

Article

Estimating the speed of icegoing ships by integrating SAR imagery and ship data from Automatic Identification System

Markku Similä¹ and Mikko Lensu¹¹ Finnish Meteorological Institute (FMI), PB 503, FI-00101, Helsinki, Finland; firstname.lastname@fmi.fi

* Correspondence: markku.simila@fmi.fi; Tel.: +358-50-4125323

Abstract: Ship speeds extracted from AIS data vary with ice conditions. We extrapolated this variation with SAR data to a chart of expected icegoing speed. The study is for the Gulf of Bothnia in March 2013 and for ships with ice class 1A Super that are able to navigate without icebreaker assistance. The speed was normalized to 0–10 for each ship. As the matching between AIS and SAR was complicated by ice drift during the time gap, from hours to two days, we calculated a set of local SAR statistics over several scales. We used random tree regression to estimate the speed. The accuracy was quantified by mean squared error (MSE), and the fraction of estimates close to the actual speeds. These depended strongly on the route and the day. MSE varied from 0.4 to 2.7 units² for daily routes. 65 % of the estimates deviated less than one unit and 82 % less than 1.5 units from the AIS speeds. The estimated daily mean speeds were close to the observations. Largest speed decreases were provided by the estimator in a dampened form or not at all. This improved when ice chart thickness was included as one predictor.

Keywords: Synthetic aperture radar, automatic identification system, ice thickness, regression

1. Introduction

Ships navigating in ice covered sea areas need ice information to support their operations. In the Baltic Sea operative ice information includes daily ice charts, ice model forecasts and ice thickness charts. The production of the information relies strongly on SAR imagery typically obtained once or twice a day from each sea area. Daily ice charts are expert estimates of the ice conditions based on the ice chart of the preceding day, on SAR images, and on observations from icebreakers and coastal stations. The charts divide the ice cover into regions characterized by set of parameters that comprises concentration, typical, maximum and minimum thickness of level ice types, and a five-degree numeral for the degree of ice ridging (DIR). Ice thickness chart, on the other hand, is a supporting operative SAR based product [18].

It seeks to enhance the level ice thickness information of the ice charts by refining the thickness type regions using SAR segmentation. The product has been operative for several years. It is validated annually using ship reports and more extensively in [19].

It seeks to enhance the level ice thickness information of the ice charts by refining the thickness type regions and thickness distribution using SAR segmentation. The product has been operative for several years. It is validated annually using ship reports and more extensively in [19].

The difficulty of ice navigation is dominated by ice ridges. Also when assisted by an icebreaker the ships must be able to proceed through the thick ice channel rubble that remains from the ridges broken by the icebreaker. The traffic restrictions to Northern Baltic ports are set in terms of the Finnish-Swedish ice classes where class 1A ships may navigate to all ports throughout the winter season. Ships in a more powerful subclass 1A Super are capable to independent navigation unless the conditions are exceptionally difficult. This may be due to very difficult ridging conditions, brash barrier, or ice compression. Brash barrier (windrow) is a typically 2 NM wide and 10 m thick band of

ice rubble created by wave and wind action. During compression the ice cover is in a prevailing stress state that may manifest as channel closing, added resistance, or even pileup of ice against the ship hull.

Also during normal conditions ridge fields complicate of the navigation system and especially the icebreaker assistance by inducing performance variation between ships of different capabilities and speed variation for each individual ship. This includes occasional besetting which in convoy operations increases collision risks and causes delays and disorder in port logistics. Better information on the distribution of ridged ice would improve the predictability of arrival times and make the winter navigation system generally more efficient and economic. For this reason the detection and quantification of ridging from SAR imagery has been investigated from different aspects since late 1980's, for the Baltic Sea studies see, e.g., [16], [?], [4], [30], [6], [10].

On the surface the ridge rubble is usually arranged into elongated sails and the usual parameters to describe ridging are ridge density, or the number of ridge sails along a linear transect, and sail height [25]. These parameters can be linked to the statistics of the subsurface ridge keels which can in turn be used to estimate ship speed reduction. Ridge density is related to the horizontal coverage of ridge rubble which in contributes to the magnitude of σ^o in SAR images. Field campaigns carried out in the Baltic Sea ice have shown that the variation of the C-band σ^o is mainly due to large-scale surface roughness, like ridge rubble [4],[5], [37], [41]. Based on the SAR studies it has been observed that on average the C-band σ_{HH}^o increases when the intensity of ridging increases. The relationship is not unambiguous. Also small-scale roughness, salinity and incidence angle contribute to the backscattering coefficient, e.g., new ice can produce a strong backscattering [29].

However, results reliable enough for operational implementation are still lacking. Attempts to derive ridge height or surface rubble volume per unit area have been inconclusive [7] although the coverage of individual sails increases with sail height. The resolution of SAR imagery used in ice information production is usually too low to make individual ridges visible and the aggregated contribution of ridges to σ^o is from an unknown horizontal sail arrangement in the pixel area. Recently learning methods like random forest classification [10] have showed promise in the determination of DIR numeral from SAR images.

