Fine Resolution Classification of New Ice, Young Ice, and First-Year Ice Based on Feature Selection from Gaofen-3 Quad-Polarization SAR

Kun Yang 1, Haiyan Li 1,2,*, William Perrie 3, Randall Kenneth Scharien 4, Jin Wu 1,5, Menghao Zhang 1 and Fan Xu 1

1 College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
2 Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China
3 Fisheries & Oceans Canada, Bedford Institute of Oceanography, Dartmouth, NS B2Y 4A2, Canada
4 Department of Geography, University of Victoria, Victoria, BC V8P 5C2, Canada
5 Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
* Correspondence: lihaiyan@ucas.ac.cn

Abstract: A new method of sea ice classification based on feature selection from Gaofen-3 polarimetric Synthetic Aperture Radar (SAR) observations was proposed. The new approach classifies sea ice into four categories: open water (OW), new ice (NI), young ice (YI), and first-year ice (FYI). Seventy parameters that have previously been applied to sea ice studies were re-examined for sea ice classification in the Okhotsk Sea near the melting point on 28 February 2020. The ‘separability index (SI)’ was used for the selection of optimal features for sea ice classification. Full polarization parameters (the backscatter intensity contains the horizontal transmit-receive intensity (σ^0_{hh}), Shannon entropy (SE_i)), the spherical scattering component of Krogager decomposition (Ks)), and hybrid polarization parameters (horizontal receive intensity (σ^0_{rh}), hybrid-pol Shannon entropy (CPSE_i), the correlation coefficient (ρ_{rh-rv}) between the σ^0_{rh} and σ^0_{rv}, and the surface scattering component of m − α decomposition (α_s)) were determined as the optimal parameters for the different work modes of SAR. The selected parameters were used to classify sea ice by the random forest classifier (RFC), and classification results were validated by manually interpreted ice maps derived from Landsat-8 data. The classification accuracy of OW, NI, YI and FYI reached 95%, 96%, 98% and 85%, respectively.

Keywords: fine sea ice classification; random forest classifier; polarization SAR; Gaofen-3; feature selection

1. Introduction

Sea ice in the Arctic is an indicator of climate change and has undergone dramatic changes because of recent global warming [1]. The changes in Arctic sea ice mainly include the continuous decreases of sea ice extent, reductions in multi-year ice (MYI) thickness, and the conversion of increasingly more MYI into first-year ice (FYI) [2]. FYI is becoming more complex, especially during the summer months. The types and distributions of sea ice directly affect the Arctic energy budget. For example, the surface albedos of melting ice, snow-covered MYI, and open water (OW) are, respectively, 65%, 85%, and 7% [3]. Increasing areal fractions of melting ice and OW boost solar radiation input to the Arctic Ocean, intensifying sea ice melt via a positive sea ice—albedo feedback effect [4]. As the sea ice grows or melts, various types of sea ice, such as new ice (NI), young ice (YI), and FYI, will emerge. Referring to the definitions of ice type by the Canadian Ice Service (Sea ice: types and forms—Canada.ca (https://www.canada.ca/en/environment-climate-change/services/ice-forecasts-observations/latest-conditions/educational-resources/sea/types-forms.html, accessed on 20 October 2022)), NI is the initial stage of sea ice formation, characterized by a thickness of less than 10 cm. YI follows NI and is thicker, ranging from 10 to 30 cm, with
a rough texture. FYI is sea ice that has survived one melt season and is between 30 and 150 cm thick. The properties and characteristics of sea ice can change as it ages and thickens. For example, FYI may have more brine channels and be more permeable to seawater than NI, which can affect its reflectivity and, therefore, its contribution to the Earth’s energy balance [5,6]. Additionally, the thickness and age of sea ice can affect its interactions with ocean currents and the atmosphere as well as its roles in the carbon and nutrient cycles [7,8]. Accurately classifying sea ice and understanding the differences between these ice types are important for accurate modeling and predictions of the behavior of the sea ice, and its impact on global climate [9,10].

Synthetic aperture radar (SAR) is a relatively practical tool for sea ice monitoring due to its low sensitivity to clouds, rain, and fog as well as its capability for high-resolution earth observation in daylight or darkness [11]. In recent years, SAR systems have developed progressively from single polarization (single-pol) to dual polarization (dual-pol) and quad-polarization (quad-pol). A quad-pol SAR system simultaneously transmits and receives both linearly horizontal (H) and vertical (V) polarized electromagnetic waves. Amplitude and phase information of the backscattered signals are recorded for four transmit/receive channels (HH, HV, VH, and VV). Dual-pol SAR contains two channels (e.g., HH and HV or VV and VH), whereas single-pol contains only one channel (HH or VV) without phase information [12]. Dual-pol SAR scenes, normally with HH and HV channels, are favored for monitoring sea ice over vast areas due to their wide coverage [13–15]. For example, the RADARSAT-2 ScanSAR Wide mode images have a swath width of approximately 500 km. Compared with dual-pol, quad-pol observations with phase information and higher spatial resolutions can provide more information on sea ice [13,15,16]. However, the unprecedented level of detailed scattering information given by quad-pol SAR observations comes at the cost of a doubled average transmission power, low width coverage, and a restricted range of acceptable incidence angles, compared with single-pol or dual-pol SAR observations [17,18]. Moreover, the narrow swath of quad-pol SAR (e.g., RADARSAT-2, 50 km; TerraSAR-X, 15 km; and ALOS PALSAR, between 20 and 60 km) limits the practical applications of these observations [19].

More recently, the compact polarimetry (CP) SAR mode has been proposed [17,18]. The essence of the CP system is a coherent dual-pol SAR with one polarization transmitted and two orthogonal polarizations received. Though not fully polarimetric, CP SAR has been shown to obtain similar results as quad-pol SAR systems in sea ice classification [20,21]. CP SAR systems simplify system design and maintenance while expanding the swath width compared with quad-pol systems [17,18]. This system is an optimal compromise between resolution and coverage. For example, the available swath width of the CP SAR from the RADARSAT Constellation Mission (RCM) is 350 km with a spatial resolution of 50 m. These values are still suitable for the application of CP SAR to sea ice observation, given that it is a large-scale feature [17,18,22]. Presently, three Earth-observing satellite systems are equipped with the advanced hybrid-polarity (hybrid-pol) CP design, which transmits circular polarization and receives two orthogonal linear polarizations. They are the Radar Imaging Satellite (RISAT-1), launched in April 2012 (India), the ALOS-2, launched in May 2014, and the RCM satellites, launched in June 2019. Prospective applications of hybrid-pol SAR observation have generated considerable attention in sea ice monitoring [16,19,21,23,24].

