remote sensing

Article

Classification of Arctic Sea Ice Type in CFOSAT Scatterometer Measurements Using a Random Forest Classifier

Xiaochun Zhai [1,2], Rui Xu [3], Zhixiong Wang [4], Zhaojun Zheng [1,2], Yixuan Shou [1,2], Shengrong Tian [5], Lin Tian [1,2], Xiuqing Hu [1,2], Lin Chen [1,2] and Na Xu [1,2,*]

1 Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center (National Center for Space Weather), China Meteorological Administration, Beijing 100081, China
2 Innovation Center for FengYun Meteorological Satellite (FYSIC), Beijing 100081, China
3 Department of Marine Technology, Ocean University of China, Qingdao 266100, China
4 School of Marine Sciences, Nanjing University of Information Science and Technology, Nanjing 210044, China
5 Beijing Huayun Shinetek Science and Technology Co., Ltd., Beijing 100081, China
* Correspondence: xuna@cma.gov.cn; Tel.: +86-010-6840-0907

Abstract: The Ku-band scatterometer called CSCAT onboard the Chinese–French Oceanography Satellite (CFOSAT) is the first spaceborne rotating fan-beam scatterometer (RFSCAT). A new algorithm for classification of Arctic sea ice types on CSCAT measurement data using a random forest classifier is presented. The random forest classifier is trained on the National Snow and Ice Data Center (NSIDC) weekly sea ice age and sea ice concentration product. Five feature parameters, including the mean value of horizontal and vertical polarization backscatter coefficient, the standard deviation of horizontal and vertical polarization backscatter coefficient and the copol ratio, are innovatively extracted from orbital measurement for the first time to distinguish water, first-year ice (FYI) and multi-year ice (MYI). The overall accuracy and kappa coefficient of sea ice type model are 93.35% and 88.53%, respectively, and the precisions of water, FYI, and MYI are 99.67%, 86.60%, and 79.74%, respectively. Multi-source datasets, including daily sea ice type from the EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF), NSIDC weekly sea ice age, multi-year ice concentration (MYIC) provided by the University of Bremen, and SAR-based sea ice type released by Copernicus Marine Environment Monitoring Service (CMEMS) have been used for comparison and validation. It is shown that the most obvious difference in the distribution of sea ice types between the CSCAT results and OSI SAF sea ice type are mainly concentrated in the marginal zones of FYI and MYI. Furthermore, compared with OSI SAF sea ice type, the area of MYI derived from CSCAT is more homogeneous with less noise, especially in the case of younger multiyear ice. In the East Greenland region, CSCAT identifies more pixels as MYI with lower MYIC values, showing better accuracy in the identification of areas with obvious mobility of MYI. In conclusion, this research verifies the capability of CSCAT in monitoring Arctic sea ice classification, especially in the spatial homogeneity and detectable duration of sea ice classification. Given the high accuracy and processing speed, the random forest-based algorithm can offer good guidance for sea ice classification with FY-3E/RFSCAT, i.e., a dual-frequency (Ku and C band) scatterometer called WindRAD.

Keywords: CFOSAT; CSCAT; scatterometer; sea ice classification; random forest classifier

Citation: Zhai, X.; Xu, R.; Wang, Z.; Zheng, Z.; Shou, Y.; Tian, S.; Tian, L.; Hu, X.; Chen, L.; Xu, N. Classification of Arctic Sea Ice Type in CFOSAT Scatterometer Measurements Using a Random Forest Classifier. *Remote Sens.* **2023**, *15*, 1310. https://doi.org/10.3390/rs15051310

Academic Editors: Tonghua Wu, Florent Garnier, Lin Liu and Yubao Qiu

Received: 5 January 2023
Revised: 14 February 2023
Accepted: 20 February 2023
Published: 27 February 2023

1. Introduction

According to the length of time that sea ice exists, it is usually divided into young ice, first-year ice (FYI) and multi-year ice (MYI). FYI refers to the sea ice that only exists in winter within one year and completely melts in summer, while MYI refers to the sea ice that has survived at least one summer melting period. Relevant studies have found [1] that the decrease in Arctic sea ice thickness and volume is mainly caused by the reduction of MYI, which further increases the sensitivity of Arctic sea ice to climate change [2] and has a far-reaching impact on climate and weather changes in the northern hemisphere [3]. There are significant differences in the effects of radiation and air–sea exchange between different ice surface types in different regions of the Arctic, thus affecting the Arctic atmospheric circulation, especially the polar vortex. Therefore, the study of Arctic sea ice type and its change laws is of great significance to accurately understand the evolution of polar sea ice on the global climate system and ecosystem.

Microwave remote sensing has become the most effective means of sea ice monitoring and ice condition assessment because of its advantages of all-weather detection [4,5]. As a typical active microwave remote sensing payload, the scatterometer is most widely used in the sea ice type monitoring. The scatterometer distinguishes FYI and MYI according to the difference in their backscatter coefficients. The physical mechanism is that after repeated melting and freezing, the salt within the MYI precipitates continuously, leaving more bubbles in the ice layer. In contrast, the FYI generally has a higher salinity and dielectric constant, and the penetration depth is smaller, resulting in MYI having stronger volume scattering than FYI [4,6,7]. In addition, the surface of MYI is rougher than that of FYI, resulting in stronger surface scattering of MYI than FYI. Based on the above two characteristics, the backscattering coefficient of MYI is significantly higher than that of FYI in winter.

Due to the differences in the scatterometer frequency and scanning mechanism, various sea ice classification algorithms have been proposed in previous studies using scatterometer measurements. Table 1 summarizes the spaceborne scatterometers and related studies on sea ice classification from 1991 to early 2021.

Table 1. Summary of spaceborne scatterometers and related studies on sea ice classification from 1991 to early 2021.

Mission	ERS-1/2	ADEOS-1	QuikSCAT	METOP	OceanSAT-II	HY-2A	CFOSAT
Scatterometer	AMI	NSCAT	SeaWinds	ASCAT	OSCAT	SCAT	CSCAT
Date	1991.7–2000.3 1995.4–2011.5	1996.8–1997.6	1999.6–2009.11	2007.6-now	2009.10–2014.2	2011.8–2020.11	2018.10-now
Institute	ESA	JAXA and NASA	NASA	ESA	ISRO	NSOAS	NSOAS and CNES
Frequency (band)	5.3 GHz (C)	13.995 GHz (Ku)	13.4 GHz (Ku)	5.3 GHz (C)	13.515 GHz (Ku)	13.255 GHz (Ku)	13.256 GHz (Ku)
Beam type	Fixed fan-beam	Fixed fan-beam	Rotating pencil-beam	Fixed fan-beam	Rotating pencil-beam	Rotating pencil-beam	Rotating fan-beam
polarization	3 VV	3 VV $\times$ 2 1 HH $\times$ 2	HH (inner) VV (outer)	3 VV $\times$ 2	HH (inner) VV (outer)	HH (inner) VV (outer)	HH VV
Incidence angles	18–59°	17–60°	46°, 54.4°	25–65°	49°, 57°	41°, 48°	28–51°
Sea ice type algorithm	-	K-means	Fixed threshold algorithm; dynamic threshold algorithm; ECICE; K-means	Bayesian classification algorithm; improved ECICE; K-means	Improved dynamic threshold algorithm	Dynamic threshold algorithm; BP Neural network classification algorithm	K-means; random forest classification algorithm; tree augmented naive Bayesian sea ice classification algorithm
References	[8–10]	[9,11–13]	[14–19]	[20–22]	[20,23]	[24,25]	[26–28]

Considering sea ice classification using a C-band scatterometer and the successive damage to QuikSCAT/SeaWinds and OSCAT, Lindell and long [20] studied the C-band ASCAT scatterometer data in order to continuously update the records of sea ice type datasets. It was found that the discrimination of sea ice classification using the C-band scatterometer is lower than that of Ku-band scatterometer in winter, and the backscatter coefficient in C-band does not have the characteristics of bimodal distribution for different sea ice types. Therefore, a Bayesian classification method combining ASCAT scatterometer and SSMIS 37 GHz brightness temperature data was proposed for sea ice classification. In order to solve the problem that the FYI with high surface roughness in the marginal ice zone is easily misjudged as MYI in C-band scatterometer measurements, a Bayesian classification cost function and a marginal ice zone correction algorithm were proposed. Compared with the ice chart results, the MYI retrieved by ASCAT/SSMIS is consistent with a sea ice concentration of more than 50%.

Based on the ASCAT backscatter coefficient and AMSR-2 brightness temperature data [21], the polar sea ice research team of Bremen University produced and released daily MYI concentration (MYIC) products for winter in the years 2009–2018 by using the improved Environment Canada ice concentration extractor algorithm (ECICE) (https://seaice.uni-bremen.de/, accessed on 20 Febuary 2023). Furthermore, OSI SAF (Ocean and sea ice satellite application facility) adopted a Bayesian classification method to produce and release daily Arctic sea ice classification products since 2005 by using ASCAT and SSMIS/AMSR-2 data [22]. In addition to distinguishing FYI and MYI, the product also classes sea ice as ambiguous based on a low predicted classification accuracy. Its disadvantage is that the misclassification error easily accumulates, thus affecting the continuity and stability of the subsequent results.

Regarding sea ice classification using Ku-band scatterometers, Kwok [14] proposed a fixed threshold method to distinguish FYI from MYI by comparing the backscatter coefficient of QuikSCAT/SeaWinds with high-resolution RADARSAT image data. Swan and long (SL method for short) [15], obtained the interannual dynamic threshold by analyzing and fitting the statistical histogram of the daily backscatter coefficient of QuikSCAT/SeaWinds from 2002 to 2009, from which Arctic sea ice classification datasets from 2002 to 2009 were generated. The comparison with the ice chart provided by the Canadian ice service (CIS) shows that the average error of sea ice classification from 2006 to 2008 is less than 6%, which is an improvement on the fixed threshold classification result. The problem of the SL method is that the interannual sea ice type difference is weakened to some extent due to the fixed use of the interannual threshold. Lindell and Long [23] processed the backscatter coefficients of QuikSCAT/SeaWinds from 1999 to 2009 and OceanSAT-II/OSCAT from 2009 to 2014 and generated Arctic sea ice classification datasets for 15 years from 1999 to 2014 by adjusting the SL method in two aspects: the ice-water discrimination and dynamic threshold calculation. In addition, for the cases when the high roughness FYI in the marginal ice zone is misjudged as MYI, they proposed a method to remove the MYI noise using an image expansion algorithm. The evaluation results show that in the marginal ice zone with broken new ice and FYI (such as in the Barents Sea), the misclassification of FYI as MYI is significantly reduced, and for the marginal ice zone with more MYI (such as in the Greenland Sea), the correct MYI distribution information can be well retained.

Shokr and Agnew [16,17] proposed an ECICE algorithm that combines the QuikSCAT/SeaWinds and AMSR-E measurements. The algorithm divides the sea ice types into new ice, FYI, and MYI, providing the MYIC and a classification confidence level per pixel. The comparison results with the ice chart show that the ECICE algorithm can effectively improve the classification accuracy of water and MYI, but the MYI tends to be misclassified as FYI and vice versa when warm air advection events occur in autumn and early spring. Ye et al. improved the ECICE algorithm based on atmospheric temperature [18] and sea ice drift data [29], generating a daily Arctic MYIC dataset for winter in the years 2003–2009.

Zhang et al. [19] constructed a K-means sea ice classification algorithm based on the backscatter coefficient of the QuikSCAT/SeaWinds and ASCAT scatterometers, and the brightness temperature data of the AMSR-E, SSMI/S and AMSR-2 radiometers, where the resolution enhancement algorithm [30] was used to generate a daily Arctic sea ice classification dataset with a spatial resolution of 4.45 km for winter in the years 2002–2017. Compared with the interpretation results of high-resolution SAR images, the overall classification accuracy in Canadian Arctic Archipelago was better than 93%.

The China–France oceanography satellite (CFOSAT) was launched on 29 October 2018, carrying the first-ever spaceborne, Ku-band, rotating, fan-beam scatterometer (CSCAT). The CSCAT combines the characteristics of the fixed fan-beam and rotating pencil-beam scatterometers with various incidence and azimuth angle information. Zhang et al. [26] systematically compared differences between Ku-band scatterometers (including QuikSCAT/SeaWinds and CSCAT) and C-band scatterometers (including ASCAT) in sea ice classification for the first time. Their results show that the distinction between MYI and FYI is more obvious in Ku-band scatterometers in winter. Furthermore, compared with the ice chart and SAR image interpretation results, adding AMSR-2 microwave radiometer data can further improve the accuracy of sea ice classification compared with the results using scattering data alone. Xu et al. [28] constructed a tree augmented naive Bayesian sea ice classification algorithm based on CSCAT data and AMSR-2 data, generating a daily Arctic sea ice classification dataset for winter in the years 2019–2021. The comparison with the ice chart and the operational sea ice type products released by OSI SAF shows that it has higher classification accuracy for lower MYIC regions than that of OSI SAF.

It should be noted that in these two studies, the incidence angle normalization correction was used based on the orbital measurement for each day before feature extraction was used for sea ice classification. This procedure can be time-consuming, but the effect of the scatterometer incidence angle normalization on the improvement of sea ice classification accuracy should be further studied in depth. In this study, a CSCAT sea ice classification algorithm based on machine learning-aided classification methods is proposed. We innovatively extract effective but simpler feature parameters based on orbital rather than daily measurement to distinguish water and two different sea ice types: FYI and MYI. This paper is organized as follows. The datasets used in this article are described and preprocessed in Section 2. Section 3 introduces the sea ice type classification algorithm in detail. The sea ice classification results are presented and discussed in Sections 4 and 5, respectively. The conclusions are finally addressed in Section 6.

2. Datasets

2.1. CFOSAT Scatterometer (CSCAT)

CFOSAT was launched on 29 October 2018. It carries the first-ever spaceborne Ku-band rotating fan-beam scatterometer, that is, CSCAT [31,32]. The main goal of CSCAT is to monitor ocean surface wind for improving numerical weather forecasts, with the additional potential of monitoring sea ice parameters. CSCAT Level 2A data has two kinds of wind vector cell configurations in the 1000 km swath with a sampling resolution of 25 km and 12.5 km, respectively. In this study, CSCAT Level 2A 25 km global orbital datasets from 1 January 2019 to 14 April 2022 were used for Arctic sea ice classification (available at ftp://osdds-ftp.nsoas.org.cn, accessed on 20 Febuary 2023). The preprocessing of the orbital data and related feature parameters extraction is specifically described in Section 3.

