Automatic Selection of Relevant Attributes for Multi-Sensor Remote Sensing Analysis: A Case Study on Sea Ice Classification

Eduard Khachatrian, *Student Member, IEEE*, Saloua Chlaily, *Member, IEEE*, Torbjørn Eltoft, *Member, IEEE*, Wolfgang Dierking, Frode Dinessen, and Andrea Marinoni, *Senior Member, IEEE*

***Abstract*—It is of considerable benefit to combine information obtained from different satellite sensors to achieve advanced and improved characterization of sea ice conditions. However, it is also true that not all the information is relevant. It may be redundant, corrupted, or unnecessary for the given task, hence decreasing the performance of the algorithms. Therefore, it is crucial to select an optimal set of image attributes which provides the relevant information content to enhance the efficiency and accuracy of the image interpretation and retrieval of geophysical parameters. Comprehensive studies have been focused on the analysis of relevant features for sea ice analysis obtained from different sensors, especially synthetic aperture radar. However, the outcomes of these studies are mostly data and application-dependent and can, therefore, rarely be generalized. In this article, we employ a feature selection method based on graph Laplacians, which is fully automatic and easy to implement. The proposed approach assesses relevant information on a global and local level using two metrics and selects relevant features for different regions of an image according to their physical characteristics and observation conditions. In the recent study, we investigate the effectiveness of this approach for sea ice classification, using different multi-sensor data combinations. Experiments show the advantage of applying multi-sensor data sets and demonstrate that the attributes selected by our method result in high classification accuracies. We demonstrate that our approach automatically considers varying technical, sensor-specific, environmental, and sea ice conditions by employing flexible and adaptive feature selection method as a pre-processing step.**

***Index Terms*—Graph Laplacians, multi-sensor remote sensing, sea ice, unsupervised information selection.**

Manuscript received February 22, 2021; revised May 5, 2021; accepted July 17, 2021. Date of publication July 26, 2021; date of current version September 20, 2021. This work was supported in part by the Centre for Integrated Remote Sensing and Forecasting for Arctic Operations (CIRFA), Research Council of Norway, under Grant 237906, in part by the Automatic Multi-sensor Remote Sensing for Sea Ice Characterization (AMUSIC) "Polhavet" flagship project 2020, and in part by the European Union's Horizon 2020 Research and Innovation Programme ExtremeEarth Project under Grant 825258. *(Corresponding author: Eduard Khachatrian.)*

Eduard Khachatrian, Saloua Chlaily, Torbjørn Eltoft, and Andrea Marinoni are with the Department of Physics and Technology, UiT The Arctic University of Norway, 9037 Tromsø, Norway (e-mail: eduard.khachatrian@uit.no; saloua.chlaily@uit.no; torbjorn.eltoft@uit.no; andrea.marinoni@uit.no).

Wolfgang Dierking is with the Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, 27570 Bremerhaven, Germany, and also with the Department of Physics and Technology, UiT The Arctic University of Norway, 9037 Tromsø, Norway (e-mail: Wolfgang.Dierking@awi.de).

Frode Dinessen is with the Division for Remote Sensing and Data Management at MET-Norway, 9293 Tromsø, Norway (e-mail: froded@met.no).

Digital Object Identifier 10.1109/JSTARS.2021.3099398

I. Introduction

In THE last decades, sea ice research has become a focus of Earth observation, especially in the Arctic region where sea ice extent and volume are declining rapidly [1]. Sea ice plays an essential role in the polar ecosystem [2]. It is one of the main indicators of global climate change, and it also considerably affects the indigenous population's lifestyle and welfare [2]. Moreover, sea ice poses a great challenge to navigation in the polar seas; therefore, the knowledge about its type, concentration, thickness, deformation, and extent is extremely important for various activities, such as marine transportation and offshore operations, and for stakeholders from the oil and gas industry, fisheries, and tourism, among others.

Sea ice observations have a long history of more than a century. They were carried out visually from coastal stations, ships, and aircraft [3], while they were spatially and temporally limited. Regular sea ice monitoring over larger regions became possible in the late 1970s using image data from satellites [3]. Since then, the technologies for acquiring and analyzing sea ice data have been considerably improved and extended.

Optical sensors provide information about sea ice in visible and infrared channels. The main advantage of this kind of information is an easier visual interpretability. However, optical sensors are limited to cloud-free and favorable light conditions, which causes a significant problem since dense cloud covers and long periods of darkness prevail in the polar regions for several months of the year. Therefore, spaceborne synthetic aperture radar (SAR), which combines high spatial resolution and independence of cloud and light conditions, is the main source from which detailed maps of sea ice conditions are produced. However, the automatic interpretation of remote sensing data, especially SAR data, is challenging and strongly relies on expert's knowledge. Passive microwave radiometers are another type of sensor that can be used for sea ice observations. However, in comparison to the aforementioned techniques, it has a significantly coarser spatial resolution and is, therefore, preferably used for global or large-scale observations [3]. The increasing amount of available satellite data together with more and more activities in sea ice covered waters requires a greater effort for supplementing the production of ice charts by employing fully automated methods of information selection and image analysis [3], [4].

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

Multi-sensor remote sensing refers to the use of different sensors usually operating at different frequencies and/or spatial resolutions and coverage and with different time intervals between data acquisitions over a given region. Therefore, diverse remote sensing systems grasp various properties of sea ice by using different physical principles. Thus, combining the information from multiple sensors allows better characterization of sea ice [5]. Nevertheless, although different data sources provide complementary information, they can potentially also include redundant, corrupted, or simply unnecessary information for a given task. Accordingly, combining these data can significantly deteriorate the performance by decreasing the accuracy and increasing the computation time and structural complexity of the algorithm. Thus, the selection of relevant information from the original data set is an essential step of multi-source data fusion that enables reliable and efficient performance [6], [7].

We note that the term "feature" is commonly used in fields such as classification methodologies, pattern recognition, and texture analysis. In our study, however, we use the notation "attribute" from information theory which refers to directly measured quantities such as, e.g., radar intensities at different frequencies and polarizations, optical reflectance, and parameters such as textural or polarimetric features. The pertinence of different attributes for sea ice characterization has been investigated for both SAR [4], [7]–[14] and optical imagery [15], [16]. In some studies, e.g., the entropy was found to be well suited for separating sea ice types [13]. In other studies, however, the same parameter was found to be less relevant for sea ice classification [8], [11] and less useful for detection of leads in the ice [9], [12]. In fact, it has been shown that the relevance of single attributes varies with the observation conditions such as incidence angle and season [7], [9]. Accordingly, the results of such analyses cannot be generalized and emphasize the need for an automatic attribute selection algorithm.

The selection of relevant information can be achieved using dimensionality reduction methods that are generally divided into two main categories: attribute extraction and attribute selection [17]. Extraction methods generate a prominent set of attributes by projecting the original set into a lower-dimensional space. Among such methods, we can highlight the principal component analysis (PCA) [18]. Selection algorithms search for the most relevant elements, according to the given criteria, within the original set, without applying any transformation. One example of this category is the forward selection (FS) [19] that determines a subset of attributes incrementally, by adding at each step the attribute that leads to the largest improvement of the information content. The transformation applied by the attribute extraction approaches improves the separation of different data classes but at the expense of reducing the physical interpretability, which may be a disadvantage for remote sensing data analyses. Unlike the attribute extraction approach, attribute selection preserves the data's physical interpretability by working in the same space. The approaches mentioned above can be classified as supervised [20], if they require training data, or unsupervised [21]. In the case of remote sensing data, especially when dealing with complex scenes or considering modalities (i.e., various sensors characterized by different acquisition techniques) that are difficult to interpret such as sea ice SAR images, providing accurate labels is challenging even with the assistance of an expert. Therefore, the unsupervised attribute selection method is the right choice in the case of multi-sensor remote sensing sea ice analysis.

In this article, we use a recently developed graph-based method (referred to as GKMI) [22] that relies on information theory metrics to capture the most relevant attributes for different sea ice classes. The acronym GKMI recalls the Gaussian kernel (GK) and the mutual information (MI) used as similarity measures within this approach. GKMI combines the advantages of both attribute extraction and selection approaches since it generates a new representation of the attributes that can be reassigned to the original ones. As such, it increases their separability (even if they are non-linearly separable) while preserving their physical interpretability. Moreover, it selects relevant attributes for separate zones of an image that might belong to different ice classes and/or are measured under different conditions (e.g., different radar incidence angles, varying sun elevation angles). Accordingly, it enhances the classifier's ability to separate different sea ice classes even in challenging scenarios.

The considered method was paired with a supervised classification approach [23]. Regions that appear homogeneous in a single image and can also be identified in images acquired from different sensors are labeled as a certain ice type with the help of sea ice experts. In our study, we focus on data that were acquired under freezing conditions. Hence, the resulting attribute selections cannot be generalized for the melting season or freeze–melt cycles. However, the GKMI method can be employed in the same manner to other than freezing conditions.

The rest of this article is organized as follows. Section II describes the data sets used in this study. Section III provides details of the proposed architecture. Section IV presents an experimental validation of the proposed method. Finally, the discussion and conclusion are presented in Section V.

II. Data Sets

This section describes the different data sets, consisting of SAR and optical images that we used in our analysis. To increase the validity range of our conclusions, we consider several data sets of different spatial resolutions representing various sea ice conditions: a multi-frequency data set from an airborne SAR measurement campaign (ICESAR), two multi-sensor data sets combining Radarsat-2 and Landsat-8 imagery, and Sentinel-1 and Sentinel-2 imagery. The data sets were labeled based on direct field observations (ICESAR) or by analysts from operational ice services and sea ice experts from institutions contributing to this article.

Table I provides general information about the data sets, and Fig. 1 illustrates their geographical location. SAR images were acquired in linear polarization, and the notation "HV" means that a horizontally polarized wave was transmitted and the backscattered signal was received at vertical polarization.

A more detailed description of each data set is presented in the following subsections.