With the progression in SAR systems from single-pol to dual-pol, quad-pol and hybrid-pol, a large number of parameters have been proposed for sea ice classification. Originally, the backscatter intensity and polarization ratio (PR) were used to discriminate OW from sea ice [25,26] or to identify ice types [27–30]. Moreover, the parameters using polarization decomposition or eigenvector decomposition from quad-pol SAR have been used to improve the accuracy of sea ice classification [31–35]. Studies using simulated hybrid-pol and real hybrid-pol data to obtain parameters have also proven to be effective for sea ice classification [16,19,21–23]. Many of these works in the studies above concentrated on the separation of OW and sea ice, classification between FYI and MYI, and analyses of various sea ice types and properties using different polarimetric parameters [19,22,26].
However, classification studies for fine sea ice with high spatial resolution and many subtypes of sea ice are rather few. Liu et al. [15] used the dual-pol RADARSAT-2 ScanSAR (HH + HV) data to classify sea ice as OW, NI, gray ice, second-year ice, and MYI based on the support vector machine methodology. Zhang et al. [30] employed deep convolutional neural networks to classify sea ice into four categories, NI, thin first-year ice, thick first-year ice, and old ice, by using dual-pol Gaofen-3 (GF-3) data (VH + VV). Both of these works [15,30] used the ice maps provided by the CIS as a reference and obtained high accuracy (over 90%) for the classification results. As mentioned before, the traditional dual-pol data could not provide phase information. They are used to classify ice types only based on the backscatter intensity or texture feature. Moreover, quad-pol and hybrid-pol can provide both intensity and phase information, related to the physical scattering mechanism of sea ice, surface roughness, dielectric coefficient, and structure, which can improve the classification accuracy.

Therefore, there are two facts that hinder fine sea ice classification. Firstly, sea ice is complex, dynamic, and interconnected with the environment, particularly FYI [36,37], and it is difficult to collect actual sea ice observations with both large area coverage and high spatial resolution. Secondly, even though a large number of SAR characteristics have been used to classify sea ice, it remains unclear which parameters are the most effective for different types of sea ice and different environmental conditions.

In this study, a methodology is developed for fine sea ice classification, particularly OW, NI, YI, and FYI based on feature selection from GF-3 quad-pol SAR. All the parameters related to both the intensity and the phase information from quad-pol and compact-pol are taken into account. The ice types are identified in SAR images using visible images as a reference. The key point of this work is to identify the optimal parameter combination for fine sea ice classification from 70 potential parameters using the separability index (SI) method.

The paper is organized as follows: Section 2 presents the dataset, including the Landsat-8 optical images, GF-3 quad-pol SAR images, and weather reanalysis data. In Section 3, the 70 parameters used for sea ice classification are summarized. Section 4 demonstrates how to use the separability index (SI) to select the optimal sea ice classification parameters and how to incorporate these parameters into the training of a random forest classifier (RFC) for subsequent sea ice classification. The classification results are given in Section 5. Section 6 is the discussion of our results and the conclusions are presented in Section 7.

2. Dataset

The study utilizes two distinct types of remote sensing data. The first type is GF-3 quad-pol SAR data, and the second type is Landsat-8 optical data. We collocated all GF-3 and Landsat-8 data from a worldwide latitude region bounded by 50–90°N and 50–90°S from August 2018 to August 2022. In addition to imaging time and area, during the time lag between optical and SAR data, sea surface temperature and wind speed were also taken into account. Six SAR images and two Landsat-8 images were finally selected, as shown in Figure 1a, over the Okhotsk Sea. Landsat-8 scenes, acquired at 01:41 (UTC) on 28 February 2020, are denoted by labels L1 and L2 and their extents are shown by blue rectangles. GF-3 images, acquired at 21:06 (UTC) also on 28 February 2020, are denoted by labels S1–S6 and their extents are shown by red rectangles. The background color in Figure 1a is the hourly average wind speed between the SAR imaging time and the Landsat-8 imaging time; arrows show the wind vectors. Figure 1b displays the hourly average wind speed at the center of each SAR scene on the day of imaging. The S1, S2, and S3 images are accompanied by wind speeds greater than 6 m/s, whereas wind speeds for S4, S5, and S6 are typically less than 5.5 m/s. On that day, we reviewed the sea surface temperature data, which was almost 0 degrees, and sea ice concentration data, from the University of Bremen (https://seaice.uni-bremen.de/databrowser/, accessed on 20 October 2022), which showed a slight decrease (from 76.4% to 76.1%) in the study area from 28 February to 29 February 2020. This implies that the sea ice may be close to melting,
by combining the surface temperature and the variation trend of the sea ice concentrations. All meteorological data are ECMWF reanalysis data (https://cds.climate.copernicus.eu/, accessed on 20 October 2022).

2.1. Landsat-8

Landsat-8 was launched on 11 February 2013, with two science instruments—the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). These two sensors provide systematic coverage of the global landmass with a spatial resolution of 30 m for visible data, namely the NIR, and SWIR; 100 m for thermal data; and 15 m for panchromatic data. The data has been widely used in crop classification [38], water body extraction [39], and sea ice identification [40].

In this study, the two Landsat-8 images are a Level-1T product with 30 m pixel spacing and less than 10% cloud cover (Figure 2). The image preprocessing, including atmospheric and radiometric corrections, was performed by ENVI 5.3. From Figure 2, we see that most of the imaged areas are covered with YI and FYI. For example, the upper left area of Figure 2a appears gray, which corresponds to the relatively dark appearance of YI in the optical image; by comparison, white areas in Figure 2a,b are FYI. The black area in the lower right of Figure 2b is land. The data in the red rectangle, which covers an area with a length of 12 km and a width of 9.6 km in Figure 2b, has been chosen for more detailed investigation.

2.2. GF-3 SAR Images

GF-3, with a C-band SAR capable of full polarization mode, was launched in August 2016. It is China’s first satellite carrying a multi-pol C-band SAR (https://spaceflight101.com/chinas-long-march-4c-lifts-off-with-gaofen-3-radar-satellite/, accessed on 20 October 2022). Employing a multi-polarized active-phased array antenna, the GF-3 SAR can achieve a nominal resolution as high as 1 m in Spotlight mode and a large swath width of 500 km in Wide ScanSAR mode. There are three quad-pol modes: Stripmap I (QPSI), Stripmap II (QPSII), and wave mode [41]. In our study, QPSI mode data was used, with a 30-km ground swath and an 8-m nominal resolution. The data is single-look complex, with incidence angles ranging from 20° to 41°. The GF-3 data have been used in retrievals of sea state parameters, ship detection, and sea ice monitoring [42].
For each of the six GF-3 SAR images (Figure 1), the assumed incidence angle was taken as the mid-point of the swath, ranging from 35.23° to 37.20°. There were no grazing or shadow effects. The HH intensities of the six SAR images are shown in Figure 3. Compared with Landsat-8 data, these SAR images exhibit more complex distributions and have higher resolutions, which indicates that richer sea ice information can be retrieved. The red rectangular areas of panel S5 in Figure 3 and panel L2 in Figure 2b are in the same geographical location. Figure 4 enlarges this area to show the different types of sea ice more clearly.

