Article

Retrieval of Summer Sea Ice Concentration in the Pacific Arctic Ocean from AMSR2 Observations and Numerical Weather Data Using Random Forest Regression

Hyangsun Han [1,*] , **Sungjae Lee** [2] , **Hyun-Cheol Kim** [2] **and Miae Kim** [3]

1   Department of Geophysics, Kangwon National University, Chuncheon 24341, Korea
2   Center of Remote Sensing and GIS, Korea Polar Research Institute, Incheon 21990, Korea;
    sungjae@kopri.re.kr (S.L.); kimhc@kopri.re.kr (H.-C.K.)
3   School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology,
    Ulsan 44919, Korea; kimmiae@unist.ac.kr
*   Correspondence: hyangsun@kangwon.ac.kr

Abstract: The Arctic sea ice concentration (SIC) in summer is a key indicator of global climate change and important information for the development of a more economically valuable Northern Sea Route. Passive microwave (PM) sensors have provided information on the SIC since the 1970s by observing the brightness temperature (T_B) of sea ice and open water. However, the SIC in the Arctic estimated by operational algorithms for PM observations is very inaccurate in summer because the T_B values of sea ice and open water become similar due to atmospheric effects. In this study, we developed a summer SIC retrieval model for the Pacific Arctic Ocean using Advanced Microwave Scanning Radiometer 2 (AMSR2) observations and European Reanalysis Agency-5 (ERA-5) reanalysis fields based on Random Forest (RF) regression. SIC values computed from the ice/water maps generated from the Korean Multi-purpose Satellite-5 synthetic aperture radar images from July to September in 2015–2017 were used as a reference dataset. A total of 24 features including the T_B values of AMSR2 channels, the ratios of T_B values (the polarization ratio and the spectral gradient ratio (GR)), total columnar water vapor (TCWV), wind speed, air temperature at 2 m and 925 hPa, and the 30-day average of the air temperatures from the ERA-5 were used as the input variables for the RF model. The RF model showed greatly superior performance in retrieving summer SIC values in the Pacific Arctic Ocean to the Bootstrap (BT) and Arctic Radiation and Turbulence Interaction STudy (ARTIST) Sea Ice (ASI) algorithms under various atmospheric conditions. The root mean square error (RMSE) of the RF SIC values was 7.89% compared to the reference SIC values. The BT and ASI SIC values had three times greater values of RMSE (20.19% and 21.39%, respectively) than the RF SIC values. The air temperatures at 2 m and 925 hPa and their 30-day averages, which indicate the ice surface melting conditions, as well as the GR using the vertically polarized channels at 23 GHz and 18 GHz (GR(23V18V)), TCWV, and GR(36V18V), which accounts for atmospheric water content, were identified as the variables that contributed greatly to the RF model. These important variables allowed the RF model to retrieve unbiased and accurate SIC values by taking into account the changes in T_B values of sea ice and open water caused by atmospheric effects.

Keywords: summer sea ice concentration; Pacific Arctic Ocean; AMSR2; ERA-5; Random Forest regression

1. Introduction

Sea ice concentration (SIC), defined as the portion of sea ice coverage within a given area, is a key indicator of climate change [1–4]. The decreasing summer Arctic sea ice extent, the sum of areas with at least 15% SIC, is the most representative indicator of global warming [5–8]. Moreover, SIC has a profound influence on ecosystems, biological habitats and human activities in the polar oceans [9–11]. The Arctic summer SIC is important information for the sailing of vessels on the Northern Sea Route (NSR) [11]. Decreasing Arctic

summer sea ice extent suggests the possibility of the development of a more economically valuable NSR. In recent years, a rapid reduction of Arctic sea ice in summer has been reported, and it is expected to vanish by the middle of this century [12–15], which could have significant impacts on the climate and ocean environment, as well as human activities and economics in the Arctic. Therefore, the accurate estimation of summer SIC in the Arctic Ocean is very important.

Satellite passive microwave (PM) sensors have provided information regarding the SIC since the 1970s by observing the microwave radiation characteristics of sea ice and open water [1,16]. The Special Sensor Microwave Imager/Sounder (SSMIS) and Advanced Microwave Scanning Radiometer 2 (AMSR2) are representative PM sensors that are currently observing sea ice and have been operated since 2008 and 2012, respectively [17–19]. For the operational estimation of SIC from the SSMIS measurements of brightness temperatures, the NASA Team (NT) [20] and Bootstrap (BT) [1,21] algorithms were developed. These algorithms use the brightness temperatures (T_B), a measure of the emitted radiance of microwave radiation from the surface, measured at 19 and 37 GHz channels to produce SIC with a grid spacing of a few kilometers or tens of kilometers, which is attributed to the instantaneous field of view (IFOV) of the used channels. For the AMSR2, the BT and Arctic Radiation and Turbulence Interaction STudy (ARTIST) Sea Ice (ASI) [22] algorithms have been used as operational SIC estimation algorithms. The ASI algorithm uses the 89 GHz channel, which has a smaller IFOV than lower-frequency channels and provides information on SIC with a grid spacing of 3.125 km to 12.5 km thanks to the fine IFOV. In addition to the above algorithms, many SIC estimation algorithms have been developed and applied for various satellite PM sensors. The existing algorithms estimate SIC by considering the different microwave emissivities between sea ice and open water. Most algorithms have shown good performance in estimating SIC in winter and in highly ice-concentrated regions [16,20,21,23,24].

In summer, however, the algorithms typically estimate SIC inaccurately because atmospheric effects such as high water content, strong winds, and high air temperature can make the difference between the microwave radiation characteristics of sea ice and open water small [1,23,25–27]. The sea ice surface melts in summer due to the increasing air temperature, which leads to varying the ice emissivity in the microwave range that makes it similar to open water. The atmospheric water content in summer, which is relatively higher than winter, can make the atmospheric signal over the sea ice more dominant relative to the surface signal in microwave range and cause the PM SIC algorithms to produce inaccurate SIC values [25,26]. Furthermore, the surface of open water can be roughened by strong winds. The wind-roughened open water surface can increase the T_B over open water, which can be a source of the overestimation of SIC from the PM observations. [25]. The area of open water increases greatly in summer, and erroneous SIC values could be estimated for a wider area than other seasons by the PM SIC algorithms. Andersen et al. [25] revealed that the values of SIC estimated from the PM SIC algorithms can vary greatly depending on atmospheric effects. The inaccuracy of the summer SIC values even varies from algorithm to algorithm [16,23], which acts as a hindrance to accurately analyzing the declining trends in the Arctic summer sea ice. The inaccuracy of summer Arctic SIC retrieved from the PM SIC algorithms was reported to be up to ±20% [26,28,29], which is closely related to the atmospheric contributions. To compensate for the atmospheric effects on the T_B values of sea ice and open water measured by the PM sensors, the SIC algorithms implement weather filters that use criteria based on combinations of T_B values [20–22,30]. However, large errors in summer PM SIC imply that the weather filters have limitations in terms of correcting the atmospheric contamination of T_B and suggest that it is difficult to estimate SIC accurately in summer using PM observations only.

Machine learning techniques, including deep learning, have recently been used to develop SIC retrieval models from various remote sensing data [31–36]. For the development of machine learning models, SIC values from ice charts, high-resolution satellites, airborne images, and in situ observations are used as a reference dataset, and various

remote sensing-derived parameters (i.e., backscattering and texture features from synthetic aperture radar (SAR), reflectance from optical sensors, T_B values from PM sensors, etc.) are used as input variables. Thanks to the increase in the available remote sensing data, machine learning models can be developed by learning patterns from vast amounts of training data, and they can show good potential in SIC estimation. In particular, satellite SARs that are hardly affected by weather conditions and sun altitudes but provide high-resolution (a few meters to tens of meters) images have been effectively used to develop SIC estimation models in summer based on machine learning approaches. Han et al. [37] developed a new method for the classification of sea ice and open water in summer from the Korean Multi-purpose Satellite-5 (KOMPSAT-5) X-band SAR images by implementing Random Forest, a rule-based machine learning approach, in which the gray level co-occurrence matrix (GLCM) features computed from the SAR imagery were used as input variables for the classification. The values of SIC estimated from the classifications were used as reliable data for the assessment of the PM SIC algorithms [26]. Wang et al. [31] extracted normalized backscattering coefficients from RADARSAT-2 dual-polarimetric SAR (HH and HV) and applied them to a convolutional neural network to estimate SIC during the ice melting season. Karvonen [32] used a combination of Sentinel-1 SAR and AMSR2 features as input variables for multi-layer perceptron (MLP) deep learning and successfully estimated the SIC in winter in the Baltic Sea. The machine learning models for SAR images typically showed higher performance in SIC estimation than the PM-based operational algorithms. However, due to the limitations in the spatial and temporal coverages of satellite SAR, the previously developed machine learning models are not sufficient to produce consecutive SIC products for a wide area.

PM observation data can be used to develop models for temporally continuous SIC retrieval for a wide area. Chi et al. [34] proposed an MLP-based SIC estimation model for AMSR2 T_B values in order to retrieve multi-temporal SIC over the whole Arctic. They reported that the estimated SIC values showed better agreement with moderate resolution imaging spectroradiometer (MODIS) SIC values, which were used as the reference dataset for the model development, than those from the BT and ASI algorithms. However, the SIC estimation model developed by Chi et al. [34] might produce erroneous SIC values when the atmospheric contributions are large enough to contaminate PM-observed T_B values, as their neural networks were trained using the MODIS-derived SIC values obtained only under clear sky conditions. For the retrieval of operational and accurate SIC amounts based on machine learning approaches, a training dataset covering various atmosphere and ice conditions is required—particularly in summer, when the atmospheric effects on PM-measured T_B values are great. If various atmospheric parameters and PM T_Bs are used as a set of input variables for the development of an SIC retrieval model, more accurate SIC values can be retrieved than those estimated by the existing PM SIC algorithms.

The Pacific Arctic Ocean, including the East Siberian Sea, Chukchi Sea, and Beaufort Sea (Figure 1), is the gateway of the Northwest and Northeast Passage. Sea ice in the region exhibits earlier melting and faster retreating in summer than other regions because of continuous heat transportation from the Pacific Ocean through the Bering Strait [38], which is closely linked to local climate change [39]. In the region, operational PM SIC products have been used as important data for ship navigation and climate research. However, the PM SIC products have been reported to be very inaccurate in summer over the region due to the atmospheric contributions to the T_B contamination of sea ice and open water [25,26].