Another line in the utilization of the SAR in winter navigation has been to provide the imagery to the ships navigating in Polar ice covered seas [17], [23], [15]. The ship can use the image as an aid of route selection and tactical navigation by identifying leads, areas of low concentration, or areas of low degree of ice ridging. The ice conditions ahead can be anticipated from the SAR basing on what has been experienced in near past when navigating through ice fields with similar SAR texture. This can be supported by state of art SAR classifications and other ice information products supplied by the land base. The main problems are limited communication bandwidth and the fact that in the presence of ice drift it can be impossible to pinpoint the ship's location in reference to the image even when only few hours old.

Finnish and Swedish icebreakers in the Baltic provide a working example of SAR-based operation planning. The SAR images used to prepare ice charts are also received to an onboard terminal and presented with other selectable layers of navigational and environmental information [23].

As the icebreakers may operate weeks or months in the same sea area the crew learns by heart to link the SAR texture with degrees of navigational difficulty. This is facilitated by the fact that a large part of the SAR texture remains largely unchanged between two images and the ice drift patterns are not too complicated in the semi-enclosed Baltic basins. However, this personal understanding is not easily expressed quantitatively or communicated to others, for example to the Ice Services. The two-way communication between FIS and the icebreakers have been discussed in [9].

The idea that a ship's observed response to the image texture and features is used to interpret the image is inherent in all SAR assisted navigation. We seek to develop this idea into a consistent methodology of SAR classification that can be used to complement other methods to extract navigationally relevant information from the images. The ship response is extracted from AIS (Automatic Information System) messages which report, among other things, the location and speed

of the ships. After the AIS system became mandatory in 2002 the data has been used extensively to study maritime traffic and its impacts.

AIS data is particularly applicable in the Baltic Sea with its intense and steadily growing ship traffic. There are around 2000 sizeable ships at sea at any given moment heading to or from its 200 commercial ports. The annual amount of cargo is about 1 billion tonnes and the ferries carried about 100 million passengers. The main research applications have been general spatiotemporal characterization and statistical analysis of the traffic system in some sea area [1] and the quantification of ship encounters and associated collision risks, e.g., [42].

The cargo and passenger flow continues without interruption throughout the ice season when on the average half of the Baltic area becomes ice covered. Icebreakers are used to keep the ports open also in the northernmost parts of the sea. The presence of ice modifies both the traffic patterns and the risks but also introduces traffic system features not present in open water conditions, especially icebreaker assistance and speed reduction and besetting due to ice conditions. AIS data has been used to assess the winter navigation system [23], navigation risks [11], icebreakers and convoy operations [?], ship ice performance [?], [34], ice resistance modeling [24], and route selection [?].

We shall widen this domain of applications by a new approach where AIS data is used to interpret the SAR images in terms of attainable ship speed and locations of difficult navigation. In the simplest rendering speed reduction observed at a certain location of the SAR image is taken to indicate that the location is navigationally difficult, and if several ships show similar response it is almost certainly so. As such this would already be valuable information along frequently operated ice channels and easily obtained if the ice is not drifting. We aim however to the SAR based classification of whole sea areas by studying the ship response to SAR feature vector constructed for this purpose. The methodology is based on random forest regression, the demonstration is for the ice covered part of the Gulf of Bothnia in March 2013, and the response is described in terms of relative speed of independently navigation 1A Super ships.

The paper is organized as follows. The data sets comprising AIS, SAR and ice chart data are described in Section II. The random tree ensemble method to estimate ship speeds using SAR features and sea ice parameters extracted from the manual ice charts is presented in Section III. The results and quantitative analysis of their accuracy are the topic of Section IV, and the concluding Sections V and VI discusses the results and the future potential of the approach.

2. Data sets and processing

2.1. Ice conditions during the ice season and study period

Our study area in the Baltic Sea is northward from the latitude of 61°N, covering the entire Bay of Bothnia and largely the Sea of Bothnia. The time period is March 2013. The maximum ice extent of the ice season 2012-13 was average but the date to finally attain this was late. In the Bay of Bothnia the pack ice zone started expanding during the first week of January and the basin had complete ice cover on 26 December. However, after a few days winds decreased the coverage to 20 % of the basin area. Full coverage was attained again on 17 January. This was followed by a sequence of cycles of windy periods with opening and deformation followed by more calm periods of thermodynamic growth. Apart from the Eastern Gulf of Finland the ice pack remained close to the coast in other sea areas. Towards the end of February the weather cooled down and new ice was formed also in the Western Gulf of Finland. In the beginning of March cold arctic air started to flow to Scandinavia and the extent of ice began to grow. The whole March was extremely cold. On the March 15 the extent of ice reached 177000 km², which was the maximum of the season although the Sea of Bothnia became fully ice covered first at the end of the month. In April the ice retreated rapidly as the ice types formed in March did not have large thickness. In March the maximum average thicknesses of level ice types were above 50 cm in the Bay of Bothnia, see Figure 1. Also the recurring flaw leads affect the regional thickness variation. Although the ice pack was mobile throughout month the winds were not particularly strong.

