Acta Oceanol. Sin., 2025, Vol. 44, No. 3, P. 84–101
https://doi.org/10.1007/s13131-025-2467-3

Sea ice classification in the Arctic transition zone using Haiyang-2B microwave scatterometer and radiometer data

Shiyu Wu[1, 2, 3], Tingting Liu[1, 2, 3*], Mohammed Shokr[4], Ruibo Lei[5], Fan Yang[1, 2, 3], Yachao Li[1, 2, 3]

[1] Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan 430079, China
[2] Key Laboratory of Polar Environment Monitoring and Public Governance, Ministry of Education, Wuhan 430079, China
[3] School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
[4] Science and Technology Branch, Environment and Climate Change Canada, Toronto M3H5T4, ON, Canada
[5] Key Laboratory for Polar Science of the Ministry of Natural Resources, Polar Research Institute of China, Shanghai 200136, China

Received 27 August 2024; accepted 12 December 2024
© Chinese Society for Oceanography and Springer-Verlag GmbH Germany, part of Springer Nature 2025

Abstract

Loss of multiyear ice (MYI) is of great importance for Arctic climate and marine systems and can be monitored using active and passive microwave satellite data. In this paper, we describe an upgraded classification algorithm using the data from the scatterometer and radiometer sensors onboard the Chinese Haiyang-2B (HY-2B) satellite to identify MYI and first-year ice (FYI). The proposed method was established based on K-means and fuzzy clustering (K-means+FC) and was used to focus on the transition zone where the ice condition is complex due to the highly commixing of MYI and FYI, leading to the high challenge for accurate classification of sea ice. The K-means algorithm was applied to preliminarily classify MYI using the combination of scatterometer and radiometer data, followed by applying fuzzy clustering to reclassify MYI in the transition zone. The HY-2B K-means+FC results were compared with the ice type products [including the Ocean and Sea Ice Satellite Application Facility (OSI SAF) sea ice type product and the Equal-Area Scalable Earth-Grid sea ice age dataset], and showed agreement in the time series of MYI extent. Intercomparisons in the transition zone indicated that the HY-2B K-means+FC results can identify more old ice than the OSI SAF product, but with an underestimation in identifying second-year ice. Comparisons between K-means and K-means+FC results were performed using regional ice charts and Sentinel-1 synthetic aperture radar (SAR) data. By adding fuzzy clustering, the MYI is more consistent with the ice charts, with the overall accuracy (OA) increasing by 0.9%–6.5%. Comparing against SAR images, it is suggested that more scattered MYI floes can be identified by fuzzy clustering, and the OA is increased by about 3% in middle freezing season and 7%–20% in early and late freezing season.

Key words sea ice classification, Arctic multiyear ice, Haiyang-2B, transition zone, microwave remote sensing, fuzzy clustering

Citation Wu Shiyu, Liu Tingting, Shokr Mohammed, Lei Ruibo, Yang Fan, Li Yachao. 2025. Sea ice classification in the Arctic transition zone using Haiyang-2B microwave scatterometer and radiometer data. Acta Oceanologica Sinica, 44(3): 84–101, doi: 10.1007/s13131-025-2467-3

1 Introduction

Multiyear ice (MYI) is defined as sea ice surviving at least one summer melt. It used to be the dominant ice type in the Arctic, making up about 75% of the total ice extent in the 1980s (Maslanik et al., 2011). However, with the climate warming and ice loss in the Arctic, MYI is now occupying less than one-third of the Arctic Ocean (Kwok,

Foundation item: the National Key Research and Development Program of China under contract No. 2021YFC2803301; the Fundamental Research Funds for the Central Universities, China under contract Nos 2042024kf0037 and 2042022dx0001; the Natural Science Foundation of Wuhan under cocntract No. 2024040701010030.
*Corresponding author, E-mail: ttliu23@whu.edu.cn

2018). The remarkable decline of MYI has had a profound impact on the Arctic climate and marine systems, and the human activities in ice-rich Arctic waters (Meier et al., 2014, 2021; Bird et al., 2008; IPCC, 2022; Hovelsrud et al., 2011). These impacts have motivated studies on Arctic MYI identification and monitoring, which also weaves the foundation for thoroughly understanding the mechanism of Arctic sea ice loss, associated with the Arctic amplification.

Microwave remote sensing is the primarily tool for sea ice type and parameter retrieval as its signal can penetrate through atmosphere and clouds, while not affected by polar night (Gloersen et al., 1973; Comiso, 1986; Shokr and Sinha, 2023). Emitted radiation is measured in the form of brightness temperature (T_b) by passive microwave (PM) radiometers, while radar backscatter (σ^0) is measured by active microwave (AM) systems such as scatterometer. Although the spatial resolution of the radiometer and scatterometer systems is coarse, they have a high temporal resolution that permits daily monitoring of sea ice geophysical variables almost covering the entire Arctic Ocean. MYI emits lower T_b and higher σ^0 than the first-year ice (FYI) (Shokr and Sinha, 2023; Comiso, 1983). These differences make the discrimination between MYI and FYI possible using radiometric and scattering signatures.

Many studies have utilized only PM or AM data to classify sea ice types. For example, the AM data from the SeaWinds scatterometer onboard the QuikSCAT satellite has shown promising results for sea ice classification (Anderson and Long, 2005; Swan and Long, 2012; Lindell and Long, 2016a). AM sensor on board the OceanSat-2 has been used to extend the time span of ice type dataset (Lindell and Long, 2016b). PM data from the Special Sensor Microwave/Imager (SSM/I) and the Advanced Microwave Scanning Radiometer for Earth Observing System (EOS) (AMSR-E) data have also been applied to identify ice types (Wensnahan et al., 1993; Comiso, 2012). However, it is evaluated that the classification results of using only AM data shows an overestimation of MYI area and decreasing of overall classification accuracy (Yu et al., 2009; Zhang et al., 2022), which can be attributed to the fact that AM data are sensitive to wind-wave roughness, inducing anomalously higher backscatter intensities for sea ice, especially in the marginal ice zone with fragmented open water (OW). Meanwhile, using only PM data tends to underestimate area of MYI, due to the lack of surface distinguishable features obtained from T_b (Zhang et al., 2019).

In many studies, it has been found that the combination of AM and PM data can produce better results for sea ice type identification by offsetting the errors caused by the anomalies from each data source (Walker et al., 2006; Yu et al., 2009). This multi-sensor data strategy has then been applied to many ice type retrieval studies. Shokr et al. (2008) utilized AMSR-E and QuikSCAT data to retrieve the MYI concentration from 2003 to 2009 in the Arctic. Lindell and Long (2016b) employed Special Sensor Microwave Imager/Sounder (SSMIS) and Advanced Scatterometer (ASCAT) data to classify MYI and FYI. Advanced Microwave Scanning Radiometer 2 (AMSR2) and ASCAT data have also been used to produce a daily MYI concentration dataset by University of Bremen (Melsheimer and Spreen, 2022). The Ocean and Sea Ice Satellite Application Facility (OSI SAF) operational system combines data from SSMIS and AMSR2, along with ASCAT, to retrieve sea ice types (Aaboe et al., 2021a). Ye (2022) uses data from the Microwave Radiation Imager (MWRI) carried onboard the FengYun-3 satellites together with other AM data to retrieve sea ice types. The asynchronous observations from different satellites may pose the challenges to data harmony and sea ice type identification caused by the variations of orbit settings, including the inconsistent timing of equatorial crossings, the variations of orbital inclination, etc. (Ignatov et al., 2004). The Chinese HY-2B satellite is the second Chinese ocean dynamic environment satellite with multiple sensors onboard, including a microwave scatterometer (SCAT), a scanning microwave radiometer (SMR), a dual-frequency radar altimeter (ALT), and a calibration microwave radiometer (CMR). Data from HY-2B sensors have been evaluated by different studies and used in many applications on sea ice monitoring in the polar regions (Yu et al., 2022; Zhang et al., 2021; Han et al., 2023; Gao et al., 2019; Chen et al., 2024; Jiang et al., 2019, 2023). Due to the advantage of carrying both scatterometer and radiometer, the HY-2B provides a good potential in ice type retrieval by applying the data derived from the multi-sensor synchronous observation. Therefore, using the combination of HY-2B scatteromter and radiometer data should be an intentional attempt for sea ice type retrieval in the Arctic Ocean.

Algorithms that combine radiometer and scatterometer data to distinguish MYI and FYI in the Arctic Ocean employ a variety of methods, including time-dependent thresholding method (Swan and Long, 2012), the Bayesian method (Lindell and Long, 2016b), and the K-means algorithm (Zhang et al., 2019). For the time-dependent thresholding method, it used an average of the yearly sea ice backscatter measurements and a curve fitting method to separate MYI and FYI. Although the threshold curve can reflect on the general seasonal changes in sea ice backscatter caused by the seasonal evolution of sea ice, the interannual variabilities, determined by differences in ice age and the spatial distribution patterns of MYI, FYI, and OW, cannot be caught by this method. Meanwhile, using the Bayesian method, a so-called supervised algorithm, the classification errors would cumulate in the iterative process, leading to a reduced classification accuracy. K-means is an unsupervised algorithm featured with relatively simple implementation, high efficiency, and feasible classification accuracy. This algorithm has shown promising ability in sea ice classification (Zhang

et al., 2019, 2022; Ye et al., 2023). Therefore, we would propose the K-means algorithm in this study as a basis.

The MYI in Arctic Ocean is mainly distributed in north of Greenland and the Canadian Arctic Archipelago, i.e., the so-called switchyard of sea ice in the Arctic Ocean (Newton et al., 2021). The sea ice circulation derived by wind and surface current stress continuously brings the MYI to the southern part of the western Arctic Ocean through the Beaufort Gyre or to the Fram Strait through the Transpolar Drift Stream (Rigor and Wallace, 2004). Therefore, there is no fixed and clear boundaries between MYI and FYI in the Arctic Ocean. Instead, they are mixed together to varying degrees in a strip like manner, forming transition zones, with a strong seasonal and interannual variability, generally in the Canadian Basin and in north of East Siberia and the Laptev Seas. A comparison between the classification results of sea ice from different algorithms showed large differences in these transition zones (Ye et al., 2023). Numerous discrete MYI floes are usually found within a matrix of FYI in these complex sea ice transition zones, combined with the widely distributed leads even in the winter (Marsan et al., 2004), hence it results in a strong spatial heterogeneity in the microwave radiation and backscattering characteristics, and makes the identification of the MYI signature more difficult by the coarse-resolution radiometer or scatterometer instrument onboard the satellite. The EASE-Grid sea ice age (SIA) dataset uses a Lagrangian parcel tracking method to catch MYI floes that can partly solve the difficulty brought by the spatial heterogeneity in microwave signature for coarse-resolution satellite data (Tschudi et al., 2019a). However, it is suggested that the ice age product overestimates the MYI area because it assigns each grid with the age of the oldest ice type within it, and this kind of overestimation is pronounced in the transition zone, where the MYI concentration is usually around 50% (Babb et al., 2023). Moreover, with the proportion of younger MYI increasing and the ice continuous thinning, the microwave signature of MYI has become more similar to FYI, especially in the later ice growth season due to the close ice thickness. Also, the highly dynamic nature of the transition zone could lead to more deformed FYI, resulting in an increased surface roughness and an enhanced scatterometer radar signature similar as the MYI. These influence factors for the discrimination between these two ice types have been observed in K-means classification result (Ye et al., 2023).

Due to the great significance of total MYI area of the Arctic Ocean for evaluating the mechanism of changes in Arctic sea ice and parameterizing thermodynamic and dynamic parameters in sea ice numerical models, in this study, we developed a technique to improve MYI identification in the Arctic Ocean, especially focusing on the transition zones, through introducing a fuzzy clustering module integrated with the K-means classification approach. The proposed method, i.e., the so-called K-means+FC method was applied to the HY-2B radiometer and scatterometer data from two freezing seasons from November to April of 2019/2020 and 2020/2021 to assess the applicability of the proposed method. Ice type products and high spatial resolution imagery data were employed to validate the proposed method.

The rest of this paper is organized as follows. All the data used in this study are described in Section 2, including the HY-2B data and validation datasets. In Section 3, details of the classification algorithm are provided. Section 4 provides the classification results and the comparisons with validation datasets. Our conclusions are presented in Section 5.

2 Datasets

In this section, we describe the data from the SCAT and SMR onboard the HY-2B satellite, which were used to classify MYI and FYI over the Arctic. The validation data, including the Canadian Ice Service (CIS) Arctic regional sea ice charts, the OSI SAF global sea ice type product, the National Snow and Ice Data Center (NSIDC) SIA dataset (Version 4), and synthetic aperture radar (SAR) data from Sentinel-1, will also be briefly introduced.