The imaging time of the Landsat-8 and GF-3 images was on 28 February 2020, which corresponds to the transition period between late winter and early spring at these relatively low latitudes (about 55–56°N). Under these conditions, with above-freezing temperatures, the appearance of liquid water on the sea ice surface is possible, which complicates sea ice classification. Our analysis of the high-resolution collocation of Landsat-8 and GF-3 observations in Figure 4 indicates different types of sea ice, specifically OW, NI, YI, and FYI. As shown in Figure 4a, in the optical image, the dark area is NI, often near OW. YI is the grey color and FYI is displayed with white color due to its high albedo. In contrast, Figure 4b, depicting the HH images of GF-3, shows that YI has the highest intensity, while those of OW, NI, and FYI have similar return intensities. This means that OW, NI, and FYI cannot be differentiated only by intensity. Visual evaluation of these two categories of data shows that optical data outperforms the SAR data in its ability to discriminate sea ice types and OW. Therefore, in our study, optical data were used as a vital reference to help us interpret the sea ice types in the SAR images.
Figure 3. Backscatter intensity in HH polarization from the six GF-3 SAR images. The red rectangle in S5 is investigated in Figure 4 to elucidate the different types of sea ice.

Figure 4. Different types of sea ice in coincident landsat-8 (a) and GF-3 (b) images. The area is an enlargement of the red rectangular areas in panel L2 in Figure 2b and S5 in Figure 3 and includes OW, NI, YI, and FYI. The scale is located to the right of (b).
3. SAR Parameters

In this section, the 70 parameters derived from the SAR images are summarized, and then optimal parameters for input to the sea ice classification are selected using the SI method. Later, the sea ice classification is illustrated using RFC.

Numerous SAR parameters have been employed to classify ground objects. Here, we explore 70 typical SAR parameters for the possible classification of sea ice. These parameters can be divided into two major categories according to SAR systems: quad-pol parameters and hybrid-pol parameters. The quad-pol parameters include backscatter intensity of different polarizations, co-polarized (co-pol) phase difference, co-pol correlation coefficients, backscattering ratios of different polarizations, total power, secondary parameters with decomposition, and parameters from eigenvalue-decomposition. The hybrid-pol parameters contain backscatter intensity, backscattering ratios, correlation coefficients, Stokes vector parameters, Stokes child parameters, and decomposition parameters.

3.1. Quad-Pol SAR Parameters

The backscatter intensity contains the horizontal transmit-receive intensity \((σ_{0}^{hh})\), vertical transmit-receive intensity \((σ_{0}^{vv})\), and the horizontal transmit-vertical receive intensity \((σ_{0}^{hv})\), which are widely used in the inversion of sea ice concentration and sea ice classification [43,44]. The co-pol phase difference \((∅_{hh−vv})\), co-pol correlation coefficient \((ρ_{hh−vv})\), and backscattering ratios co-pol ratio \((R_{hh−vv})\), cross-pol ratio \((R_{hh−hv}, R_{hv−vv})\), and depolarization ratio \((R_{depol})\), are sensitive to sea ice thickness and differences between sea ice and water [26,45,46]. Total power (Span) indicates the total backscattered power and is a basic parameter that is used to differentiate sea ice classes [45,47]. Huynen [48] was the first to formalize target decomposition theorems in 1970, but their origins may be traced to Chandrasekhar’s research on light scattering by small anisotropic particles. Since the publication of this initial paper, there have been numerous other proposed decompositions [48–56] that can be categorized into four categories [57].

Decompositions were based on the dichotomy of the Kennaugh matrix K [48,53,56]. In this study, we select three parameters based on Huynen decomposition, which are Huynen_T11 \((T_{11})\), Huynen_T22 \((T_{22})\), and Huynen_T33 \((T_{33})\). \(T_{11}\) represents the total scattered power from the regular, smooth, and convex parts of the scatterer. \(T_{22}\) denotes total symmetric or irregular depolarized power, and \(T_{33}\) stands for total non-symmetric depolarized power.

Decompositions were based on “model-based” decompositions of the covariance C3 matrix or the coherency matrix T3 [52,55,58]. In the study, we select two sets of parameters from the three-component Freeman decomposition and the four-component Yamaguchi decomposition. They are Freeman_dbl \((F_{d})\), Freeman_vol \((F_{v})\), Yamaguchi_dbl \((Y_{d})\), Yamaguchi_vol \((Y_{v})\), Yamaguchi_odd \((Y_{s})\), and Yamaguchi_hlx \((Y_{h})\). The postfix ‘dbl’ denotes double-bounce scattering power; ‘odd’ denotes surface scattering power; ‘vol’ denotes volume scattering power; and ‘hlx’ denotes helix scattering power.

Decompositions used an eigenvector or eigenvalues analysis of the covariance C3 matrix or coherency matrix T3 [51,53,59,60]. In this study, two sets of parameters are selected from van Zyl methods and the entropy (H)/anisotropy (A)/alpha (α) decomposition method proposed by Cloude and Pottier in 1996 and 1997. One set contains Zyl1, Zyl2, and Zyl3, which represent odd bounce reflections, even bounce reflections, and other reflections, respectively. The other contains four parameters: H, A, α, λ, where H reflects the randomness degree of the scattering of the targets; A is complementary to H and improves the discrimination of different types of the scattering processes. The parameter α corresponds to a continuous change from surface scattering to double bounce scattering and λ represents the average scattered power.

Decompositions employed a coherent decomposition of the scattering matrix S [54,61,62]. We choose the parameters from the Krogager decomposition, which has three components: \(K_{s}, K_{d},\) and \(K_{h}\). The component \(K_{s}\) denotes the spherical scattering component, \(K_{d}\) denotes the double-bounce scattering component, and \(K_{h}\) represents the helix scattering component.
In addition to the above parameters, there are also some parameters proposed by researchers based on eigenvalues, including Shannon entropy (SE, SEp), target randomness (PR), polarimetric asymmetry (PA), polarization fraction (PF), pedestal height (PH), and relative kurtosis (PK). These features complement the above eigenvalue-based decomposition, which not only reflects the intensity information but also includes the polarization and depolarization information. These have been verified to be closely related to sea ice surface roughness, applicable to the distinction among sea ice, OW, and snow [46].