Specially in March, the wind speeds fluctuated during the first week. After this the winds settled around 5 m/s with occasional spikes reaching 10 m/s. The wind direction varied much however. In the Bay of Bothnia this opened a flaw lead alternatingly along Finnish and Swedish coasts, see Fig. 1. The lead became covered by new ice that was somewhat deformed as the lead changed to the other side. The main pack of older and thicker ice suffered little changes but was rather shuffled back and forth. In the Sea of Bothnia the ice cover started to expand more rapidly after the first week of March as an interplay of freezing and drifting, and during the last ten days southward ice drift accelerated freezing of remaining open water areas. The ice ridging conditions were not difficult. The highest degree of ice ridging, which would be Ridged 3 in Figure 1, was atypically not present even in the Bay of Bothnia, and the Sea of Bothnia became only slightly ridged due to the moderate winds.

2.2. AIS data

The AIS is a VHF system for broadcasting information on the ship state and it has been mandatory for vessels over 300GT and for all passenger vessels since 2002. Finnish Traffic Administration (FTA) receives Baltic AIS data on terrestrial stations, conducts a quality check, and redirects the data to the VTS centers and other civil end users. The AIS data consists of different types of messages many of which are optional. However, the basic navigational data is automatically coded in position reports (Type 1 – 3 messages) which include ship identifier (MMSI number), location, speed, heading, and course over ground.

This AIS data stream has been received at FMI since 2007, arranged into a database, and used for research purposes. The database links the position reports with marine environmental data and ship particulars data and provides search utilities that can be conditioned by time, geographic area, environmental parameters, and ship particulars. The Type 1 – 3 message interval for a steaming ship is about 10 seconds and number of position reports for years 2007-2016 amounts to 6 billion. The main use of the database has been winter navigation research in which context the environmental data consists of gridded ice charts and ice model forecasts.

The main shortcoming of the AIS data is that the changes of propulsion power are not known. When studying the ship response to ice conditions there is no guarantee that the observed speed reduction is due to increased difficulty of ice conditions and not due to change of power setting. The power setting may be reduced when safe encounter with other ships requires slower speed, increased if the ice conditions ahead pose a besetting risk, or changed either way in order to adjust to convoy speed. To minimize this uncertainty we selected the group of independently navigating ships having Finnish-Swedish ice class 1A Super. The criterion of independency is that there is neither another ship within 3 nautical miles nor any ship expected to come within 3 miles during the next 10 minutes. The 1A Super ships, on the other hand, are able to navigate without icebreaker assistance unless the ice conditions are exceptionally difficult. They usually do not change their power setting much either when navigating independently [26].

The range of FTA terrestrial stations is limited. The coverage of the data decreases when moving close to the western coastline of the Sea of Bothnia. For the Bay of Bothnia the coverage is complete. For the Bay of Bothnia, AIS message intensity in the Figure 2 is measured by the number of position reports from 1x1 NM square per day for the ship class 1A Super. The concentration of the traffic to the main ice channels in the Bay of Bothnia is discernible. Outside these routes the ships have visited almost all locations within the pack ice except an area in the middle of Bay of Bothnia where the ice was very thick (Figure 1). In the fast ice zone the traffic is constrained to narrow channels leading to ports. The average traffic speed is the monthly average of speed records from a 1X1 NM square for the ship class 1A Super. It is a summarizing descriptor that also includes ships that have been idling. It is seen that the traffic speed generally increases with decreasing intensity in the Bay of Bothnia. This is due to powerful ships selecting routes outside the main ice channels occupied by convoys of weaker ships led by icebreakers.

Totally 21 1A Super cargo ships operated in March 2013 in our test area north of latitude 61. There were sister ships and the number of ships with different particulars was 13. They comprised 0.7 million position reports of independent navigation, and made altogether 55 voyages. Size ranged from GRT 4000 to GRT 28000 and the length from 105 to 215 meters. We imposed threshold speed of 1 knot to eliminate from the data the occasions when the ship was idle and just drifted with ice. We further selected from the voyages continuous test routes that constitute the basic data for our analysis. If a voyage took place in more than one calendar day, the track covered during each different day was regarded as a different route. The data for the ships contained gaps whenever the independency condition did not hold any more, and less frequently if the AIS messages were not received from the Western part of the Sea of Bothnia. To eliminate short AIS records it was required that only such transects for which at least 2000 AIS measurements were recorded consecutively during one day were included in the analysis. There were finally 139 routes in the data set with total length of 38500 km. The fraction of the data used in the training of the regression model was about 23 %, the rest of the data was used in the testing. The training data was created by randomly selecting 30 routes from the AIS-data set.

In the analysis we normalized the speed for each ship to 0-10 so that 10 indicate a high speed. A speed limit of 18 knots was set so that all speeds exceeding 18 knots per hour got the value 10. In the AIS measurements about 2 % of all the measurements in the training set got the value 10. After the normalization these relative few extreme speeds, like 20 knots per hour, were not hampering the estimation process as outliers.

