Automated Sentinel-1 ice type mapping and in-situ validation during the CIRFA-22 cruise

Johannes LOHSE,[1] Catherine TAELMAN,[1] Alistair EVERETT,[2] Nick HUGHES[3]

[1] *UiT The Arctic University of Norway, Tromsø, Norway*

[2] *Norwegian Meteorological Institute, Oslo, Norway*

[3] *Norwegian Meteorological Institute, MET Norway Ice Service, Tromsø, Norway*

Correspondence: Johannes Lohse <johannes.p.lohse@uit.no>

ABSTRACT. We present a fully-automated workflow to map sea ice types from Sentinel-1 data and transfer the results in near real-time to the research vessel Kronprins Haakon (KPH) in order to support tactical navigation and decision-making during a research cruise conducted towards Belgica Bank in April and May 2022. We used overlapping SAR and optical imagery to train a pixel-wise classifier for the required season and region, and implemented a processing chain with the Norwegian Ice Service at MET Norway that automatically classifies all Sentinel-1 images covering the area of interest. During the cruise, classification results were available on KPH within hours after image acquisition, which is significantly faster than manually produced ice charts. We evaluate the results both quantitatively, based on manually selected validation regions, and qualitatively in comparison to in-situ observations and photographs. Our findings show that open water, level ice, and deformed ice are classified with high accuracy, while young ice remains challenging due to its variable small-scale surface roughness. This work presents one of the first attempts to transfer automated ice type classification results into the field in near real-time and contributes to bridging the gap between research and operations in automated sea ice mapping.

INTRODUCTION

Synthetic aperture radar (SAR) is the main data source for year-round, high-resolution sea ice monitoring and for the production of operational sea ice charts by national ice services around the world (Zakhvatkina and others, 2019).

The most commonly used data format is wide-swath imagery such as Sentinel-1 (S1) data acquired in extra-wide swath (EW) mode, which is typically distributed at 40×40 m pixel spacing and covers a ground range of approximately 410 km. The resulting ice charts provide information that is crucial for navigational support and to ensure the safety of vessels in the Arctic. While operational ice charts are at present still based on manual analysis of SAR imagery by expert sea ice analysts, considerable progress has been made in the field of automated mapping of both sea ice concentration (SIC) and sea ice type/stage of development (SoD). This has resulted in a variety of published algorithms which can potentially increase automation in operational ice charting or ice type mapping for navigation support inside the pack ice (e.g. Ochilov and Clausi, 2012; Leigh and oterhs, 2014; Zakhvatkina and others, 2017; Boulze and others, 2020; Malmgren-Hansen and others, 2021; Khaleghian and others, 2021; Pires de Lima and others, 2023). However, most of these algorithms are only used within academia and evaluation of their classification results is usually done in the traditional way, i.e. based on independent training and test sets (e.g. Murashkin and Frost, 2021; Stokholm and others, 2023). Running an automated algorithm in the operational procedures at the ice services requires more thorough and representative "real-world" in-situ validation that must go hand-in-hand with further improvement of the algorithms to ensure that they can uphold or even improve the quality standards of manual image analysis by trained experts.

In this study, we take a step towards bridging the gap between research and operations in automated ice type mapping, using a research cruise conducted by the *Centre for Integrated Remote Sensing and Forecasting for Arctic Operations* (CIRFA-22 cruise) as an example for the application and validation of a supervised algorithm in a fully-automated processing chain. The main goals of this study can be summarized as follows:

1. Automatically classify sea ice types in the area of interest for the cruise and demonstrate that we can transfer classification results in near real-time (NRT) to the ship.

2. Validate the classification results in the field.

3. Assess which ice types can be mapped reliably based on manually selected validation regions and compare

this traditional quantitative assessment to a qualitative evaluation based on in-situ observations.

It should be emphasized that the goal of our automated support for the CIRFA-22 cruise was not to reproduce operational ice charts but to enable fast and efficient navigation within the pack ice and close to the Greenland fast ice. While standard ice charts are very useful to safely navigate close to the ice edge, they do not always provide sufficient spatial detail for navigation within the ice. Once a vessel is in a region of high SIC, the best options available are (a) the direct transfer and interpretation of satellite imagery, (b) manual analysis with finer detail than the usual ice charts, or (c) automated products that provide information on the individual lead and floe scale level. While option (a) requires trained personnel on board the ship to interpret the SAR imagery, option (b) creates additional daily workload for the ice services' sea ice analysts. Option (c), on the other hand, requires preparation work such as the setup of the processing and data transfer chain and the training of the algorithm before a cruise or operation. The technical part of this preparation work is of course directly transferable between different operations. However, the training of the algorithm and selection of ice types may depend on the region, time of year, and user requirements and abilities, such as the demands on the mobility within the pack ice or the icebreaker class of the vessel.

The remainder of this article is structured as follows: In the following section, we give an overview of the study area and environmental conditions during the cruise, followed by a description of the remote sensing data sets and in-situ observations used in this study. Afterwards, we describe our selection of training data for different ice types, the classification algorithm, data processing chain, and the setup for the NRT data transfer to the vessel. We then present the results and discuss them with respect to the main goals of the study stated above and in comparison to standard operational ice charts. Finally, we summarize our conclusions and outline recommendations for future work.

STUDY AREA AND DATA SETS

Study area

The CIRFA-22 cruise was conducted on the research vessel *Kronprins Haakon* (KPH) in April and May 2022. The main purpose of the cruise was "to perform measurements and make observations which allow for validation of information and forecast products resulting from CIRFA's work" (Dierking and others, 2022), making it an ideal test scenario for this study. The cruise started in Longyearbyen, Svalbard, on April 22nd and the ship spent approximately three weeks in the Belgica Bank area outside the north-east
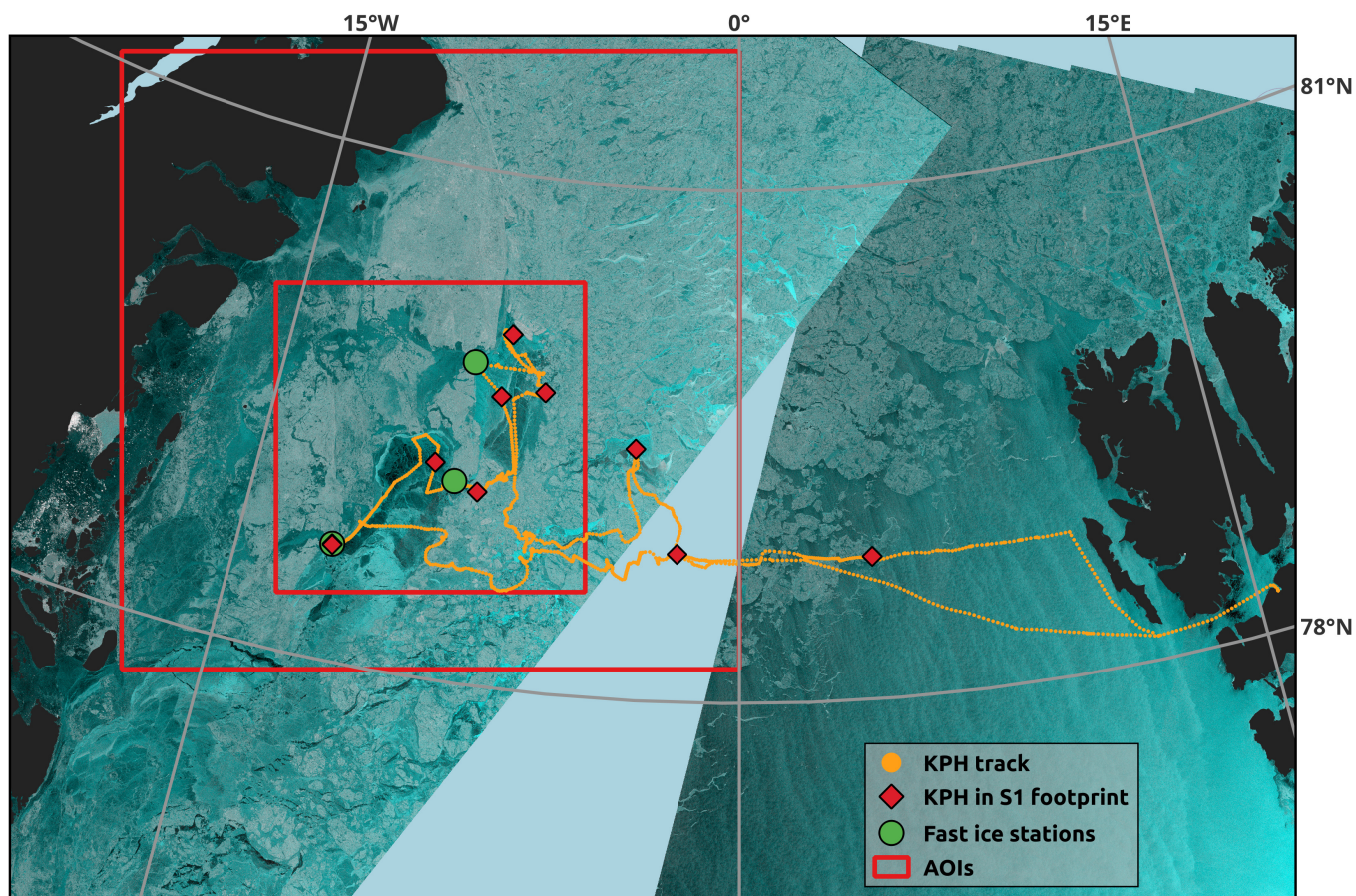
Fig. 1. Overview of the CIRFA-22 cruise conducted between April 22nd and May 9th 2022. The map shows the KPH ship track (orange), the locations of three multi-day ice stations (green markers) along the fast ice edge, and the positions of KPH whenever the vessel was within the footprint of an S1 scene at the time of image acquisition (red markers). The red squares indicate the small and large AOIs that were used for sub-setting the imagery sent to the ship. The S1 image in the background (rgb: HV, HH, HH) was acquired during two satellite overpasses on May 3rd at 08:26 (left) and 06:48 (right) UTC, and shows representative sea ice conditions during the cruise.