3.2. Hybrid-Pol SAR Parameters

The hybrid-pol Stokes vector (g0, g1, g2, g3) can be directly derived from the quad-pol correlation matrix. Each parameter has a distinct physical significance: g0 is the total power of the wave; g1 is the power difference between the horizontal and vertical polarized components; g2 is the power in the linearly polarized components with tilt angles of 45° or 135°; and g3 is the power in the left- and right-handed circularly polarized components of the plane wave [63]. The parameters of hybrid-pol backscatter intensity contain the right-circular transmit-receive intensity (σ_rr), the right-circular transmit and left-circular receive intensity (σ_rl), the left-circular transmit-receive intensity (σ_lr), the right-circular transmit and horizontal receive intensity (σ_rh) and the right-circular transmit and vertical receive intensity (σ_rv). These parameters represent the power of various transmission and reception modes [64]. They can all be simulated from the quad-pol SAR data without developing new hardware [19].

Similar to quad-pol SAR, hybrid-pol SAR data also contains correlation coefficients (ρ_rh-rr, ρ_rh-rv, ρ_rr-ll). ρ_rh-rr denotes the correlation coefficient between the σ_rh and σ_rr, and ρ_rr-ll denotes the correlation coefficient between σ_rr and σ_ll. We calculated (R_rh-rr) as the ratio of σ_rh and σ_rr, (R_rr-ll) as the ratio of σ_rr and σ_ll, (R_rl-rr), (R_rv-rr) as the ratio of σ_rl and σ_rr, and (R_rv-ll) as the ratio of σ_rv and σ_ll [64].

With the advance of hybrid-pol SAR, many decompositions have been developed [63,65]. They are m − χ, m − δ, m − ψ and m − α, which have demonstrated utility in oil slick detection [66] and sea ice melt monitoring [16,64]. Here, m indicates the degree of polarization, which is the ratio of the polarized power to the total power, χ is the ellipticity, δ is the relative phase, ψ is the orientation and α is the polarization angle.

The four decompositions can be divided into three components. Division by m − χ decomposition gives (χ, χd, χv), where χd indicates single-bounce (Bragg) backscatter, χv corresponds to randomly polarized backscatter. Division by m − δ decomposition gives (δs, δd, δv), where δs is related to double-bounce scattering, δd is related to volumetric scattering, and δv is related to surface scattering. Division by m − ψ gives (ψs, ψd, ψv), where ψs is sensitive to orientations of −90 degrees, ψd is sensitive to orientations of +90 degrees, and ψv represents the randomly polarized constituent. Division by m − α decomposition gives (αs, αd, αv), where αs is sensitive to a polarization angle of 0 degrees, αd is sensitive to a polarization angle of 90 degrees, and αv is sensitive to dominantly depolarized backscatter. It is worth noticing that the third component (χv, δv, ψv, αv) of the proposed four decomposition methods is expressed in the same way.

Additionally, Charbonneau et al. (2010) proposed the right circular polarization ratio (μC) and ellipticity (μE) [65]. They are sensitive to sea ice characteristics. In a similar way to quad-pol SAR, hybrid-pol SAR can also define the Shannon entropy intensity and polarimetric components (CPSEq, CPSEP) [22], which contain similar information as given by quad-pol SAR data.

A total of 70 parameters from quad-pol and hybrid-pol SAR are used in this study, as summarized with relevant references in Table 1.
### Table 1. Polarimetric parameters in this study.

<table>
<thead>
<tr>
<th>SAR Mode</th>
<th>Sub-Class</th>
<th>Parameters</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quad-pol</strong></td>
<td>Based on backscatter intensity and polarization features</td>
<td>$\sigma_{hh}^0$, $\sigma_{vv}^0$, $\sigma_{hv}^0$, $\varnothing_{hh-vv}$</td>
<td>backscatter intensity (dB)</td>
<td>[45]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\rho_{hh-vv}$</td>
<td>co-pol phase difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R_{hh-vv}$, $R_{hh-hv}$, $R_{hv-vv}$</td>
<td>co-pol correlation coefficients</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Span</td>
<td>$R_{depol}$</td>
<td>backscattering ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Span</td>
<td></td>
<td>depolarization ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Span</td>
<td></td>
<td>total power</td>
<td></td>
</tr>
<tr>
<td><strong>Quad-pol</strong></td>
<td>Based on polarization target decomposition</td>
<td>$T_{11}$, $T_{22}$, $T_{33}$</td>
<td>Huynen decomposition</td>
<td>[48]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$F_d$, $F_v$, $F_s$</td>
<td>three-component freeman decomposition</td>
<td>[52]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Y_{d}$, $Y_v$, $Y_s$, $Y_h$</td>
<td>four-component Yamaguchi decomposition</td>
<td>[55]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$Z_{yl_1}$, $Z_{yl_2}$, $Z_{yl_3}$</td>
<td>vanZyl decomposition</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$H$, $A$, $\alpha$, $\lambda$</td>
<td>$H/A/\alpha$ decomposition</td>
<td>[51]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_s$, $K_d$, $K_h$</td>
<td>Krogager decomposition</td>
<td>[54]</td>
</tr>
<tr>
<td><strong>Other parameters</strong></td>
<td></td>
<td>$\text{SEi}, \text{SEP}$</td>
<td>Shannon entropy</td>
<td>[67]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{PR}$</td>
<td>target randomness</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{PA}$</td>
<td>polarimetric asymmetry</td>
<td>[69]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{PF}$</td>
<td>polarization fraction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{PH}$</td>
<td>pedestal height</td>
<td>[70]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\text{PK}$</td>
<td>relative kurtosis</td>
<td>[71]</td>
</tr>
<tr>
<td><strong>Stokes parameters</strong></td>
<td>$g_0$, $g_1$, $g_2$, $g_3$</td>
<td>Stokes vector elements</td>
<td></td>
<td>[63]</td>
</tr>
<tr>
<td><strong>Hybrid-pol</strong></td>
<td>Based on hybrid-pol and circle-pol backscatter intensity</td>
<td>$\sigma_{rr}^0$, $\sigma_{rl}^0$, $\sigma_{ll}^0$, $\rho_{rr-rl}$, $\rho_{rl-ll}$</td>
<td>hybrid-pol backscatter intensity correlation coefficients</td>
<td>[22]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R_{rll}$, $R_{rlr}$, $R_{llr}$, $R_{lrl}$</td>
<td>hybrid-pol backscattering ratio</td>
<td>[22,65]</td>
</tr>
<tr>
<td><strong>Hybrid-pol</strong></td>
<td>Hybrid-pol decomposition</td>
<td>$m$, $\chi$, $\chi_r$, $\chi_r$, $\chi_v$, $\delta$, $\delta_r$, $\delta_v$</td>
<td>$m - \chi$ decomposition</td>
<td>[17]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m$, $\psi$, $\psi_r$, $\psi_r$, $\psi_v$</td>
<td>$m - \delta$ decomposition</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$m$, $\alpha$, $\alpha_r$, $\alpha_r$, $\alpha_v$</td>
<td>$m - \alpha$ decomposition</td>
<td></td>
</tr>
<tr>
<td><strong>Other parameters</strong></td>
<td>$\mu_C$, $\mu_E$</td>
<td>circular (right) polarization ratio and ellipticity</td>
<td></td>
<td>[65]</td>
</tr>
<tr>
<td></td>
<td>CPSEi, CPSEP</td>
<td>hybrid-pol Shannon entropy</td>
<td></td>
<td>[67]</td>
</tr>
</tbody>
</table>

### 4. Methods

It is important to reduce the parameter set and determine which parameters are optimal for sea ice classification [64]. If the spatial dimensions of the features are too small, it may not be possible to identify the characteristics of sea ice, which makes it difficult to achieve the desired classification results. Conversely, using a considerable number of parameters will lead to data redundancy, which not only reduces the classification accuracy but also increases the computational cost [72]. Therefore, it is important to select the optimal features, from the high-dimensional feature space. Here, we introduce the Separability Index (SI) method, which is widely used for feature extraction [72–74].