2.3. RADARSAT-2 SAR imagery

The SAR imagery used in this study is RADARSAT-2 ScanSAR Wide (RS-2 SCWA) dual-polarized imagery with the HH/HV polarization combination. The nominal size of an RS-2 SCWA image is around 500 by 500 km, and the pixel size is 50 m. The spatial resolution is around 73-163 m by 100 m (range by azimuth). The incidence angle (θ_0) varies from 20 to 49 degrees. The equivalent number of looks (ENL) is larger than six. The noise equivalent σ^0 at both HH- and HV-polarization is around -28 ± 2.5 dB and the absolute accuracy of σ^0 is better than 1 dB [33].

Using the acquired RS-2 SAR imagery a SAR mosaic covering the whole Baltic Sea was formed by employing so many SAR images that no holes in the SAR coverage existed. This Baltic Sea SAR mosaic was updated twice daily. The number of images used in to update the SAR mosaic over the study area varied from 0 to 2 SAR updates per day in March 2013. On average, the study area was partly covered by one new SAR image each day. For a specific location in the study area the SAR mosaic data was 0-2 days old.

The preprocessing of the RS-2 SCWA images consisted of calibration of σ_{HH}^0 and σ_{HV}^0 , georectification, calculation of the incidence angle θ_0 , and land masking. First, the data were rectified into the Mercator projection with 100 m pixel size and reference latitude 61°40'N. This projection is also used by the graphic ice charts of Finnish Ice Service (FIS). This georectification is compatible with the FIS ice charts and the navigation system of the Finnish and Swedish icebreakers. In this Mercator projection the reference latitude is 61°40' N. At the HV-polarization the SAR σ^0 values are close to the instrument noise floor (around -28 dB for RS-2 SCWA mode), and the noise floor (noise equivalent σ^0) varies along the across-track direction. The noise floor modulates the HV channel signal leading to clearly visible stripes (artifacts) in the HV imagery. The calculation of the incidence angle correction is described in detail in [36] for the HH-channel and in [20] for the HV-channel. Due to the large size of the SAR images in the SAR mosaics the resolution was downsampled into 500 m. The daily SAR mosaics were constructed by overlaying all of the SAR data available for each day, i.e. the latest data is shown in the mosaic.

2.4. Ice charts

In our analysis we used also the daily manual ice charts provided by FIS over the Baltic Sea. The ice charting software divides the ice cover into polygons described by quantitative parameters and qualitative characterizations. The polygons and their quantitative parameters in ice charts are also saved as numerical latitude-longitude grids using the ice charting software. The usual grid resolution is 1/60 degree of latitude and 1/30 degree of longitude which is approximately 1 nautical mile (NM) in both directions at 60°N. In the grid format gridded parameters include typical, minimum and maximum thickness of level ice and, ice concentration, and the degree of ice ridging. Level ice thickness and sea ice concentration grids over the study area were derived from the grids. These grids were converted to the Mercator projection used in the rectification of the SAR imagery. When combining the ice chart grids with the SAR mosaic the nearest neighbor interpolation was used to find the chart value for the SAR pixels.

The information in the FIS ice charts is typically reported for rather large areas (hundreds of square kilometers). The level ice thickness values are based on the observations made by the operating icebreakers. Hence, the uncertainty and the variation of these values is considerable. In the graphic FIS charts the thickness values are given as a variation range between minimum and maximum thickness. The gridded charts include in addition typical thickness which is the mean of the range. This value is also rounded to nearest 5 cm if the result is at least 15 cm. We chose to use the mean value without the rounding in order to have denser gradation of the thickness. Sea ice concentration at FIS is based mainly on the SAR imagery and, if the conditions allow, also optical data (MODIS). Also this variable has a significant uncertainty.

3. Methodology for estimation of ship speed

The construction of the regression model consists of two steps: defining a feature vector and then training the regression model. Also the importance of the different features will be evaluated to decide if the statistics in question will be included into the feature vector or not. For the feature vector we first extracted from SAR data selected statistics calculated over sliding windows of different sizes. The center pixel coordinates of the windows corresponded the AIS location. We completed the feature vector using level ice thickness and ice concentration from the ice charts. We connected each AIS position record with the feature vector at the respective location. The ship speed acts as the response variable in regression. A major advantage by including sea ice thickness in the feature vector was that we could separate between open water and ice cover without a separate discrimination phase.

3.1. Features

There is an unavoidable time gap of varying length between the SAR acquisition and the AIS messages. How accurately the ice conditions around the ship at the time of the message and those reflected by the SAR signatures from the same location correspond depends on the magnitude of ice drift during the time gap. In the Baltic, if the prevailing wind increases ice concentration that already exceeds 70 % the ice cover is usually static unless the wind speeds exceed 10-15 m/s. If the concentration is lower or the wind decreases concentration the onset of ice drift may occur at any wind speed. As a rule of thumb, the ice drift speed after drift onset is 1-2 % from the wind speed and speeds exceeding 50 cm/sec may be observed [27]. Thus the ice may drift as much as 2 km per hour although 100-300 m per hour is typical. However, even then the position difference between the AIS message location and its SAR reference can grow to several kilometers during 24 hours. This is partly compensated by the fact that the ice surface can remain statistically rather similar over relative large areas. We can see from the results presented later in Section 4 that the ship speed may exhibit small variation along relatively long transect and substantial variation during both short and long time periods.