TABLE I
CHARACTERISTICS OF REMOTELY SENSED DATA SETS USED IN THIS STUDY. INTENSITIES ARE GIVEN AS BACKSCATTERING COEFFICIENT SIGMA-NOUGHT

Name	Date	Region	Measured Quantities	Pixel Size [m]	Coverage [km]
ICESAR	March 20, 2007	Fram Strait	L-band : HV, HH, VV Intensities C-band : VH, VV Intensities	1.5	5.3×1.5
Sentinel-1/Sentinel-2	April 08, 2018	North-East of Svalbard	SAR : HH and HV Intensities Optical : 13 bands from 0.443 µm - 2.19 µm	60	87.5×109.8
Radarsat-2/Landsat-8	April 25, 2018	North-East of Svalbard	SAR : HH and HV Intensities Optical : 11 bands from 0.433 µm - 12.5 µm	100	102.5×35.5

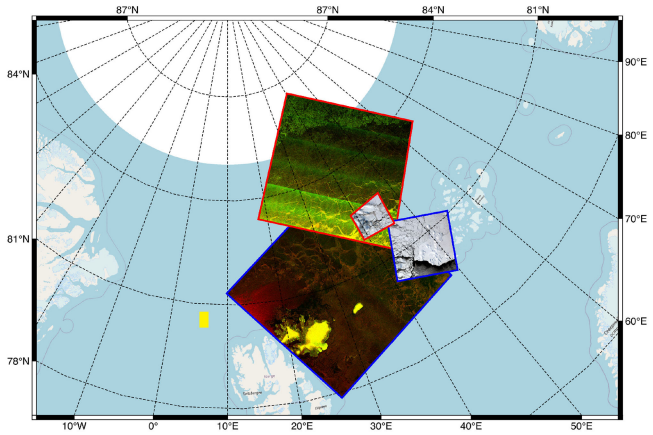


Fig. 1. Location of the data sets used in this study. Boundary colors refer to a specific data set: Sentinel-1/Sentinel-2 (Red), Radarsat-2/Landsat-8 (Blue). Given the small area of coverage of the ICESAR data set, an approximation of its location is given by a yellow square. It should be noted that only the overlapped area of the multi-sensor data sets, mainly Sentinel-1/Sentinel-2 and Radarsat-2/Landsat-8, are used.

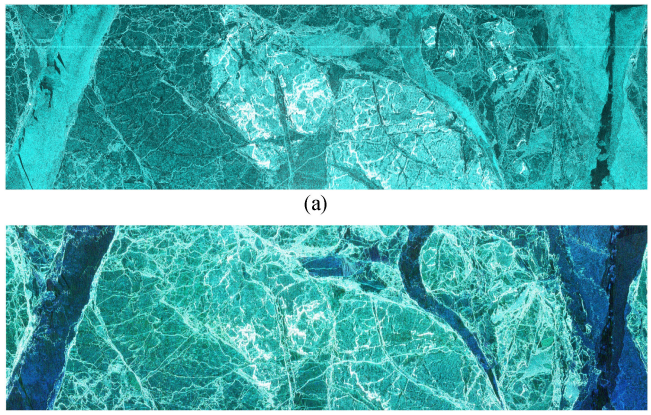


Fig. 2. False-color composite representation of ICESAR data set. (a) C-band (VH, VV, and VV as RGB). (b) L-band (HV, HH, and VV as RGB).

A. *ICESAR*

From the ICESAR measurement campaign, we used images acquired by the airborne SAR of the German Aerospace Center (DLR) and an optical scanner operated on an aircraft of the Alfred Wegener Institute (AWI). The SAR data set was recorded at C-band (dual-polarization, VH, and VV) and L-band (quad-polarization, HV, HH, and VV). The time difference between C-band and L-band measurements varies approximately between 10 and 30 min. A more detailed description can be found in [4] and [24]. The ICESAR data set includes six classes which are open water, nilas, grey ice, grey-white ice, level first-year ice (FYI) and deformed FYI according to the WMO sea ice nomenclature [25]. Fig. 2 shows the false-color composite of one scene at C- and L-bands.

B. *Sentinel-1/Sentinel-2*

The Sentinel-1/Sentinel-2 multi-sensor data set consists of SAR and optical data obtained from Sentinel-1 and Sentinel-2 satellites. The time difference between SAR and optical data was less than an hour. Sentinel-1 imagery was acquired in extra-wide (EW) swath mode at dual polarization (HH and HV), which is commonly used for sea ice monitoring. From Sentinel-2, we used 13 bands in the visible, near-infrared, and shortwave infrared part of the spectrum. Sentinel-1 data have a pixel size of 40 m, while the pixel size of Sentinel-2 varies from 10 to 60 m depending on the spectral band. Both Sentinel-1 and Sentinel-2 data sets were downsampled to the same pixel size of 60 m. The data set includes several classes, such as thick FYI, thin FYI, grey ice, grey-white ice and open water. Fig. 3 shows the SAR false-color composite and optical natural color composite images of the Sentinel-1/Sentinel-2 data set. In Fig. 3(b), clouds can be recognized, which is quite common for optical data.

C. *Radarsat-2/Landsat-8*

The Radarsat-2/Landsat-8 consists of optical and SAR data acquired from Landsat-8 and Radarsat-2, respectively. The time difference between SAR and optical data was less than an hour. The SAR images were acquired at HH and HV polarizations. From Landsat-8, we have images for 11 different spectral bands. The Radarsat-2 ScanSAR Wide A image product has a pixel size of 50 m, and the pixel size of Landsat-8 data varies in interval from 15 to 100 m. Both SAR and optical imagery were downsampled to the same pixel size (100 m) by the nearest neighbor resampling method and projected onto the same coordinate system. The data set includes several sea ice types (thick FYI, thin FYI, young ice, and nilas) and open water. Fig. 4 shows the SAR false-color composite image along with the optical natural color composite image of the Radarsat-2/Landsat-8 overlapped area. In contrast to the Sentinel-2 image of Fig. 3, the optical scene of the Radarsat-2/Landsat-8 data set is not affected by clouds [Fig. 4(b)].

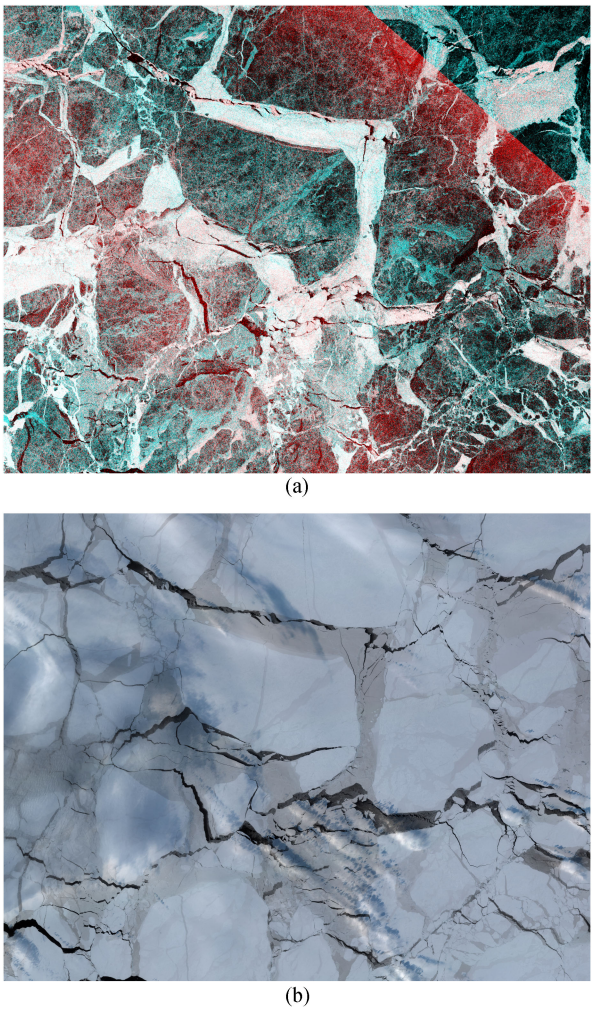


Fig. 3. Color representation of Sentinel-1/Sentinel-2 data set. (a) SAR false-color composite (HV, HH, and HH as RGB). (b) Optical natural color composite (RGB).

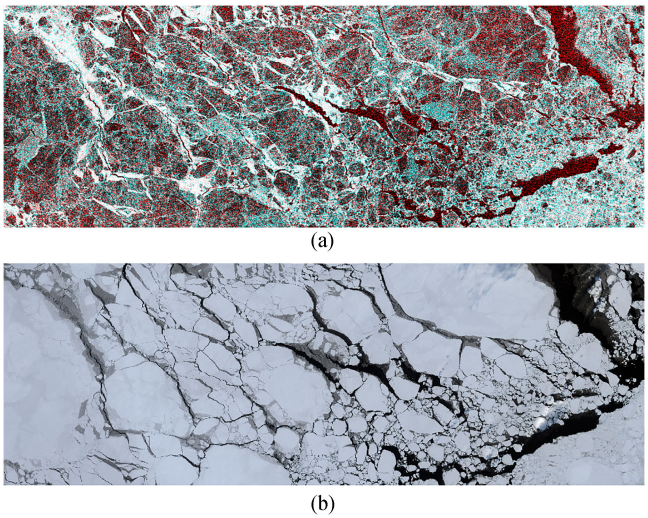


Fig. 4. Color representation of Radarsat-2/Landsat-8 data set. (a) SAR false-color composite (HV, HH, and HH as RGB). (b) Optical natural color composite (RGB).

III. Methods

Fig. 5 represents the flowchart of the proposed approach that consists of pre-processing (collocation and downsampling of the original data sets), the generation and combined extraction and selection of attributes, and, finally, the classification.

In the following sections, random scalars are denoted by lower case letters, e.g., z. Random vectors are designated by bold lower case letters, e.g., $\mathbf{z}$. Bold upper case letters refer to matrices, e.g., $\mathbf{A}$. $|\mathbf{A}|$ denotes the determinant of the matrix $\mathbf{A}$. $\text{diag}\{d_1, \ldots, d_N\}$ refers to a diagonal matrix whose diagonal elements are $d_1, \ldots, d_N$ starting from upper left. The $\text{ddiag}(\mathbf{A})$ operator set to zero the off-diagonal entries of $\mathbf{A}$.