2.1 *HY-2B SCAT data*

The HY-2B satellite was launched on 25 October 2018 and currently uses a sun-synchronous orbit with 99.3° inclination and 970.73-km average altitude. It is a fixed and circular orbit with a local equator crossing time with the ascending node at 06:00 am of local time. The repeat cycle of this orbit is 14 d. More specifications and applications of the instruments onboard the HY-2B and its products are available at the Chinese National Satellite Ocean Application Service (NSOAS) website (https://osdds.nsoas.org.cn).

Table 1 represents the main parameters of the HY-2B SCAT instrument. The HY-2B SCAT employs a rotating pencil beam antenna, with the transmiting and receiving frequency at Ku-band of 13.256 GHz. Measurements of the normalized backscatter coefficient (σ^0) are acquired in the two co-polarization transmit-receive modes, i.e., horizontal-horizontal (HH) and vertical-vertical (VV) at incidence angles of 41.5° and 48.6°, respectively. The swath widths of HY-2B SCAT are 1 350 km and 1 700 km for HH (inner beam) and VV (outer beam), respectively. The

Table 1. Technical specifications of the HY-2B SCAT

Parameter	Value
Instrument frequency/GHz	13.256
Bandwidth/MHz	1
Polarization mode	HH, VV
Spatial resolution/km	25
Swath width/km	H polarization: 1 350 V polarization: 1 700
Sigma-0 measurement accuracy/dB	0.5
Backscattering range/dB	–40 to 20

observational area effectively covers 90% of the global ocean surface within 48 h. The HY-2B SCAT L1B σ^0 data are provided by NSOAS at a spatial resolution of 25 km. About 16–17 swath image files are produced over the Arctic daily. In this study, the SCAT L1B data in the freezing seasons from November to April of two years of 2019/2020 and 2020/2021 were used to avoid the surface melt conditions during the warm season. For each day, the swath images were gridded into a 25 km × 25 km polar stereographic grid via the "drop-in-the-bucket" method.

2.2 *HY-2B SMR data*

The SMR onboard the HY-2B satellite operates at 6.6 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz, and 37 GHz. Except for 23.8 GHz, which is only vertically polarized, the other four frequencies have both vertical (V) and horizontal (H) polarization channels. The SMR sensor scans the Earth over a swath width of 1 600 km with a scan angle of 53°, and the footprints (instantaneous field of view) vary with frequency. The sensitivity is better than 0.8 K in the 37 GHz band and better than 0.5 K for other frequency bands (Yu et al., 2022). More specifications are shown in Table 2.

In this study, the HY-2B SMR L2A_TC swath data from November to April in two years of 2019/2020 and 2020/2021 were utilized. Approximately 29–30 daily swaths were gridded into a 25 km × 25 km grid. The data are provided as a corrected T_b product after conducting inter-calibration. The data have a superior stability, with the annual drift of T_b less than 0.1 K (Wang et al., 2022). Data from the 18.7 GHz and 36.5/37 GHz, both in V-polarization, were used in this study for the purpose of flagging out the open waters and classifying sea ice types, respectively.

2.3 *CIS Arctic regional sea ice charts*

The CIS produces digital Arctic regional sea ice charts with a relatively high spatial resolution for marine navigation, forecasting, and climate research. The charts are generated through manual analysis of various satellites images (with SAR data as the primary source), aerial reconnaissance information, and ship reports. The ice charts are provided in SIGRID-3 format, containing polygons identified subjectively by ice analysts. Each polygon is coded with sea ice attributes such as stage of development (SoD), concentration, and surface form. Ice information is presented in the Egg Code format and color-coded using the World Meteorological Organization (WMO) Standard (WMO and IOC-UNESCO, 2014).

In this study, the CIS Arctic regional sea ice charts of Western Arctic were used to validate the HY-2B MYI classification results. We selected one ice chart for each month to evaluate the classification results. The maps of SoD information were considered as the reference. When using ice charts, the old ice, multiyear ice, and second-year ice documented in the ice charts were grouped as MYI, while all the other ice types were grouped as FYI. For polygons that contain MYI and FYI of the equivalent partial concentration, we followed the colour graphic chart published by the CIS using the WMO standard. We then rasterized the MYI polygons to the same projection as the HY-2B image data.

2.4 *OSI SAF global sea ice type product*

The Ocean and Sea Ice Satellite Application Facility (OSI SAF) global sea ice type product is provided by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). This product is generated by combining PM and AM remote sensing data in a Bayesian classification scheme (Aaboe et al., 2021a). Sea ice is separated into three classes of FYI, MYI, and ambiguous class. In addition, due to melting at the ice surface in the warm season from May to September, the ice type classification becomes ambiguous. This product has been validated by evaluating the temporal variation of the MYI area coverage (Aaboe et al., 2021b), which reveals it can be used as a reliable validation dataset for MYI classification.

2.5 *NSIDC EASE-Grid SIA dataset*

The Equal-Area Scalable Earth-Grid sea ice age (EASE-Grid SIA) dataset (Version 4) from the National Snow and Ice Data Center (NSIDC) provides weekly ice age information at a spatial resolution of 12.5 km since 1984 (Tschudi et al., 2019a). A Lagrangian parcel tracking method is used to estimate sea ice age based on the NSIDC Polar Pathfinder Sea Ice Motion dataset (Tschudi et al., 2019b, 2020). The age of the ice in the product is

Table 2. Technical specifications of the HY-2B SMR

Parameter	Value				
Frequency/GHz	6.925	10.7	18.7	23.8	37.0
Bandwidth/MHz	350	100	200	400	1 000
Polarization mode	V, H	V, H	V, H	V	V, H
Swath width/km	⩾1 600	⩾1 600	⩾1 600	⩾1 600	⩾1 600
Ground inclination	53°	53°	53°	53°	53°
Ground track/km	90 × 150	70 × 110	36 × 60	30 × 52	20 × 35
Sensitivity/K	⩽0.5	⩽0.6	⩽0.5	⩽0.5	⩽0.8
Dynamic range/K	3–350	3–350	3–350	3–350	3–350
Calibration accuracy/K	1	1	1	1	1

categorized as FYI (0–1 year old), second-year ice (1–2 years old), and so forth, based on how many summer melt seasons the ice parcel has survived. Note that grid cells with less than 15% sea ice concentration are treated as open water. In this study, the ice age product was utilized to assess the proposed retrieval algorithm.

2.6 *Sentinel-1 SAR images*

The Sentinel-1 mission comprises a constellation of two polar-orbiting satellites, Sentinel-1A launched on 3 April 2014 and Sentinel-1B on 25 April 2016, both with C-band SAR instrument onboard. The two satellites have a revisit period of 12 d for the Arctic Ocean. In this study, images in the Extra-Wide swath (EW) mode with a medium resolution of 40 m in HH/HV polarization were used. Each scene covers approximately 400 km × 400 km. For validating of the algorithm, we selected six images in the Beaufort Sea. To illustrate the performance of algorithm in various stages of freezing season, the images are chosen from the following dates of November 28, 2019 and November 3, 2020 for early freezing season, 15 January 2020 and 10 January 2021 for middle freezing season, and 14 April 2020 and 18 April 2021 for late freezing season. The ice cover in Beaufort Sea usually features a mixture of FYI and MYI, due to the clockwise ice circulation derived by Beaufort Gyre (Moore et al., 2022). Images of HV polarization were mainly used, with the images of HH polarization as supplement, as the backscattering signals of MYI and FYI are more differentiated from the cross-polarized image (Onstott et al., 1979).

In Fig. 1, the shown SAR image was acquired on day of the year (DOY) 314, 2019 over a section of the Beaufort Sea. The MYI is identified based on histogram thresholding and visual interpretation from the Sentinel-1 images, as the MYI usually has higher backscatter than FYI. A threshold that separates the two peaks of the histogram of the SAR backscatter coefficient was used to generate an initial map of MYI. Visual interpretation was then implemented to improve the MYI map based on the following principles: (1) back scattering of FYI is usually lower, with brightly linear structures over the dark background; (2) MYI are usually brighter, and the MYI floes have more rounded boundaries (Comiso, 1983). Some misclassifications can be found in the automatically generated MYI map, thus visual interpretation were further used to correct these errors, including those caused by (1) the leads that had thin ice covered with frost flowers would show a bright feature, which could be misclassified as MYI (Nghiem et al., 1997), and (2) the low intensity spots in the interior of MYI floe, which could be misclassified (Shokr and Dabboor, 2024).

3 Methods

The MYI classification using HY-2B data consisted of three steps, including to filter out open water pixels, to obtain the initial MYI classification results using the K-means algorithm, and to reclassify the ice types using fuzzy clustering to improve the accuracy. The procedure of the whole method is shown in Fig. 2.

3.1 *OW filter*

The emissivity contrast between OW and sea ice is high in low-frequency PM data (Eppler et al., 1992). This is because the sea water is highly reflective, emitting little energy when compared with sea ice. Simultaneously, the emissivity of OW increases rapidly with higher frequency, while that of sea ice (i.e., dry FYI) shows almost consistency over all frequencies (Shokr and Sinha, 2023). Thus, the T_b obtained from the 18.7 GHz V-polarization channel was used to filter out OW pixels. We used 18.7 GHz because of its higher spatial resolution due to the smaller footprint size (36 km × 60 km), compared to that (90 km × 150 km) for 6.9 GHz. The difference between the T_b from ocean and sea ice is generally obvious in the freezing season, with the values of approximately 175 K and 250 K, respectively (Fig. 3). Thus, we have the feasibility to apply the Ostu's method presented in Otsu (1975) for the ice-water segmentation. The basic principle of Ostu's method is to find the threshold that minimizes the intra-class variance. Given two classes as C_0 and C_1, and the threshold as t. Assume $\sigma_B^2(t)$ as the classes variance when the threshold is t. Then the method can be transferred to find the maximum of $\sigma_B^2(t)$, which is denoted by

$$\sigma_B^2(t) = \omega_0(t)[\mu_0(t) - \mu_T]^2 + \omega_1(t)[\mu_1(t) - \mu_T]^2, \quad (1)$$

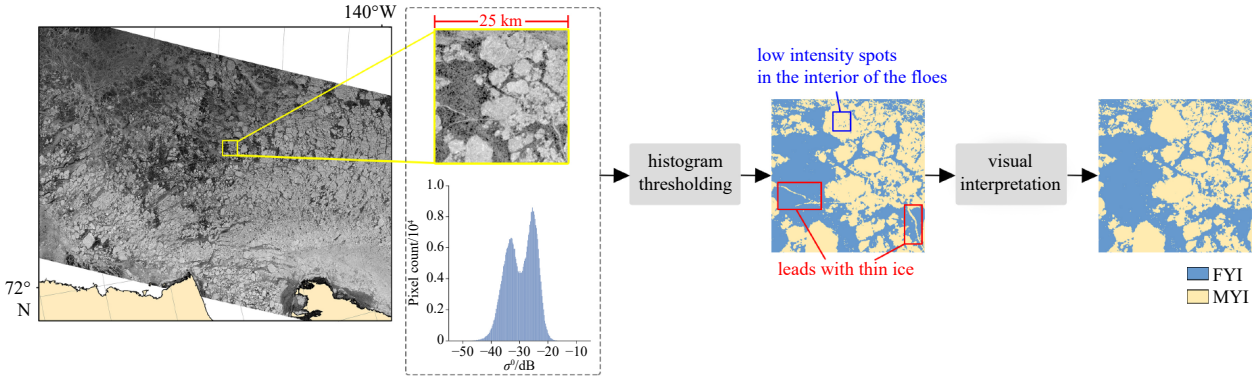


Fig. 1. Procedure of MYI identification from Sentinel-1 SAR images. MYI: multiyear ice; FYI: first-year ice.

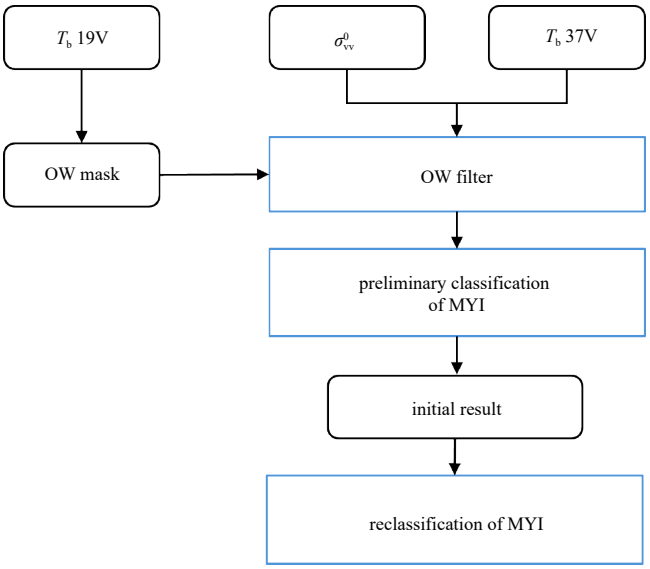


Fig. 2. Flowchart of the K-means+FC method. MYI: multiyear ice; OW: open water; H: horizontal; V: vertical.