2.2. National Snow and Ice Data Center Sea Ice Concentration and Age Product

Similar to the previous studies mentioned in [27], the random forest classifier is used to distinguish sea ice type because of its overall high accuracy and efficiency. As a kind of supervised learning classifier, training datasets with labeled or a priori information are necessary for training a random forest classification model. In this study, the sea ice concentration [33] and sea ice age (SIA) products [34,35] released from National Oceanic and Atmospheric Administration/National Snow and Ice Data Center (NOAA/NSIDC) are used for distinguishing water, FYI and MYI (available at https://nsidc.org/data, accessed on 20 Febuary 2023).

It should be noted that NOAA/NSIDC sea ice concentration is used as the training data in this paper because it has the same spatial resolution of 25 km as the CSCAT data, and the data has good continuity, which can ensure the stability of a priori information acquisition of the algorithm used in this paper. Furthermore, the SIA product is derived from passive and active microwave observations, as well as auxiliary data such as drifting buoys. It is weekly-updated data with a spatial resolution of 12.5 km. In our study, the nearest neighbor resampling method, which is suitable for representing the discrete data of classification [28], was used to adjust the NSIDC SIA product with a resolution of 12.5 km to a resolution of 25 km in order to match the CSCAT orbital data. Specifically, the pixel value nearest to a pixel position in the image is taken as the new value of the pixel. The advantages of this method are that it is simple and efficient, has a fast operation speed, and does not change the original image grid value. For comparison, grid cells with ice age older than one year are regarded as MYI.

2.3. OSI SAF Sea Ice Type and Drift Product

The sea ice type product provided by OSI SAF (available at https://osi-saf.eumetsat.int/products, accessed on 20 Febuary 2023) is used as the main validation data source in this study [36]. It is a multi-sensor daily-updated product derived from passive and active microwave remote sensing data combined in a Bayesian approach with a spatial resolution of 10 km. To differentiate between FYI and MYI in the northern hemisphere, the Bayesian algorithm computation is trained against data from regions with "known ice type" of FYI and MYI. In the latest version of the algorithm, the training data is continuously daily-updated from the preceding 15 days. Since the Bayesian approach describes the probability of occurrence of the most likely surface class, the probability itself can be an indicator of statistical uncertainty of the classification. Therefore, except for FYI and MYI, the OSI SAF sea ice type product also includes an ambiguous ice type classification for probabilities of less than 75% or in the summer period. Furthermore, the OSI SAF low-resolution ice drift product (OSI-405 series) is also used to analyze some case studies in Section 4.2.

2.4. IUP Multiyear Ice Concentration Product

The Arctic MYIC product provided by the University of Bremen, Institute of Environmental Physics (IUP) is used as another validation data source (available at https://seaice.uni-bremen.de, accessed on 20 Febuary 2023). The ECICE algorithm is firstly used to retrieve young ice, FYI, and multiyear ice concentration, respectively, where the microwave radiometer data of the sensors AMSR-E or AMSR2 and scatterometer data from ASCAT are used as inputs [16]. Then several correction schemes, including the so-called temperature correction [18] and drift correction [29], are used to correct misclassification by melt and refreezing and so forth, where surface temperature and sea ice motion or drift products are used as inputs. It is noted that under melting conditions, the surface properties of those three ice types become similar, resulting in the ECICE algorithm being available only from autumn to spring [21].

2.5. Synthetic-Aperture Radar (SAR)-Based Sea Ice Type Products

In order to validate CSCAT sea ice classification results using SAR products from a long-term perspective, the SAR-based sea ice type products provided by Copernicus Marine Environment Monitoring Service (CMEMS) from 1 January 2021 to 31 December 2021 are used in this study (https://resources.marine.copernicus.eu/product-detail/SEAICE_ARC_PHY_AUTO_L4_NRT_011_015/DATA-ACCESS, accessed on 20 Febuary 2023). This high-resolution operational sea ice type is derived from Sentinel-1 SAR data using a convolutional neural network (CNN) [37]. Validation shows that the overall accuracy exceeds 90% and has proved to be more efficient in computing time and less sensitive to noise in SAR data. In this algorithm, the prior information comes from reference ice charts produced by human experts. The SAR-based sea ice type provides pixel-by-pixel classification of SAR data into four different types, that is, water, young ice, FYI, and MYI. The young ice and FYI are grouped as FYI in this study.

3. Methodology

Figure 1 gives the flowchart of sea ice classification based on the machine learning approach with CSCAT measurement in this study. It can be seen that three main aspects need to be considered and assessed carefully.

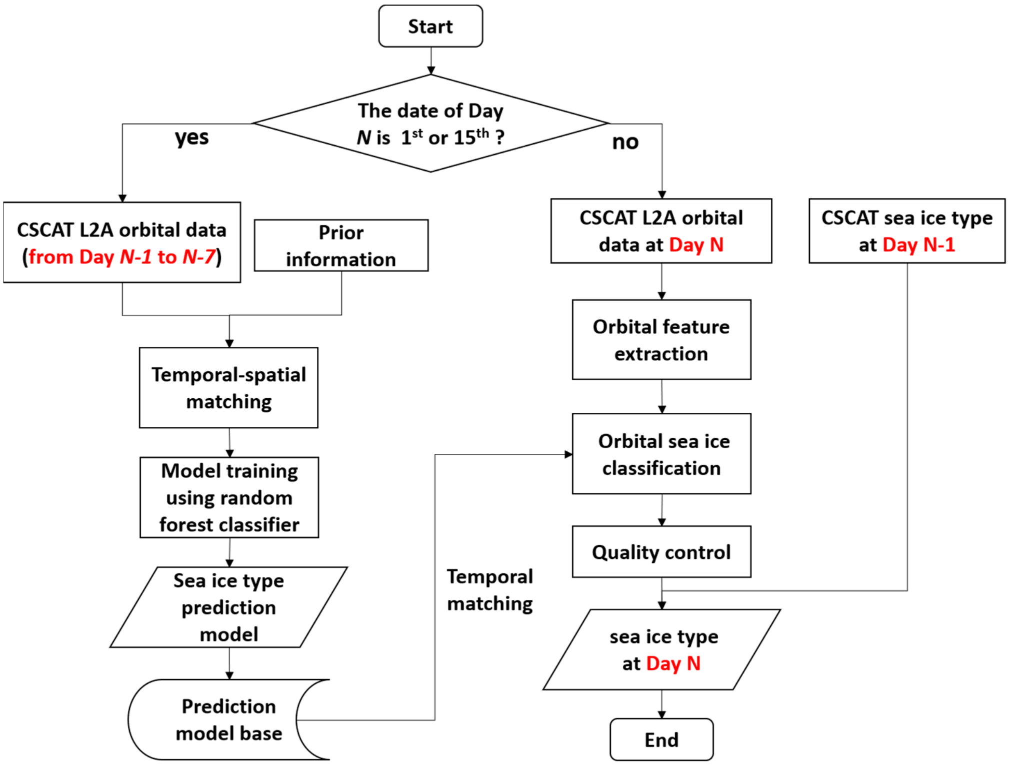

Figure 1. The flowchart of sea ice classification based on the machine learning approach with CSCAT measurement.

(1) The selection of machine-learning-aided sea ice classification method.
(2) The determination of the prior knowledge for sea ice type and the optimization of the prediction model, including its update frequency.
(3) The choice of feature parameters for sea ice classification based on the CSCAT orbital dataset.

As for the first aspect, five machine learning classifiers, including logistic regression, naïve Bayes, random forest, gradient boosting, and support vector machine (SVM), have been evaluated for sea ice distribution in a previous study [27]. It was concluded that the random forest classifier is the best option for sea ice monitoring because of its high overall accuracy. Following the sea ice monitoring study in [27], the random forest classifier is also used in this study [38]. It is a supervised learning algorithm that is used for both classification as well as regression. Figure 2 shows the structural schematic of the random forest classifier used in this study. It is a bagging algorithm with the decision tree as an estimator, the sample and feature parameters of which are randomly selected every time. A single decision tree is sensitive to the noise of the training dataset, but the random forest classifier can reduce the correlation between the trained multiple decision trees, effectively reducing the overfitting problem. In addition, the importance of variables can be assessed, and data scaling is not required in the random forest algorithm. The disadvantage of the random forest classifier is that it needs more computational resources, and the construction of the random forest is much harder and more time-consuming than the decision tree.

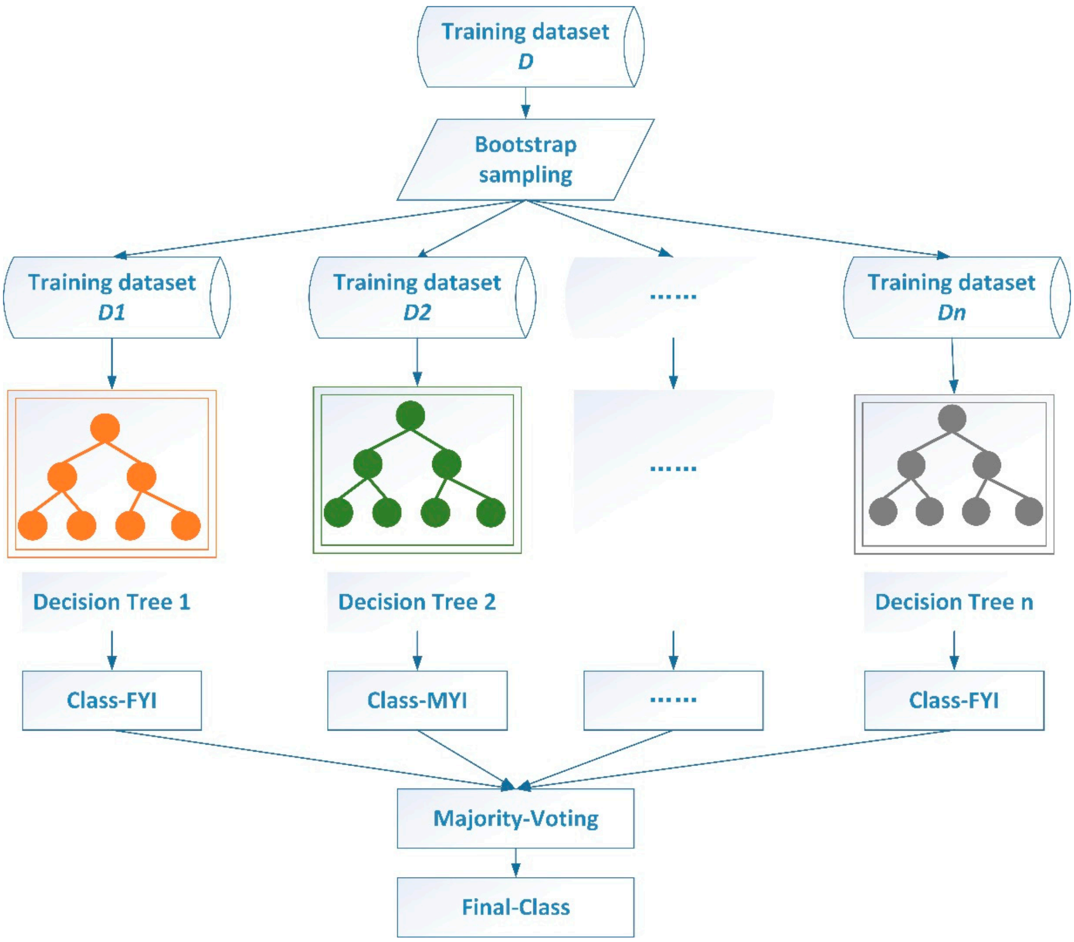

Figure 2. The structural schematic of random forest classifier used in this study.

As for the second aspect, to differentiate between first-year ice and MYI, different training datasets have been used in previous studies [28,36]. For instance, in the OSI SAF sea ice type procedure, specific regions are defined as MYI and FYI [36]. In this study, sea ice concentration and SIA products released from NSIDC are used for prior information definition. Specifically, water is defined as the area with a sea ice concentration of less than 40% and SIA less than 1 year, and FYI is defined as the pixel where the sea ice concentration is higher than 40% and the SIA product is less than 1 year. MYI corresponds to pixels with a sea ice concentration higher than 40% and SIA older than 1 year. The advantage of applying 40% as the threshold is that most of marginal ice zone area can be excluded, where the backscatter coefficient of first-year ice appears to be similar to that of multi-year ice [19,20]. Furthermore, the training model is updated on the 1st and 15th of each month, the assessment of which will be described later.

As for the third aspect, similar to [27], five feature parameters are defined based on the CSCAT orbital backscatter coefficient datasets, that is, the mean value of horizontal and vertical polarization backscatter ($\overline{\sigma}_{hh}$ and $\overline{\sigma}_{vv}$), the standard deviations of horizontal and vertical polarization measurements ($\Delta\sigma_{hh}$ and $\Delta\sigma_{vv}$), and the copol ratio ($\gamma = \overline{\sigma}_{vv}/\overline{\sigma}_{hh}$), the equations of which are listed as below:

$$\overline{\sigma}_{hh} = \frac{1}{N}\sum_{i=1}^{N}\sigma_{hh,i}(\theta_i,\phi_i), \tag{1}$$

$$\overline{\sigma}_{vv} = \frac{1}{N}\sum_{i=1}^{N}\sigma_{vv,i}(\theta_i,\phi_i), \tag{2}$$

$$\Delta\sigma_{hh} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\sigma_{hh,i}(\theta_i, \phi_i) - \overline{\sigma}_{hh}\right]^2}, \tag{3}$$

$$\Delta\sigma_{vv} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\sigma_{vv,i}(\theta_i, \phi_i) - \overline{\sigma}_{vv}\right]^2}, \tag{4}$$

$$\gamma = \overline{\sigma}_{vv} / \overline{\sigma}_{hh}, \tag{5}$$

where N is the number of observations at the grid cell after stereographic projection, $\sigma_{hh,i}(\theta_i, \varphi_i)$ and $\sigma_{vv,i}(\theta_i, \varphi_i)$ are the ith horizontal and vertical polarization backscatter, respectively, and θ_i and φ_i are corresponding incidence and azimuth angles, respectively.

