



Deep learning based retrieval algorithm for Arctic sea ice concentration from AMSR2 passive microwave and MODIS optical data



Junhwa Chi^{a,*}, Hyun-cheol Kim^a, Sungjae Lee^a, Melba M. Crawford^b

^a Unit of Arctic Sea-Ice Prediction, Korea Polar Research Institute, Incheon, Republic of Korea

^b School of Civil Engineering, Purdue University, West Lafayette, IN, USA

ARTICLE INFO

Edited by Menghua Wang

Keywords:

Endmember extraction
Machine learning
Neural network
Spectral mixture analysis
Spectral unmixing

ABSTRACT

This study applies deep learning (DL) to retrieve Arctic sea ice concentration (SIC) from AMSR2 data. MODIS-derived SICs are calculated based on spectral unmixing with a new ice/water endmember extraction algorithm that exploits global/local representatives, and then used to train a DL network with AMSR2 data. The resulting SIC maps outperform popular SIC products both regionally and globally. The RMSE of the proposed DL model is 5.19, whereas those of the widely used Bootstrap and ASI-based SIC images are 6.54 and 7.38, respectively, with respect to MODIS-derived SICs at global scale. In particular, our proposed method better describes regions of low-SIC and melting ice in summer, which are generally difficult-to-estimate. As the DL-based model consistently generates accurate SIC values that are not time- or region-dependent, it is considered to be an operational system. Additionally, our SICs can be used to generate initial conditions facilitating development of more accurate climate models.

1. Introduction

Sea ice plays a vital role in the global climate system as the high albedo of ice and snow make them the biggest reflectors of solar radiation. Moreover, sea ice reduces heat transfer between the ocean and the atmosphere (Vinnikov et al., 1999; Liu et al., 2004; Spreen et al., 2008). Thinning ice in polar regions accelerates global warming, which is one of the most serious issues currently facing our planet (Stroeve et al., 2012; Ivanova et al., 2014). However, the high variability and uncertainty of sea ice over both time and space limit the accurate retrieval of sea ice parameters for climate research and data assimilation in prediction models (Cho and Naoki, 2015; Karvonen, 2017).

In 1972, information about the distribution of sea ice was first retrieved by a passive microwave (PMW) sensor, the ESMR (Electrically Scanning Microwave Radiometer) on the Nimbus-5. Since the launch of the SMMR (Scanning Multichannel Microwave Radiometer) on board Nimbus-7 in 1978, global sea ice observations have been continuously acquired from various PMW radiometers such as the SSM/Is (Special Sensor Microwave/Imagers) and SSMIS (Special Sensor Microwave Imager/Sounder) on the DMSP (Defense Meteorological Satellite Program) satellites, AMSR (Advanced Microwave Scanning Radiometer) on the ADEOS (Advanced Earth Observing Satellite) II, AMSR-E on Aqua, and AMSR2 on the GCOM-W1 (Global Change Observation Mission 1st-Water) (Spreen et al., 2008; Ivanova et al.,

2014; Cho and Naoki, 2015; Cavalieri et al., 1996; Comiso, 1995). The sea ice concentration (SIC), which is the fraction of a given area covered with sea ice, is one of the most important polar ocean parameters, and is often generated from PMW data (Comiso, 1995). Because of the importance of SIC information in both industry and research, many retrieval algorithms for estimating more accurate SIC values have been proposed in recent decades (Ivanova et al., 2014; Andersen et al., 2006). The SIC values generated using frequencies around 19 and 37 GHz have a grid size of 25 km, but recent advances in sensor technology and new retrieval algorithms using high-frequency channels have improved the grid size of 3.125 km (Spreen et al., 2008).

Because of the large emissivity differences between water and ice, most SIC retrieval algorithms employ linear combinations of brightness temperatures (TBs) at different frequencies and polarizations to distinguish open water from first- and multi-year ice (Cho and Naoki, 2015; Svendsen et al., 1983; Cavalieri et al., 1984; Swift and Cavalieri, 1985). Cavalieri et al. (1984) developed the first SIC retrieval algorithm from Nimbus-7 SMMR data, and this technique was later applied by the NASA Sea Ice Algorithm Working Group to SSM/I data. The NASA Team (NT) algorithm uses three SSM/I channels (19.4 GHz horizontal and vertical polarizations, and 37 GHz vertical polarization) and comprises three parts: 1) calculation of the polarization and spectral gradient ratios; 2) estimation of first- and multi-year ice concentrations using coefficients derived from tie-points over known areas; and 3) use

* Corresponding author.

E-mail addresses: jhchi@kopri.re.kr (J. Chi), kimhc@kopri.re.kr (H.-c. Kim), sungjae@kopri.re.kr (S. Lee), mcrawford@purdue.edu (M.M. Crawford).

<https://doi.org/10.1016/j.rse.2019.05.023>

Received 17 August 2018; Received in revised form 15 May 2019; Accepted 15 May 2019

Available online 04 June 2019

0034-4257/ © 2019 Elsevier Inc. All rights reserved.

of weather filters to remove anomalous SIC values over the open ocean caused by weather effects (Cavalieri et al., 1996; Swift and Cavalieri, 1985; Cavalieri, 1994; Gloersen and Cavalieri, 1986). Although the NT algorithm is widely used, it often underestimates the SIC values because of layering in the snow and young/thin ice types (Andersen et al., 2007). The NT algorithm was extended to develop the NT2 algorithm using high-frequency channels (85 GHz for SSM/Is; 89 GHz for AMSR), which allowed for atmospheric effects over open water to be investigated (Markus and Cavalieri, 2000; Markus and Cavalieri, 2009). Another popular algorithm is the Bootstrap (BT) developed by Comiso (Comiso, 1986) for SMMR data, and then enhanced for new sensors such as SSM/I and AMSR (Cho and Naoki, 2015; Comiso, 1995). The BT algorithm assumes that there are two surface types: consolidated ice regions (where the SIC is near 100%) and ice-free regions. It exploits the distribution of unique clusters of TB in multichannel data, and tends to be seasonally stable because it is less sensitive to thin ice and layering effects with polarization (Comiso, 1995; Andersen et al., 2007; Comiso, 1986). Svendsen et al. (1987) proposed taking advantage of the high-spatial resolution of the 85 GHz channels of SSM/I. It was later extended by Kaleschke et al. (2001) into the ARTIST (Arctic Radiation and Turbulence Interaction Study) Sea Ice (ASI) algorithm. The ASI algorithm combines the model for SIC retrieval from high-frequency channels (85 GHz for SSM/I; 89 GHz for AMSR) proposed by Svendsen et al. (1987) with an ocean mask derived from low-frequency channels (19, 22, and 37 GHz for SSM/I; 18, 23, and 37 GHz for AMSR) using the NT algorithm and weather filters (Spreen et al., 2008; Kern et al., 2003). Although high-frequency data are sensitive to atmospheric influences, they result in higher spatial resolution products.

Comparative studies to quantify the differences between the various SIC retrieval algorithms have been conducted by Emery et al. (1994), Comiso et al. (1997), Andersen et al. (2006), Ivanova et al. (2014, 2015), Kern et al. (2016), and Han and Kim (2018). In particular, Ivanova et al. (2015) compared 13 SIC retrieval algorithms over low and high SICs. Their study showed that there is no one superior algorithm, and a combination of algorithms takes advantages of each algorithm. Additionally, implementations of dynamic tie-points and atmospheric correction of input TBs reduced systematic bias and seasonal variability in SIC. In Han and Kim (2018), they addressed the fact that SIC retrieval algorithms exhibit regional variations and are inaccurate in summer using high-resolution KOMPSAT (Korea Multipurpose Satellite)-5 SAR (Synthetic Aperture Radar) images. Despite many comparisons and assessments, concluding which algorithms are superior is difficult because of their different characteristics and the inherent difficulties in obtaining true SICs. The differences among retrieval algorithms are attributed to the selection of different frequency and polarization modes, tie-points, weather filters, and land-ocean spillover masks (Ivanova et al., 2014; Andersen et al., 2006; Emery et al., 1994; Comiso et al., 1997; Ivanova et al., 2015; Kern et al., 2016; Han and Kim, 2018). The absolute error of such algorithms is mainly due to sensitivities to changes in the emissivity and temperature of the surface, effect of atmospheric constituents, and melt ponds on Arctic summer sea ice (Ivanova et al., 2015). Thus, although SIC estimates from PMW data contain significant uncertainties, satellite-based SIC data remain the best means of providing invaluable and continuous information for many commercial, navigational, and operational applications over extended areas.

In recent years, interest in deep learning (DL) methodologies has increased dramatically in the machine learning community, helped by advances in computer hardware, algorithmic techniques, and the era of big data. DL hierarchically learns representative features from large, complex datasets. Compared to the shallow architectures employed by artificial neural networks (ANNs) in the 1980s, the ability of DL to use larger and deeper networks has promoted rapid growth in its use across diverse fields of study (Zhang et al., 2016; LeCun et al., 2015). With increasing quantities of remote sensing (RS) data available from multiple sources, deep architectures have also led to a proliferation of

promising results in classification, regression, and feature extraction tasks (Hinton and Salakhutdinov, 2006; Zuo and Wang, 2014; Zou et al., 2015; Chen et al., 2014; Hu et al., 2015a; Hu et al., 2015b; Chi and Kim, 2017). Chen et al. (Chen et al., 2014) proposed a new classification approach for hyperspectral data using autoencoders (AEs). In this study, single-layer and multi-layer stacked AEs were exploited to extract shallow (abstract) and deep (complex) features, resulting in improved accuracy compared to traditional methods. Hu et al. (2015a) employed deep convolutional neural networks (CNNs) to classify hyperspectral RS data directly in the spectral domain, and investigated how to transfer features for classification tasks with high-resolution RS data using CNNs (Hu et al., 2015b). In a technical tutorial by Zhang et al. (2016), various DL studies in RS data analysis were extensively reviewed from the perspectives of image preprocessing, pixel-based classification, target recognition, and scene understanding, and the nature of future challenges was discussed. Several DL studies related to Arctic sea ice have recently been conducted. Chi and Kim (2017) proposed a DL-based SIC prediction method using long-term SIC time series. While traditional SIC predictions have been performed using environmental parameters from complex numerical, statistical, or ensemble models, Chi and Kim used SIC data without incorporating any physical data, and obtained comparable results to other prediction models. Regional SIC estimation using a DL approach was considered by Karvonen (2017), who successfully trained a multilayer perceptron DL model using a combination of Sentinel-1 SAR and AMSR2 data in the Baltic Sea. Although the Karvonen's study applied a DL architecture to calculate SICs and yielded better estimates, it is not a global method for multi-temporal datasets over the entire Arctic ocean.

In this paper, we propose an accurate and robust DL-based Arctic SIC retrieval algorithm incorporating the TB data from the AMSR2 PMW sensor and high-resolution MODIS (Moderate Resolution Imaging Spectroradiometer) images for practical, operational use. To develop more accurate supervised models, obtaining sufficient true labeled data of suitable quality for use as a reference is critical. This study is composed of two phases: 1) the development of an accurate labeled dataset; and 2) construction of a DL architecture for pan-Arctic SIC retrieval. The first part of this study constructs reference labels by calculating accurate SIC values from MODIS images corresponding to the area of a pixel in low-resolution AMSR2 data. The retrieval of SIC values from high-resolution images is often used to evaluate low-resolution SIC data using a threshold method from single- or multi-band images and classification. However, determining the appropriate threshold value is problematic for multi-temporal images because of the subtle spectral differences inherent in multi-temporal images; classification methods that determine the class membership of each pixel may result in the overestimation or underestimation of continuous SIC values. To compute more accurate, continuous, and quantitative SIC values at the sub-pixel level of MODIS data, we propose a SIC retrieval approach based on spectral mixture analysis with a new endmember extraction algorithm appropriate for the identification of pure sea ice and water components in multispectral data. We then show that the proposed method outperforms pixel classification-based SIC retrieval approaches. In the second part of this study, we develop the network topology of a DL model for pan-Arctic SIC retrievals for operational purposes, and conduct quantitative and qualitative comparisons with popular BT and ASI-based SIC products at global and local scales, as well as for long-term evaluations.

2. Datasets

The primary goal of this study is to develop a DL-based pan-Arctic SIC retrieval model using AMSR2 TB data. Obtaining accurate reference SIC labels, which is the most critical task in successful model development, can be accomplished from field observations, ice charts, and high-resolution images. Field observations generally have the most accurate references, but are limited and costly over large regions such

as Arctic oceans. The ice charts provided by the Russian Arctic and Antarctic Research Institute (AARI) may provide useful sea ice information over extended areas periodically, but low-spatial/temporal resolutions and high levels of uncertainty limit their use as true labels. For these reasons, satellite RS data are commonly used to obtain reference SIC data and evaluate SIC retrieval algorithms because of their extensive coverage. In this study, we use 500 m MODIS surface reflectance images, which are collected daily over the Arctic region, except in winter, to compute accurate SIC values for the reference labels.

2.1. AMSR2 sensor data

Following the AMSR-E on the Aqua satellite launched in 2002, AMSR2 was successfully placed in orbit by the Japan Aerospace Exploration Agency (JAXA) in May 2012 aboard the GCOM-W1 satellite, and began its operational observations in July. GCOM-W1's sun-synchronous, near-polar circular orbit is timed to cross the equator from north to south (descending node) at approximately 01:30 local time, and from south to north (ascending node) at approximately 13:30 local time. AMSR2 acquires measurements at seven different frequencies (6.9, 7.3, 10.6, 18.7, 23.8, 36.5, and 89 GHz) with both horizontal (H) and vertical (V) polarization (Okuyama and Imaoka, 2015). Compared to the past-generation SSM/I sensor, AMSR2 offers a significant improvement in spatial resolution because of its large antenna size, with a spatial resolution at 89 GHz of approximately $3 \times 5 \text{ km}^2$ compared with approximately $14 \times 16 \text{ km}^2$ for SSM/I at 85 GHz. This enhanced spatial resolution allows AMSR2-based SIC products to observe more detailed sea ice distributions than SSM/I. The spatial resolution of SIC products depends on the channel frequency, but AMSR2 generally offers higher resolution products than SSM/I (Cho and Naoki, 2015; Karvonen, 2017; Han and Kim, 2018; Okuyama and Imaoka, 2015).

Most SIC retrieval algorithms for AMSR2 use a subset of PMW channel frequencies and polarizations. For example, the NT algorithm uses 18.7 V, 18.7 H, and 36.5 V channels (with NT2 additionally using 85 V and 85 H); the BT algorithm uses 18.7 V, 36.5 H, and 36.5 V; the ASI algorithm uses 89 H and 89 V. The 23.8 GHz channel is often used for weather filters. In this study, all channels of the AMSR2 level 3 TB product provided by JAXA, which is daily gridded data at a 10 km spatial resolution in the polar stereographic projection, are used as input data for a DL-based SIC retrieval model. It is unclear whether exploiting other low-frequency channels that are not currently used in existing SIC retrieval algorithms could significantly contribute to SIC estimations. However, the use of more attributes and properties (e.g., frequency and polarization in our data) generally yields better outcomes in DL studies, as novel topologies, coupled with computational advances enable more complex and abstract decisions than in the past (LeCun et al., 2015; Chi and Kim, 2017). As baseline SIC maps, BT and ASI-based SIC images provided by the University of Bremen (available at seaice.uni-bremen.de) are used. Both products include weather filters (Spreen et al., 2008).

2.2. MODIS sensor data

MODIS is designed to provide measurements of large-scale global dynamics in oceans, on land, and in the atmosphere. Its wide swath (approximately 2330 km) and high-temporal/spectral resolution are useful for many products and applications in diverse fields. The images from the two MODIS instruments on the Terra and Aqua satellites observe the same regions three hours apart and both have sun-synchronous, near-polar circular orbits. Terra's orbit is timed to cross the equator from north to south at approximately 10:30 local time, Aqua is timed to cross the equator from south to north at approximately 13:30 local time. In particular, Aqua passes the same orbital track 4 min after the observation by GCOM-W1. Unfortunately, due to the failure of 70% of the band-6 detectors on MODIS Aqua, which is the important band for discriminating sea ice and clouds (Gladkova et al., 2012), this study

uses MODIS Terra data to create reference SIC values for DL model training and evaluation of AMSR2-retrieved SIC values. Although SAR has the advantage of acquiring images regardless of weather and season, MODIS can observe the entire Earth every 1–2 days, except in winter, resulting in spatially well-distributed labeled data. Seven 500 m surface reflectance (MOD09GA) bands centered at 648, 858, 470, 555, 1240, 1640, and 2130 nm (bands 1–7, respectively) are used to compute the reference SIC values for model training and evaluation. This product is an estimate of the surface spectral reflectance that is corrected for the effects of atmospheric gases and aerosols at ground level based on the assumption that there is no atmospheric scattering or absorption (Vermote et al., 2011). The data are downloaded from the Level-1 and Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC), and processed using the MODIS Reprojection Tool to create daily mosaic images (Vermote and Wolfe, 2015).