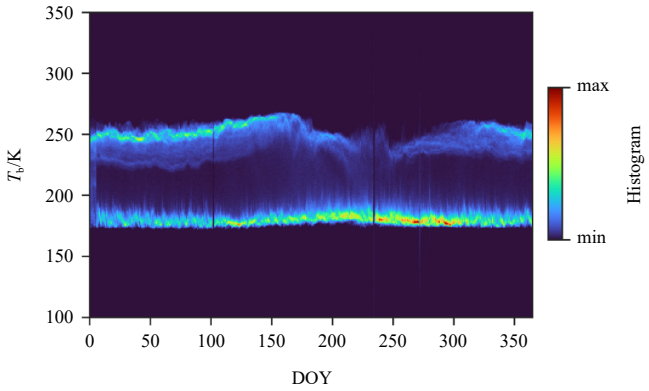


Fig. 3. Time series of the histograms (in color) of HY-2B SMR 18.7V T_b over the Arctic Ocean during 2020. The lower clustered line is from OW, and the top one is from sea ice. SMR: synthetic aperture radar; OW: open water; DOY: day of the year.

where ω_0 and ω_1 are class probabilities, μ_0 and μ_1 are class means, and μ_T is the total mean. This is a nonparametric unsupervised method for computing an automatic threshold in the images with a high day-to-day variability for the T_b distribution.

3.2 *Preliminary classification of MYI*

Preliminary classification of the MYI was carried out using the K-means algorithm. Data from two bands, T_b from HY-2B SMR 37 GHz and the σ^0 from HY-2B SCAT, are jointly used. For the choose of PM data, 37 GHz illustrates the highest sensitivity to differentiate the FYI and MYI (Comiso, 1983; Lindell and Long, 2016b). Joint histogram of T_b versus σ^0 is plotted in Fig. 4. In dry freezing season, the surface scattering layer of MYI engenders relatively higher σ^0 due to strong volume scattering and lower T_b due to the low emissivity than FYI, which has thinner surface scattering layer and higher salinity for the lower layer. The relatively high brine content within the FYI column would result in a higher electromagnetic absorption (Shokr and Sinha, 2023; Onstott, 1992). Thus, MYI and FYI are expected to reveal different clusters in the T_b and σ^0 distribution. As shown in Fig. 4, the distance between cluster centers in the two-dimensional (2D) space of σ^0 VV and T_b from 37V is obviously larger than other paired cases, thus we chose the data from these two channels for the classification between MYI and FYI. In this study, K-means was applied to the 2D data space, in which each point was denoted as a feature vector (σ^0, T_b). The initial number of k-groups was set to 2 as MYI and FYI, and then each feature vector was assigned to the cluster with the closest centroid, according to the Euclidean distance.

3.3 *Reclassification of MYI using fuzzy clustering*

To further refine the classification of the MYI in the transition zone, with a complex challenge because of a disorder mixture with the FYI this region, we reclassify the MYI using a fuzzy clustering approach.

Fuzzy clustering is an unsupervised clustering technique where each data point belongs to a cluster (class), yet with a degree that is specified by a membership grade (Ruspini, 1970). Let $\boldsymbol{X}=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N\}$ be a sample of N observations in $\mathbb{R}^n$. Clustering partitions the dataset into c clusters. For traditional hard partitioning, each feature can belong to only one class, with a membership degree u of 1 or 0 for each cluster. On the contrary, for fuzzy partitioning, each feature vector x can belong to more than one class, with a membership function $u_i(x_k)$ that has a real value between 0 and 1, depending on the "degree of belonging". Fuzzy clustering can be considered as an optimization problem which minimizes a cost function. The most commonly-used method to date is associated with the generalized least-squared errors functional as follows:

$$J(U,v)=\sum_{k=1}^{N}\sum_{i=1}^{c}(u_{ik})^m d_{ik}^2,\quad 1\leqslant m<\infty, \tag{2}$$

where $\boldsymbol{U}=[u_{ik}]$ is the membership matrix, and $\boldsymbol{v}=(\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_c)$ is the vector of cluster centers. m is the weighting exponent, which controls the fuzziness of the clusters. d_{ik}^2 is the squared distance between $\boldsymbol{x}_k$ and $\boldsymbol{v}_i$, which is computed as

$$d_{ik}^2=\|\boldsymbol{x}_k-\boldsymbol{v}_i\|_{\boldsymbol{A}}^2=(\boldsymbol{x}_k-\boldsymbol{v}_i)^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}_k-\boldsymbol{v}_i), \tag{3}$$

where $\boldsymbol{A}$ is the norm-induced matrix controlling the shape of the optimal clusters. For $m=1$, the clustering becomes a hard partitioning with membership $u_{ik}=1$ or 0 (Bezdek et al., 1984). As the value of m increases, the partitions become fuzzier. No theoretical or computational evidence distinguishes an optimal m. However, for most datasets,

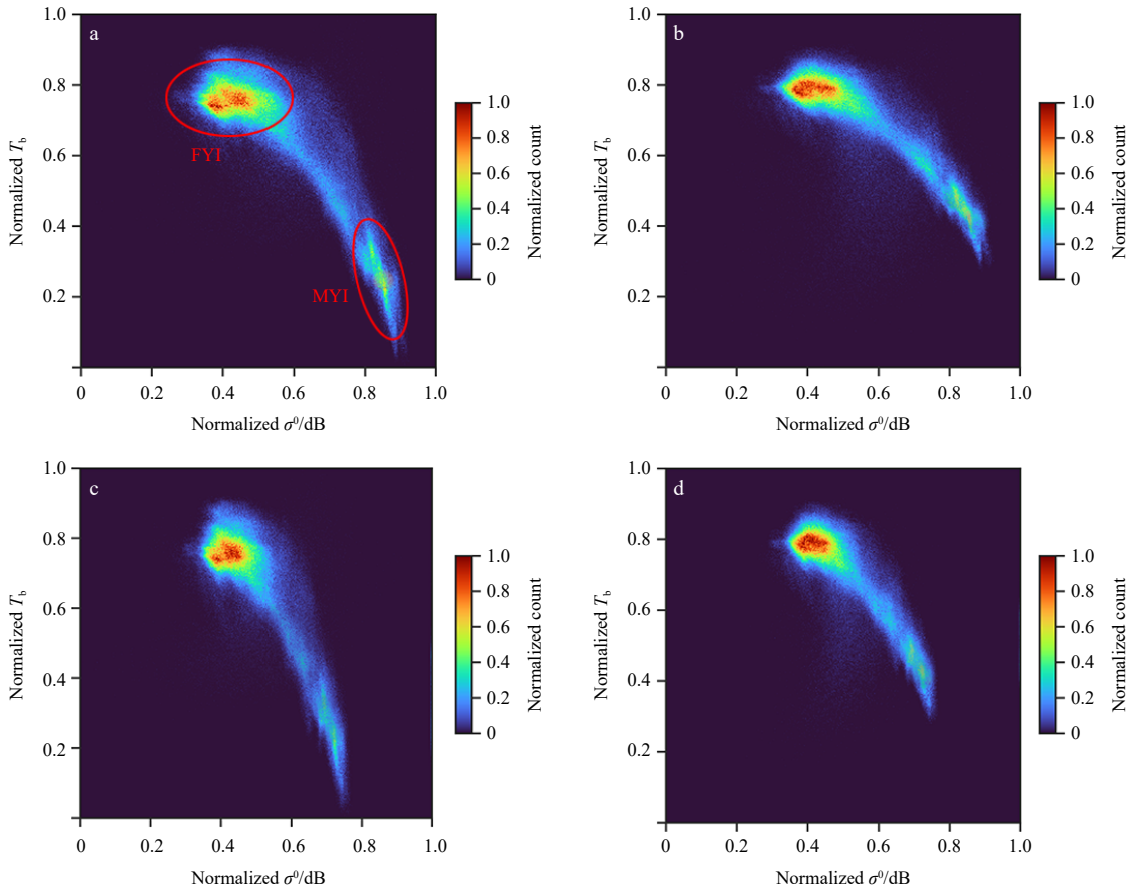


Fig. 4. Joint histograms of $\sigma^0$and T_b from the combination of HY-2B scatterometer and radiometer data over the Arctic Ocean. a. σ^0 VV and T_b from 37V. b. σ^0 VV and T_b from 37H. c. σ^0 HH and T_b from 37V. d. σ^0 HH and T_b from 37H. The histograms are normalized and shown in color. Measurements over land and OW are excluded. MYI: multiyear ice; FYI: first-year ice; OW: open water; HH: horizontal-horizontal; VV: vertical-vertical; H: horizontal; V: vertical.

the $m \in [1.5, 2.5]$ can provide good results (Pal and Bezdek, 1995). Thus, we use m of 2 in this study, which has been proved suitable for most cases (Pal et al., 1993).

For d_{ik}^2, the most-commonly used is the Euclidean distance, which is suitable for hyper-spherical clusters. However, in this study, the classification was based on the fusion of σ^0 and T_b. As shown in Fig. 4, the shapes of the MYI and FYI clusters in the feature space are different, and they do not satisfy the hyper-spherical type. Meanwhile, the ice surface condition is influenced by the environment (e.g., air temperature, wind speed, precipitation, etc.) and historical ice processes (desalination, thickening, and deformation, etc.), which can change the microwave signature of a given ice type. This can cause the distribution of $\sigma^0$and T_b to be unstable. Therefore, it is not appropriate to apply a fixed distance norm in the clustering space. Gustafson and Kessel (1979) extended the standard fuzzy clustering algorithm by employing an adaptive distance norm to detect clusters of different geometrical shapes in a dataset. In this algorithm, each cluster is associated with its own norm-inducing matrix $\boldsymbol{A}_i$, and the expression is given as

$$\boldsymbol{A}_i = |\boldsymbol{F}_i|^{\frac{1}{n}} \boldsymbol{F}_i^{-1}, \tag{4}$$

where $\boldsymbol{F}_i$ is the fuzzy covariance matrix of the i-th cluster, denoted by

$$\boldsymbol{F}_i = \frac{\sum_{k=1}^{N} (u_{ik})^m (\boldsymbol{x}_k - \boldsymbol{v}_i)(\boldsymbol{x}_k - \boldsymbol{v}_i)^{\mathrm{T}}}{\sum_{k=1}^{N} (u_{ik})^m}. \tag{5}$$

Note that the substitution of Eqs (4) and (5) into Eq. (3) gives a generalized squared Mahalanobis distance norm, where the covariance is weighted by the membership degree in U. When computing the algorithm, given the termination tolerance $\varepsilon > 0$ and the iterating time as l. Then the stop criteria of the fuzzy clustering can be set as

$$\|U^l - U^{l-1}\| < \varepsilon. \tag{6}$$

We adopted the fuzzy clustering algorithm to reclassify the ice types in the transition zone to reduce the misclassifications induced by the highly dynamic nature of the ice in this region. The new method is referred to as K-means+FC. Figure 5 shows a case of ice classification de-