In the present study, we propose a new summer (July to September) daily SIC retrieval model for AMSR2 observation over the Pacific Arctic Ocean based on a machine learning approach by using SAR-derived SIC for various weather and ice conditions and information of the atmosphere from a numerical weather prediction (NWP) model. The machine learning model was developed by implementing Random Forest regression, and its performance was evaluated statistically. The feasibility of the developed machine learning-based SIC retrieval model was evaluated through a performance comparison with the operationally used PM SIC algorithms.
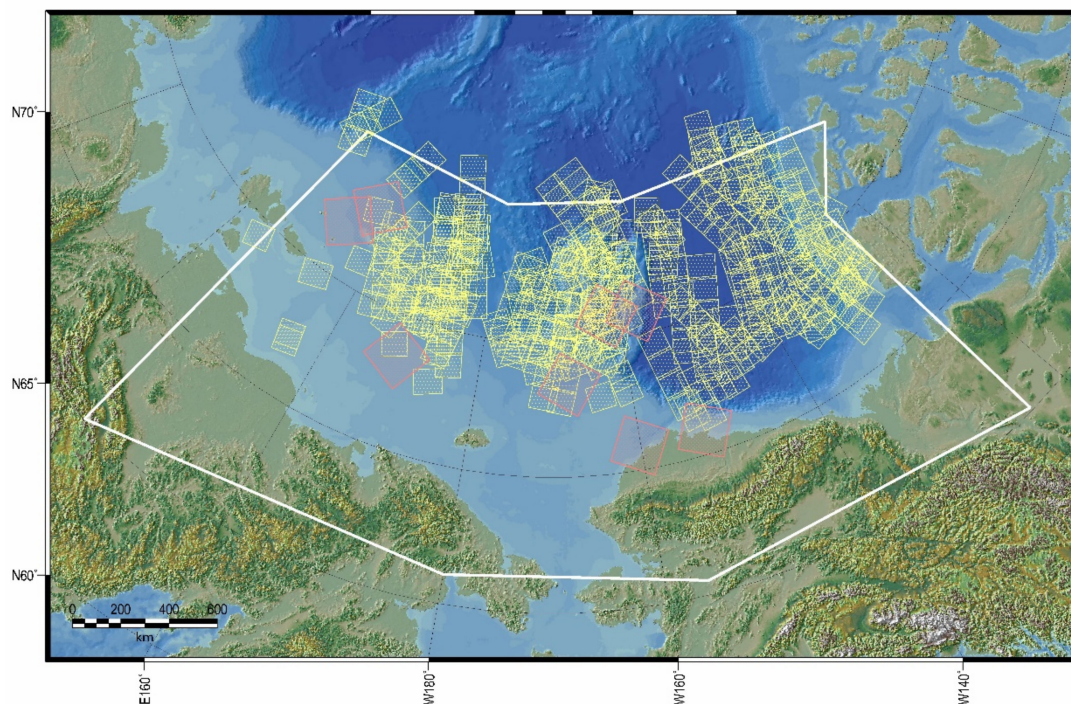
Figure 1. A map of the East Siberian Sea, Chukchi Sea, and Beaufort Sea (the Pacific Arctic Ocean). Yellow and red boxes represent the coverages of KOMPSAT-5 EW SAR and Landsat-8 OLI images used in this study. The white polygon represents the area of the Pacific Arctic Ocean defined by Meier et al. [40].

2. Materials

2.1. AMSR2 Data

AMSR2 is a passive microwave sensor onboard the Global Change Observation Mission–Water (GCOM-W) satellite, launched in 2012, which is a replacement and successor for the Advanced Microwave Scanning Radiometer (AMSR) and Advanced Microwave Scanning Radiometer–Earth Observing System (AMSR-E). AMSR2 is composed of 6.925, 7.3, 10.65, 18.7, 23.8, 36.5, and 89.0 GHz dual-polarized (horizontal (H) and vertical (V) polarization) channels [18,19]. We used AMSR2 Level 3 daily averaged ascending and descending T_B data provided by Japan Aerospace Exploration Agency (JAXA), which are fully calibrated and gridded into a polar stereographic projection with a grid spacing of 10 km by matching the different resolution of each frequency channel by resampling the T_B of L1B swath data. For use in the retrieval of daily SIC values, the AMSR2 daily ascending and descending T_Bs were averaged for each channel every day. The T_B values in the 7.3 GHz channel were not used in this study. The T_B values in the 7.3 GHz channel of AMSR2 are used to detect radio-frequency interference and calibrate the T_B values at 6.9 GHz [41], and are unnecessary information for the development of an SIC retrieval model.

2.2. SAR-Derived Ice/Water Maps

KOMPSAT-5 is equipped with X-band SAR with a center frequency of 9.66 GHz. A total of 454 KOMPSAT-5 SAR images over the Pacific Arctic Ocean were acquired in HH polarization at the Enhanced Wide (EW) swath mode from 6 August to 5 September in 2015, from 8 August to 17 August in 2016, and from 15 July to 25 September in 2017 (Figure 1) and were provided by Korea Aerospace Research Institute (KARI). The KOMPSAT-5 EW SAR images cover an area of 100 km × 100 km with a spatial resolution of 6.25 m (1-look). Sea ice and open water were classified from the KOMPSAT-5 SAR images by using the sea ice mapping model developed by Han et al. [37]. The sea ice mapping model was developed for KOMPSAT-5 EW SAR images in HH polarization based on the classification of GLCM textures of the images by RF, which produces a sea ice map with a grid size of 125 m. The model showed a very high overall classification accuracy and kappa coefficient

(98% and 99%). The overall accuracy is computed by dividing the number of samples that were correctly classified by the total number of samples. The kappa coefficient measures the degree of agreement between classification and reference data considering a change in agreement as occurring by chance, which is widely used as a criterion for the accuracy assessment of classification. The SIC values computed from the SAR-derived ice/water maps were validated by using the Russian Arctic and Antarctic Research Institute (AARI) ice charts. The mean value of the difference between the SIC values from the RF ice/water maps and the ice charts was −8.85%, which could possibly be caused by the uncertainty in the SIC values of the ice charts given as coarse ice concentration categories (10% or 20% SIC increments) in large polygons [37]. Such high performance of the classification is attributed to the striking differences in SAR intensities of ice and water in X-band SAR images. Detailed methodologies for generating ice/water maps from the KOMPSAT-5 EW SAR images are described in Han et al. [37] and Han et al. [26]. We generated ice/water maps from the KOMPSAT-5 EW SAR images by using the sea ice mapping model developed by Han et al. [37] (Figure 2), which was used to compute reference SIC values for training and validating the machine learning based summer SIC retrieval model.
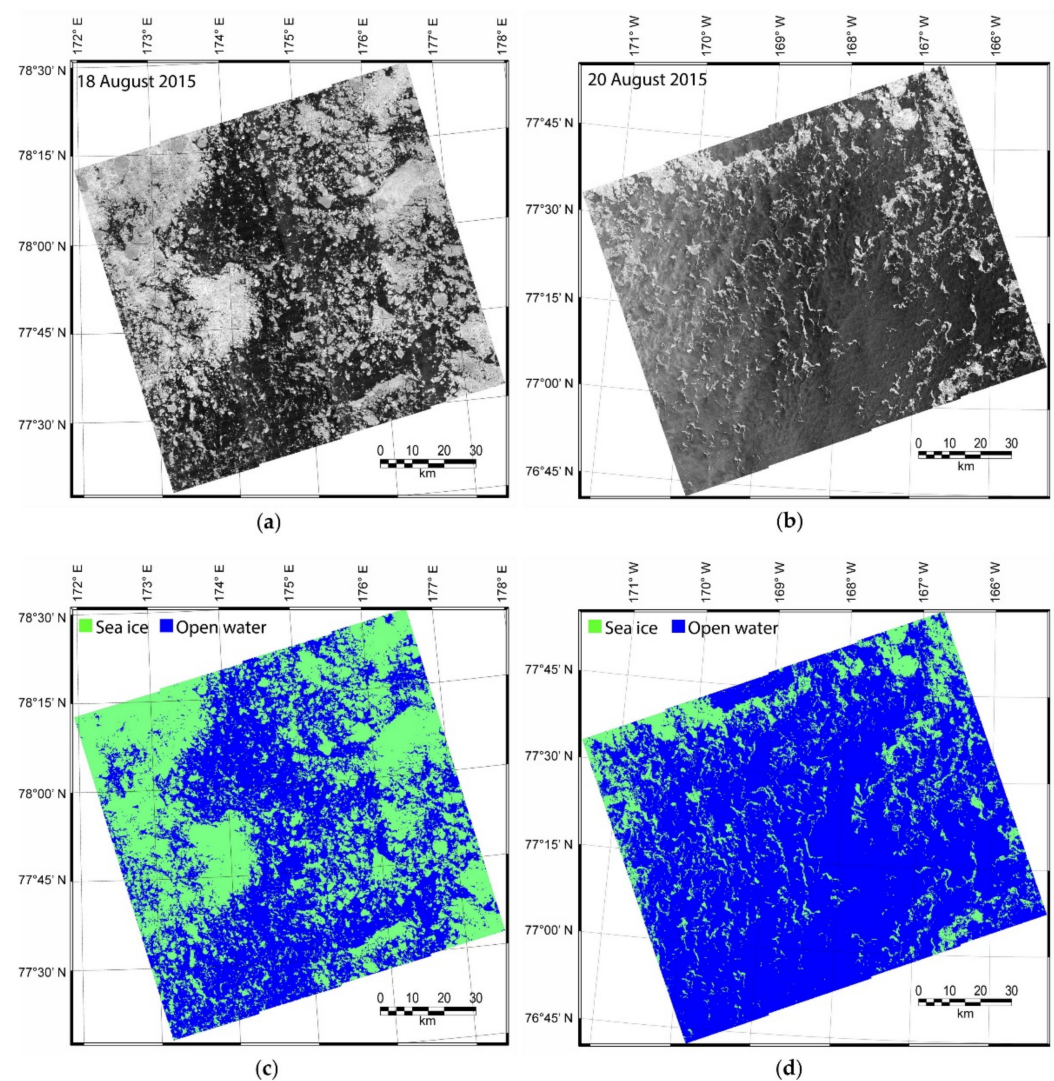

Figure 2. Examples of (**a**,**b**) KOMPSAT-5 EW SAR images and (**c**,**d**) corresponding ice/water maps, modified from Han and Kim [25].

2.3. ERA-5 Reanalysis Data

The ERA-5 reanalysis is generated by the European Centre for Medium-Range Weather Forecasts (ECMWF), and provides multi-decadal reanalysis fields of atmospheric, oceanic, and land variables with a 31 km spatial resolution globally and 137 vertical pressure levels from ground to 0.01 hPa by using a 4D-Var data assimilation system from 1979 to the present [42]. ERA-5 replaced the ERA-Interim, featuring several improvements such as a much higher spatial and temporal resolution, more information about quality assessment and better performance in estimating variables. The reanalysis fields used in this study comprised hourly predicted 2 m air temperature, air temperature at 925 hPa, wind speed at 10 m, and total column water vapor (TCWV) predicted by the ERA-5, which were reported to reflect atmospheric effects on microwave radiation characteristics of sea ice and open water [26]. The air temperature at 925 hPa can indicate the thermal state of the lower troposphere and is appropriate to help assess ice surface melting condition [43]. The hourly reanalysis fields were averaged daily as we used daily averaged AMSR2 T_B values. We also computed the 30-day average of 2 m air temperature and air temperature at 925 hPa from a particular day to account for the ice surface condition by accumulating the effects of air temperature on sea ice surface melting [26]. These reanalysis fields were used as input variables for the machine learning models along with the AMSR2 observation data.

