$H_s = \rho c_\mathrm{p} C_\mathrm{sh} U_r \left(T_s - T_r \right),$$ (3)

$$H_l = \rho L_\mathrm{w} C_\mathrm{le} U_r \left(q_s - q_r \right),$$ (4)

where ρ is the air density, c_p is the specific heat of air, and L_w is the latent heat of evaporation; these are constants in the bulk formula. In addition, U_r, T_r, and q_r represent the air velocity, temperature, and specific humidity, respectively, at the certain

225 height r=2 m; T_s and q_s are the surface temperature and specific humidity at the leads' surface, respectively; C_sh and C_le are transfer coefficients. All of these variables can be acquired or calculated from the SST image and meteorological data. A more detailed description can be found in the original article (Goosse et al., 2001). Once H_s and H_l have been calculated, the overall THF can be obtained by summing up them.

3.3. Accuracy assessment

230 The outputs obtained from DeepSTHF were compared with those from the bicubic interpolation-based subpixel-scale method (termed here as CubicSTHF) and the pixel-scale method (termed here as OriTHF). For the CubicSTHF method, the coarse-resolution SST image was first super-resolved by cubic interpolation. Then the resulting super-resolved SST image was used to produce a corresponding lead map with a pixel-based classification approach (Willmes and Heinemann, 2015). Finally, the THF over leads was calculated using the super-resolved SST image and lead map. For the OriTHF method, the THF over leads

235 was estimated using the original coarse-resolution SST image and the corresponding lead map, which was also produced from the coarse-resolution SST image with the pixel-based approach.

The results of image SR reconstruction of the SST, SR lead mapping, and estimated THF over leads were all assessed as follow. (1) For the result of SST image SR reconstruction in a simulated test, the root mean square error (RMSE), and the Pearson coefficient (R) were calculated using the real fine resolution SST image as a reference. (2) For the SR lead mapping,

240 the fine-resolution lead map drawing from Landsat-8 imagery was taken as a reference; we employed the overall accuracy, the omission error, the commission error, and the mean intersection over union (MIOU) to evaluate the results in both simulated and real tests. (3) For the estimation of THF, the overall error (OE) and RMSE were calculated for the result of simulated test using the result calculated from the fine-resolution images as a reference.

4. Experiments and Results

4.1. Experiment with simulated MODIS imagery

A simulated experiment with Landsat-8 data was first applied to explore the strengths of DeepSTHF as well as to avoid the uncertainty of co-registration and temperature estimation differences between Landsat-8 and MODIS data. In this experiment, MODIS SST images were simulated from the Landsat-8 SST images using the pixel aggregate method and were used as the input coarse-resolution data. The original Landsat-8 SST images and corresponding lead maps were used as the fine-resolution data as a reference.

4.1.1. Training and testing data

The training samples of the CNN models were generated from the SST images and lead maps derived from ten Landsat-8 images acquired during 2013–2017. The SST images and lead maps were clipped into image patches, each of which had a size of 80×80 pixels; the step size for the clipping was 40. The clipped SST image patches were degraded to 1000 m to simulate the MODIS SST imagery. A total of 36,000 image patches, comprising degraded SST image patches, original Landsat-8 SST image patches, and lead map patches, were randomly selected to form the training data. The degraded SST image patches, and the corresponding original SST image patches were used to train the SR SST image reconstruction CNN. Similarly, the degraded SST image patches, and the corresponding lead map patches were used to train the SR lead mapping CNN.

During the testing process, an SST image and lead map obtained from the Landsat-8 scene (p071r010) acquired on 31 March 2020 were used. We degraded the SST image to 1000 m and used it as the input of the trained CNNs. The original SST image and lead map were used as real data to validate the results.

4.1.2. Results

The simulated coarse-resolution MODIS SST image, SR SST images produced by CubicSTHF as well as DeepSTHF, the reference fine-resolution Landsat-8 SST image, and the error images for the coarse resolution image and SR results are shown in Fig. 6. Note that the overall spatial texture of the output from DeepSTHF is more like the reference Landsat-8 SST image than the coarse-resolution image and that from CubicSTHF. The SR result for CubicSTHF is blurred in lead areas. Though the SR results for CubicSTHF and DeepSTHF methods are similar with the reference data for areas covered by ice, the DeepSTHF method produced more accurate results in areas with leads. From the error images, the original coarse-resolution SST image has the largest RMSE, which is more than twice as much as those of CubicSTHF and DeepSTHF. Even though CubicSTHF generated a smaller number of errors when compared with the coarse-resolution image, the errors in the lead areas were more significant than those of DeepSTHF.

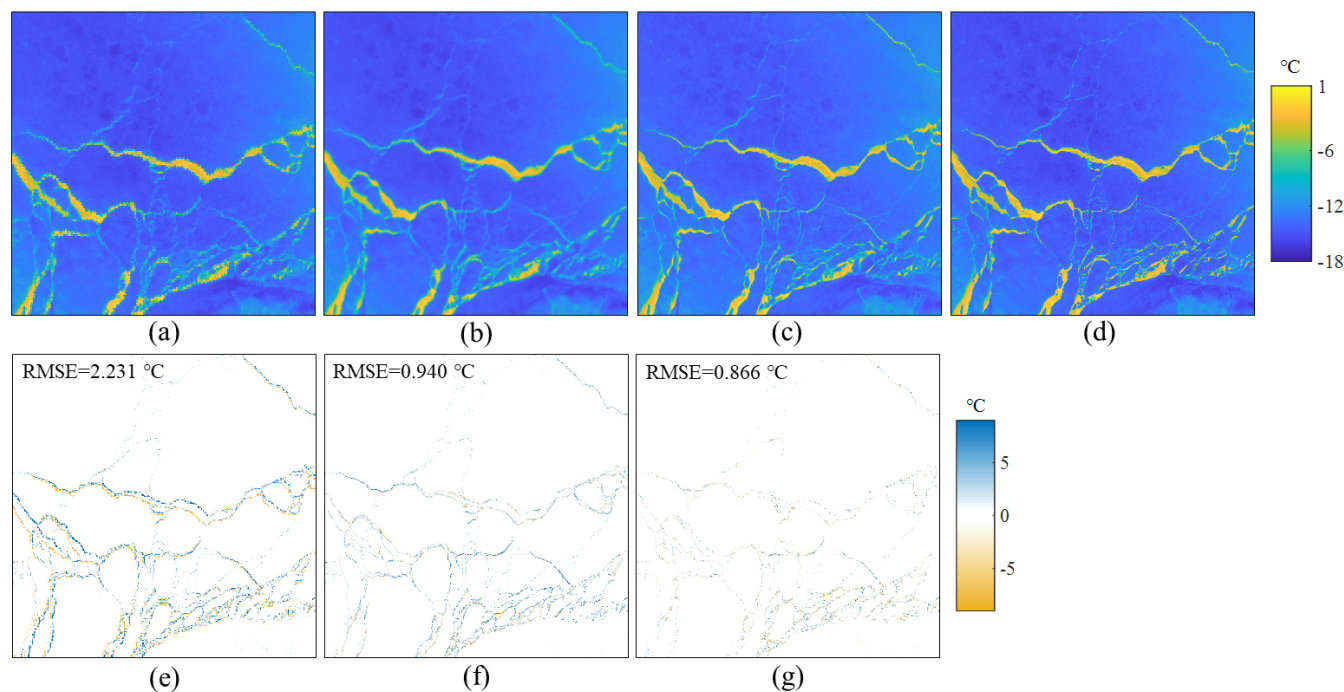



Figure 6. (a) Simulated moderate resolution imaging spectroradiometer (MODIS) sea surface temperature (SST) image with super-resolution results using different methods, specifically (b) the bicubic interpolation-based subpixel-scale method (CubicSTHF) or (c) by a deep residual convolutional neural network-based framework used to estimate THF over leads at the subpixel scale (DeepSTHF); (d) the corresponding Landsat-8 SST image and the error images for (e) simulated MODIS image and super-resolution results, (f) the CubicTHF method, and (g) the DeepSTHF method. The red rectangle indicates the large error of CubicSTHF.

Fig. 7 shows the scatter plots allowing a comparison between reference SST in generated lead areas against corresponding SST from the coarse-resolution image and the super-resolved images of CubicSTHF and DeepSTHF. For the plot with OriTHF, the R^2 was only 0.274. Much higher coefficients were observed for CubicSTHF and DeepSTHF, indicating the results of SR methods have a much stronger correlation with the fine-resolution image. Among the two SR methods, DeepSTHF resulted in a higher R^2, as well as a lower RMSE than CubicSTHF. Additionally, it is apparent that the CubicSTHF method underestimated most pixels with a reference temperature > −6°C because substantial data points fall below the diagonal line (Fig.7b). This problem has been improved by using the DeepSTHF method; with this method data points fall closer to the diagonal line and comparable data points are located in both sides of the line when comparing the two methods. Overall, the DeepSTHF method achieved the most accurate SST image SR results.