#### 4.1. Separability Index (SI)

SI is a binary classification feature selection method that can be used to assess the separability of object features and select the appropriate features for target identification and classification [73,74]. The SI method between a pair of class-specific separations is defined as the following Equation (1):

$$\text{SI} = \frac{|\mu_1 - \mu_2|}{\sigma_1 + \sigma_2}$$  \hspace{1cm} (1)

where $\mu_1$, $\sigma_1$, and $\mu_2$, $\sigma_2$ denote the mean and the standard deviation of two categories of a feature, respectively. In general, the features with a SI value between 0.8 and 1.5 are possible...
candidates to be used in classification. When the SI value of a feature is over 2, it indicates that this feature can be completely separated from all categories [72,75]. The method has been widely used as an effective feature selection methodology in crop identification and oil spill detection [74,76], and it is logical to expand this approach for the determination of ideal parameters for classifying sea ice.

The SI value is calculated for any two combinations of the following four types of sea ice: OW and NI, OW and YI, OW and FYI, NI and YI, NI and FYI, and YI, and FYI. Parameters where SI is greater than 0.8 in any combination was utilized as valid classification features for sea ice.

4.2. Parameters Selection

For the selection of class-specific data, we determined that the SAR image S5 (shown in Figures 1 and 3) provided the best match to the reference optical data from Landsat-8, that also included a diversity of ice types (see Figure 4). The match between S5 and the optical data was ideal since the variance of wind speed in the S5 coverage area was the least for all SAR images, as shown in Figure 1b. Sample areas of sea ice types in S5 were determined by matching optical data (Figure 5).

![Figure 5. Different types of sea ice, with the Landsat-8 image on the left and the S5 SAR image of the corresponding region on the right. Red regions represent OW, blue regions are NI, orange regions are YI, and green regions are FYI.](image_url)

The distribution areas and coverage of the four ice types in Figure 5 are dramatically different. Among the four ice types, FYI covered the most area, while OW covered the least. Initially, we sampled 255,379 pixels of FYI and 3056 pixels of OW. To ensure the same size of each type of ice and OW sample, pixel counts of the three types of sea ice and OW were sub-sampled to 3000 pixels using an isometric sampling method. Based on the sub-sampled pixels of each type, the SI between each ice type for the 70 parameters was calculated. The results are shown in Figure 6.

![Figure 6.](image_url)

The SI values for parameters between six combinations of two different types of sea ice and OW are shown in Figure 6. The parameters’ names are on the left of each row, and the SI values between ice type pairs are presented from left to right, in each column: OW and NI; OW and YI; OW and FYI; NI and YI; NI and FYI; and YI and FYI, respectively. In each row, if the SI is greater than 0.8, then the parameters represented by that row can distinguish ice-type pairs in the corresponding column; but if all SI values are greater than or equal to 0.8 (thereby highlighted in yellow), then this parameter satisfies the fine sea
ice classification. We analyzed the total number of parameters with SI values above the 0.8 threshold between the ice-type pairs; results are shown in Table 2. The first column in Table 2 shows the ice type pairs; the second and third columns contain the possible quad-pol and hybrid-pol parameters, respectively; and the final column displays the number of parameters and the average SI value for the parameter group. It is found that 52 parameters can be used to distinguish YI from OW, but only 25 can be used to distinguish NI from FYI. These results can aid in the resolution of pair-wise OW and different types of sea ice classification problems.

<table>
<thead>
<tr>
<th>Type Pair</th>
<th>Hybrid-Pol Parameters</th>
<th>Total Parameters</th>
<th>SI Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>OW, NI</td>
<td>15</td>
<td>15</td>
<td>2399</td>
</tr>
<tr>
<td>OW, YI</td>
<td>16</td>
<td>16</td>
<td>2401</td>
</tr>
<tr>
<td>OW, FYI</td>
<td>15</td>
<td>15</td>
<td>2402</td>
</tr>
<tr>
<td>NI, YI</td>
<td>15</td>
<td>15</td>
<td>2398</td>
</tr>
<tr>
<td>NI, FYI</td>
<td>15</td>
<td>15</td>
<td>2398</td>
</tr>
<tr>
<td>YI, FYI</td>
<td>15</td>
<td>15</td>
<td>2398</td>
</tr>
</tbody>
</table>

From Figure 6, we find that 19 parameters satisfy the SI threshold requirement of 0.8, including the quad-pol parameters ($\sigma_{hh}^0$, $T_{11}$, $\lambda$, $P_F$, $Y_s$, $K_s$, SE, Span) and hybrid-pol parameters ($g_0$, $g_3$, $\sigma_{hh}^0$, $\sigma_{hh}^3$, $\sigma_{hh}^0$, $\rho_{hh-rr}$, $\alpha_s$, $\alpha_d$, $\psi_d$, $\psi_d$, CPSEi). All of these parameters are able to provide fine sea ice classification. The distributions of the 19 parameters that satisfy the SI threshold requirement of 0.8 are shown as boxplots, according to the four different types: OW and NI; OW and FYI; NI and YI; NI and FYI; and YI and FYI, with the quad-pol parameters on the left and hybrid-pol parameters on the right. The names of parameters are on the left side of the table. The color indicates the value of SI, and all values greater than or equal to 0.8 are highlighted in yellow.

![Figure 6: SI value for parameters between six combinations of two different types: OW and NI; OW and YI; OW and FYI; NI and YI; NI and FYI; and YI and FYI, with the quad-pol parameters on the left and hybrid-pol parameters on the right. The names of parameters are on the left side of the table. The color indicates the value of SI, and all values greater than or equal to 0.8 are highlighted in yellow.](image-url)
Table 2. Summary of feasible parameters for different types of sea ice classification.