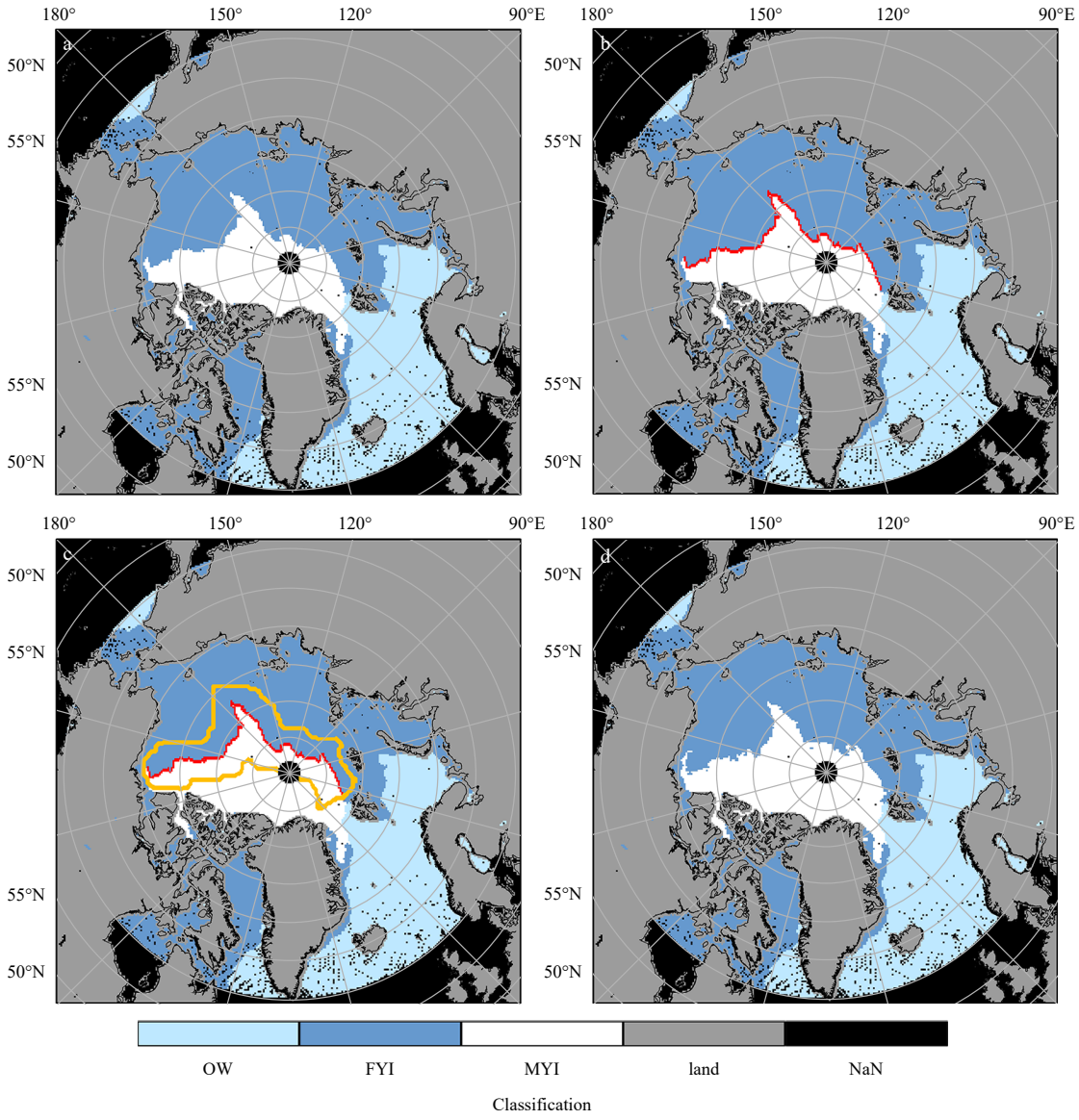


Fig. 5. Initial ice classification of the HY-2B Arctic-wide scene on February 10, 2020 (a), extracted MYI boundaries shown with red pixels (b), the buffer (orange polygon) created by dilation based on the MYI boundaries (c), and final result of ice classification after reclassifying MYI (d). MYI: multiyear ice; FYI: first-year ice; OW: open water. NAN means no data.

rived from the traditional K-means classification or that with additional fuzzy reclassification.

After the K-means classification (Fig. 5a), the fuzzy reclassification involved three steps as follows.

(1) Firstly, the boundaries of the MYI extent were extracted from the initial classification results (Fig. 5b). Here, we ignored the MYI in the south of Greenland Sea because the misclassification of ice types in this region is common due to anomalous backscatter caused by highly deformed ice surface triggered by strong dynamic environments (Lindell and Long, 2016b), and that in the Canadian Arctic Archipelago (CAA) because the coarse-resolution radiometer and scatterometer data is not appropriate for this region with the interlaced narrow waterway.

(2) Secondly, we defined the transition zone of MYI and FYI as a buffer created using inward and outward dilation of the MYI boundaries, with a radius of 250 km (Fig. 5c). The reason for selecting 250 km was that, by combining the initial classification results with the NSIDC SIA dataset, we found that the second-year ice edge extended to about 250 km away from the initial MYI boundary. This means that, within this distance, MYI possibly mixes with FYI at a small scale. Here, we assumed that this mixing behavior shared an inward MYI boundary of the same distance.

(3) Finally, the fuzzy clustering algorithm was applied to the data of $\sigma^0$and T_b in the transition zone to reclassify the MYI. Pixels with a membership degree u to MYI cluster > 0.5 were labeled as MYI; otherwise, pixels were labeled as FYI. Subsequently, the newly classified pixels of MYI and FYI were used to correct the initial

classification map and obtain the final result, as presented in Fig. 5d.

4 Results

Time series of sea ice extent in the Arctic Ocean from HY-2B K-means+FC classifications are shown in Fig. 6. The extent of sea ice is calculated as the summation of the areas of the pixels assigned to each sea ice types. The total sea ice extent illustrates the same pattern for the two freezing seasons, where the ice continues to grow until reaching the maximum in March, which is about 9.4 × 10^6 km^2. However, at the beginning of November 2020, the total ice extent is 5.2 × 10^6 km^2, about 0.8 × 10^6 km^2 lower than that of November 2019. For MYI, it is expected that there should be no increasing trend for MYI extent during freezing season, as the transformation of FYI to MYI only happens in fall (September and October), and the MYI can only decrease through the outflow across various gates. As presented in Fig. 6, the MYI extent shows a declining trend for both freezing seasons. For 2019/2020, MYI extent reduces from about 2.8 × 10^6 km^2 in November to about 2.2 × 10^6 km^2 in April, while the reducing trend is more plateau for 2020/2021. In addition, comparing with freezing season of 2019/2020, the averaged MYI extent for 2020/2021 is 2.3 × 10^6 km^2, dropped by about 0.3 × 10^6 km^2. On the other hand, FYI shows an increasing trend over the whole freezing season, while the rising speed gradually slows. For freezing season of 2019/2020, the FYI extent grows to reach the peak in March at about 7 × 10^6 km^2, then almost plateaus afterwards and even drops a bit at the end of April. However, the FYI extent also peaks at 7 × 10^6 km^2 in 2021, but about one month earlier. After that, it witnesses no more increase. In addition, at the beginning of freezing season 2019/2020, the FYI extent is approximately 3.5 × 10^6 km^2, which is about 0.6 × 10^6 km^2 higher than that of 2020/2021. Nevertheless, the FYI extent for both freezing seasons reach the same gratitude at the end of April, indicating that the FYI grows faster during freezing season 2020/2021.

Figure 7 shows the spatial distribution of sea ice over the Arctic Ocean from 2019 to 2021. At the beginning of the freezing season in 2019/2020, the Kara Sea is covered by FYI, and the FYI coverage in the Nansen Basin is extended to 80°N. On the contrary, there is only small area of FYI along the coast of Kara Sea in November 2020, and almost no sea ice at the north of Svalbard and Fanz Josef Land. This causes the total sea ice extent in November 2020 to be lower than November 2019, as suggested in Fig. 6. During the middle freezing season, it is shown that there is always no sea ice at the north coast of Svalbard. This can be explained that the mean sea ice concentration for freezing season 2021 is much lower and the air temperature is higher than freezing season 2020 at this region (Schmitt and Lüpkes, 2023), leading to the more melt of sea ice.

The spatial distribution of MYI also shows distinctive pattern between the two freezing seasons. Firstly, there is MYI distributed near the north of the East Siberian Sea in winter 2019/2020, but no MYI for winter 2020/2021. Secondly, the MYI extent in Beaufort Sea is higher in 2020/2021, which covers most part of this water and even extents to the east of Chukchi Sea. Finally, there is less MYI at the north of Svalbard and Fram Strait for winter 2020/2021. The difference between the MYI distribution for the two freezing seasons can be explained by the atmospheric circulation anomalies, as suggested in Moore et al. (2022): Higher sea-level pressure over the Arctic Ocean in winter 2020/2021 caused stronger wind towards the west Arctic, moving more MYI from the north of Greenland to Beaufort Sea and Chukchi Sea. In contrast,

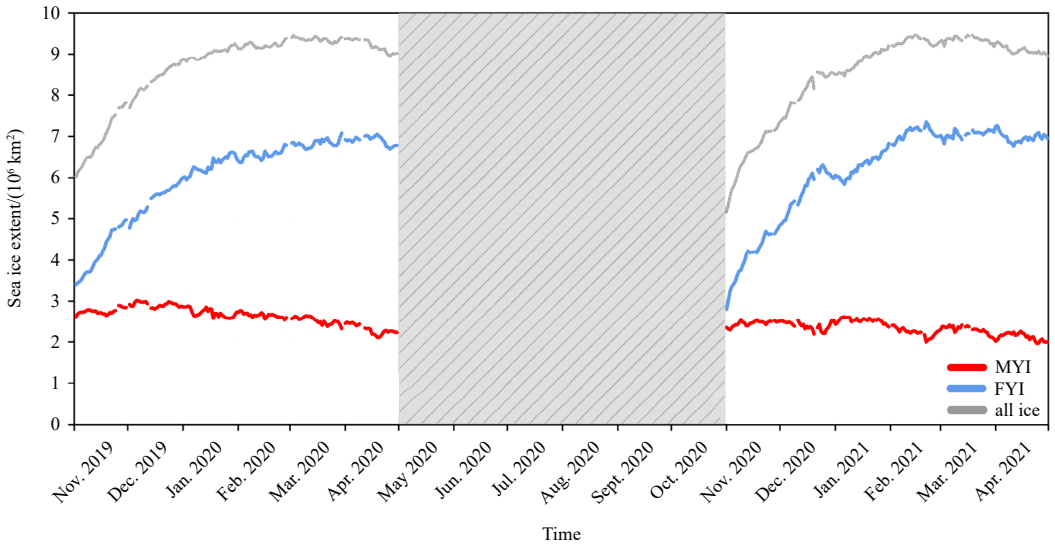


Fig. 6. Time series of MYI extent and total sea ice extent over the Arctic Ocean from November 2019 to April 2021. The data in warm season from May 2020 to October 2020 is ignored. MYI: multiyear ice; FYI: first-year ice.

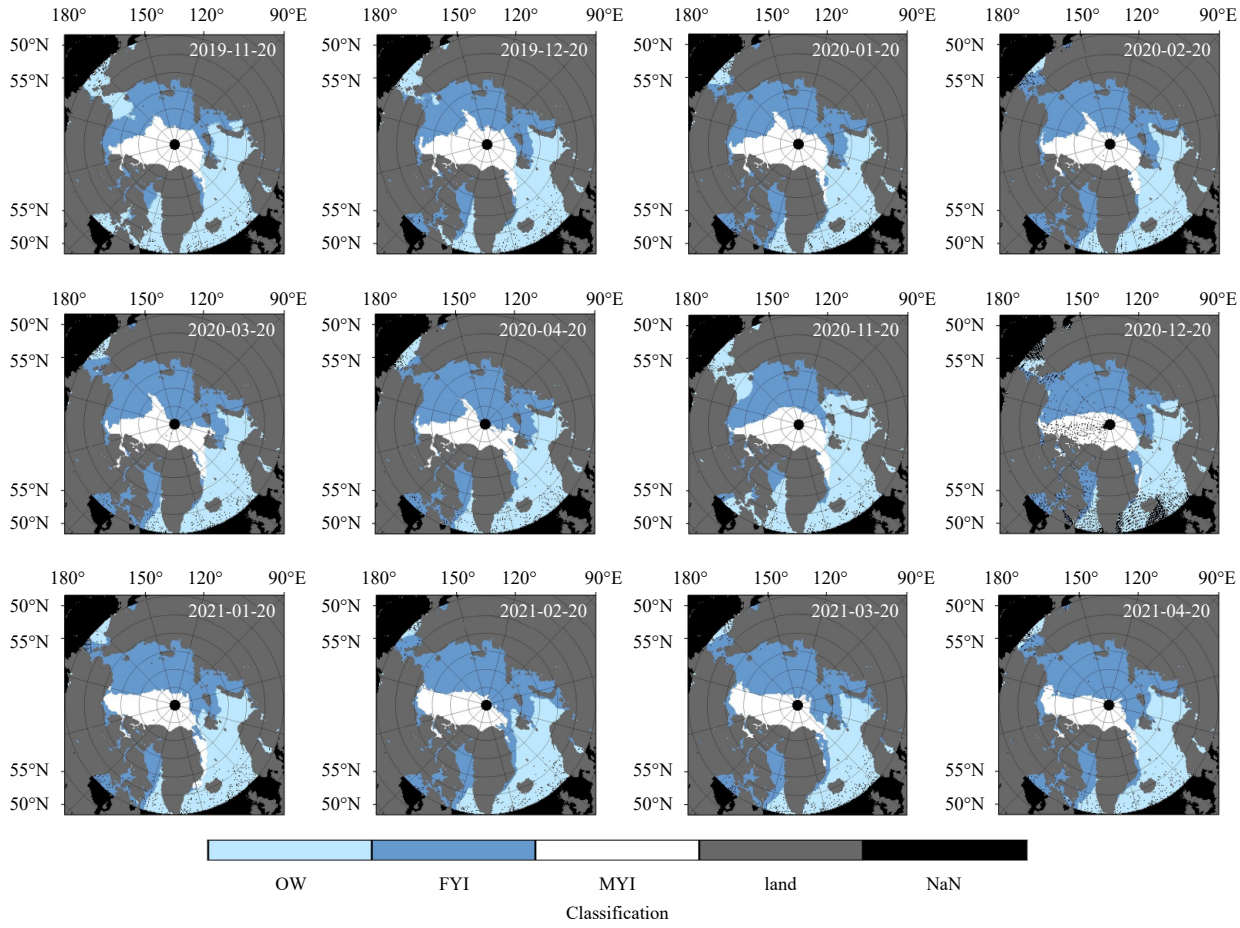


Fig. 7. Spatial distribution of sea ice for freezing seasons of 2019/2020 and 2020/2021. MYI: multiyear ice; FYI: first-year ice; OW: open water. NAN means no data.

winter 2019/2020 witnessed a collapse of Beaufort High, leading to lower sea-level pressures across the Arctic Ocean. This then resulted in a cyclonic wind towards Beaufort Sea.