3. Methodology

3.1. Limitations of classification for SIC estimates and ice mixtures at sub-pixel level

SIC is defined as the proportion of sea ice over a given area of the ocean. The popular data analysis technique of classification aims to categorize all pixels in an image into discrete, non-overlapping thematic categories according to algorithm specific rules.

As shown in Fig. 1, for example, let S be the corresponding SIC value over an area of $4 \times 4 \text{ km}^2$. Assuming that eight sea ice patches of $0.8 \times 0.8 \text{ km}^2$ are located in an image with 1 km spatial resolution, as shown in Fig. 1(a), S is 32% (i.e., $(0.8 \text{ km} \times 0.8 \text{ km} \times 8) / (4 \text{ km} \times 4 \text{ km}) = 0.32$). Applying a simple classifier, which determines the class membership based on the proportion of sea ice area in a given $1 \text{ km} \times 1 \text{ km}$ pixel, to this image to classify it as sea ice or water and then compute S for the true SIC labels in the $4 \text{ km} \times 4 \text{ km}$ area. For example, in Fig. 1(b), S is 50% (i.e., eight sea ice pixels out of 16 total pixels). Assume that the same sea ice patches moved to different locations in the same 16 km^2 area, as shown in Fig. 1(c). The S is still 32%, but the classification for S would be 0% as no pixels are classified as sea ice, as shown in Fig. 1(d). Although this is a schematic example that demonstrates the extreme cases, similar misclassifications can be found in real images, as shown in Fig. 2.

Fig. 2 is composed of a MODIS color composite image (R/G/B: bands 1/6/7) over the Fram Strait on August 1, 2017 and a classification result map using a CNN. To generate the classification map, we first remove land areas using a Land/Water Mask (MOD44W) Version 6 data product (Carroll et al., 2017), and then remove clouds using a very strict threshold ($\rho_i \geq 0.1; \forall \rho_i: i = 1, \dots, 7$, where ρ_i is the spectral reflectance of the i -th MODIS band). Identifying the pure water and ice pixels is generally simple because of the high spectral differences of these classes in visible and near-infrared wavelengths. However, determining the class membership of thin ice such as first-year ice or nilas, and melt ponds is often difficult due to the mixtures of water and ice signals. The classification results often depend on the classifier being used; in this study, a CNN is used as it is a well-suited DL strategy for the image classification tasks (Zhang et al., 2016). In Fig. 2(b), the pixels associated with land, cloud, water, and sea ice are shown in dark gray, light gray, dark blue, and dark red, respectively. As shown in Fig. 2, many pixels that may be associated with low-concentration ice or melting ice (rectangles) are classified as water, resulting in significant underestimations when computing the reference SIC labels. In other cases, some pixels in low-ice concentration zones, for which the SIC values are obviously different from old (multi-year) or consolidated ice pixels, are classified into the same class as the old ice (circles). These overestimations degrade SIC estimates derived from high-resolution images. Such under- and overestimates may be transferred to the DL model training associated with AMSR2 TB data, eventually affecting the

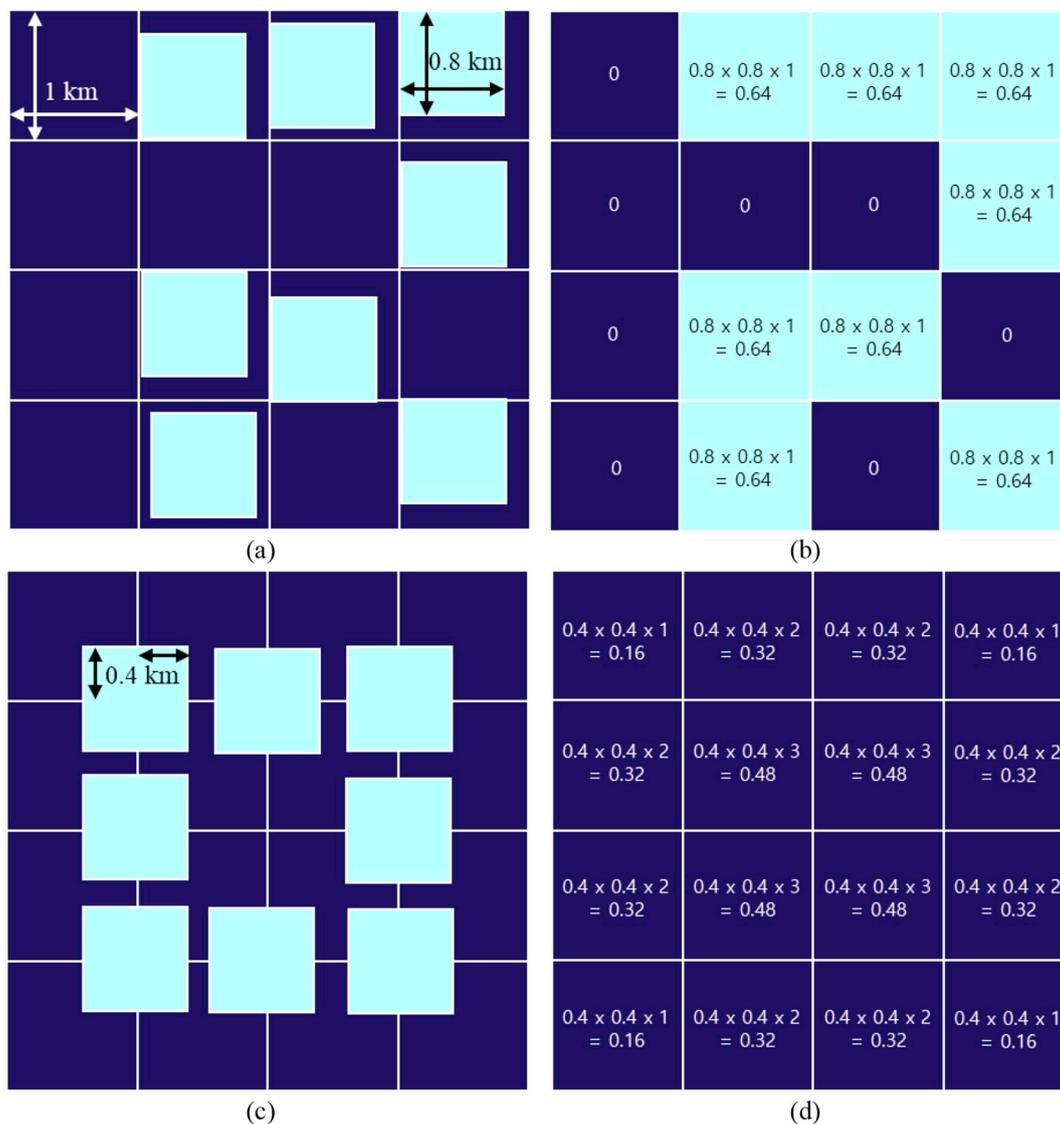


Fig. 1. Diagrams showing the differences in calculations of sea ice concentration between practical and classification results. (a) Case 1: Distributions of sea ice patches in a $4 \times 4 \text{ km}^2$ area. (b) Classification result of Case 1 for a 1 km spatial resolution image with area proportion-based classifier. (c) Case 2: Distributions of sea ice patches in the $4 \times 4 \text{ km}^2$ area. (d) Classification result of Case 2 for a 1 km spatial resolution image with area proportion-based classifier. The dark blue and cyan pixels represent ocean and sea ice pixels, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

accuracy of SIC estimates.

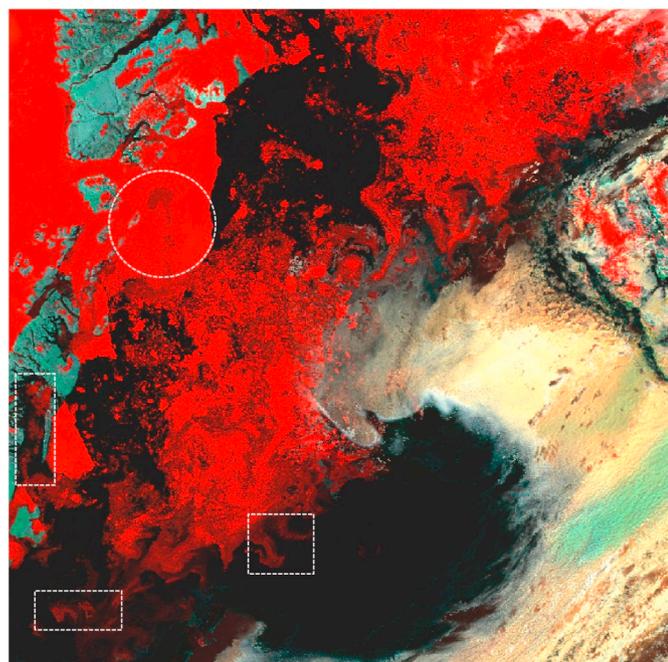
In medium-spatial-resolution RS imagery, most pixels are a mixture of more than one spectral signature owing to insufficient spatial resolution, the existence of intimate mixtures at a microscopic scale, and multiple interference factors by atmospheric attenuation (Chi and Crawford, 2014; Keshava and Mustard, 2002). In sea ice images, low-spatial resolution implies that real mixtures of ice and water can be found in one pixel area, and melting or low-concentration ice, as shown in Fig. 2, may be related to intimate mixtures. This problem is difficult to resolve by simply increasing the spatial resolution. Therefore, as shown in Fig. 3, quantification of the ice and water signals associated with each pixel is necessary for calculating more accurate SIC labels than classification-based calculations, which can then be used as training data for the DL model.

3.2. Spectral mixture analysis

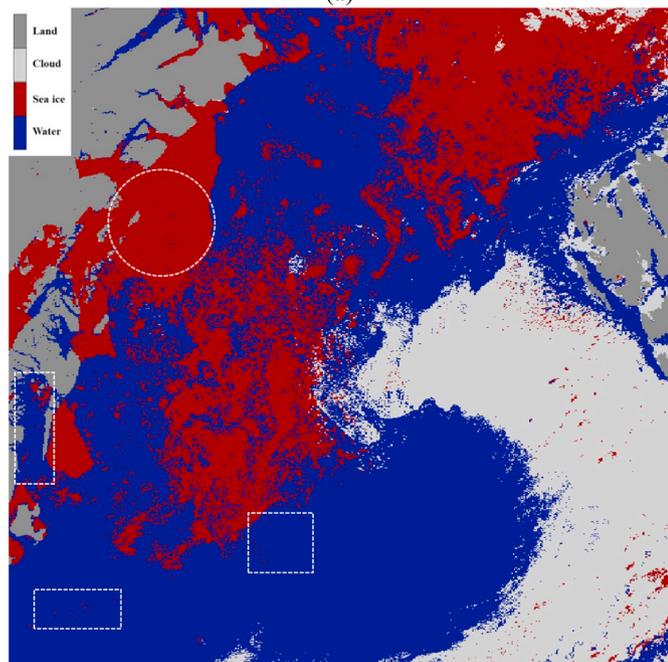
Spectral mixture analysis (SMA), also known as spectral unmixing, has been widely studied as a hyperspectral RS data analysis technique

for addressing the mixed-pixel issue. SMA decomposes a mixed pixel into a collection of individual spectral signals at sub-pixel levels (Keshava and Mustard, 2002), and usually comprises two steps. 1) Finding the spectrally pure or extreme signatures, referred to as “endmembers,” that can be used to “unmix” the remaining mixed pixels. 2) Expressing each pixel in the image as linear combinations of the endmembers, and computing the corresponding fractional abundances related to the “physical” quantity of the endmembers. Endmembers normally correspond to familiar spectrally homogeneous objects in the scene.

In this study, linear unmixing is employed because it is robust and easy to implement, where solving mixing problems for ice and water is relatively simple. Based on the assumption that endmembers are linearly independent and the pixels in the image lie in linear spaces, let \mathbf{x} be a D -dimensional vector of spectral reflectance. Then, \mathbf{x} can be modeled in terms of a linear combination of q endmember vectors $[\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_q]$ using $\mathbf{x} = \sum_{i=1}^q a_i \mathbf{e}_i$, where a_i is a fractional abundance of the endmember vector \mathbf{e}_i . Two constraints are imposed to estimate physically meaningful abundances (Heinz and Chang, 2001): 1) nonnegativity



(a)



(b)

Fig. 2. (a) Enhanced MODIS 1/6/7 band composite of the Fram Strait on August 1, 2017, (b) Classification result (land: dark gray; cloud: light gray; sea ice: dark red; water: dark blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

($a_i \geq 0; \forall a_i; 1 \leq i \leq q$) and 2) sum-to-one ($\sum_i^q a_i = 1$).

As the identification of an appropriate set of endmembers for the modeling of mixed pixel spectra is the most critical task in linear SMA, endmember extraction (EE) algorithms have been extensively investigated in the SMA community (Plaza et al., 2004). The EE algorithm used in this study is unsupervised (fully automated) and robust to anomalous pixels. Though the MODIS images used to form the SIC labels in this study are acquired daily, the number of available pixels is limited because of frequent high cloud cover. As the labels from MODIS images must be spatially and temporally well distributed, and there should be sufficient labels to develop a robust and operational model,

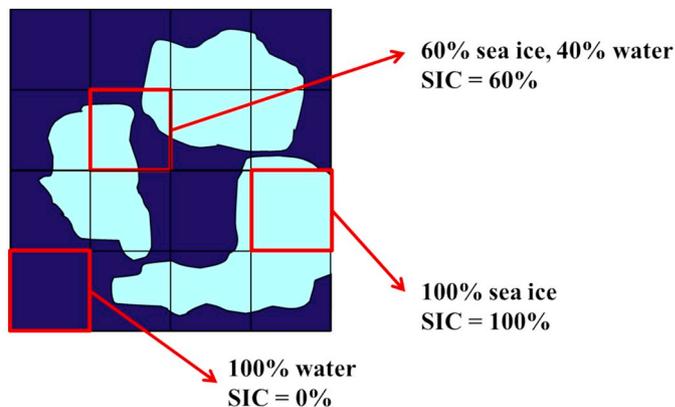


Fig. 3. Examples of ice and water mixtures at the sub-pixel level (SIC: sea ice concentration).

the collection of SIC labels from many optical images should be fully automated. Although seeking representative pixels that may be associated with extreme signals is often acceptable in a hyperspectral image, in our case of solving the ice/water endmember problem in multispectral data, selecting extreme signals can result in significant underestimations of the SIC.

Among the various EE algorithms, the popular N-finder algorithm (N-FINDR) exploits convex geometry, which is a straightforward concept in nature (Winter, 1999). Because of its popularity, there are many variations of N-FINDR. The goal of N-FINDR is to find the simplex with the maximal volume spanned by q endmembers embedded in a data space of lower dimension. First, the original data points are mapped into a lower $(q-1)$ -dimensional space using a feature extraction method; then, the simplex volume formed by randomly selected seed points is computed based on the matrix $M = \begin{bmatrix} 1 & 1 & \dots & 1 \\ e_1 & e_2 & \dots & e_q \end{bmatrix}$. The volume of the resulting simplex spanned by the endmembers $[e_1, e_2, \dots, e_q]$ is proportional to the determinant of M : $V(M) = \frac{1}{(q-1)!} \text{abs}(\det(M))$. The volume of the simplex is computed for every pixel in each endmember position by replacing the pixel and recalculating the volume when a larger simplex is found. Typically, N-FINDR does not determine the global maximum in general unmixing problems by iteratively testing different sets of initial endmembers. In our case of solving a simple EE problem with multispectral data, the effect of extreme pixels may not be mitigated. This is discussed further in the next section.

3.3. Global and local landmark endmember extraction algorithm

A new EE algorithm is proposed using the global and local landmark (GLL) approach to mitigate the impact of the global and local maxima by selecting the representative pixels as landmarks. The landmarks are a small number of points that are spectrally and spatially distinguishable pixels and are used as endmember candidates. The GLL algorithm is composed of three steps: 1) performing singular value decomposition (SVD) to determine a set of eigenvectors of image subsets; 2) determining endmember candidates (landmarks) by projecting the image data onto the eigenvectors; and 3) averaging the landmarks according to their spectral similarity. The proposed algorithm starts with a full image, and then successively divides it into smaller subsets to detect pixels that contribute to local spectral variability. The details are as follows.