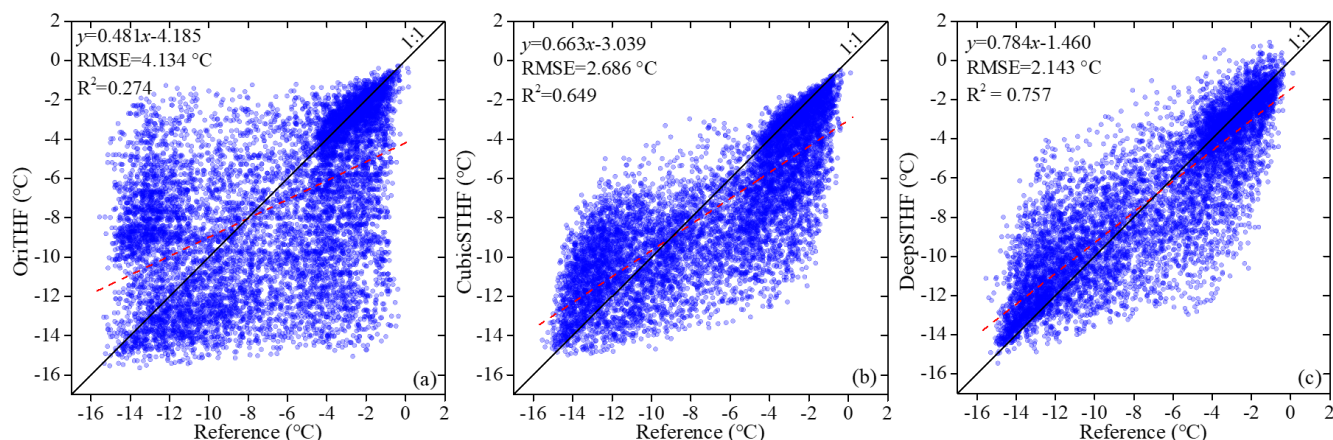



Figure 7. Scatter plots between sea surface temperature in lead areas from Landsat against that from simulated moderate resolution imaging spectroradiometer (MODIS) image at the pixel-scale (OriTHF) and the super-resolution results with the (b) bicubic interpolation-based subpixel-scale method (CubicSTHF), and (c) the deep residual convolutional neural network-based framework method used to estimate THF over leads at the subpixel scale (DeepSTHF). The red dashed lines are fitted linear regression lines of the data points

Fig. 8 shows the lead maps produced with the OriTHF, CubicSTHF, and DeepSTHF methods. The main lead networks generated from the three methods are similar to those in the reference fine-resolution lead map (Fig. 8c), especially for leads wider than several kilometers. However, the boundaries of mapped leads for OriTHF are not smooth and not visually realistic. Many narrow lead networks were not extracted by OriTHF and CubicSTHF (red ellipses in Figs. 8a and 8b). In contrast, the lead map produced using the DeepSTHF method is more visually realistic and much closer to the reference lead map. Many narrow leads were correctly mapped, and their connectivity was well-maintained. Note that some very narrow lead, especially for leads with a width smaller than 5 pixels in the fine-resolution lead map, became disconnected in the DeepSTHF model (red rectangle in Fig. 8c), because the ice lead fraction in the mixed pixels of the coarse-resolution SST image is too small to provide detailed lead information.

| (a) OriTHF | (b) CubicSTHF | (c) DeepSTHF | (d) Reference map |

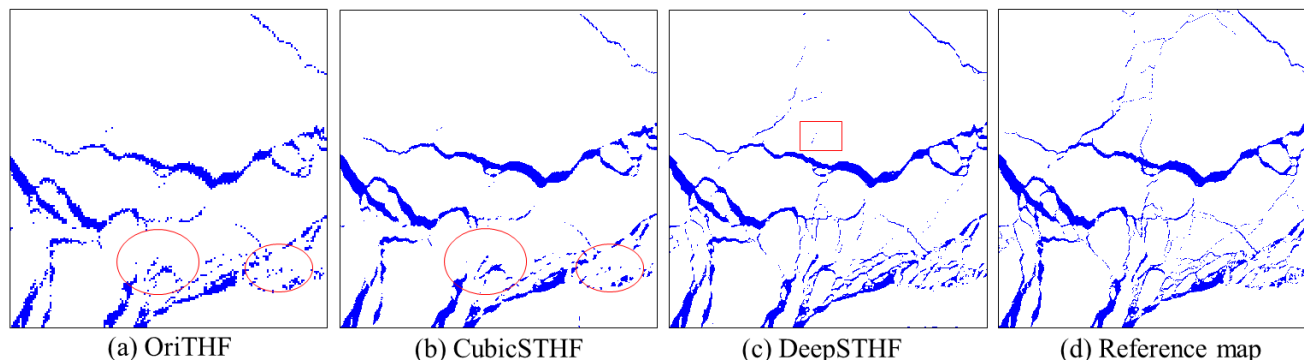



Figure 8. Lead maps generated from the simulated sea surface temperature image of coarse spatial resolution by different methods (a, b, and c) and (d) the reference lead map extracted from Landsat-8 operational land imager data. Lead and ice-covered areas are marked as blue and white, respectively. Lead maps generated by the bicubic interpolation-based at (a) pixel scale (OriTHF) and (b) bicubic interpolation-based sub-pixel scale (CubicSTHF) with the temperature anomaly threshold approach, (c) is the lead map generated by the deep residual convolutional neural network-based framework method used to estimate THF over leads at the subpixel scale (DeepSTHF) based on convolutional neural network model. The red ellipses represent lead networks that have not been mapped by (a) OriTHF and (b) CubicSTHF. (c) The red rectangle indicates an area with very narrow leads that have not been mapped by DeepSTHF.

The quantitative assessment results for lead mapping of the OriTHF, CubicSTHF, and DeepSTHF methods is demonstrated in Table 1. The DeepSTHF method has a greater overall accuracy and MIOU, and a smaller omission error than OriTHF and CubicSTHF. The omission errors for the OriTHF and CubicSTHF methods were 0.341 and 0.240, much greater than that for the DeepSTHF, indicating that many lead pixels were not extracted, which is consistent with the visual performance (Fig. 8). Additionally, although more lead pixels have been identified by the DeepSTHF method, it did not increase the rate of commission errors. Overall, the DeepSTHF method allowed the production of the most accurate lead map.

Table 1. Accuracies of leads produced by the threshold approach and convolutional neural network model. (The most accurate results are highlighted in bold text)

Method	Overall accuracy	Commission error	Omission error	MIOU
OriTHF	0.961	0.015	0.341	0.756
CubicSTHF	0.975	**0.008**	0.240	0.834
DeepSTHF	**0.980**	**0.008**	**0.171**	**0.865**

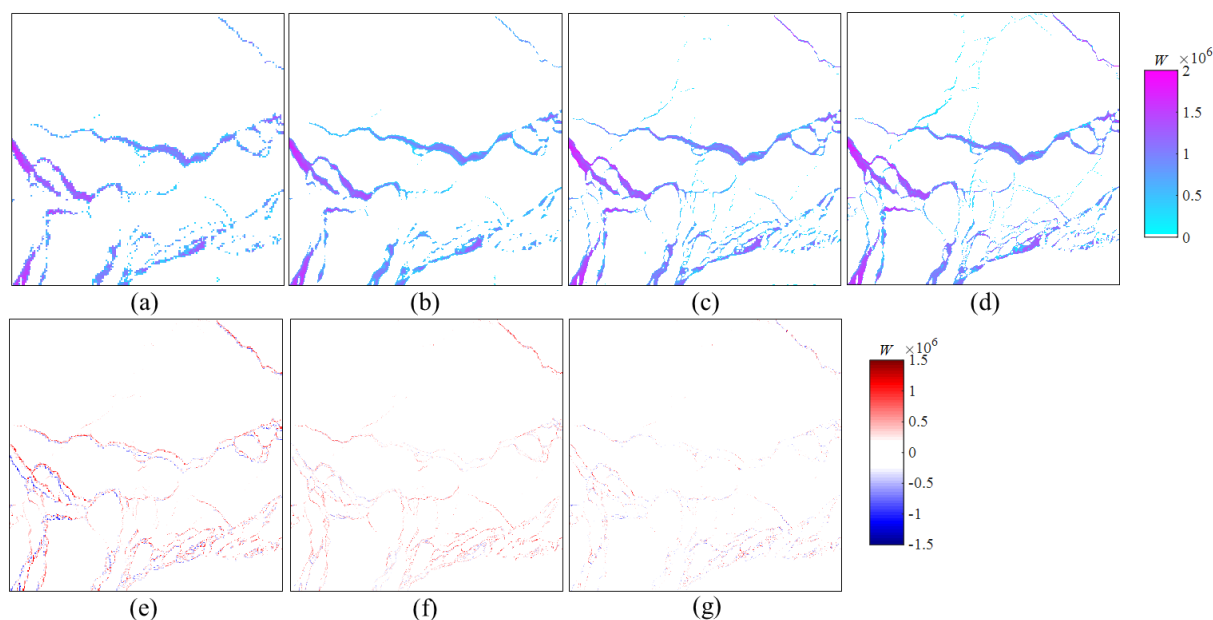
Note: MIOU, mean intersection over union; OriTHF, the pixel-scale turbulent heat flux estimation method; CubicSTHF, the bicubic interpolation-based sub-pixel scale heat flux estimation method; DeepSTHF, the deep residual convolutional neural network-based sub-pixel heat flux estimation method.