5 Discussion

This section starts with the temporal and spatial comparisons of the K-means+FC ice classification against ice type products. The OSI SAF global sea ice type product (Section 5.1) and NSIDC SIA dataset (Section 5.2) are compared with K-means+FC results separately over the Arctic Ocean, and an intercomparison between ice type products and K-means+FC results at the transition zone is presented in Section 5.3. Then, we carry out comparisons between K-means and K-means+FC result using ice charts (Section 5.4) and SAR images (Section 5.5) to assess the performance of fuzzy clustering in improving the MYI classification accuracy in the transition zone.

5.1 *Comparison with the OSI SAF global sea ice type product*

Figure 8 compares the MYI extent derived from the K-means+FC method using HY-2B data and the OSI SAF global sea ice type product. The lines represent the MYI extent. Only the MYI in the Central Arctic region is considered. The annual standard deviation (STD) and mean of the MYI extent were calculated for both HY-2B K-means+FC and the OSI SAF product, as shown in Table 3, where the bias refers to the difference of the HY-2B K-means+FC MYI extent minus the OSI SAF MYI extent.

The MYI extent of HY-2B K-means+FC and the OSI SAF product show consistently decreasing trends within both freezing seasons in 2019/2020 and 2020/2021. The decrease of MYI extent in 2020/2021 is slow from November to April. The freezing season of 2020/2021 featured an extremely strong Beaufort High and anticyclonic winds in the Central Arctic. Therefore, the ice floes were more easily captured by the Beaufort Gyre, rather than advected out of the Arctic Ocean through transpolar stream (Mallett et al., 2021). It is expected that no increasing trend of MYI extent should be observed during freezing season. However, for OSI SAF products, there are some distinctive exceptions of increasing trend for MYI extent. For example, a clear increase can be seen in December 2019 for OSI SAF MYI extent. Such anomaly was also observed in Ye et al. (2023) and might be explain by the fact that OSI SAF product uses C-band scatterometer data from ASCAT, which can be less sensitive to second-year ice. A similar decreasing trend of MYI extent is ob-

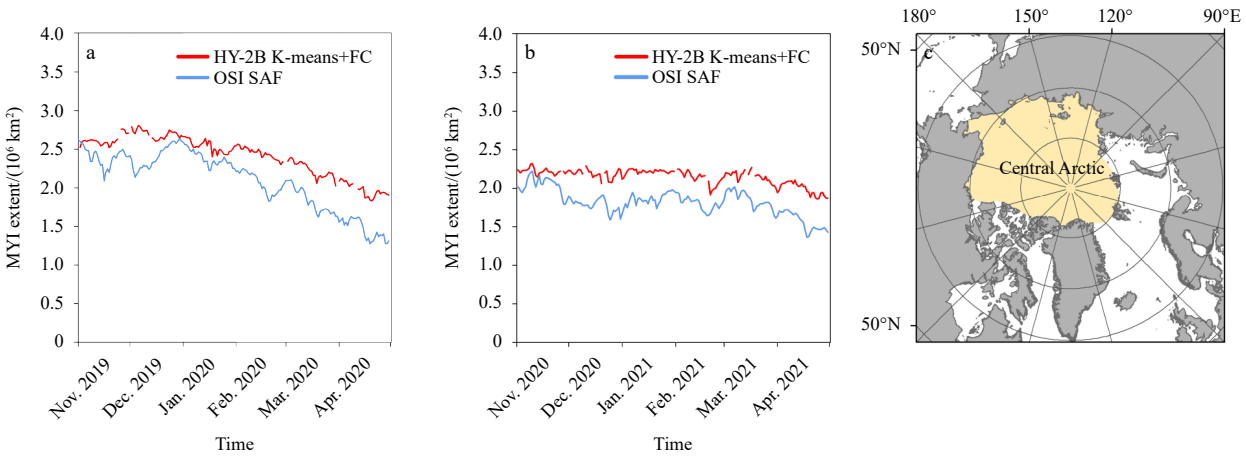


Fig. 8. Time series of the MYI extent from the OSI SAF and HY-2B data using the K-means+FC algorithm in freezing season 2019/2020 (a) and freezing season 2020/2021 (b), and the region of central Arctic (c).

Table 3. Analysis of the MYI extent for HY-2B K-means+FC and OSI SAF

Year	HY-2B K-means+FC			OSI SAF	
	Mean/(10^6 km^2)	STD/(10^6 km^2)	Bias/(10^6 km^2)	Mean/(10^6 km^2)	STD/(10^6 km^2)
2019/2020	2.42	0.26	0–0.65	2.09	0.38
2020/2021	2.14	0.10	0.03–0.64	1.82	0.18

served in both datasets. Meanwhile, the K-means+FC method has a lower STD for both freezing seasons, compared to the OSI SAF product, indicating that fewer day-to-day fluctuations exist when using the proposed method.

The MYI extent from HY-2B K-means+FC is larger than that for the OSI SAF product for both 2019/2020 and 2020/2021. For freezing season 2019/2020, the average MYI extent from HY-2B K-means+FC is 2.42 × 10^6 km^2, exceeding that from OSI SAF by 0.33 × 10^6 km^2. Figure 8a shows that the largest bias in 2019/2020 occurred in April (0.65 ×10^6 km^2). From February to April, the MYI extent difference gradually increased. This can be observed for both freezing seasons, but more significantly in the freezing season of 2019/2020. By the time the end of freezing season approaches, MYI becomes less stable as the surface becomes wetter and may even be covered by melt ponds. This leads to lower separability between MYI and FYI in microwave observations. Fluctuations are more evident in the OSI SAF product, which is derived from the C-band data of ASCAT. In general, the C-band has been found to be less sensitive to the difference between MYI and FYI, compared to the Ku-band of HY-2B (Ezraty and Cavanié, 1999). Consequently, there are more MYI pixels ignored in the OSI SAF product in late freezing season, causing an increasing underestimation of MYI extent.

5.2 *Comparison with the NSIDC SIA dataset*

The MYI extent from the HY-2B K-means+FC algorithm and the NSIDC SIA dataset is shown in Fig. 9. Grid cells classified with an age value of two or greater are labeled as MYI. In general, the MYI extent from the HY-

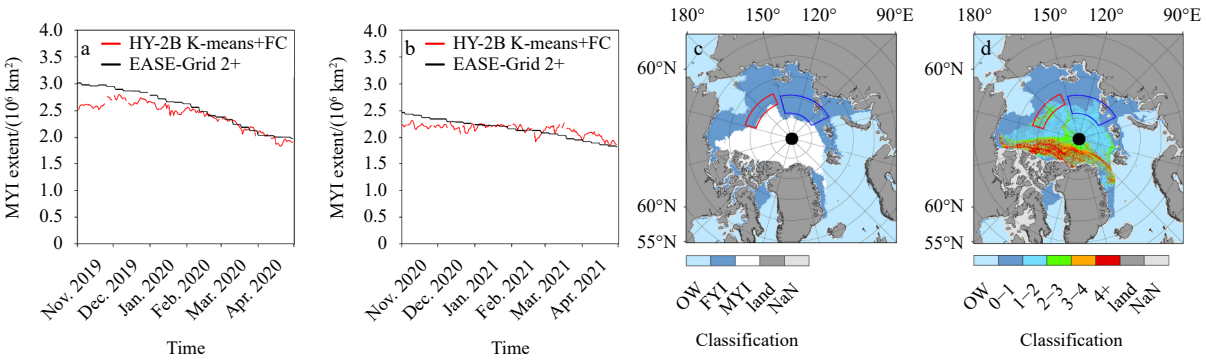


Fig. 9. Central Arctic MYI extent variation from the HY-2B K-means+FC result and the NSIDC SIA dataset result. a. Freezing season of 2019/2020. b. Freezing season of 2020/2021. c. HY-2B K-means+FC result for DOY 321, 2019. d. Sea ice age dataset result for DOY 316 to DOY 322, 2019. The contents of the colored fan-shaped frames are explained in the text. MYI: multiyear ice; FYI: first-year ice; OW:open water; DOY: day of the year. NAN means no data.

2B K-means+FC classification agrees well with that from the SIA dataset, with expected decreasing trends in freezing season being observed. The SIA dataset demonstrates an averaged MYI extent of 2.55×10^6 km^2 and 2.16×10^6 km^2 for freezing season 2019/2020 and 2020/2021, respectively. The averaged bias of HY-2B K-means+FC with the SIA dataset is about 6% for 2019/2020 and about 10% for 2020/2021. The largest difference happens in the fall. For the further analysis, we used the classification result for DOY 321, 2019 to illustrate the difference of the MYI spatial distribution. We note that the SIA dataset reports large areas of MYI extending south to approximately 75°N, whereas the MYI derived from HY-2B K-means+FC is mostly within 75°N.

Two regions are highlighted in Figs 9c and d by fan-shaped frames. For the red frame region, a patch of older ice (of more than three years old) is shown in the SIA dataset result, while almost no MYI is classified in the HY-2B K-means+FC result. The difference can be explained by the fact that the large packs of MYI break up into smaller floes when drifting within the domain of the Beaufort Gyre and become fully mixed with FYI. For coarse-resolution satellite images (e.g., HY-2B), the fraction of MYI within a grid cell in the red frame region is small, and thus this cell might be classified as FYI. Nevertheless, these older ice floes can be identified using the SIA dataset because it uses a Lagrangian parcel tracking algorithm to measure the age of sea ice. Each grid cell in the product is assigned the age of the oldest parcel within it at that time, even if the fraction of this oldest ice is relatively small. However, this phenomenon can in turn bring about an inherent overestimation of MYI area within the SIA dataset, especially for the transition zone of MYI and FYI (Korosov et al., 2018; Tschudi et al., 2016). Another distinct difference is in the blue frame region, where a certain amount of second-year ice is reported in the SIA result but no MYI is seen in the HY-2B K-means+FC result. We attribute this difference partly to the overestimation of the SIA dataset. On the other hand, second-year ice has been reported as having a closer microwave signature to that of FYI than MYI (Comiso, 2006). This means that this type of MYI is more difficult to discriminate from FYI by using only microwave satellite data. Thus, the difference of the MYI distribution in the blue frame region can also be ascribed to an underestimation of second-year ice in the HY-2B K-means+FC result.

5.3 *Intercomparison with ice type products in the transition zone*

Intercomparisons of the HY-2B K-means+FC result, the OSI SAF global sea ice type product, and the SIA dataset result are provided in this section. Considering that the proposed algorithm focuses on the improvement of the MYI classification accuracy around the MYI boundary (i.e., the transition region), these intercomparisons also concentrate on the same region. These three products show an overall agreement with each other for sea ice type distribution. Nevertheless, we found that the largest discrepancy usually appears at the MYI boundary in both Beaufort Sea region and Nansen Basin region.