SVD is a very efficient projection technique commonly used in RS to obtain a set of eigenvectors that explain most of the spectral variability of the data. The SVD of X can be defined by $X = USV^T$, where V^T contains the unit row eigenvectors of $X^T X$ sorted in order of descending significances, S is of the form $\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}$, C is diagonal, the square roots of the eigenvalues of $X^T X$ are sorted in descending value, and U denotes a

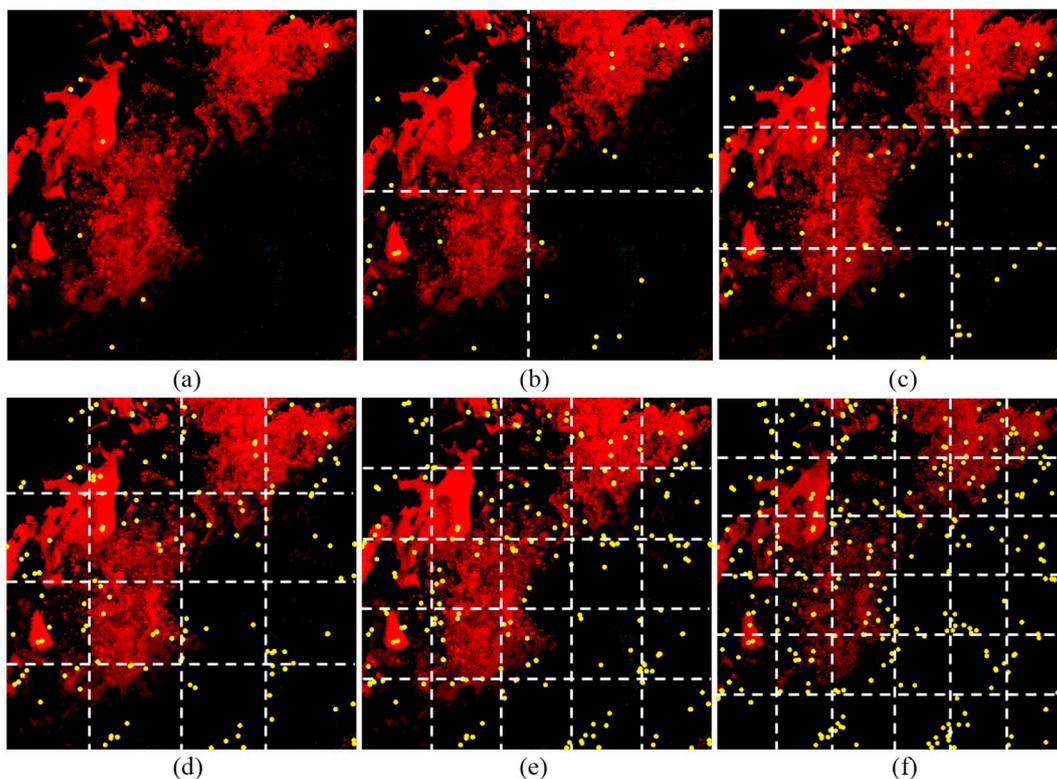


Fig. 4. Examples of spectrally and spatially meaningful pixels, referred to as landmarks (yellow dots) from different subset-sizes: (a) 1×1 , (b) 2×2 , (c) 3×3 , (d) 4×4 , (e) 5×5 , and (f) 6×6 from an enhanced MODIS 1/6/7 band composite image of the Fram Strait on August 1, 2017. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

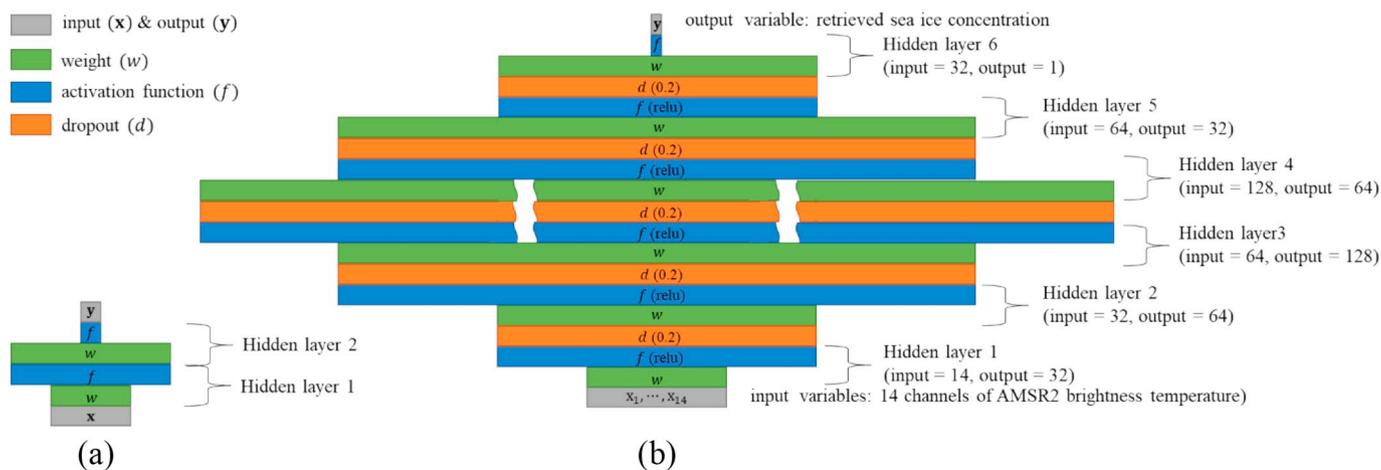


Fig. 5. Architectures of a multilayer perceptron deep learning model. (a) Simple model architecture with two hidden layers. (b) Tuned model topology for Arctic sea ice concentration retrieval.

matrix containing the unit column eigenvectors of XX^T (Danaher and Omngain, 1992).

The proposed GLL algorithm uses sequential, non-overlapping, equally sized subsets of the image (see Fig. 4). The eigenvectors accounting for 99% of the total spectral variance are retained from each image subset and compiled into a single eigenvector matrix M . These vectors represent high-spectral-variability landmarks in the local spatial subsets. In the case of larger subsets, landmarks might have relatively high spectral variability, indicating that they are useful for identifying strong signals. The use of smaller subsets may lead to more representative landmarks for local areas that cannot be detected from the global image, although a large number of eigenvectors is ultimately

retained.

The projection P of the entire set of data points onto one of the eigenvectors M is expressed by $P = M^T X$. The pixels lying at either extreme of the projection are more likely to be good landmarks, and these are identified from each subset. As the number of pixels of the subsets decreases (i.e., the number of the subsets increases), the number of landmark pixels increases, indicating that the spectral redundancy will also increase (see Fig. 4).

In the final combining step, this large number of redundant data points in the local images may play a role in mitigating the impact of the extreme signals detected in the global images by averaging landmarks depending on their spectral similarity.

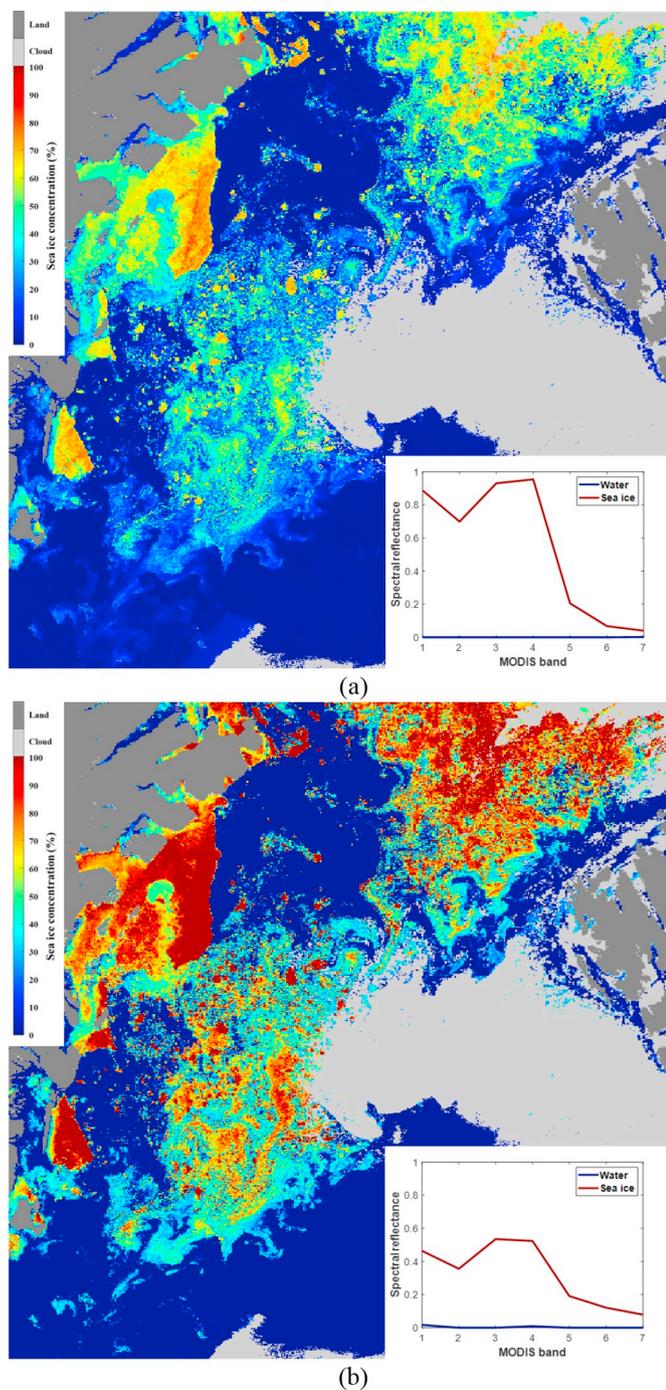


Fig. 6. Identified endmembers and sea ice concentration maps derived from fractional abundance maps using (a) N-FINDER (N-finder) and (b) GLL (global and local landmark) endmember extraction algorithms.

3.4. Deep learning model

Deep learning is a fast-growing subfield of machine learning based on artificial neural network, and its application in various fields has been enabled by the growing availability of big data and solutions to the problems of overfitting and high computational demand. While traditional ANN has only one hidden layer, DL may have more than two hidden layers in the network. The first hidden layer of the DL architecture makes simple decisions, such as the weighting of the input variables, while the deeper hidden layers engage in more complicated decision-making (LeCun et al., 2015). Among the many types of DL models, in this study, a multilayer perceptron (MLP), also known as a

“feedforward neural network”, is used to develop a DL-based regression model for SIC retrievals from PMW data. The differences between the outputs of an MLP and the desired values are used to iteratively update the weights in the network, and then the MLP eventually estimates the optimal results.

Fig. 5(a) illustrates the architecture of a simple MLP network with two hidden layers for learning a mapping function using the hierarchical or multilayered structure of the networks. A neuron is a computational unit that has weighted input signals and produces output signals using an activation function. The neurons are then arranged into a network topology. A simple neural network comprises input (visible), hidden, and output layers. The input layer is the bottom layer that is exposed to the network (x in Fig. 5(a)). After the input layer, several hidden layers that are not directly exposed to the network play a role in solving problems, such as using the weight (w) and activation function (f) in Fig. 5(a). The final layer, called the output layer, provides the prediction values of the resulting variables (y in Fig. 5(a)). To learn the representation in datasets and predict the output variable, developing an appropriate DL network topology through suitable parameter configurations is critical. Although there are many options, some preferred configurations, such as small random numbers for weight initialization, a rectified linear unit (relu) activation function, and an Adam (Adaptive Moment Estimation) gradient descent optimizer are typically used. A dropout layer is one of the most effective and commonly used regularization techniques for ANNs where arbitrarily selected neurons are ignored during training, and is often added after each hidden layer to prevent overfitting (Srivastava et al., 2014). However, the numbers of hidden layers and neurons should be tuned for each dataset, because it is important to determine the optimal network topology for the dataset (Bashiri and Geranmayeh, 2011). To select the optimal parameters that maximize the model scores, a grid search is used (Bergstra and Bengio, 2012). The final network topology for the DL-based SIC retrieval model is described in Fig. 5(b). The first five hidden layers of the developed network use the relu activation function, meaning zero out negative values, and have dropout layers with a rate of 0.2, i.e., one in five inputs are randomly excluded from each update cycle. Moreover, a saturation threshold value of the relu function in the last hidden layer is set to one as the output variable is the SIC. Therefore, the output values predicted by training the proposed network have a range of 0 to 1, which are then scaled from 0 to 100.

To create high-resolution-based SIC labels for DL-MLP model training and testing, we generate daily mosaics of MODIS surface reflectance images in the polar stereographic projection. Seventy-two images from April to September for each of the years 2016 and 2017 are used for model training and evaluation, respectively. Each week's data are composed of three consecutive images (e.g., 4/10, 11, 12; 9/11, 12, 13). First, ice/water classification or fractional abundance maps from each MODIS image are generated using CNN-classification and SMA-based approaches. To generate label datasets composed of MODIS SIC values (y_s, ρ) and AMSR2 TB values (x_s, ρ), we then find the pixel in the 500 m MODIS image nearest to the center coordinate (s, t) of each 10 km AMSR2 pixel using a Euclidean distance calculation. Because of the significant resolution difference between the two sensors, MODIS pixels with > 500 m distance from the center of the AMSR2 pixel are removed from the label sets. Let $x_{s, \rho}$ be a 14×1 vector of TBs recorded at the 6.9, 7.3, 10.6, 18.7, 23.8, 36.5, and 89 GHz channels with horizontal and vertical polarization. Using all MODIS pixels within 10 km of the IFOV (Instantaneous field of view) of AMSR2, we calculate $y_{s, \rho} = A_{ice}/A_{IFOV}$, where A_{ice} and A_{IFOV} are the areas of MODIS sea ice pixels and total MODIS pixels in the AMSR2's IFOV, respectively. The training process runs for a number of iterations in which the entire training dataset is exposed to the network until the loss function reaches its minimum error. The model score converges to the maximum after approximately 10,000 iterations and trains for around 14 h on an Intel Xeon (2.20 GHz, 22 cores) with a NVIDIA Titan X (3584 CUDA cores).

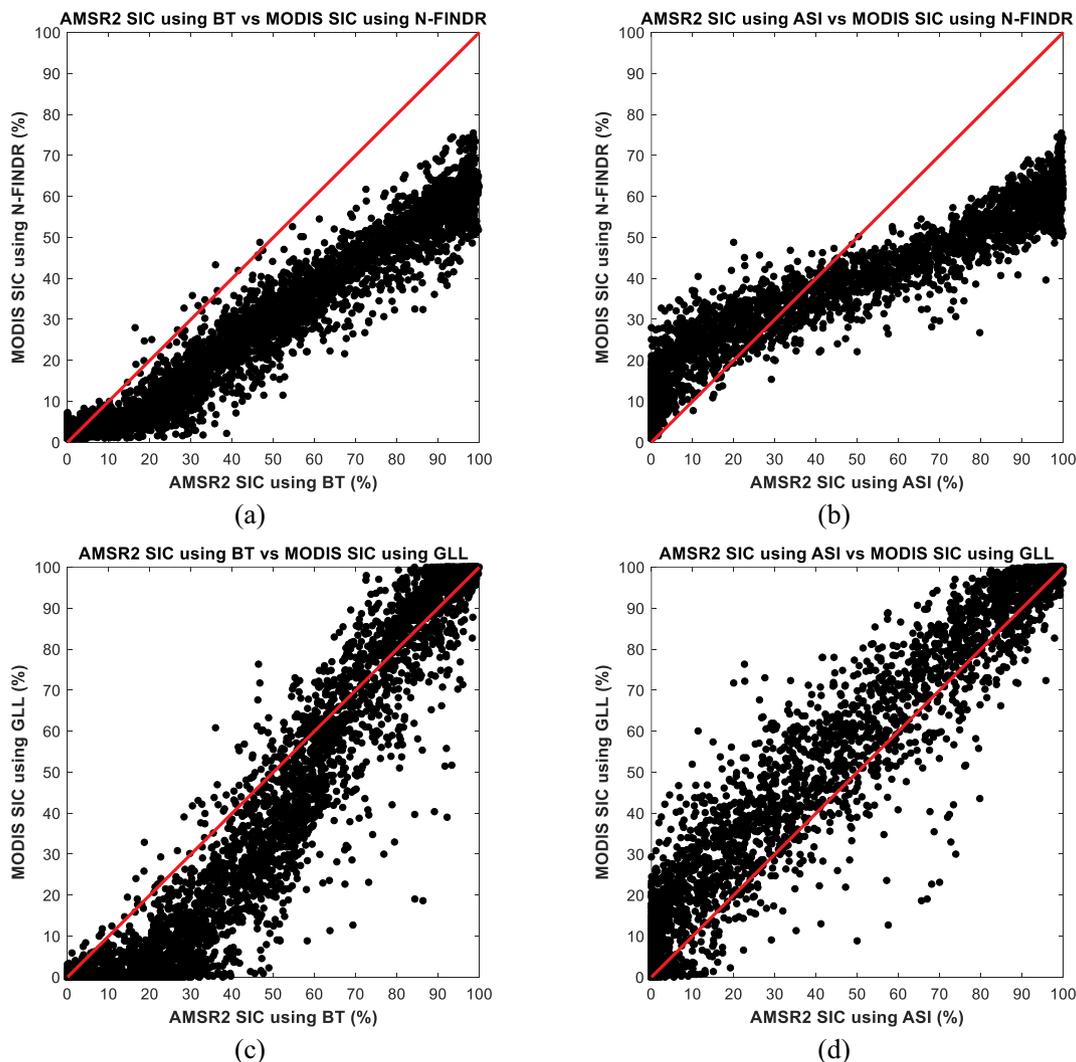


Fig. 7. Comparison of the performance of endmember extraction algorithms for SMA (spectral mixture analysis)-based SIC (sea ice concentration) calculations: (a) $SIC_{MODIS/SMA-NFINDR}$ vs $SIC_{AMSR2/BT}$; (b) $SIC_{MODIS/SMA-NFINDR}$ vs $SIC_{AMSR2/ASI}$; (c) $SIC_{MODIS/SMA-GLL}$ vs $SIC_{AMSR2/BT}$; (d) $SIC_{MODIS/SMA-GLL}$ vs $SIC_{AMSR2/ASI}$ (assuming that the $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/ASI}$ are the true values; N-FINDR: N-finder; GLL: global and local landmark; BT: Bootstrap; ASI: ARTIST sera ice).