Greenland coast before returning to Longyearbyen on May 9th (Fig. 1). The sea ice situation around Belgica Bank can be challenging for navigation at this time of the year. The ice cover typically consists of both level and heavily deformed landfast ice close to the Greenland coast, as well as drift ice at various SoD further east in Fram Strait (Hughes and others, 2011). More detailed information on the actual ice conditions in the area in 2022 can be found in the cruise report (Dierking and others, 2022) or in Eltoft and others (2023).

The local air temperature measured on board KPH was consistently cold during the entire cruise period, mostly between -10 and -15°C and never exceeding -5°C. This is in agreement with several sea ice mass balance buoys deployed during the cruise, which recorded temperatures rising above -5°C for the first time in the second half of May 2022. Surface melt was not observed at any time during the cruise. We therefore consider the entire cruise period as "winter conditions" and can hence apply a classifier that was trained for cold winter conditions and dry snow. The training selection is described in more detail in the *Method* section of this paper.

Satellite data

Sentinel-1

Our processing chain and classification algorithm (described in the *Method* section) is based entirely on S1 data. S1 operates at C-band frequency (5.405 GHz) in either single- or dual-polarization mode. All data is freely available and can be accessed for example through the Copernicus dataspace platform (https://dataspace.copernicus.eu/). Here we use S1 images acquired at dual polarization (HH and HV channels) in EW mode and work with the Level-1 product in ground-range detected format at medium resolution (GRDM). The EW GRDM product is provided at a pixel spacing of 40×40 m with an actual spatial resolution of approximately 93×87 m (Aulard-Macler, 2011). The full swath width of 410 km is divided into five sub-swaths EW1 to EW5, with incident angles (IA) ranging from 18.9° in the near-range to 47.0° in the far-range. The pixel values are multi-looked intensities with 18 looks in the first sub-swath EW1 and 12 looks in the remaining sub-swaths EW2 to EW5. The noise-equivalent sigma zero (NESZ) of the S1 EW GRDM product, also known as the system noise floor, decreases across the swath. While its maximum value is equal to -23.1 dB in sub-swath EW1, the NESZ is mostly in the range between -27 and -33 dB in sub-swaths EW2 to EW5 (Aulard-Macler, 2011).

Thanks to its fine spatial resolution at wide coverage, its all-day and all-weather imaging capability, and

the free data availability, S1 wide-swath imagery is one of the most important data sources in operational ice charting.

Sentinel-2

In cloud-free conditions during daylight, optical sensors can provide valuable complementary information to the SAR data and thus aid in the interpretation of SAR signatures from different sea ice types. In this study, we use optical imagery acquired by Sentinel-2 (S2) to guide the selection of ice classes and training data. The S2 high-resolution multispectral instrument provides data at 13 spectral channels. For our visual interpretation in combination with S1 SAR data, we only use the visible channels (B4, B3, B2), which are provided at a pixel spacing of 10x10 m.

In-situ data

A large set of in-situ measurements was conducted during the cruise, including ship-based ice observations, on-ice measurements of physical snow and ice properties, drift observations using buoys with GPS sensors, and drone-based observations with both optical and radar sensors. Details about all the acquired data sets can be found in the official cruise report (Dierking and others, 2022) and in the online publications of the individual data sets.

For the work presented here, our in-situ validation of the classification results in the field is based on visual observations from the ship. These include the regular IceWatch (Hutchings and others, 2020) observations during the cruise (available at https://icewatch.met.no/cruises/130), as well as additional visual observations and photographs from the bridge and observation deck that were specifically timed to coincide with the timing of overlapping S1 image acquisitions. During the cruise, there were ten occasions at which KPH was located within the footprint of an S1 scene at the time of image acquisition (Fig. 1, Table 5 in Appendix A). Finally, most of the analysis presented here uses photographs that are taken each minute by a camera mounted in the crow's nest of KPH (hereafter denoted as "monkeytop camera"). Compared to the hand-held photographs, the monkeytop camera offers the advantage of a fixed imaging geometry and does not require manual operation.

Table 1.   Overview of ice types for the classifier.

Class index	Ice type	Acronym
1	Open Water / New Ice	OW
2	Young Ice	YI
3	Level Ice	LI
4	Deformed Ice	DI

METHOD

Training data selection

Before the start of the expedition, we used overlapping SAR (S1) and optical (S2) data to assess the typical sea ice situation in the cruise's target area around Belgica Bank during April and May. For this purpose, we studied ice charts and satellite images from the previous years (2020 and 2021) and from the months in the lead-up to the cruise in 2022 (February, March, and first half of April). Based on this analysis, we identified four main ice types: Open Water/New Ice (OW), Young Ice (YI), Level Ice (LI), and Deformed Ice (DI) (Table 1). Fig. 2 and 3 show examples of overlapping S1 and S2 images and indicate areas with the identified ice types. We used multiple such image pairs to select the classes (ice types) that we needed to separate in our automated processing chain during the cruise.

The choice of these four ice types is motivated by the goal of this project, which is to provide a product for navigational purposes inside the pack ice. As a consequence, we do not include a separate class for large open water areas. The OW class is specifically trained to identify small leads and openings (on the scale of tens or hundreds of meters up to a few kilometers) within large areas of high SIC. By our definition, this class includes entirely open leads as well as refrozen leads that may be covered by grease ice or very thin sheets of nilas in the earliest stages of sea ice formation. Adding an additional class for large areas of open water is possible, but since the cruise was planned to mostly operate close to the landfast ice far away from the marginal ice zone, we did not consider it necessary for the given task. An outlook on how to best include this additional class is given in the final section of this paper.

Furthermore, we separate between YI, LI, and DI. For the latter two we do not distinguish between FYI and MYI. While the separation of FYI and MYI is important for several applications, it is also a challenging task based on individual SAR intensities only. In most cases, the history of a particular ice

floe or region needs to be considered to make a correct and unambiguous decision. Because of its overall strong backscatter and particularly its expected high backscatter in HV channel due to volume scattering (e.g. Onstott and Carsey, 1993; Komarov and Buehner, 2019; Lohse and others, 2019), most of the MYI will fall into the DI class. This class should be interpreted as an area that will be difficult to navigate through even with an icebreaker such as KPH, hence making the classification result with our selected ice types valuable for navigational purposes. Furthermore, as the DI areas are likely to be considerable thicker than the LI areas, classification results for the selected ice types can also be useful for data assimilation in numerical models for sea ice forecasts.

It should be noted here that the ice types used in this study are consistent with the World Meteorological Organization (WMO) Sea Ice Nomenclature (WMO, 2014) definitions, and hence do not exactly match the SoD ice types provided in some operational ice charts. However, here we are interested in providing useful information for navigational support and route planning inside the pack ice. Separation of lead areas with OW from lead areas covered with thin nilas is not required for navigational purposes, as an icebreaker will travel equally easily through both. Since both of these classes are also difficult to distinguish because of their weak backscatter signatures caused by a smooth surface, it is reasonable to combine them into one class here. For the Norwegian icebreakers KPH and *KV Svalbard*, similar arguments hold for the separation of deformed FYI and MYI. To save time and fuel, both ships would avoid deformed FYI as much as MYI. However, this may be different for other cruises or operations on ships with a higher ice class. Hence, the class and training data selection should be tailored to the planned operations and the abilities of the involved vessels. Combining these classes (OW and new ice as well as deformed FYI and MYI) will of course lead to higher scores when evaluating classification results. This must be kept in mind when comparing classification scores to other algorithms that try to separate these ice types.