2.4. BT and ASI Sea Ice Concentration Products

The BT and ASI SIC values were assessed by using the reference dataset, and their performance was compared with that of the developed machine learning model. The BT and ASI algorithms are currently operationally used for the AMSR2 T_B values. The BT algorithm utilizes T_B values in 19 V and 37 V channels [1,21]. The gradient difference between 19 V and 37 V channels is effectively used for detecting the seasonal ice area near the ice edge and open water. Therefore, the BT algorithm has the advantage when estimating the SIC of the ice edge. The ASI utilizes 89 GHz dual-polarized channels for SIC estimation [22]. The ASI algorithm estimates SIC with a finer grid spacing than the BT algorithm, as a result of the higher spatial resolution of the 89 GHz channels, however, the 89 GHz measurements are more sensitive to atmospheric water content than lower-frequency channels [22,23]. Sea ice shows small differences between the T_B measured at 89 V and 89 H channels, while open water shows large differences. Moreover, the T_B measured at high frequency channels are less influenced by the snow layer on the ice surface than that at lower frequency channels. We used the daily averaged BT SIC products (with a grid size of 10 km) provided by the JAXA and ASI SIC products (grid size of 6.25 km) provided by the University of Bremen.

2.5. Landsat-8 OLI Images

The Landsat-8 Operational Land Imager (OLI) Level 1 images listed in Table 1 were acquired in the Pacific Arctic Ocean (Figure 1) to evaluate the performance of the developed machine learning model. The dates of Landsat-8 OLI images were different from the dates of the training dataset for the machine learning in order to evaluate the developed model independently. The Landsat-8 OLI provides multispectral imagery in the visible, near infrared, and shortwave infrared bands with a spatial resolution of 30 m for the swath width of 185 km. A panchromatic image with a spatial resolution of 15 m is also captured by the OLI. All the OLI images were obtained under mostly clear sky conditions, radiometrically calibrated, and coordinated in the Universal Transverse Mercator (UTM) projection. From the Landasat-8 OLI panchromatic images, we classified sea ice and open water and computed SIC values, which were compared with those from the developed machine learning model.

Table 1. Details of the Landsat-8 OLI data used in this study.

Date	Path	Row
12 July 2013	77	10
11 August 2014	90	6
11 August 2014	90	8
6 September 2014	105	8
8 July 2015	87	6
13 July 2018	82	10
27 July 2018	173	239
9 August 2018	168	240

3. Methodology

This section describes the construction of the reference dataset and input variables for the development of the machine learning-based summer daily SIC retrieval model. The machine learning approach used in this study and the methods for the evaluation of the model performance are also presented in this section.

3.1. Construction of Reference Dataset and Input Variables for Machine Learning

The ice/water maps generated from the SAR images had a grid size of 125 m and the AMSR2 T_B data had a 10 km grid size. The SIC values were computed in an 80×80 grid cell window of the ice/water maps to produce a grid cell of size 10 km, where the area overlapped with the AMSR2 T_Bs. A total of 35,240 SIC values were computed from the ice/water maps, which were used as a reference dataset for the development of the machine learning-based summer daily SIC retrieval model.

The T_B values at each channel of AMSR2 showed different ranges of values for sea ice and open water due to atmospheric and ice melting conditions. In many PM SIC retrieval algorithms, parameters based on the ratios of T_Bs are adopted to estimate SIC [1,20–22,44]. The representative parameters of the T_B ratios used for the SIC estimation were the polarization ratio (*PR*) at 18 GHz (*PR*(18)), the spectral gradient ratios (*GR*) between 37 GHz and 18 GHz at vertical polarization (*GR*(37V18V)) and 23 GHz and 18 GHz at vertical polarization (*GR*(23V18V)), and the difference between *GR*(89H18H) and *GR*(89V18V) (ΔGR), which are computed as

$$PR(18) = \frac{TB(18V) - TB(18H)}{TB(18V) + TB(18H)} \tag{1}$$

$$GR(f_1 p f_2 p) = \frac{TB(f_1 p) - TB(f_2 p)}{TB(f_1 p) + TB(f_2 p)} \tag{2}$$

$$\Delta GR = GR(89H18H) - GR(89V18V) \tag{3}$$

where f is the frequency and p is the polarization of the PM channel. The *PR*(18) and *GR*(37V18V) are used for the discrimination of sea ice types (first-year ice and multiyear ice in the Arctic) and open water, as implemented in the NT algorithm [19]. The difference in T_B measured at 18 V and 18 H channels is greater in open water than in sea ice. Therefore, open water has a higher *PR*(18) value than sea ice, which is used to distinguish sea ice from open water. *GR*(37V18V) is useful for distinguishing first-year sea ice from multiyear sea ice, because the difference in T_B measured at 37 V and 18 V channels is very small, close to zero for first-year ice, whereas it is negative in multiyear ice. In open water, T_B measured on the 37 V channel is higher than that measured on the 18 V channel, and *GR*(37V18V) of open water is calculated as a positive value. The ΔGR is used in the enhanced NT (NT2) algorithm and enables the identification of sea ice with an inhomogeneous surface layer such as surface glaze and layering based on a decreasing T_B at 18 GHz and stable 89 GHz channels by increasing inhomogeneity of the surface layer [44]. The *GR*(23V18V) is used in conjunction with *GR*(37V18V) to correct the influences of weather on SIC estimation from PM SIC algorithms such as NT, NT2, and ASI [1,20–22,44]. The parameters derived

from the T_B ratios have the advantage of being less sensitive to variations in the physical temperature of the snow/ice layer [16].

A total of 24 features were extracted from the AMSR2 (T_B values in vertically and horizontally polarized channels of each frequency, *PR*(18), *GR*(37V18V), *GR*(23V18V), *GR*(89H18H), *GR*(89V18V), and ΔGR) and from the ERA-5 reanalysis (2 m air temperature, air temperature at 925 hPa, wind speed at 10 m, TCWV, and 30-day average of 2 m air temperature and air temperature at 925 hPa) for the same dates and areas with the reference SIC grids for the model development. These features were set to be the independent variables, and the SIC values computed from the SAR ice/water maps were considered to be the dependent variable in the machine learning model. A total of 35,240 samples were constructed, of which 80% (28,192 samples) were used for the training of the machine learning model and the rest (7048 samples) were used to validate the developed model. The training and validation samples were extracted from different ice/water maps, and thus they were independent of each other.

3.2. Random Forest Regression for SIC Retrieval

Random Forest (RF) [45] regression was used to retrieve summer daily SIC from the AMSR2 observation data and ERA-5 reanalysis fields. The RF is a popular rule-based machine learning algorithm due to its good prediction capability for high-dimensional data in remote sensing fields [46]. In the remote sensing of sea ice, the RF has been widely used for the SIC retrieval [36], melt pond fraction estimation [47], ice type classification [37,48–50], and lead detection [51,52]. The RF is an ensemble classifier and regressor that creates multiple bootstrapped samples of the original training data and builds a set of no pruning classification and regression trees (CART) from each set of bootstrapped samples, which is an ensemble of rule-based decision trees [45]. The numerous independent decision trees are created by randomly selecting a subset of the training samples with replacement for each tree and a subset of recursively splitting input variables at each node of the tree. This process can improve the predictability of response variables by reducing the learning dependence on the quality and configuration of training samples [46]. For regression, the RF predicts value at a node by averaging the response variable of all observations in the node. The predictions from the independent decision trees are aggregated (averaged), and then a final conclusion is determined. The RF provides the statistical measure of the relative importance of input variables in terms of the prediction accuracy, called the mean decrease in accuracy, which represents the decrease in the accuracy when the values of variables are randomly permuted.

The performance of the RF-based SIC retrieval model was evaluated in terms of the correlation coefficient (*R*) between the predicted SIC values and the reference SIC values, and the mean bias (mean error), the standard deviation of error (SDE) and the root mean square error (RMSE), in which errors were calculated by subtracting the corresponding predicted SIC values from the corresponding reference values. The BT and ASI SIC values were also assessed using the same measures. Prior to the assessment, the SIC products were resampled to 10 km using the nearest-neighbor scheme to match the grid size of the reference SIC.

The accuracy of SIC retrieved from the RF model, ASI, and BT algorithms was also evaluated by using the Landsat-8 OLI panchromatic images. According to Cavalieri et al. [53], the broadband albedo (ρ) of the panchromatic image of Landsat-7 Enhanced Thematic Mapper Plus (ETM+) with a wavelength range of 0.52–0.90 μm is useful to classify surface types into open water ($\rho < 0.1$), new ice ($0.1 \leq \rho < 0.4$), young ice ($0.4 \leq \rho < 0.6$), and first-year ice ($0.6 < \rho$). The SIC values calculated from this ice/water classification based on the thresholds of the albedo were effectively used for the assessment of the AMSR-E NT2 SIC product for the Arctic [54]. In Cavalieri et al. [54], the broadband albedo from the three visible bands of MODIS (0.46–0.67 μm) was computed using different weights for each band [55] and classified into open water and sea ice types by using the same thresholds proposed in Cavalieri et al. [53]. The MODIS broadband albedo covers approximately the

same spectral range as the ETM+ panchromatic band; thus SIC values calculated based on the albedo thresholds for the classification of sea ice and open water can be used for the AMSR-E NT2 SIC product in the Antarctic winter [54].