Figure 3 shows the spatial distribution and probability density distribution of these five parameters for three surface classes (water, FYI, and MYI) based on collocated data from 1 April 2020 to 7 April 2020 in the whole Arctic region. It can be seen from Figure 3(a1,a2,b1,b2) that the difference between $\overline{\sigma}_{hh}$ and $\overline{\sigma}_{vv}$ for MYI and FYI is quite significant, where the MYI area is generally higher than the FYI and water area, proving the feasibility of CSCAT orbital feature analysis for sea ice classification. Furthermore, as shown in Figure 3(c1,c2,d1,d2,e1,e2), all of these three parameters, including $\Delta\sigma_{vv}$, $\Delta\sigma_{hh}$ and γ, can well distinguish water and ice, but cannot classify sea ice type very well. It should be noted that as for the CSCAT orbital measurements, the diversity of incidence angles and azimuth angles of measurements is much lower in the outer swath than in other swath regions, and the values of $\Delta\sigma_{vv}$ and $\Delta\sigma_{hh}$ are much lower in the outer swath area. In conclusion, the combination of these five parameters is expected to effectively distinguish water and different sea ice type.

Figure 4 shows the statistical results of the sea ice type feature parameters based on the random forest classification during the period from 1 January 2019 to 31 December 2020. The red, blue, and black lines represent the mean values of the Gaussian distribution of feature parameters derived from the water, FYI, and MYI, respectively. The corresponding shaded part is the standard deviation of the Gaussian distribution. It can be seen that the feature parameters have obvious differences except for June, July, August, and September of each year. The corresponding date of the model update is represented by red solid dots shown in Figure 4a, that is, the 1st and 15th of each month. Specifically, the model on the 1st day of each month is used for sea ice classification from the 2nd to 15th day of the current month, and the model on the 15th day of each month is used for sea ice classification from the 16th day of the current month to the 1st day of the next month. This model updating frequency setting can reasonably represent the difference in feature parameters. Additionally, it can effectively reduce the impact of prior information fluctuations on the results and ensure the stability of the results.

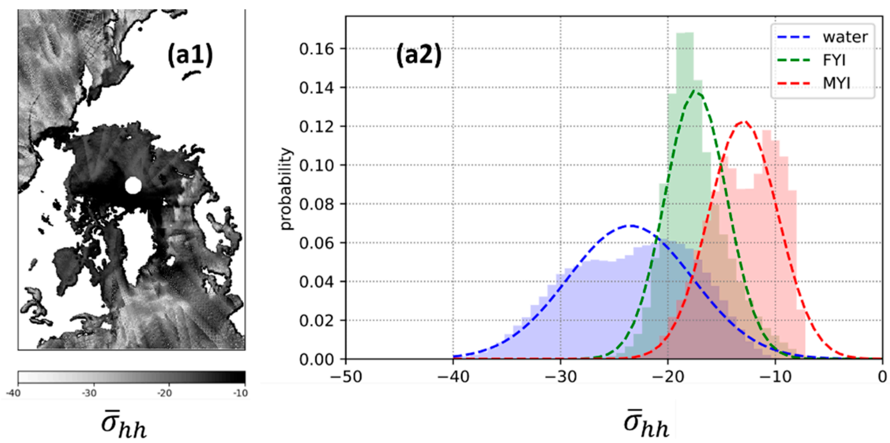

Figure 3. *Cont.*

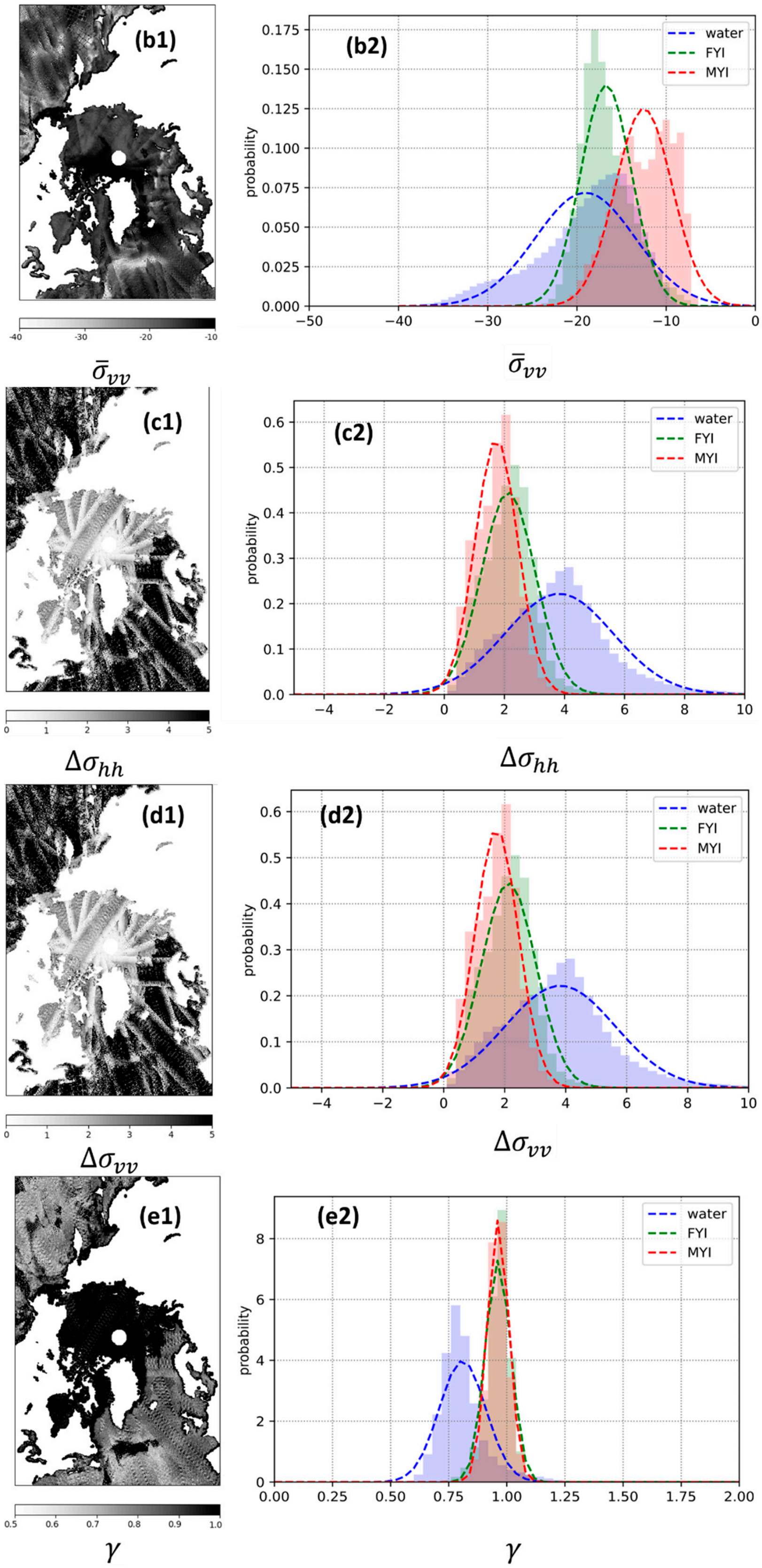

Figure 3. Parameters derived from one day of CSCAT measurements in the Arctic on April 2020. (**a1**) $\overline{\sigma}_{hh}$, (**b1**) $\overline{\sigma}_{vv}$, (**c1**) $\Delta\sigma_{hh}$, (**d1**) $\Delta\sigma_{vv}$ and (**e1**) γ. Arctic images contain 448 $\times$ 304 pixels with a pixel resolution of 25 km. The central white circular area represents no observations. Density plot for parameters (**a2**) $\overline{\sigma}_{hh}$, (**b2**) $\overline{\sigma}_{vv}$, (**c2**) $\Delta\sigma_{hh}$, (**d2**) $\Delta\sigma_{vv}$, and (**e2**) γ for water (blue), FYI (green), and MYI (red).

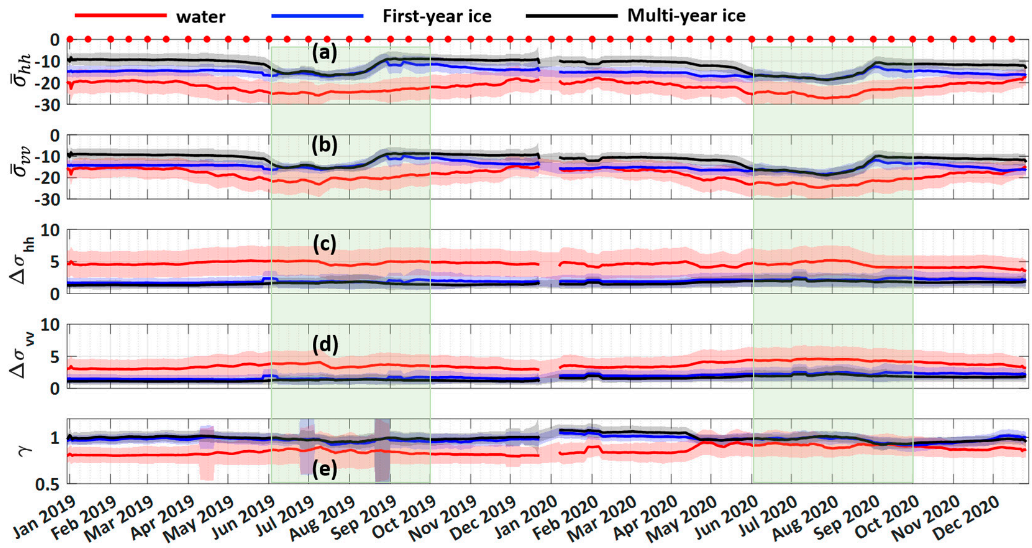

Figure 4. The mean values (solid lines) and standard deviations (shaded areas) of Arctic sea ice distribution feature parameters from 1 January 2019 through 31 December 2019: (**a**) $\overline{\sigma}_{hh}$, (**b**) $\overline{\sigma}_{vv}$, (**c**) $\Delta\sigma_{hh}$, (**d**) $\Delta\sigma_{vv}$ and (**e**) γ. The red, blue, and black marked results represent water, FYI, and MYI, respectively. The red solid points correspond to the updating time of the prediction model.

4. Results

4.1. Evaluation of Sea Ice Classification Model Precision

The sea ice type model based on the random forest classifier is quantitatively assessed using a confusion matrix through a comparison with the NSIDC SIA product as reference data. Table 2 defines the confusion matrix in this study, each column and row of which represents a specific prediction and real category, respectively.

Table 2. Definition of sea ice type training model confusion matrix.

		Prediction		
		Water	**FYI**	**MYI**
True	**Water**	a	b	c
	FYI	d	e	f
	MYI	g	h	i

The accuracy (*accuracy*) is defined, as shown in Equation (6), to evaluate the model's overall performance. However, the number of samples between each category is often not balanced for practical classification problems. If the imbalanced dataset is not adjusted, the model tends to favor large categories and discard small categories. In this situation, the accuracy can be high, but some categories are not recalled at all. Therefore, the Kappa coefficient (*kappa*) is defined, as shown in Equation (7), to take sample number difference into consideration. It can be seen that the more unbalanced the confusion matrix is, the higher the PE (p_e), and the lower the kappa value, which results in a low score for the model with a strong bias.

$$accuracy = \frac{a+e+i}{a+b+c+d+e+f+g+h+i}, \tag{6}$$

$$kappa = \frac{accuracy - p_e}{1 - p_e}, \tag{7}$$

$$p_e = \frac{(a+d+g)\times(a+b+c)+(b+e+h)\times(d+e+f)+(c+f+i)\times(g+h+i)}{(a+b+c+d+e+f+g+h+i)^2}, \tag{8}$$

As for the specific category, precision and recall are used to evaluate each category's performance. Taking the type of water as example, TP_{water} is defined as the number of actual water and predicted water values. FP_{water} is defined as the number of actual sea ice values (FYI or MYI) predicted as water, and FN_{water} is defined as the number of actual water values predicted as sea ice. Therefore, the precision and recall of water are defined as:

$$precision_{water} = \frac{TP_{water}}{TP_{water}+FP_{water}} = \frac{a}{a+d+g}, \tag{9}$$

$$recall_{water} = \frac{TP_{water}}{TP_{water}+FN_{water}} = \frac{a}{a+b+c}, \tag{10}$$

where $TP_{water} = a$, $FP_{water} = d + g$, and $FN_{water} = b + c$, respectively.

Figure 5 shows the time series of the parameters described above to evaluate the sea ice type training model in the Arctic from 1 January 2019 through 1 March 2020. The precisions of water, FYI, and MYI are 99.67%, 86.60%, and 79.74%, respectively. The recalls of water, FYI, and MYI are 99.77%, 84.78%, and 80.59%, respectively. Both precision and recall illustrate that the classification model has the highest prediction accuracy for water, followed by FYI and MYI. The average accuracy and the kappa coefficient are 93.35% and 88.53%, respectively. The reason why accuracy is much higher than the kappa coefficient is that the sample number of water is much higher than that of sea ice. The overall accuracy is thus more easily affected by water precision.

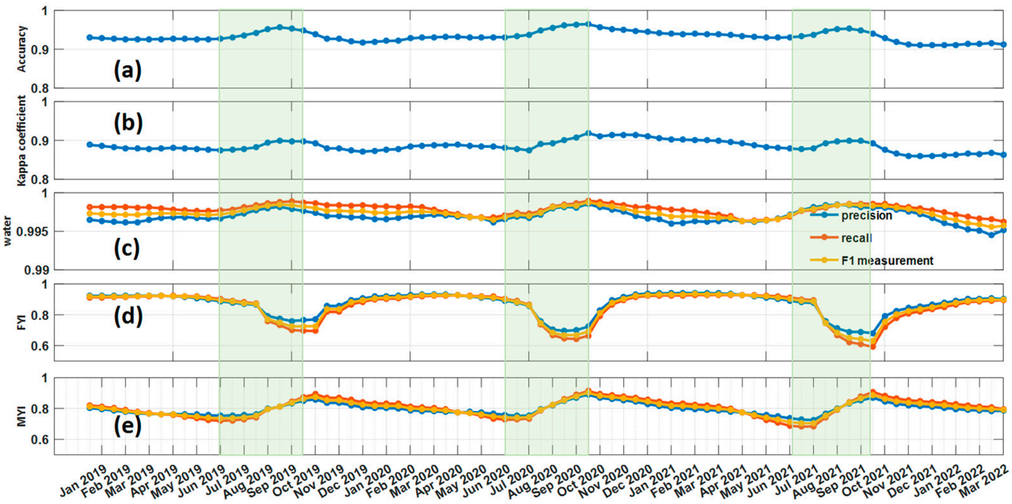

Figure 5. The time series of the evaluation parameters of the sea ice type training model in the Arctic from 1 January 2019 through 1 March 2020 for (**a**) accuracy, (**b**) kappa coefficient, (**c**) water, (**d**) FYI, and (**e**) MYI.