Classification algorithm

Based on the manually selected training regions for the ice types in Table 1, we trained a pixel-wise classification algorithm introduced by Lohse and others (2020) that uses both HH and HV intensity together with the IA to classify the S1 images. The method accounts for class-dependent differences in the variation of backscatter intensity with IA (Mäkynen and Karvonen, 2017; Guo and others, 2022), assuming a linear decrease of backscatter in decibel (dB) with increasing IA. This is achieved by using a two-dimensional Gaussian distribution with a linearly variable mean vector $\underline{\mu} = \underline{a} + \underline{b} \cdot \Theta$, where $\underline{\mu}$ is the mean vector, Θ is
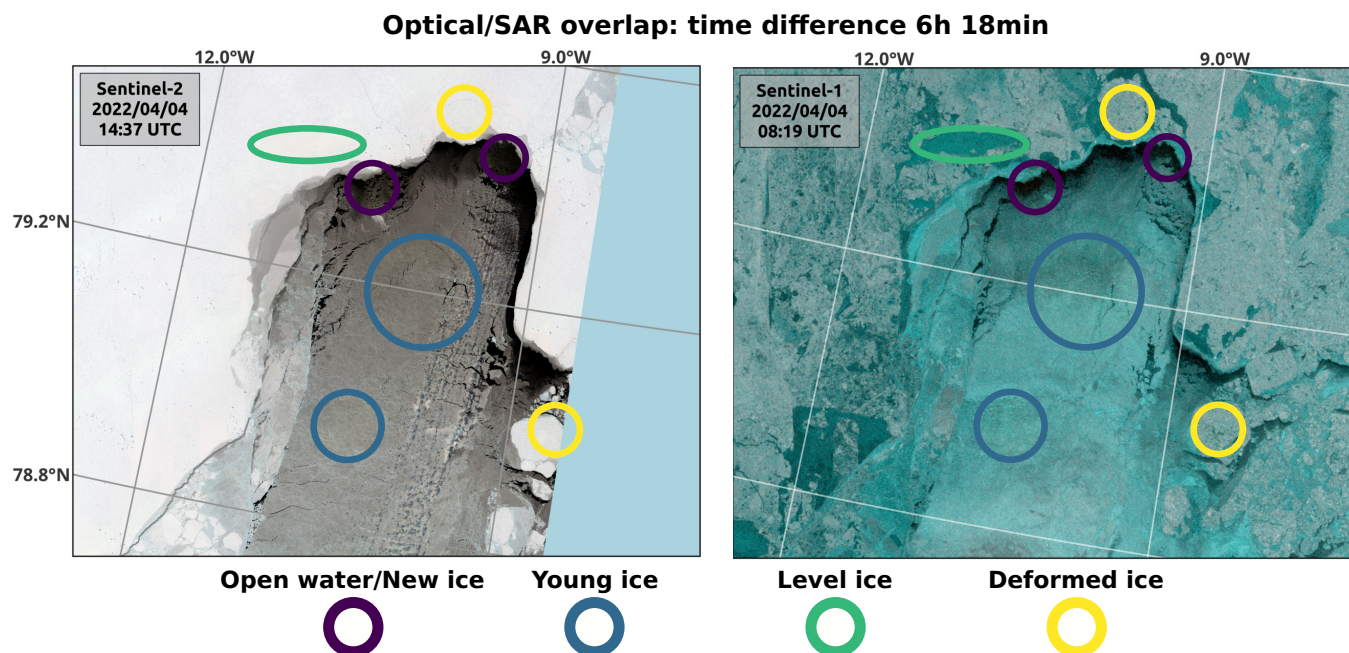
Fig. 2. Example of overlapping optical (left) and SAR (right) images in the Belgica Bank area, acquired on April 4th 2022, several weeks before the cruise. Selected ice types are marked by colored circles and ellipses. Note that the markers are drawn large for better visibility. The actually selected training regions are smaller and more precisely drawn to ensure that they do not contain mixed classes. To see the difference between LI and DI in the optical image, the dynamic range of the image must be adjusted (Fig. 3).
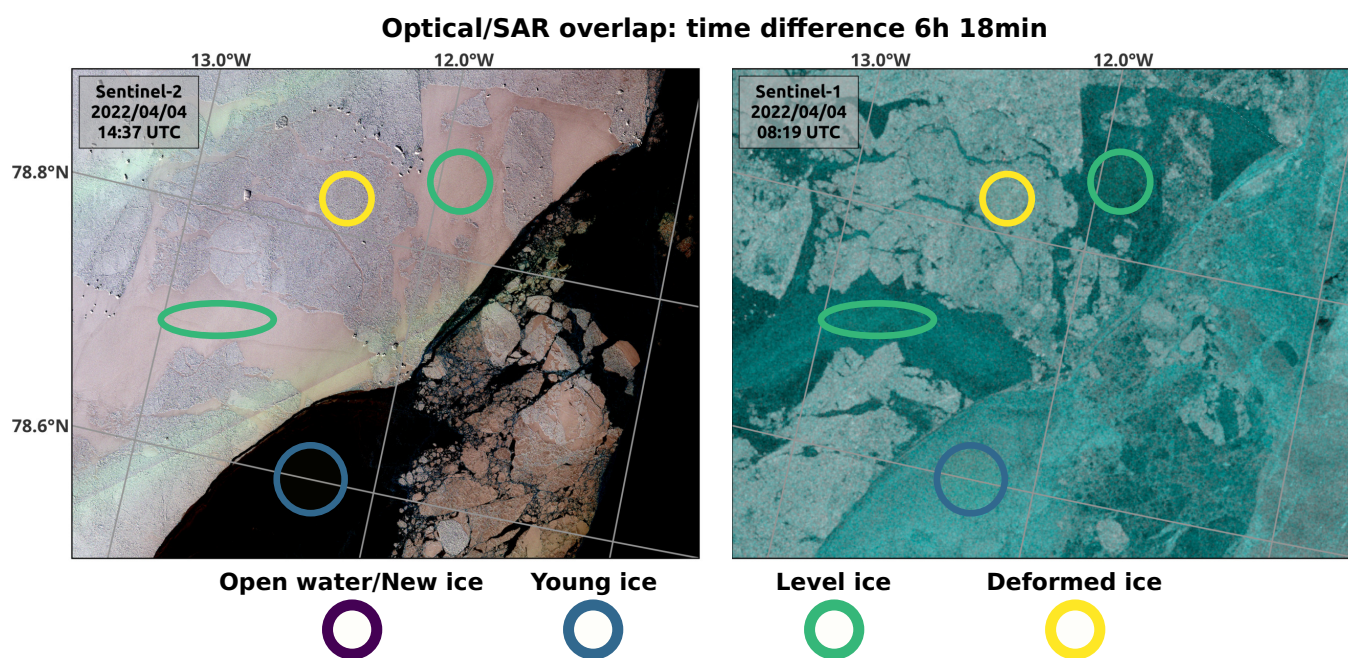
Fig. 3. Close-up of a different region from the same image pair as shown in Fig. 2, after adjusting the dynamic range of the optical image. Differences between LI and DI are now clearly visible in the optical image, while OW and YI both appear dark.

the IA, and $\underline{b}$ is a vector with the linear slopes for HH and HV. The concept of linear IA dependency of the input features can in principle be extended to include texture features. However, the decision whether or not to use texture features is a trade-off between computation time, spatial resolution of the results, and the gain in classification accuracy (CA). As shown in Lohse and others (2021), the algorithm requires large texture windows (>21x21 pixels) for a significant improvement of CA, which effectively decreases the spatial resolution of the resulting ice type maps. Furthermore, the largest improvement is found for the separation of OW and FYI or MYI. Since we require fine spatial resolution in this study and do not include a separate OW class, we use only the intensity channels as input features here. Fig. 4 illustrates the per-class IA dependency of HH and HV backscatter after training the algorithm for the relevant region and the season of the cruise. We see that in the near-range the HV intensity of LI, YI, and OW is close to or partly below the system noise floor. However, since the HH intensity is above the noise floor for all classes, we do not expect thermal noise to significantly affect the classification in this study. This is in good agreement with previous studies (e.g. Dierking, 2010).
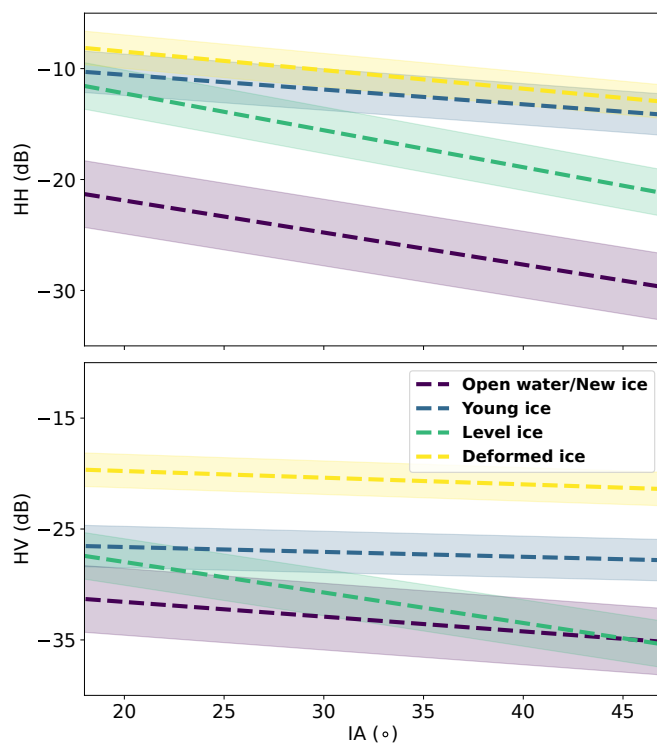
Fig. 4. Illustration of the per-class IA dependency of HH and HV backscatter intensity after training the algorithm for the relevant region and the season of the cruise. The dashed lines show the linearly variable mean values and the shaded areas correspond to two standard deviations.

Table 2.   Overview of settings for processing with various levels of multi-looking (ML), pixel spacing after geocoding, and approximate resulting file sizes for the larger AOI (400x400 km) shown in Fig. 1 after compression.