A. Pre-processing

The images acquired by different sensors can have different characteristics, such as units of measurements, spatial resolution, image, and geographical coordinate systems. The first step of our analysis consists of making the data compatible by means of subsampling, alignment on the same coordinate system, collocating, and, finally, extracting the overlapping area.

We note that we did not use any additional filters or correction schemes usually applied to optical data, such as atmospheric correction, cloud masking or to SAR images, such as thermal noise removal, incidence angle compensation.

During the ICESAR campaign, the radar images at L- and C-bands were acquired with a time difference between 10 and 30 min, while for Sentinel-1/Sentinel-2 and Radarsat-2/Landsat-8, the time difference between multi-sensor data was less than 1 h. Nevertheless, after a detailed analysis of the images, especially focusing on individual ice formations and areas of open water, we can assume that there was no significant drift effect that needs to be considered.

B. Attributes Extraction

In addition to the original radar intensities or optical reflectances, we consider the textural attributes in this study. It is noteworthy that GKMI is independent of the data type; hence, it can also be applied to other attributes such as polarimetric features.

For each image layer (frequency/wavelength band and/or polarization), we extract the textural features using the gray-level co-occurrence matrix (GLCM) [26], [27]. We use the directional average for 0°, 45°, 90°, and 135° which is common practice to account for the possible rotation of sea ice floes, leads, or any other sea ice structure or roughness pattern on the ocean surface. Moreover, it reduces the number of GLCM matrices. We also considered the average of the distance/displacement parameter that we varied from 1 to 5 (corresponding to the half of the window's size that was set to 11×11). It is noteworthy that minor changes of window size (±5) do not significantly affect the classification accuracy. Table II illustrates the extracted features as well as their mathematical definitions.

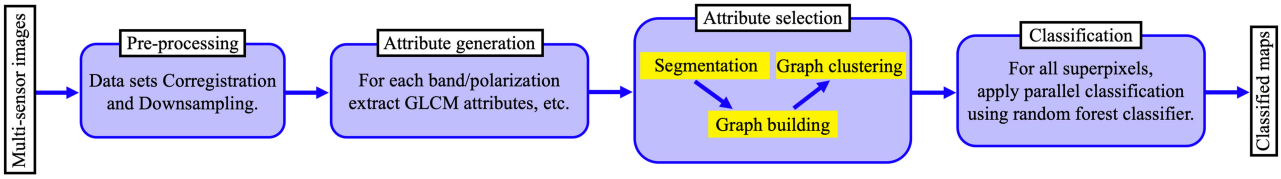


Fig. 5. Flowchart of the proposed approach.

TABLE II
MATHEMATICAL DEFINITION OF GLCM FEATURES

Texture Features	Definition	Texture Features	Definition
Angular Second Moment/ Energy (ASM)	$\sum_{i,j=0}^{Q-1} g_{i,j}^2$	**Maximum Correlation Coefficient (MCC)**	$\sqrt{\lambda}$
Correlation (CORR)	$\sum_{i,j=0}^{Q-1} g_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j}\right]$	**Variance (VAR)**	$\sum_{i,j=0}^{Q-1} g_{i,j}(i-\mu)^2$
Contrast (CST)	$\sum_{i,j=0}^{Q-1} g_{i,j} \lvert i-j\rvert^2)$	**Sum Average (SAVE)**	$\sum_{n=2}^{2Q} g_{x+y}(n)$
Sum Variance (SVAR)	$\sum_{n=2}^{2Q} \left(n - \sum_{n=2}^{2Q} g_{x+y}(n)^2\right) g_{x+y}(n)$	**Sum Entropy (SENT)**	$-\sum_{n=2}^{2Q} g_{x+y}(n) \log\left(g_{x+y}(n)\right)$
Entropy (ENT)	$\sum_{i,j=0}^{Q-1} g_{i,j} \log\left(g_{i,j}\right)$	**Difference Variance (DVAR)**	$\sum_{n=0}^{Q-1} \left(n - \mu_{x-y}\right)^2 g_{x-y}(n)$
Difference Entropy (DENT)	$-\sum_{n=0}^{Q-1} g_{x-y}(n) \log\left(g_{x-y}(n)\right)$	**Information Correlation (IC)**	$-\frac{\sum_{i,j}^{Q-1} g_{i,j} \log \frac{g_{i,j}}{g_x(i) g_y(j)}}{\max\left(H_x, H_y\right)}$
Inverse Different Moment/ Homogeneity (IDM)	$\sum_{i,j=0}^{Q-1} \frac{g_{i,j}}{1+(i-j)^2}$	**Coefficient of Variation (CV)**	$\frac{\sqrt{\sum_{i,j=0}^{Q-1} g_{i,j}(i-\mu)^2}}{\sum_{i,j=0}^{Q-1} i g_{i,j}}$

Note: $g_{i,j}$ denotes the elements of the GLCM matrix $\mathbf{G}$. Q is the number of gray levels used. $g_x = \sum_i g_{i,j}$, $g_y = \sum_j g_{i,j}$, $g_{x+y}(n) = \sum_{i+j=n} g_{i,j}$, $g_{x-y}(n) = \sum_{|i-j|=n} g_{i,j}$. $\mu_{x-y} = \sum_{n=0}^{Q-1} n g_{x-y}(n)$. $H_x = -\sum_i g_x(i) \log(g_x(i))$, and $H_y = -\sum_i g_y(i) \log(g_y(i))$. λ is the second largest eigenvalue of $\mathbf{A}$, where $\mathbf{A}(i,j) = \sum_k \frac{g_{i,k} g_{j,k}}{g_x(i) g_y(k)}$.

C. Attributes Selection

In this section, we briefly describe the GKMI method that we employ to find relevant attributes [22]. This approach consists of three steps: segmentation, graph building, and graph clustering.

Segmentation: GKMI finds different attributes for separate zones of an image (superpixels) to reflect their particularity. In fact, different parts of an image might represent separate entities and may require different types of attributes to be well represented. Moreover, even if the various parts represent the same entities, they might be observed under different conditions (different noise levels, light conditions, incidence angles, etc.). As such, the first step of GKMI consists of determining the superpixels using the Watershed segmentation method [28], [29].

Graph building: Let L be the number of superpixels extracted using the segmentation method and N the number of initial attributes (see Table III).

In order to select K relevant attributes, among the initial N attributes, GKMI employs two similarity measures, GK and MI. GK permits to preserve the structure of the original set and is defined as follows:

$$w_{i,j}^{\mathrm{GK}} = \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_j||^2}{2\sigma^2}\right) \tag{1}$$

where $||.||$ is the Frobenius norm, $\mathbf{x}_i$ and $\mathbf{x}_j$ are the vectors corresponding to the i-th and j-th attribute, respectively, and $\sigma > 0$ is a parameter that controls the measure of similarity of the attributes [30]. A large value of σ will indicate a larger similarity even if the Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ is relatively large, whereas, conversely, a lower value of σ weakens the resemblance judged from only the Euclidean distance. Accordingly, σ might be interpreted as a scale factor that controls the strength of the similarity measure between attributes. In this work, we set σ to the default value 1. It is worth noting that we tried other values of σ, but the performance of the analysis did not change significantly (less than 1% of accuracy).

MI quantifies the shared information between two attributes and is defined as follows [31]:

$$w_{i,j}^{\mathrm{MI}} = D_{\mathrm{KL}}\left(P(\mathbf{x}_i, \mathbf{x}_j) || P(\mathbf{x}_i) P(\mathbf{x}_j)\right) \tag{2}$$

where $D_{\mathrm{KL}}(.||.)$ is the Kullback–Leibler divergence, $P(\mathbf{x}_i, \mathbf{x}_j)$ is the joint density function of $\mathbf{x}_i$ and $\mathbf{x}_j$, and $P(\mathbf{x}_i)$ and $P(\mathbf{x}_j)$ are the marginals. Small values of $w_{i,j}^{\mathrm{MI}}$ exhibits independency of $\mathbf{x}_i$ and $\mathbf{x}_j$ which means that both attributes encompass different information. Conversely, high values of $w_{i,j}^{\mathrm{MI}}$ show dependency between the attributes, which means that both reflect similar information.

For each superpixel l, we build a multigraph $\mathcal{G}_l(V, E)$ with N vertices corresponding to the N attributes connected by two edges. The weights of the edges (strength of the connections) are given by GK (1) and MI (2). The MI is measured image-wise, considering all pixels in the image, while the GK is calculated superpixel-wise, using only the pixels within each superpixel.

Graph clustering: By partitioning the graph $\mathcal{G}_l$ into subgraphs, GKMI groups similar attributes together, according to MI and GK measures. As such by selecting a representative attribute from each subgraph, we obtain a subset that preserves the structure and the information content of the original set of attributes within the l superpixel.

Two Laplacian matrices, $\mathbf{L}^{\mathrm{GK}}$ and $\mathbf{L}^{\mathrm{MI}}$, are associated with the graph $\mathcal{G}_l$, and are defined as follows:

$$\mathbf{L}^{\mathrm{GK}} = \mathbf{I} - \mathbf{D}^{\mathrm{GK}^{-1/2}} \mathbf{W}^{\mathrm{GK}} \mathbf{D}^{\mathrm{GK}^{-1/2}} \tag{3}$$

$$\mathbf{L}^{\mathrm{MI}} = \mathbf{I} - \mathbf{D}^{\mathrm{MI}^{-1/2}} \mathbf{W}^{\mathrm{MI}} \mathbf{D}^{\mathrm{MI}^{-1/2}} \tag{4}$$

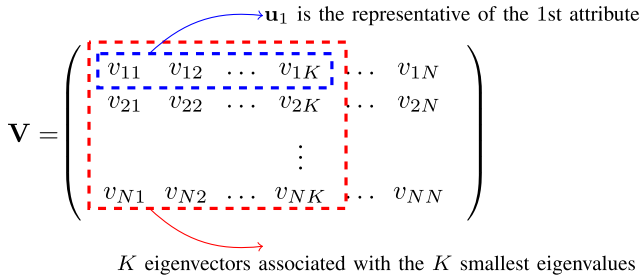
where $\mathbf{I}$ is the identity matrix. $\mathbf{W}^{\mathrm{GK}} = (w_{ij}^{\mathrm{GK}})_{ij}$ and $\mathbf{W}^{\mathrm{MI}} = (w_{ij}^{\mathrm{MI}})_{ij}$ are the adjacency matrices of the graph $\mathcal{G}_l$, and $\mathbf{D}^{\mathrm{GK}} = \mathrm{diag}(\sum_{i \neq j} w_{lij}^{\mathrm{GK}})$ and $\mathbf{D}^{\mathrm{MI}} = \mathrm{diag}(\sum_{i \neq j} w_{ij}^{\mathrm{MI}})$ are their corresponding degree matrices, respectively.