In addition, the identification accuracy of FYI, as shown in Figure 5d, has obvious seasonal characteristics, with a significant decrease from June to September of each year. Combined with the results in Figure 4, it can be seen that the feature parameters of FYI and MYI are almost the same and cannot be distinguished very well from June to September of each year, resulting in a higher tendency to be misclassified. Correspondingly, the FYI is easily misclassified as MYI at the end of the summer and during the whole autumn season, as shown in Figure 6b. On the contrary, as shown in Figure 5e, the identification accuracy of MYI from June to September of each year is higher than that in other months. This is likely related to the prior information we use in this study. Generally, the NSIDC SIA product is more conservative in the identification of MYI, which means that more pixels are classified as MYI, resulting in the MYI model error characteristics shown in Figure 6c.

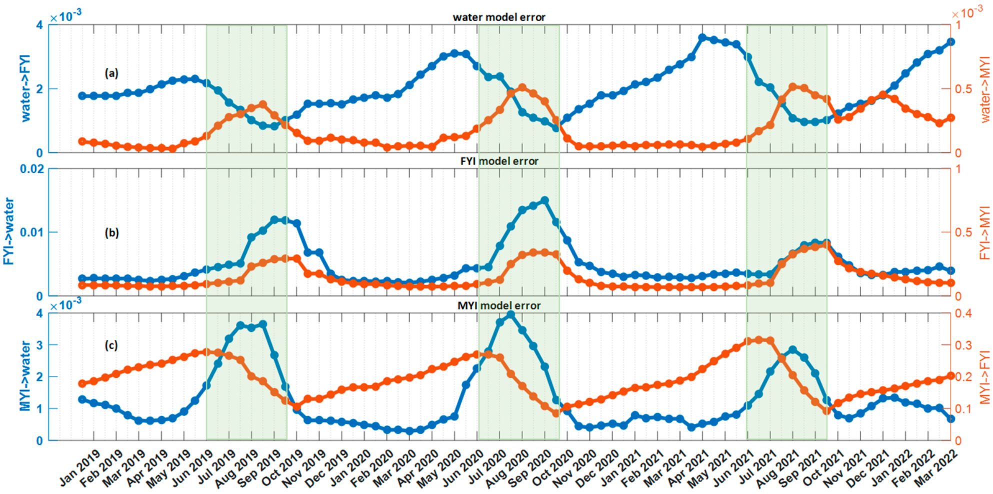

Figure 6. The error analysis of the Arctic sea ice type training models from 1 January 2019 through 1 March 2020: (**a**) water is misclassified as FYI (blue dotted line) and MYI (red dotted line), (**b**) FYI is misclassified as water (blue dotted line) and MYI (red dotted line), and (**c**) MYI is misclassified as water (blue dotted line) and FYI (red dotted line).

4.2. Comparison of Spatial Distribution of Sea Ice Type with Validation Products

In order to analyze the spatial characteristics of the Arctic sea ice in more detail, eight subregions that are consistent with previous studies are defined, as shown in Figure 7 [26,39]. The eight subregions include the central Arctic (CA), Chukchi Sea/Beaufort Sea (CBS), Laptev Sea/East Siberian Sea (LESS), Kara Sea/Barents Sea (KBS), East Greenland (EG), Hudson Bay/Baffin Bay (HBS), Canadian Arctic Archipelago (CAA), and Bering Sea (BS).

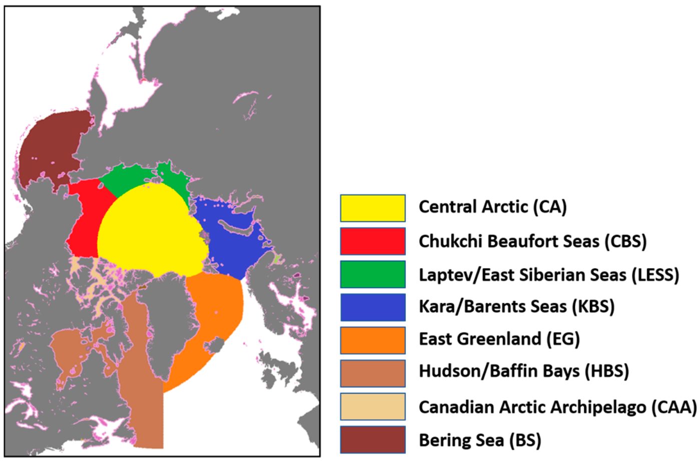

Figure 7. Map of the eight subregions of the Arctic used in this study.

Figure 8 shows the spatial distribution of sea ice type on selected dates from 2019 to 2021. LESS, KBS, HBB, and BS are mainly FYI throughout the year. The MYI of CA, EG, and CAA has obvious seasonal changes. The MYI of CA is mainly reduced through several routes, the most obvious of which is that the transpolar drift stream (TDS) causes MYI to flow into EG. On the two days of 15 May 2019 and 15 May 2020, the presence of MYI was seen in the Chuikui sea and the East Siberian Sea area. According to the analysis of the results of the previous and next days, and a comparison with other products, such as MYIC products, it was found that MYI from other regions had drifted into the region, and therefore, this was not a misclassification. The specific analysis is discussed later.

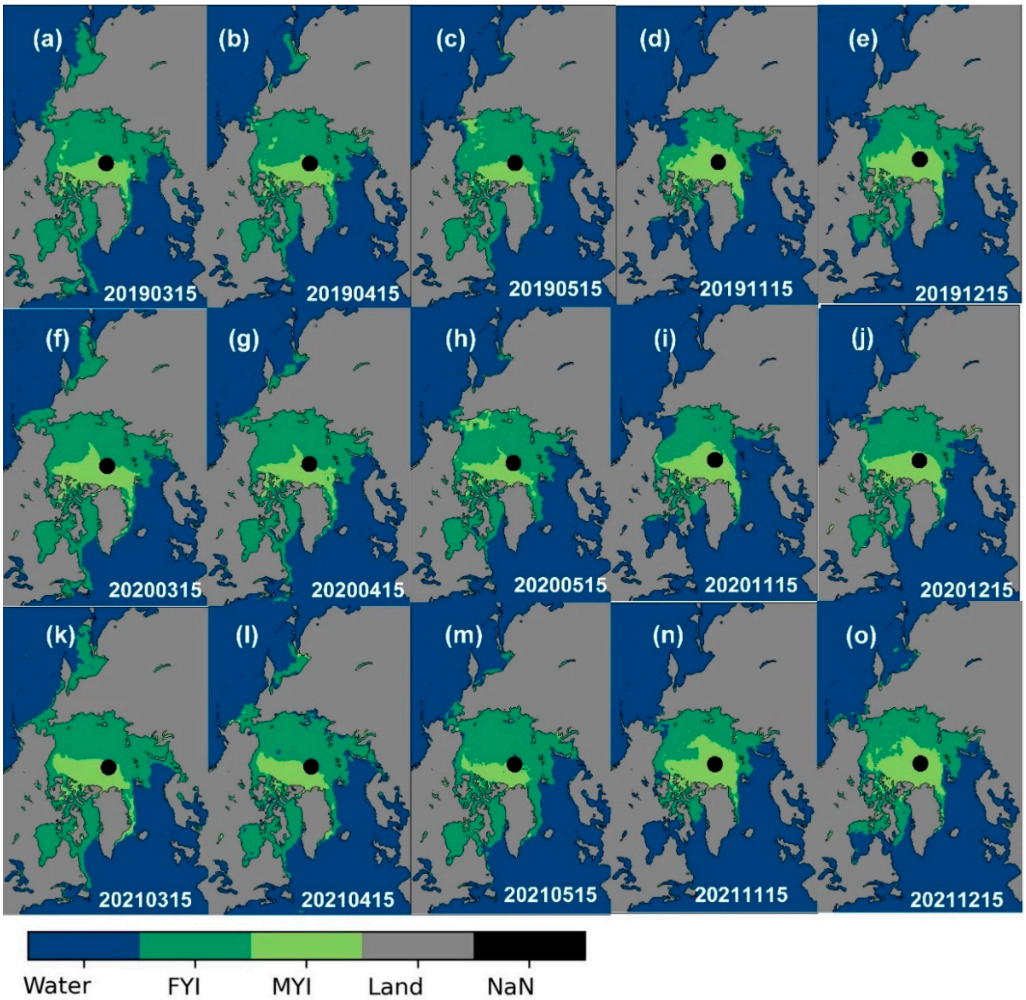

Figure 8. Spatial distribution of sea ice types on (**a**) 15 March 2019, (**b**) 15 April 2019, (**c**) 15 May 2019, (**d**) 15 November 2019, (**e**) 15 December 2019, (**f**) 15 March 2020, (**g**) 15 April 2020, (**h**) 15 May 2020, (**i**) 15 November 2020, (**j**) 15 December 2020, (**k**) 15 March 2021, (**l**) 15 April 2021, (**m**) 15 May 2021, (**n**) 15 November 2021, and (**o**) 15 December 2021.

The spatial distribution differences between the CSCAT results and the OSI SAF sea ice type products on selected dates from 2019 to 2021 are shown in Figure 9. It can be seen that the most obvious differences in the distribution of sea ice types between these two results are mainly concentrated in the marginal zones of FYI and MYI, where some areas identified as MYI by the CSCAT classification were classified as FYI by the OSI SAF sea ice type product. A similar phenomenon can be seen in previous studies [26,28], and this may be caused by the different responses of microwave signals of different frequencies to the mixed pixel detection of FYI and MYI.

As shown in Figure 9d,e, a speckled line crossing over the north pole from Frans Josef Land toward the Beaufort Sea can be seen in the OSI SAF sea ice type results on both 15 November 2019 and 15 December 2019. This misclassification is due to the insensitivity of ASCAT backscatter on younger multiyear ice. Related upgrades for OSI SAF sea ice type (Product-d) have dealt with strong disagreement in classifications from multi-sensor measurement [40]. In contrast, the CSCAT results in this study have no similar problem, and the area of MYI is more homogeneous with less noise, especially in the case of younger multiyear ice, such as second-year ice.

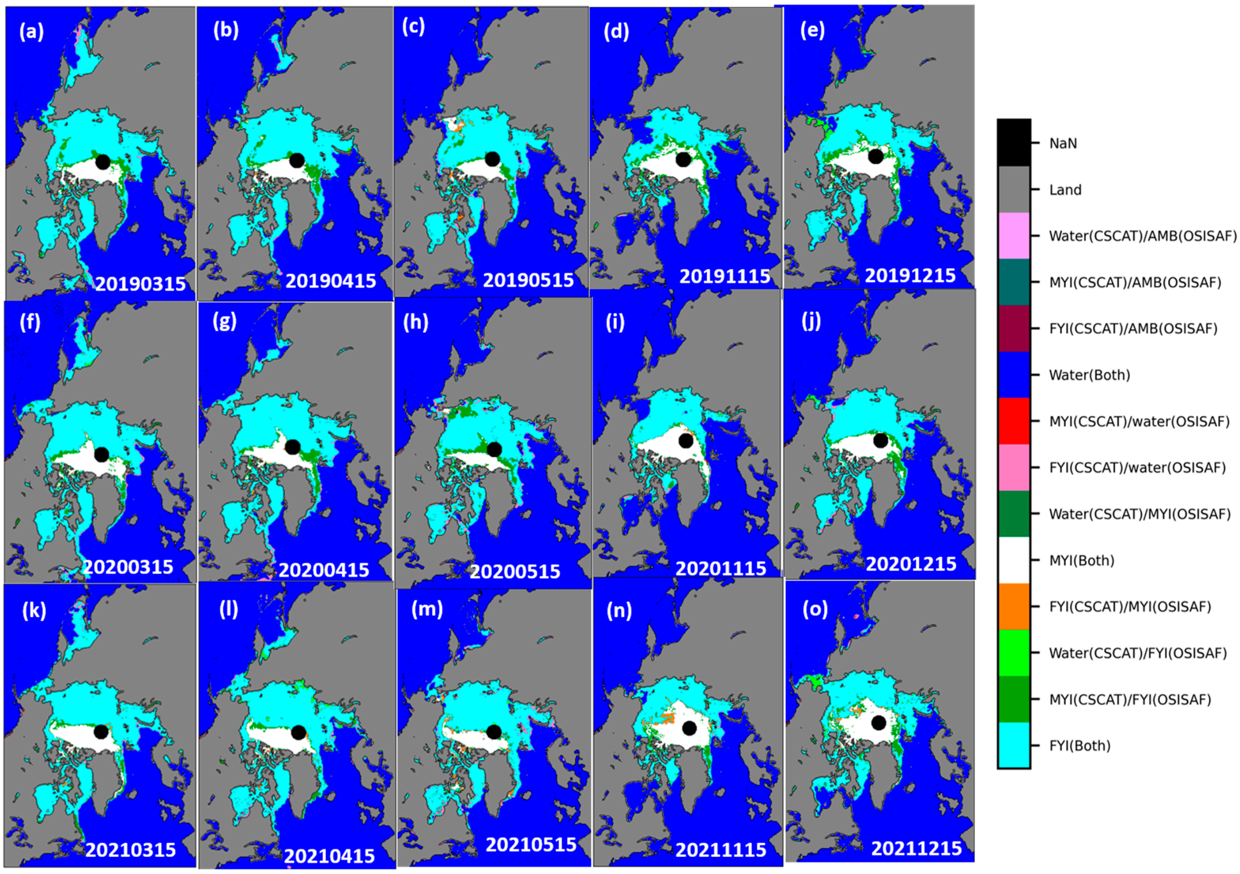

Figure 9. Distribution differences between the CSCAT results and the OSI SAF sea ice type products on (**a**) 15 March 2019, (**b**) 15 April 2019, (**c**) 15 May 2019, (**d**) 15 November 2019, (**e**) December 15, 2019, (**f**) 15 March 2020, (**g**) 15 April 2020, (**h**) 15 May 2020, (**i**) 15 November 2020, (**j**) 15 December 2020, (**k**) 15 March 2021, (**l**) 15 April 2021, (**m**) 15 May 2021, (**n**) 15 November 2021, and (**o**) 15 December 2021.