We obtained the top-of-the-atmosphere reflectance of the OLI panchromatic images and resampled it to a 100 m grid for ease of handling. Then, the albedo images were reprojected into the polar stereographic projection. Although the wavelength range of the OLI panchromatic band (0.50–0.68 μm) was almost the same as that of the MODIS broadband albedo derived in Cavalieri et al. [54], we used the albedo of 0.15 as a threshold to separate water from ice. Through visual inspection of the classification results, it was found that the open water observed in the OLI panchromatic images was misclassified as new ice when the albedo threshold was set to 0.1, while most was correctly classified as sea ice with an albedo threshold of 0.15. In the visible wavelength (<0.7 μm), new ice can have an albedo value larger than 0.11 [56]. We also performed the visual inspection on the classification results when different albedo thresholds were applied. It was confirmed that the misclassification when using the albedo threshold values of 0.13 and 0.17 was greater than when the albedo threshold of 0.15 was used. However, the classifications with albedo threshold values of 0.14 and 0.16 were not visually different from when the albedo threshold value of 0.15 was used. The changes in the open water fraction in the Landsat-8 images were only 1% for every 0.1 increment of the albedo threshold from 0.14 to 0.16. Figure 3 shows the classification result of the Landsat-8 OLI panchromatic albedo on 8 August 2018, which demonstrates that the separation of open water from sea ice is clearer with an albedo threshold of 0.15 than 0.10. The grid cells with a panchromatic albedo value greater than 0.15 were classified into sea ice, and others were defined as open water. The grids of the albedo covered by clouds were removed before the classification using the quality assessment band of the Landsat products. From the classified images, SIC values were calculated with a 10 km grid, over which the SIC results retrieved from the RF model and ASI and BT algorithms were overlapped. We calculated the values of *R*, mean bias, SDE, and RMSE of the RF, ASI, and BT SICs by comparing them with the Landsat-8 OLI SICs.
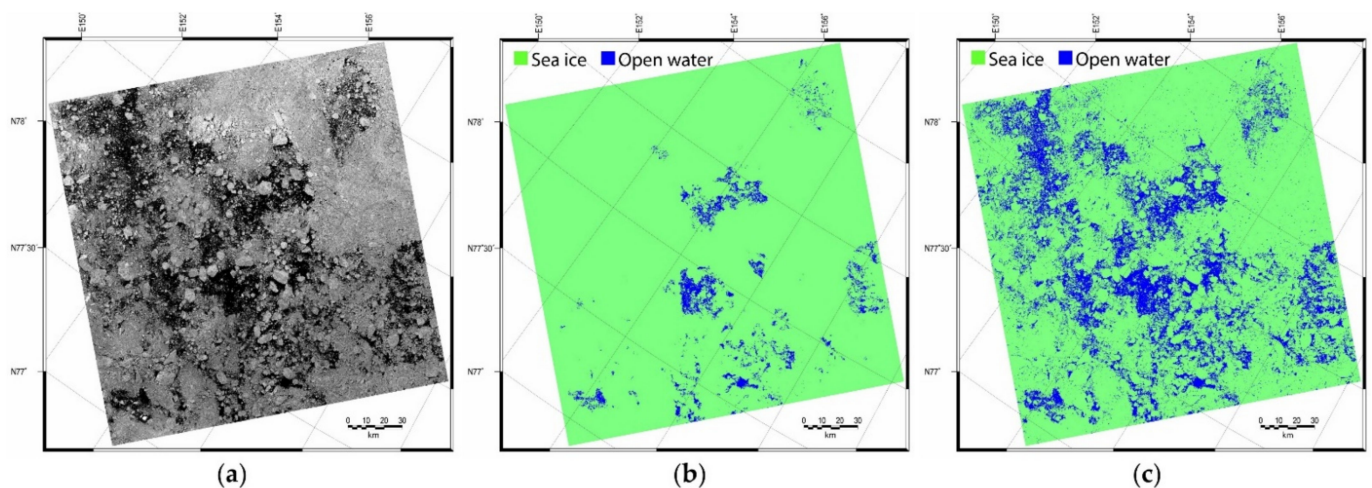

Figure 3. An example of (**a**) a Landsat-8 OLI panchromatic image resampled to a grid of 100 m and corresponding ice/water classification map, in which the open water was classified from albedo values of less than (**b**) 0.1 and (**c**) 0.15.

4. Results and Discussion

Table 2 shows the descriptive statistics of the samples used for the development of a summer SIC retrieval model by RF regression. The samples were extracted under a variety of weather conditions, which might help the RF model to produce SIC values without relying on the atmospheric contamination of the AMSR2 observations.

Table 2. Descriptive statistics of the input variables used in the training and validation samples for the RF model (T_B: brightness temperature, *H*: horizontally polarized channel, *V*: vertically polarized channel, Q1: 25th percentile, Q3: 75th percentile).

Variable / Statistics	Mean	Median	Std.	Min.	Max.	Q1	Q3
T_B 6H (K)	168.23	171.34	38.94	87.87	241.45	133.82	200.65
T_B 6V (K)	218.23	222.73	24.40	165.68	259.93	196.63	238.40
T_B 10H (K)	173.73	178.11	37.72	96.00	240.02	140.25	206.11
T_B 10V (K)	223.53	228.48	22.42	175.06	259.60	203.72	242.84
T_B 18H (K)	183.35	189.06	32.29	113.03	241.38	155.61	211.49
T_B 18V (K)	230.08	234.0	17.11	190.47	258.33	215.35	245.09
T_B 23H (K)	198.75	203.93	26.33	131.93	250.18	176.80	220.52
T_B 23V (K)	236.13	237.41	13.70	202.38	263.07	224.65	248.38
T_B 36H (K)	195.44	198.90	24.23	141.35	250.18	175.15	215.40
T_B 36V (K)	232.28	231.83	12.73	194.31	262.47	222.15	242.47
T_B 89H (K)	220.86	218.88	16.06	179.16	268.38	209.31	231.33
T_B 89V (K)	243.05	243.66	12.18	194.90	272.15	236.61	250.06
*PR*18	1.12	1.11	0.06	1.03	1.26	1.07	1.16
GR(36V18V)	1.01	1.01	0.02	0.91	1.06	0.99	1.02
GR(23V18V)	1.01	1.01	0.01	0.98	1.05	1.00	1.02
GR(89H18H)	2.07	2.06	0.03	2.00	2.17	2.04	2.10
GR(89V18V)	1.10	1.09	0.08	0.92	1.29	1.04	1.16
ΔGR	1.03	1.03	0.05	0.88	1.13	0.99	1.07
TCWV (kg/m^2)	11.99	11.35	3.45	3.83	29.28	9.98	13.34
Wind speed (m/s)	4.88	4.68	2.44	0.04	13.09	2.94	6.24
2 m air temperature (°C)	−0.28	−0.12	1.73	−10.81	6.78	−0.81	0.70
925 hPa air temperature (°C)	−0.32	−0.58	4.25	−10.98	12.98	−3.10	2.08
30-day average of 2 m air temperature (°C)	0.63	0.69	1.02	−3.88	5.57	0.10	1.30
30-day average of 925 hPa air temperature (°C)	1.07	0.31	3.01	−5.87	8.11	−1.22	3.72
Reference SIC (%)	61.06	71.13	34.95	0.00	100.00	26.44	95.36

4.1. Performance of Summer SIC Retrieval Model Based on RF Regression

The summer daily SIC retrieval model for the Pacific Arctic Ocean was developed using RF regression. Figure 4 shows the scatterplots of the RF model-derived SIC values and reference SIC values for the training and validation datasets. For the training dataset, the RF SIC values matched well with the reference values with a very small RMSE (3.41%), mean bias (0.02%) and SDE (3.41%) and a very high *R* value (0.992). The RF SIC values for the validation dataset, which was selected independently of the training dataset, were also strongly correlated with the reference SIC values (*R* value of 0.959), showing small values of RMSE (7.89%), bias (0.11%), and SDE (7.89%). The values of RMSE and SDE for the validation dataset were about twice as great as those for the training dataset, but they were still small and the value of the mean bias was close to 0%. The performance of the developed RF model was greatly superior to that of the ASI and BT algorithms (Figure 5). The BT and ASI SIC values showed lower *R* values (0.876 and 0.864, respectively) compared to the reference SIC values for the validation dataset, and they had three times greater values of RMSE (20.19% and 21.39%, respectively) and SDE (19.11% and 19.67%, respectively) than the RF SIC values. Furthermore, the ASI and BT algorithms overestimated SIC values with a mean bias of 6.49% and 8.40%, respectively.

The deviations of the ASI and BT SIC values from the reference values were mainly caused by the atmospheric effects on the T_B values of sea ice and open water observed by AMSR2 [26]. Even though the developed RF model uses the same observations, it can retrieve more accurate SIC values by reducing the physical change of the T_B values of sea ice and open water due to atmospheric effects by considering the weather conditions predicted by the NWP model.
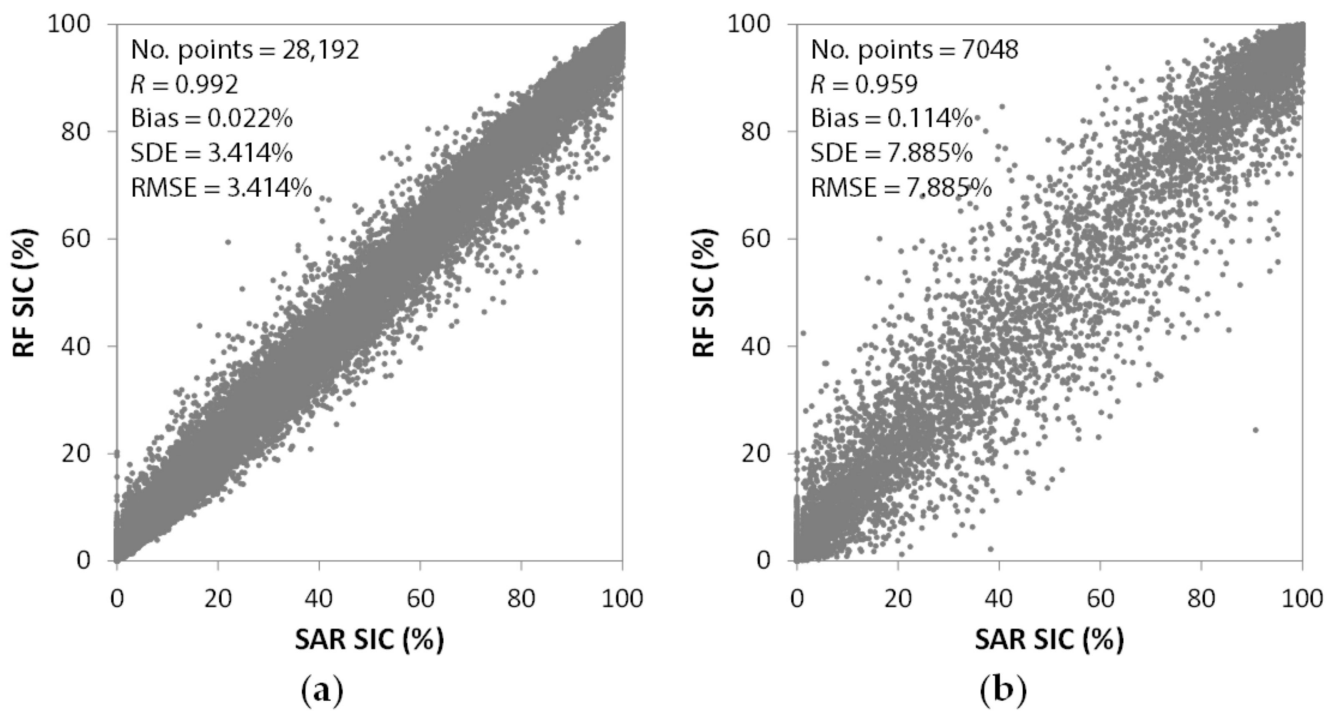
Figure 4. Comparisons of SIC values computed from the KOMPSAT-5 ice/water maps with those predicted by the RF model for the (**a**) training and (**b**) validation datasets.
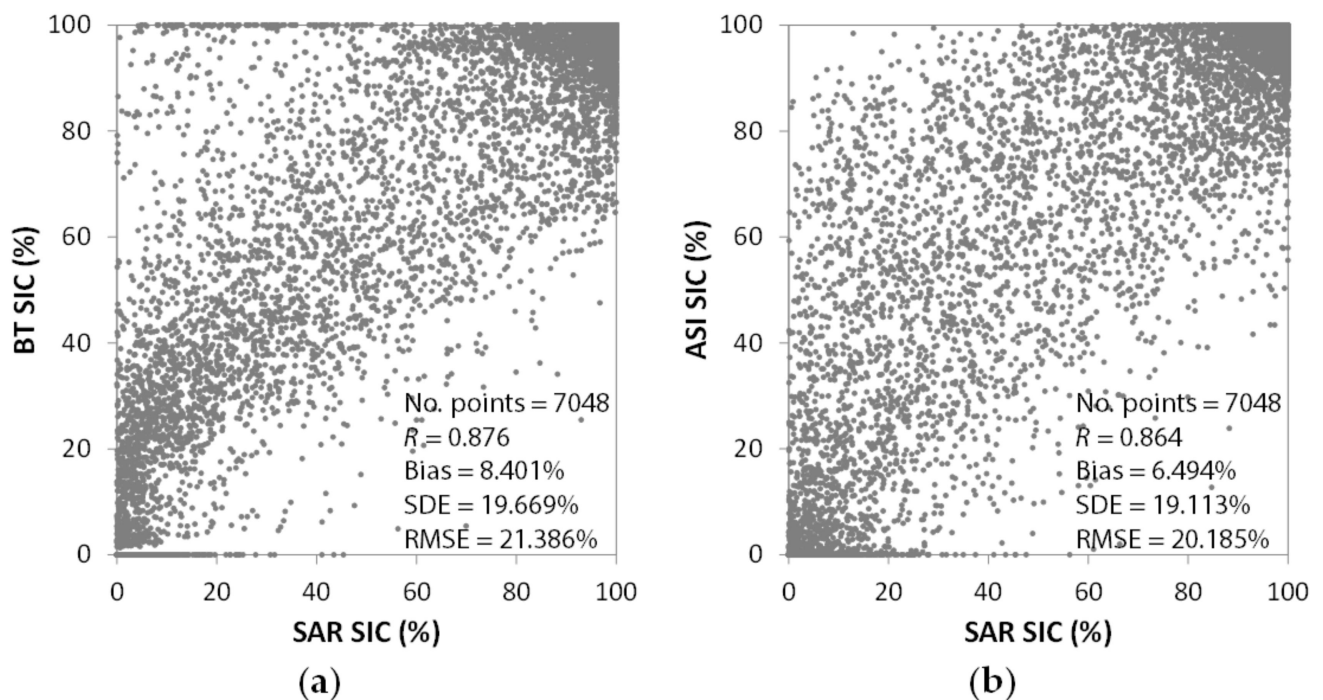