The backtracking of SAR information can be attempted by ice drift models and the analysis of the SAR time series itself. However, as this is rather complicated and has its own sources of uncertainty, we adopted for the present study an approach where the statistics are calculated in many scales. The statistical quantities are obtained in four different resolutions: 7 x7 pixels (3.5 km scale), 5x5 pixels (2.5 km scale), 3x3 pixels (1.5 km scale), 1 pixel (0.5 km scale), where the window sizes and the corresponding scales are indicated. Hence a window of 7x7 pixels covers an area of 12.25 km² in the scale of 3.5 km.

In our earlier studies we have observed that spatial autocorrelation and entropy are effective features for the SAR image classification, e.g., [40], [21]. We tested them in the present context with larger windows up to scale of 5.5 km but these features showed too much variation in this scale to be useful components of the feature vector. We specify the applied feature vector later in Section 3.2.2 after first presenting the regression method in more general terms.

3.2. Random forest regression method

As indicated earlier the time gap between the AIS measurements and the SAR data acquisition makes the estimation problem hard. Our first attempt was to fit feedforward neural network to the data to make predictions, see e.g. [2]. This approach gave ship speed v estimates where the mean speed along an individual route was close to the measured one. The weakness of the method was that the values of v were concentrated on very narrow range, the high ship speeds (> 8 units) were missing as were the low speeds (< 6 units).

In search of a method which would yield a larger spread for the estimated speed values v we found that the random forest regression (RFR) [3] worked better in this respect. Random forest is an ensemble learning method which can be applied to classification and regression. Nowadays it is a rather popular method in the data analysis (e.g. [28], [38], [14], [10]). In RFR we artificially generate several training sets from the original training set using bootstrapping, grow a regression tree for each individual training set, perform regression for each tree and then aggregate the results. This is called bagging or bootstrap aggregation [3], [13]. For the analysis we divided our data into the training and the test data sets, see Section 2.2.

To gain intuitive understanding how a regression tree is grown we discuss shortly the basic principles underlying such tree. Our feature space is a p -dimensional rectangle X . Using recursive binary partition the regression tree divides X into M disjoint rectangles $\{R_m, m = 1, \dots, M\}$, which rectangles are represented by the terminal nodes of the tree. In each p -dimensional rectangle R_m we model the response as a constant c_m . Our training data consists of N observations with a p -dimensional feature vector x_i and a response y_i , i.e. the regression is performed using pairs (x_i, y_i) where $i = 1, 2, \dots, N$, with $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$. We denote by $g(x)$ the tree's prediction at input vector x . The residual sum of squares (RSS) is adopted as our criterion function in the automatic splitting a node in a regression tree and also in the selection of the splitting variable. The minimization of $\sum_{i=1}^N (y_i - g(x_i))^2$ implies that the best estimate of \hat{c}_m in Eq. 1 is the average of all y_i in region R_m , i.e. $c_m = \text{ave}(y_i | x_i \in R_m)$ [13]. The regression tree is:

$$\begin{aligned} g(x) &= \sum_{m=1}^M c_m I_{x \in R_m} \\ &= T(x, \Theta), \end{aligned} \quad (1)$$

Hence the structure of the regression function is a step function. The parameter Θ characterizes the tree T in terms of split variables, split points at each node, and terminal node values.

3.2.1. Description of the algorithm

One of the advantages of RFR is that the feature vector can consists of many kinds of variables, e.g. continuous, discrete or categorical. An example of an categorical variable is level ice thickness which

has six possible categories in our data set. In the previous Section the structure of a single regression tree was shown. Here we outline the algorithm to calculate the regression tree ensemble in terms of bootstrap samples from the training set Z . A bootstrap sample from Z is denoted by Z^* . Our bootstrap sample Z^* is of the same size as the original sample, so on average the fraction 63% of the original samples of Z belong to it, the rest being duplicates [8]. The samples of Z left out from Z^* (about 37% of the samples) are called out-of-bag (OOB) samples.

We generate B bootstrapped training sets Z_b^* , $b = 1, \dots, B$ and relying on every training set we grow a regression tree $T_b(\Theta_b)$. We obtain an ensemble of trees $T_b(\Theta_b)$ with $b \in \{1, \dots, B\}$ where for every tree a regression estimate is computed as in Eq. 1. The regression estimate by $T_b(\Theta_b)$ depends on the feature vector x which is used as input for the tree. This is denoted by $T_b(x; \Theta_b)$. A regression tree often achieves a rather low bias if it is grown deep with many nodes without pruning [13]. On the other hand, a single tree has a large variance.

In regression we record all the regression estimates provided by the ensemble of B trees for a specific feature vector, i.e. we have B regression estimates at our disposal. If we average the regression estimates, then the ensemble estimate has a smaller uncertainty than a single regression tree, because an average has a smaller variance than a single variable. This is true also for the correlated variables. The problem with the presented bagging approach is that the grown trees are correlated. To reduce this correlation RFR has a randomization step [3]. When building a tree, each time a split in a tree is considered, a random sample of m predictors is chosen as split candidates from the full set of p predictors. A new sample of m predictors is taken at each split. This step prevents the same features to dominate the ensemble. The flow of the random forest algorithm is presented below.