In addition, cases that CSCAT classifies as MYI but OSI SAF classifies as FYI often occur all year round in the EG region. To further analyze this situation, the results of SIA, MYIC, CSCAT, and OSI SAF in this region are analyzed. Figure 10(a1–a3) provide three weekly SIA spatial distributions, that is, from 12 November 2019 to 18 November 2019, from 12 February 2022 to 18 February 2020, and from 15 April 2020 to 21 April 2020, respectively. Figure 10(b1–b3) show the results of the IUP MYIC spatial distribution on the three days of 15 November 2019, 1 February 2020, and 15 April 2020, respectively. Corresponding to these three days, Figure 10(c1–c3) show the spatial distributions of IUP MYIC marked by CSCAT MYI pixels, and Figure 10(d1–d3) show the spatial distributions of the IUP MYIC marked by OSI SAF MYI pixels. It can be seen that in the EG region, CSCAT identifies more pixels as MYI with lower MYIC values, whereas FYI and MYI are identified by OSI SAF as sea ice type and SIA, respectively. This indicates that CSCAT tends to identify the pixels with lower MYIC as MYI compared with OSI SAF sea ice types, showing better accuracy in the identification of areas with obvious mobility of MYI, such as in the EG.

It is shown in Figure 9h that the disagreement occurs in the Chuikui and East Siberian Seas, where the MYI area identified by CSCAT was classified as FYI by the OSI SAF sea ice types. Results for the OSI SAF sea ice drift and type products, IUP MYIC, and CSCAT, on 15 May 2020 were selected for further analysis. Figure 11a shows the OSI SAF LR drift product from 13 May 2020 to 15 May 2020, and Figure 11b–d show the daily MYIC, CSCAT result and OSI SAF sea ice type distribution on 15 May 2020, respectively. It can be seen from Figure 11a,b that the MYI of the Beaufort Sea drifted towards the Chukchi Sea and East Siberian Sea areas at a relatively high speed during this period, resulting in the relatively dispersed spatial distribution of the MYI. This phenomenon is reflected in the results of the

CSCAT and OSI SAF sea ice types, but the area identified as MYI by CSCAT is obviously higher than that of the OSI SAF sea ice types, which is more consistent with the distribution of MYIC.

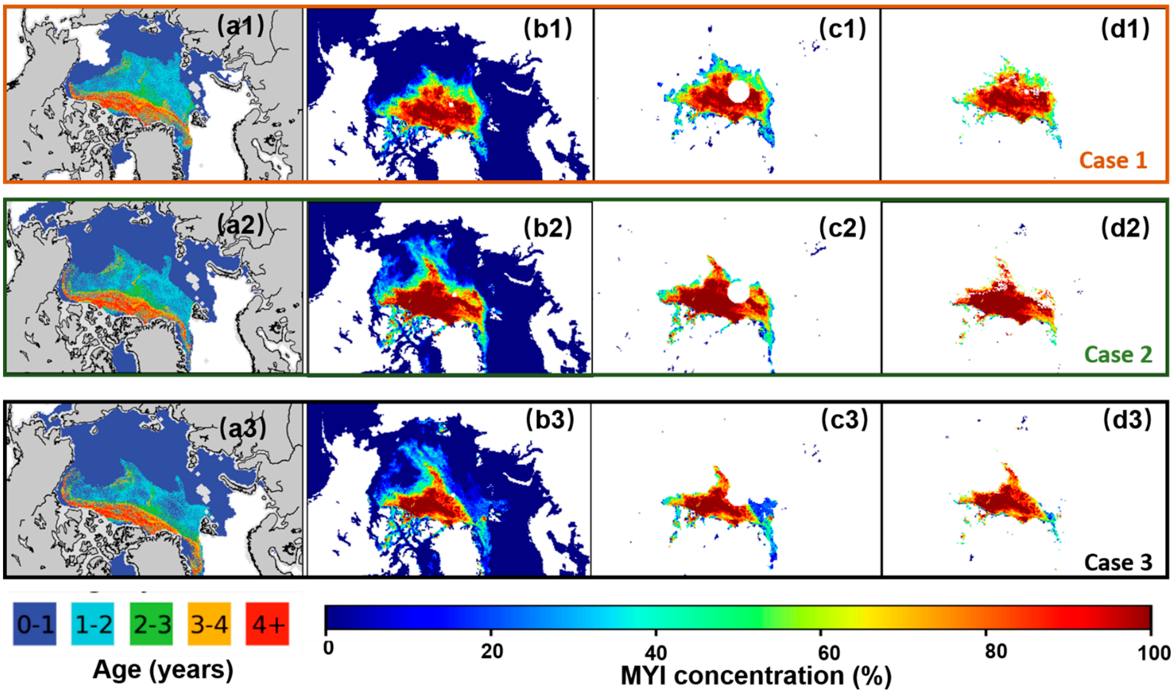

Figure 10. The spatial distribution of the NSIDC weekly SIA product (**a1**) from 12 November 2019 to 18 November 2019, (**a2**) from 12 February 2020 to 18 February 2020, and (**a3**) from 15 April 2020 to 21 April 2020. The spatial distribution of IUP daily MYIC on (**b1**) 15 November 2019, (**b2**) 14 February 2020, and (**b3**) 15 April 2020. The corresponding spatial distribution of IUP MYIC marked as CSCAT MYI pixels and OSI SAF MYI pixels on three selected days are shown in (**c1**–**c3**), and (**d1**–**d3**), respectively.

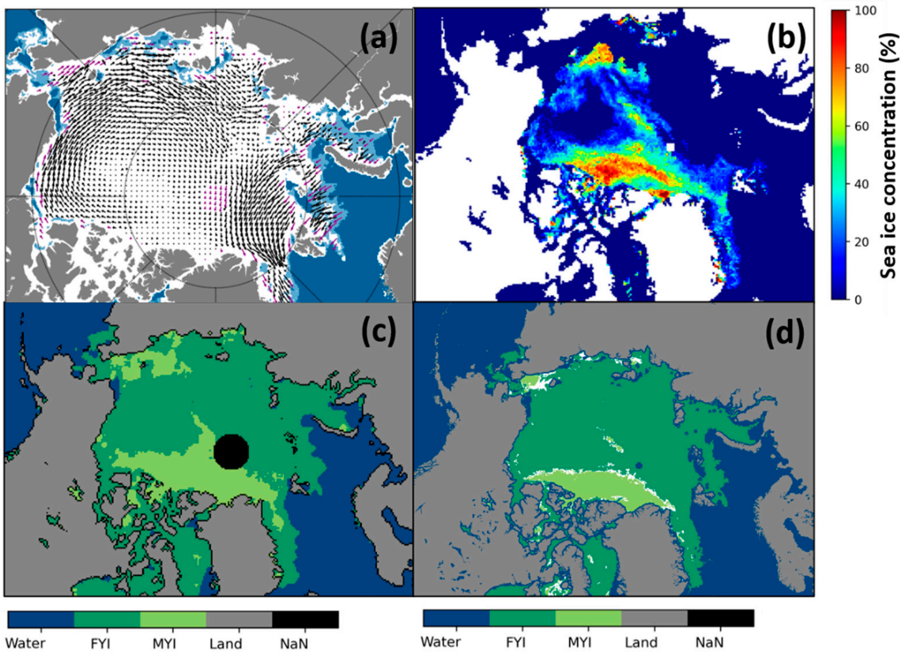

Figure 11. The spatial distribution of (**a**) LR sea ice drift from 13 May 2020 to 15 May 2020, (**b**) IUP daily MYIC on 15 May 2020, (**c**) CSCAT sea ice type on 15 May 2020, and (**d**) OSI SAF sea ice type on 15 May 2020.

5. Discussion

5.1. Comparison of Arctic FYI and MYI Extent with Validation Products

Figure 12a,b show the time series of FYI and MYI extent derived from the CSCAT and OSI SAF sea ice type products, respectively, from 1 January 2019 to 14 April 2022. It should be noted that the time period marked with a green-shaded area is classed by OSI SAF sea ice type as ambiguous ice, for which no analysis is presented in Section 5.1. It can be seen from Figure 12a that the FYI extent of CSCAT has an excellent consistency with that of OSI SAF sea ice type, the root-mean-square-error (RMSE) and correlation coefficient of which are 6.87×10^5 km^2 and 0.989, respectively. Table 3 lists the statistics of the maximum and minimum FYI extent and corresponding dates in each year during the analyzed period derived from CSCAT and OSI SAF sea ice type products, respectively. It can be seen that the maximum values of FYI from CSCAT and OSI SAF occur in March every year, and the minimum values occur from the end of September to the beginning of October every year.

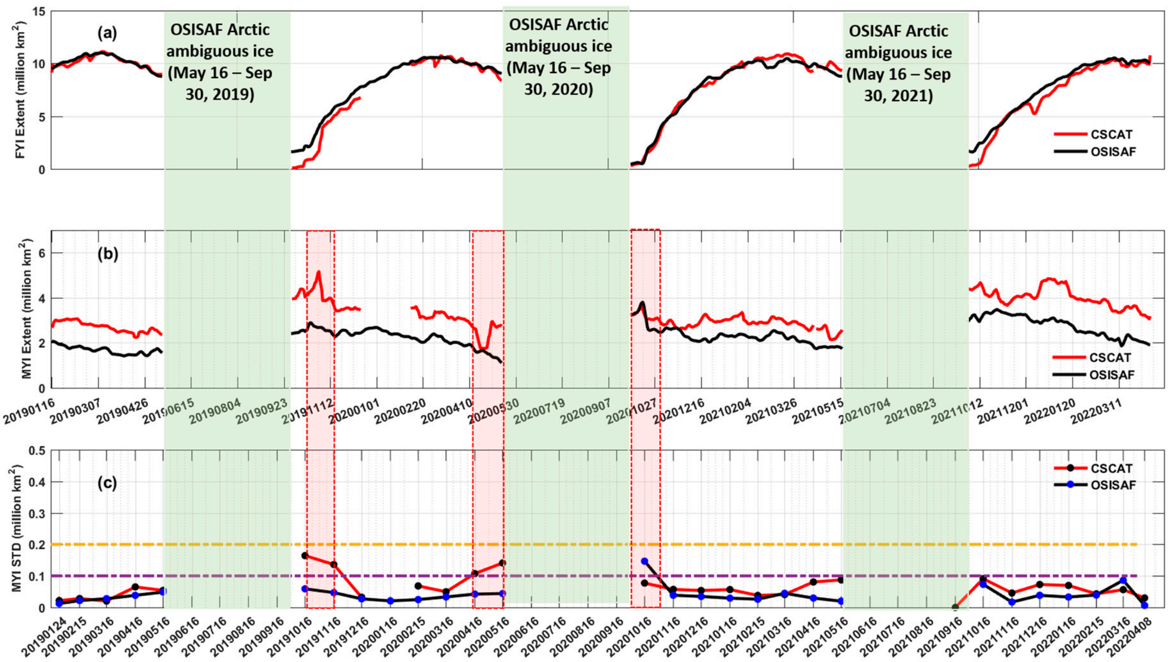

Figure 12. Time series of (**a**) FYI extent and (**b**) MYI extent derived from CSCAT (red line) and OSI SAF (black line), respectively. MYI monthly standard deviations (STD) derived from CSCAT (red dotted line) and OSI SAF (black dotted line) are shown in (**c**).

Table 3. Statistics of the maximum and minimum of FYI extent and corresponding dates in each year during the analyzed period derived from CSCAT and OSI SAF sea ice type product.

	FYI Extent Maximum		FYI Extent Minimum	
	Date	**Value (Million km^2)**	**Date**	**Value (Million km^2)**
CSCAT	11 March 2019	11.6837	2 October 2019	1.6368
	18 March 2020	11.2110	3 October 2020	0.4953
	19 March 2021	11.515	4 October 2021	1.6061
	5 March 2022	11.1344		
OSI SAF	11 March 2019	11.001	4 October 2019	0.1243
	3 March 2020	10.589	1 October 2020	0.3437
	19 March 2021	10.462	30 September 2021	0.2067
	6 March 2022	10.513		

Furthermore, it can be seen from Figure 12a that the CSCAT FYI extent in October 2019 and October 2021 is significantly less than that of the OSI SAF sea ice type product, and corresponding to Figure 12b, the CSCAT MYI extent is much higher than that of OSI SAF MYI in this time period. A comparison with the weekly SIA product is made to further analyze this phenomenon, as shown in Figure 13. Specifically, Figure 13(a1–a3) show the spatial distribution difference between the CSCAT results and the OSI SAF sea ice type products on the first day of the end of OSI SAF ambiguous ice in 2019–2021, that is, 4 October 2019, 1 October 2020, and 1 October 2021, respectively. The weekly SIA corresponding to the three dates are shown in Figure 13(b1–b3), respectively. It can be seen from Figure 13(a1,a3) that on the two days of 4 October 2019 and 1 October 2021, in the CA area, the area determined as FYI by OSI SAF was identified as MYI by CSCAT. During the two one-week periods from 1 October 2019 to 7 October 2019, and from 1 October 2021 to 7 October 2021, as shown in Figure 13(b1,b3), almost all the sea ice was identified as MYI by the SIA product. This indicates that the OSI SAF sea ice type product underestimates the extent of MYI in these cases, and compared with the OSI SAF sea ice type product, CSCAT can better describe the distribution of sea ice type at the initial stage of the freezing-up period.

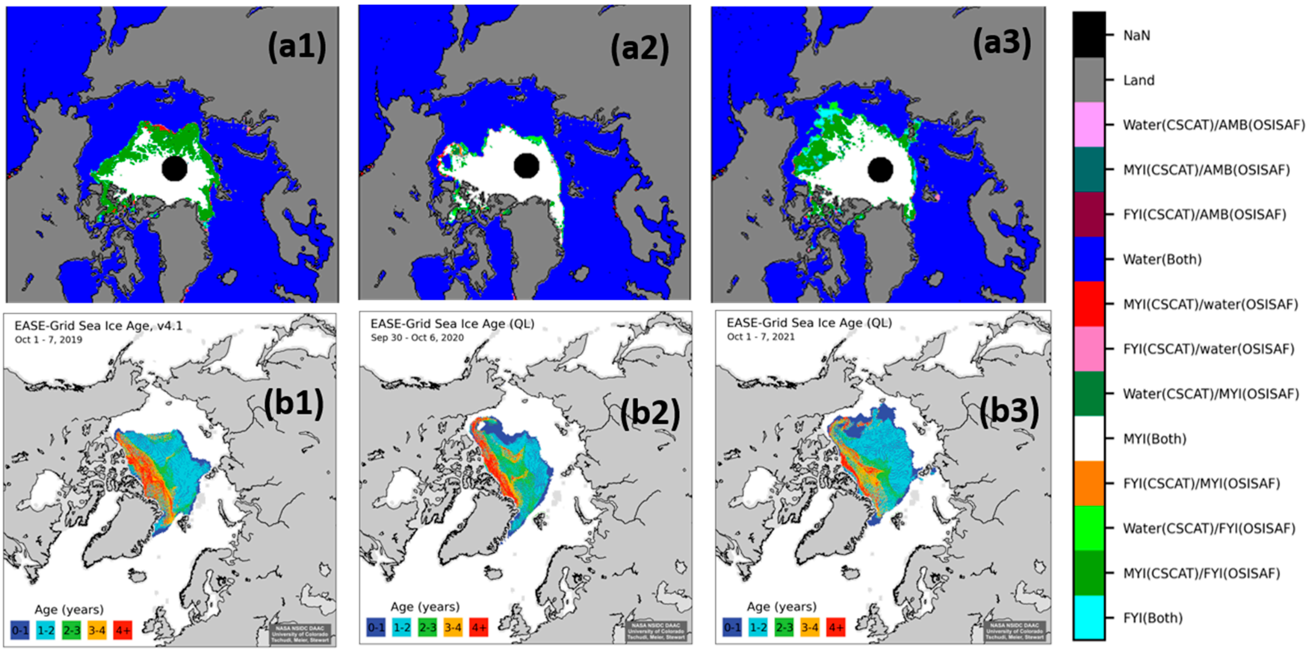

Figure 13. Spatial distribution differences between the CSCAT results and the OSI SAF sea ice type products on (**a1**) 4 October 2019, (**a2**) 1 October 2020, and (**a3**) 1 October 2021. The spatial distribution of the NSIDC weekly SIA product (**b1**) from 1 October 2019 to 7 October 2019, (**b2**) from 30 September 2020 to 6 October 2020, and (**b3**) from 1 October 2021 to 7 October 2021.