4. Experimental results and discussion

We conduct experiments to test the performance of our proposed approaches, and to address and discuss issues in existing SIC products. This section describes the following experiments: 1) Test of the proposed GLL algorithm to select the best ice/water endmembers; 2) Comparison of classification using CNN and SMA to determine SIC estimates from MODIS; 3) Quantitative validations of the DL-based SIC retrieval model; 4) Feature importance test; 5) In-depth regional comparisons; and 6) Long-term cross-validations for an operational system.

4.1. Evaluation of GLL endmember extraction algorithm

We evaluate the proposed EE algorithm using a subset of MODIS images from over the Fram Strait acquired on August 1, 2017 (Fig. 2(a)). Using the SMA approach, SIC is calculated as $SIC_{MODIS/SMA} = (1/n) \sum_{i=1}^n a_{ice, i} \times 100$, $i = 1, \dots, n$, where $a_{ice, i}$ is the fractional abundance of the sea ice endmember at the i -th MODIS pixel and n is the number of MODIS pixels in the IFOV of AMSR2. Based on the assumption that, after removing land and cloud pixels, there are only ice and water pixels in the image, we first find ice and water endmembers using the well-known N-FINDR algorithm, and then compute the fractional abundances of the ice endmember to retrieve the SIC values, as

illustrated in Fig. 6(a). As mentioned in Section 3.2, because N-FINDR seeks extreme pixels as endmembers, the ice endmember identified by N-FINDR has high reflectance values in the visible and near-infrared wavelengths (bands 1–4), which may be associated with anomalous pixels in the image. If we use this anomalous signal as the ice endmember, the fractional abundance of the sea ice could be underestimated. In our test image, most sea ice pixels that are obviously consolidated ice show SIC values of 50%–60%. However, the proposed GLL algorithm finds a more representative ice endmember than N-FINDR, although both methods find similar water endmembers. As shown in Fig. 6(b), the resulting fractional map of the ice component is in better agreement with the MODIS reflectance image (see Fig. 2(a)), and captures more detailed sea ice characteristics than N-FINDR. Note that water and ice endmembers in SMA are similar to tie-points for low and high SIC in sea ice community, respectively, but they are not necessarily identical.

Fig. 7 compares SMA-based SICs using the N-FINDR ($SIC_{MODIS/SMA-NFINDR}$) and GLL ($SIC_{MODIS/SMA-GLL}$) EE algorithms, and the SICs from the BT ($SIC_{AMSR2/BT}$) and ASI ($SIC_{AMSR2/ASI}$)-based products. It is worthwhile to include results of such baseline SIC products for this experiment to evaluate the performance of the EE algorithms, as they are widely used and the SICs in the consolidated ice and ice-free regions are generally in good agreement, unlike in marginal ice zones. As

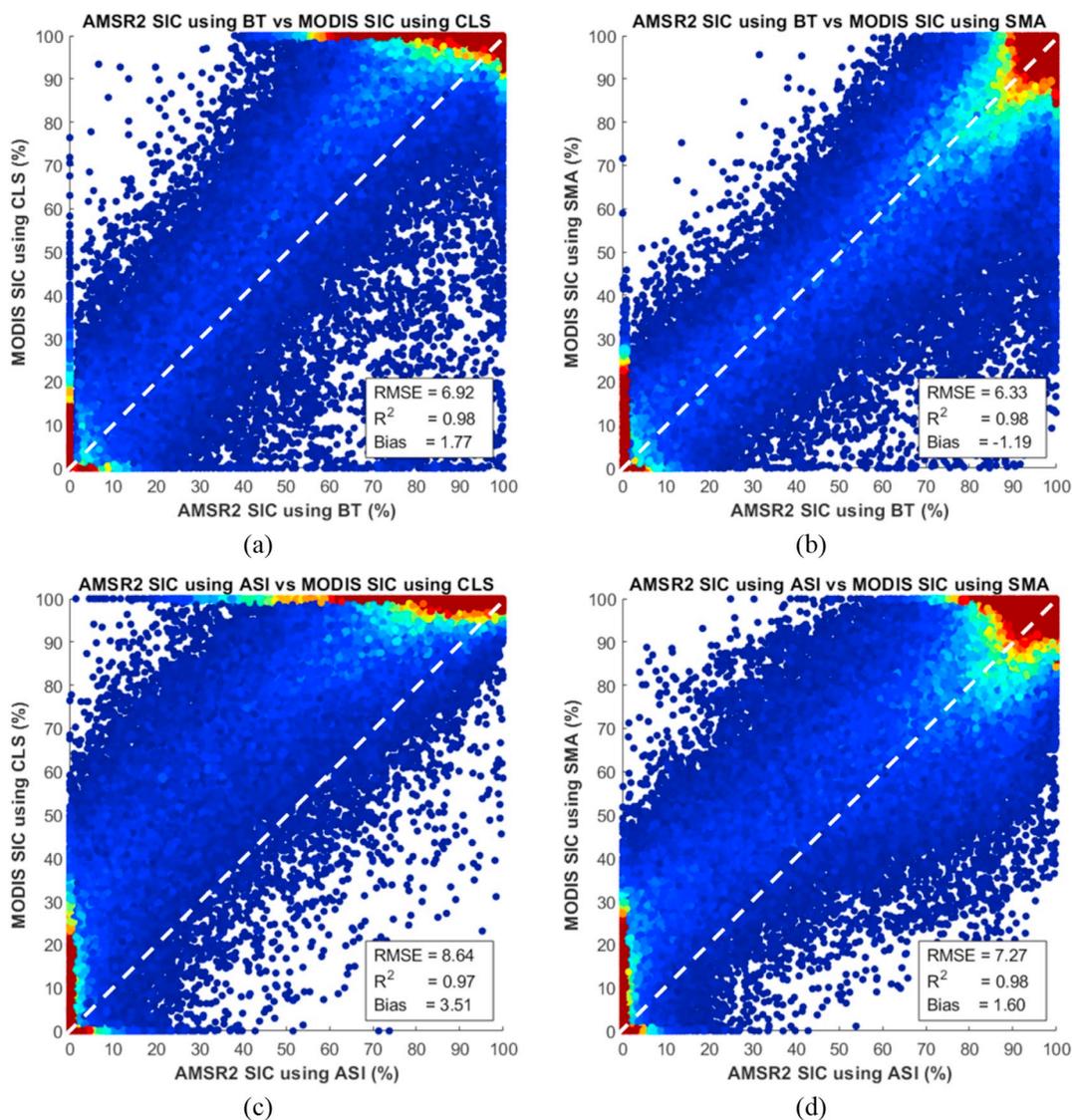


Fig. 8. Comparison of the performance of SIC (sea ice concentration) retrieval approaches using 2016 MODIS datasets. (a) $SIC_{AMSR2/BT}$ vs $SIC_{MODIS/CLS}$; (b) $SIC_{AMSR2/BT}$ vs $SIC_{MODIS/SMA-GLL}$; (c) $SIC_{AMSR2/ASI}$ vs $SIC_{MODIS/CLS}$; (d) $SIC_{AMSR2/ASI}$ vs $SIC_{MODIS/SMA-GLL}$ (assuming that the $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/ASI}$ are the true values; CLS: classification; SMA: spectral mixture analysis; GLL: global and local landmark; BT: Bootstrap; ASI: ARTIST sea ice).

shown in the scatter plots, the $SIC_{MODIS/SMA-NFINDR}$ values are significantly underestimated in high-SIC regions when compared with $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/ASI}$ values (see Fig. 7(a) and (b)). In contrast, the SICs given by our proposed GLL method are in reasonable agreement with those from the BT and ASI-based SIC images (see Fig. 7(c) and (d)), especially in the high-concentration ranges. As the proposed GLL algorithm mitigates extreme signals through the spectral averaging of globally and locally representative signals, it selects more robust and insensitive signals as endmembers than N-FINDR, resulting in more accurate fractional abundances (SICs) than N-FINDR. Therefore, in this study, the GLL algorithm is used to retrieve SMA-based SIC estimates from MODIS ($SIC_{MODIS/SMA-GLL}$) to determine the accurate labels.

4.2. Comparison of SIC calculations based on classification and spectral mixture analysis

As we address the limitations of classification for estimating high-resolution SICs in Section 3.1, the experiments described in this section are intended to evaluate whether SMA-based SIC retrieval is more accurate than the classification approach. A visual inspection of Figs. 2(b) and 6(b) shows that the classification approach assigned pixels to either

the sea ice or water class, regardless of ice types, conditions, and concentrations (see Fig. 2(b)), whereas each pixel identified by the SMA approach has different and continuous fractions that seem to be highly related to ice concentrations (see Fig. 6(b)). Quantitative comparisons are conducted using the 2016 MODIS datasets used for DL model training. As this experiment aims to determine the SIC retrieval approach from MODIS images, we assume that the $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/ASI}$ values are the true SICs in these experiments. For classification-based SIC calculations, supervised classification using CNN is performed to generate classification maps, and the SICs are then computed using classified MODIS pixels in the 10 km IFOV of AMSR2 as follows: $SIC_{MODIS/CLS} = \frac{\text{number of MODIS ice pixels in AMSR2's IFOV}}{\text{total number of MODIS pixels in AMSR2's IFOV}} \times 100$.

Fig. 8 shows scatter plots of the AMSR2-retrieved SICs by popular algorithms ($SIC_{AMSR2/BT}$, $SIC_{AMSR2/ASI}$; x-axis) and MODIS-retrieved SICs by classification/SMA ($SIC_{MODIS/CLS}$, $SIC_{MODIS/SMA-GLL}$; y-axis). Statistical comparisons are conducted using three difference metrics: 1) root mean square error (RMSE); 2) coefficient of determination (R^2); and 3) bias (mean error). In statistics, the bias is not usually as informative as the RMSE because positive and negative errors cancel each other out; however, in SIC evaluations, it is often used to determine

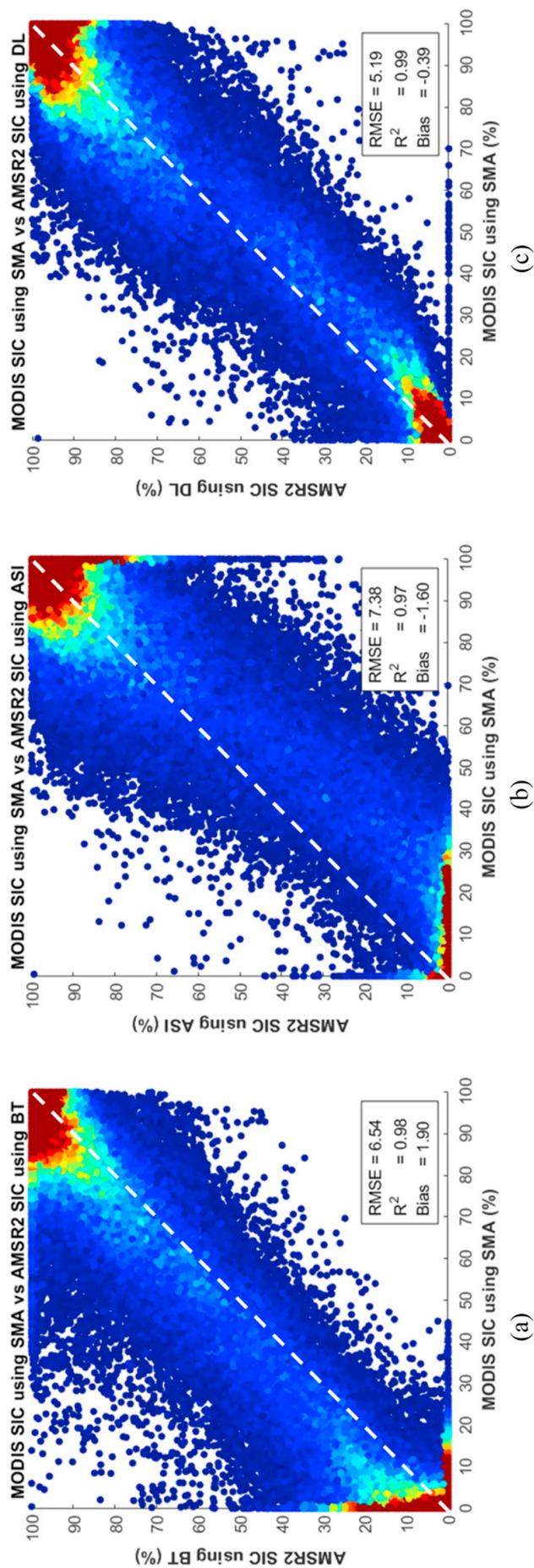


Fig. 9. Comparison of three SIC (sea ice concentration) retrieval algorithms with $SIC_{MODIS/SMA-GLL}$ using 2017 MODIS datasets: (a) BT ($SIC_{AMSR2/BT}$); (b) ASI ($SIC_{AMSR2/ASI}$); and (c) DL ($SIC_{AMSR2/DL}$) (assuming that the $SIC_{MODIS/SMA-GLL}$ values are the true values; SMA: spectral mixture analysis; GLL: global and local landmark; BT: Bootstrap; ASI: ARTIST sea ice; DL: deep learning).

Table 1
Statistical accuracies of SIC (sea ice concentration) products depending on the SIC ranges.

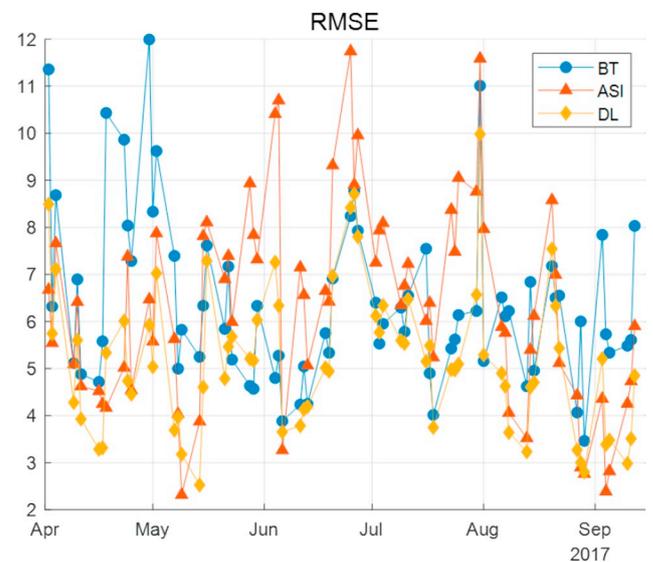
		Bootstrap	ARTIST sea ice	Deep learning
SIC < 5% & SIC > 95%	RMSE	2.83	2.75	1.87
	Bias	0.36	-0.99	-0.66
20% < SIC < 80%	RMSE	10.83	14.29	8.46
	Bias	6.32	-7.84	-0.40

whether the estimated values are overestimated or underestimated compared to the observed ones. As shown in Fig. 8(a) and (c), the distribution of points in the scatter plots and the positive bias values of SIC_{MODIS/CLS} indicate that the CNN classifier identifies more ice pixels than what actually existed, while the negative values in the comparison of the SIC_{MODIS/SMA-GLL} and SIC_{AMSR2/BT} in Fig. 8(b) indicate underestimations. Although the bias between SIC_{MODIS/SMA} and SIC_{AMSR2/ASI} in Fig. 8(d) also shows positive values, it decreases from 3.51 to 1.60. Furthermore, the absolute bias values of SIC_{MODIS/SMA-GLL} are smaller than those of SIC_{MODIS/CLS} in both comparisons of the BT- and ASI-based SIC images. Overall, the SIC_{MODIS/SMA-GLL} generally shows better statistical agreement with SIC_{AMSR2/BT} and SIC_{AMSR2/ASI} than SIC_{MODIS/CLS}. None of the difference metrics significantly improves as there are many samples in the open sea and consolidated ice regions with quite accurate SIC estimates from most SIC retrieval algorithms, along with inherent errors and uncertainties in the retrieval algorithms. However, this comparison is valuable, as it indicates that SMA-based SIC estimates from high-resolution data outperform the classification approach in terms of creating reference SIC labels. Therefore, in this study, the SMA approach is used to estimate SIC values from MODIS images for use as the ground truth in DL model training.