Figure 5. Comparisons of SIC values computed from the KOMPSAT-5 ice/water maps with those from the (**a**) BT and (**b**) ASI algorithms.

Summer SIC mapping results from the RF model were compared to the reference SIC maps, which were used as the validation dataset (Figures 6 and 7). The BT and ASI SIC maps were also compared with the reference maps. From the SAR intensity images (Figures 6a and 7a) and ice/water maps (Figures 6b and 7b), we could confirm that the accurate classification of sea ice and open water was performed by the ice/water mapping model developed by Han et al. [37]. For the comparison on 18 August 2015 (Figure 6),

the RF model, BT, and ASI had RMSE values of 9.14%, 12.96%, and 16.81%, respectively. The BT and ASI SIC values were largely overestimated in the low SIC region (Figure 6h,i), where the TCWV was higher than 11 kg/m^2 (Figure 6j), which is a main cause of the overestimation of SIC by the BT and ASI algorithms [26]. In the high SIC region, the ASI SIC values were largely underestimated (Figure 6i) due to ice surface melting, which could be confirmed by the low backscattering of sea ice (Figure 6a) and the averaged 2 m air temperature being above 0 °C for 30 days (Figure 6l). The 30-day average of 2 m air temperature in the north, where there was sea ice, was higher than in the south, where there was open water (Figure 6l). Sea ice continues to move, and the 30-day average of 2 m air temperature could be higher in the sea ice region than open water region on a particular day. The differences between the RF-retrieved SIC values and the reference values were mostly less than 10% (Figure 6g), regardless of atmospheric and ice surface melting conditions.

In a small patch of ice surrounded by open water south of the ice edge with a strongly wind-roughened ocean on 4 September 2015 (Figure 7o), the RF model, BT, and ASI had a RMSE values of 9.17% (Figure 7g), 12.74% (Figure 7h), and 15.63% (Figure 7i), respectively. The BT SIC values were greatly overestimated (Figure 7h), which could be caused by the effect of the water content and the surface roughness of open water. High atmospheric water vapor and strong winds over the open water increase T$_B$ [25], which could be a reason for the overestimation of SIC by the algorithm. The ASI algorithm estimated SICs slightly more accurately than the BT, which could be attributed to the higher spatial resolution of the 89 GHz channel than others, but it still overestimated SICs for fragmented ice and underestimated those for some ice floes (Figure 7i) showing low backscattering due to surface melting. The RF model-retrieved SIC values were much more accurate than the BT and ASI SIC values, showing small deviations from the reference values (Figure 7g). These comparisons showed that the RF model can retrieve accurate summer SIC values under various atmospheric and ice surface melting conditions by considering weather conditions in its learning process.