In regards to MYI extent, it can be seen from Figure 12b that the pattern of MYI extent of CSCAT has a good consistency with that of the OSI SAF sea ice types, the RMSE and correlation coefficient of which are 6.84×10^5 km^2 and 0.792, respectively. Basically, the CSCAT MYI extent is higher than that of OSI SAF sea ice types as a whole. It is known from the validation report of OSI SAF sea ice types that the OSI SAF sea ice type product (OSI-403-c) used for comparison in this study does not use the latest updated algorithm [40], resulting in an underestimation of MYI extent to some degree. The most significant difference between the latest algorithm and the current algorithm of OSI SAF sea ice type is the prioritization of the passive microwave radiometer (PMW) classification in cases where the PMW and ASCAT probabilities disagree completely in the final classification. Since the ASCAT mean backscatter strongly captures the older part of the pack ice, and the PMW brightness temperature channel combinations also capture the younger MYI, such as second-year ice, this is better mapped by the PMW brightness temperature than the pure ASCAT backscatter signal. Therefore, the updated algorithm can capture more of the younger MYI, so the difference between OSI SAF sea ice type using the latest algorithm and CSCAT should be smaller. In addition, it can be seen from the analysis in Figure 10 that CSCAT marks the lower MYIC pixels as MYI pixels, which is one of the main reasons why CSCAT MYI extent is higher than that of OSI SAF.

Similar to the analysis in OSI SAF sea ice type validation report [40], the MYI is assumed not to have rapid changes in horizontal coverage, and therefore, the monitoring of the temporal variation of the MYI extent can be used to assess sea ice type precision. Figure 12c gives the monthly standard deviation (STD) of the daily differences from a running mean of CSCAT and OSI SAF MYI extent, respectively. Overall, the MYI monthly STD of the daily differences calculated from CSCAT and the OSI SAF sea ice type product is less than 1×10^5 km^2 in most of the analyzed period, indicating that the sea ice type classification accuracy is reliable as a whole. However, in October 2019 and April 2020, the deviation of CSCAT is relatively large, and in October 2020, the deviation of OSI SAF is relatively large. The results of these three periods are analyzed in detail.

In October 2019, CSCAT MYI extent showed obvious fluctuation, reaching the highest value on 31 October 2019. According to the analysis of the spatial distribution of sea ice type on 25 October 2019, 31 October 2019 and 6 November 2019, as shown in Figure 14, sea ice in the LESS area is misclassified as MYI by CSCAT, leading to the overestimation of MYI. This also corresponds to the results in Figures 5 and 6, that is, on 15 October 2019, the prediction model accuracy and recall of FYI are only 0.73 and 0.65, respectively, and the probability of FYI misclassification as MYI is 0.26.

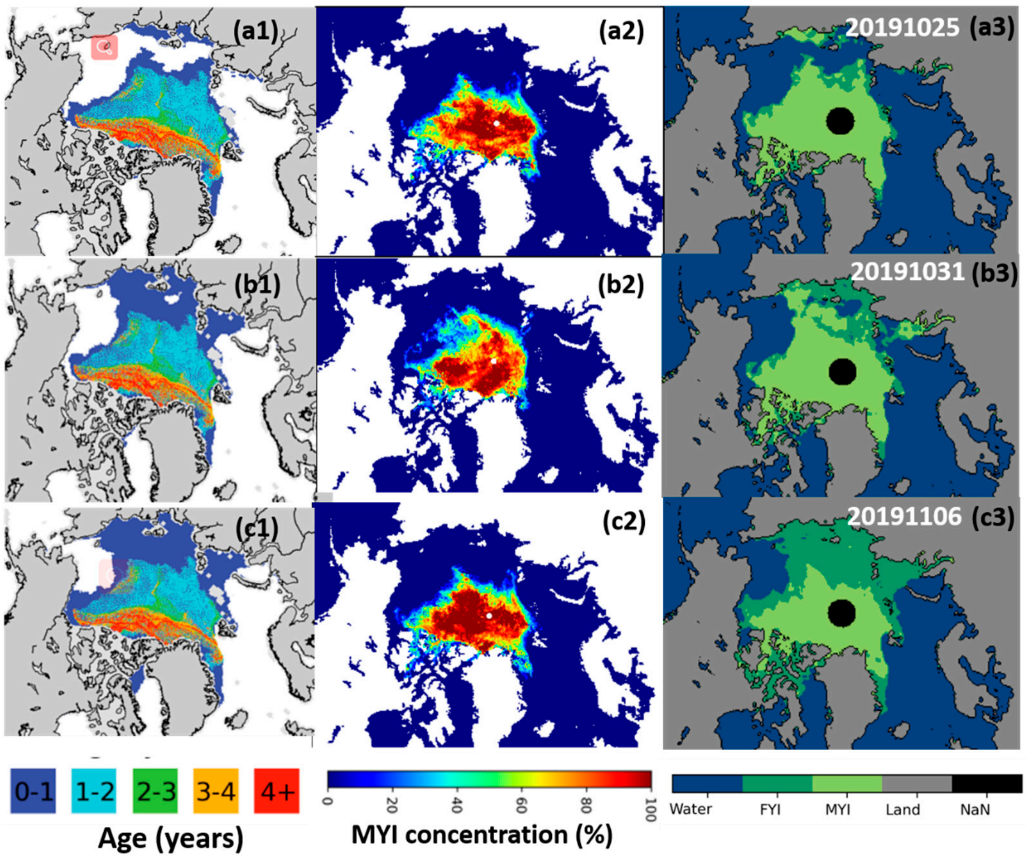

Figure 14. The spatial distributions of NSIDC weekly SIA product (**a1**) from 22 October 2019 to 28 October 2019, (**a2**) from 29 October 2019 to 4 November 2019, and (**a3**) from 5 November 2019 to 11 November 2019. The spatial distribution of IUP daily MYIC on (**b1**) 25 October 2019, (**b2**) 31 October 2019, and (**b3**) 6 November 2019. The corresponding spatial distribution of CSCAT sea ice type on (**c1**) 25 October 2019, (**c2**) 31 October 2019, and (**c3**) 6 November 2019.

Warm air intrusion (WAI) occurred between 17 April 2020 and 22 April 2020, resulting in the reduction of the observed MYI area [40]. On the one hand, this was due to the end of the WAI, on the other hand, as shown in Figure 11, the sea ice drift increased the MYI area, thus recovering the MYI area after the WAI.

In addition, MYI overestimation and fluctuation was observed in OSI SAF sea ice types in October 2020. This due to the lack of training data for FYI for most of the month in October. This caused a generation of dynamical PDFs that favored MYI more than was realistic. The situation stabilized at the end of October because of the availability of training data within the target regions.

5.2. Comparison of FYI and MYI Extent at Arctic Sub Regions with Validation Products

The monthly mean FYI and MYI extent derived from the CSCAT and OSI SAF sea ice type products in eight subregions are shown in Figures 15 and 16, respectively. In the CA region, as shown in Figure 15a, the FYI extent and change trend of CSCAT and OSI SAF are very consistent with each other, both showing a continuous increasing trend from October to May of the next year. Overall, the FYI extent derived from OSI SAF is larger than that of CSCAT. CSCAT also provides results in the period when OSI SAF identified ambiguous ice. It can be seen that the FYI extent reached the maximum in June and then decreased to October. As for the MYI extent in the CA region, as shown in Figure 16a, both the CSCAT results and OSI SAF sea ice type product show a continuous decreasing trend from October to May of the next year, and the MYI extent derived from CSCAT is larger than that of OSI SAF sea ice type. From June to August, there was almost no MYI shown in the CSCAT results, but the MYI extent increased in September. Specifically, in middle and late September of each year, CSCAT determined most of the sea ice as MYI, and the prediction accuracy is quite high, as shown in Figures 5 and 6. This indicates that the CSCAT results have a better representation even in the period when the OSI SAF sea ice type is classed as ambiguous.

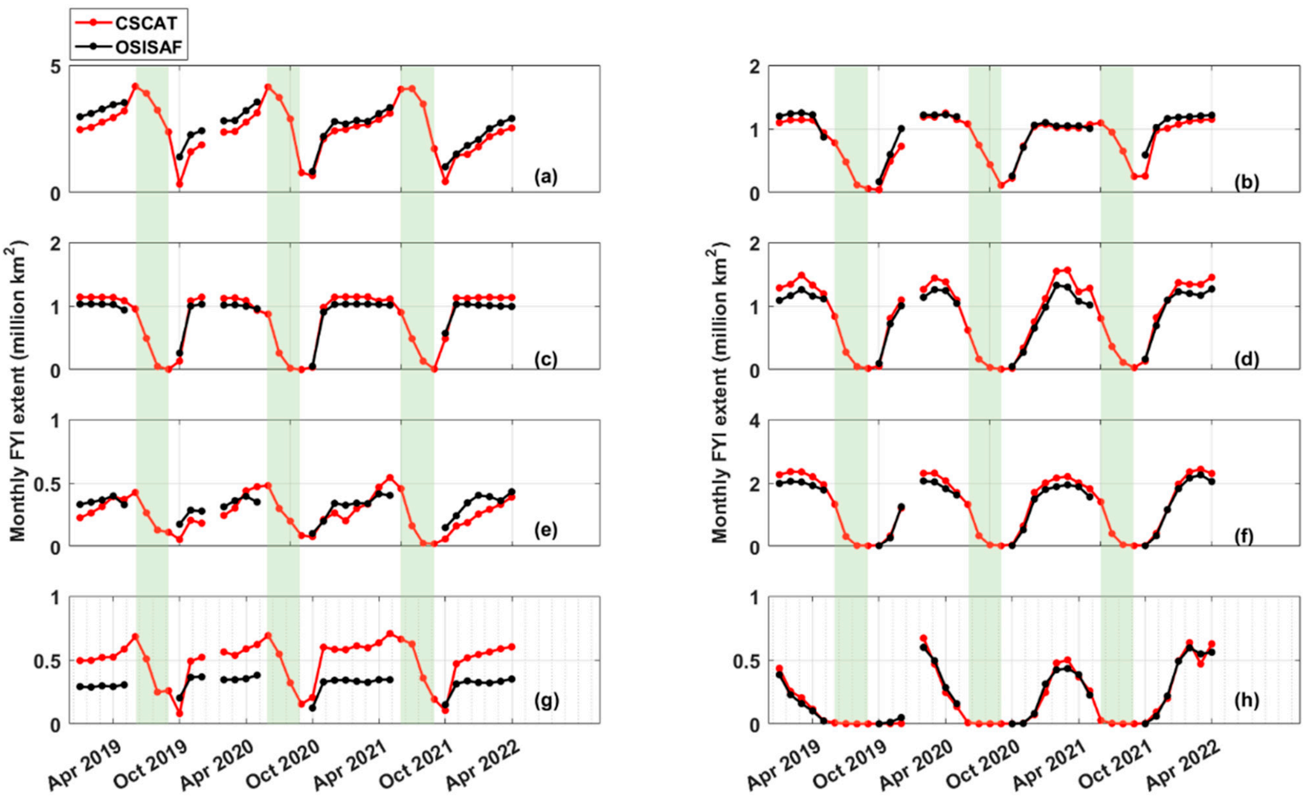

Figure 15. Monthly mean FYI extent derived from CSCAT (red dotted line) classification and OSI SAF sea ice type product (black dotted line) from January 2019 to April 2022 in the (**a**) CA, (**b**) CBS, (**c**) LESS, (**d**) KBS, (**e**) EG, (**f**) HBB, (**g**) CAA, and (**h**) BS regions.