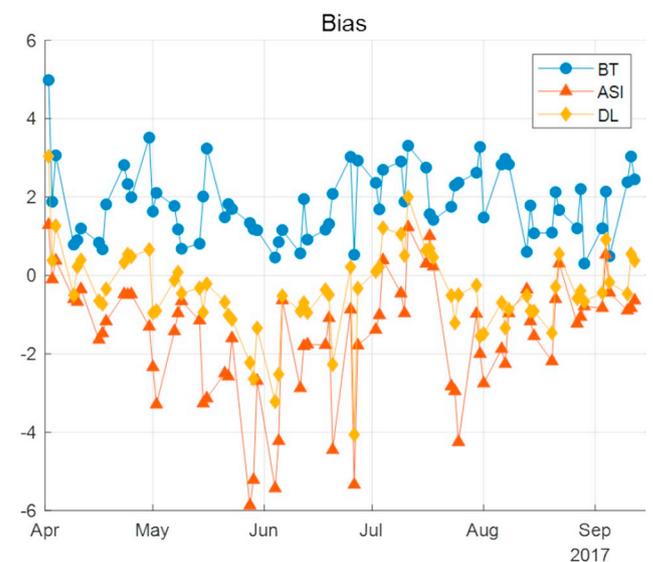
4.3. Training and validation of AMSR2 SIC retrieval model using MLP DL model

As described in Section 3.4, independent TB datasets from AMSR2 acquired in 2016 are used to train DL models for ascending and descending paths based on the network topology illustrated in Fig. 5(b). The SIC values retrieved from the DL models for ascending and descending paths are then averaged to calculate the daily SIC values (SIC_{AMSR2/DL}). As the development of an operational DL-based SIC retrieval model is the ultimate goal of this study, the retrieval model must be robust in time. We use 72 independent datasets collected in 2017 that are not used for model training to qualitatively and quantitatively validate the proposed retrieval model.

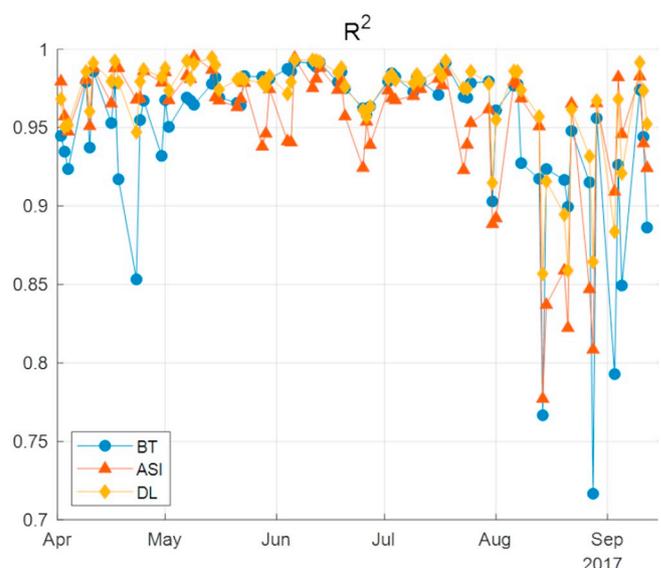
First, the overall accuracy of three SIC images retrieved by the BT, ASI, and DL algorithms are compared using quantitative evaluations, as shown in Fig. 9. Statistical comparisons are conducted using the SIC_{MODIS/SMA-GLL} values (x-axis) and the corresponding AMSR2-derived SICs given by the three retrieval algorithms (SIC_{AMSR2/BT}, SIC_{AMSR2/ASI}, SIC_{AMSR2/DL}; y-axis). For the SIC_{AMSR2/BT} values (see Fig. 9(a)), the distribution of points is mostly above the one-to-one line (white dashed line) and shows a positive bias, whereas with the SIC_{AMSR2/ASI} values (see Fig. 9(b)), many points are below the white dashed line showing a negative bias. Note that the SIC_{AMSR2/BT} values tend to overestimate the SIC, whereas the SIC_{AMSR2/ASI} values underestimate the SIC compared to the SIC_{MODIS/SMA-GLL} values. Overall, the statistical accuracy of the BT-based retrievals is slightly better than that of ASI (i.e., smaller RMSE and larger R²). In statistical terms, the proposed DL-based retrieval model generally outperforms both BT- and ASI-based retrievals. As shown in Fig. 9(c), the data points in the scatter plot are located around the one-to-one line, and the bias is much closer to zero than either BT or ASI, indicating that the DL model neither underestimates nor overestimates the SIC values compared to SIC_{MODIS/SMA-GLL}. This direct comparison without tie-point tuning of the BT and ASI algorithms may



(a)



(b)



(c)

(caption on next page)

Fig. 10. Daily statistical comparisons of BT (blue), ASI (red), and DL (yellow) algorithms. (a) RMSE; (b) Bias; and (c) R^2 (BT: Bootstrap; ASI: ARTIST sea ice; DL: deep learning). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Summary of statistical accuracies according to sea ice concentration retrieval algorithms on validation datasets from April–September 2017 (note: standard deviations are in parentheses.)

	Bootstrap	ARTIST sea ice	Deep learning
RMSE	6.39 (1.79)	6.37 (2.14)	5.19 (1.55)
Bias	1.84 (1.19)	-1.52 (1.55)	-0.42 (1.11)
R^2	0.95 (0.05)	0.95 (0.05)	0.97 (0.03)

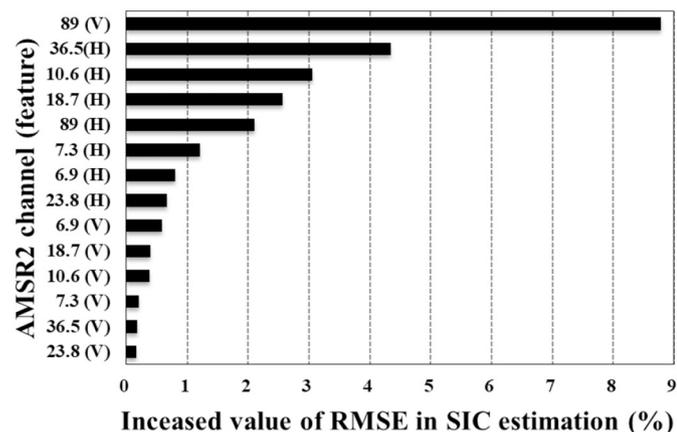


Fig. 11. Increased values of RMSE by feature importance test of AMSR2 channels to statistically analyze which channel contributes most to the sea ice concentration estimations (i.e., more important feature has higher increased value of RMSE).

not be adequate as the DL model is trained on SIC labels generated from MODIS datasets, and all results are relative to $SIC_{MODIS/SMA-GLL}$ values. Although the tie-points in the low and high SIC ranges can mostly be controlled (Ivanova et al., 2015), as shown in Table 1, at the extremes of the SIC (< 5% and > 95%), differences in all statistical metrics among the retrieval algorithms are not significant and the $SIC_{AMSR2/BT}$ values show the smallest bias. The differences and errors probably result from the marginal ice zones (20% < SIC < 80%) as most SIC algorithms are developed based on open sea and consolidated ice regions. In particular, the distributions of the $SIC_{AMSR2/DL}$ values (Fig. 9(c)) in the 20%–80% range are smaller and closer to the one-to-one line than values from other algorithms. However, some pixels are significantly misestimated compared to the reference $SIC_{MODIS/SMA-GLL}$ values in all retrieval algorithms. For example, there are zero SIC values estimated by all retrieval algorithms along the x-axis in Fig. 9. These significant differences may be caused by acquisition time differences between AMSR2 (01:30 for descending node; 13:30 for ascending node) and Terra MODIS (10:30) or inherent errors in the SMA-based SIC calculations from MODIS. Although $SIC_{MODIS/SMA-GLL}$ values are quantitatively estimated at the sub-pixel level, in some cases, such as clouds containing ice particles, the pixels can be detected as ice mixtures, which could be included in training and test datasets. We use a strict cloud filter (see Section 3.1) to minimize such errors, but this is not completely effective. However, these inaccurate labels are relatively few compared to the accurate labels. In addition, the DL model is trained using sufficient data and learning epochs. Therefore, these errors might be mitigated during the training process, resulting in good

agreement with the $SIC_{MODIS/SMA-GLL}$ values for most AMSR2 pixels.

Whereas the previous comparisons consider the overall accuracy for the test sets, Fig. 10 and Table 2 present statistical comparisons of daily test data to observe the seasonal performance of each retrieval algorithm. Table 2 shows that the proposed DL retrieval algorithm generally has better statistical performance with respect to $SIC_{MODIS/SMA-GLL}$ values than $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/ASI}$ values. Better mean statistics indicate that the proposed DL algorithm produces more accurate SIC products, and low-standard deviations in the daily data suggest more robust SICs in the time domain. Consistent with the previous overall comparison, $SIC_{AMSR2/ASI}$ is statistically less accurate than $SIC_{AMSR2/BT}$ for the data in this study. In the daily comparisons (Fig. 10), ASI-based SIC images underestimate the overall SIC values, but are exceptionally accurate compared with the $SIC_{AMSR2/BT}$ values in summer (August–September) when many melting ice and melt ponds exist. SIC values can be underestimated in summer as the existing PMW SIC retrieval algorithms consider the melt ponds as open water due to their similar radiometric signatures. Therefore, they often yield the percentage of the ice covered surface excluding the area of melt ponds (Ivanova et al., 2015; Kern et al., 2016). This is owing to the fact that the uncertainties and errors in the SIC retrieval algorithms are much greater in summer than in other seasons (Ivanova et al., 2015; Kern et al., 2016; Han and Kim, 2018). In optical images, the melt ponds have similar spectral signatures of open water, and this may result in mixtures of ice and water signals in a 500 m MODIS pixel. However, the proposed SMA approach accurately quantifies the amount of ice and water (either the open water or the melt ponds) fractions in a pixel, and then estimate accurate ice concentrations that are used for the DL model training. Although BT-based SIC images outperform ASI overall in this study, ASI-based retrievals are in better agreement with the DL results for summer sea ice estimates (Table 2). As the proposed DL retrieval model generates SIC outcomes that are statistically more accurate than those of both BT and ASI throughout the validation period of our data, there is a strong justification for further investigation of DL for estimating Arctic SIC values from AMSR2 PMW data.

4.4. Feature importance of AMSR2 channels in DL-based SIC retrieval model

Unlike the BT and ASI algorithms that exploit physical characteristics of ice and water signals across certain frequency and polarization channels of PMW data, in neural networks, quantifying the importance of single channels on the SIC estimations is difficult. This is because neural networks often solve problems by exploiting nonlinear relationships between multiple input features via hidden layers; hence, they are often called a “black box”. However, permutation importance tests can indicate the features having the biggest statistical impact on the retrievals. The permutation importance tests are performed using a trained model, where a single feature of the test data is randomly shuffled, while all other channels are kept constant. This results in less accurate retrievals as the data no longer corresponds to the fitted model. The magnitude of the errors as each channel is varied can be used to indicate its importance.

Fig. 11 shows the increased values of RMSE obtained by the permutation importance tests of the AMSR2 channels, which are calculated after 100 replications of tests. More important feature has higher increased value of RMSE. The most important feature is the 89 (V) channel. In this study, although we use resampled 10 km AMSR2 TB data as inputs, the original resolution of the high-frequency channel is nearly three times higher than that of the low-frequency channels. This may be one reason that the 89 (V) channel has the biggest impact on the retrievals. Four out of the first five most important features are the channels widely used in other SIC retrieval algorithms. However, the 10.6 (H) channel, which is often used in wind speed retrieval algorithms (Reddy et al., 2018), statistically shows the third-largest contribution to the model. Although current SIC retrieval algorithms do not

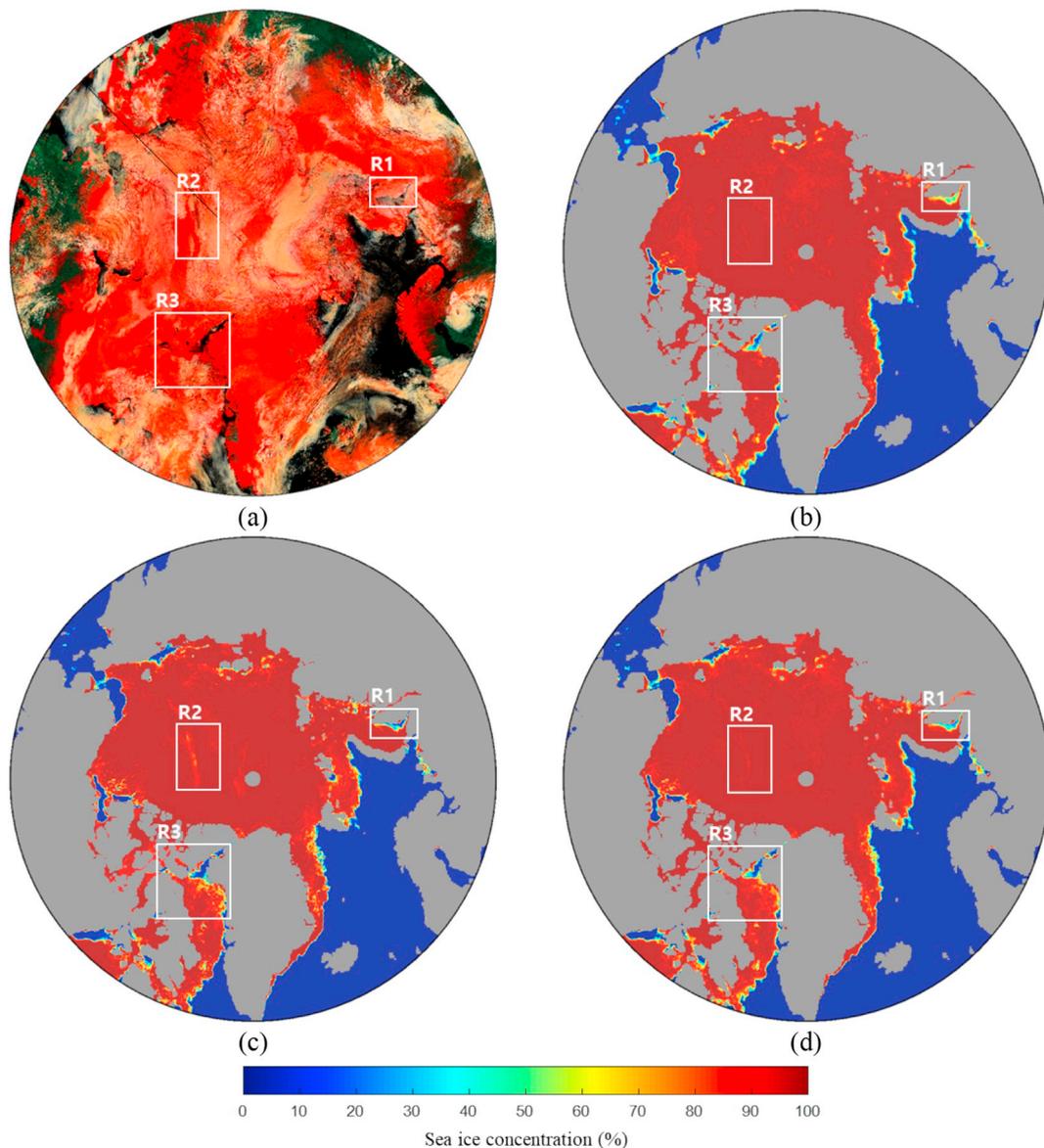


Fig. 12. Regional comparison on May 15, 2017, over Kara Sea (R1), Central Arctic ocean (R2), and Baffin Bay (R3). (a) Enhanced MODIS 1/6/7 band composite; sea ice concentration images retrieved by (b) BT (Bootstrap), (c) ASI (ARTIST sea ice), and (d) DL (deep learning).

exploit data from the 10.6 GHz channels, (Zhang et al., 2018) has shown a negative correlation between wind and SIC. Therefore, such data could be used to improve sea ice retrievals. The 23.8 GHz channels, often used in weather filters of the BT and ASI-based SIC products, are not statistically important, but this may have been due to the small number of pixels corrected by the weather filters. Overall, the horizontal polarization channels may statistically be more important than vertical polarizations, but there is no negative feature in the feature importance tests. It should be noted that all frequency and polarization channels of AMSR2 play a certain role in improving the accuracy of SIC retrievals, irrespective of whether they are statistically significant.