In Beaufort Sea region (Figs 10a–d), compared with OSI SAF, HY-2B K-means+FC can generally identify more MYI at the transition zone. However, both HY-2B K-means+FC and OSI SAF demonstrate an underestimation of MYI when compared with the SIA dataset, for which the MYI edge extends to 150°W. We attribute this difference to the fact that the SIA dataset tracks ice parcels in the NSIDC Polar Pathfinder Sea Ice Motion dataset. For the OSI SAF product, many areas are identified as the ambiguous class at the MYI boundary, as labeled by the green pixels. In addition, some scattered MYI floes are identified, which are surrounded by FYI. This condition can also be observed in the SIA dataset result, whereas the spatial continuity of the MYI in the HY-2B K-means+FC result is higher. This can be explained by the coarser resolution of the HY-2B images.

We focused our attention on the two regions marked by boxes in Figs 10a–c. For the region in the blue box, compared to the OSI SAF product, the MYI coverage of the HY-2B K-means+FC result is more consistent with that of the SIA dataset result, where more MYI is derived by HY-2B K-means+FC. It is worth noting that this MYI contains a "long tail" of ice older than three years, which is a typical feature of the ice distribution in the Beaufort Sea, indicating that HY-2B K-means+FC can identify more old ice than the OSI SAF product. In the region marked by the black box, HY-2B K-means+FC classifies younger-aged MYI than the OSI SAF product at the transition zone, suggesting the potential of the fuzzy clustering algorithm in improving sea ice type retrieval at the MYI boundary. However, those areas of FYI in between, identified by the SIA dataset result and the OSI SAF product, are ignored by HY-2B K-means+FC, which can again be attributed to the low spatial resolution of the satellite data.

For Nansen Basin region (Figs 10e–h), HY-2B K-means+FC result again show better agreement with SIA dataset, especially in the second-year ice region (marked by red rectangle in Figs 10e–g). However, for OSI SAF product, there are large amount of ambiguous type in the same region. Areas of FYI are also observed, which indicates that OSI SAF product understates MYI area. This condition is the same to the classification results in Beaufort Sea region. Another place we paid attention to is denoted by black rectangle in Figs 10e–g, near the Svalbard Island. There are MYI or ambiguous type being identified in HY-2B K-means+FC and OSI SAF result at latitude lower than 79°N, but no MYI in SIA dataset. It might be explained that in this region, which is near the shore, sea ice is affected by wind and wave. As a result, the roughening of the ice surface leads to anomalous microwave signature in satellite data, thus causing misclassification in

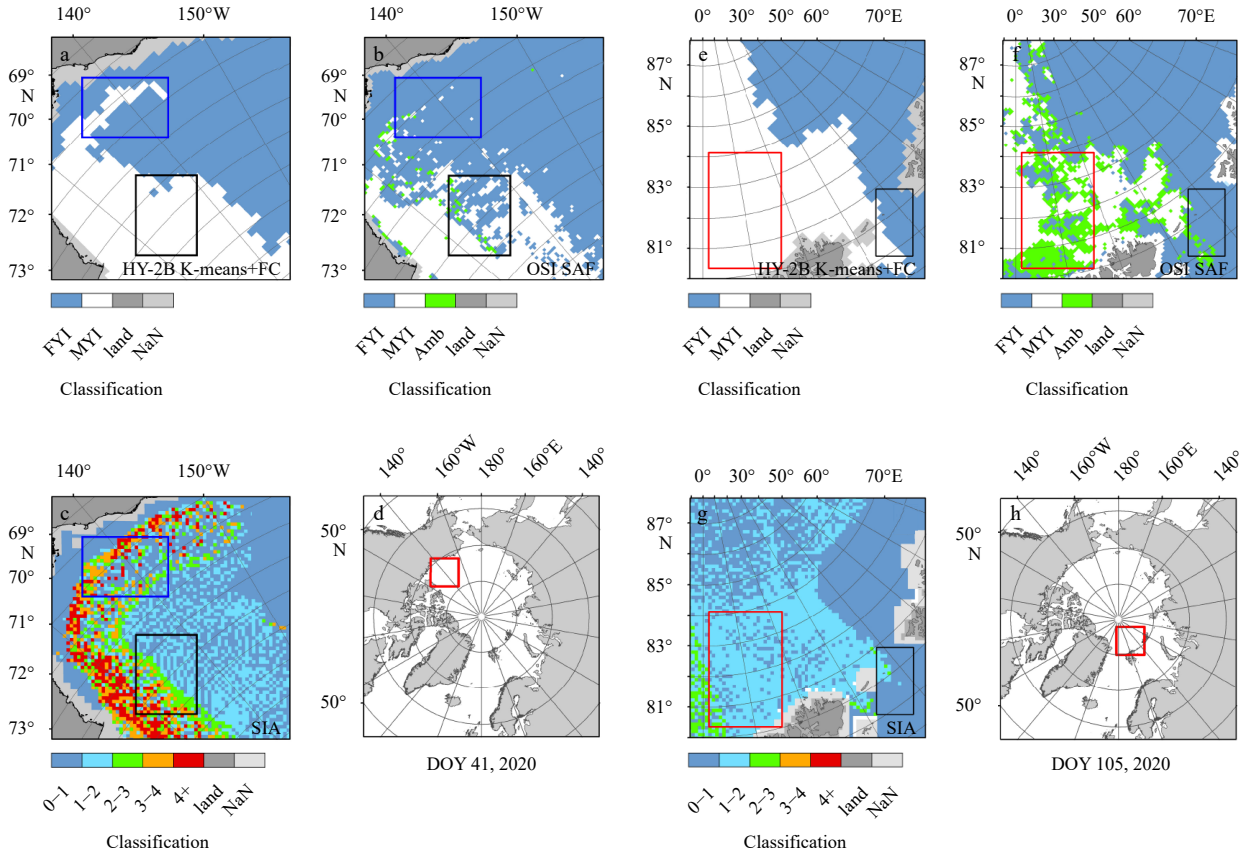


Fig. 10. Regional MYI boundary comparison. a–d. Boundary at Beaufort Sea. e–h. Boundary at Nansen Basin. Red rectangles in d and h represent the locations of the comparing regions. MYI: multiyear ice; FYI: first-year ice; DOY: day of the year. Amb: ambiguous. Blue frames and black frames in a–c are the regions paid special attention to in Beaufort Sea case. Red frames in e–g are the region paid special attention to in Nansen Basin case. NAN means no data.

HY-2B K-means+FC and OSI SAF results.

5.4 *Comparison with CIS ice charts*

The comparison is quantified here using the overall accuracy (OA) and Kappa coefficient, to estimate the agreement of the K-means and K-means+FC results with ice chart MYI. The OA is calculated by summing the number of correctly classified values and dividing by the total number of values, as presented in the following equation:

$$\mathrm{OA} = \frac{\sum_{j=1}^{z} m_{jj}}{Z}, \tag{7}$$

where m_{jj} is the number of items that belong to class j and are correctly classified as class i. Z is the total number of items, and z is the number of classes. The Kappa coefficient measures the agreement between classification and truth values, which is denoted by

$$\mathrm{Kappa} = \frac{\mathrm{OA} - p_\mathrm{e}}{1 - p_\mathrm{e}}, \tag{8}$$

where p_e is the expected agreement which can be calculated as

$$p_\mathrm{e} = \frac{\sum_{j=1}^{z} G_j C_j}{Z^2}, \tag{9}$$

where G_j is the total number of items that truely belong-

Table 4. OA and Kappa coefficient of the HY-2B classification using the two algorithms, compared with the CIS ice charts at the transition zone

Date	K-means		K-means+FC	
	OA/%	Kappa/%	OA/%	Kappa/%
2019-11-10	91.38	82.33	91.12	81.91
2019-12-16	93.18	86.38	94.24	88.49
2020-01-13	88.15	76.41	92.30	84.19
2020-02-10	88.72	77.24	92.33	84.01
2020-03-16	83.55	67.13	89.98	78.60
2020-04-20	88.52	77.09	90.11	80.10
2020-11-10	78.97	59.38	85.52	69.34
2020-12-26	86.78	73.37	87.70	74.85
2021-01-11	81.58	60.33	86.08	68.47
2021-02-15	89.90	72.97	90.75	74.36
2021-03-08	89.93	72.72	91.95	77.21
2021-04-12	91.83	73.57	92.84	75.94

ing to class j, and C_j is the total number of items that predicted to class j.

The results are presented in Table 4. The OA of the MYI classification using K-means+FC is over 89% and 85% in freezing season 2019/2020 and 2020/2021, respectively. There is an increase by approximately 0.9%–6.5% when compared to the classification using K-means. The Kappa coefficient of the K-means+FC results shows an increase of 1.4%–11.5% as the consistency of the MYI classification with the CIS ice charts is improved at the MYI boundaries.

Nevertheless, the OA and Kappa coefficient from the scene of 10 November 2019 decrease slightly after using the fuzzy clustering. A further analysis using Sentinel-1 SAR images was implemented to explain this difference (Fig. 11). Comparing the K-means+FC MYI classification results against the SAR image classification for this day, we found that the decrease of accuracy is due to mislabeling FYI pixels in the ice chart as MYI using the fuzzy clustering algorithm version. These mislabeled pixels are identified by green color within the red rectangle in Fig. 11b. A total of 13 pixels are numbered and labeled by black rectangles in Fig. 11c. For each pixel (grid point of the algorithm), a Sentinel-1 SAR image is displayed in Fig. 11d. Each image covers a 25 km × 25 km area, containing 625 × 625 SAR pixels. We calculated the MYI fraction in each SAR image using histogram thresholding and visual interpretation for each image independently. There are many MYI floes with smaller sizes within each image, of which eight out of 13 have an MYI concentration of over 50%. This indicates that the K-means+FC algorithm can identify smaller and scattered MYI floes within the transition zone. Therefore, this algorithm provides a more reasonable classification

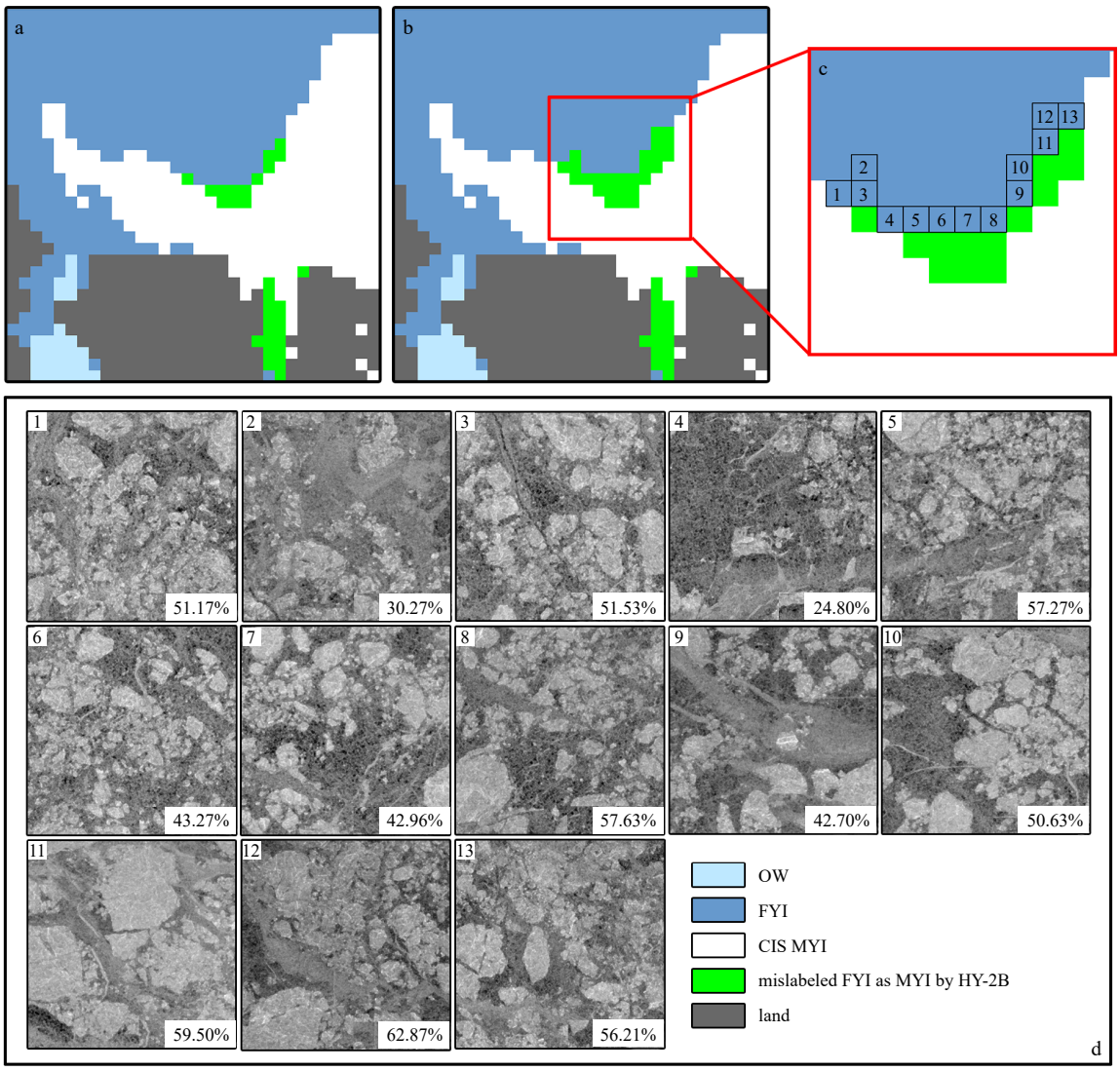


Fig. 11. Data to explain the exception in the scene of 10 November 2019. a. CIS ice charts with misclassified FYI as MYI by the K-means results (green color). b. CIS ice charts with misclassified FYI as MYI by the K-means+FC results (green color). c. MYI pixels causing the decrease of OA and Kappa coefficient, labeled 1–13. d. Sentinel-1 SAR images of these 13 pixels in c, with the MYI concentration shown in the lower right. MYI: multiyear ice; FYI: first-year ice; DOY: day of the year.

than K-means, even though it agrees less with the CIS charts.