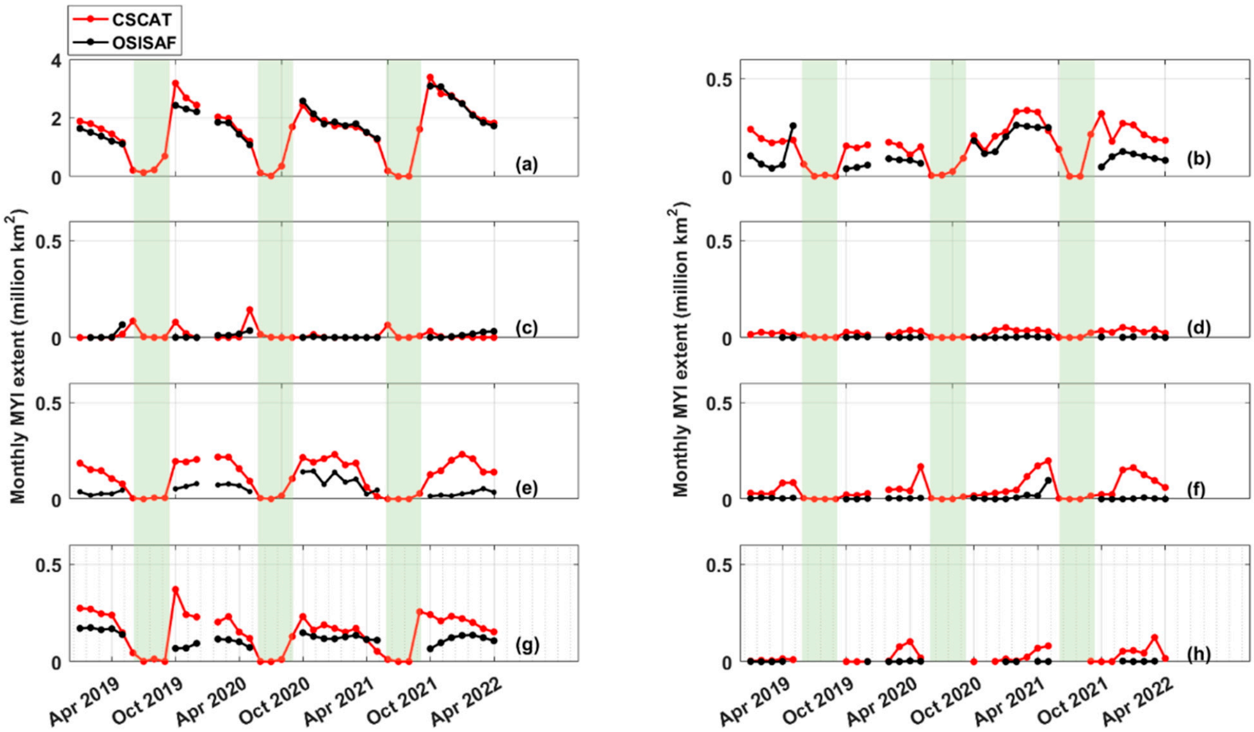

Figure 16. Monthly mean MYI extent derived from CSCAT (red dotted line) classification and OSI SAF sea ice type product (black dotted line) from January 2019 to April 2022 in the (**a**) CA, (**b**) CBS, (**c**) LESS, (**d**) KBS, (**e**) EG, (**f**) HBB, (**g**) CAA, and (**h**) BS regions.

In the CBS region, as shown in Figure 15b, good consistency between CSCAT results and OSI SAF sea ice type can be seen in the time series of FYI extent. The FYI extent increases continuously before December, and stabilizes from December to May of the next year under the impact of the Beaufort Gyre. In Figure 16b, the MYI extent derived from CSCAT is higher than that of OSI SAF sea ice type in this region. In the LESS region, as shown in Figure 15c, the FYI extent in October of each year increases to November and remains stable until May of the next year. During the period when OSI SAF identifies ambiguous ice, the CSCAT FYI decreases continuously. As for the MYI extent shown in Figure 16c, because of the inflow from other subregions, both OSI SAF and CSCAT showed a significant increase in MYI extent in May 2019 and May 2020. The MYI extent increase in October 2019 is related to CSCAT misclassification, as shown in Figure 14.

In the KBS region, as shown in Figure 15d, the trends in FYI from CSCAT and OSI SAF are identical, continuously increasing from October to March of the next year. In the period when sea ice is defined as ambiguous ice by OSI SAF, the FYI extent from CSCAT has a continuous decreasing trend with the minimum extent in September. It can be seen in Figure 16d that there is almost no MYI in the KBS region throughout the year, which is consistent with previous studies. In the EG region, as shown in Figure 15e, the FYI extent shows a gradually increasing trend from October to May of the next year. In the month of OSI SAF ambiguous ice, the CSCAT FYI decreased continuously, reaching its lowest value in September. As shown in Figure 16e, the MYI extent from CSCAT is much higher than that of OSI SAF, which is consistent with the results in Figures 9 and 10. This indicates that in such a highly dynamic ice regime, CSCAT tends to identify the areas with lower MYIC as MYI.

In the HBB region, as shown in Figure 15f, the FYI extent shows a gradually increasing trend from October to March of the next year. The MYI in this region should be rather limited, and the MYI overestimation from CSCAT is due to the poor quality of CSCAT L2A, which should be noted in further analysis. In the CAA region, as shown in Figures 15g and 16g, both FYI and MYI maintain a stable extent from November to March of the next year.

Obviously, the MYI extent from CSCAT is much higher than that from OSI SAF sea ice type in the CAA region. Considering the data from the BS region, as shown in Figures 15h and 16h, FYI increases from October to February of the next year, and then decreases continuously. During the OSI SAF ambiguous ice period, the FYI extent from CSCAT is almost zero. Similar to the HBB region, the MYI in the BS region should be rather limited, the MYI overestimation from CSCAT is due to the poor quality of CSCAT L2A, which should be noted in further analysis.

5.3. Comparison with SAR-Based Sea Ice Type Products

In order to use high-resolution products to validate our results from a long-term period, the SAR-based sea ice type products in Sentinel-1 SAR data using a convolutional neural networks (CNN) algorithm released by CMEMS are used in this study. Figure 17 shows the time series of a three-day moving average MYI extent derived from SAR-based sea ice type, CSCAT, and OSI SAF sea ice type over the region covered by SAR images all-year round during 2021. The grey-shaded area represents the time period with an ambiguous ice type classification from the OSI SAF sea ice type product. Good consistency in the trend from these three results can be seen in Figure 17. The correlation coefficient of SAR and the CSCAT result, and SAR and the OSI SAF are 0.74 and 0.73, respectively. Additionally, the correlation coefficient for the CSCAT and OSI SAF results is up to 0.96. The averaged MYI extents in the analyzed period derived from SAR, CSCAT, and OSI SAF are 5.12×10^5, 4.69×10^5 and 2.89×10^5 km^2, respectively. During the period from Day 26 to Day 140, the MYI extents derived from SAR are much higher than the other two results, whereas in the Day 279–Day 318 period, the MYI extent from CSCAT is the largest compared with the other two results. From Day 318 to the end of 2021, the MYI extents from CSCAT have quite good consistency with those of SAR.

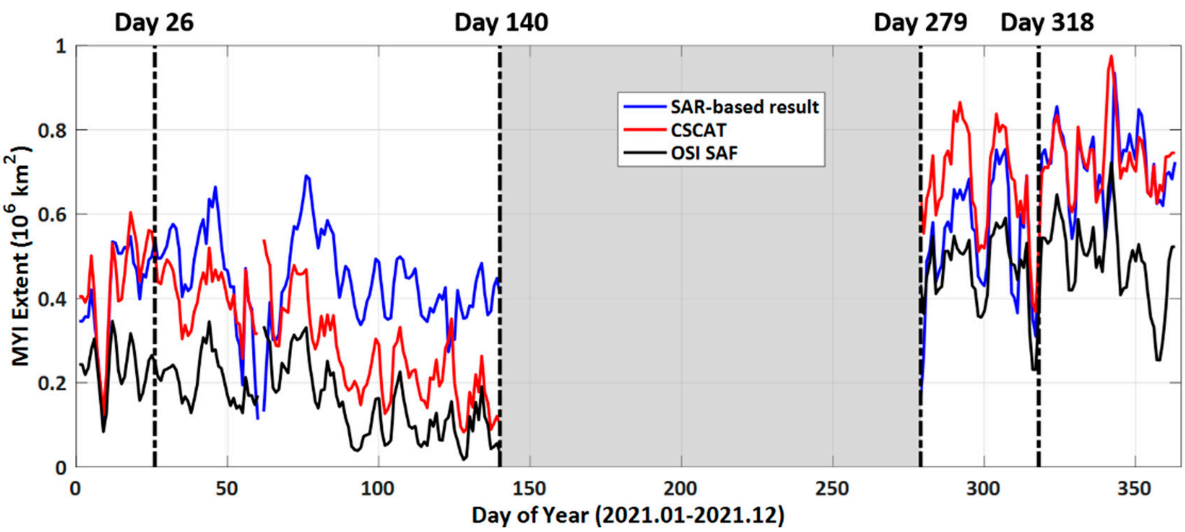

Figure 17. Time series of 3-day moving average MYI extent derived from SAR-based sea ice type (blue line), CSCAT (red), and OSI SAF sea ice type (black) over the region covered by SAR images all year round in 2021.

Figure 18 shows SAR-based sea ice type images superimposed with MYI edges of the CSCAT results and OSI SAF sea ice type product on 17 January 2021, 26 February 2021, 19 October 2021, and 9 November 2021. Generally, the MYI edges derived from CSCAT are much more consistent with the SAR-based sea ice type distribution than that of the OSI SAF sea ice type product. Especially in the area of East Greenland, more area is classified as MYI by CSCAT in such a highly dynamic ice regime, which is consistent with the case in Figure 16e.

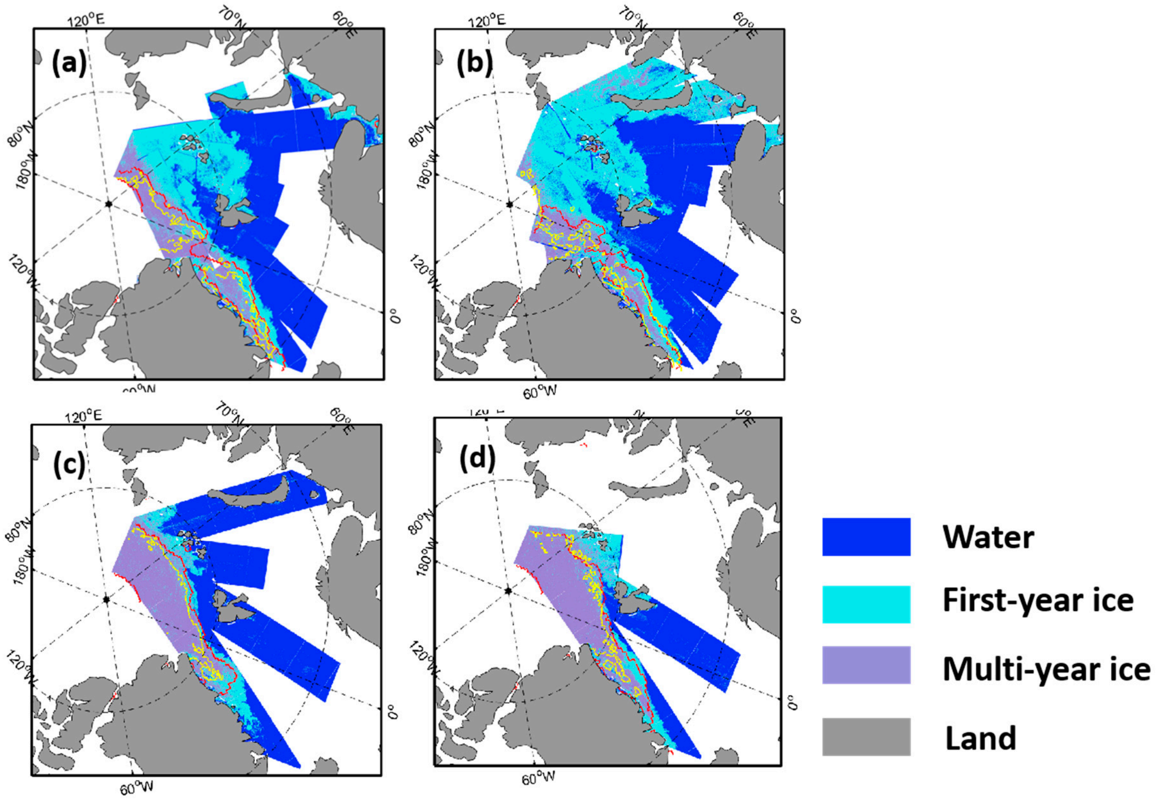

Figure 18. SAR-based sea ice type images superimposed with MYI edges of CSCAT results (red line) and OSI SAF sea ice type product (yellow line) on (**a**) 17 January 2021 (Day 17 of 2021), (**b**) 26 February 2021 (Day 57 of 2021), (**c**) 19 October 2021 (Day 292 of 2021), and (**d**) 24 November 2021 (Day 328 of 2021).

6. Conclusions

The main goal of CSCAT is to monitor ocean surface wind to improve numerical weather forecasting, with the additional potential of monitoring sea ice parameters. In this paper, a new algorithm for classification of Arctic sea ice types on CSCAT measurement data using the random forest classifier is presented.

The random forest classifier is trained on NSIDC weekly SIA and SIC products, where the prediction model is basically updated twice a month. Different from previous studies where the incidence angle normalization correction was used based on daily measurement, five feature parameters, including the mean value of horizontal and vertical polarization backscatter ($\overline{\sigma}_{hh}$ and $\overline{\sigma}_{vv}$), the standard deviations of horizontal and vertical polarization measurements ($\Delta\sigma_{hh}$ and $\Delta\sigma_{vv}$), and the copol ratio ($\gamma = \overline{\sigma}_{vv}/\overline{\sigma}_{hh}$) are innovatively extracted from orbital measurement for the first time to distinguish water, FYI, and MYI. The spatial distribution and probability density distribution of feature parameters indicate that the differences between $\overline{\sigma}_{hh}$ and $\overline{\sigma}_{vv}$ for FYI and MYI are quite significant, where the MYI area is generally higher than the FYI and water area. Furthermore, $\Delta\sigma_{vv}$, $\Delta\sigma_{hh}$ and γ can distinguish water and ice well, but cannot classify sea ice type very well, proving the feasibility of CSCAT orbital feature analysis for sea ice classification. The sea ice type model based on the random forest classifier is quantitatively assessed using a confusion matrix through a comparison with the NSIDC weekly SIA product. During the period from 1 January 2019 through 1 March 2020, the averaged overall accuracy and kappa coefficient are 93.35% and 88.53%, respectively, and the precisions of water, FYI, and MYI are 99.67%, 86.60%, and 79.74%, respectively.

The CSCAT sea ice classification in this study is compared against the sea ice type product (OSI-403-c) from EUMETSAT OSI SAF, the NSIDC weekly SIA product, MYIC from IUP, and the SAR-based sea ice type product from CMEMS. The spatial distribution differences between the CSCAT results and the OSI SAF sea ice type products on selected dates from 2019 to 2021 indicate that the most obvious difference in the distribution of sea ice types between these two results is mainly concentrated in the marginal zones of FYI and MYI, where some areas identified as MYI by the CSCAT classification were classified as FYI by the OSI SAF sea ice type product. This may be caused by the different response of microwave signals with different frequencies to the mixed pixel detection of FYI and MYI. Furthermore, compared with the OSI SAF sea ice type product (OSI-403-c), the area of MYI derived from CSCAT is more homogeneous with less noise, especially for younger multiyear ice such as second-year ice. In the EG region, CSCAT identifies more pixels as MYI with lower MYIC values, showing better accuracy in the identification of areas with obvious mobility of MYI.