4.5. Regional in-depth comparisons

In this section, we investigate selected regions on specific dates to provide in-depth quantitative and qualitative comparisons of the retrieval methods. The comparisons are conducted over Baffin Bay in spring and the Fram Strait in summer. These data are selected as they exhibited high local variations in SIC values.

4.5.1. Baffin Bay (May 15, 2017)

Fig. 12 illustrates the MODIS 1/6/7 band composite image and SIC images generated by the BT, ASI, and DL algorithm for data captured on May 15, 2017. This experiment is designed to allow in-depth quantitative and qualitative comparisons of Baffin Bay, but several other interesting areas are also identified. In the Kara Sea (R1), the characteristics of each retrieval algorithm (as described in the previous section with respect to DL, i.e., BT: overestimate; ASI: underestimate; DL: neutral) are again observed. Due to insufficient number of pixels for statistical comparisons by a strict cloud filter (see Section 3.1), only visual inspections are conducted in this region. As shown in Fig. 13(a), it is difficult to discriminate ice from cloud pixels, but there are few low-SIC pixels that may be associated with mixtures of ice and water pixels. Visual interpretation indicates that $SIC_{AMSR2/ASI}$ and $SIC_{AMSR2/DL}$ in this region are similar to the ice distributions in high-resolution image, whereas the BT-based SIC values overestimate the SIC compared to the others. In the central Arctic ocean (R2 in Fig. 12), significant underestimations in the ASI-derived SIC map are observed in a narrow band running from top to bottom (70%–80% SIC range), unlike the BT and DL-based retrievals. In the rest of R2 of the ASI results and in the

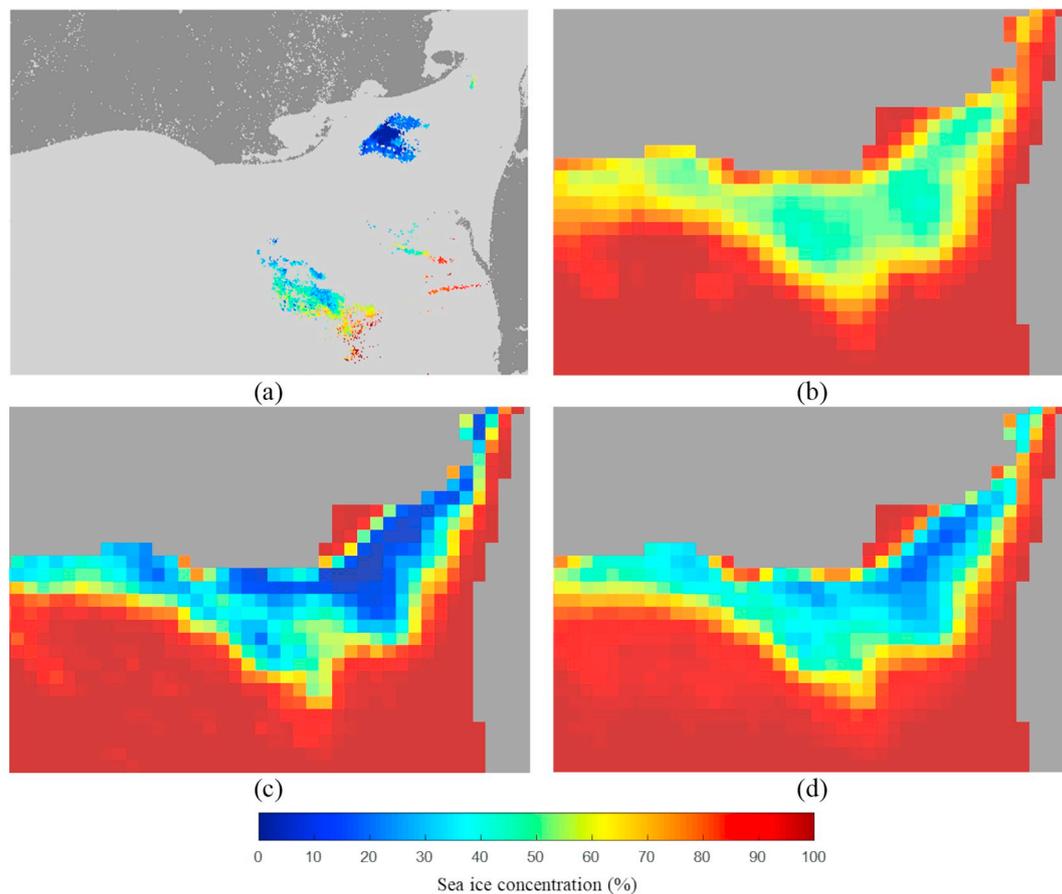


Fig. 13. Enlarged sea ice concentration images over Kara Sea (R1 in Fig. 12) retrieved by (a) SMA (spectral mixture analysis) from MODIS; (b) BT (bootstrap), (c) ASI (ARTIST sea ice), and (d) DL (deep learning)-based approach from AMSR2.

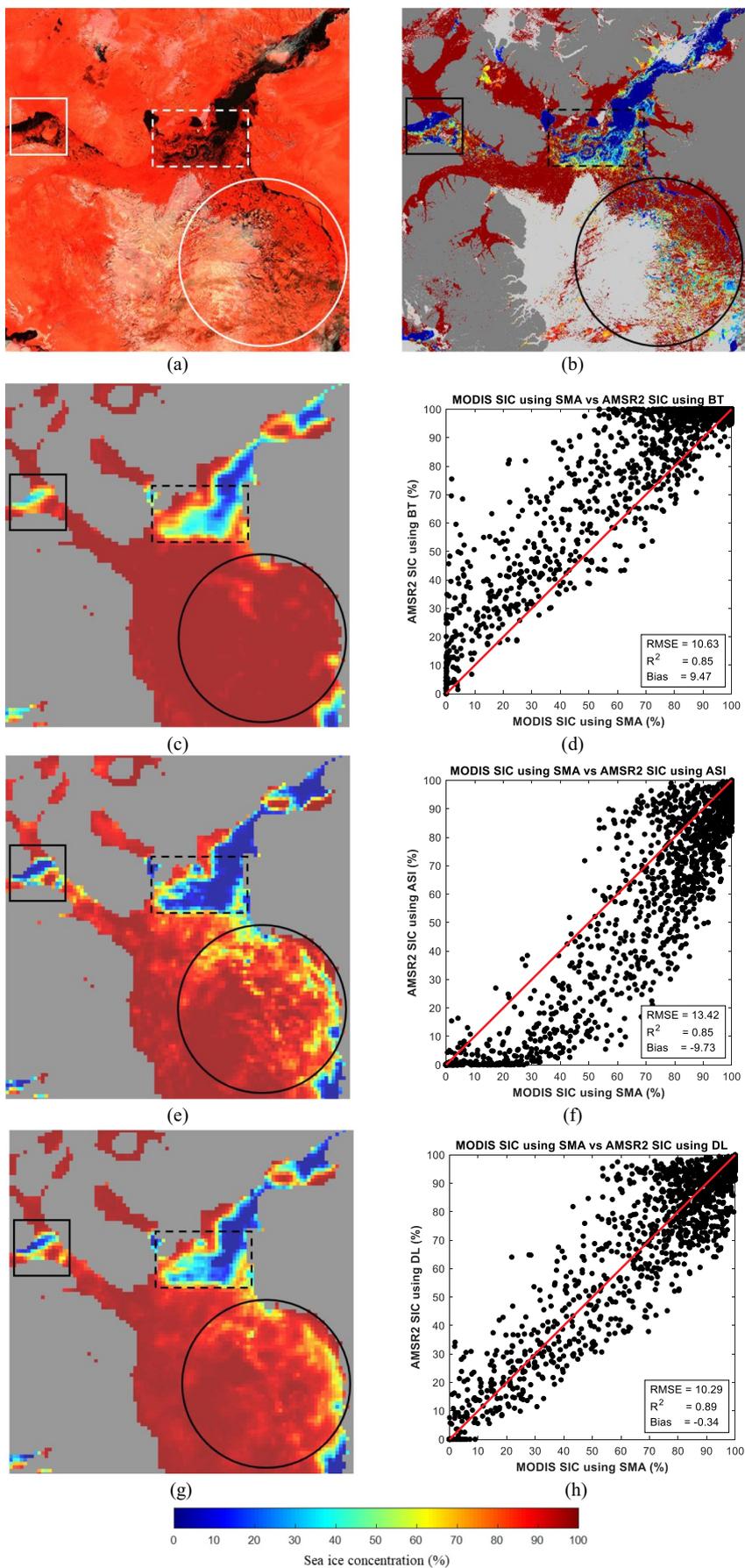
complete BT and DL results in R2, the SIC is homogeneously high or higher than that observed in the MODIS image (Fig. 12(a)). From an overall analysis using adjacent areas and dates, and high-resolution images, we additionally confirm that this region is definitely covered by consolidated ice. This type of obviously incorrect estimation may be attributable to weather events. The ASI algorithm uses high-frequency channels to provide high-resolution SIC products, but contributing factors from other frequencies and polarizations cannot be discerned. However, the proposed DL-based retrieval algorithm exploits all frequencies and polarizations for SIC estimates. Although it is difficult to physically explain how each channel contributes to the SIC estimations in the DL model, these inaccurate estimates in other retrieval algorithms are properly handled by the DL approach.

Fig. 14 illustrates in-depth comparisons over Baffin Bay (R3 in Fig. 12) and is composed of SIC images and scatter plots of considered true (x-axis) and estimated (y-axis) SICs retrieved by the BT, ASI, and DL algorithms. As shown in the visual comparisons of the resulting images (areas indicated by solid and dotted rectangles in Fig. 14(a)–(c)), the BT-retrieved image overestimates SIC values in regions where sea ice is not observed or covered only very small portions of the MODIS image. In contrast, the ASI-based SIC image neglects low-concentration or melting ice, and yields relatively large areas of ice-free pixels (see the dotted rectangle in Fig. 14(e)). However, the SIC image generated by our DL approach is in better agreement with the spatial distributions of sea ice in the MODIS SIC image. The SIC image by the ASI algorithm seems exhibiting greater SIC variability and capturing more detailed characteristics of the sea ice than other retrieval algorithms, but it underestimates high-concentration pixels in the visual inspection (see circles in Fig. 14(a)–(c), (e), (g)). In statistical comparisons based on the scatter plots, the $SIC_{AMSR2/BT}$ data show a large

positive bias of 9.47, whereas $SIC_{AMSR2/ASI}$ data have a large negative bias of -9.73 , compared to the $SIC_{MODIS/SMA-GLL}$ values. It exhibits similar incorrect estimation patterns as those using the entire validation dataset (see Fig. 14(d) and (f)). The DL-based retrieval model achieves a higher R^2 value for the linear regression model between the reference and estimated values, and a lower RMSE value than the BT and ASI-based products. As shown in Fig. 14(h), the samples are mainly distributed near the one-to-one line. A small bias also indicates that there are no systematic over-/underestimation, unlike the BT and ASI-based retrievals. In addition, pixels in the difficult-to-estimate range (20%–80% SIC) exhibit higher accuracy outcomes in the scatter plot.

4.5.2. Fram Strait (August 1, 2017)

Fig. 15 shows another MODIS 1/6/7 band composite alongside SIC products generated by the BT, ASI, and DL algorithms for data recorded on August 1, 2017. Similar to the previous validations, we first examine several local areas with visual inspections, and then conduct in-depth visual and statistical investigations over the Fram Strait (R4). Unlike the ASI and DL-retrieved images, BT generates a wide spread of low-SIC in the ice-free Kara Sea (R1). These errors may be dominated by TB changes caused by weather influences. Although weather filters are usually applied to SIC retrieval algorithms, as in this case, the artifacts are not always properly removed. An advantage of the proposed DL model is that it does not use any weather filters. As our proposed method uses all frequency and polarization channels to train the DL model, the channels used in the weather filters are inherently exploited as input attributes for the DL model training (see Section 4.4 and Fig. 11). In the Bering Sea (R2), small ice patches are reported by the BT and ASI-based SIC maps, unlike DL. Although the ice patches are not observed in the MODIS images because of high cloud cover, the



(caption on next page)

Fig. 14. Qualitative and quantitative comparisons over Baffin Bay (R3 in Fig. 12). (a) Enhanced MODIS 1/6/7 band composite; (b) MODIS-derived SIC (sea ice concentration) image; SIC images retrieved by (c) BT (Bootstrap), (e) ASI (ARTIST sea ice), and (g) DL (deep learning); scatter plots of $SIC_{MODIS/SMA}$ values (x-axis) and estimated SIC values (y-axis) retrieved by (d) BT, (f) ASI, and (h) DL.

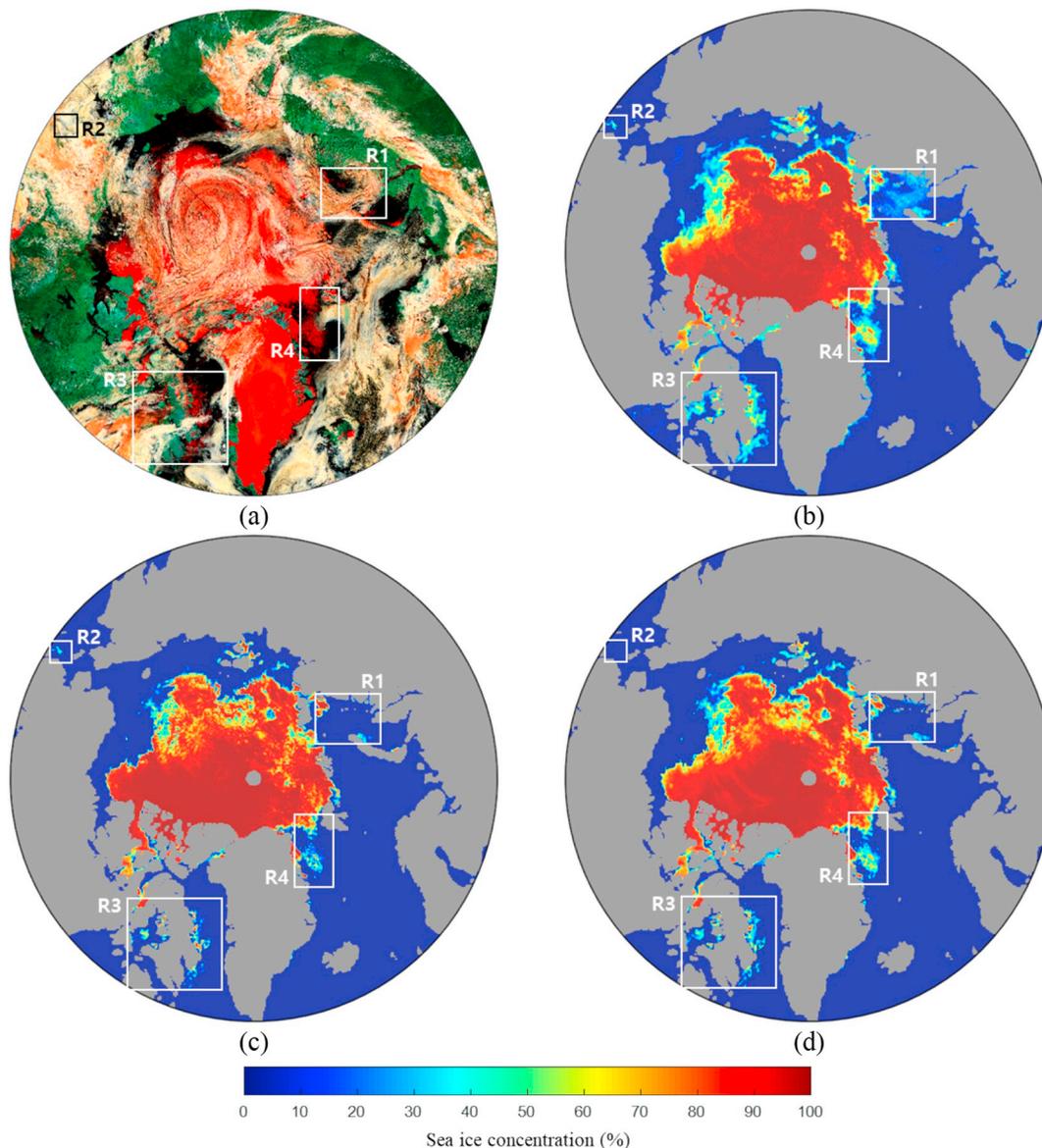


Fig. 15. Regional comparison on August 1, 2017, over Kara Sea (R1), Bering Sea (R2), Baffin Island (R3), and Fram Strait (R4). (a) Enhanced MODIS 1/6/7 band composite; sea ice concentration images retrieved by (b) BT (Bootstrap), (c) ASI (ARTIST sea ice), and (d) DL (deep learning).

presence of sea ice in the Bering Sea ($\sim 60^\circ$ latitude) in summer is unlikely and these patches are not observed on adjacent dates. The BT-based SIC image also produces large overestimations along the coastline of Baffin Island (R3) compared with the ASI and DL-based SIC values, although quantitative comparisons could not be conducted due to cloud cover. Overall, the BT-based SIC image visually overestimates the SICs compared with the ASI and DL-based SIC values in the summer season.