For both SST image SR reconstruction and SR lead mapping, the DeepSTHF method generated the most accurate results. This mainly occurred because the CNN-based DeepSTHF model has the ability to extract a potential spatial pattern in the coarse- and fine-resolution SST images/lead maps through learning and built an appropriate nonlinear relationship between them to implement subpixel analysis, which is essential for producing reliable SR results. Meanwhile the CubicSTHF method generated a value for each pixel in the SST image SR result that is a linear combination of surface temperature of neighboring

pixels, which makes it difficult to represent the complex nonlinear relationship between fine- and coarse-resolution images. Additionally, the threshold approach was applied in the CubicSTHF method to extract leads; this is a pixel-based method and

335   cannot achieve subpixel analysis.

Fig. 9 shows the distribution of the THF over leads estimated by the three methods. The DeepSTHF method preserved abundant spatial texture and achieved a more accurate result than OriTHF and CubicSTHF. Because OriTHF and CubicSTHF failed to retrieve many small leads, the THF over these leads was not calculated (Figs. 9a and 9b) and thus significant errors in the corresponding areas are shown in the error map (Fig.9e). Additionally, even though the estimated THF for large leads

340   with OriTHF and CubicSTHF was close to the reference data, the obtained THF along boundaries was much lower than the true value and resulted in large errors (Figs. 9e and 9f), especially for the error image of OriTHF. In contrast, DeepSTHF produced the smallest overall error.

345   **Figure 9. Spatial distribution of turbulent heat flux (THF) calculated by different methods (a–c) and (d) the reference distribution and the error maps of the distribution of THF estimated by (e) the pixel-scale turbulent heat flux estimation method, (f) the bicubic interpolation-based sub-pixel scale heat flux estimation method, and (g) the deep residual convolutional neural network-based sub-pixel heat flux estimation method.**

The scatter plots of the reference THF data plotted against those estimated from OriTHF, CubicSTHF, and DeepSTHF are

350   shown in Fig. 10. Generally, data points of plot created with the DeepSTHF method fall closer to the diagonal line than those created with OriTHF and CubicSTHF. Few calculated THF values of OriTHF and CubicSTHF were less than 0.25×10^6 W; this mainly occurred because small leads were not mapped and thus the corresponding THFs were not estimated. The R^2 for the plots of CubicSTHF and DeepSTHF are both much higher than that of OriTHF. Even though the result of CubicSTHF has a higher correlation with the reference data than that of OriTHF, it is evident that most of the pixels' values were underestimated

355 (Fig. 10 b). Note that, for all of the plots, several data points fell on the vertical and horizontal axes; this resulted from the omission (points on the horizontal axis) and misclassification (points on the vertical axis) of lead pixels of the three methods during the lead mapping.

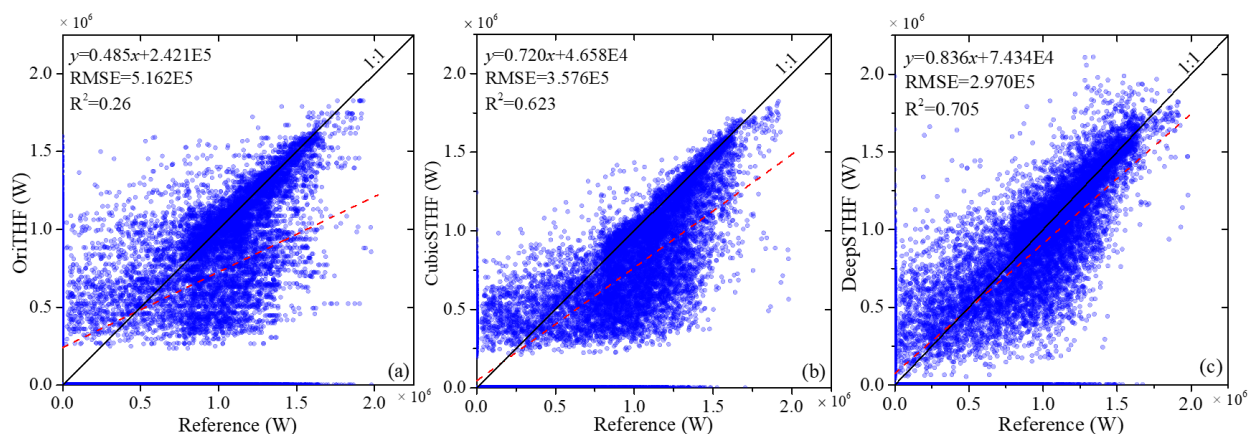



Figure 10. Scatter plots of the turbulent heat flux calculated from the reference data and those from the estimated from (a) the pixel-
360 **scale turbulent heat flux estimation method (OriTHF), (b) the bicubic interpolation-based sub-pixel scale heat flux estimation method (CubicSTHF), and (c) the deep residual convolutional neural network-based sub-pixel heat flux estimation method (DeepSTHF). The red dashed lines are fitted linear regression lines of the data points.**

The total THF estimated with OriTHF, CubicSTHF, and DeepSTHF, and the accuracies of each method are listed in Table 2. Generally, the total estimated THF estimated by the DeepSTHF method was the closest value to the reference data. Although
365 the THF calculated from OriTHF was relatively closer to the reference value than that from the CubicSTHF, the RMSE of it was much larger. In contrast, the THF estimated from DeepSTHF had the smallest RMSE, greatest R^2, and smallest OE, especially for the OE; the OE from OriTHF and CubicSTHF was almost three times than that from DeepSTHF. The THF error (real value minus estimated value) distributions of the three methods are further demonstrated in Fig. 11. More than 70% of the errors found in DeepSTHF data are located in the small error bin ([$-0.25 \times 10^{11}W$, $0.25 \times 10^{11}W$]), which is much greater
370 than those of OriTHF (less than 50%) and CubicSTHF (less than 60%). For the DeepSTHF method, the errors have close to a normal distribution and the rate of positive errors is close to that of negative errors. Meanwhile, for CubicSTHF, the THF of most pixels was underestimated (Fig. 10); the rate of positive errors was significantly larger than the corresponding negative errors, and the errors were biased. Therefore, compared with CubicSTHF, although the improvement of DeepSTHF in RMSE was not very large, the total estimated THF of it was much closer to the reference data because the most positive and negative
375 errors tended to cancel each other out, which is statistically good. These findings indicated that the proposed method, DeepSTHF, can produce a more favorable result and is able to estimate THF over leads at a subpixel scale.