The number of subgraphs within the graph $\mathcal{G}_l$ is equal to the multiplicity of the null eigenvalues of $\mathbf{L}^{\mathrm{GK}}$ and $\mathbf{L}^{\mathrm{MI}}$. Moreover, each eigenvector associated with a null eigenvalue is an indicator of a subgraph [30]. Accordingly, the partition of the graph $\mathcal{G}_l$ into subgraphs can be performed by embedding the attributes into a new manifold spanned by the joint null eigenvectors (eigenvectors associated with the null eigenvalues) of the Laplacian matrices. The joint eigenvectors of $\mathbf{L}^{\mathrm{GK}}$ and $\mathbf{L}^{\mathrm{MI}}$ are obtained by a joint approximate diagonalization [32]

$$\mathbf{L}^{\mathrm{GK}} = \mathbf{V} \mathbf{\Lambda}^{\mathrm{GK}} \mathbf{V}^T \tag{5}$$

$$\mathbf{L}^{\mathrm{MI}} = \mathbf{V} \mathbf{\Lambda}^{\mathrm{MI}} \mathbf{V}^T \tag{6}$$

where $\mathbf{V} = (v_{i,j})_{0 \leq i,j \leq N}$ is the matrix of eigenvectors, and $\mathbf{\Lambda}^{\mathrm{GK}} = \mathrm{diag}(\lambda_1^{\mathrm{GK}}, \ldots, \lambda_N^{\mathrm{GK}})$ and $\mathbf{\Lambda}^{\mathrm{MI}} = \mathrm{diag}(\lambda_1^{\mathrm{MI}}, \ldots, \lambda_N^{\mathrm{MI}})$ are diagonal matrices of the corresponding eigenvalues. The K first eigenvectors corresponding to the smallest eigenvalues will define a new representation of the attributes, i.e., the vector $\mathbf{u}_k = [v_{k,1}, \ldots, v_{k,N}]^T$ is the new representation of the k-th attribute $\mathbf{x}_k$, as follows:



The final step consists of performing a clustering on the new attributes $\mathbf{u}_1, \ldots, \mathbf{u}_K$. In our case, we perform the clustering using K-means [18]. The subset of K relevant attributes correspond to the closest attributes to the centroids of the clusters. It is worth noting that the embedding of the attributes increases their separability and accordingly gives better results than when performing the clustering on the original attributes [30]. Moreover, unlike the feature extraction approaches, the attributes embedding can be mapped back to the original set (since the original attributes and their new representations have the same indices), hence preserving their physical interpretability.

TABLE III
INITIAL ATTRIBUTES CONSIST OF THE ORIGINAL BANDS/POLARIZATIONS AND 14 TEXTURAL FEATURES (TABLE II)

Data set	Original	Derived	N
ICESAR	2 × C-band + 3 × L-band	Textural	75
Sentinel-1/Sentinel-2	2 × SAR + 13 × optical	Textural	225
Radarsat-2/Landsat-8	2 × SAR + 11 × optical	Textural	195

TABLE IV
CLASSES DETERMINED BY SEA ICE EXPERTS FROM VISUAL INSPECTION, AND NUMBER OF TRAINING SAMPLES AND REGIONS OF INTEREST FOR ALL THE DATA SETS USED IN THIS ARTICLE

	Class Name	Class Description	Training Samples	ROI
ICESAR	OW	Open Water	2398	3
	NI	Nilas	12946	8
	GI	Grey Ice	2342	5
	GWI	Grey-White Ice	10640	9
	LFY	Level FYI	14233	8
	DFY	Deformed FYI	6356	12
S-1/S-2	OW	Open Water	911	3
	GI	Grey Ice	9906	3
	GWI	Grey-White Ice	6205	3
	TNFY	Thin FYI	5054	3
	TKFY	Thick FYI	8009	3
R-2/L-8	OW	Open Water	401	3
	NI	Nilas	774	6
	YI	Young Ice	889	5
	TNFY	Thin FYI	746	4
	TKFY	Thick FYI	1381	6

Note: Training samples refers to the number of pixels available for classification performance evaluation.

D. Classification

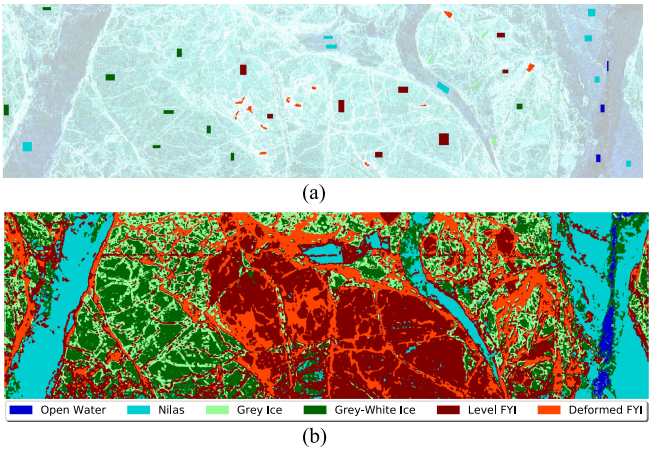
It is noteworthy that the main focus of this article is the application of the GKMI method to determine an optimal set of attributes, which we here demonstrate for the separation of ice types. We perform the sea ice classification using the random forest method, which is a widely applied classifier in remote sensing [23].

Table IV illustrates the sea ice and water classes for all the data sets used in this article as well as the number of regions of interest (ROIs) and available training samples that have been applied for creating the training and test data sets to evaluate the classification performance. Additionally, Figs. 6(a), 7(a), and 8(a) show the distribution of ROIs that were used for the classification of all the data sets.

In all the experiments, we randomly choose 20% of the samples from each label as a training set, while the remaining 80% of samples are used as a test set for performance evaluation.

IV. EXPERIMENTS

In this section, we study the relevance of various attributes, described in the previous section, for the characterization of different sea ice types employing GKMI. To quantitatively evaluate the result of sea ice classification, we use two measures: the overall accuracy (OA) index and Cohen's Kappa coefficient (k). The OA shows the percentage of correctly classified samples,



(a)

(b)

Fig. 6. (a) ROI and (b) classified map for ICESAR data set.

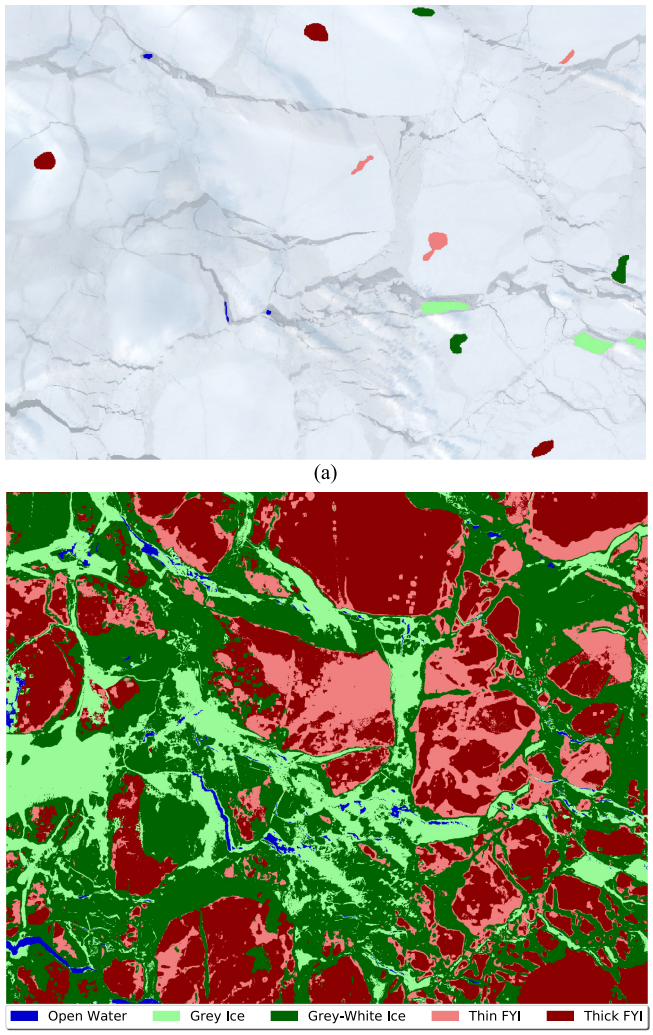
while Kappa measures the agreement between the classification and the reference data [33].

In the following, using the ICESAR, Radarsat-2/Landsat-8, and Sentinel-1/Sentinel-2 data sets, we run different analyses to demonstrate the relevance of automatic feature selection in Section IV-A, the importance of the combination of several imaging modes in Section IV-B, and the data dependency of the relevant attributes in Section IV-C. Finally, in Section IV-D, we compare GKMI to the commonly automatic approaches for attributes selection, PCA and FS.

A. *Performance Analysis*

Before conducting our analysis of the relevant attributes for the characterization of different ice types, we evaluate the performance of GKMI.