Random forest algorithm for regression

1. For $b = 1$ to B :
 - (a) Draw a bootstrap sample Z_b^* of size N from the training data.
 - (b) Grow a random-forest tree $T_b(\Theta_b)$ to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size is reached.
 - i. Select m variables at random from the p variables.
 - ii. Determine the best variable and split-point among the m variables using the RSS criterion.
 - iii. Split the node into two daughter nodes.
2. Output the ensemble of trees $\{T_b(\Theta_b)\}_1^B$.

Regression estimate for a new feature vector x_n :

Let $T_b(\Theta_b)$, $b = 1, \dots, B$ be the ensemble of all the regression trees constructed as described in the steps i to iii . Then the random forest predictor is

$$g^B(x_n) = \frac{1}{B} \sum_{b=1}^B T(x_n; \Theta_b). \quad (2)$$

In the presented algorithm we have taken the average of all regression estimates. It is possible to use also other statistics. In fact, we have applied in our inference the quantiles computed from the set of regression estimates.

We run all our random forest regressions with the same set of tuning parameters using the routine *TreeBagger* [31]. From the set of sixteen ($p = 16$) features we randomly chose $m = 4$ features to be

used in a single split. The components of the feature vectors as well as the value of m were selected on the basis of our experimental runs and the evaluation of the importance of a feature component, see 3.2.2. If B is large enough then the random forest algorithm avoids the tendency of overfitting the model which often occurs in the context of the decision trees. In our analysis we used $B = 200$ trees. We did not observe any notable improvement in the results when the amount of trees was larger.

3.2.2. Importance of the features

The OOB samples are used to estimate the predictor importance. All OOB feature vectors, i.e., about $N/3$ vectors, are run through the trees. Consider the importance of the statistics x_k in the feature vector. All splits in a single tree are examined where the variable x_k is used in the binary partition of an internal node. The RSS before the split and after the split are recorded. The change in RSS is divided by the number of the internal nodes. This procedure is repeated for every statistics x_k for all the OOB feature vectors. Finally all the changes are accumulated over the ensemble of the regression trees separately for each variable. This gives us the predictor importance values shown in Fig. 4

Into Table 1 we have collected all the variables in the feature vector whose predictor importance value exceeded 2.0. There were totally 8 such statistics extracted from the SAR imagery and 2 ice descriptors extracted from the FIS charts, ice concentration and level ice thickness (h_i). They have been ordered according to their importance. For the SAR statistics we have included the used scale and polarization and its importance value. We have also indicated which feature component the statistics corresponded in Fig. 4.

The SAR statistics belonging to a feature vector but not included in Table 1 were similar to the included variables but calculated either in smaller scale or/and in different polarization. We tested the addition of the coefficient of variation in the scales 3.5 km and 2.5 km using both polarization modes but the results did not improve. They were left out as were the standard deviations (std) of the σ_{HH}^o values in all scales for their negligible importance. Hence only the very basic SAR statistics were used in regression: mean, standard deviation, maximum and minimum. We note that the importance of one variable depends which other variables are included in the feature vector. We performed an experimental RFR run using only the 10 most important features. The results were, however, more erroneous than the ones obtained using the combined feature vector with all the 16 components.

It is worth noting that the scales 2.5 km and 3.5 km yielded the best results. This is a reasonable result because the locations of the AIS measurements in the SAR image had a great uncertainty. We observe that the maximum and minimum values in the HH-channel in the 2.5 km scale were highly important. From the HV-channel statistics other important measures were both the mean and the std values in the 3.5 km scale as well as the HV minimum in the 2.5 km scale and the mean HV values in the same scale. The most important HH-channel mean value was in the scale of the 3.5 km.

4. Results

5. Estimation accuracy

The main statistics to measure the estimation accuracy was the mean squared error (MSE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (v_i - \hat{v}_i)^2, \quad (3)$$

where we denoted the response variable by v_i , its estimate by \hat{v}_i and n refers to the sample size.

The mean absolute percentage error (MAPE), also known as mean absolute percentage deviation (MAPD), is a measure of prediction accuracy used often in forecasting. It usually expresses accuracy as a percentage, and is defined by the formula:

$$M = \frac{100\%}{n} \sum_{i=1}^n \frac{|v_i - \hat{v}_i|}{v_i}, \quad (4)$$

where the symbols are the same as in Eq. 3. The interesting feature in MAPE is that it gives a large weight to a situation where v_i is small and the prediction \hat{v}_i is essentially greater. This kind of situations are of particular interest to us.

5.1. Estimation accuracy in general

In Figure 5 are shown the distributions of estimated and measured speeds for the whole data set. The mean speeds of distributions were almost equal, 7.57 speed units for measurements and 7.55 units for estimates. Of significance is that the estimated values are concentrated in narrower range than the estimated values. About 6 % of the measurements were below 5 units but only 1 % of the estimates. Estimates below 3.5 units did not appear at all. In the upper tail of the distribution 16 % of the estimates and 27 % of the measurements exceeded 8.5 speed units. No estimated speed value was above 9 units but 11 % of the measurements were. This limited range of the estimates compared to the measurements is a typical limitation of the model of this kind.