<table>
<thead>
<tr>
<th>Type Pair</th>
<th>Quad-Pol Parameters</th>
<th>Hybrid-Pol Parameters</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>OW, NI</td>
<td>$c_{hh}^0, c_{vv}^0, R_{hh-hv}, R_{hv-vv}, R_{depol}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, R_{rr-rt}, R_{rt-rr}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>36 SI average is 1.10</td>
</tr>
<tr>
<td>OW, YI</td>
<td>$c_{hh}^0, c_{vv}^0, c_{hv}^0, R_{hh-hv}, R_{hv-vv}, R_{depol}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, R_{rr-rt}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>52 SI average is 2.04</td>
</tr>
<tr>
<td>OW, FYI</td>
<td>$c_{hh}^0, c_{vv}^0, c_{hv}^0, R_{hh-hv}, R_{hv-vv}, R_{depol}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, R_{rr-rt}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>45 SI average is 1.70</td>
</tr>
<tr>
<td>NI, YI</td>
<td>$c_{hh}^0, c_{vv}^0, c_{hv}^0, R_{hh-hv}, R_{hv-vv}, R_{depol}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>46 SI average is 1.48</td>
</tr>
<tr>
<td>NI, FYI</td>
<td>$c_{hh}^0, c_{vv}^0, c_{hv}^0, R_{hh-hv}, R_{hv-vv}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>25 SI average is 0.91</td>
</tr>
<tr>
<td>YI, FYI</td>
<td>$c_{hh}^0, c_{vv}^0, c_{hv}^0, R_{hh-hv}, R_{hv-vv}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>34 SI average is 1.08</td>
</tr>
<tr>
<td>(OW, NI, YI, FYI)</td>
<td>$c_{hh}^0, c_{vv}^0, c_{hv}^0, R_{hh-hv}, R_{hv-vv}, R_{depol}$</td>
<td>$g_{00}, g_{32}, c_{rr}^0, c_{ll}^0, R_{rr-ll}, m, \mu, \sigma, \rho, \delta, \alpha, \psi, \psi_d$, CPSEi</td>
<td>19 SI average is 1.53</td>
</tr>
</tbody>
</table>

Since applying all 19 parameters to the classifier will greatly increase the complexity, further reduction is required due to redundancy. To achieve this, we analyzed the ways in which these parameters were obtained and the corresponding scattering mechanism. The quad-pol parameters can be divided into three categories [47, 64]:

1. Co-pol intensity parameters ($c_{hh}^0, T_{11}$).
2. Total power parameters ($\lambda$, SE, Span).
3. Parameters that indicate surface scattering ($F_s$, $Y_s$, $K_s$).

The hybrid-pol parameters can be divided into four categories:

4. Circle-polarized intensity parameters ($g_{32}, c_{rr}^0, c_{ll}^0, c_{rh}^0$) whose dominant scattering is surface scattering.
5. Total power parameters ($g_{00}$, CPSE), which are similar to the corresponding quad-pol,
6. Depolarization parameters ($\rho_{\text{hh-rr}}$) due to multi-scattering
7. Parameters representing the orientation and polarization angle ($\alpha_s, \alpha_d, \psi_s, \psi_d$).

From each of the seven categories, the parameter with the maximum SI sum for ice type, for each pair-wise combination was selected (Figure 8).

According to Figure 8, the parameters with the highest SI sums are $c_{hh}^0$, SE, $K_s$, $c_{rh}^0$, CPSE, $\rho_{\text{hh-rr}}$, and $\alpha_s$ among the above seven groups. It is worth noting that several available satellites (RISAT-1, ALOS-2, and RCM) carry CP SAR with the right-circular transmit and coherent linear polarization receive work modes. However, when we compared the results of $c_{rl}^0$ and $c_{rh}^0$, we found that the SI sum difference between $c_{rl}^0$ and $c_{rh}^0$ was only 0.07. Therefore, in order to make a realistic implementation for CP SAR, we replaced $c_{rl}^0$ with $c_{rh}^0$ as the representative parameter. In order to compare the advantages of various parameters based on the work modes of the SAR system in the classification, we employed three distinct groups of parameters to classify sea ice. Group (1) used only quad-pol parameters ($c_{hh}^0$, SE, $K_s$); Group (2) used only hybrid-pol parameters ($c_{rh}^0$, CPSE, $\rho_{\text{hh-rr}}$, $\alpha_s$); Group (3) used both quad-pol and hybrid-pol parameters ($c_{hh}^0$, SE, $K_s$, $c_{rh}^0$, CPSE, $\rho_{\text{hh-rr}}$, $\alpha_s$).
Figure 7. Distributions of the 19 parameters. The legend is shown at the lower right. Black dotted lines represent the range, the pool blue asterisk * represents the mean value, and the red lines denote the median. The top of the blue rectangle represents the upper quartile, and the bottom represents the lower quartile.

4.3. Random Forest Classifier

The RFC is an ensemble classifier that generates multiple decision trees using randomly selected subsets of training samples and variables [77]. There are several advantages to using RFC in remote sensing applications, including the handling of a large number of variables, identifying missing data and outliers, providing unbiased estimates of out-of-bag errors, optimizing feature space using variable importance functions, and being relatively robust to noise and outliers. Thus, RFC has been previously used in sea ice extent monitoring and sea ice classification [78–80].
Figure 8. Sum of the SI value for ice type, for each pair-wise combination, for the 19 parameters shown by the groupings described in the text. The y-axis represents the SI sum, and the different colors denote the SI values for the separations of different ice types.

In the study, we compared the accuracy of five machine learning methods using the matched data from Section 4.2, with 3000 pixels per subclass, and a 9:1 ratio between the training set and the test set. These corresponding results show that both the training and testing accuracy of the RFC methodology are higher than those of the other four methods. The training and testing accuracy of the five methods are shown in Table 3. Zhai et al. [78] compared the five classifiers in sea ice extent monitoring using scatterometer observations, and found that RFC exhibits the highest overall accuracy. Our results are consistent with those of [78]. Therefore, it is reasonable to select RFC as the sea ice classification algorithm.

Table 3. Training and test accuracy of sea ice classification using different machine learning classifiers.

<table>
<thead>
<tr>
<th>The Classifier</th>
<th>Training Accuracy</th>
<th>Test Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compared Logistic Regression</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Naive Bayes</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td>Random Forest Classifier</td>
<td>0.99</td>
<td>0.92</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>0.84</td>
<td>0.83</td>
</tr>
</tbody>
</table>

We used the RFC function in the sklearn package with Python 3.8 and adjusted the n_estimators and random_state parameters while keeping other parameters at their default values. The n_estimators parameter determines the number of trees in the forest, i.e., the number of base estimators. The random_state parameter controls the mode of forest generation, with base classifiers being mutually independent. Generally, the larger the number of the n_estimators and random_state, the better the model’s performance. However, the accuracy of a random forest methodology may stop improving or even fluctuate when n_estimators and random_state reach a certain level. In our study, we found that the training accuracy and the test accuracy can reach 100% and 92%, respectively, with n_estimators = 64 and random_state = 45, for the training and testing sets. Therefore, in this study, the RandomForestClassifier function was used with n_estimators = 64 and random_state = 45, while keeping other parameters at their default values.