ML window	pixel spacing (m)	file size (KB)
1x1	80	3200
3x3	200	700
9x9	400	200
21x21	800	50

Processing chain and data transfer

We implemented a processing chain at the Norwegian Meteorological Institute (MET Norway) as part of the Norwegian Ice Service's (NIS) daily production that automatically downloads, pre-processes, classifies, and geocodes all S1 EW data covering an area of interest (AOI) for the cruise. The pre-processing includes the standard noise correction implemented in the Sentinel Application Platform (SNAP) as well as calibration of the data to normalized radar cross section σ_0, followed by various levels of multi-looking (ML) with increasing window sizes, and finally the conversion of σ_0 to dB. The pixel-wise classification result for each ML level was then geocoded to a suitable corresponding pixel size in polar stereographic projection (EPSG:3996), sub-set to two differently sized AOIs (red squares in Fig. 1), compressed, and uploaded to an ftp server that can be accessed from the ship. Table 2 gives an overview of the different processing settings for ML and pixel spacing after geocoding. Note that larger ML windows result in a smoother classification result that can hence be geocoded to coarser pixel spacing. The final file sizes are much smaller than the images at the original pixel spacing of 40m (Table 2) and can be downloaded to the ship more easily. However, smoothing and re-sampling comes at the cost of losing spatial detail. Processing the data with multiple ML levels allowed us to download the finest spatial resolution possible to the ship at any time, while considering the limited bandwidth and internet connection on board KPH. Sub-setting to two separate AOIs furthermore enabled us to download a coarser resolution product for the larger AOI (400x400 km) and finer resolution product with more spatial detail for the smaller AOI (200x200 km).

RESULTS

NRT data transfer

The fully automated processing chain at MET Norway and the data transfer to the vessel during the cruise worked well. All relevant images were successfully downloaded and processed, and classification results were available on KPH within 2 to 5 hours after the image acquisition, which was sufficient for navigation support and decision making during the cruise. The sea ice information was available significantly faster compared to manually produced operational ice charts, which are at best issued once per day for the Belgica Bank area. Furthermore, our automated product contained more spatial detail than the standard sea ice charts, including for example the exact location of large floes or of leads that were favourable for efficient travel (see last subsection of this *Results* section).

Classification time series

During the cruise, we classified all S1 imagery acquired over the large AOI indicated in Fig. 1 and transferred results to KPH in NRT. After the cruise, we extended the time series to cover the area for the entire time period from March 1st until May 31st 2022. Fig. 5 shows selected examples of the S1 imagery and the corresponding classification results.

Visual inspection of classification results

A visual inspection of the SAR imagery and the classification results shows that most ice types are success-fully identified by our algorithm. The stationary fast ice area is classified consistently over time and the results are independent of changes in imaging geometry such as IA, the radar look direction, or ascending and descending orbits of the satellite. Classification errors occur close to the Greenland coast in landfast LI areas with an untypical radar signature. The ice here presumably grows under protected conditions and forms a very smooth surface. This results in a low backscatter signal in both polarization channels which is easily confused with OW (Fig. 5 d)) and poses a known challenge for automated classification algorithms (e.g. Wang and others, 2023). At large enough distance to the coast, passive microwave radiometer (PMR) data could help to mitigate classification errors (Malmgren-Hansen and others, 2021), but in narrow fjords and close to the coast the PMR has too coarse spatial resolution. However, for our task of navigation support, these classification errors are not critical, as the affected regions are too far in the landfast ice to
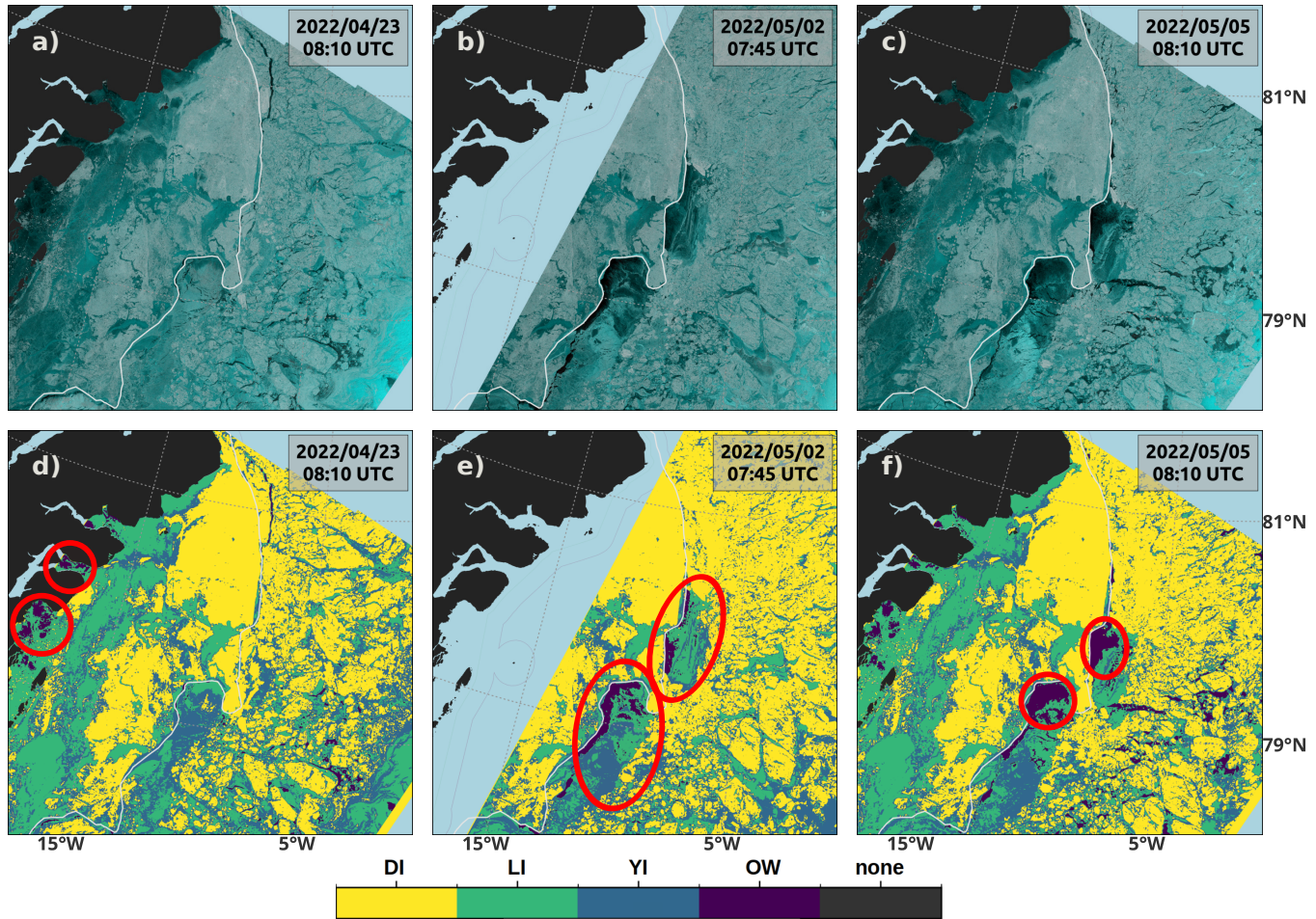
Fig. 5. Selected examples of S1 images (a), b), c)) and corresponding classification results (d), e), f)) covering the large AOI around Belgica Bank during the CIRFA-22 cruise. The white line indicates the fast ice edge at the end of April 2022. The fast ice region is classified consistently over time. Two polynya areas at the fast ice edge, marked by the red ellipses in e), are clearly visible in the classification result. Inside the polynyas, we find some misclassification of YI as LI. OW areas, marked by ellipses in f), are identified correctly. Classification errors close to the Greenland coast (LI is classified as OW) are highlighted in d).

be reached by KPH.

LI and DI are generally mapped correctly by the classifier. While the landfast ice is largely classified as either LI (western part) and DI (eastern part), the drift ice entering the area from the north consists mostly of deformed floes, intersected by smaller areas of YI or LI. Two polynyas repeatedly opened up at the fast ice edge in the time period between March and May 2022 (Fig. 5 e)). As the overall ice drift in the polynyas was towards the south and temperatures were well below freezing point until the middle of May, the polynya areas were mostly covered by YI, with smaller fraction of OW directly at the fast ice edge. While these patterns are often identified correctly by our algorithm, we also observe some classification errors of YI as LI. This is most likely caused by variations in the small-scale surface roughness of YI, due to the absence or the presence (and density) of frost flowers as well as finger rafting and the beginning of ridging.

Quantitative assessment of classification results

For a quantitative assessment of our algorithm, we evaluate the classification results over manually selected validation regions of interest (ROI)s and report per-class CA [%] in the form of a confusion matrix. We consider landfast and drift ice separately and include only images from the time period during the cruise. The selection of the validation ROIs is based on a comparison of S1 and optical data in combination with ship-based IceWatch and in-situ observations during the cruise. We therefore consider our manual selection to be reliable ground-truth data for a quantitative validation. For the landfast ice, we defined ten ROIs each for LI and DI. Each of the ROIs covers 8x8 pixels, corresponding to an area of 640x640 m. As the landfast ice did not undergo major changes during the cruise, we can use the same ROIs for all images. For the drift ice, we defined five ROIs per image for each class (OW, YI, LI, DI). Because some of the classes in the drift ice, in particular OW, often cover only small contiguous areas, we chose a smaller size of 5x5 pixels, corresponding to an area of 400x400 m. For a common evaluation of the different ML settings (Table 2) in our processing chain, we re-sampled the results from the 9x9 and 21x21 ML to 80 m pixel spacing.