In Fig. 6 the assessment of the estimation accuracy is based on pointwise (pixelwise) differences between individual speed measurements and their estimates. We observe that in the pixel scale 0.5 km 65 % of all estimates are within 1 speed unit, 75 % within 1.25 units and 82 % within 1.5 units from the correct value (colored blue in Fig. 6). If we label all the errors with absolute deviation value over 2 units in scale 0.5 km as severe, then the amount of severe errors is 9 %.

As our first analysis using the random forest regression underestimated on the average slightly the routewise mean speeds we began using the mean of two quantile regressions for quantiles 0.2 and 0.85. This improved somewhat the results and was used in all reported results. This includes also the results shown in Figures 5 and 6.

5.2. Estimation accuracy in different ice categories

In Figures 7 and 8 we see the estimated and measured daily mean speeds plotted around the blue line which corresponds to an exact match between the speeds. The parallel green lines deviate from central line by one unit. The quantile regression can also be used to create an approximative confidence band for the estimates. This was not utilized, however, because the width of the confidence band did not indicate well whether the prediction was close the measured value or not.

Looking at Fig. 7 or the Table 2 in Section 5.5 we note that the mean speed error exceeds 0.5 speed units only in 3 routes out of 29. The maximum mean speed error was -0.8 units, i.e., the estimated speed was larger. When taking average over all routes the estimated mean speed was only 0.013 units higher than the AIS based mean speed.

To obtain a better understanding on how the ice thickness affects the accuracy of the daily mean speeds we generated similar scatter plots for the following ice thickness classes: cl10 (1-10cm), cl20 (11-20cm), cl30 (21-30cm), cl40 (31-40cm), cl50 (41-50cm) and cl60 (51-60cm). We observe from Fig. 8 that for thin ice categories cl10 and cl20 almost all the estimation results were within 1 unit from the measured means. In cl30 the mean measured speeds and estimates were quite close to each other except for some cases of lower measured speed when overestimation took place. In two cl40 cases high measured speed (> 8 units) resulted in underestimation. As discussed earlier our regression procedure could not reproduce very low speeds which occurred in thicker ice classes. The measured mean speeds were in cl50 mostly below the mean speed of the whole data set (7.5 units) and all the measured mean speeds were less than 8 units in cl60. On the other hand, the presence of very small mean speeds (< 5 units) in cl50 and cl60 resulted in overestimation.

If the ships use channels made by the icebreakers, their speed in thick ice can be rather high in comparison with the icebreaking mode. The presence of open channels made speed estimation less accurate in thick ice than in thin ice in the Bay of Bothnia, manifesting as underestimation in the categories cl50 and cl60. In cl60 over half of the cases were such that RFR overestimated the mean speed. If the regression model underestimated the mean speed, the recorded speed was rather high (> 6 units) considering ice thickness. It is likely that the ships used in these cases the existing channels although the exact information is lacking.

5.3. Estimation of speed variation

Much of the information of ice conditions is contained in the speed variation. One index for the navigability of a certain route is its standard deviation of speed. Hence it is of interest look how well the variation is preserved in the estimated speeds along the route. The estimated speed variation is significantly smoother the recorded speeds, see Figures 11 and 13. Over all routes the mean std of measured speed variation is 1.31 units but the mean std of the estimated speed is only 0.83 units for all routes, i.e., 63 % of the measured variation. The respective ranges for the standard deviations are from 0.86 units to 1.78 units for measurements and from 0.52 units to 1.13 units for estimates.

Another measure for the preservation of the speed variation is how close the histograms of the measurements and the estimates are each other. We investigated their relationship using the Kullback-Leibler distance (KLD) which for discrete distributions P (measurements) and Q (estimates) is defined as follows [22]:

$$D_{KL}(P||Q) = - \sum_i P(i) \log \frac{Q(i)}{P(i)}. \quad (5)$$

Hence KLD is the average difference in log probability between the true (P) and estimated (Q) distributions, i.e., its value tells how much information is lost when Q is used instead of P . If we have a pair of estimated distributions, then the candidate that minimizes KLD will be closer to the true distribution [32]. KLD is not symmetric, i.e. $D_{KL}(P||Q) \neq D_{KL}(Q||P)$. The daily KLD values can be found from the Table 2.

When we inspect Fig. 9 (left panel), it is evident that on average the estimation accuracy decreased when the standard deviation of measurements increased. For this data set the prediction uncertainty according to MAPE increased from 6 to 24 %. There was not such clear connection between KLD and MAPE except for very large KLD values. We believe that the lack of the correspondence between KLD and MAPE is due to the fact that no spatial information is incorporated in the KLD value unlike in the MAPE measure.

5.4. Speed charts

The major goal in this study was to create an estimated speed chart over the test area where every pixel value represents a typical speed for the specific ship class 1A Super. We will now discuss two examples of such charts, one successful and one less successful.