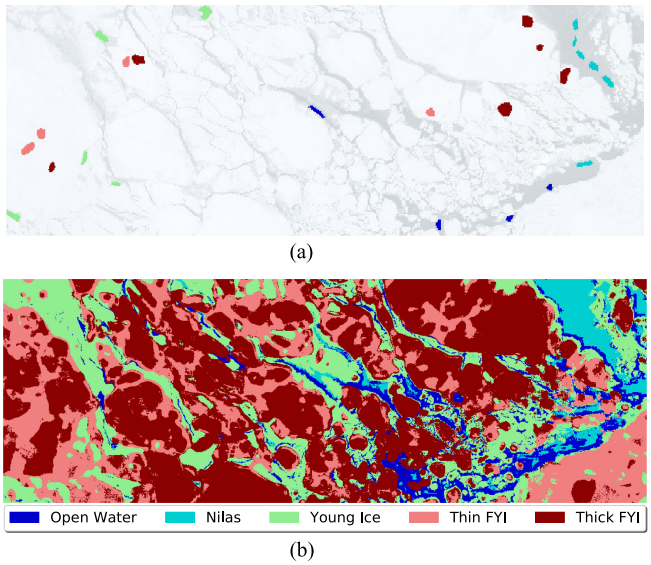
Fig. 9 illustrates the overall accuracies of the GKMI attribute selection method dependent on the number of selected attributes for the ICESAR, Sentinel-1/Sentinel-2, and Radarsat-2/Landsat-8 data sets. The red line indicates the accuracies obtained with the RF classifier for ICESAR, the blue line refers to Sentinel-1/Sentinel-2, and the black line illustrates the Radarsat-2/Landsat-8 data set. The stars show the points where the accuracies reach their peak. It can be clearly seen that all curves rise sharply until the number of attributes reaches 20 for ICESAR and Sentinel-1/Sentinel-2 and 30 for Radarsat-2/Landsat-8, whereupon the accuracies become stable and high for Sentinel-1/Sentinel-2 and Radarsat-2/Landsat-8 and slightly decreasing for ICESAR. Moreover, the maximum accuracy was reached with almost half of the attributes for ICESAR, less than one-fifth for Radarsat-2/Landsat-8, and less than one-seventh of the attributes for the Sentinel-1/Sentinel-2 data set. Table V shows the OA and Kappa coefficient (k) obtained for the optimal number of selected attributes K, and the ones obtained for the total number of attributes N. After reaching the point of the highest accuracy, the inclusion of additional attributes in the classification does not provide any further information which could improve the classification performance for Sentinel-1/Sentinel-2 and Radarsat-2/Landsat-8 data sets. In case of the ICESAR data set, more attributes slightly decline the classifier performance.



(a)

(b)

Fig. 7. (a) ROI and (b) classified map for Sentinel-1/Sentinel-2 data set.



(a)

(b)

Fig. 8. (a) ROI and (b) classified map for Radarsat-2/Landsat-8 data set.

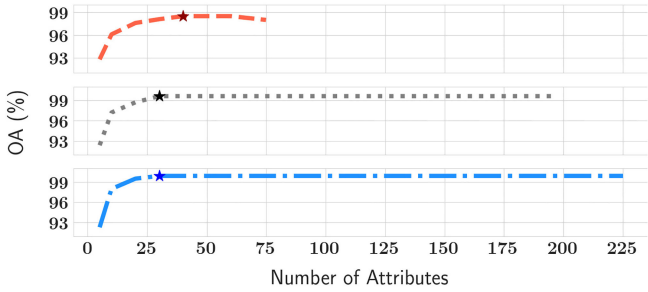


Fig. 9. Overall accuracies of GKMI as a function of the number of selected attributes for ICESAR, Sentinel-1/Sentinel-2, and Radarsat-2/Landsat-8 data sets.

TABLE V
CLASSIFICATION PERFORMANCE OBTAINED OVER DIFFERENT DATA SETS

	K	OA	k	N	OA	k
ICESAR	40	**98.5**	**98.0**	75	98.0	97.5
Radarsat-2/Landsat-8	30	**99.6**	99.3	195	**99.6**	**99.5**
Sentinel-1/Sentinel-2	30	**99.9**	99.8	225	**99.9**	**99.9**

Note: K refers to the optimal number of selected attributes, N shows the total number of attributes, k refers to the Kappa coefficient, and OA (%) corresponds to the overall accuracy. The maximum values of the OA and k are shown in bold. Note that the table shows the classification performance for both optimal and total number of attributes.

TABLE VI
CLASSIFICATION PERFORMANCE OBTAINED USING C-BAND, L-BAND, AND MULTI-FREQUENCY ATTRIBUTES FOR ICESAR DATA SET

L-band			**C-band**			**Multi-frequency**		
K	OA	k	K	OA	k	K	OA	k
40	93.2	91.2	30	90.1	87.2	40	**98.5**	**98.0**
20	92.4	90.2	20	89.6	86.5	20	**97.6**	**96.9**

These results clearly demonstrate the prominence of information selection.

Figs. 6(b), 7(b), and 8(b) show the classified maps for multi-frequency and multi-sensor data sets used in this work. Since the ICESAR data were acquired with a higher spatial resolution, narrow structures such as ice ridges, cracks, and small leads are easier to identify in the classification maps.

B. Multi-sensor vs. Single-Sensor

To demonstrate the advantage of combining data obtained from various sensors, we compare the performance of the classification algorithm when only using the attributes of each imaging mode individually and when using their attributes combined.

1) ICESAR: Table VI shows the OA and Kappa coefficient k obtained for K, the optimal number of selected attributes among the attributes extracted for L-band, C-band, and multi-frequency data set (i.e., when we are using them together). Note that the optimal number is different for each case. To better show the difference in performance, we also added results obtained for a fixed amount of selected attributes $K = 20$. It is evident that the joint use of both data sets increases the accuracy significantly. The OA of the multi-frequency data set reaches its peak with 40 attributes and is equal to 98.5%, while the L-band achieves 93.2% and C-band 90.1%. In fact, even with only 20 attributes, the multi-frequency data set achieves higher performance than the maximum achievable performance using a single band with its optimal number of attributes. Our result indicates a slightly better accuracy at L-band than at C-band. Since the penetration depth into the ice is larger at L-band, deformation structures are easier to recognize in the corresponding SAR images. In cases for which deformation structures comprise an important criterion for classification (like for our data set here), better accuracies are often achieved at L-band.

Furthermore, we analyze the interclass performance for the single-frequency and the multi-frequency approaches by calculating the normalized confusion matrices depicted in Table VII(a)–(c). The results show that the NI (nilas) and DFY (deformed FYI) were the easiest to distinguish for single-frequency and multi-frequency data sets, compared to OW (open water), GWI (grey-white ice), and GI (grey ice). Besides, we notice that GI has the lowest accuracy among all classes for both L-band and C-band data. In fact, GI is highly misclassified as GWI, as can be visible on the confusion matrices. This misclassification

TABLE VII
NORMALIZED CONFUSION MATRIX FOR CLASSIFICATION RESULTS OBTAINED WITH ICESAR DATA SET

		Actual class					
		OW	NI	GI	GWI	LFY	DFY
Assigned class	OW	0.722	0	0.065	0.03	0	0
	NI	0	0.978	0	0.001	0.011	0
	GI	0.117	0	0.664	0.111	0	0
	GWI	0.161	0	0.271	0.854	0	0
	LFY	0	0.022	0	0.004	0.893	0.021
	DFY	0	0	0	0	0.096	0.979

(a)

		Actual class					
		OW	NI	GI	GWI	LFY	DFY
Assigned class	OW	0.961	0.007	0.009	0	0	0
	NI	0.008	0.974	0.042	0.019	0	0
	GI	0.026	0.001	0.792	0.05	0.011	0
	GWI	0.005	0.018	0.139	0.867	0.065	0
	LFY	0	0	0.018	0.064	0.924	0
	DFY	0	0	0	0	0	1

(b)

		Actual class					
		OW	NI	GI	GWI	LFY	DFY
Assigned class	OW	0.964	0	0.005	0	0	0
	NI	0	0.999	0	0	0.003	0
	GI	0.021	0	0.926	0.035	0	0
	GWI	0.015	0	0.069	0.962	0	0
	LFY	0	0.001	0	0.003	0.997	0
	DFY	0	0	0	0	0	1

(c)

Note: OW denotes the open water, NI the nilas, GI the grey ice, GWI the grey-white ice, LFY the level FYI, and DFY the deformed FYI class. (a) C-band attributes. (b) L-band attributes. (c) Multi-frequency attributes.

TABLE VIII
CLASSIFICATION PERFORMANCE OBTAINED USING SAR, OPTICAL, AND MULTI-SENSOR ATTRIBUTES FOR SENTINEL-1/SENTINEL-2

	SAR			Optical			Multi-sensor		
N	K	OA	k	K	OA	k	K	OA	k
225	30	98.2	97.6	100	99.8	99.7	30	**99.9**	**99.8**
43	20	97.8	97.1	13	99.4	99.1	10	**99.7**	**99.6**

results from the significant overlapping of GI and GWI signatures because of their high similarities in many radar signature attributes. Moreover, the L-band data set shows higher accuracy than the C-band data set for all classes. However, the highest accuracies were always achieved by the multi-frequency data set. The prominence of the multi-frequency approach is more apparent for the class GI (grey ice). In fact, by combining L-band and C-band, the accuracy of their most challenging class GI has significantly improved.

To show the flexibility and robustness of the attribute selection method and multi-sensor approach, we also perform a classification comparison of SAR and optical attributes for Sentinel-1/Sentinel-2 and Radarsat-2/Landsat-8. The optical data that were used as complementary source for the manual classification by experts could not be geometrically registered to the radar images because the time gap between acquisitions was too large to compensate for sea ice deformations. The Radarsat-2/Landsat-8 and Sentinel-1/Sentinel-2 data reveal a clear dominance of the optical attributes over radar attributes. Since this is dependent on the special measurement conditions (e.g., sensor characteristics and ice properties, for optical sensors sun elevation), it is difficult to generalize this result, and, correspondingly, we expect varying priorities of single attributes. Nevertheless, the dominance in the number of attributes does not always lead to better accuracy or more information content, as shown in Fig. 9. We emphasize that one of the main ideas of this study is to show that the GKMI method can be easily adapted to any sensor and image combination and for any environmental conditions.