380

Table 2. Accuracies of the estimated turbulent heat flux (THF) estimated by different methods

Method	Reference THF (*W*)	Estimated THF (*W*)	OE	RMSE (*W*)	R^2
OriTHF		1.871×10^{11}	3.72×10^{10}	5.162×10^5	0.260
CubicSTHF	2.243×10^{11}	1.765×10^{11}	4.78×10^{10}	3.576×10^5	0.623
DeepSTHF		2.115×10^{11}	1.28×10^{10}	2.970×10^5	0.705

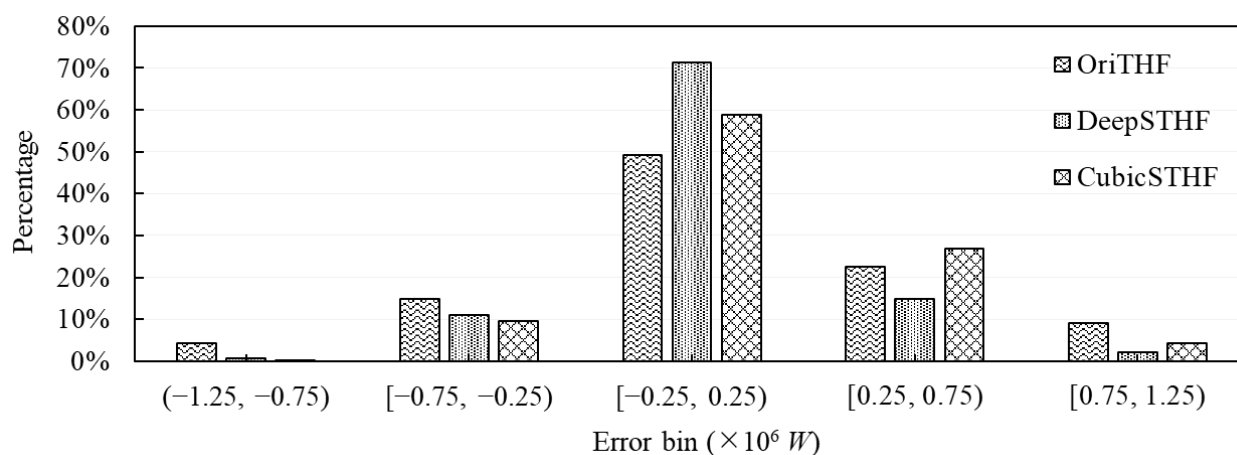
Note: OriTHF, the pixel-scale turbulent heat flux estimation method; CubicSTHF, the bicubic interpolation-based sub-pixel scale heat flux estimation method; DeepSTHF, the deep residual convolutional neural network-based sub-pixel heat flux estimation method.

385

Figure 11. Error distribution of estimated turbulent heat flux for three different methods: OriTHF, the pixel-scale turbulent heat flux estimation method; DeepSTHF, the deep residual convolutional neural network-based sub-pixel heat flux estimation method; and CubicSTHF, the bicubic interpolation-based sub-pixel scale heat flux estimation method.

4.2. Experiment with real MODIS imagery

390 To assess the performance of the proposed DeepSTHF model in practical applications, an experiment using real MODIS SST imagery was conducted. In this experiment, MODIS images were used to form the input coarse-resolution SST images, while Landsat-8 images were used to produce the reference fine-resolution SST images and lead maps.

4.2.1. Training and testing data

Ten MODIS SST images and the corresponding Landsat-8 SST images and lead maps acquired during 2013–2017 were used
395 to create training samples of the CNN models. The Landsat-8 images were converted to a MODIS geolocation grid to achieve accurate co-registration between Landsat-8 and MODIS imagery. Using the same method used in the simulated SST experiment, we clipped these images into image subsets with a size of 80 × 80 pixels at a step size of 40. For SR SST image

reconstruction using a CNN and image SR lead mapping with a CNN, a total of 36,000 randomly selected MODIS SST image patches along with the corresponding Landsat-8 SST image and lead map patches were used as the training data sets.

400   Three MODIS SST images, one each acquired on 25 April 2018, 9 May 2019, and 31 March 2020, were used to estimate the THF at a subpixel scale. For each MODIS image, a subset containing leads with different widths and lengths was selected for the experiment. Three corresponding Landsat-8 scenes (p057r010, p111r240, and p071r010) were employed to provide reference fine-resolution data used to validate the results. The generalization ability of the proposed model could be accurately validated because the test images were located in different regions and observed on different dates.

405 Note that, large temperature differences could be observed between MODIS and Landsat-8 SST images, which had mainly resulted from the different overpass times of the MODIS and Landsat-8 satellites. A possible way to reduce this inconsistency is to apply a temporal correction method (Zhao et al., 2019). In practice, however, the temporal correction method would bring about an additional layer of error. Considering this factor, MODIS and Landsat-8 SST were normalized for training by the min–max normalization method. During the testing stage, the MODIS SST image SR results were not evaluated quantitatively

410 due to a lack of true fine resolution SST reference imagery. For lead mapping, it was assumed that the range of leads varies little from the MODIS observation time to Landsat-8 observation time, so that the lead maps produced from Landsat-8 OLI data could be used to validate the SR lead mapping results.

4.2.2. Results

The MODIS SST images and the result of image SR reconstruction for them with CubicSTHF and DeepSTHF are shown in

415 Fig. 12. A visually assessment of the results from CubicSTHF and DeepSTHF show those results are more realistic than those of the original MODIS SST images, which are not smooth along the boundaries of the lead networks. Compared with the results generated by the CubicSTHF method, finer spatial textures can be observed in the images produced by DeepSTHF. For lead networks, the SR result is blurred to some extent for CubicSTHF, but it is more sharp for DeepSTHF. The temperature difference between lead and ice areas along boundary for DeepSTHF is more significant than that for CubicSTHF.

420

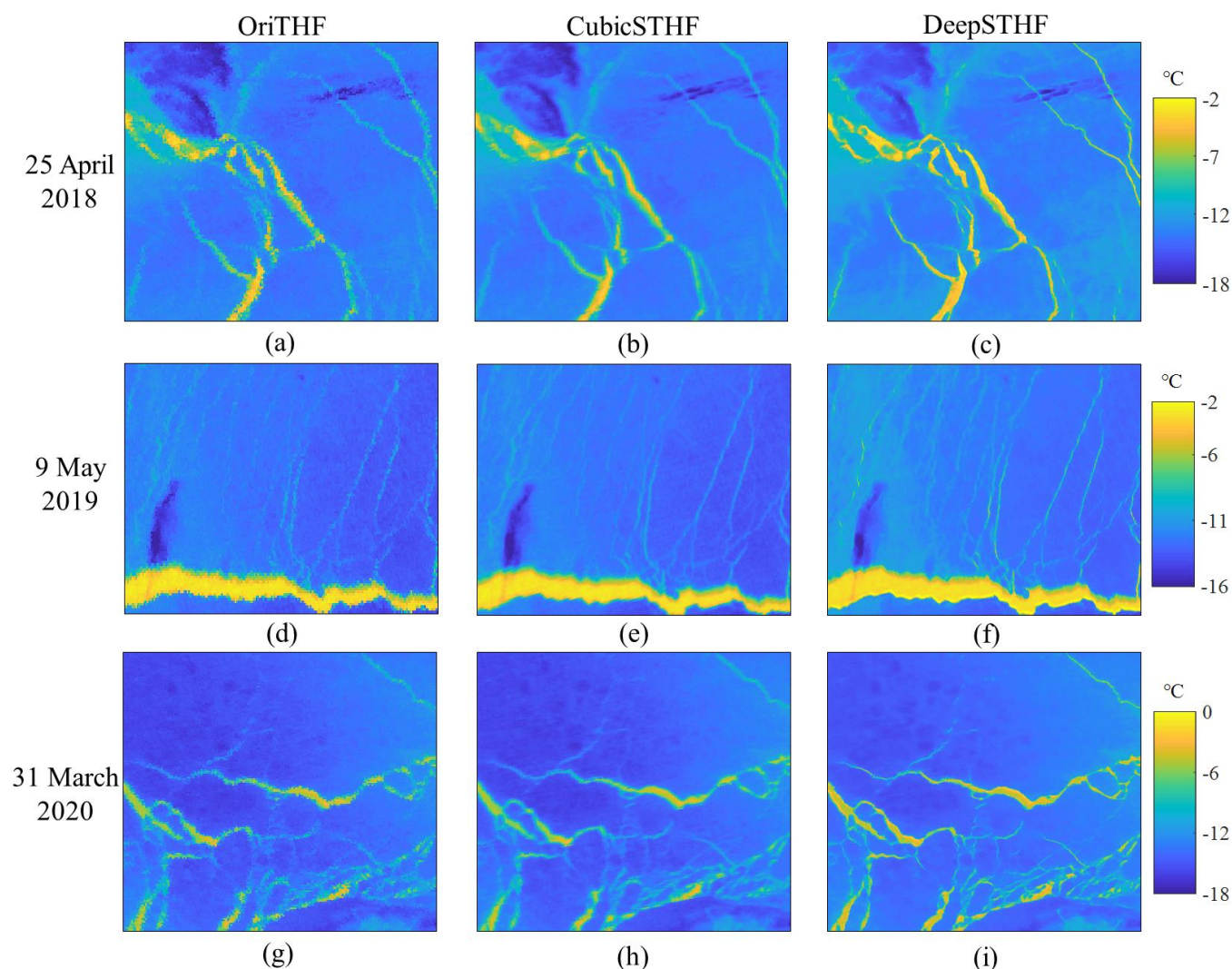



Figure 12. Sea surface temperature images acquired on: (a)–(c) 25 April 2018, (d)–(f) 9 May 2019, and (g)–(f) 31 March 2020 and the super-resolution result of each set of images. Each set of three images include the (a, d, g) original moderate resolution imaging spectroradiometer image (OriTHF), and the super-resolution results for sea surface temperature estimated with the (b, e, h) bicubic interpolation-based sub-pixel scale heat flux estimation method (CubicSTHF) and (c, f, i) the deep residual convolutional neural network-based sub-pixel heat flux estimation method (DeepSTHF), respectively.