We additionally tested the RF SIC model using the SIC values computed from the Landsat-8 OLI panchromatic images. The BT and ASI SIC values were also compared with the Landsat-8 SIC values. Figure 8 shows the comparisons of Landsat-8 SIC values with the RF, BT, and ASI-derived SIC values. RF-based SIC values matched well with the Landsat-8 SIC values with small values of RMSE (9.21%), mean bias (−2.07%) and SDE (8.97%), and a very high *R* value (0.955) (Figure 8a). The *R* values for the Landsat-8 SIC values and the BT and ASI SIC values were also high (0.879 and 0.924, respectively). However, the algorithms estimated SIC values with a larger RMSE, mean bias, and SDE than the RF model (Figure 8b,c), which shows the superior performance of the RF model for SIC estimation. The ASI values were underestimated compared to the Landsat-8 SIC values, which was possibly caused by ice surface melting due to the very high temperature at 925 hPa (Table 3). The ASI algorithm produced SIC values close to 0% when the Landsat-8 SIC values less than 20%, i.e., near the ice edge. This is the same result as reported in Radhakrishnan et al. [57], presenting that the ASI algorithm tends to underestimate SIC in the vicinity of the ice edge. Meanwhile, the BT algorithm estimated slightly positively biased SIC values. This might be attributed to the fact that the underestimation of SIC from the algorithm due to ice surface melting may be compensated by the water vapor content [26]. Figure 8 confirms once again that the BT and ASI algorithms estimate SIC values inaccurately in summer due to the contaminated T$_B$ of sea ice and open water caused by atmospheric effects, while the RF model developed in this study retrieves more accurate SIC values by considering the atmospheric effects through the NWP data.
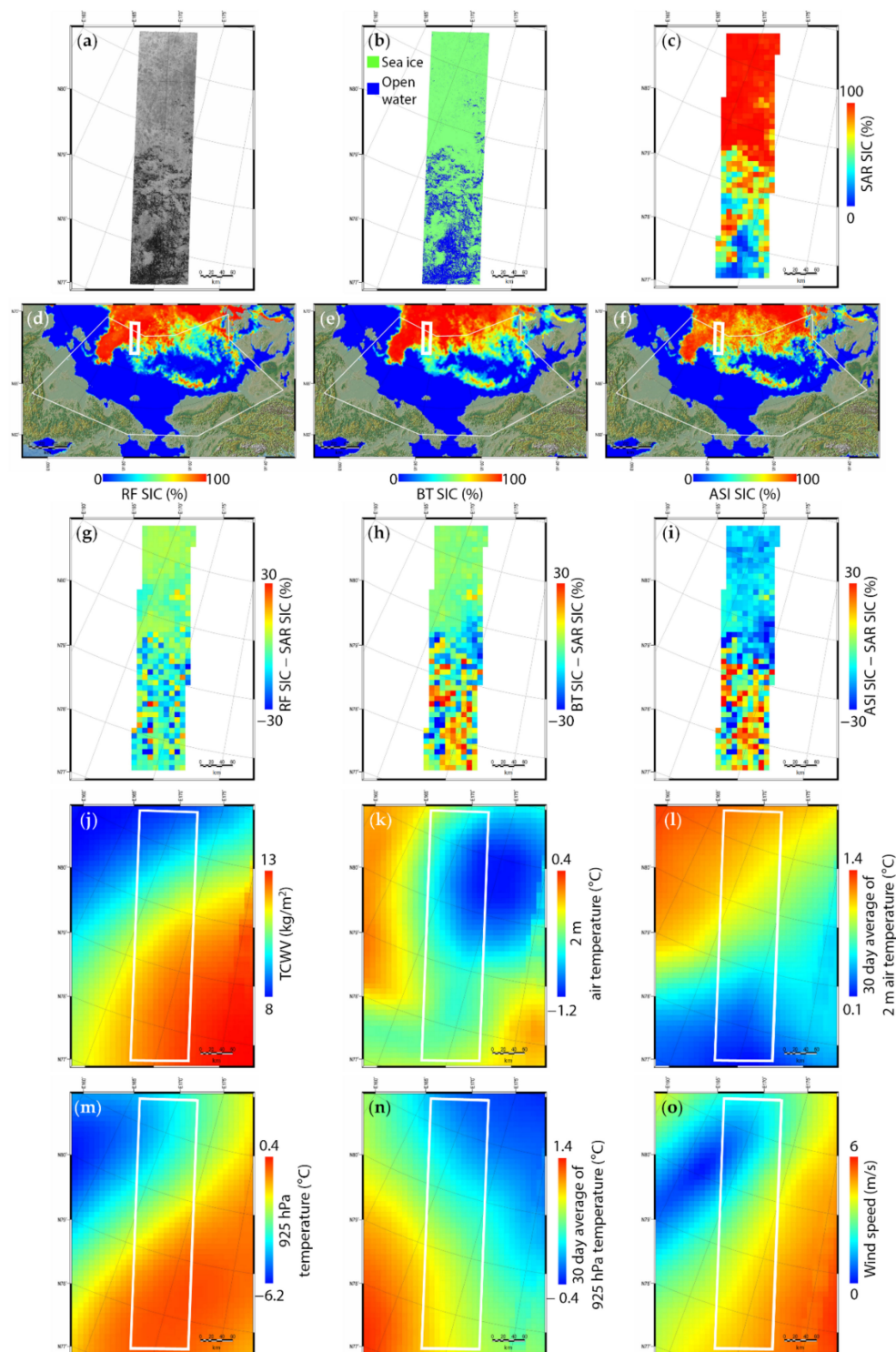
Figure 6. (**a**) KOMPSAT-5 EW SAR image acquired on 18 August 2015, and corresponding (**b**) ice/water map and (**c**) SIC with a grid size of 10 km. (**d**–**f**) RF, BT and ASI SIC map for the Pacific Arctic Ocean (the white polygon) on the same date as the SAR image. The difference between SAR SIC and (**g**) RF, (**h**) BT, and (**i**) ASI SIC for the same area of the SAR image (the white box in (**d**–**f**)). For the same date, (**j**) TCWV, (**k**) 2-m air temperature, (**l**) 30-day average of 2 m temperature, (**m**) air temperature at 925 hPa, (**n**) 30-day average of air temperature at 925 hPa and (**o**) wind speed. The white box in (**j**–**o**) represents the area of SAR image.
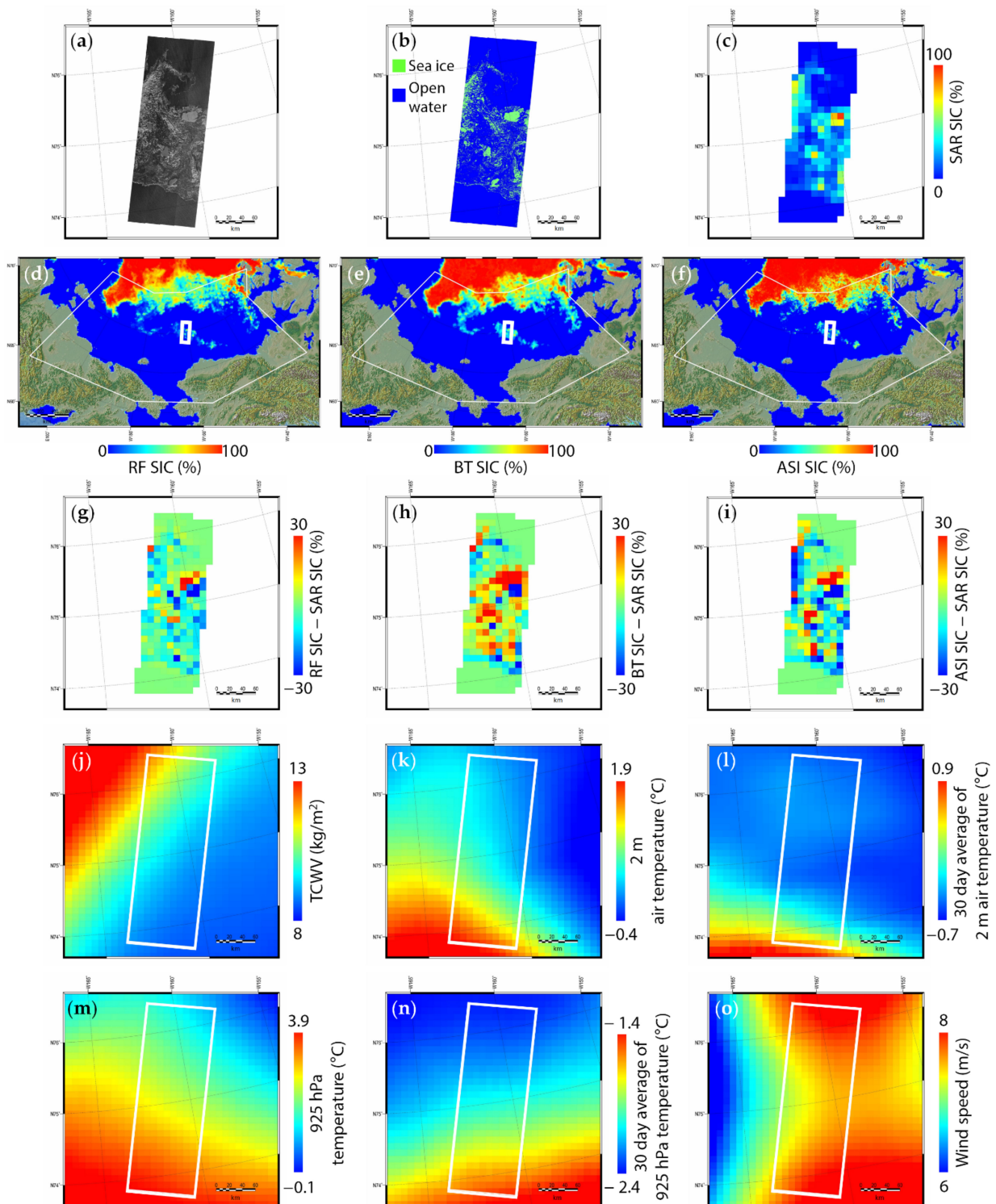
Figure 7. (**a**) KOMPSAT-5 EW SAR image acquired on 4 September 2015, and corresponding (**b**) ice/water map and (**c**) SIC with a grid size of 10 km. (**d**–**f**) RF, BT, and ASI SIC map for the Pacific Arctic Ocean (the white polygon) on the same date as the SAR image. The difference between SAR SIC and (**g**) RF, (**h**) BT, and (**i**) ASI SIC for the same area of the SAR image (the white box in (**d**–**f**)). For the same date, (**j**) TCWV, (**k**) 2-m air temperature, (**l**) 30-day average of 2 m temperature, (**m**) air temperature at 925 hPa, (**n**) 30-day average of air temperature at 925 hPa and (**o**) wind speed. The white box in (**j**–**o**) represents the area of SAR image.
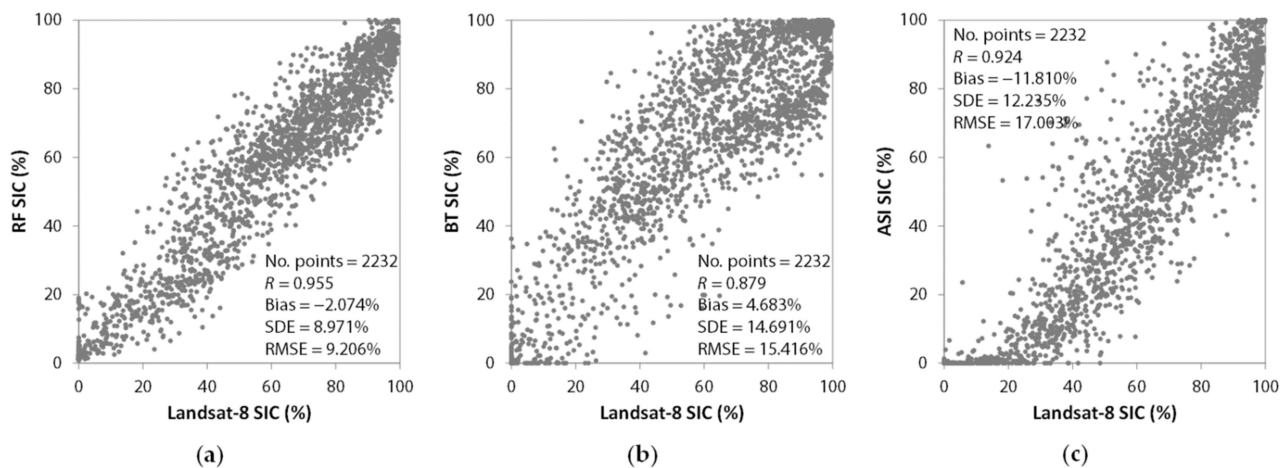
Figure 8. Comparisons of SIC values computed from Landsat-8 ice/water maps with those from the (**a**) RF model and (**b**) BT and (**c**) ASI algorithms.

Table 3. Descriptive statistics of atmospheric conditions used in the RF SIC retrieval corresponding to the Landsat-8 SIC (T_B: brightness temperature, H: horizontally polarized channel, V: vertically polarized channel, Q1: 25th percentile, Q3: 75th percentile).

Variable / Statistics	Mean	Median	Std.	Min.	Max.	Q1	Q3
TCWV (kg/m^2)	14.19	13.99	3.07	10.54	26.02	11.01	16.15
Wind speed (m/s)	5.67	5.76	1.26	3.03	8.39	4.84	6.60
2 m air temperature (°C)	0.65	0.58	0.49	−0.28	4.05	0.26	0.9
925 hPa air temperature (°C)	3.84	5.44	3.42	−3.54	7.74	0.88	6.19
30 days average of 2 m air temperature (°C)	1.09	0.75	0.87	−0.09	7.62	0.55	1.46
30 days average of 925 hPa air temperature (°C)	2.43	2.01	2.71	−1.06	9.96	−0.09	3.25

4.2. Variable Importance of the RF Model

The relative importance of the 10 most important input variables used in the development of the summer SIC retrieval model based on RF regression is shown in Figure 9. The most important variable is the 30-day average of the temperature at 925 hPa, and the second and third most important variables are the temperatures at 925 hPa and 2 m, respectively. All three of the most important variables are related to air temperature and account for the ice surface melting condition in summer. The air temperature at 925 hPa indicates the thermal state of the lower troposphere and that at 2 m potentially reflects the thermal properties of the ice surface [43]. The ice surface melting in summer makes it difficult to distinguish between sea ice and open water from the T_B values observed by the PM sensors and has a great influence on the underestimation of the SIC [16,26,58]. In particular, when high air temperatures are maintained for a long time (e.g., for a month), the ice surface melting becomes more severe. Therefore, it can be concluded that the three most important variables related to the air temperature have been used to compensate for the effect of ice surface melting included in the PM observations in the SIC retrieval by the RF model. The average of the air temperature at 2 m for 30 days was the sixth most important variable of the model.