Tables 3 and 4 show the confusion matrices for the classification results from the validation ROIs over landfast ice and over drift ice, respectively. The results support our visual inspection of the classified images. OW, LI, and DI are all identified with high accuracy. YI proves to be the most challenging class and only achieves an accuracy of around 50 %. It is often incorrectly classified as LI, and sometimes as

Table 3. Confusion matrices for classification results from validation regions over landfast ice for three distinct ML levels. The landfast ice validation labels (True class) only contain LI and DI ice types. Values are given in percentage [%].

			Predicted class			
			OW	**YI**	**LI**	**DI**
ML	True	**LI**	2.4	0.4	97.2	0.0
3x3	class	**DI**	0.0	6.0	0.4	93.6
ML	True	**LI**	1.1	0.0	98.9	0.0
9x9	class	**DI**	0.0	0.6	0.0	99.4
ML	True	**LI**	1.1	0.0	98.9	0.0
21x21	class	**DI**	0.0	0.0	0.0	99.9

OW or DI. More smoothing with larger ML windows increases the CA for DI, both in the landfast and the drift ice region. For OW, YI, and LI this is only true when we increase the ML window from 3x3 to 9x9. The further increase to 21x21 results in a constant or lower accuracy for these classes.

Comparison to ship-based in-situ photographs

For the in-situ validation of our algorithm, we compare the classification results to visual observations from the monkeytop camera of KPH at the time of image acquisition. There were in total ten occasions during the cruise at which KPH was in areas of high SIC and within the footprint of an S1 image. Fig. 6 shows four representative examples of the monkeytop photographs together with 25x25 km close-ups of the coincident S1 images and the corresponding classification results, centered around the position of KPH. The comparison with in-situ data shows that large areas of both LI and DI are correctly classified (Fig. 6 a) and c)). Both ice types are also mapped correctly in more heterogeneous regions with smaller floe sizes and a mixture of classes (Fig. 6 d)). Note that in-situ ice cores taken during the cruise (not shown here) indicate that the deformed ice areas do in fact contain a mixture of FYI and MYI, which is in agreement with our initial assumption during ice class and training data selection.

For YI, the comparison of in-situ observations and classification results reveals significant classification errors. The region around KPH in Fig. 6 b) is mostly classified as LI, while the monkeytop photograph clearly shows YI. Additional visual observations and manual photographs (not shown here) confirm that the YI in this example has little small-scale (mm to dm) or large-scale (m) surface roughness , hence its

Table 4. Confusion matrices for classification results from validation regions over drift ice for three distinct ML levels. Values are given in percentage [%].

			Predicted class			
			OW	**YI**	**LI**	**DI**
ML 3x3	True class	**OW**	99.6	0.0	0.4	0.0
		YI	8.6	47.0	41.4	2.7
		LI	1.7	8.5	89.5	0.3
		DI	0.0	8.0	1.1	90.9
ML 9x9	True class	**OW**	99.9	0.0	0.0	0.0
		YI	6.6	54.2	38.7	0.6
		LI	0.0	5.7	94.3	0.0
		DI	0.0	2.7	0.0	97.3
ML 21x21	True class	**OW**	99.4	0.0	0.6	0.0
		YI	5.3	52.6	42.1	0.0
		LI	0.0	11.9	88.1	0.0
		DI	0.0	0.3	0.0	99.7

backscatter signal is relatively weak. Further south in the same image, we find rougher YI with a stronger backscatter signature that is mapped correctly. These results are in good agreement with the significantly lower CA scores for the YI class.

Comparison to operational ice charts

Operational ice charts are the standard product issued by national ice services to support maritime navigation in the Arctic. For the Belgica Bank area at the time of the CIRFA-22 cruise, the NIS produced charts on weekdays showing SIC while the Greenland Ice Service at the Danish Meteorological Institute (DMI) produced weekly ice charts with total SIC, partial concentrations for different SoD, and floe size distributions. Fig. 7 shows an example comparison of original SAR imagery, NIS and DMI ice charts, and our classification result. While the polygons in both ice charts generally reflect the large-scale (tens of kilometer) patterns in the SAR imagery and the classification result, the pixel-wise ice type labels from our algorithm provide much finer spatial information on the individual lead and floe scale level, which is not present in either of the ice charts. The two large polynya areas along the fast ice edge are identified in both ice charts. The NIS characterizes the polynyas as a mixture of *Very Open Drift Ice* and *Close Drift Ice*,
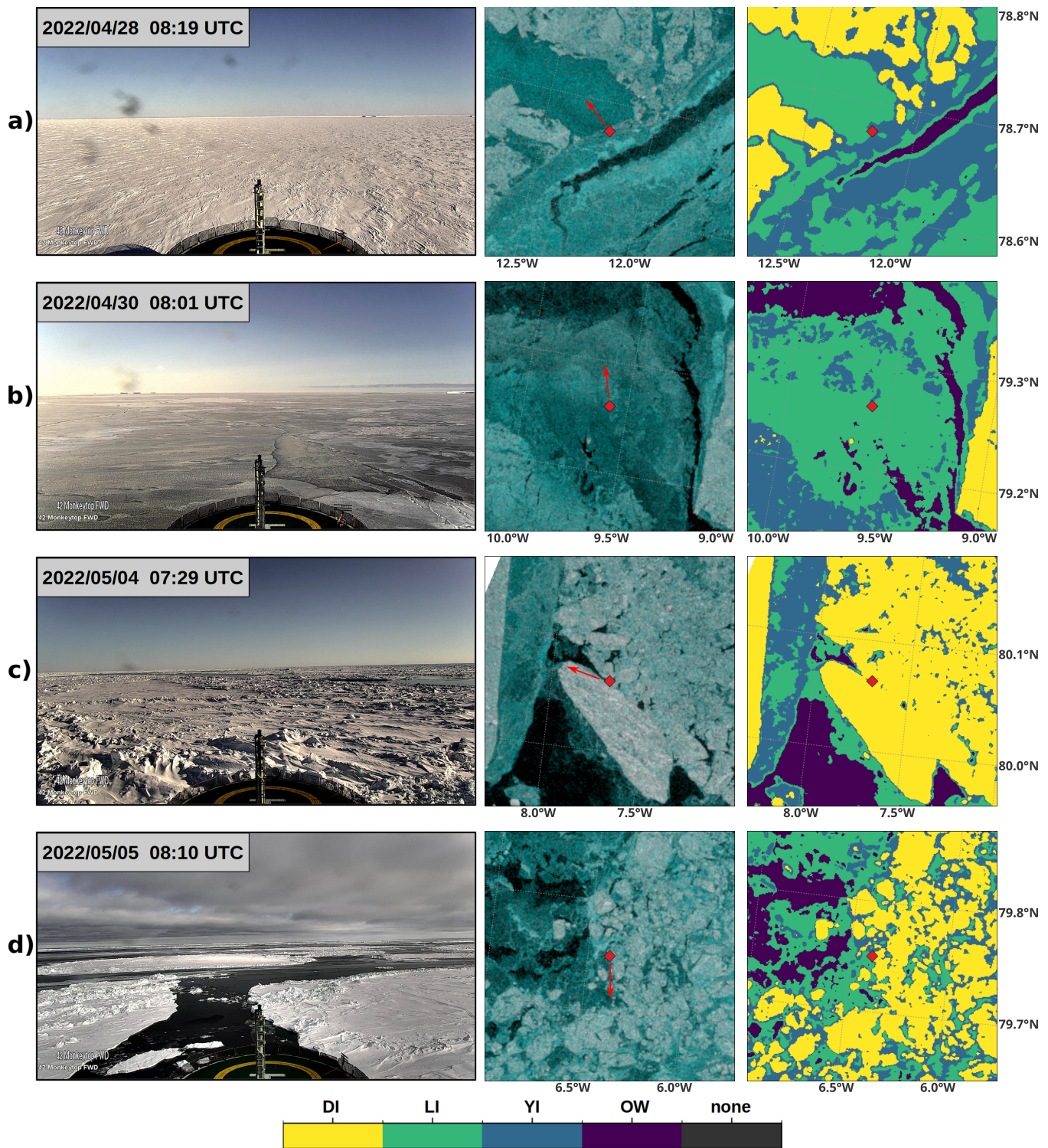
Fig. 6. Coincident monkeytop camera photographs (left), S1 images (middle), and classification results (right) for four selected examples during the CIRFA-22 cruise. SAR images are cropped to an area of 25 by 25 km around KPH's position (indicated by the red marker in the center) at the time of image acquisition. LI (row 1), DI (row 3), and a mixture of DI and LI floes and YI (row 4) are identified correctly by the classifier. Large areas of YI (row 2) in the polynya area are partly misclassified as LI.

while the DMI ice chart denotes them as *Close Drift Ice* with YI, nilas, and thin FYI as the predominant SoD. Neither of the ice charts provides the precise location or orientation of refrozen leads or deformed ice floes.