For parameter selection, 3000 pixels were selected for each type of sea ice and OW, which were used to calculate the seven parameters ($\sigma_h^{0}$, SE, $\sigma_h^{0}$, CPSE, $\rho_{rh-rv}$, $\alpha_s$). These parameters were grouped into three categories: using only quad-pol parameters ($\sigma_h^{0}$, SE, $\sigma_h^{0}$, $\rho_{rh-rv}$, $\alpha_s$), using only hybrid-pol parameters ($\sigma_h^{0}$, CPSE, $\rho_{rh-rv}$, $\alpha_s$), and using both quad-pol and hybrid-pol parameters ($\sigma_h^{0}$, SE, $\sigma_h^{0}$, CPSE, $\rho_{rh-rv}$, $\alpha_s$). RFCs were trained
for each of the work modes of the SAR systems: quad-pol, hybrid-pol, and both quad-pol and hybrid-pol. Three well-trained models were obtained, corresponding to three different work modes of SAR. The three trained models were used for classifying the six SAR images S1–S6 in the subsequent analysis.

5. Results

5.1. Classification Results

Three classification results (using only the quad-pol parameters, using only hybrid-pol parameters, and using both the quad-pol and hybrid-pol parameters) were obtained for each of the six SAR images S1–S6 (Figure 9), where the OW is denoted in dark blue, the NI in green, the YI in light blue, and the FYI in light brown. From these results, we qualitatively observed that the distribution of leads was clearly visible in each SAR image, regardless of the input parameters. However, for these three ice types, the classification results from the three groups of input parameters showed significant differences. Comparing the results among the columns, the largest fraction of FYI appears in the first column with the quad-pol input parameters. Similarly, the largest fractions of NI and YI appear in the second column with hybrid-pol input parameters. Finally, results using both quad-pol and hybrid-pol parameters are between the other two. The red rectangles in (S1–S6) were used to verify the classification accuracy in the next section.

5.2. Verification

The collocated SAR images and optical data were used to verify the classification results, specifically in the regions of the five SAR images that are independent of the regions used to derive input data for the classification, namely S1, S2, S3, S4, and S6. These regions are marked by red rectangles in Figure 9. The specific operational process was as follows: for each of the five images, a region was selected, and using the optical image as a reference, the Span image of the selected region was manually classified. The classification accuracy for each type of sea ice was calculated as the proportion of correctly classified pixels to all pixels. In this study, the manual classification was carefully carried out, under the supervision of an expert (Lijian Shi, National Satellite Ocean Application Service, personal communication). The results are shown in Figure 10.

The manual classification in Figure 10 is generally smoother, with fewer isolated regions. The manually classified SAR images in S1, S2, and S3 images do not contain FYI, and S6 only includes YI and FYI. The number of pixels that were validated using the manual classification is shown in Table 4, and Table 5 shows the classification accuracy of the three categories of input parameters for the five validated SAR images.

### Table 4. Numbers of validated pixels (Each pixel is eight meters in range and azimuth directions).

<table>
<thead>
<tr>
<th>SAR Images</th>
<th>OW</th>
<th>NI</th>
<th>YI</th>
<th>FYI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>3198</td>
<td>9914</td>
<td>44,688</td>
<td>0</td>
<td>57,800</td>
</tr>
<tr>
<td>S2</td>
<td>518</td>
<td>7545</td>
<td>62,437</td>
<td>0</td>
<td>70,500</td>
</tr>
<tr>
<td>S3</td>
<td>7946</td>
<td>19,180</td>
<td>83,208</td>
<td>0</td>
<td>110,334</td>
</tr>
<tr>
<td>S4</td>
<td>3410</td>
<td>8160</td>
<td>6976</td>
<td>28,842</td>
<td>47,388</td>
</tr>
<tr>
<td>S6</td>
<td>0</td>
<td>0</td>
<td>32,086</td>
<td>113,514</td>
<td>145,600</td>
</tr>
<tr>
<td>Total</td>
<td>15,072</td>
<td>44,799</td>
<td>229,395</td>
<td>142,356</td>
<td>431,622</td>
</tr>
</tbody>
</table>

### Table 5. Sea ice and open water classification accuracy by different SAR parameter sets.

<table>
<thead>
<tr>
<th>Parameter Set</th>
<th>OW</th>
<th>NI</th>
<th>YI</th>
<th>FYI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quad_pol</td>
<td>95.29%</td>
<td>81.65%</td>
<td>98.76%</td>
<td>77.61%</td>
<td>88.36%</td>
</tr>
<tr>
<td>Hybrid_pol</td>
<td>95.58%</td>
<td>84.29%</td>
<td>98.39%</td>
<td>85.95%</td>
<td>92.30%</td>
</tr>
<tr>
<td>Quad + Hybrid_pol</td>
<td>91.35%</td>
<td>96.88%</td>
<td>98.89%</td>
<td>83.35%</td>
<td>91.53%</td>
</tr>
</tbody>
</table>
Figure 9. Results for sea ice classification for six SAR images using different SAR imaging modes. The first, second, and third columns present the results from quad-pol, hybrid-pol, and the combination of both imaging modes, respectively. The colors of dark blue, green, light blue, and light brown denote OW, NI, YI, and FYI. Red rectangles (S1, S2, S3, S4, and S6) were used to verify the classification accuracy in the next section.
5.2. Verification

The collocated SAR images and optical data were used to verify the classification results. The accuracy of the classification results was calculated using the manual classification results as the reference. The results are shown in Figure 10. The interpretation results for verification. Panels in the left column are Landsat-8 images, taken as reference, i.e., truth. Panels in the second column are sub-section images of SAR, which match well with the optical images. The third column contains the manual classification results. The last three columns show the classification results using the quad-pol parameters, the hybrid-pol parameters, and finally, both the quad-pol and the hybrid-pol parameters, respectively.

![Figure 10](image_url)

The total accuracy of fine sea ice classification using the three parameter sets reached more than 88%. The overall accuracy using quad-pol parameters was 88.36%, while the classification accuracy for OW, NI, YI, and FYI was 95.29%, 81.65%, 98.76%, and 77.61%, respectively. When hybrid-pol parameters were used, the overall classification accuracy rose to 92.30%, and the classification accuracy for OW, NI, YI, and FYI was 95.58%, 84.29%, 98.39%, and 85.95%, respectively. Using both quad-pol and hybrid-pol parameters, the overall classification accuracy was 91.53%, with OW, NI, YI, and FYI classification accuracy of 91.35%, 96.88%, 98.89%, and 83.35%, respectively.