The generated lead maps produced using the OriTHF, CubicSTHF, and DeepSTHF methods are displayed in Fig. 13. Generally, the lead maps yielded by DeepSTHF were more like the reference data maps than those produced by the other methods. The results produced by OriTHF and CubicSTHF failed to identify many narrow leads and the corresponding lead networks became disconnected (red rectangles in Figs. 13b, 13f, and 13j). Some parts of ice-covered regions surrounded by leads were misclassified as leads by OriTHF and CubicSTHF (red ellipses in Figs.13b and 13j). Even though the main lead networks were mapped by the OriTHF method, the boundaries of the produced lead networks were jagged (Figs. 13q and 13u). The results from CubicSTHF are smoother than those from OriTHF, but some parts of large lead networks are also

discontinuous (Figs. 13r and 13v), which increased lead widths in them to some extent. In contrast, the proposed method,

435   DeepSTHF, generated more promising results. First, it identified most small leads and maintained their connectivity. Second, the boundaries of the segmented leads were smooth and much closer to those of reference leads (subareas r1–r3). Third, most ice-covered areas surrounded by leads were correctly classified, even if the areas were relatively small (dashed red rectangle in Fig. 13 k). Note that, although DeepSTHF extracted most leads accurately, it failed to perform well on some very narrow leads (especially those with widths of < 5 pixels in the fine-resolution map), because the lead fraction for these leads in the

440   coarse image was too small and could hardly be mapped them the fine resolution lead map through the CNN model. Additionally, the results of DeepSTHF were influenced by some abnormal pixels in the input data. For instance, some pixels were misclassified in a narrow rectangular area (red dashed ellipse in Fig. 13 c) because the temperature of this area did not correctly reflect the actual case in the ocean (Fig. 12 a).

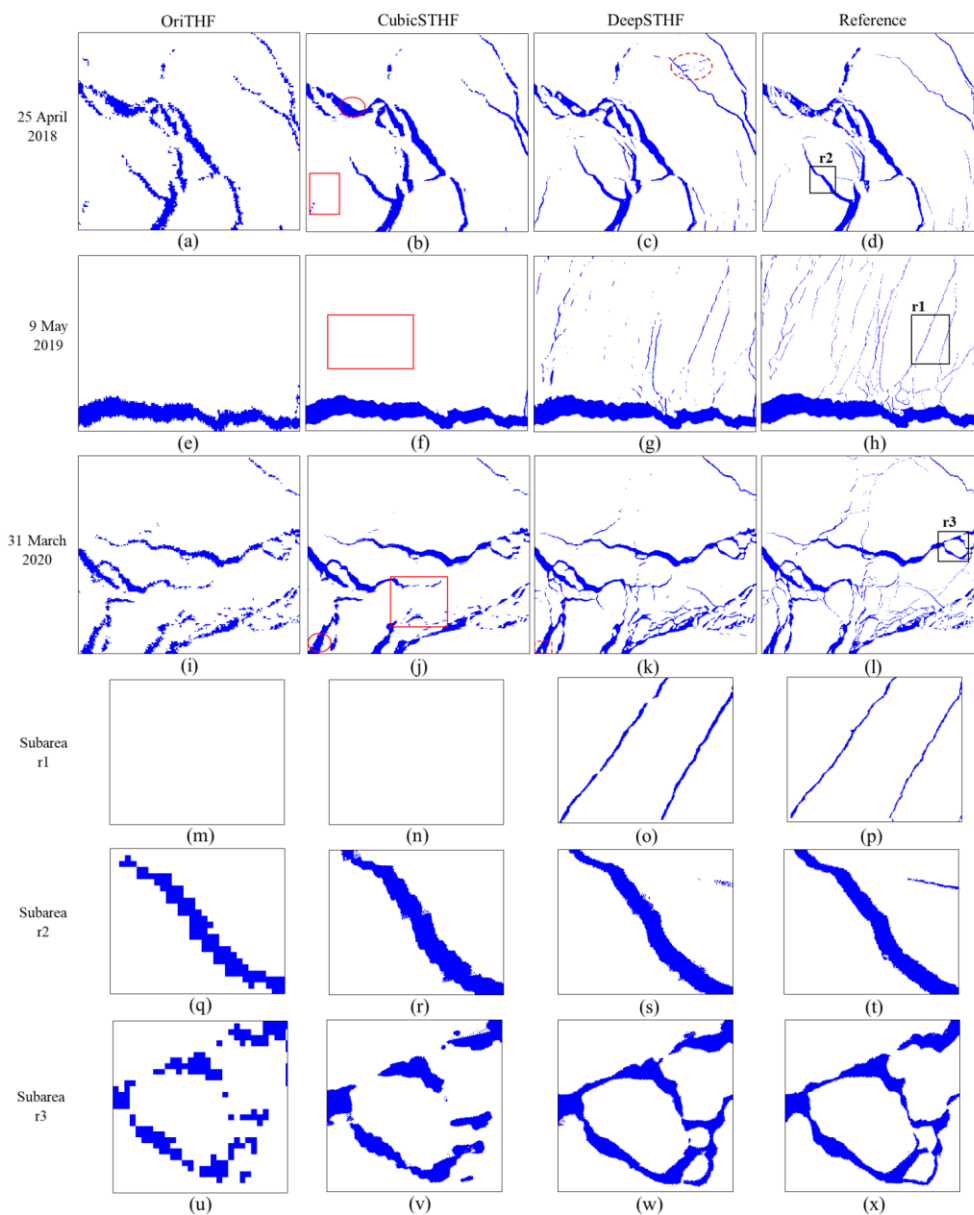



Figure 13. Lead maps generated from moderate resolution imaging spectroradiometer sea surface temperature imagery with the three methods acquired on: (a)–(d) 25 April 2018, (e)–(h) 9 May 2019, and (i)–(l) 31 March 2020. Lead and ice-covered areas are marked as blue and white, respectively. Each set of four parts of the figure (a)–(d), (e)–(h), and (i)–(l) are results from the (pixel-scale turbulent heat flux estimation method (OriTHF), the bicubic interpolation-based sub-pixel scale heat flux estimation method (CubicSTHF), the deep residual convolutional neural network-based sub-pixel heat flux estimation method (DeepSTHF), and the reference lead maps based on data acquired on 25 April 2018, 9 May 2019, and 31 March 2020, respectively. In addition, each set of four parts of the figure (m)–(p), (q)–(t), and (u)–(x) are the results for OriTHF, CubicSTHF, DeepSTHF, and the reference data, respectively, for subareas r1, r2, and r3, respectively. The black rectangles in d, h, and i represent subareas. The red rectangles in b, f, and j represent lead networks that have been mapped by CubicSTHF. The red ellipses in b and j show the ice-covered regions that have been misclassified as leads by CubicSTHF. The red dashed rectangle in k represents ice-covered areas that have been correctly classified by DeepSTHF. The red dashed ellipse in c shows the pixels misclassified as leads by DeepSTHF due to the errors in the sea surface temperature.

The quantitative performances of the generated lead maps by the three methods were compared (Table 3). The DeepSTHF
460   method provided the highest overall accuracy and MIOU, as well as a lowest rates of errors of commission and omission on
25 April 2018 and 31 March 2020. Although the commission error rate of DeepSTHF on 9 May 2019 is higher than those of
OriTHF and CubicSTHF, the overall accuracy and MIOU of DeepSTHF is larger. The accuracy of OriTHF is the lowest except
for 9 May 2019. The commission error rates of the three methods on 9 May 2019 is small, because the lead areas mainly
consisted of a large lead network that could be more easily extracted from this data. For all dates, the omission error rates for
465   OriTHF and CubicSTHF were much larger than those for the CNN, demonstrating that many lead pixels were not correctly
classified, which was consistent with the visual results analyzed subjectively (Fig. 13). Therefore, the proposed framework
was also effective in image SR lead mapping with real MODIS data.