DISCUSSION

The automated ice type mapping and information transfer to the vessel was clearly successful. To our best knowledge, this is the first time that such results were actually sent into the field in NRT. While having high-resolution classification images available on KPH within a few hours after image acquisition was beneficial for navigation and route planning, the information was also used to guide scientific questions and decisions about the locations of ice stations and in-situ measurements during the cruise. An example case of how the SAR data and classified images were used for tactical decisions is given at the end of this section. Using the fully-automated processing chain, the classification results were available significantly faster than traditional ice charts and contained more detailed spatial information than the standard ice charts by the NIS, which only provide SIC for the AOI. The DMI ice charts contain some information on ice types (SoD and floe size) that could potentially be used for route planning, yet it is not provided with the same spatial detail as our classification product. Furthermore, the DMI charts for the AOI were issued only once per week at the time of the cruise. For the tactical navigation within the ice, the timeliness as well as the temporal and spatial resolution of our classification product is clearly preferable. It should be noted, however, that the setup for automated support during the cruise required a considerable amount of preparation work, most importantly the training of a pixel-wise classifier for the specific area and season of interest. While we were able to do this successfully for this demonstration example, it cannot necessarily be directly transferred to other ice regions or times of the year. Furthermore, a different cruise or operation with another research vessel may have different requirements on the mapped ice types, both because of scientific research questions and because of the ice-breaking capabilities of the vessel.

The quantitative evaluation of the classification results shows that OW, LI, and DI are generally mapped well by the classifier, achieving maximum accuracies of 99.9 %, 94.3 %, and 99.7 %, respectively. It is noteworthy that these maxima for the different classes are achieved at different ML levels. For DI, the CA improves steadily with increasing ML windows and achieves its maximum value for ML 21x21 (Tables 3 and 4). This indicates that classification errors of DI are largely caused by class-internal speckle variation. A larger number of looks reduces speckle and results in a tightened class distribution around the mean,
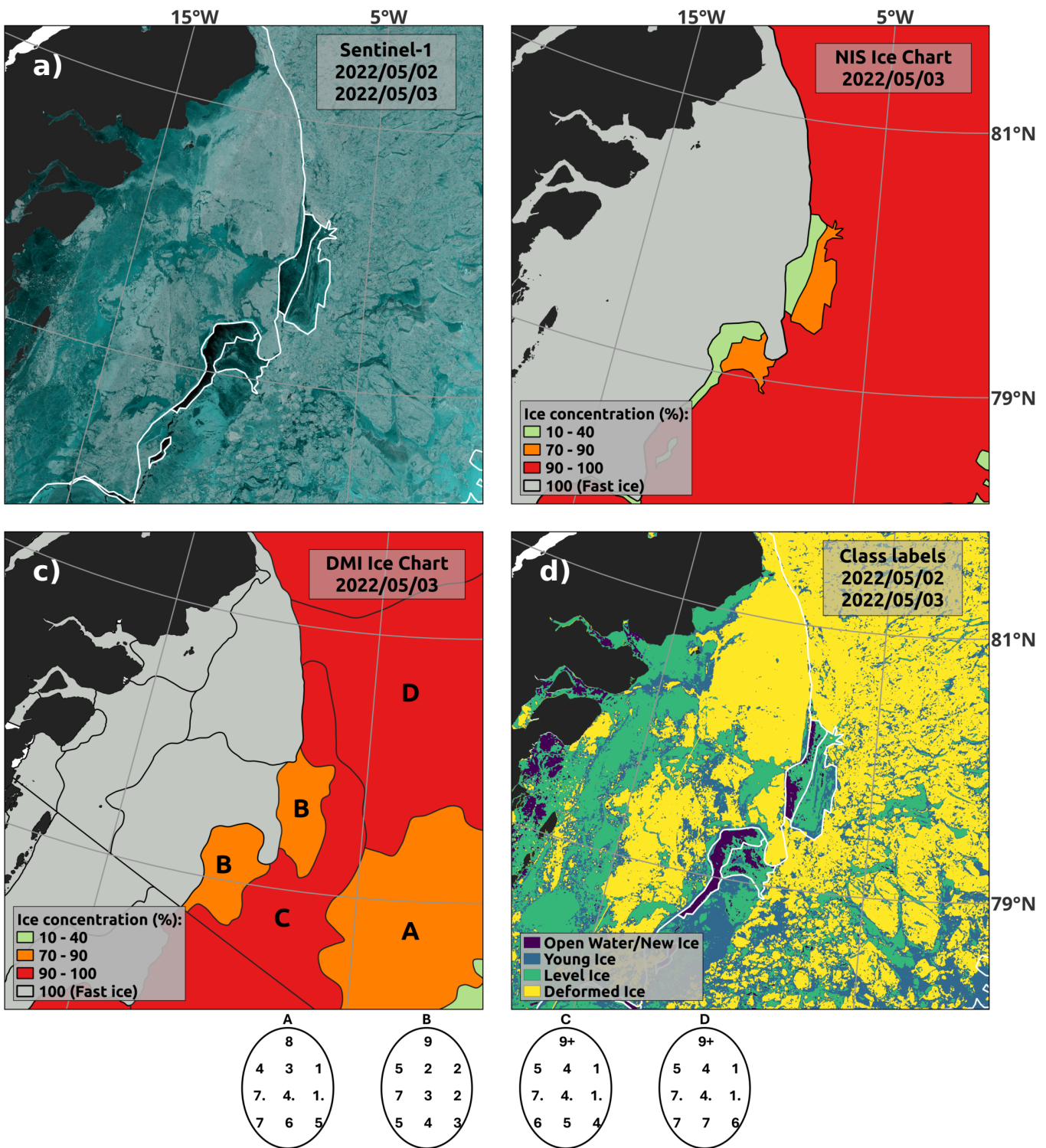
Fig. 7. Comparison of original SAR imagery acquired on May 2nd and 3rd 2022 (a), corresponding ice chart polygons from NIS (b) and DMI (c), and our pixel-wise classification result (d), shown for the large AOI (Fig. 1). For orientation, the NIS polygon outlines are overlaid on the SAR imagery and the classification result. The NIS ice chart contains SIC only, the DMI ice chart also provides information on partial ice type concentration, SoD, and form/floe size. This additional information is provided by the egg codes according to the WMO sea ice nomenclature (WMO, 2014) for sea ice charts for the four main drift ice polygons (A-D in c)) in the shown example. Note that polygons A and B have a similar total SIC, but are dominated by different SoD (A: old ice (7.), thick FYI (4.), medium FYI (1.); B: thin FYI (7), YI (3), Nilas (2)).

hence reducing the number of misclassified pixels. For OW and LI, however, maximum CA is achieved at ML 9x9 (Table 4). This can be explained by two competing effects: On the one hand, similar to the DI case, a larger number of looks decreases the class-internal speckle and hence improves the CA. On the other hand, edge effects and mixing of classes in large ML windows can lead to classification errors close to the boundaries between two different classes. As the regions covered by OW or LI within our study AOI are often significantly smaller than the areas covered by DI, these boundary effects are more prominent for OW and LI. Note also that within the landfast ice, there are larger contiguous regions of LI, and the maximum CA of LI within the landfast area is achieved with ML 21x21 (Table 3).

YI is the most difficult ice type to classify and achieves the lowest score of all classes in this study. Its maximum CA is at 54.2% with a ML window of 9x9. The main challenge for the classification of YI is its highly variable small-scale surface roughness with respect to the radar wavelength. For C-band, this small-scale roughness is on the order of millimeters to centimeters, which corresponds to changes of the surface caused for example by frost flowers or snow crusts (Isleifson and others, 2013). While very smooth YI with low-backscatter is misclassified as LI (38.7%) or OW (6.6%), very rough YI with strong-backscatter can be misclassified as DI (0.6%). These fractions of misclassified YI indicate that our initial selection of YI training data before the cruise was biased towards rough YI. One way to mitigate this issue could be to introduce a second YI class that is trained on smoother YI. While this should increase the YI CA, it would also increase the false positives and wrongly classify more LI as YI. This is unwanted, in particular in the landfast ice area. Separately trained classifiers for the landfast ice and the drift ice can be used to overcome this issue, but will require either a manual delineation or an automated detection of the fast ice edge. This is beyond the scope of the present study, but we are planning to investigate it in future work.

It should also be noted that the absolute CA numbers must be interpreted carefully, as they are dependent on the subjective selection of validation ROIs. This is a common problem in any traditional evaluation of a classifier that is based on a train and test set. In this study, we therefore also qualitatively compare the classification results with in-situ observations during the cruise. Overall, the quantitative accuracy assessment is in good agreement with the qualitative comparison of in-situ observations and classification results. OW, LI, and DI are correctly classified whenever the monkeytop photograph shows the respective ice type (Fig. 6). Furthermore, the monkeytop photographs and IceWatch observations also confirm the challenges of YI classification, as for example the relatively smooth YI in Fig. 6 b) is misclassified as LI.