It is evident that the hybrid-pol parameter set resulted in the highest overall accuracy, and higher classification accuracy for OW, NI, and FYI than the quad-pol parameter set. On the other hand, the quad-pol parameter set had better performance for YI classification. With both quad-pol and hybrid-pol parameters, the overall accuracy was greater than with only quad-pol parameters and less than with only hybrid-pol parameters. The classification accuracy of NI and YI was greater than that of hybrid-pol, whereas the classification accuracy of FYI was greater than that of quad-pol and less than that of hybrid-pol. The classification accuracy of OW was less than that of quad-pol and hybrid-pol.

In this study, YI classification accuracy was significantly higher than the other ice types. This is likely due to the consistently unique and strong response of the examined parameters for YI conditions, relative to the other ice types and open water, as shown in Figure 2. We expect that a rough surface, combined with a high dielectric constant due to high salinity, which is characteristic of YI, produced the observed behavior at the C-band frequency of GF-3 [81].

Overall, if the sea ice classification objective is primarily to focus on the OW and YI, all three sets of parameters yielded good results. If the numerous forms of FYI are to be sorted, the hybrid-pol parameters can be considered as providing improved classification accuracy.
compared with quad-pol. Moreover, the simultaneous use of quad-pol parameters and hybrid-pol parameters can significantly improve the accuracy of identification of NI, which is very important for the study of sea ice growth.

6. Discussion

In this study, we compared the quad-pol and hybrid-pol parameters for sea ice classification. The results show that using only the hybrid-pol parameter set can achieve a higher classification accuracy compared with the other two parameter sets. This is an interesting issue. We speculate that the reason for this result might be that the selected hybrid-pol parameters are more sensitive to thin sea ice than the selected quad-pol parameters. Raney [17,82] has presented the advantages of the hybrid-pol SAR system, based on measurement potential, ease of calibration, and favorable implementation. He also emphasized that any non-circular illuminating polarization would impose preferential selectivity on the backscattered polarizations in response to the relative alignment between the principal axis of the incoming elliptically polarized field and the angular orientation of scene constituents. Charbonneau et al. [65] mentioned that the hybrid-pol outperforms the quad-pol in discriminating sea ice types and open water. Our conclusion is coincident with the work of Charbonneau [65]. We also checked this conclusion by using different classification algorithms: Support Vector Machine and RFC. Both results are in agreement. The results obtained by using the Support Vector Machine classifier are not shown here. The reasons why the sea ice classification results of hybrid-pol parameters are better than those of quad-pol parameters is an issue that clearly needs further study. In a future investigation, we plan to further validate the effectiveness of the selected parameters in sea ice classification using real hybrid-pol SAR data, such as that collected by RCM.

It should be noted that the sea ice examined in this study was on, or near the beginning, of its ice-melt phase, in the late winter/early spring conditions in the Okhotsk Sea, on 28 February 2020. This potentially complicates the sea ice classification, in particular, if the above-freezing conditions lead to the presence of liquid water in the snow and on the sea ice. A further consideration may be imposed by the time difference in imaging between the Landsat-8 and the SAR images (~19 h). Therefore, some physical differences in sea ice and open water areas, and some possible position differences due to drift, are inherently present between SAR and optical data, thereby providing possible errors in the classification.

Another consideration is related to the SAR incidence angle. In this study, the swath width of the quad-pol SAR data was narrow, and there was essentially no influence of the incidence angle on the SAR parameters. Since the hybrid-pol data used in this study was simulated from the quad-pol SAR data, the hybrid-pol parameters also only had a small variation in the range of incidence angles. However, real hybrid-pol SAR images have a wide swath width and a much larger range of incidence angles [23]. Thus, variation of the incidence angle across these latter SAR images may have an implication on the classification results for the hybrid-pol parameters and should be considered in the future.

7. Conclusions

In this study, a method is proposed for fine sea ice classification according to different polarimetric SAR modes, based on quad-pol GF-3 SAR acquisitions of ice in the Okhotsk Sea. The method can clearly discriminate OW, NI, YI, and FYI with an overall accuracy of over 90%. The precise classification of OW, NI, and YI can reach 95%, 96%, and 98%, respectively. The improvement in sea ice classification is a key factor in increasing the accuracy of sea ice monitoring and charting, with impacts on weather and climate forecasts in polar regions.

We selected the optimal quad-pol and hybrid-pol parameters from 70 possible parameters for sea ice classification using SAR data. The availability of these parameters in sea ice classification was evaluated using the Separability Index (SI) between different sea ice classes and open water. Adopting a general threshold of 0.8 for the SI, there are 19 parameters that are effective features for classifying various ice types. These are quad-pol
parameters: $\sigma_{hh}^0, T_{11}, \lambda, F_s, Y_s, K_s, SE_i$. Span and hybrid-pol parameters: $g_0, g_3, \sigma_{hh}^0, \sigma_{hh}^0, \rho_{10-20}, \alpha_s, \beta_d, \psi, \Phi, \rho_{hh}, CPSE_i$.

In this study, we compared the quad-pol and hybrid-pol parameters for sea ice classification. Higher classification accuracy (92.3%) was achieved as a result of using hybrid-pol parameters, compared with quad-pol parameters (88.36%). When using both parameter sets simultaneously, thin ice (NI and YI) can be classified more effectively, with an accuracy of over 96%. Based on these findings, we recommend combining the hybrid-pol and quad-pol parameters for thin sea ice classification, as it leads to superior accuracy. The reasons why the sea ice classification results for the hybrid-pol parameters outperform those of the quad-pol parameters need further study.

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

Funding: This work was supported by the National Key R&D Program of China (No. 2022YFC3104900/2022YFC3104904), the Fundamental Research Funds for the Central Universities, the Canadian Ocean Frontier Institute, the Canadian Space Agency program for SWOT satellite, the Government of Canada Competitive Science Research Fund, and a Natural Sciences and Engineering Research Council of Canada Discovery Grants Program project RGPIN-2022-05217 to Scharien.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the ECMWF for providing ERA5 reanalysis data, and the National Satellite Ocean Application Service for providing GF-3 data. The U.S. Geological Survey (USGS), NASA National Snow and Ice Data Center and Goddard Space Flight Center provided Landsat-8 images, and the University of Bremen provided sea ice concentration data. The authors would like to thank Lijian Shi (National Satellite Ocean Application Service, Beijing 100081, China) for sea ice manual classification.

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

References


73. Moreno Navas, J.; Telfer, T.C.; Ross, L.G. Separability Indexes and Accuracy of Neuro-Fuzzy Classification in Geographic Information Systems for Assessment of Coastal Environmental Vulnerability. *Ecol. Inform.* 2012, 12, 43–49. [CrossRef]

74. Tao, J.; Zhang, X.; Wu, Q.; Wang, Y. Mapping Winter Rapeseed in South China Using Sentinel-2 Data Based on a Novel Separability Index. *J. Integr. Agric.* 2022, S2095311922002581. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.