5.5 *Comparison with Sentinel-1 SAR images*

Sentinel-1 SAR images were used to compare with the K-means results and K-means+FC results and validate the effect of the fuzzy clustering algorithm. To generate MYI maps from the SAR images with the same spatial resolution as the HY-2B classification results, the following steps were employed: (1) a grid with a cell size of 25 km × 25 km was created within the SAR images, with the same coordinate reference as HY-2B. (2) MYI cells were identified as such when the fraction of MYI exceeded 50%. The K-means+FC classification results were then compared with the SAR-derived MYI maps. The MYI identified in the SAR images covers most parts of the MYI floes, but still some small sizes of MYI are left out due to the coarse resolution of the MYI maps and the generation steps we followed. The OA and Kappa coefficient were calculated to quantify the classification accuracy and showed in Table 5. Six case studies are presented in Fig. 12.

As shown in Table 5, the MYI results from K-means+FC are generally more consistent with SAR MYI, with the averaged OA and Kappa coefficient increase by about 10% and about 24%, respectively. For early freezing season, the sea ice is going through a transition process, in which FYI surviving the last summer will become second-year ice. As shown in SAR images from Figs 12a and b, in the transition zone, MYI floes are small and surrounded by FYI. For the two cases in early freezing season, the MYI in November 2020 are distributed in a narrow region near the coast of the CAA due to a strong westward transport of thick ice from the north of Greenland, which was discussed in Moore et al., 2022. Clear MYI tails are seen in SAR images and successfully identified in K-means+FC results (Figs 12a and b, forth row). However, in K-means results, only MYI near the north boundary of the image scene are recognized, where the MYI floes are relatively large and have less FYI around them, and the tail feature is omitted by K-means algorithm. This makes the OA of K-means results lower than

Table 5. OA and Kappa coefficient of the HY-2B classification using the two algorithms, compared with the Sentinel-1 SAR images at the Beaufort Sea

Period	Date	K-means		K-means+FC	
		OA/%	Kappa/%	OA/%	Kappa/%
Early freezing season	2019-11-28	75.63	22.07	86.97	69.06
	2020-11-03	72.05	30.04	86.46	71.86
Middle freezing season	2020-01-15	85.77	70.82	87.87	75.40
	2021-01-10	92.86	74.87	95.54	83.13
Late freezing season	2020-04-14	82.38	63.40	89.43	73.42
	2021-04-18	62.95	33.63	84.38	66.86

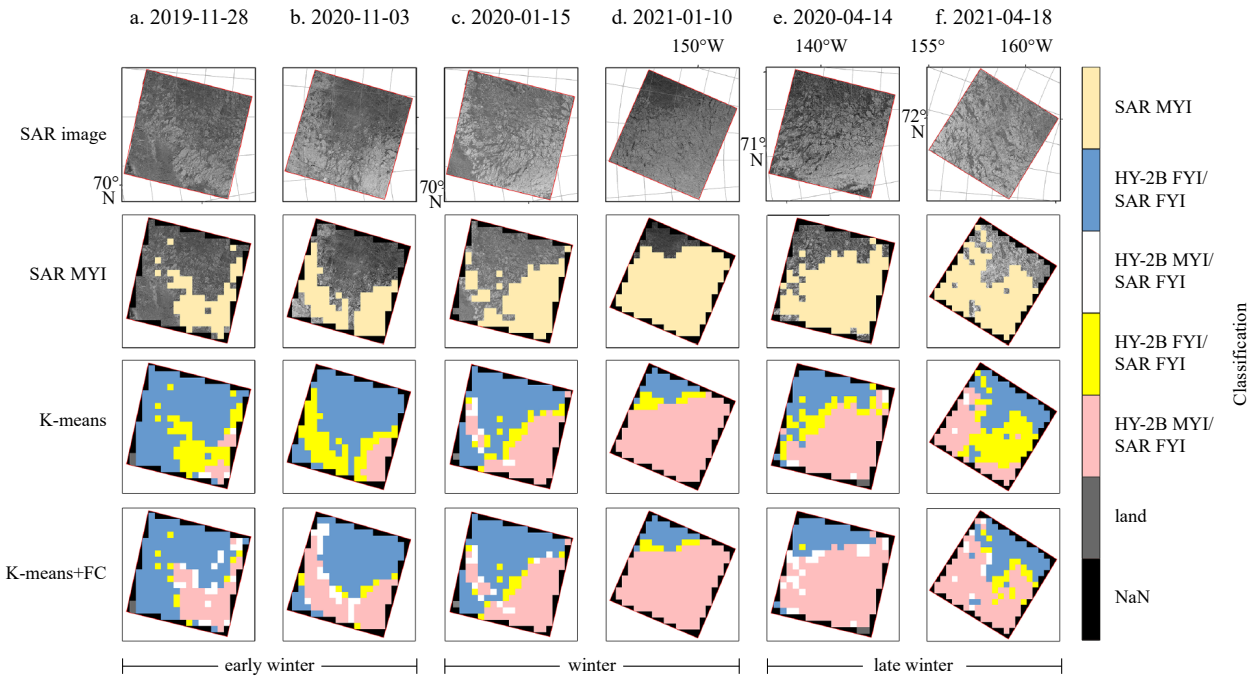


Fig. 12. HY-2B MYI classification results compared with Sentinel-1 SAR images. MYI: multiyear ice; FYI: first-year ice. NAN means no data.

80% and the Kappa coefficient not ideal.

During freezing season (Figs 12c and d), MYI floes have become larger and the boundaries of which are rounded due to continuous collisions. For freezing season 2019/2020 and 2020/2021, the OA and Kappa coefficient of K-means and K-means+FC results are higher than 85% and 70%, respectively. This means that the two algorithms can both correctly identify most part of the MYI. However, by adding fuzzy clustering, the classification of MYI in the transition zone has seen a little improvement, with small and scattered MYI being classified.

For late freezing season, owing to the unusually high sea level pressure over the central Arctic Ocean in the freezing season of 2020/2021 (Mallett et al., 2021), large amount of MYI is continuously transported westward and out of Beaufort Sea, making the MYI boundary be closer to Chukchi Sea. As presented in Figs 12e and f, MYI mix with FYI further, therefore leading to a wider transition zone. This then has some negative impact on the MYI classification accuracy for K-means algorithm. However, by using fuzzy clustering, this situation is largely improved, due to the ability of fuzzy clustering to identify more MYI in the transition zones.

In general, K-means+FC algorithm has shown better performance in MYI classification in the transition zone, especially for early freezing season and late freezing season. During freezing season, the improvement brought by fuzzy clustering is limited, as the boundary between MYI and FYI is much more distinctive, thus the K-means algorithm has the ability to identify MYI with relatively high accuracy. However, there are some misclassifications in K-means+FC results, where some FYI are classified as MYI (white pixels in Fig. 12). This can be attribute to the weighting exponent value we chose, which controls the fuzziness of the clusters. In addition, the FYI in the transition zone also undergo collisions, making the surface of which roughen. Consequently, the microwave signature of these FYI is similar to MYI, causing it to be misclassified.

6 Conclusions

In this study, HY-2B scatterometer and radiometer data were used in Arctic MYI classification for the freezing seasons of 2019/2020 and 2020/2021. An algorithm based on K-means with an additional component of fuzzy clustering (called K-means+FC) was developed to solve the problem of FYI/MYI misclassification at the transition zone, where the two types coexist at a small scale. We compared the HY-2B K-means+FC classification result with the OSI SAF global sea ice type product and the NSIDC SIA dataset result. Generally speaking, the MYI derived from HY-2B K-means+FC agreed well with the OSI SAF product, whereas the MYI extent was observed to be larger than that of the OSI SAF product, especially in late freezing season. The main reason for this was that more MYI in the transition zone was classified by the fuzzy clustering algorithm. Furthermore, the HY-2B scatterometer uses the Ku-band, which has been found to be more sensitive to MYI, and thus can identify more MYI than the C-band data used by the OSI SAF product. The MYI distribution of HY-2B K-means+FC was in line with that of ice with an age of not less than two years in the SIA dataset result, but an underestimation occurred in the identification of second-year ice at the MYI boundary. One possible explanation for this is that the microwave signature of second-year ice is similar to that of FYI, and thus it is more difficult to discriminate this type of ice from FYI (Comiso, 2006). Comparison with the CIS ice charts showed that, by adding fuzzy clustering to reclassify the transition zone, the MYI derived from the HY-2B data is more consistent with the MYI of the CIS ice charts, with the OA increasing by 0.9%–6.5%. Meanwhile, the comparison with interpretation of SAR images indicated that the introduction of the fuzzy clustering module can lead to the identification of smaller MYI floes, which consequently improves the accuracy of MYI classification.

Overall, this study has provided some valuable insights into sea ice type retrieval using HY-2B satellite data. We have paid special attention to the MYI identification in the transition zone, a region with highly dynamic ice conditions, with the accurate identification of MYI in which is crucial for the monitoring of MYI annual cycles and budget (Kwok, 2004). Another concern is that the spatial resolution of HY-2B data is relatively coarse, which will cause the mixed pixel problem and has impact on ice type retrieval. However, study on the spatial resolution enhancement of HY-2B data have been proposed and indicates the possibility of increasing the spatial resolution of HY-2B data while still obtaining the data quality (He et al., 2023). Therefore, the study on the application of HY-2B data in ice type retrieval is worthy and practical. Our further work might be concentrated on using the HY-2B data with higher spatial resolution to produce ice type product.

Acknowledgements

The authors would like to thank the Chinese National Satellite Ocean Application Service for providing HY-2B data, Canadian Ice Service for providing the Arctic Regional Sea Ice Charts, Ocean and Sea Ice Satellite Application Facility for providing the Global Sea Ice Type Product, the American National Snow and Ice Data Center for providing EASE-Grid Sea Ice Age dataset (Version 4), and the European Space Agency for providing Sentinel-1 data.