2) Sentinel-1/Sentinel-2: Table VIII illustrates the classification performance of SAR, optical, and multi-sensor attributes for Sentinel-1/Sentinel-2 data set. This analysis was conducted using two initial sets of attributes. The first set consists of the original bands and polarizations of optical and SAR images in addition to their corresponding textural features. Conversely, in the second set, optical images' textural features were omitted (in this scenario, the data set consists of 43 attributes: 13 optical and 30 SAR). The second set was considered to appropriately evaluate each sensor's contribution. The results show that using only optical attributes, one reaches the same high accuracy as multi-sensor attributes for Sentinel-1/Sentinel-2 data set. However, when using multi-sensor attributes, accuracy is achieved with less number of attributes.

Additionally, we also extracted confusion matrices, like we did for the ICESAR data set, to better analyze the interclass performance of the single-sensor and multi-sensor approaches. Table IX(a)–(c) demonstrates the results for the Sentinel-1/Sentinel-2 data set. We notice that the optical data set achieves a better separation for all classes compared to the SAR data set. However, there is no significant accuracy superiority in any of the classes. Moreover, the multi-sensor approach results in higher accuracies than the single-sensor.

TABLE IX
NORMALIZED CONFUSION MATRIX FOR CLASSIFICATION RESULTS OBTAINED WITH THE SENTINEL-1/SENTINEL-2 DATA SET

		Actual class				
		OW	GI	GWI	TNFY	TKFY
Assigned class	OW	0.977	0	0	0	0.013
	GI	0	0.994	0.008	0	0
	GWI	0	0.006	0.959	0.032	0
	TNFY	0.014	0	0.033	0.968	0.006
	TKFY	0.009	0	0	0	0.981

(a)

		Actual class				
		OW	GI	GWI	TNFY	TKFY
Assigned class	OW	0.993	0	0	0	0
	GI	0.007	0.998	0.005	0	0
	GWI	0	0.002	0.995	0	0
	TNFY	0	0	0	1	0
	TKFY	0	0	0	0	1

(b)

		Actual class				
		OW	GI	GWI	TNFY	TKFY
Assigned class	OW	1	0	0	0	0
	GI	0	1	0.003	0	0
	GWI	0	0	0.997	0	0
	TNFY	0	0	0	1	0.001
	TKFY	0	0	0	0	0.999

(c)

Note: OW denotes the open water, GI the grey ice, GWI the grey-white ice, TNFY the thin FYI, and TKFY the thick ice class. (a) SAR attributes. (b) Optical attributes. (c) Multi-sensor attributes.

TABLE X
CLASSIFICATION PERFORMANCE OBTAINED USING SAR, OPTICAL, AND MULTI-SENSOR ATTRIBUTES FOR RADARSAT-2/LANDSAT-8 DATA SET

	SAR			Optical			Multi-sensor		
N	K	OA	k	K	OA	k	K	OA	k
195	20	91.0	88.3	120	98.5	98.1	30	**99.6**	**99.3**
41	20	91.0	88.3	11	95.3	93.9	10	**98.6**	**98.2**

3) Radarsat-2/Landsat-8: Now, we evaluate the difference in performance between optical and SAR sensors for Radarsat-2/Landsat-8 data set. As for the Sentinel-1/Sentinel-2 data set, we consider two sets of initial attributes. In the second set, the textural features of the optical bands were excluded to have a comparable number of attributes for each sensor. Accordingly, for this scenario, the Radarsat-2/Landsat-8 data set includes 41 attributes (11 optical and 30 SAR). Table X illustrates the performance comparison using SAR, optical, and multi-sensor attributes. Unlike the Sentinel-1/Sentinel-2 data

TABLE XI
NORMALIZED CONFUSION MATRIX FOR CLASSIFICATION RESULTS OBTAINED WITH RADARSAT-2/LANDSAT-8 DATA SET

		Actual class				
		OW	NI	YI	TNFY	TKFY
Assigned class	OW	0.857	0.004	0	0	0.01
	NI	0.031	0.848	0	0.005	0.147
	YI	0	0	1	0	0
	TNFY	0.017	0.016	0	0.993	0.001
	TKFY	0.095	0.133	0	0.002	0.842

(a)

		Actual class				
		OW	NI	YI	TNFY	TKFY
Assigned class	OW	0.974	0.024	0.015	0	0
	NI	0.016	0.976	0.007	0	0
	YI	0.01	0	0.978	0	0
	TNFY	0	0	0	0.998	0.001
	TKFY	0	0	0	0.002	0.999

(b)

		Actual class				
		OW	NI	YI	TNFY	TKFY
Assigned class	OW	0.985	0.013	0	0	0
	NI	0.015	0.987	0	0	0
	YI	0	0	1	0	0
	TNFY	0	0	0	1	0
	TKFY	0	0	0	0	1

(c)

Note: OW denotes the open water, NI the nilas, YI the young ice, TNFY the thin FYI, and TKFY the thick ice class. (a) SAR attributes. (b) Optical attributes. (c) Multi-sensor attributes.

set, where the single-sensor performance for both sensors was high, the Radarsat-2/Landsat-8 single-sensors show less accurate performance than multi-sensor attributes.

Table XI(a)–(c) shows the confusion matrices for Radarsat-2/Landsat-8. The results reveal that the YI (young ice) was the easiest to distinguish using SAR, while TNFY (thin FYI) and TKFY (thick FYI) showed high classification accuracy using optical data. Unlike optical, the SAR part shows significantly lower accuracies for OW (open water), NI (nilas), and TKFY (thick FYI). However, the multi-sensor attributes combine the advantages of both sensors and ensure the highest separability for all classes.

It should be noted that even though, for this particular example, optical attributes result in high accuracies, they cannot serve as basic source for operational monitoring. This is due to several familiar factors, such as limitations due to light and cloud conditions that influence the separability of rough and smooth level ice, and snow cover that prevents recognition of sea ice types.

The results obtained for the investigated data sets reflect different situations: the dominance of one frequency or one sensor for ice type separation (ICESAR and Sentinel-1/Sentinel-2), and the complementarity of sensors (Radarsat-2/Landsat-8). In all cases, the combination of data from different sensors guarantees better performance. In particular, for certain ice conditions and combinations of available data, the multi-sensor approach can increase the classification accuracy significantly if compared to the single-sensor case (in Table VII(c), e.g., the ICESAR data).

C. Relevant Attributes

Because the selection by GKMI is performed in superpixel space, different attributes are selected for different classes. The selected attributes can also differ between the superpixels of one ice-class due to the influence of several factors, such as incidence angle, range-dependent noise, light and cloud conditions, etc. However, we can still estimate the relevant attributes for a class by identifying its frequently selected attributes. To this end, we identify all superpixels belonging to a class and extract their chosen attributes.

Fig. 10 shows the histograms of the five most selected attributes for each class for the different data sets.

Fig. 10(a) shows the relevant attributes for each class of ICESAR data set. The histograms show that, in general, C-band and L-band attributes were often selected equally. However, for some classes, there is a clear predominance of L-band (open water, level FYI, and grey-white ice) or C-band attributes (nilas). Moreover, several attributes, such as information correlation and inverse different moment derived from L-band VV polarization (L VV IC and L VV IDM), were selected for several classes of the ICESAR data set.

Fig. 10(b) demonstrates the relevant attributes selected for the different classes of Sentinel-1/Sentinel-2 data set. Optical attributes are predominant for this data set, especially for thick and thin FYI. Moreover, the intensity of band B1 (B1 INTST) was frequently selected for three classes (open water, grey-white ice, and thick FYI).

Fig. 10(c) illustrates the relevant attributes for the classes of the Radarsat-2/Landsat-8 data set. Optical attributes were selected more frequently than SAR attributes; however, unlike the Sentinel-1/Sentinel-2 data set, all the classes have a mixed set of selected attributes. The most predominantly selected attribute for Radarsat-2/Landsat-8 was the maximum correlation coefficient derived from HV polarization (HV MCC), that was chosen for open water, nilas, young ice, and thick FYI.

From the results mentioned above, it is possible to conclude that it is hard to find any generalization among the attributes and corresponding sea ice types. Moreover, for a given combination of data types, selected attributes vary even between the superpixels that belong to the same ice class. This indicates that the selection procedure is sensitive to the technical setup (e.g., incidence and sun elevation angle) and slight variations of the appearance of a given ice class (e.g., varying degrees of deformation and ice concentration). Therefore, it is crucial to have a flexible attribute selection method that allows selecting the relevant attributes for different data combinations and sea ice classes even under various conditions to enhance the accuracy of classification.

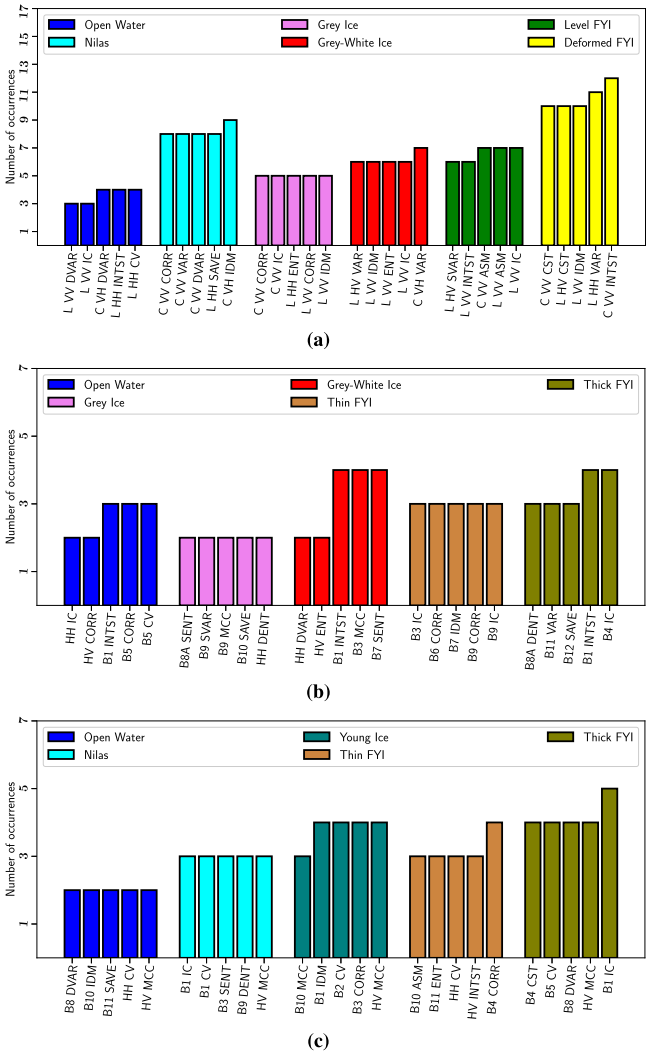


Fig. 10. Five most frequently selected attributes for each class of (a) ICESAR, (b) Sentinel-1/Sentinel-2, and (c) Radarsat-2/Landsat-8. Note that the height of the bars (number of attributes' occurrences) is different for the various classes because the number of used ROI is different for each class. INTST refers to intensity. For radar data, intensity refers to the backscattering coefficient.