**Table 3. Accuracies of the produced lead maps produced by the pixel-scale turbulent heat flux estimation (OriTHF), bicubic
interpolation-based sub-pixel scale heat flux estimation (CubicSTHF), and deep residual convolutional neural network-based sub-**
470   **pixel heat flux estimation (DeepSTHF) methods represented by the first, second, and third number in each entry, respectively. The
most accurate results are in bold text.**

Date	Overall accuracy	Commission error	Omission error	MIOU
25 April 2018	0.960/0.966/**0.978**	0.027/0.024/**0.015**	0.247/0.185/**0.126**	0.745/0.779/**0.843**
9 May 2019	0.968/0.966/**0.973**	**0.004**/0.006/0.010	0.254/0.265/**0.162**	0.843/0.833/**0.873**
31 March 2020	0.956/0.960/**0.972**	0.022/0.021/**0.018**	0.319/0.281/**0.156**	0.742/0.764/**0.831**

Note: MIOU, mean intersection over union

The THF over mapped leads was estimated by the OriTHF, CubicSTHF, and DeepSTHF methods (Fig. 14). Because the
estimated THF was dependent on the generated lead maps, the spatial distribution of the estimated THF was consistent with
475   the lead maps. Specifically, in the plots of the OriTHF and CubicSTHF, the THF of most small leads has not been depicted
and the boundaries of lead were also not smooth, especially for the plot based on the OriTHF method. Additionally, the
estimated THF of pixels along the boundaries of lead networks was relatively small for CubicSTHF, which may be a result of
the underestimation of temperature in the SR SST image. Meanwhile, for DeepSTHF, the THF over many of the small leads
was estimated, and the overall spatial pattern of the estimated THF in the plot was much finer than OriTHF and CubicSTHF.

480

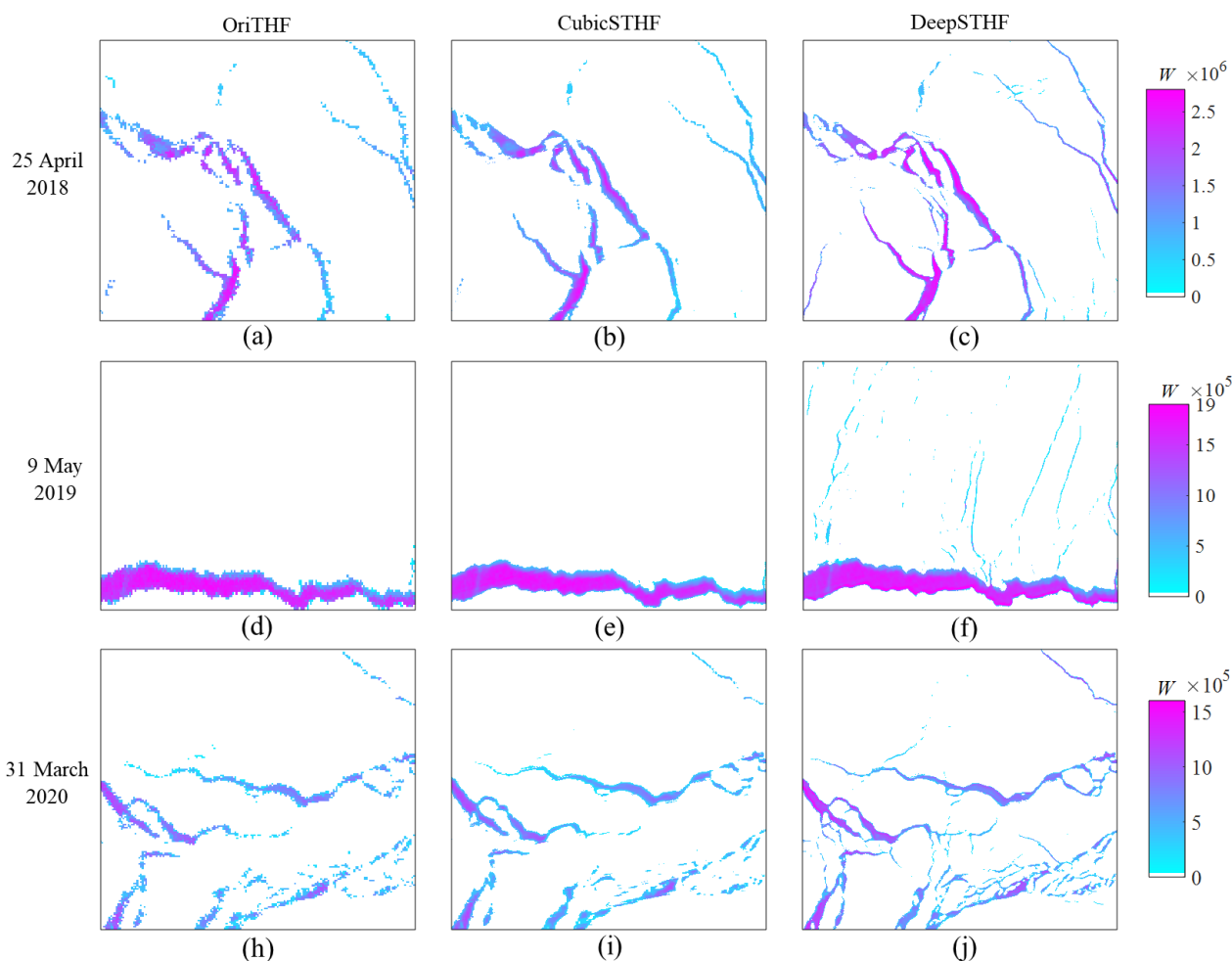



Figure 14. Spatial distribution of turbulent heat flux (THF) calculated by the (a, d, h) pixel-scale turbulent heat flux estimation method (OriTHF), (b, e, i) the bicubic interpolation-based sub-pixel scale heat flux estimation method (CubicSTHF), and (c, f, j) the deep residual convolutional neural network-based sub-pixel heat flux estimation method (DeepSTHF) on: (a)–(c) 25 April 2018, (d)–(f) 9 May 2019, and (h)–(j) 31 March 2020.

The quantities of estimated THF on different dates estimated using the three methods are shown in Table 4. Although more leads were identified using the CubicSTHF method than with OriTHF (indicated by the smaller number of omission errors) on 25 April 2018 and 31 March 2020, the estimated THF from CubicSTHF was slightly smaller than that from OriTHF on both dates. This mainly occurred because the temperature of lead pixels along lead networks was lowered in the SR process with the CubicSTHF method (Fig. 12). Meanwhile, for DeepSTHF, because more leads have been mapped and the reconstructed temperature of lead pixels along lead nets was close to the pixels in the central part of lead networks, the total calculated THF of DeepSTHF on all dates were greater than those of OriTHF and CubicSTHF. It is evident that the THF difference between DeepSTHF and the other two methods on 31 March 2020 was the largest, followed by that on 25 April 2018, while that on 9 May 2019 was the smallest. A main reason for this is that the test area for 31 March 2020 comprised many small leads that

were correctly classified by OriTHF and CubicSTHF, and the test area for 9 May 2019 was mainly consisted of a large lead network that was successfully extracted by all three methods, where only a few small lead networks were mapped by the DeepSTHF method (Fig. 13 g). Therefore, DeepSTHF can achieve more accurate results in areas comprised of leads with abundant widths and lengths. A large difference was observed for the estimated THF with the three methods from the simulated and real MODIS images on 31 March 2020; for example, the THF obtained by DeepSTHF from simulated and real MODIS images were 2.115×10^{11} W and 1.747×10^{11} W, respectively. This may have resulted from the temperature differences between MODIS and Landsat-8 data.

Table 4. Total estimated turbulent heat flux measured with the three methods on 25 April 2018, 9 May 2019, and 31 March 2020 (unit: W)

Method	25 April 2018	9 May 2019	31 March 2020
OriTHF	2.114×10^{11}	2.074×10^{11}	1.348×10^{11}
CubicSTHF	2.006×10^{11}	2.096×10^{11}	1.287×10^{11}
DeepSTHF	2.426×10^{11}	2.337×10^{11}	1.747×10^{11}

Note: OriTHF, the pixel-scale turbulent heat flux estimation method; DeepSTHF, the deep residual convolutional neural network-based sub-pixel heat flux estimation method; CubicSTHF, the bicubic interpolation-based sub-pixel scale heat flux estimation method.

5. Discussion

The experiments involving simulated and real data show that the DeepSTHF method could increase the accuracy of THF estimation when compared with the method that uses the original MODIS data (OriTHF), and the method based on conventional bicubic interpolation (CubicSTHF). In practice, however, three issues that are related to the performance of DeepSTHF should be considered including the CNN architecture, parameter settings, and the uncertainty of SR lead mapping.