The ten occasions during the CIRFA-22 cruise when KPH was located within an S1 footprint are not sufficient to use the in-situ observations for a quantitative evaluation of the classification results. Yet our qualitative comparison here shows the potential of such ship-based photographs to be used for the assessment of automated ice type classification. Similar to the monkeytop camera, manual IceWatch observations and photographs can be used in the same way. However, to facilitate the use of IceWatch data, the observations need to be aligned with the S1 acquisition schedule, which is not always feasible given the many different tasks that are carried out on a cruise. Hence, automatically acquired photographs provide a more practical solution. In the future, it would be beneficial to install additional monkeytop cameras looking not just to the front, but also to the sides of the ship. Especially in variable sea ice conditions, this is an easy way to increase the amount of validation data. In a large-scale study using monkeytop photographs for quantitative accuracy assessment, the photographs taken in the different directions at fixed geometries can then be warped onto a map and directly compared to classification results. This can also be useful to validate not only ice type classification, but also pixel-wise ice-water mapping in the marginal ice zone, such as the method introduced by Wang and others (2023).

Example of guided tactical decision-making during the cruise

In Fig. 8 we show an illustrative example of how the classified imagery was used on board KPH to support tactical navigation during the cruise. The figure shows a time series of four SAR images over the small AOI between May 2nd and May 5th 2022. KPH's position is indicated by a red marker within each image and the ship track between the previous and the current image acquisition is shown by an orange line. Older ship tracks are displayed as gray lines.

On May 2nd, the scientific work within the "southern polynya" was finished and the cruise plan was to go north through an area of DI to reach a planned fast ice station adjacent to the northern tip of the "northern polynya". At the time of image acquisition on May 3rd (Fig.8b and 8f) KPH was traversing through the northern polynya with fast progress. Given the information from the imagery on this day (SAR imagery or classification result) it became clear that the polynya was kept open by a large DI floe (marked by a red ellipse) that blocked the smaller DI floes drifting in from the north. As the large floe slowly drifted southward, the goal to reach the planned fast ice station with enough time left to conduct valuable work became unfeasible. Instead, various in-situ measurements were conducted within the northern polynya and on the drift ice just next to the large floe. The imagery on May 4th (Fig.8c and 8g) indicated that the drift
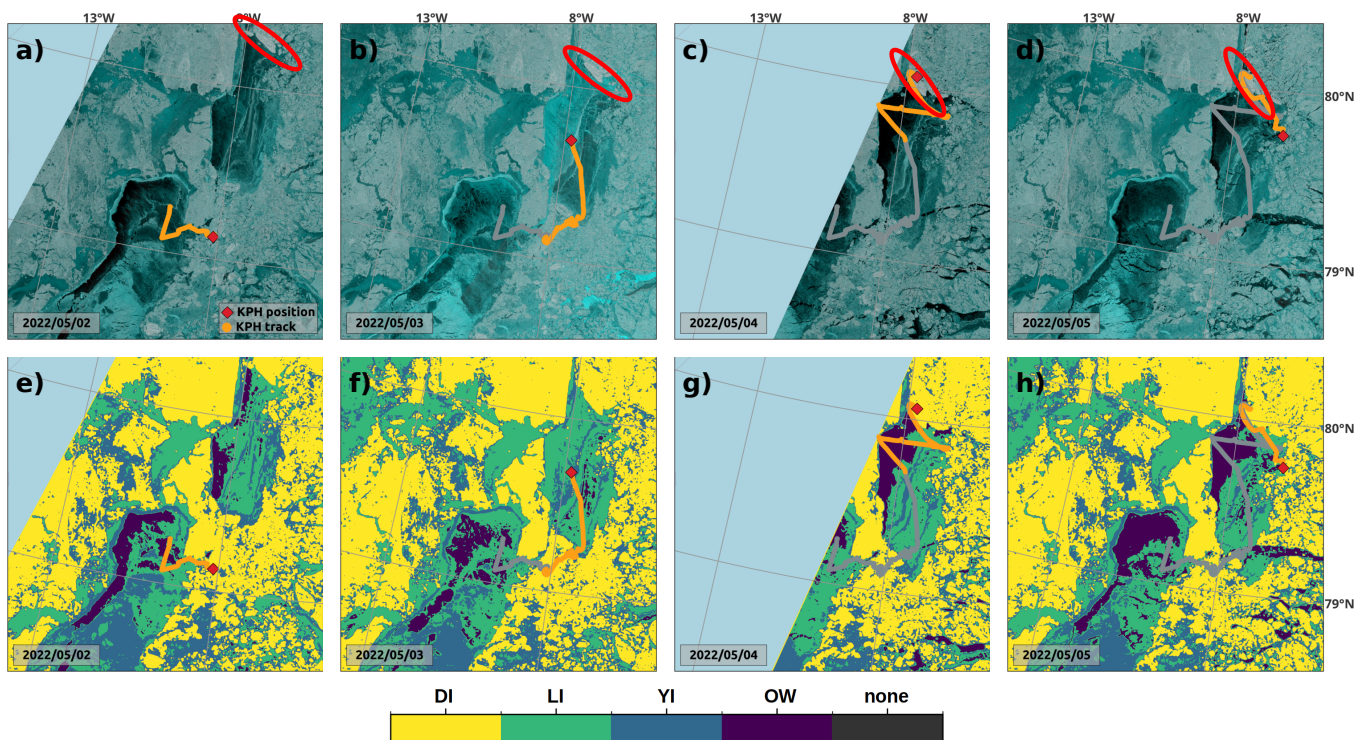
Fig. 8. Time series (May 2nd until May 5th 2022) of SAR imagery (a-d) and corresponding classification results (e-h) over the small AOI (Fig. 1). The red marker indicates KPH's position at the time of each image acquisition and the gray and orange lines show the ship track, with the orange part representing the track between the previous and the current image acquisition. A large deformed ice floe blocked the southward drift of the deformed ice further north and kept the northern polynya almost ice free. The floe is approximately 20 km wide and highlighted by the red ellipse in the SAR imagery (a-d).

of the large floe had slowly turned towards the fast ice edge in the west, constituting the risk of trapping KPH between the floe and the fast ice edge. This led to the decision to escape back around the large floe (see the ship track Fig.8d and 8h) before getting stuck. Without the high-resolution information available on board and the fine spatial detail provided by the pixel-wise classification result in comparison to the ice charts (Fig. 7), these considerations and decisions would not have been possible and much cruise time and fuel would likely have been wasted trying to follow the original cruise plan.

CONCLUSION

In this study, we have taken a step to bridge the gap between research and operations in automated ice type mapping. We have successfully demonstrated the application of a fully-automated sea ice type classification workflow at MET Norway and transferred classification results in NRT to a vessel in the Arctic, providing

detailed sea ice information at fine spatial resolution in support of tactical navigation and decision making. We have evaluated the classification results for individual ice types using validation ROIs that are confirmed from observations in the field. The results show that OW, LI, and DI are mapped with high accuracy, while YI remains challenging due to variable small-scale (mm to dm) and large-scale (m) surface roughness. Finally, we have used ship-based in-situ photographs to qualitatively evaluate the classification results. The comparison shows good agreement between observations and classification of OW, LI, and DI, while it also reflects the challenges for reliable identification of YI, caused by its small-scale surface roughness variability.

For the next steps of this work, we plan to classify ice types within the landfast ice and drift ice areas separately, using either a manually or automatically detected fast ice edge (Selyuzhenok and Demchev, 2021; Wang and others, 2021). This will reduce classification errors within the fast ice and furthermore allow us to test the incorporation of multiple YI classes in the drift ice, without compromising the landfast ice type classification. While the reliable separation of ice types within the fast ice is only of minor importance for ship traffic, it can be critical at inhabited coasts where people move on fast ice for hunting, fishing, and access to islands (Segal and others, 2020).

Furthermore, recall that we did not train a separate class for large open water areas. While it has been shown that the algorithm used in this study can classify OW using a combination of intensity features and image texture (Lohse and others, 2021), other publications indicate the convolutional neural networks perform ice-water separation faster and more reliable (e.g. Malmgren-Hansen and others, 2021; Stokholm and others, 2022; Chen and others, 2023; Wang and others, 2023). Hence, in future work, the fine-resolution ice type classification presented in this study can be applied in regions that are identified as high SIC by a CNN.

Finally, the monkeytop photographs can be warped onto the geocoded SAR classification results and thus allow large-scale qualitative evaluation of the retrieved ice types independent of manually selected validation ROIs. We expect that this "real-world" validation based on a comparison between automatically mapped ice types and in-situ observations will increase stakeholders' trust in the automated products and thus facilitate the transition of algorithms from research into operations at the ice services.

ACKNOWLEDGEMENTS

primary and altered Sentinel data products (@Copernicus data). We would like to thank the anonymous reviewers for their helpful comments and feedback. The suggested changes have significantly improved this manuscript. We also thank Wolfgang Dierking for his invaluable input during the writing process, that improved the quality of this manuscript.