The time series of FYI extent derived from the CSCAT and OSI SAF shows that the FYI extent of CSCAT has an excellent consistency with that of OSI SAF sea ice type, the RMSE and correlation coefficient of which are 6.87×10^5 km^2 and 0.989, respectively. Case studies on the first day of the end of OSI SAF ambiguous ice in 2019–2021 show that CSCAT can better describe the distribution of sea ice type at the initial stage of the freezing-up period. The time series of MYI extent derived from CSCAT and the OSI SAF sea ice type product shows the patten of MYI extent of CSCAT has good consistency with that of OSI SAF sea ice type, the RMSE and correlation coefficient of which are 6.84×10^5 km^2 and 0.792, respectively. There are mainly two factors leading to CSCAT MYI extent being higher than that of OSI SAF sea ice type as a whole. Firstly, CSCAT marks the lower MYIC pixels as MYI pixels, and secondly, OSI SAF sea ice type (OSI-403-c) underestimates MYI due to the insensitivity of ASCAT measurement to the younger ice type. The MYI monthly STD of the daily differences calculated from CSCAT and OSI SAF sea ice type product is less than 1×10^5 km^2 in most of the analyzed period, indicating that the sea ice type classification accuracy is reliable as a whole. However, in October 2019 and April 2020, the deviation of CSCAT is relatively large due to the overestimation of MYI in the LESS area and the warm air intrusion effect, respectively. Furthermore, MYI overestimation and fluctuations in October 2020 derived from OSI SAF sea ice type are mainly due to the lack of training data for FYI for most of the month in October, which will be improved in the updated product (OSI-403-d).

In order to analyze the spatial characteristics of the Arctic sea ice in more detail, eight subregions are used in this study. The monthly mean FYI and MYI extent derived from CSCAT and the OSI SAF sea ice type product in the eight subregions show that the MYI extent derived from CSCAT is larger than that of OSI SAF sea ice type in the CA, CBS, EG, and CAA regions. Regarding the HBB and BS regions, the MYI in the BS region is rather limited, and the MYI overestimation from CSCAT is due to the poor quality of CSCAT L2A, which should be analyzed in further study. The SAR-based sea ice type products in Sentinel-1 SAR data were used to validate our results over a long-term period. The comparison with SAR-based sea ice type product found that the correlation coefficient for the SAR and CSCAT, and SAR and OSI SAF are 0.74 and 0.73, respectively. The averaged MYI extents in the analyzed period derived from SAR, CSCAT and OSI SAF are 5.12×10^5, 4.69×10^5 and 2.89×10^5 km^2, respectively. The MYI edges derived from CSCAT are much more consistent with SAR-based sea ice type distribution than those of OSI SAF. Especially in the area of East Greenland, which is a highly dynamic ice regime, more area is classified as MYI by CSCAT.

In conclusion, this research verifies the capability of CSCAT in monitoring Arctic sea ice classification, especially in the spatial homogeneity and detectable duration of sea ice. Given the high accuracy and processing speed, the random forest-based algorithm can offer good guidance for sea ice classification using FY-3E/RFSCAT, i.e., a dual-frequency (Ku and C band) scatterometer called WindRAD. In future work, different machine learning models, such as support vector machine, gradient boosting, and K-means unsupervised learning algorithm and so forth, should also be evaluated for sea ice classification, which will lead to further algorithm improvements.

Author Contributions: Conceptualization, X.Z. and N.X.; methodology, X.Z.; software, X.Z.; validation, X.Z., R.X., Y.S., L.T. and S.T.; formal analysis, L.T., Z.W. and S.T.; investigation, X.Z. and R.X.; resources, N.X., X.H. and L.C.; data curation, X.Z., R.X. and Z.Z.; writing—original draft preparation, X.Z.; writing—review and editing, Z.W., Z.Z., R.X. and Y.S.; visualization, X.Z.; supervision, Z.Z., N.X., X.H. and L.C.; project administration, N.X.; funding acquisition, X.Z., N.X. and Z.Z.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China Meteorological Administration Innovation and Development Project: Remote sensing monitoring and analysis system for long-term dynamic changes of Arctic sea ice, the Fengyun Satellite Application Advance Plan Project of the China Meteorological Administration (No. FY-APP-2021.0505), and the China Youth Foundation of the National Satellite Meteorological Center (No. 412672).

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the Chinese National Satellite Ocean Application Services (NSOAS) for providing the CSCAT data to develop the sea ice classification algorithm, the National Snow and Ice Data Center (NSIDC) for providing sea ice concentration and weekly sea ice age products to train the random forest-based prediction model, the Ocean and Sea Ice Satellite Application Facility (OSI SAF) for providing the sea ice type and LR sea ice drfit products, the University of Bremen, Institute of Environmental Physics (IUP) for providing the multi-year ice concentration product, and the Copernicus Marine Environment Monitoring Service for proving the SAR-based sea ice type product for comparison and validation. The authors also would like to thank the editor and reviewers whose insightful comments offered significant improvements to this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kwok, R. Arctic sea ice thickness, volume, and multiyear ice coverage: Losses and coupled variability (1958–2018). *Environ. Res. Lett.* **2018**, *13*, 105005. [CrossRef]
2. You, Q.; Cai, Z.; Pepin, N.; Chen, D.; Ahrens, B.; Jiang, Z.; Wu, F.; Kang, S.; Zhang, R.; Wu, T.; et al. Warming amplification over the Arctic Pole and Third Pole: Trends, mechanisms and consequences. *Earth-Sci. Rev.* **2021**, *217*, 103625. [CrossRef]
3. Serreze, M.C.; Meier, W.N. The Arctic's sea ice cover: Trends, variability, predictability, and comparisons to the Antarctic. *Ann. N. Y. Acad. Sci.* **2019**, *1436*, 36–53. [CrossRef] [PubMed]
4. Sandven, S.; Johannessen, O.M. Sea ice monitoring by remote sensing. In *Manual of Remote Sensing: Remote Sensing of the Marine Environment*, 3rd ed.; James Gower, F.R., Ed.; The American Society for Photogrammetry & Remote Sensing: Bethesda, MD, USA, 2006; Volume 6, pp. 241–283.
5. Pope, A.; Rees, W.G.; Fox, A.J.; Fleming, A. Open Access Data in Polar and Cryospheric Remote Sensing. *Remote Sens.* **2014**, *6*, 6183–6220. [CrossRef]
6. Shokr, M.E.; Sinha, N.K. Arctic Sea Ice Microstructure Observations Relevant to Microwave Scattering. *Arctic* **1994**, *47*, 265–279. [CrossRef]
7. Sinha, N.K.; Shokr, M. *Sea Ice: Physics and Remote Sensing*, 1st ed.; John Wiley & Sons: Hoboken, NJ, USA, 2015; pp. 412–424.
8. Gohin, F.; Cavanié, A. A first try at identification of sea ice using the three beam scatterometer of ERS-1. *Int. J. Remote Sens.* **1994**, *15*, 1221–1228. [CrossRef]
9. Ezraty, R.; Cavanié, A. Intercomparison of backscatter maps over Arctic sea ice from NSCAT and the ERS scatterometer. *J. Geophys. Res. Oceans* **1999**, *104*, 11471–11483. [CrossRef]
10. Grandell, J.; Johannessen, J.; Hallikainen, M. Development of a synergetic sea ice retrieval method for the ERS-1 AMI wind scatterometer and SSM/I radiometer. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 668–679. [CrossRef]
11. Ezraty, R.; Cavanie, A. Construction and evaluation of 12.5-km grid NSCAT backscatter maps over Arctic sea ice. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1685–1697. [CrossRef]
12. Long, D.; Drinkwater, M. Cryosphere applications of NSCAT data. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1671–1684. [CrossRef]
13. Remund, Q.P.; Long, D.G. Sea ice extent mapping using Ku band scatterometer data. *J. Geophys. Res. Oceans* **1999**, *104*, 11515–11527. [CrossRef]
14. Kwok, R. Annual cycles of multiyear sea ice coverage of the Arctic Ocean: 1999–2003. *J. Geophys. Res. Atmos.* **2004**, *109*, 1–13. [CrossRef]
15. Swan, A.M.; Long, D.G. Multiyear Arctic Sea Ice Classification Using QuikSCAT. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3317–3326. [CrossRef]
16. Shokr, M.; Lambe, A.; Agnew, T. A New Algorithm (ECICE) to Estimate Ice Concentration From Remote Sensing Observations: An Application to 85-GHz Passive Microwave Data. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 4104–4121. [CrossRef]
17. Shokr, M.; Agnew, T.A. Validation and potential applications of Environment Canada Ice Concentration Extractor (ECICE) algorithm to Arctic ice by combining AMSR-E and QuikSCAT observations. *Remote Sens. Environ.* **2013**, *128*, 315–332. [CrossRef]

18. Ye, Y.; Heygster, G.; Shokr, M. Improving Multiyear Ice Concentration Estimates With Reanalysis Air Temperatures. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 2602–2614. [CrossRef]
19. Zhang, Z.; Yu, Y.; Li, X.; Hui, F.; Cheng, X.; Chen, Z. Arctic Sea Ice Classification Using Microwave Scatterometer and Radiometer Data During 2002–2017. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5319–5328. [CrossRef]
20. Lindell, D.B.; Long, D.G. Multiyear Arctic Ice Classification Using ASCAT and SSMIS. *Remote Sens.* **2016**, *8*, 294. [CrossRef]
21. Melsheimer, C.; Spreen, G. *AMSR-E ASI Sea Ice Concentration Data, Arctic, Version 5.4 (NetCDF) (June 2002–September 2011)*; PANGAEA: Bremen, Germany, 2020.
22. Aaboe, S.; Down, E.J.; Eastwood, S. *Product User Manual for the Global Sea-Ice Edge and Type Product*; Norwegian Meteorological Institute: Oslo, Norway, 2021.
23. Lindell, D.B.; Long, D.G. Multiyear Arctic Sea Ice Classification Using OSCAT and QuikSCAT. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 167–175. [CrossRef]
24. Li, M.; Zhao, C.; Zhao, Y.; Wang, Z.; Shi, L. Polar Sea Ice Monitoring Using HY-2A Scatterometer Measurements. *Remote Sens.* **2016**, *8*, 688. [CrossRef]
25. Xu, R.; Zhao, C.; Zhai, X.; Zhao, K.; Shen, J.; Chen, G. Polar Sea Ice Identification and Classification Based on HY-2A/SCAT Data. *J. Ocean Univ. China* **2022**, *21*, 331–346. [CrossRef]
26. Zhang, Z.; Yu, Y.; Shokr, M.; Li, X.; Ye, Y.; Cheng, X.; Chen, Z.; Hui, F. Intercomparison of Arctic Sea Ice Backscatter and Ice Type Classification Using Ku-Band and C-Band Scatterometers. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–18. [CrossRef]
27. Zhai, X.; Wang, Z.; Zheng, Z.; Xu, R.; Dou, F.; Xu, N.; Zhang, X. Sea Ice Monitoring with CFOSAT Scatterometer Measurements Using Random Forest Classifier. *Remote Sens.* **2021**, *13*, 4686. [CrossRef]
28. Xu, R.; Zhao, C.; Zhai, X.; Chen, G. Arctic Sea Ice Type Classification by Combining CFOSCAT and AMSR-2 Data. *Earth Space Sci.* **2022**, *9*, e2021EA002052. [CrossRef]
29. Ye, Y.; Shokr, M.; Heygster, G.; Spreen, G. Improving Multiyear Sea Ice Concentration Estimates with Sea Ice Drift. *Remote Sens.* **2016**, *8*, 397. [CrossRef]
30. Lindsley, R.D.; Long, D.G. Enhanced-Resolution Reconstruction of ASCAT Backscatter Measurements. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 2589–2601. [CrossRef]
31. Lin, W.; Dong, X.; Portabella, M.; Lang, S.; He, Y.; Yun, R.; Wang, Z.; Xu, X.; Zhu, D.; Liu, J. A Perspective on the Performance of the CFOSAT Rotating Fan-Beam Scatterometer. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 627–639. [CrossRef]
32. Liu, J.; Lin, W.; Dong, X.; Lang, S.; Yun, R.; Zhu, D.; Zhang, K.; Sun, C.; Mu, B.; Ma, J.; et al. First Results From the Rotating Fan Beam Scatterometer Onboard CFOSAT. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8793–8806. [CrossRef]
33. Meier, W.N.; Fetterer, F.; Windnagel, A.K.; Stewart, S. *NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 4*; National Snow and Ice Data Center: Boulder, CO, USA, 2021. [CrossRef]
34. Tschudi, M.; Meier, W.N.; Stewart, J.S.; Fowler, C.; Maslanik, J. *EASE-Grid Sea Ice Age, Version 4*; NASA National Snow and Ice Data Center Distributed Active Archive Center: Boulder, CO, USA, 2019. [CrossRef]
35. Tschudi, M.; Meier, W.N.; Stewart, J.S. *Quicklook Arctic Weekly EASE-Grid Sea Ice Age, Version 1*; NASA National Snow and Ice Data Center Distributed Active Archive Center: Boulder, CO, USA, 2019. [CrossRef]
36. Aaboe, S.; Down, E.J.; Eastwood, S. *Algorithm Theoretical Basis Document for the Global Sea-Ice Edge and Type Product, Version 3.3*; EUMETSAT Ocean and Sea Ice SAF: Darmstadt, Germany, 2021; Available online: https://osisaf-hl.met.no/sites/osisaf-hl.met.no/files/baseline_document/osisaf_cdop3_ss2_atbd_sea-ice-edge-type_v3p3.pdf (accessed on 3 January 2023).
37. Boulze, H.; Korosov, A.; Brajard, J. Classification of Sea Ice Types in Sentinel-1 SAR Data Using Convolutional Neural Networks. *Remote Sens.* **2020**, *12*, 2165. [CrossRef]
38. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]
39. Mortin, J.; Howell, S.E.; Wang, L.; Derksen, C.; Svensson, G.; Graversen, R.G.; Schrøder, T.M. Extending the QuikSCAT record of seasonal melt–freeze transitions over Arctic sea ice using ASCAT. *Remote Sens. Environ.* **2014**, *141*, 214–230. [CrossRef]
40. Aaboe, S.; Down, E.J.; Eastwood, S. *Validation Report for the Global Sea-Ice Edge and Type Product, Version 3.1*; EUMETSAT Ocean and Sea Ice SAF: Darmstadt, Germany, 2021; Available online: https://osisaf-hl.met.no/sites/osisaf-hl.met.no/files/validation_reports/osisaf_cdop3_ss2_svr_sea-ice-edge-type_v3p1.pdf (accessed on 3 January 2023).