5.1. The CNN architecture

The enhanced performance of DeepSTHF relative to CubicSTHF arose primarily from the ability of DeepSTHF to automatically learn the complicated nonlinear relationships between the coarse resolution SST image and corresponding fine resolution SST image and lead map with the CNN models. This study applied two CNNs with different architectures for the SST image SR and SR lead mapping, because a single CNN architecture cannot achieve the different objectives of the two CNN models. We tested the performance of the multi-level feature fusion residual CNN, which was used for lead mapping, on a MODIS SR SST image (Fig. 15). Although most the fine spatial information has been recovered (red rectangles in Fig. 15b), the retrieved surface temperature of lead pixels along boundaries were greater than those in the central regions and these were not visually continuous (red ellipses in Fig. 15b), which was unsatisfactory. This may have occurred because the multi-level feature fusion residual CNN mainly focused on the semantic information of lead nets through the down-sampling layers, which may have caused the loss of spatial texture of the input data to some extent. Additionally, the very deep CNN used for

SR SST has also been proved to be invalid in lead mapping because it could not identify any lead networks. A major reason for this is that it did not comprise any down-sampling layers and therefore the semantic information of lead nets can only be
525   extracted with some difficulty.

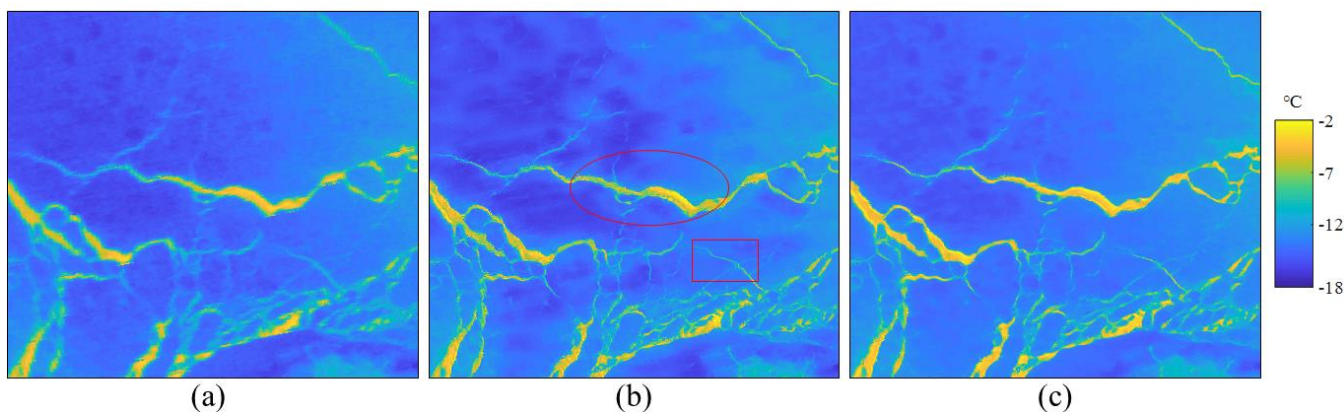



Figure 15. Results of moderate resolution imaging spectroradiometer super-resolution (SR) sea surface temperature imaging for data acquired on 31 March 2020 by: (a) the cubic interpolation method, (b) a convolutional neural network (CNN) model with the backbone of the CNN used in SR lead mapping, and (c) the very deep residual CNN model. The red rectangles represent the fine
530   **spatial pattern has been recovered by the multi-level feature fusion residual CNN. The red ellipses represent the recovered surface temperature of lead pixels by the multi-level feature fusion residual CNN that were not visually continuous.**

5.2. Parameter settings

Like many classic algorithms, the proposed DeepSTHF is not totally automatic; there are some customized parameters for the CNN models used in the DeepSTHF method. Overall, the batch size of the training samples, optimization method, and learning
535   rate should be set in advance. The optimization method (or optimizer) is a major approach used for training a CNN model. Recently, Stochastic Gradient Descent (SGD) and Adaptive Moment Estimation (Adam) optimizers have been widely used; they are both gradient-based methods that update the parameters of CNN along the direction that would minimize the error rate. The two optimizers were tested in the present study; the results show that the Adam algorithm had a much faster convergence speed and performed better than the SGD algorithm. As a result, Adam was selected as the optimization method
540   in the present study. Additionally, the exponential decay rates β_1 and β_2 of Adam were set to 0.9 and 0.999, respectively, which is typically recommended in practice (Reddi et al., 2019). The learning rate determines the step size at each iteration while moving toward a minimum of a loss function; a favorable value mainly depends on the training data set and the architecture of the CNN model. In this study, 10^{-4} was empirically set for the learning rate through substantial experiments. Batch size defines the number of training samples used in one iteration for the training process; Masters and Luschi (2018) suggested a
545   batch size between 2 and 32 because this size can provide stable convergence. In the present study, the batch size was set to 24 considering the trade-off between training speed and computational speed of the computer. The experiment showed that the DeepSTHF method was able to generate accurate subpixel THF data for data which were observed on different dates and covered different areas with these parameters. In fact, however, better parameters may be selected in the future and thus provide

more accurate results. Knowing how to acquire the best values for these parameters is a hot topic in deep learning and needs
550   further investigation.

5.3 Uncertainty of SR lead mapping

The main limitation of the DeepSTHF method when used for the estimation of THF is that it was limited by the accuracy SR
lead mapping. First, a large spatial resolution gap exists between Landsat-8 and MODIS surface temperature imagery, while
some very narrow lead nets in the fine resolution image (especially those with width < 5 pixels), were not mapped with
555   DeepSTHF. The reliability of DeepSTHF would decrease as the numbers of these very narrow leads increase. Second, we
assumed that the lead networks do not change between the overpass time of MODIS and Landsat-8 satellites and used the lead
maps produced from Landsat-8 data as the reference data in the MODIS imagery experiment; doing so may bring about some
errors if an abrupt change occurs in the ice pack.

6. Conclusions

560   The turbulent heat flux over leads is an important variable for climate studies in the Arctic which has been estimated using
remotely sensed data acquired from satellite sensors. Fine spatial resolution data is required for accurate calculation of the
THF, although sometimes it is of limit use relative to coarse resolution data for some reasons including data availability.
However, many mixed pixels along the edges of a lead will greatly decrease the accuracy of THF estimation when traditional
methods are used. This paper proposes a deep learning-based method to calculate THF over leads at a subpixel scale
565   (DeepSTHF) to address this problem. Specifically, two CNN models were first applied to generate a fine spatial resolution
surface temperature image and a corresponding fine resolution lead map was produced from a coarse resolution surface
temperature image; next, the fine spatial resolution data were then used for THF estimation.

The results of two experiments showed that the proposed DeepSTHF can model the spatial pattern and relationship between
coarse and fine resolution data quite well and achieve reliable results with a high level of accuracy. The main reason for the
570   good performance of the DeepSTHF method is the potential of CNN models to specify complex nonlinear relationships
between data. For SR SST image reconstruction, the bicubic interpolation-based method obtain the values of interpolated
pixels by linearly combining the neighboring pixels. The spatial textures between coarse and fine resolution pixels, however,
is not linear in some conditions (especially for pixels along lead boundaries). Therefore, the interpolated SST images
commonly lacked a fine spatial pattern; the same problem could be seen in the lead maps produced by the threshold method
575   since it is a pixel-based method. In contrast, the proposed CNN-based method learns the spatial patterns automatically from
existing data, and so achieves a more powerful SR of data. The proposed method, DeepSTHF, is promising for calculating the
THF over leads at a subpixel scale based on remotely sensed data.

References

580   Atkinson, P. M.: Downscaling in remote sensing, International Journal of Applied Earth Observation and Geoinformation, 22, 106-114, 10.1016/j.jag.2012.04.012, 2013.

Badrinarayanan, V., Kendall, A., and Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE transactions on pattern analysis and machine intelligence, 39, 2481-2495, 10.1109/TPAMI.2016.2644615, 2017.

585   Barber, D. G., Ehn, J. K., Pućko, M., Rysgaard, S., Deming, J. W., Bowman, J. S., Papakyriakou, T., Galley, R. J., and Søgaard, D. H.: Frost flowers on young Arctic sea ice: The climatic, chemical, and microbial significance of an emerging ice type, Journal of Geophysical Research: Atmospheres, 119, 11,593-511,612, 10.1002/2014jd021736, 2014.