REFERENCES

Aulard-Macler M (2011) Sentinel-1 product definition. *Document Reference MPC-0240*

Boulze H, Korosov A and Brajard J (2020) Classification of Sea Ice Types in Sentinel-1 SAR Data Using Convolutional Neural Networks. *Remote Sensing*, **12**(13) (doi: 10.3390/rs12132165)

Chen X, Patel M, Pena Cantu F, Park J, Noa Turnes J, Xu L, Scott A and Clausi DA (2023) MMSeaIce: Multi-task Mapping of Sea Ice Parameters from AI4Arctic Sea Ice Challenge Dataset. *EGUsphere*, **2023**, 1–17 (doi: 10.5194/egusphere-2023-1297)

Dierking W (2010) Mapping of different sea ice regimes using images from Sentinel-1 and ALOS synthetic aperture radar. *IEEE Transactions on Geoscience and Remote Sensing*, **48**(3), 1045–1058 (doi: 10.1109/TGRS.2009.2031806)

Dierking W, Schneider A, Eltoft T and Gerland S (2022) CIRFA Cruise 2022, Cruise report (doi: 10.5281/zenodo.7314066)

Eltoft T, Johansson M, Lohse J and Ferro-Famil L (2023) Analysis of Time Series of Polarimetric SEA ICE Signatures Observed In Fast ICE in the Belgica Bank Area. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, 176–178, IEEE (doi: 10.1109/IGARSS52108.2023.10281665)

Guo W, Itkin P, Lohse J, Johansson M and Doulgeris AP (2022) Cross-platform classification of level and deformed sea ice considering per-class incident angle dependency of backscatter intensity. *The Cryosphere*, **16**(1), 237–257 (doi: 10.5194/tc-16-237-2022)

Hughes NE, Wilkinson JP and Wadhams P (2011) Multi-satellite sensor analysis of fast-ice development in the Norske Øer Ice Barrier, northeast Greenland. *Annals of Glaciology*, **52**(57), 151–160 (doi: 10.3189/172756411795931633)

Hutchings S, Delamere J and Heil P (2020) The Ice Watch Manual. *https://icewatch.met.no/Ice_Watch_Manual_v4.1.pdf*

Isleifson D, Galley RJ, Barber DG, Landy JC, Komarov AS and Shafai L (2013) A study on the C-band polarimetric scattering and physical characteristics of frost flowers on experimental sea ice. *IEEE Transactions on Geoscience and Remote Sensing*, **52**(3), 1787–1798 (doi: 10.1109/TGRS.2013.2255060)

Khaleghian S, Ullah H, Kræmer T, Hughes NE, Eltoft T and Marinoni A (2021) Sea ice classification of SAR imagery based on convolution neural networks. *Remote Sensing*, **13**(9) (doi: 10.3390/rs13091734)

Komarov AS and Buehner M (2019) Detection of first-year and multi-year sea ice from dual-polarization sar images under cold conditions. *IEEE Transactions on Geoscience and remote sensing*, **57**(11), 9109–9123 (doi: 10.1109/TGRS.2019.2924868)

Leigh S, Wang Z and Clausi DA (2014) Automated Ice–Water Classification Using Dual Polarization SAR Satellite Imagery. *IEEE Transactions on Geoscience and remote sensing*, **52**(9), 5529–5539 (doi: 10.1109/TGRS.2013.2290231)

Lohse J, Doulgeris AP and Dierking W (2019) An optimal decision-tree design strategy and its application to sea ice classification from SAR imagery. *Remote Sensing*, **11**(13), 1574 (doi: 10.3390/rs11131574)

Lohse J, Doulgeris AP and Dierking W (2020) Mapping sea-ice types from Sentinel-1 considering the surface-type dependent effect of incidence angle. *Annals of Glaciology*, **61**(83), 260–270 (doi: 10.1017/aog.2020.45)

Lohse J, Doulgeris AP and Dierking W (2021) Incident angle dependence of Sentinel-1 texture features for sea ice classification. *Remote Sensing*, **13**(4), 552

Mäkynen M and Karvonen J (2017) Incidence angle dependence of first-year sea ice backscattering coefficient in Sentinel-1 SAR imagery over the Kara Sea. *IEEE Transactions on Geoscience and Remote Sensing*, **55**(11), 6170–6181 (doi: 10.1109/TGRS.2017.2721981)

Malmgren-Hansen D, Pedersen LT, Nielsen AA, Kreiner MB, Saldo R, Skriver H, Lavelle J, Buus-Hinkler J and Krane KH (2021) A convolutional neural network architecture for Sentinel-1 and AMSR2 data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, **59**(3), 1890–1902 (doi: 10.1109/TGRS.2020.3004539)

Murashkin D and Frost A (2021) Arctic Sea ICE Mapping Using Sentinel-1 SAR Scenes with a Convolutional Neural Network. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 5660–5663, IEEE (doi: 10.1109/IGARSS47720.2021.9553206)

Ochilov S and Clausi DA (2012) Operational SAR Sea-Ice Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **50**(11), 4397–4408 (doi: 10.1109/TGRS.2012.2192278)

Onstott RG and Carsey FD (1993) SAR and scatterometer signatures of sea ice. In *The Remote Sensing of Sea Ice*, 73–104, American Geophysical Union Washington, DC

Pires de Lima R, Vahedi B, Hughes NE, Barrett AP, Meier W and Karimzadeh M (2023) Enhancing sea ice segmentation in Sentinel-1 images with atrous convolutions. *International Journal of Remote Sensing*, **44**(17), 5344–5374 (doi: 10.1080/01431161.2023.2248560)

Segal RA, Scharien RK, Duerden F and Tam CL (2020) The best of both worlds. *Arctic*, **73**(4), 461–484 (doi: 10.14430/arctic71567)

Selyuzhenok V and Demchev D (2021) An application of sea ice tracking algorithm for fast ice and stamukhas detection in the Arctic. *Remote Sensing*, **13**(18), 3783 (doi: 10.3390/rs13183783)

Stokholm A, Wulf T, Kucik A, Saldo R, Buus-Hinkler J and Hvidegaard SM (2022) AI4SeaIce: Toward solving ambiguous SAR textures in convolutional neural networks for automatic sea ice concentration charting. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1–13 (doi: 10.1109/TGRS.2022.3149323)

Stokholm A, Buus-Hinkler J, Wulf T, Korosov A, Saldo R, Arthurs D, Solberg R, Longépé N and Kreiner M (2023) The AutoICE Competition: Automatically Mapping Sea Ice in the Arctic. Technical report, Copernicus Meetings (doi: 10.5194/egusphere-egu23-13038)

Wang Q, Lohse J, Doulgeris AP and Eltoft T (2023) Data augmentation for SAR sea ice and water classification based on per-class backscatter variation with incidence angle. *IEEE Transactions on Geoscience and Remote Sensing* (doi: 10.1109/TGRS.2023.3291927)

Wang Z, Wang Z, Li H, Ni P and Liu J (2021) A modified approach of extracting landfast ice edge based on sentinel-1A InSAR coherence image in the gulf of bothnia. *Journal of Marine Science and Engineering*, **9**(10), 1076 (doi: 10.3390/jmse9101076)

World Meteorological Organization (2014) WMO Sea Ice Nomenclature (WMO No. 259) *World Meteorological Organization*

Zakhvatkina N, Korosov A, Muckenhuber S, Sandven S and Babiker M (2017) Operational algorithm for ice-water classification on dual-polarized RADARSAT-2 images. *The Cryosphere*, **11**(1), 33–46 (doi: 10.5194/tc-11-33-2017)

Zakhvatkina N, Smirnov V and Bychkova I (2019) Satellite SAR data-based sea ice classification: An overview. *Geosciences*, **9**(4) (doi: 10.3390/geosciences9040152)

APPENDIX A - SENTINEL-1 DATA

For the classification time series, we used in total 277 S1 scenes that were acquired between March 1st and May 31st 2022 and intersect with the larger AOI shown in Fig. 1. A complete list of images can be obtained using an automated search on the Copernicus dataspace platform or by contacting the corresponding author. Additional scenes from previous years and earlier months in 2022 were used for visual inspection and the selection of training data.

Table 5. List of S1 scenes acquired during the CIRFA-22 cruise with KPH in the footprint at the time of image acquisition

Sentinel-1 name string	shown in Figure
S1A_EW_GRDM_1SDH_20220424T071244_20220424T071348_042912_051F67_D6FE	—
S1A_EW_GRDM_1SDH_20220427T073724_20220427T073828_042956_0520D1_72C2	—
S1A_EW_GRDM_1SDH_20220428T081914_20220428T082014_042971_052159_9CD0	6
S1A_EW_GRDM_1SDH_20220430T080147_20220430T080251_043000_052241_0B70	6
S1A_EW_GRDM_1SDH_20220502T074527_20220502T074631_043029_05233F_7BC7	5, 7, 8
S1A_EW_GRDM_1SDH_20220503T082621_20220503T082725_043044_0523D1_AF89	1, 7, 8
S1A_EW_GRDM_1SDH_20220504T072910_20220504T073014_043058_052441_2DD2	6, 8
S1A_EW_GRDM_1SDH_20220505T081019_20220505T081123_043073_0524C2_2C15	5, 6, 8
S1A_EW_GRDM_1SDH_20220507T075351_20220507T075455_043102_0525AC_18EB	—
S1A_EW_GRDM_1SDH_20220508T065615_20220508T065720_043116_052628_47B1	—

520 Table 5 lists the ten S1 scenes that were acquired while KPH was within the footprint of the image

521 during the CIRFA-22 cruise and indicates if (and where) the scenes are shown in this publication.