The GR(23V18V), TCWV, and GR(36V18V) were analyzed as the fourth, fifth and seventh most important variables in the RF model, respectively. In the existing PM SIC algorithms, the GRs are used to prevent the overestimation of SIC caused by the microwave radiation characteristics of open water being similar to those of sea ice due to precipitation, water vapor, and the roughening of the ocean surface by winds [59]. It has been found

that as the amount of TCWV increases, the overestimation of the SIC from the existing PM SIC algorithms becomes greater [25,26]. The summer SIC retrieval model developed in this study can be characterized as correcting the false detection of sea ice from open water attributed to the weather conditions by using the *GR*(23V18V), TCWV, and *GR*(36V18V) as important input variables. The wind speed is the 10th most important variable, and it contributes to SIC retrieval by quantifying the effect of a wind-roughened ocean surface.
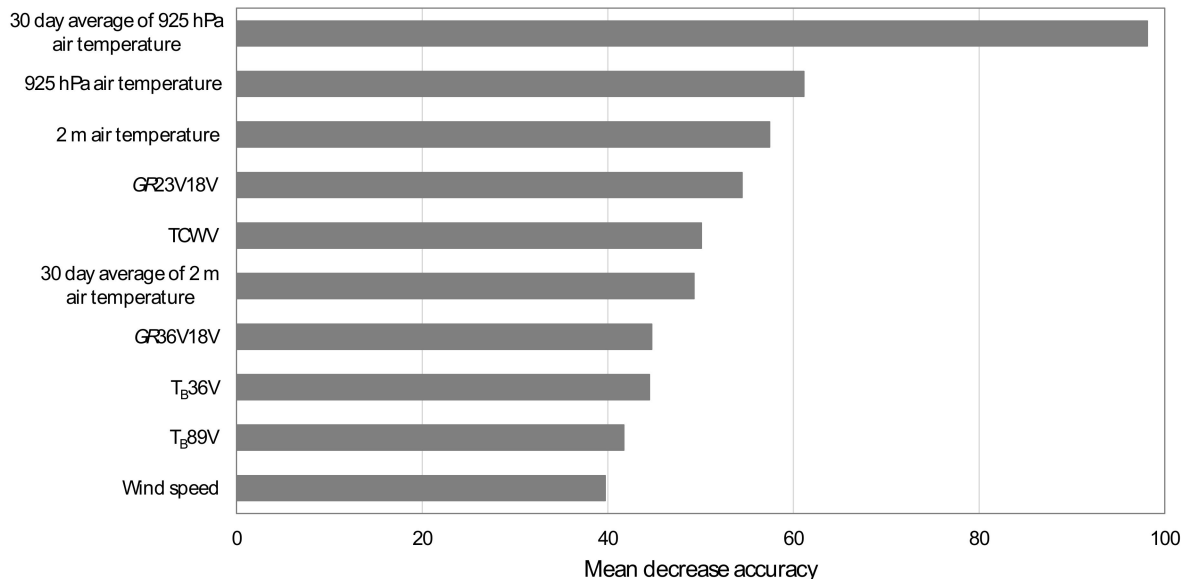

Figure 9. Mean decrease in accuracy of the RF model.

In the developed model, most variables constructed from the NWP data were more important than PM observation variables. This proves that an accurate summer SIC can be estimated by considering information on atmospheric effects together with the PM observations. In addition, it can be concluded that by inputting variables representing the atmospheric effects into the machine learning approach, it is possible to retrieve a more accurate summer SIC without the hassle of determining the weather filters based on the PM observations and T_B threshold values.

4.3. Implications for the Machine Learning Model

Sea ice in the Pacific Arctic Ocean melts rapidly in summer and the ice edge retreats dramatically, and the summer SIC in this region has been used as an important indicator of climate change. In addition, as it is the entrance of the Northwest and Northeast Passage, the accurate estimation of SIC in summer of this region is necessary. The machine learning model developed in this study adopted the AMSR2 observation data and NWP fields in combination to retrieve the summer SIC to consider the physical changes in the T_B values of sea ice and open water caused by atmospheric effects, from which more accurate SIC values were obtained than the BT and ASI algorithms. The machine learning model also has the advantage of retrieving accurate SIC values without the complication of using weather filters based on the complex thresholds of T_B values.

In the machine learning model, the daily averaged T_Bs and NWP fields were used as input variables. In summer, sea ice can move quickly and the weather can vary with time. Therefore, the use of the daily averaged data may lead to inaccurate SIC retrieval, which can be particularly highlighted at the ice edge. The ice surface melting can be affected not only by air temperature, but also by the surface energy balance components such as downwelling fluxes of longwave and shortwave radiation, latent heat flux, and heat release from the ocean. However, the energy balance components were not used in the model development. Moreover, the 2 m air temperature of ERA-5, identified as one of the important variables for retrieving summer SIC by the RF model, has a warm bias (<1 °C)

over Arctic sea ice relative to buoy observations in summer [60]. This shows that the ERA-5 near-surface temperature may not be suitable as a physical parameter for considering the sea ice melting, although it has been used as an important measure for estimating SIC through the machine learning model. Another limitation of the machine learning model developed in this study is that real-time SIC estimation is not possible because the quality-assured ERA-5 reanalysis fields are published tens of days later than real time [61]. In the machine learning model, NWP reanalysis fields were used as important variables for SIC retrieval. However, the near-real time SIC estimation would be possible if forecast fields from the real time forecast models such as the Integrated Forecasting System (IFS) provided by the ECMWF or United Model provided by the UK Met office (Global UM) were used in the model development. In future research, we will develop a more accurate summer SIC retrieval model by using individual AMSR2 swath data and real time forecast fields, including the surface energy balance components of the corresponding time, to the AMSR2 observations.

5. Conclusions

A summer SIC retrieval model for the Pacific Arctic Ocean was developed by using AMSR2 observation data and ERA-5 reanalysis fields based on a rule-based machine learning approach—RF regression. The reference dataset for the retrieval of SIC was constructed from the ice/water classification maps generated from the KOMPSAT-5 SAR images acquired from July to September in 2015–2017. The T_B values of the AMSR2 channels, the ratios of T_Bs (PR and GRs) and the NWP fields presenting water vapor content, air temperature, and wind speed were used as input variables of the machine learning model. The RF model retrieved more accurate SIC values than the BT and ASI algorithms under various atmospheric and ice surface melting conditions. The BT and ASI algorithms produced SIC values that had large errors due to the atmospheric effects, but the SIC values retrieved from the RF model had much smaller errors. The air temperatures at 2 m and 925 hPa, as well as their 30-day averages, GR(23V18V), TCWV, and GR(36V18V), were identified as more significantly contributing input variables than others by the RF model. The variables related to air temperature and the other important variables contributed to the RF model retrieving accurate SIC by taking into account the changes in the T_B values of sea ice and open water caused by ice surface melting and weather conditions.

The machine learning model proposed in this study successfully retrieved the SIC in the Pacific Arctic Ocean in summer by considering the atmospheric effects on the T_B values of sea ice and open water, and it can be used to reconstruct sea ice information with much higher accuracy than the operational algorithms for passive microwave sensors. However, the machine learning model may retrieve inaccurate SIC values in regions with severe spatiotemporal variations in sea ice and weather conditions. Future research will include developing a more improved SIC retrieval model by using individual swath data from passive microwave sensors and real time weather information including the surface energy balance components for the corresponding time of the swath data for the entire Arctic Ocean.

Author Contributions: Conceptualization, H.H. and S.L.; methodology, H.H. and M.K.; validation, H.H.; formal analysis, H.H., S.L. and M.K.; investigation, H.H.; resources, H.H.; data curation, H.H., S.L. and H.-C.K.; writing—original draft preparation, H.H.; writing—review and editing, H.H.; visualization, H.H. and S.L.; supervision, H.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

References

1. Comiso, J.C.; Cavalieri, D.J.; Parkinson, C.L.; Gloersen, P. Passive microwave algorithms for sea ice concentration: A comparison of two techniques. *Remote Sens. Environ.* **1997**, *60*, 357–384. [CrossRef]
2. Johannessen, O.M.; Bengtsson, L.; Miles, M.W.; Kuzmina, S.I.; Semenov, V.A.; Alekseev, G.V.; Nagurnyi, A.P.; Zakharov, V.F.; Bobylev, L.P.; Pettersson, L.H.; et al. Arctic climate change: Observed and modelled temperature and sea-ice variability. *Tellus A* **2004**, *56*, 328–341. [CrossRef]
3. Sun, L.; Alexander, M.; Deser, C. Evolution of the global coupled climate response to Arctic sea ice loss during 1990–2090 and its contribution to climate change. *J. Clim.* **2018**, *31*, 7823–7843. [CrossRef]
4. Ogawa, F.; Keenlyside, N.; Gao, Y.; Koenigk, T.; Yang, S.; Suo, L.; Wang, T.; Gastineau, G.; Nakamura, T.; Cheung, H.N.; et al. Evaluating impacts of recent Arctic sea ice loss on the northern hemisphere winter climate change. *Geophys. Res. Lett.* **2018**, *45*, 3255–3263. [CrossRef]
5. Vinnikov, K.Y.; Robock, A.; Stouffer, R.J.; Walsh, J.E.; Parkinson, C.L.; Cavalieri, D.J.; Mitchell, J.F.B.; Garrett, D.; Zakharov, V.F. Global warming and northern hemisphere sea ice extent. *Science* **1999**, *286*, 1934–1937. [CrossRef] [PubMed]
6. Kay, J.E.; Holland, M.M.; Jahn, A. Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world. *Geophys. Res. Lett.* **2011**, *38*, L15708. [CrossRef]
7. Yadav, J.; Kumar, A.; Mohan, R. Dramatic decline of Arctic sea ice linked to global warming. *Nat. Hazards* **2020**, *103*, 2617–2621. [CrossRef]
8. Pörtner, H.O.; Roberts, D.C.; Masson-Delmotte, V.; Zhai, P.; Tignor, M.; Poloczanska, E.; Mintenbeck, K.; Nicolai, M.; Okem, A.; Petzold, J.; et al. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*; IPCC Intergovernmental Panel on Climate Change (IPCC): Geneva, Switzerland, 2019; in press.
9. Arrigo, K.R.; van Dijken, G.; Pabi, S. Impact of a shrinking Arctic ice cover on marine primary production. *Geophys. Res. Lett.* **2008**, *35*, L19603. [CrossRef]
10. Kovacs, K.M.; Lydersen, C.; Overland, J.E.; Moore, S.E. Impacts of changing sea-ice conditions on Arctic marine mammals. *Mar. Biodivers.* **2011**, *41*, 181–194. [CrossRef]
11. Inoue, J.; Yamazaki, A.; Ono, J.; Dethloff, K.; Maturilli, M.; Neuber, R.; Edwards, P.; Yamaguchi, H. Additional Arctic observations improve weather and sea-ice forecasts for the Northern Sea Route. *Sci. Rep.* **2015**, *5*, 16868. [CrossRef]
12. Boé, J.; Hall, A.; Qu, X. September sea-ice cover in the Arctic Ocean projected to vanish by 2100. *Nat. Geosci.* **2009**, *2*, 341–343. [CrossRef]
13. Overland, J.E.; Wang, M. When will the summer Arctic be nearly sea ice free? *Geophys. Res. Lett.* **2013**, *40*, 2097–2101. [CrossRef]
14. Notz, D.; Stroeve, J. Observed Arctic sea-ice loss directly follows anthropogenic CO_2 emission. *Science* **2016**, *354*, 747–750. [CrossRef]
15. Rogelj, J.; den Elzen, M.; Höhne, N.; Fransen, T.; Fekete, H.; Winkler, H.; Schaeffer, R.; Sha, F.; Riahi, K.; Meinshausen, M. Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature* **2016**, *534*, 631–639. [CrossRef] [PubMed]
16. Ivanova, N.; Johannessen, O.M.; Pedersen, L.T.; Tonboe, R.T. Retrieval of Arctic sea ice parameters by satellite passive microwave sensors: A comparison of eleven sea ice concentration algorithms. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 7233–7246. [CrossRef]
17. Kunkee, D.B.; Swadley, S.D.; Poe, G.A.; Hong, Y.; Werner, M.F. Special Sensor Microwave Imager Sounder (SSMIS) radiometric calibration anomalies—Part I: Identification and characterization. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1017–1033. [CrossRef]
18. Imaoka, K.; Kachi, M.; Fujii, H.; Murakami, H.; Hori, M.; Ono, A.; Igarashi, T.; Nakagawa, K.; Oki, T.; Honda, Y.; et al. Global Change Observation Mission (GCOM) for monitoring carbon, water cycles, and climate change. *Proc. IEEE* **2010**, *98*, 717–734. [CrossRef]
19. Okuyama, A.; Imaoka, K. Intercalibration of Advanced Microwave Scanning Radiometer-2 (AMSR2) brightness temperature. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4568–4577. [CrossRef]
20. Cavalieri, D.J.; Gloersen, P.; Campbell, W.J. Determination of sea ice parameters with the Nimbus 7 SMMR. *J. Geophys. Res.* **1984**, *89*, 5355–5369. [CrossRef]
21. Comiso, J.C. Characteristics of Arctic winter sea ice from satellite multispectral microwave observations. *J. Geophys. Res.* **1986**, *91*, 975–994. [CrossRef]
22. Spreen, G.; Kaleschke, L.; Heygster, G. Sea ice remote sensing using AMSR-E 89-GHz channels. *J. Geophys. Res.* **2008**, *113*, C02S03. [CrossRef]
23. Ivanova, N.; Pedersen, L.T.; Tonboe, R.T.; Kern, S.; Heygster, G.; Lavergne, T.; Sørensen, A.; Saldo, R.; Dybkjær, G.; Brucker, L.; et al. Inter-comparison and evaluation of sea ice algorithms: Towards further identification of challenges and optimal approach using passive microwave observations. *Cryosphere* **2015**, *9*, 1797–1817. [CrossRef]
24. Tonboe, R.T.; Eastwood, S.; Lavergne, T.; Sørensen, A.M.; Rathmann, N.; Dybkjær, G.; Pedersen, L.T.; Høyer, J.L.; Kern, S. The EUMETSAT sea ice concentration climate data record. *Cryosphere* **2016**, *10*, 2275–2290. [CrossRef]
25. Andersen, S.; Tonboe, R.; Kern, S.; Schyberg, H. Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using numerical weather prediction model fields: An intercomparison of nine algorithms. *Remote Sens. Environ.* **2006**, *104*, 374–392. [CrossRef]