Brodeau, L., Barnier, B., Gulev, S. K., and Woods, C.: Climatologically Significant Effects of Some Approximations in the Bulk Parameterizations of Turbulent Air–Sea Fluxes, Journal of Physical Oceanography, 47, 5-28, 10.1175/jpo-d-16-
590   0169.1, 2017.

Dong, C., Loy, C. C., He, K., and Tang, X.: Image Super-Resolution Using Deep Convolutional Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 10.1109/TPAMI.2015.2439281, 2014.

Ebert, E. E., and Curry, J. A.: An intermediate one-dimensional thermodynamic sea ice model for investigating ice-atmosphere interactions, Journal of Geophysical Research: Oceans, 98, 10085-10109, 10.1029/93jc00656, 1993.

595   Eppler, D. T., and Full, W. E.: Polynomial trend surface analysis applied to AVHRR images to improve definition of arctic leads, Remote Sensing of Environment, 40, 197-218, 10.1016/0034-4257(92)90003-3, 1992.

Eythorsson, D., Gardarsson, S. M., Ahmad, S. K., Hossain, F., and Nijssen, B.: Arctic climate and snow cover trends – Comparing Global Circulation Models with remote sensing observations, International Journal of Applied Earth Observation and Geoinformation, 80, 71-81, 10.1016/j.jag.2019.04.003, 2019.

600   Foody, G. M., Muslim, A. M., and Atkinson, P. M.: Super-resolution mapping of the waterline from remotely sensed data, International Journal of Remote Sensing, 26, 5381-5392, 10.1080/01431160500213292, 2005.

Foody, G. M., and Doan, H. T. X.: Variability in Soft Classification Prediction and its implications for Sub-pixel Scale Change Detection and Super Resolution Mapping, Photogrammetric Engineering & Remote Sensing, 73, 923-933, 10.14358/PERS.73.8.923, 2007.

605   Ge, Y., Li, S., and Lakhan, V. C.: Development and Testing of a Subpixel Mapping Algorithm, IEEE Transactions on Geoscience and Remote Sensing, 47, 2155-2164, 10.1109/TGRS.2008.2010863, 2009.

Ge, Y., Jin, Y., Stein, A., Chen, Y., Wang, J., Wang, J., Cheng, Q., Bai, H., Liu, M., and Atkinson, P. M.: Principles and methods of scaling geospatial Earth science data, Earth-Science Reviews, 197, 102897, 10.1016/j.earscirev.2019.102897, 2019.

610   Gerace, A., and Montanaro, M.: Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard Landsat 8, Remote Sensing of Environment, 191, 246-257, 10.1016/j.rse.2017.01.029, 2017.

Glasner, D., Bagon, S., and Irani, M.: Super-resolution from a single image, 2009 IEEE 12th International Conference on Computer Vision, 349-356, 2009.

Goosse, H., Campin, J.-M., Deleersnijder, E., Fichefet, T., Mathieu, P.-P., Maqueda, M. M., and Tartinville, B.: Description
615   of the CLIO model version 3.0, Institut d'Astronomie et de Géophysique Georges Lemaitre, Catholic University of Louvain, Belgium, 2001.

Hall, D., and Riggs, G.: Algorithm Theoretical Basis Document (ATBD) for the MODIS Snow and Sea Ice-Mapping Algorithms, 2001.

Jia, Y., Ge, Y., Chen, Y., Li, S., Heuvelink, G. B. M., and Ling, F.: Super-Resolution Land Cover Mapping Based on the
620   Convolutional Neural Network, Remote Sensing, 11, 1815, 10.3390/rs11151815, 2019.

Kingma, D., and Ba, J.: Adam: A Method for Stochastic Optimization, International Conference on Learning Representations, 2014.

Leach, R., and Sherlock, B.: Applications of super-resolution imaging in the field of surface topography measurement, Surface Topography: Metrology and Properties, 2, 023001, 10.1088/2051-672x/2/2/023001, 2013.

625 Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W.: Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 105-114, 2017.

Lewis, B. J., and Hutchings, J. K.: Leads and Associated Sea Ice Drift in the Beaufort Sea in Winter, Journal of Geophysical Research: Oceans, 124, 3411-3427, 10.1029/2018jc014898, 2019.

630 Lindsay, R. W., and Rothrock, D. A.: Arctic sea ice leads from advanced very high resolution radiometer images, Journal of Geophysical Research: Oceans, 100, 4533-4544, 10.1029/94jc02393, 1995.

Ling, F., Du, Y., Xiao, F., Xue, H., and Wu, S.: Super-resolution land-cover mapping using multiple sub-pixel shifted remotely sensed images, International Journal of Remote Sensing, 31, 5023-5040, 10.1080/01431160903252350, 2010.

Ling, F., Boyd, D., Ge, Y., Foody, G. M., Li, X., Wang, L., Zhang, Y., Shi, L., Shang, C., Li, X., and Du, Y.: Measuring River
635   Wetted Width From Remotely Sensed Imagery at the Subpixel Scale With a Deep Convolutional Neural Network, Water Resources Research, 55, 5631-5649, 10.1029/2018wr024136, 2019.

Ling, F., and Foody, G. M.: Super-resolution land cover mapping by deep learning, Remote Sensing Letters, 10, 598-606, 10.1080/2150704X.2019.1587196, 2019.

Lüpkes, C., Vihma, T., Birnbaum, G., and Wacker, U.: Influence of leads in sea ice on the temperature of the atmospheric
640   boundary layer during polar night, 35, 10.1029/2007gl032461, 2008.

Marcq, S., and Weiss, J.: Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean and the atmosphere, The Cryosphere, 6, 143-156, 10.5194/tc-6-143-2012, 2012.

Maykut, G. A.: Energy exchange over young sea ice in the central Arctic, Journal of Geophysical Research: Oceans, 83, 3646-3658, 10.1029/JC083iC07p03646, 1978.

645   Montanaro, M., Gerace, A., and Rohrbach, S.: Toward an operational stray light correction for the Landsat 8 Thermal Infrared Sensor, Appl. Opt., 54, 3963-3978, 10.1364/AO.54.003963, 2015.

Noh, H., Hong, S., and Han, B.: Learning Deconvolution Network for Semantic Segmentation, 2015 IEEE International Conference on Computer Vision (ICCV), 1520-1528, 2015.

Qu, M., Pang, X., Zhao, X., Zhang, J., Ji, Q., and Fan, P.: Estimation of turbulent heat flux over leads using satellite thermal

650   images, The Cryosphere, 13, 1565-1582, 10.5194/tc-13-1565-2019, 2019.

Röhrs, J., and Kaleschke, L.: An algorithm to detect sea ice leads by using AMSR-E passive microwave imagery, The Cryosphere, 6, 343-352, 10.5194/tc-6-343-2012, 2012.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention, 234-241, 2015.

655   Tschudi, M. A., Curry, J. A., and Maslanik, J. A.: Characterization of springtime leads in the Beaufort/Chukchi Seas from airborne and satellite observations during FIRE/SHEBA, Journal of Geophysical Research: Oceans, 107, SHE 9-1-SHE 9-14, 10.1029/2000jc000541, 2002.

Wang, Q., Shi, W., and Atkinson, P. M.: Sub-pixel mapping of remote sensing images based on radial basis function interpolation, ISPRS Journal of Photogrammetry and Remote Sensing, 92, 1-15, 10.1016/j.isprsjprs.2014.02.012, 2014.

660   Wang, Z., Chen, J., and Hoi, S. C. H.: Deep Learning for Image Super-resolution: A Survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1-1, 10.1109/TPAMI.2020.2982166, 2020.

Willmes, S., and Heinemann, G.: Pan-Arctic lead detection from MODIS thermal infrared imagery, Annals of Glaciology, 56, 29-37, 10.3189/2015AoG69A615, 2015.

Zhang, S.: Application of Super-Resolution Image Reconstruction to Digital Holography, EURASIP Journal on Advances in

665   Signal Processing, 2006, 090358, 10.1155/ASP/2006/90358, 2006.

Zhang, Y., Tian, Y., Kong, Y., Zhong, B., and Fu, Y.: Residual dense network for image super-resolution, Proceedings of the IEEE conference on computer vision and pattern recognition, 2472-2481, 2018.

Zhao, W., Wu, H., Yin, G., and Duan, S.-B.: Normalization of the temporal effect on the MODIS land surface temperature product using random forest regression, ISPRS Journal of Photogrammetry and Remote Sensing, 152, 109-118,

670   10.1016/j.isprsjprs.2019.04.008, 2019.

Zhong, Y., and Zhang, L.: Sub-pixel mapping based on artificial immune systems for remote sensing imagery, Pattern Recognition, 46, 2902-2926, 10.1016/j.patcog.2013.04.009, 2013.