TABLE XII
PERFORMANCE COMPARISON AMONG DIFFERENT METHODS FOR ATTRIBUTE SELECTION, USING RF CLASSIFIER

	ICESAR			Sentinel-1/Sentinel-2			Radarsat-2/Landsat-8		
Method	N	OA (%)	k	N	OA (%)	k	N	OA (%)	k
FS	40	98.0	97.5	30	99.8	99.7	30	97.7	97.0
PCA	40	97.7	97.0	30	99.9	99.8	30	98.9	98.0
GKMI	40	**98.5**	**98.0**	30	99.9	99.8	30	**99.6**	**99.3**

D. Comparison of Methods

To evaluate the GKMI method's performance, we compared its results to the two mostly used automatic attributes selection methods, mainly PCA and FS. PCA is an unsupervised feature extraction approach, while FS is a supervised attribute selection. Table XII shows the accuracy and kappa for the different methods and data sets used in this article. All methods show a high classification accuracy; however, GKMI gives slightly higher accuracy for all data sets. These results validate our approach's efficiency since it offers better performance without any supervision as for FS, and while preserving the physical interpretability of attributes as opposed to PCA.

We emphasize that in this study, the major focus was on attribute selection for maximizing the information content of satellite and airborne images as a step toward improved sea ice classification. The results of attribute selection can also be applied for an assessment of dominant scattering mechanisms or for optimizing the retrieval of ice parameters such as roughness or thickness.

In comparison to other attribute selection algorithms, particularly to the ones that were used in this study, GKMI is substantially different, which is an advantage for automated classification. Hereby, the optimal subset of attributes using the GKMI method is determined according to two different criteria, which are structure (GK) and information content (MI). These criteria are applied simultaneously and account for the global and local particularities of the original data set. Moreover, the method is performed on a superpixel level, which, in combination with the use of two similarity measures explained in Section III-C, allows us to capture relevant information at different scales. Additionally, GKMI preserves the advantages of both attribute extraction and selection approaches, namely, increases the class separability and, at the same time, preserves the physical meaning of the original data.

V. CONCLUSION

In this article, we employed GKMI, a recently developed approach for attribute selection, and applied it to the task of sea ice classification using different multi-sensor data combinations. The optimal set of selected attributes, by GKMI, is not fixed for the whole image or a sea ice class but is adaptive for each region of the observed scene according to the technical and environmental conditions (such as noise level, incidence angle, cloudiness, etc.)

Furthermore, our results show the ability of GKMI to process different combinations of data sets and the importance of deploying multi-sensors for the characterization of ice types by comparing the performance of using several sensors separately and simultaneously. Since the major idea of our approach is to introduce a highly flexible and adaptive scheme, it is not necessary to recommend fixed sets of attributes for classification, which depend on sensor types, imaging modes, and environmental and ice conditions during data acquisitions. As such, the fact that optical information is more accurate in the considered data sets is not of general validity due to the limitations of optical sensors in cloudy and dark areas.

Finally, two issues still need to be addressed. For the first, we did not yet investigate in detail how our results are affected by the observation conditions, such as radar incidence angle or sun elevation angle. This is part of future investigations. Another item concerns the availability of the investigated sensor combinations for operational mapping. Since our results clearly demonstrate the advantage of multi-sensor classification (here focussing on

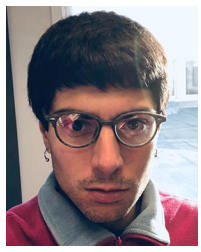
GKMI), space agencies and operational services should develop strategies that make the joint use of data from different sensor types possible. For example, it has to be considered that optical data can only be used under favorable light and cloud conditions (hence requiring a "only-add-if-usable" strategy). Moreover, the combination of different radar frequencies for operational services is recently under discussion, e.g., the Copernicus ROSE-L mission (see [34]). For joint data use, it is important to consider the drift of sea ice, which requires that different sensors acquire data with smallest possible time gaps over a given region. For the combination of C-band and L-band SAR, corresponding acquisition scenarios are under investigation in the ROSE-L project.

ACKNOWLEDGMENT

The authors would like to thank the German Aerospace Center (ESAR)for providing the ICESAR data set. The Landsat-8 images were provided by the United States Geological Survey (USGS). RADARSAT-2 data were provided by NSC/KSAT under the Norwegian-Canadian Radarsat agreement 2018. The article contains Sentinel-1 and Sentinel-2 Copernicus data (2018, 2020).

REFERENCES

[1] M. R. England, L. M. Polvani, and L. Sun, "Robust arctic warming caused by projected antarctic sea ice loss," *Environ. Res. Lett.*, vol. 15, Sep. 2020, Art. no. 104005.
[2] S. Funder *et al.*, "History of sea ice in the arctic," *Quaternary Sci. Rev.*, vol. 29, no. 15/16, pp. 1757–1778, 2010.
[3] S. Sandven, O. M. Johannessen, and K. Kloster, "Sea ice monitoring by remote sensing," in *Encyclopedia of Analytical Chemistry*. Hoboken, NJ, USA: Wiley, 2006, pp. 1–43.
[4] J. Lohse, A. P. Doulgeris, and W. Dierking, "An optimal decision-tree design strategy and its application to sea ice classification from SAR imagery," *Remote Sens.*, vol. 11, no. 13, 2019, Art. no. 1574.
[5] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, Sep. 2015.
[6] N. Longbotham *et al.*, "Multi-modal change detection, application to the detection of flooded areas: Outcome of the 2009–2010 data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 331–342, Feb. 2012.
[7] T. Geldsetzer, M. Arkett, T. Zagon, F. Charbonneau, J. J. Yackel, and R. K. Scharien, "All-season compact-polarimetry C-band SAR observations of sea ice," *Can. J. Remote Sens.*, vol. 41, no. 5, pp. 485–504, 2015.
[8] M. E. Shokr, "Evaluation of second-order texture parameters for sea ice classification from radar images," *J. Geophys. Res.: Oceans*, vol. 96, no. C6, pp. 10625–10640, 1991.
[9] D. Barber and E. LeDrew, "SAR sea ice discrimination using texture statistics : A multivariate approach," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 4, pp. 385–395, 1991.
[10] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 780–795, Mar. 1999.
[11] D. A. Clausi and M. E. Jernigan, "A fast method to determine co-occurrence texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 1, pp. 298–300, Jan. 1998.
[12] D. Murashkin, G. Spreen, M. Huntemann, and W. Dierking, "Method for detection of leads from Sentinel-1 SAR images," *Ann. Glaciol.*, vol. 59, no. 76pt2, pp. 124–136, 2018.
[13] J.-W. Park, A. A. Korosov, M. Babiker, J.-S. Won, M. W. Hansen, and H.-C. Kim, "Classification of sea ice types in Sentinel-1 synthetic aperture radar images," *Cryosphere*, vol. 14, no. 8, pp. 2629–2645, 2020.
[14] M.-A. Moen, "Analysis and interpretation of c-band polarimetric SAR signatures of sea ice," Ph.D. thesis, UiT Arctic Univ. Norway, Tromsø, Norway, 2015.
[15] N. C. Wright and C. M. Polashenski, "Open-source algorithm for detecting sea ice surface features in high-resolution optical imagery," *Cryosphere-Discussions*, vol. 12, no. 4, pp. 1307–1329, 2018.
[16] X. Miao, H. Xie, S. F. Ackley, D. K. Perovich, and C. Ke, "Object-based detection of arctic sea ice and melt ponds using high spatial resolution aerial photographs," *Cold Regions Sci. Technol.*, vol. 119, pp. 211–222, 2015.
[17] B. Ghojogh *et al.*, "Feature selection and feature extraction in pattern analysis: A literature review," 2019, *arXiv:1905.02845*.
[18] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Orlando, FL, USA: Academic Press, Inc., 2008.
[19] M. Fauvel, C. Dechesne, A. Zullo, and F. Ferraty, "Fast forward feature selection of hyperspectral images for classification with Gaussian mixture models," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2824–2831, Jun. 2015.
[20] S. H. Huang, "Supervised feature selection: A tutorial," *Artif. Intell. Res.*, vol. 4, pp. 22–37, 2015.
[21] S. Solorio-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, "A review of unsupervised feature selection methods," *Artif. Intell. Rev.*, vol. 53, pp. 907–948, 2020.
[22] E. Khachatrian, S. Chlaily, T. Eltoft, and A. Marinoni, "A multimodal feature selection method for remote sensing data analysis based on double graph Laplacian diagonalization," to be published.
[23] A. Liaw and M. Wiener, "Classification and regression by random forest," *Forest*, vol. 23, pp. 18–23, Nov. 2001.
[24] W. Dierking, "Technical assistance for the deployment of airborne SAR and geophysical measurements during the ICESAR2007; *final report - part 2: Sea ice*," 2008.
[25] A. V. Bushuyev, "Sea-ice nomenclature," WMO /OMM /BMO - No. 259. Russian Federation, 1970.
[26] R. Haralick, K. Shanmugam, and I. Dinstein, "Texture features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, no. 6, pp. 610–621, Nov. 1973.
[27] U. Kandaswamy, D. A. Adjeroh, and M. C. Lee, "Efficient texture analysis of SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2075–2083, Sep. 2005.
[28] S. Beucher, "The watershed transformation applied to image segmentation," in *Proc. 10th Pfefferkorn Conf. Signal Image Process. Microsc. Microanalysis*, 1992, pp. 299–314.
[29] P. Neubert and P. Protzel, "Compact watershed and preemptive SLIC: On improving trade-offs of superpixel segmentation algorithms," in *Proc. 22nd Int. Conf. Pattern Recognit.*, 2014, pp. 996–1001.
[30] U. Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, pp. 395–416, Dec. 2007.
[31] J. R. Vergara and P. A. Estévez, "A review of feature selection methods based on mutual information," *Neural Comput. Appl.*, vol. 24, no. 1, pp. 175–186, 2014.
[32] P. Ablin, J. Cardoso, and A. Gramfort, "Beyond Pham's algorithm for joint diagonalization," 2018, *arXiv:1811.11433*.
[33] P. Bharatkar and R. Patel, "Approach to accuracy assessment tor RS image classification techniques," *Int. J. Sci. Eng. Res.*, vol. 4, no. 12, pp. 79–86, 2013.
[34] M. Davidson *et al.*, "Copernicus L-band SAR mission requirements document," Eur. Space Agency, Paris, France, ESA-EOPSM-CLIS-MRD-3371, Tech. Rep., Oct. 2019.