References

Aaboe S, Down E J, Eastwood S. 2021a. Product User Manual for the Global Sea-ice Edge and Type Product. Oslo, Nor-

way: Norwegian Meteorological Institute

Aaboe S, Down E J, Eastwood bS. 2021. Validation Report for the Global Sea-ice Edge and Type Product. Oslo, Norway: Norwegian Meteorological Institute

Anderson H S, Long D G. 2005. Sea ice mapping method for SeaWinds. IEEE Transactions on Geoscience and Remote Sensing, 43(3): 647–657, doi: 10.1109/TGRS.2004.842017

Babb D G, Galley R J, Kirillov S, et al. 2023. The stepwise reduction of multiyear sea ice area in the Arctic Ocean since 1980. Journal of Geophysical Research: Oceans, 128(10): e2023JC020157, doi: 10.1029/2023JC020157

Bezdek J C, Ehrlich R, Full W. 1984. FCM: The fuzzy *c*-means clustering algorithm. Computers & Geosciences, 10(2–3): 191–203

Bird K J, Charpentier R R, Gautier D L, et al. 2008. Circum-Arctic Resource Appraisal: Estimates of Undiscovered Oil and Gas North of the Arctic Circle. Reston, VA, USA: U. S. Geological Survey

Chen Yizhuo, Pang Xiaoping, Ji Qing, et al. 2024. Retrieval of Antarctic sea ice freeboard and thickness from HY-2B satellite altimeter data. Acta Oceanologica Sinica, 43(3): 87–101, doi: 10.1007/s13131-023-2250-2

Comiso J C. 1983. Sea ice effective microwave emissivities from satellite passive microwave and infrared observations. Journal of Geophysical Research: Oceans, 88(C12): 7686–7704, doi: 10.1029/JC088iC12p07686

Comiso J C. 1986. Characteristics of Arctic winter sea ice from satellite multispectral microwave observations. Journal of Geophysical Research: Oceans, 91(C1): 975–994, doi: 10.1029/JC091iC01p00975

Comiso J C. 2006. Impacts of the variability of second-year ice types on the decline of the Arctic perennial sea-ice cover. Annals of Glaciology, 44: 375–382, doi: 10.3189/172756406781811673

Comiso J C. 2012. Large decadal decline of the Arctic multiyear ice cover. Journal of Climate, 25(4): 1176–1193, doi: 10.1175/JCLI-D-11-00113.1

Eppler D T, Farmer L D, Lohanick A W, et al. 1992. Passive microwave signatures of sea ice. Geophysical Monograph Series, 68: 47–71

Ezraty R, Cavanié A. 1999. Intercomparison of backscatter maps over Arctic sea ice from NSCAT and the ERS scatterometer. Journal of Geophysical Research: Oceans, 104(C5): 11471–11483, doi: 10.1029/1998JC900086

Gao Shuo, Li Zhen, Chen Quan, et al. 2019. Inter-sensor calibration between HY-2B and AMSR2 passive microwave data in land surface and first result for snow water equivalent retrieval. Sensors, 19(22): 5023, doi: 10.3390/s19225023

Gloersen P, Nordberg W, Schmugge T J, et al. 1973. Microwave signatures of first-year and multiyear sea ice. Journal of Geophysical Research, 78(18): 3564–3572, doi: 10.1029/JC078i018p03564

Gustafson D E, Kessel W C. 1979. Fuzzy clustering with a fuzzy covariance matrix. In: 1978 IEEE Conference on Decision and Control Including the 17th Symposium on Adaptive Processes. San Diego, CA, USA: IEEE, 761–766

Han Lu, Chen Haihua, Guan Lei, et al. 2023. Multiple sea ice type retrieval using the HaiYang-2B scatterometer in the Arctic. Remote Sensing, 15(3): 678, doi: 10.3390/rs15030678

He Mingyao, Yin Xiaobin, Li Yan, et al. 2023. Spatial resolution enhancement of HY-2B scanning microwave radiometer low-frequency data. IEEE Transactions on Geoscience and Remote Sensing, 61: 5302615

Hovelsrud G K, Poppel B, Van Oort B, et al. 2011. Arctic societies, cultures, and peoples in a changing cryosphere. Ambio, 40(1): 100–110, doi: 10.1007/s13280-010-0105-5

Ignatov A, Laszlo I, Harrod E D, et al. 2004. Equator crossing times for NOAA, ERS and EOS sun-synchronous satellites. International Journal of Remote Sensing, 25(23): 5255–5266, doi: 10.1080/01431160410001712981

IPCC. 2022. The Ocean and Cryosphere in a Changing Climate. Cambridge: Cambridge University Press, 203–320

Jiang Chengfei, Lin Mingsen, Cao Ruixue, et al. 2023. Classification of ice and water in the Arctic using radar altimeter and microwave radiometer data from HY-2B satellite. Acta Oceanologica Sinica, 42(5): 179–191, doi: 10.1007/s13131-022-2067-4

Jiang Chengfei, Lin Mingsen, Wei Hao. 2019. A study of the technology used to distinguish sea ice and seawater on the Haiyang-2A/B (HY-2A/B) altimeter data. Remote Sensing, 11(12): 1490, doi: 10.3390/rs11121490

Korosov A A, Rampal P, Pedersen L T, et al. 2018. A new tracking algorithm for sea ice age distribution estimation. The Cryosphere, 12(6): 2073–2085, doi: 10.5194/tc-12-2073-2018

Kwok R. 2004. Annual cycles of multiyear sea ice coverage of the Arctic Ocean: 1999–2003. Journal of Geophysical Research: Oceans, 109(C11): C11004

Kwok R. 2018. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). Environmental Research Letters, 13(10): 105005, doi: 10.1088/1748-9326/aae3ec

Lindell D B, Long D G. 2016a. Multiyear Arctic sea ice classification using OSCAT and QuikSCAT. IEEE Transactions on Geoscience and Remote Sensing, 54(1): 167–175, doi: 10.1109/TGRS.2015.2452215

Lindell D B, Long D G. 2016b. Multiyear Arctic ice classification using ASCAT and SSMIS. Remote Sensing, 8(4): 294, doi: 10.3390/rs8040294

Mallett R D C, Stroeve J C, Cornish S B, et al. 2021. Record winter winds in 2020/21 drove exceptional Arctic sea ice transport. Communications Earth & Environment, 2(1): 149

Marsan D, Stern H, Lindsay R, et al. 2004. Scale dependence and localization of the deformation of Arctic sea ice. Physical Review Letters, 93(17): 178501, doi: 10.1103/PhysRevLett.93.178501

Maslanik J, Stroeve J, Fowler C, et al. 2011. Distribution and trends in Arctic sea ice age through spring 2011. Geophysical Research Letters, 38(13): L13502

Meier W N, Hovelsrud G K, Van Oort B E H, et al. 2014. Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity. Reviews of Geophysics, 52(3): 185–217, doi: 10.1002/2013RG000431

Meier W N, Perovich D, Farrell S, et al. 2021. Sea ice. Washington, DC, USA: NOAA

Melsheimer C, Spreen G. 2022. IUP Multiyear Ice Concentration and Other Sea Ice Types, Version 1.1 (Arctic)/Ver-

sion AQ2 (Antarctic)-User Guide. Bremen, Germany: Institute of Environmental Physics, University of Bremen

Moore G W K, Steele M, Schweiger A J, et al. 2022. Thick and old sea ice in the Beaufort Sea during summer 2020/21 was associated with enhanced transport. Communications Earth & Environment, 3(1): 198

Newton R, Pfirman S, Tremblay L B, et al. 2021. Defining the "ice shed" of the Arctic Ocean's Last Ice Area and its future evolution. Earth's Future, 9(9): e2021EF001988, doi: 10.1029/2021EF001988

Nghiem S V, Martin S, Perovich D K, et al. 1997. A laboratory study of the effect of frost flowers on C band radar backscatter from sea ice. Journal of Geophysical Research: Oceans, 102(C2): 3357–3370, doi: 10.1029/96JC03208

Onstott R G. 1992. SAR and scatterometer signatures of sea ice. Geophysical Monograph Series, 68: 73–104

Onstott R G, Moore R K, Weeks W F. 1979. Surface-based scatterometer results of Arctic sea ice. IEEE Transactions on Geoscience Electronics, 17(3): 78–85, doi: 10.1109/TGE.1979.294616

Otsu N. 1975. A threshold selection method from gray-level histograms. Automatica, 11: 23–27

Pal N R, Bezdek J C. 1995. On cluster validity for the fuzzy c-means model. IEEE Transactions on Fuzzy systems, 3(3): 370–379, doi: 10.1109/91.413225

Pal N R, Bezdek J C, Tsao E C K. 1993. Generalized clustering networks and Kohonen's self-organizing scheme. IEEE Transactions on Neural Networks, 4(4): 549–557, doi: 10.1109/72.238310

Rigor I G, Wallace J M. 2004. Variations in the age of Arctic sea-ice and summer sea-ice extent. Geophysical Research Letters, 31(9): L09401

Ruspini E H. 1970. Numerical methods for fuzzy clustering. Information Sciences, 2(3): 319–350, doi: 10.1016/S0020-0255(70)80056-1

Schmitt A U, Lüpkes C. 2023. Attributing near-surface atmospheric trends in the Fram Strait region to regional sea ice conditions. The Cryosphere, 17(8): 3115–3136, doi: 10.5194/tc-17-3115-2023

Shokr M, Dabboor M. 2024. Angular variation of SAR polarimetric parameters over multiyear ice. International Journal of Remote Sensing, 45(22): 7996–8014, doi: 10.1080/01431161.2024.2394237

Shokr M, Lambe A, Agnew T. 2008. A new algorithm (ECICE) to estimate ice concentration from remote sensing observations: An application to 85-GHz passive microwave data. IEEE Transactions on Geoscience and Remote Sensing, 46(12): 4104–4121, doi: 10.1109/TGRS.2008.2000624

Shokr M, Sinha N K. 2023. Sea Ice: Physics and Remote Sensing. 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, 284–336

Swan A M, Long D G. 2012. Multiyear Arctic sea ice classification using QuikSCAT. IEEE Transactions on Geoscience and Remote Sensing, 50(9): 3317–3326, doi: 10.1109/TGRS.2012.2184123

Tschudi M A, Meier W N, Stewart J S, et al. 2019a. EASE-Grid Sea Ice Age, Version 4. Boulder, CO, USA: NASA National Snow and Ice Data Center Distributed Active Archive Center

Tschudi M A, Meier W N, Stewart J S, et al. 2019b. Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 4. Boulder, Colorado USA: NASA National Snow and Ice Data Center Distributed Active Archive Center

Tschudi M A, Meier W N, Stewart J S. 2020. An enhancement to sea ice motion and age products at the National Snow and Ice Data Center (NSIDC). The Cryosphere, 14(5): 1519–1536, doi: 10.5194/tc-14-1519-2020

Tschudi M A, Stroeve J C, Stewart J S. 2016. Relating the age of Arctic sea ice to its thickness, as measured during NASA's ICESat and IceBridge campaigns. Remote Sensing, 8(6): 457, doi: 10.3390/rs8060457

Walker N P, Partington K C, Van Woert M L, et al. 2006. Arctic sea ice type and concentration mapping using passive and active microwave sensors. IEEE Transactions on Geoscience and Remote Sensing, 44(12): 3574–3584, doi: 10.1109/TGRS.2006.881116

Wang Shishuai, Zhou Wu, Yin Xiaobin, et al. 2022. Stability of the HY-2B scanning microwave radiometer (SMR) brightness temperature using a modified vicarious cold reference. IEEE Transactions on Geoscience and Remote Sensing, 60: 5303916

Wensnahan M, Maykut G A, Grenfell T C, et al. 1993. Passive microwave remote sensing of thin sea ice using principal component analysis. Journal of Geophysical Research: Oceans, 98(C7): 12453–12468, doi: 10.1029/93JC00939

WMO, IOC-UNESCO. 2014. Ice Chart Colour Code Standard, WMO/TD-No. 1215. Geneva: World Meteorological Organization

Ye Yufang. 2022. Arctic sea ice type based on active and passive microwave data (2002–2017) (in Chinese). https://www.doi.org/10.11878/db.202210.007362[2022-10-25]

Ye Yufang, Luo Yanbing, Sun Yan, et al. 2023. Inter-comparison and evaluation of Arctic sea ice type products. The Cryosphere, 17(1): 279–308, doi: 10.5194/tc-17-279-2023

Yu P, Clausi D A, Howell S E L. 2009. Fusing AMSR-E and QuikSCAT imagery for improved sea ice recognition. IEEE Transactions on Geoscience and Remote Sensing, 47(7): 1980–1989, doi: 10.1109/TGRS.2009.2013632

Yu Rui, Lu Hailiang, Li Shujing, et al. 2022. Instrument design and early in-orbit performance of HY-2B scanning microwave radiometer. IEEE Transactions on Geoscience and Remote Sensing, 60: 5301213

Zhang Yi, Mu Bo, Lin Mingsen, et al. 2021. An evaluation of the Chinese HY-2B satellite's microwave scatterometer instrument. IEEE Transactions on Geoscience and Remote Sensing, 59(6): 4513–4521, doi: 10.1109/TGRS.2020.3008405

Zhang Zhilun, Yu Yining, Li Xinqing, et al. 2019. Arctic sea ice classification using microwave scatterometer and radiometer data during 2002–2017. IEEE Transactions on Geoscience and Remote Sensing, 57(8): 5319–5328, doi: 10.1109/TGRS.2019.2898872

Zhang Zhilun, Yu Yining, Shokr M, et al. 2022. Intercomparison of Arctic sea ice backscatter and ice type classification using Ku-band and C-band scatterometers. IEEE Transactions on Geoscience and Remote Sensing, 60: 4301718