In Fig. 16, the SICs generated by the BT, ASI, and DL algorithms over the Fram Strait (R4 in Fig. 15) are qualitatively and quantitatively compared. As indicated by the solid and dotted rectangles in Fig. 16, both $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/DL}$ exhibit good agreement with $SIC_{MODIS/SMA-GLL}$ values, whereas $SIC_{AMSR2/ASI}$ does not locate most of the low-concentration pixels, resulting in underestimations. Overall, the extents of the ice patches identified by BT better match those in the MODIS images compared to the DL-based result, but BT does not capture the small ice-free areas indicated by the solid rectangles. In

addition, as illustrated by the circles in Fig. 16, both BT and DL indicate a similar SIC variability to the high-resolution images, unlike ASI. However, the sea ice extents identified from the BT and ASI images seem to be overestimations and underestimations, respectively, compared to MODIS-derived and DL-based SIC images, although the DL result misses some low-concentration pixels that are captured by the BT-based retrieval. As seen in the scatter of plots and generated statistics, the $SIC_{AMSR2/DL}$ values agree best with the $SIC_{MODIS/SMA-GLL}$ values that we take as reference (small RMSE and large R^2) and the samples show no significant bias. However, the $SIC_{AMSR2/BT}$ significantly overestimates values in the range 0–50% SIC, while the $SIC_{AMSR2/ASI}$ shows large underestimations in the 20%–80% SIC range, similar to the global comparisons. In general, in the spring data discussed in Section 4.5.1, the BT-based SIC slightly outperforms ASI, while in the summer data, the ASI generates more accurate SIC values than BT. Note that ASI-based SIC product shows relatively good estimates during the

melting season, where this regional result is consistent with the daily comparisons discussed in Section 4.3 (see Fig. 10).

4.6. Six-year cross-validation for an operational algorithm

Although our SIC calculations from MODIS using the SMA approach are performed in an unsupervised manner to generate validation data, the use of all past data for validation is inefficient due to the large quantity of data and the lack of availability of MODIS images in winter. In addition, as the goal of this research is to develop an operational SIC

retrieval algorithm, the model should be robust to unseen future data as well as past data.

In this experiment, we evaluate daily DL-based SIC data in the entire Arctic ocean using the results generated by operational BT and ASI algorithms on six-years of historical data collected from July 3, 2012 to December 31, 2017. This comparison enables us to determine that our proposed DL-based retrieval algorithm can be used in operational systems. The cross-validations used the RMSE and bias values between $SIC_{AMSR2/DL}$ and $SIC_{AMSR2/BT}/SIC_{AMSR2/ASI}$ values. Fig. 17 shows the daily RMSE and bias changes, while Table 2 summarizes the statistics

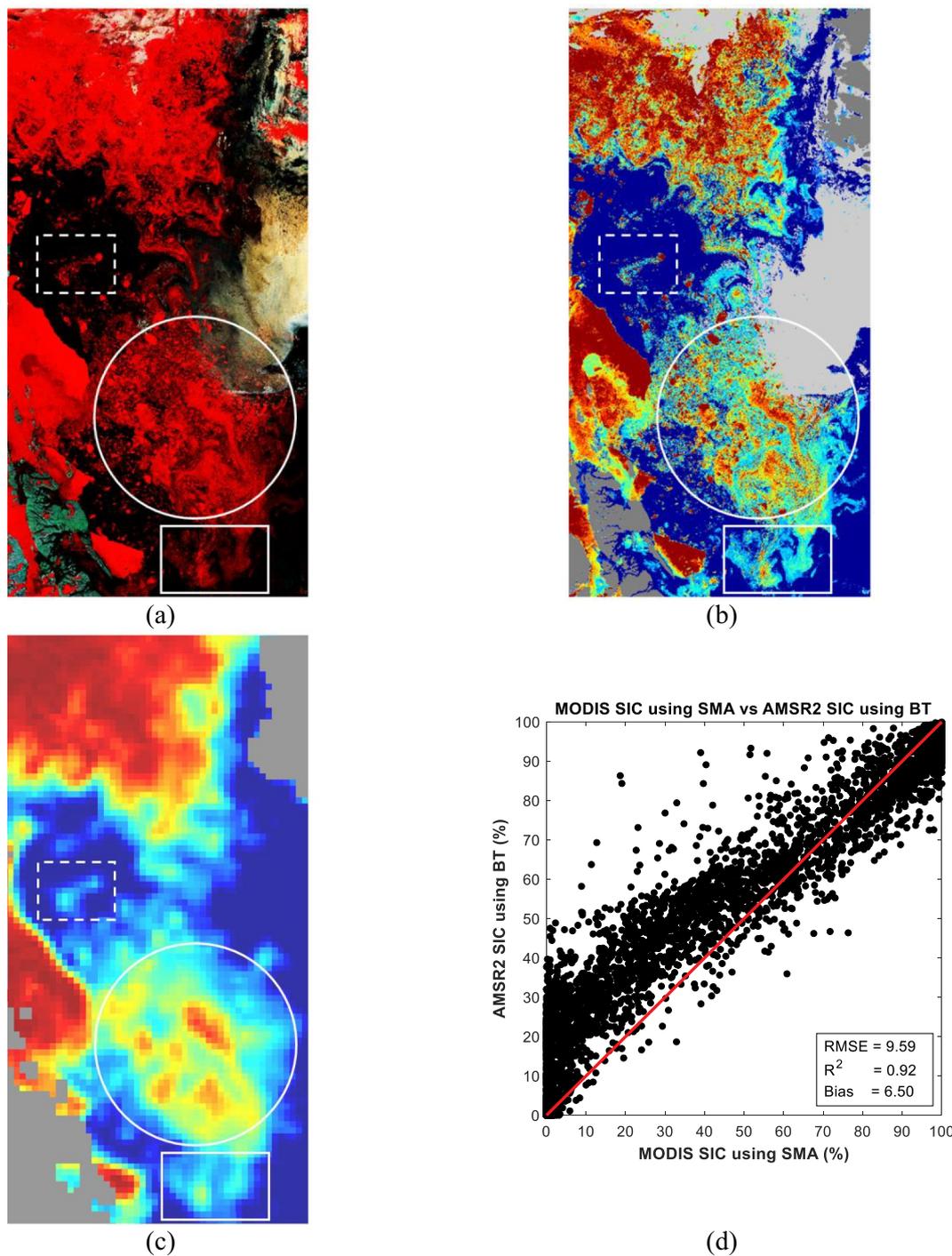


Fig. 16. Qualitative and quantitative comparisons over the Fram Strait (R4 in Fig. 15). (a) Enhanced MODIS 1/6/7 band composite; (b) MODIS-derived SIC (sea ice concentration) image; SIC images retrieved by (c) BT (Bootstrap), (e) ASI (ARTIST sea ice), and (g) DL (deep learning); scatter plots of $SIC_{MODIS/SMA}$ values (x-axis) and estimated SIC values (y-axis) retrieved by (d) BT, (f) ASI, and (h) DL.

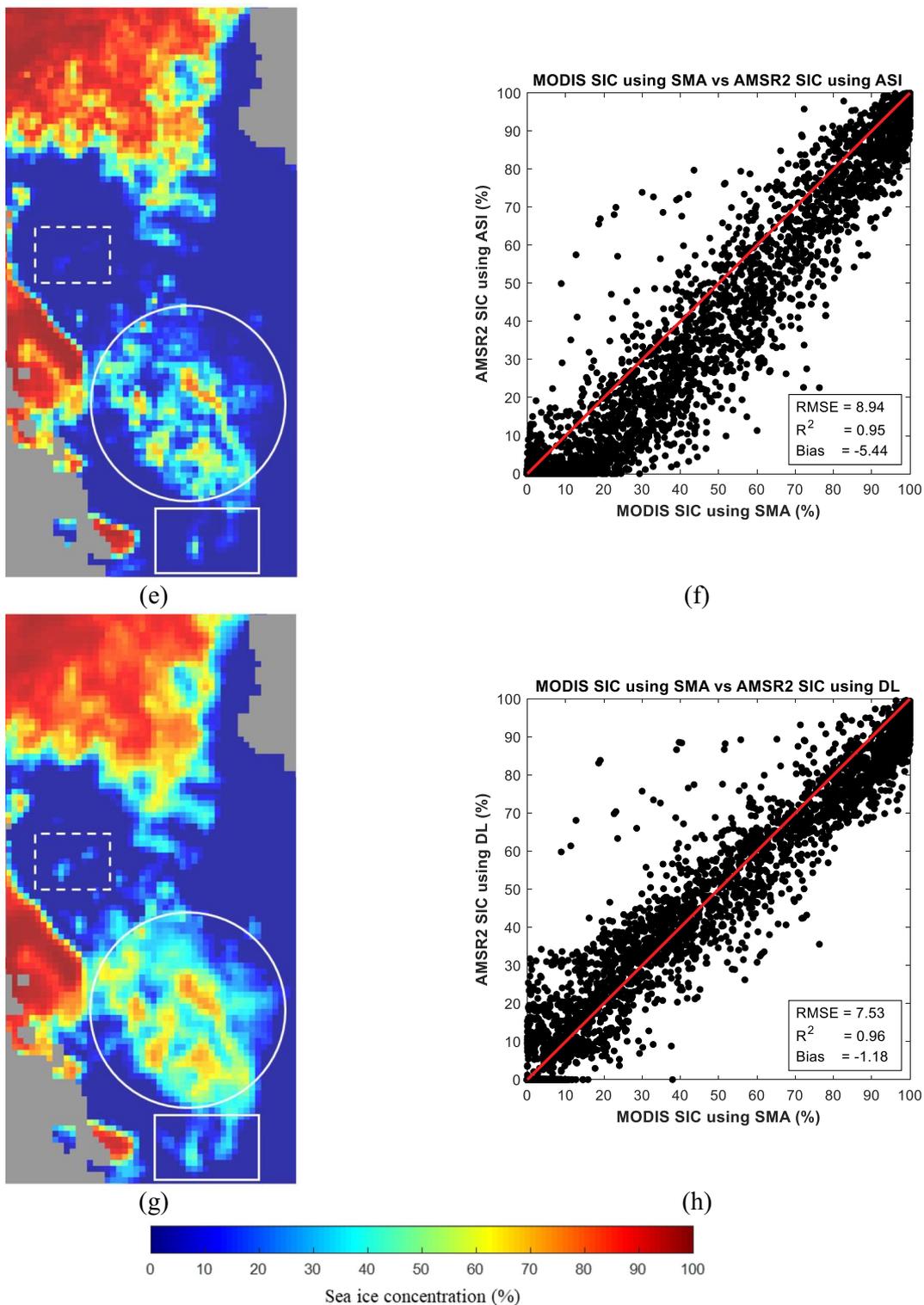


Fig. 16. (continued)

for the annual means and standard deviations of RMSE and bias values from 2012 to 2017. In the previous sections, we indicate that BT generally outperforms ASI throughout the year, except in summer. As shown in Fig. 17(a), the RMSE values of DL with respect to ASI are greater than those with respect to BT, while the RMSE values of ASI in summer are generally lower than or similar to those of the BT algorithm. For the relative bias changes in Fig. 17(b), overall, the $SIC_{AMSR2/BT}$ values slightly overestimate ice covers compared to the $SIC_{AMSR2/DL}$ values, whereas the $SIC_{AMSR2/ASI}$ values are underestimated, except in

summer, as discussed previously. However, these differences are not statistically significant. As listed in Table 3, the average RMSEs for BT and ASI in the 2016 datasets used for the DL model training are $2.87 \pm 0.41\%$ and $3.85 \pm 0.58\%$, respectively. In the pre-2017 and 2017 data, the annual mean RMSEs range from 2.61 to 2.99 for BT and 3.52–3.89 for ASI. The long-term means of BT and ASI are 2.85 and 3.68, respectively. In bias comparisons with respect to $SIC_{AMSR2/DL}$, BT and ASI have positive and negative bias, respectively, from 2012 to 2017, but the annual differences in the bias values are not significant

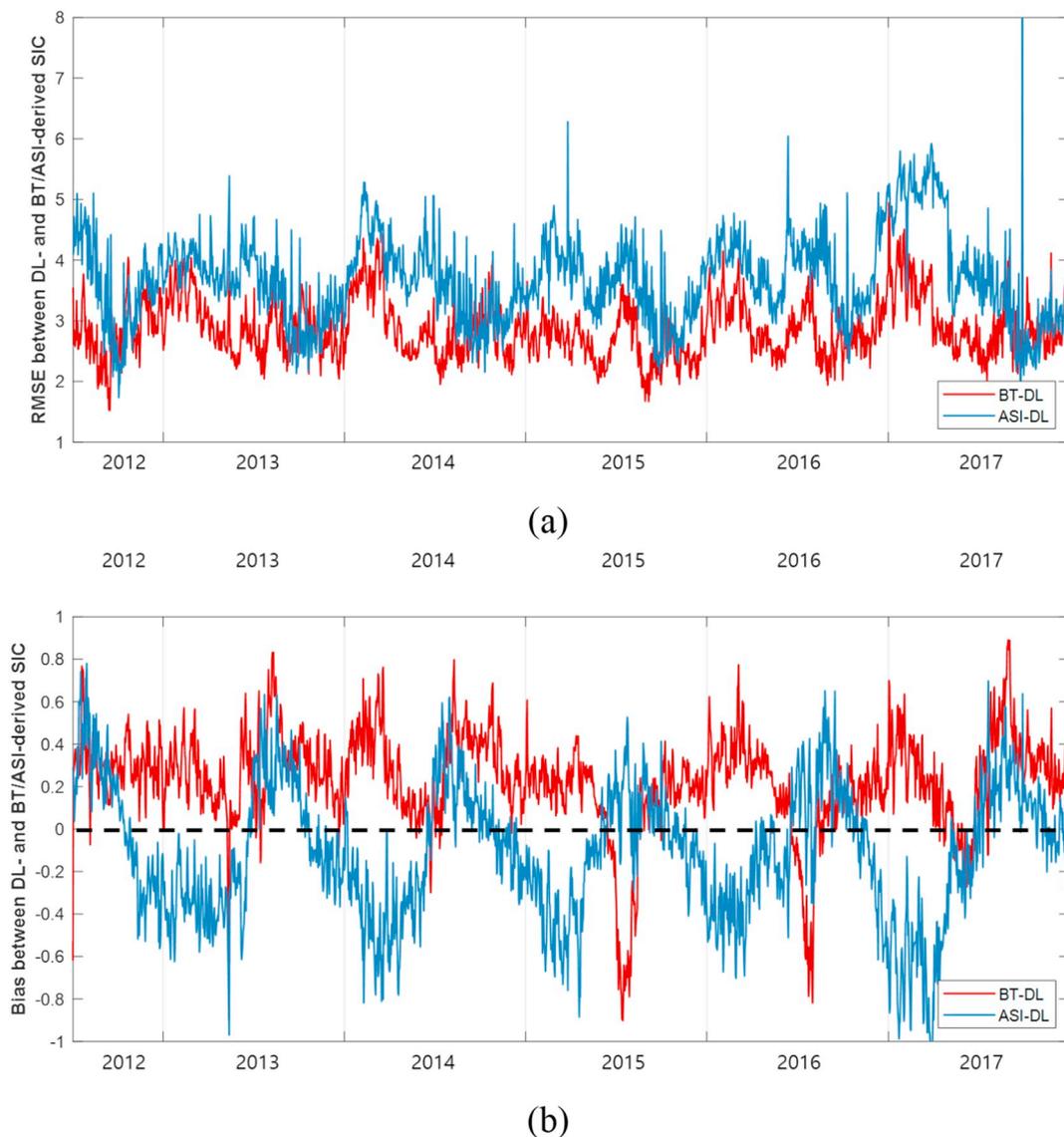


Fig. 17. Long-term (a) RMSE and (b) bias changes using cross-validation between $SIC_{AMSR2/DL}$ values and $SIC_{AMSR2/BT}$ (red)/ $SIC_{AMSR2/ASI}$ (blue) values in the entire Arctic ocean from 2012 to 2017 (SIC: sea ice concentration; BT: bootstrap; ASI: ARTIST sea ice; DL: deep learning). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

Annual means and standard deviations of RMSE and bias between DL (deep learning)-based and BT (bootstrap)/ASI (ARTIST sea ice)-based sea ice concentration values in the entire Arctic ocean from 2012 to 2017 (note: standard deviations are in parentheses.)