Eduard Khachatrian (Student Member, IEEE) received the dual M.Sc. degrees in polar and marine sciences from the Faculty of Mathematics, Informatics, and Natural Sciences, Hamburg University, Hamburg, Germany, and the Institute of Earth Science, Saint Petersburg State University, Saint Petersburg, Russia, in 2017.

From 2017 to 2018, he was a Junior Scientist with Nansen International Environmental and Remote Sensing Centre, Bergen, Norway. He is currently a Ph.D. Fellow with the Center of Integrated Remote Sensing and Forecasting for Arctic Operations, UiT The Arctic Univesity of Norway, Tromsø, Norway. His research interests include multimodal data analysis, image processing, and remote sensing of polar areas.

Saloua Chlaily (Member, IEEE) received the M.Sc. degree in electronics engineering from the National Higher School of Electrical Engineering, Electronics, Informatics, Hydraulics and Telecommunications, Toulouse, France, in 2013, the M.Sc. degree in electrical engineering from the Hassania School of Public Works, Casablanca, Morocco, in 2013, and the Ph.D. degree in signal, image, speech, and telecommunication from Grenoble Alpes University, Grenoble, France, in 2018.

She is currently a Researcher with the UiT The Arctic Univesity of Norway, Tromsø, Norway, where she is conducting her research with the Center of Integrated Remote Sensing and Forecasting for Arctic Operations. Her research interests include multimodal data analysis, signal processing, and image processing.

Dr. Chlaily is currently a Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, and *IEEE Geoscience and Remote Sensing Letters*.

Torbjørn Eltoft (Member, IEEE) received the M.Sc. and Ph.D. degrees from the University of Tromsø, Norway, in 1981 and 1984, respectively.

In 1988, he joined the Department of Physics and Technology, UiT the Arctic University of Norway (UiT), Tromsø, Norway, where he is currently a Professor and the Director of the Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, a centre for research-based Innovation awarded by the Norwegian Research Council in 2014, whose objective is to develop knowledge and remote sensing technology for arctic applications. From 2013 to 2015, he was the Head of the Department of Physics and Technology, UiT. He was an International Researcher with the University of California, Irvine, CA, USA, from 1992 to 1993 and then from 1997 to 1998, and with the University of California, San Diego, CA, USA, from 2004 to 2005. He has a significant publication record in the area of signal processing and remote sensing. His research interests include multidimensional signal and image analysis, statistical modeling, neural networks, and machine learning, with emphasis on applications in multichannel synthetic aperture radar remote sensing and multi-sensor remote sensing for the Arctic.

Dr. Eltoft was an Associate Editor for the Elsevier journal *Pattern Recognition* from 2005 to 2011 and was the Guest Editor for *Remote Sensing*'s Special Issue for the PolInSAR 2017 Conference. He was the co-recipient of the Outstanding Paper Award in Neural Networks awarded by IEEE Neural Networks Council in 2000, Honourable Mention for the 2003 Pattern Recognition Journal Best Paper Award, and the 2017 UiT Award for Research and Development from UiT The Arctic University of Norway.

Wolfgang Dierking received the M.Sc. degree in geophysics from the University of Hamburg, Hamburg, Germany, in 1985, and the Ph.D. degree in physics (ocean remote sensing) from the University of Bremen, Bremen, Germany, in 1989.

Since 2002, he has been with the Helmholtz Centre for Polar and Marine Research, Alfred Wegener Institute (AWI), Bremerhaven, Germany, and since 2016, he has been a Professor II with the Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, UiT The Arctic Univesity of Norway, Tromsø, Norway, where he has been leading a group working with remote sensing of sea ice and icebergs. Since 2015, he has been a Lecturer of ocean remote sensing with the University of Applied Sciences, Bremerhaven, Germany. From 2007 to 2019, he was the Head of the research group Earth Observation Systems, AWI. The group dealt with remote sensing of sea ice, ice sheets, and permafrost regions. In 2007, he was a Visiting Professor with the Chalmers University of Technology, Gothenburg, Sweden. From 1996 to 2002, he was an Associate Research Professor with the Danish Centre for Remote Sensing, Technical University of Denmark, Kongens Lyngby, Denmark. His research focuses on the remote sensing of the polar regions.

Prof. Dierking was a member of different advisory groups of the European Space Agency (ESA), which include Sentinel-1 Mission Advisory Group from 2011 to 2015, SAR Advisory Group from 2003 to 2010, and Cat-1 Advisory Group from 2003 to 2015. In 2019, he became a member of Copernicus L-Band SAR Mission Advisory Group, and in 2021 a member of the Sentinel-1 Next Generation Mission Advisory Group. He participated in several field expeditions to the Arctic, Antarctic, and the Baltic Sea. Since 2008, he has been the Principal Investigator with Dragon Cooperation Programme between the Ministry of Science and Technology of China and ESA with focus on sea ice remote sensing.

Frode Dinessen received the M.Sc. degree in applied physics from UiT The Arctic University of Norway, Tromsø, Norway, in 1996.

He is currently a Research Scientist with the Division for Remote Sensing and Data Management, Norwegian Meteorological Institute, Tromsø, Norway. From 1997 to 2001, he was the Leader of the Norwegian Ice Service, Forecasting Division for the Northern Norway, Norwegian Meteorological Institute. The Ice Service is responsible for delivering regular updated sea ice analysis for the public users. From 2001 to 2007, he was a Software Engineer with Kongsberg Spacetec, Tromsø, Norway, developing satellite-based applications. In 2007, he joined the Research Department, Norwegian Meteorological Institute. He is currently the Deputy of the Sea Ice Thematic Assembling Center, providing operational sea ice products to the EU Copernicus Marine Environment Monitoring Service. His research interests include development of operational multi-sensor sea ice concentration products by combining satellite data from synthetic aperture radar, passive microwave radiometers, and national data hub for sentinel data.

Andrea Marinoni (Senior Member, IEEE) received the B.S., M.Sc. (cum laude), and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 2005, 2007, and 2011, respectively.

He is currently an Associate Professor with the Earth Observation Group, Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway, and a Visiting Academic Fellow with the Department of Engineering, University of Cambridge, Cambridge, U.K. From 2013 to 2018, he was a Research Fellow with the Telecommunications and Remote Sensing Lab, Department of Electrical, Computer and Biomedical Engineering, University of Pavia. In 2009, he was a Visiting Researcher with the Communications Systems Lab, Department of Electrical Engineering, University of California - Los Angeles, Los Angeles, CA, USA. From 2015 to 2017, he was a Visiting Researcher with the Earth and Planetary Image Facility, Ben-Gurion University of the Negev, Be'er Sheva, Israel; the School of Geography and Planning, Sun Yat-Sen University, Guangzhou, China; the School of Computer Science, Fudan University, Shanghai, China; the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China; Instituto de Telecomunicaccões, Instituto Superior Tecnico, Lisbon, Portugal; and Universidade de Lisboa, Lisbon, Portugal. In 2020 and 2021, he was a Visiting Professor with the Department of Electrical, Computer and Biomedical Engineering, University of Pavia. His main research interests include efficient information extraction from multimodal remote sensing, nonlinear signal processing applied to large-scale heterogeneous records, earth observation interpretation, big data mining, and analysis and management for human–environment interaction assessment.

Dr. Marinoni was the recipient of the two-year Applied Research Grant, sponsored by the region of Lombardy, Italy, and STMicroelectronics N.V., the INROAD Grant, sponsored by the University of Pavia and Fondazione Cariplo, Milan, Italy, for supporting excellence in design of ERC proposal, the Progetto professionalitá Ivano Becchi Grant, funded by Fondazione Banco del Monte di Lombardia, Pavia, Italy, and sponsored by the University of Pavia and NASA Jet Propulsion Laboratory, Pasadena, CA, USA, for supporting the development of advanced methods of air pollution analysis by remote sensing data investigation, and the Åsgard Research Programme Grants funded by Institut Franccais de Norvége, Oslo, Norway, in 2019 and 2020, respectively, for supporting the development of scientific collaborations between French and Norwegian research institutes. He is the Founder and the chair of the IEEE GRSS Norway Chapter. He is also an Ambassador of the IEEE Region 8 Humanitarian Activities and a Research Contact Point for the Norwegian Artificial Intelligence Research Consortium. He was a Topical Associate Editor of machine learning for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the Guest Editor of three special issues on multimodal remote sensing and sustainable development for the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. He is the Leader of the GR4S Committee within IEEE GRSS, coordinating the organization of schools and workshops sponsored by IEEE GRSS worldwide.