26. Han, H.; Kim, H.-C. Evaluation of summer passive microwave sea ice concentrations in the Chukchi Sea based on KOMPSAT-5 SAR and numerical weather prediction data. *Remote Sens. Environ.* **2018**, *209*, 343–362. [CrossRef]
27. Shin, D.-B.; Chiu, L.S.; Clemente-Colon, P. Effects of atmospheric water and surface wind on passive microwave retrievals of sea ice concentration: A simulation study. *Int. J. Remote Sens.* **2008**, *29*, 5717–5731. [CrossRef]
28. Meier, W.N. Comparison of passive microwave ice concentration algorithm retrievals with AVHRR imagery in Arctic peripheral seas. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1324–1337. [CrossRef]
29. Meier, W.; Notz, D. A note on the accuracy and reliability of satellite-derived passive microwave estimates of sea-ice extent. In *Clic Arctic Sea Ice Working Group Consensus Document*; World Climate Research Program: Geneva, Switzerland, 2010.
30. Cavalieri, D.J.; Germain, K.M.S.; Swift, C.T. Reduction of weather effects in the calculation of sea-ice concentration with the DMSP SSM/I. *J. Glaciol.* **1995**, *41*, 455–464. [CrossRef]
31. Wang, L.; Scott, K.A.; Xu, L.; Clausi, D.A. Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4524–4533. [CrossRef]
32. Karvonen, J. Baltic sea ice concentration estimation using SENTINEL-1 SAR and AMSR2 microwave radiometer data. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2871–2883. [CrossRef]
33. Wang, L.; Scott, K.A.; Clausi, D.A. Sea ice concentration estimation during freeze-up from SAR imagery using a convolutional neural network. *Remote Sens.* **2017**, *9*, 408. [CrossRef]
34. Chi, J.; Kim, H.-C.; Lee, S.; Crawford, M.M. Deep learning based retrieval algorithm for Arctic sea ice concentration from AMSR2 passive microwave and MODIS optical data. *Remote Sens. Environ.* **2019**, *231*, 111204. [CrossRef]
35. Fritzner, S.; Graversen, R.; Christensen, K.H. Assessment of high-resolution dynamical and machine learning models for prediction of sea ice concentration in a regional application. *J. Geophys. Res.* **2020**, *125*, e2020JC016277. [CrossRef]
36. Kim, Y.J.; Kim, H.-C.; Han, D.; Lee, S.; Im, J. Prediction of monthly Arctic sea ice concentrations using satellite and reanalysis data based on convolutional neural networks. *Cryosphere* **2020**, *14*, 1083–1104. [CrossRef]
37. Han, H.; Hong, S.-H.; Kim, H.-c.; Chae, T.-B.; Choi, H.-J. A study of the feasibility of using KOMPSAT-5 SAR data to map sea ice in the Chukchi Sea in late summer. *Remote Sens. Lett.* **2017**, *8*, 468–477. [CrossRef]
38. Woodgate, R.A.; Weingartner, T.; Lindsay, R. The 2007 Bering Strait oceanic heat flux and anomalous Arctic sea-ice retreat. *Geophys. Res. Lett.* **2010**, *37*, L01602. [CrossRef]
39. Stroeve, J.C.; Markus, T.; Boisvert, L.; Miller, J.; Barrett, A. Changes in Arctic melt season and implications for sea ice loss. *Geophys. Res. Lett.* **2014**, *41*, 1216–1225. [CrossRef]
40. Meier, W.N.; Stroeve, J.; Fetterer, F. Whither Arctic sea ice? A clear signal of decline regionally, seasonally and extending beyond the satellite record. *Ann. Glaciol.* **2007**, *46*, 428–434. [CrossRef]
41. Maeda, T.; Taniguchi, Y.; Imaoka, K. GCOM-W1 AMSR2 level 1R product: Dataset of brightness temperature modified using the antenna pattern matching technique. *IEEE Trans. Geosci. Remote Sens.* **2015**, *54*, 770–782. [CrossRef]
42. Hersbach, H.; de Rosnay, P.; Bell, B.; Schepers, D.; Simmons, A.; Soci, C.; Abdalla, S.; Alonso-Balmaseda, M.; Balsamo, G.; Bechtold, P.; et al. Operational Global Reanalysis: Progress, Future Directions and Synergies with NWP. *ECMWF Re-Anal. Proj. Rep. Ser.* **2018**, *27*, 1–63.
43. Serreze, M.C.; Crawford, A.D.; Stroeve, J.C.; Barrett, A.P.; Woodgate, R.A. Variability, trends, and predictability of seasonal sea ice retreat and advance in the Chukchi Sea. *J. Geophys. Res.* **2016**, *121*, 7308–7325. [CrossRef]
44. Markus, T.; Dokken, S.T. Evaluation of late summer passive microwave Arctic sea ice retrievals. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 348–356. [CrossRef]
45. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
46. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
47. Han, H.; Im, J.; Kim, M.; Sim, S.; Kim, J.; Kim, D.-J.; Kang, S.-H. Retrieval of melt ponds on arctic multiyear sea ice in summer from terrasar-x dual-polarization data using machine learning approaches: A case study in the Chukchi Sea with mid-incidence angle data. *Remote Sens.* **2016**, *8*, 57. [CrossRef]
48. Kim, M.; Im, J.; Han, H.; Kim, J.; Lee, S.; Shin, M.; Kim, H.-C. Landfast sea ice monitoring using multisensor fusion in the Antarctic. *GIScience Remote Sens.* **2015**, *52*, 239–256. [CrossRef]
49. Shen, X.; Zhang, J.; Zhang, X.; Meng, J.; Ke, C. Sea ice classification using Cryosat-2 altimeter data by optimal classifier–feature assembly. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1948–1952. [CrossRef]
50. Kim, M.; Kim, H.-C.; Im, J.; Lee, S.; Han, H. Object-based landfast sea ice detection over West Antarctica using time series ALOS PALSAR data. *Remote Sens. Environ.* **2020**, *242*, 111782. [CrossRef]
51. Lee, S.; Im, J.; Kim, J.; Kim, M.; Shin, M.; Kim, H.-C.; Quackenbush, L.J. Arctic sea ice thickness estimation from CryoSat-2 satellite data using machine learning-based lead detection. *Remote Sens.* **2016**, *8*, 698. [CrossRef]
52. Murashkin, D.; Spreen, G.; Huntemann, M.; Dierking, W. Method for detection of leads from Sentinel-1 SAR images. *Ann. Glaciol.* **2018**, *59*, 124–136. [CrossRef]
53. Cavalieri, D.J.; Markus, T.; Hall, D.K.; Gasiewski, A.J.; Klein, M.; Ivanoff, A. Assessment of EOS Aqua AMSR-E Arctic sea ice concentrations using Landsat-7 and airborne microwave imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3057–3069. [CrossRef]
54. Cavalieri, D.J.; Markus, T.; Hall, D.K.; Ivanoff, A.; Glick, E. Assessment of AMSR-E Antarctic winter sea-ice concentrations using Aqua MODIS. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3331–3339. [CrossRef]

55.   Liang, S.; Strahler, A.; Walthall, C. Retrieval of land surface albedo from satellite observations: A simulation study. In Proceedings of the 1998 IEEE International Geoscience and Remote Sensing Symposium IGARSS '98, Seattle, WA, USA, 6–10 July 1998; Volume 1283, pp. 1286–1288.
56.   Brandt, R.E.; Warren, S.G.; Worby, A.P.; Grenfell, T.C. Surface albedo of the Antarctic sea ice zone. *J. Clim.* **2005**, *18*, 3606–3622. [CrossRef]
57.   Radhakrishnan, R.; Scott, A.; Clausi, D.A. Sea ice concentration estimation: Using passive microwave and SAR data with a U-net and curriculum learning. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**. [CrossRef]
58.   Kern, S.; Rösel, A.; Pedersen, L.T.; Ivanova, N.; Saldo, R.; Tonboe, R.T. The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations. *Cryosphere* **2016**, *10*, 2217–2239. [CrossRef]
59.   Meier, W.N.; Ivanoff, A. Intercalibration of AMSR2 NASA Team 2 algorithm sea ice concentrations with AMSR-E slow rotation data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 3923–3933. [CrossRef]
60.   Wang, C.; Graham, R.M.; Wang, K.; Gerland, S.; Granskog, M.A. Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: Effects on sea ice thermodynamics and evolution. *Cryosphere* **2019**, *13*, 1661–1679. [CrossRef]
61.   Di Napoli, C.; Barnard, C.; Prudhomme, C.; Cloke, H.L.; Pappenberger, F. ERA5-HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. *Geosci. Data J.* **2020**, 1–9. [CrossRef]