RMSE	2012	2013	2014	2015	2016	2017	2012–2017
BT – DL	2.79 (0.48)	2.90 (0.44)	2.93 (0.52)	2.61 (0.38)	2.87 (0.41)	2.99 (0.54)	2.85 (0.48)
ASI – DL	3.52 (0.75)	3.55 (0.57)	3.67 (0.66)	3.55 (0.60)	3.85 (0.58)	3.89 (1.07)	3.68 (0.74)
Bias	2012	2013	2014	2015	2016	2017	2012–2017
BT – DL	0.31 (0.14)	0.28 (0.18)	0.30 (0.19)	0.09 (0.28)	0.16 (0.25)	0.27 (0.22)	0.23 (0.23)
ASI – DL	0.07 (0.31)	-0.11 (0.29)	-0.13 (0.30)	-0.12 (0.26)	-0.11 (0.29)	-0.17 (0.38)	-0.11 (0.32)

(see Fig. 17(b) and Table 3). Generally, the comparisons of $SIC_{AMSR2/DL}$ values with respect to those from $SIC_{AMSR2/BT}$ and $SIC_{AMSR2/ASI}$ do not deviate considerably from the annual ranges. Note that the TBs from the AMSR2 PMW sensor are well calibrated, and demonstrate that the DL

model is robust and can be operationally used to generate daily SIC products without additional model re-training or updating processes when new data are acquired.

5. Conclusions

In this study, SMA (spectral mixture analysis; see Section 3.2) and a state-of-the-art DL (deep learning; see Section 3.4) are successfully incorporated to develop a new operational SIC (sea ice concentration) retrieval algorithm for the entire Arctic ocean using AMSR2 passive microwave data and high-resolution MODIS images. The primary contributions of this study are summarized as follows.

First, high-resolution imagery is commonly used to create reference SIC values, considered as ground truth. While pixel classification or thresholding is widely used to discriminate ice and water pixels in these images, determining robust thresholds is difficult in multi-temporal images and classification can overestimate or underestimate SICs compared to the actual amounts of ice due to sub-pixel mixtures of ice and water. These limitations motivated us to propose an SMA approach, which resulted in improved high-resolution SIC maps. Second, the

selection of proper endmembers (i.e., similar to tie-points in sea ice community) is critical for a successful SMA application. Solving a two-endmember (ice and water) problem with multispectral data seems easy, but extreme pixels or outliers can be selected as the endmembers. To address this, a new GLL (global and local landmark) algorithm that exploits both globally and locally representative pixels as endmember candidates is proposed (see Section 3.3), which resulted in accurate and continuous reference SIC labels. Third, based on high-resolution SIC maps given by the SMA approach and GLL method, we develop a DL architecture using MLP (multilayer perceptron) for pan-Arctic SIC estimates. The high-resolution SIC labels are successfully incorporated into the DL model training with AMSR2 data. Hyperparameters for determining the network topology for Arctic SIC retrievals are found through an iterative parameter optimization process that maximizes the model scores. Quantitative and qualitative evaluations with ice concentration map obtained from popular SIC retrieval algorithms are performed on both global and regional scales. Overall, the proposed DL-based retrieval model generates visually and statistically accurate SIC images, unlike the overestimates from BT (Bootstrap) and underestimates from ASI (ARTIST sea ice)-generated SIC images with respect to MODIS-derived SICs. In comparison with high-resolution regional images, the proposed method shows better agreement than the other algorithms, especially in the 20%–80% SIC zones. Additionally, there are some incorrect estimations given by the BT and ASI-retrieved SIC products (see Figs. 9, 14, 16). These are likely caused by severe weather conditions, although these products use weather filters to prevent such errors. However, we presume that the DL model inherently uses all frequencies and polarizations of AMSR2 data, so these errors are minimized without using weather filters. Although it is difficult to physically explain which channel most effectively reduces the errors, feature importance tests provide this information by calculating the change of the outcome when permuting channels. As the most important channel for the DL-based SIC retrieval model, the 89 GHz (V) channel is found. For a cross-validation, SIC images given by the proposed approach are consistently generated for six-year historical data, demonstrating sufficient statistical accuracy; thus, our proposed DL-based model is considered for operational use as a new retrieval method for providing improved SIC information in the Arctic.

Although the DL framework is successfully integrated into a new SIC retrieval algorithm, several challenges motivate further research. First, this study exploits all channels of the passive microwave data, although some channels and polarizations may be redundant. Efficient band selection or the projection of data onto a transformed domain (i.e., feature extraction) should be investigated with the aim of increasing the performance of the retrieval model. Second, combining environmental or image data acquired from different sensors, such as atmospheric parameters could enable better results for thin sea ice or for the marginal ice zones. Third, as the DL model proposed in this research is designed for the Arctic ocean, and supervised models are often highly dependent on training data, this retrieval model may not be appropriate for use with Antarctic data. Thus, development of additional retrieval models may be required to provide Antarctic SIC maps. Finally, we demonstrate that the results of the proposed retrieval model are stable in time because the AMSR2 brightness temperature products are well calibrated and our model is temporally robust. However, additional model re-training may be required if AMSR2 processing or calibration algorithms are updated in the future.

Acknowledgements

This study was supported by the Korea Polar Research Institute (KOPRI) Grant PE19120 (Research on analytical technique for satellite observation of Arctic sea ice).

References

- Andersen, S., Tonboe, R., Kern, S., Schyberg, H., 2006. Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using numerical weather prediction model fields: an intercomparison of nine algorithms. *Remote Sens. Environ.* 104 (4), 374–392.
- Andersen, S., Tonboe, R., Kaleschke, L., Heygster, G., Pedersen, L.T., 2007. Intercomparison of passive microwave sea ice concentration retrievals over the high-concentration Arctic sea ice. *J. Geophys. Res.-Oceans* 112 (C8), C08004.
- Bashiri, M., Geranmayeh, A.F., 2011. Tuning the parameters of an artificial neural network using central composite design and genetic algorithm. *Scientia Iranica* 18 (6), 1600–1608.
- Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13, 281–305.
- Carroll, M.L., DiMiceli, C.M., Wooten, M.R., Hubbard, A.B., Sohlberg, R.A., Townshend, J.R.G., 2017. MOD44W MODIS/Terra Land Water Mask Derived From MODIS and SRTM L3 Global 250m SIN Grid V006, 7 ed. NASA EOSDIS Land Processes DAAC (2017).
- Cavalieri, D.J., 1994. A microwave technique for mapping thin sea-ice. *J. Geophys. Res.-Oceans* 99 (C6), 12561–12572.
- Cavalieri, D.J., Gloersen, P., Campbell, W.J., 1984. Determination of sea ice parameters with the Nimbus-7 SMMR. *J. Geophys. Res.-Atmospheres* 89 (Nd4), 5355–5369.
- Cavalieri, D.J., Parkinson, C.L., Zwally, P.G.H.J., 1996. Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data. NASA National Snow and Ice Data Center, Boulder, Colorado USA.
- Chen, Y.S., Lin, Z.H., Zhao, X., Wang, G., Gu, Y.F., 2014. Deep learning-based classification of hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7 (6), 2094–2107.
- Chi, J., Crawford, M.M., 2014. Spectral unmixing-based crop residue estimation using hyperspectral remote sensing data: a case study at Purdue University. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7 (6), 2531–2539.
- Chi, J., Kim, H.C., 2017. Prediction of Arctic sea ice concentration using a fully data driven deep neural network. *Remote Sens.* 9 (12), 1305.
- Cho, K., Naoki, K., 2015. Advantages of AMSR2 for monitoring sea ice from space. In: *Asian Conference on Remote Sensing*, Manila, Philippines.
- Comiso, J.C., 1986. Characteristics of Arctic winter sea ice from satellite multispectral microwave observations. *J. Geophys. Res.-Oceans* 91 (C1), 975–994.
- Comiso, J.C., 1995. SSM/I Sea Ice Concentrations Using the Bootstrap Algorithm. National Aeronautics and Space Administration, Goddard Space Flight Center, USA.
- Comiso, J.C., Cavalieri, D.J., Parkinson, C.L., Gloersen, P., 1997. Passive microwave algorithms for sea ice concentration: a comparison of two techniques. *Remote Sens. Environ.* 60 (3), 357–384.
- Danaher, S., Omogain, E., 1992. Singular value decomposition in multispectral radiometry. *Int. J. Remote Sens.* 13 (9), 1771–1777.
- Emery, W.J., Fowler, C., Maslanik, J., 1994. Arctic sea-ice concentrations from special sensor microwave imager and advanced very high-resolution radiometer satellite data. *J. Geophys. Res.-Oceans* 99 (C9), 18329–18342.
- Gladkova, I., Grossberg, M., Bonev, G., Romanov, P., Shahrir, F., 2012. Increasing the accuracy of MODIS/aqua snow product using quantitative image restoration technique. *IEEE Geosci. Remote Sens. Lett.* 9 (4), 740–743.
- Gloersen, P., Cavalieri, D.J., 1986. Reduction of weather effects in the calculation of sea ice concentration from microwave radiances. *J. Geophys. Res.-Oceans* 91 (C3), 3913–3919.
- Han, H., Kim, H.C., 2018. Evaluation of summer passive microwave sea ice concentrations in the Chukchi Sea based on KOMPSAT-5 SAR and numerical weather prediction data. *Remote Sens. Environ.* 209, 343–362.
- Heinz, D.C., Chang, C.I., 2001. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* 39 (3), 529–545.
- Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. *Science* 313 (5786), 504–507.
- Hu, W., Huang, Y.Y., Wei, L., Zhang, F., Li, H.C., 2015a. Deep convolutional neural networks for hyperspectral image classification. *J. Sens.* 2015, 258619.
- Hu, F., Xia, G.S., Hu, J.W., Zhang, L.P., 2015b. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* 7 (11), 14680–14707.
- Ivanova, N., Johannessen, O.M., Pedersen, L.T., Tonboe, R.T., 2014. Retrieval of Arctic sea ice parameters by satellite passive microwave sensors: a comparison of eleven sea ice concentration algorithms. *IEEE Trans. Geosci. Remote Sens.* 52 (11), 7233–7246.
- Ivanova, N., et al., 2015. Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations. *Cryosphere* 9 (5), 1797–1817.
- Kaleschke, L., et al., 2001. SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis. *Can. J. Remote. Sens.* 27 (5), 526–537.
- Karvonen, J., 2017. Baltic Sea ice concentration estimation using SENTINEL-1 SAR and AMSR2 microwave radiometer data. *IEEE Trans. Geosci. Remote Sens.* 55 (5), 2871–2883.
- Kern, S., Kaleschke, L., Clausi, D.A., 2003. A comparison of two 85-GHz SSM/I ice concentration algorithms with AVHRR and ERS-2 SAR imagery. *IEEE Trans. Geosci. Remote Sens.* 41 (10), 2294–2306.
- Kern, S., Rösel, A., Pedersen, L.T., Ivanova, N., Saldo, R., Tonboe, R.T., 2016. The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations. *Cryosphere* 10 (5), 2217–2239.
- Keshava, N., Mustard, J.F., 2002. Spectral unmixing. *IEEE Sig. Proc. Mag.* 19 (1), 44–57.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444.

- Liu, J.P., Curry, J.A., Martinson, D.G., 2004. Interpretation of recent Antarctic sea ice variability. *Geophys. Res. Lett.* 31 (2), L02205.
- Markus, T., Cavalieri, D.J., 2000. An enhancement of the NASA Team sea ice algorithm. *IEEE Trans. Geosci. Remote Sens.* 38 (3), 1387–1398.
- Markus, T., Cavalieri, D.J., 2009. The AMSR-E NT2 sea ice concentration algorithm: its basis and implementation. *J. Remote Sens. Soc. Japan* 29 (1), 216–225.
- Okuyama, A., Imaoka, K., 2015. Intercalibration of Advanced Microwave Scanning Radiometer-2 (AMSR2) brightness temperature. *IEEE Trans. Geosci. Remote Sens.* 53 (8), 4568–4577.
- Plaza, A., Martinez, P., Perez, R., Plaza, J., 2004. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* 42 (3), 650–663.
- Reddy, B.N.K., Venkatesan, R., Osuri, K.K., Mathew, S., Kadiyam, J., Joseph, K.J., 2018. Comparison of AMSR-2 wind speed and sea surface temperature with moored buoy observations over the northern Indian Ocean. *J. Earth Syst. Sci.* 127 (14), 1–11.
- Spren, G., Kaleschke, L., Heygster, G., 2008. Sea ice remote sensing using AMSR-E 89-GHz channels. *J. Geophys. Res.-Oceans* 113 (C2), C02S03.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.
- Stroeve, J.C., Serreze, M.C., Holland, M.M., Kay, J.E., Malanik, J., Barrett, A.P., 2012. The Arctic's rapidly shrinking sea ice cover: a research synthesis. *Clim. Chang.* 110 (3/4), 1005–1027.
- Svendsen, E., et al., 1983. Norwegian remote sensing experiment: evaluation of the nimbus 7 scanning multichannel microwave radiometer for sea ice research. *J. Geophys. Res.: Oceans* 88 (C5), 2781–2791.
- Svendsen, E., Matzler, C., Grenfell, T.C., 1987. A model for retrieving Total Sea ice concentration from a spaceborne dual-polarized passive microwave instrument operating near 90 GHz. *Int. J. Remote Sens.* 8 (10), 1479–1487.
- Swift, C., Cavalieri, D., 1985. Passive microwave remote sensing for sea ice research. *Eos, Trans. Amer. Geophys. Union* 66 (49), 1210–1212.
- Vermote, E., Wolfe, R., 2015. MOD09GA MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006, 6 ed. NASA EOSDIS Land Processes DAAC.
- Vermote, E., Kotchenova, S., Ray, J., 2011. MODIS surface reflectance user's guide. In: MODIS Land Surface Reflectance Science Computing Facility, Version 1.
- Vinnikov, K.Y., Robock, A., Stouffer, R.J., Walsh, J.E., Parkinson, C.L., Cavalieri, D.J., Mitchell, J.F.B., Garrett, D., Zakharov, V.F., 1999. Global warming and Northern Hemisphere sea ice extent. *Science* 286 (5446), 1934–1937.
- Winter, M.E., 1999. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In: *Proc. SPIE Image Spectrometry V.* 3753. pp. 266–277.
- Zhang, L.P., Zhang, L.F., Du, B., 2016. Deep learning for remote sensing data: a technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* 4 (2), 22–40.
- Zhang, J., Stegall, S.T., Zhang, X., 2018. Wind–sea surface temperature–sea ice relationship in the Chukchi–Beaufort seas during autumn. *Environ. Res. Lett.* 13 (3), 034008.
- Zou, Q., Ni, L.H., Zhang, T., Wang, Q., 2015. Deep learning based feature selection for remote sensing scene classification. *IEEE Geosci. Remote Sens. Lett.* 12 (11), 2321–2325.
- Zuo, Z., Wang, G., 2014. Learning discriminative hierarchical features for object recognition. *IEEE Sig. Proc. Lett.* 21 (9), 1159–1163.