In Fig. 10 we present March 8 results. The SAR mosaic of the test area is in panel (a) and the histograms are in the panel (b). The corresponding speed chart created by estimating every SAR pixel using the RFR method is in the panel (c) where the speed range is restricted from 5.5 to 10 units for sharper visualization. The along route measured speed is indicated by three colors: red for the AIS speeds from 1 to 6 units, blue for speeds from 6 to 8 units, white for speeds over 8 units. Upon the speed chart in panel (d) is marked how close the regression estimates are to the measured AIS speeds. If the difference is less than 1.5 units, the corresponding pixel is colored by green. Otherwise it is marked by red. From the Table 2, see Section 5.5, we can read that MSE was 0.95 units² on March 8 2013 which is a good result because the speed variation along the route was large. The AIS histogram in the panel (b) shows that the speeds were distributed rather evenly for a wide range, from 5.5 to 9.5 units. Also lower AIS speeds were recorded. The estimated histogram showed also a wide spread with two peaks, at 6 units and at 8-8.5 units. The std of the estimated speed was close the std of the measurements, the values being 1.13 units and 1.47 units, respectively. The Kullback-Leibler distance describing the similarity of histogram had a value 1.24.

In the panels of Figure 11 three different lines are drawn. The black line represents level ice thickness. It is scaled so that the actual thickness value in centimeters is divided by 10. Hence the value 2 refers to thickness 20 cm and likewise the other values. The blue color represents the measurements and red the estimates. When we examine the estimated and measured speeds along the route of length 750 km in Fig. 11, we note that usually these two speed values are close to each other. Most of differences occurred in rather thin ice areas, panels (c) and (d), when the level ice thickness was either 10 cm (c) or 20 cm (d). Otherwise the agreement was very good, e.g., in panel (a) the speed estimates also exhibited variation in thin ice as when the measurements did due to the SAR statistics. The speed variation increased often when the ship moved from one ice regime to another. This can be seen clearly in Figure 11. It means that transitions from one ice category to another were accompanied often with speed change. As we saw in Fig. 8 the most problematic areas from the point of view of estimation are thick ice areas.

In Fig. 12 we have a similar panel than in Fig. 10 with identical ordering. This time the date is 20 March. It illustrates a case where the estimation was less successful. The AIS speed variation was high with a large amount of measured AIS speeds exceeding 9 units. Also very low AIS speeds, less than 5 units, were recorded. The estimated speeds had a more limited range from 6 to 9 units. The std of the AIS speeds was 1.78 units and the std of the estimates was significantly lower, just 0.93 units. Hence the measured std was almost twice as large than the estimated. Also the Kullback-Leibler distance had a very large value ($KLD = 3.55$) indicating a large discrepancy between the measured and estimated histograms. We see from lower right panel in Fig. 12 that large estimation errors (red color) occurred both in thin and thick ice areas.

The total length of the transects was again about 750 km in Fig. 13. We inspect now in more detail where the estimation failed. In transect (a) the agreement is pretty good both in thin and thick ice. In transect (b) both large and low ship speeds were recorded in thick ice (cl50). This lead first an underestimation of ship speed in cl50 and later overestimation. The AIS recordings showed speeds close to 10 units in part of the cl50 transect. The panel (b) is an example of the ambiguous nature of ship speed in thick ice. In the third transect (c) the AIS recordings were again close to 10 units in cl50. In the lower panels(d-e) the largest differences between measured and estimated speeds took place in cl40 and cl30, i.e. in thinner ice. The occurrence of short cl50 intervals did not have a large effect on the measured or estimated ship speed. Instead the somewhat thinner ice (cl40) slowed the ship speed although this was not reflected in the estimated speed. The discrepancy between the measurements and estimates continued in transect (e) in thin ice areas. Finally, we note that the average speeds

taken over the daily data were quite close each other, 7.59 for the AIS measurements and 7.67 for the estimates, see Fig. 12 (lower left panel).

5.5. Collected results

We collected in Table 2 several different daily statistics. We have already discussed the mean speed estimation and its accuracy. Here those values are given for each day. Another measure of the estimation accuracy is how large fraction of the speed estimates were close to the respective AIS measured speeds. We define that the estimate is close to the measurement if their difference is less than 1 unit. The daily fluctuation of this fraction was large ranging from 41 % to 87 %. The mean for the whole month was 64 %. The spatial information is incorporated in this concept. If we consider estimates with an allowed deviation of 1.5 units from the correct value, the fraction of the included estimates varied from 63 % to 96 % per route, the average being 81 %.

We looked also how the lack of fresh SAR data affected the estimation. The results varied strongly, see Table 2, depending on how fast the local ice conditions were changed.

6. Discussion

6.1. Uncertainties

Our approach is a theoretical development of the basic idea that ships respond with their speed to the change of ice conditions in a consistently similar manner. For certain ice type the speed response is described by average speed and speed variation and the starting assumption is that these correlate with the average ice thickness and thickness variation. There are however three missing pieces of information that make the straightforward implementation of this idea problematic. First of all, the propulsion power is not included to AIS messages. With such data a rapid progress would be possible using the engineering models linking propulsion power, hull particulars and the along route ice profile to the speed of a specific ship type. Such models could also provide theoretical basis for more accurate prediction of the ice performance of some ship type from the observed performance of another ship type. The very different icegoing capabilities of ships is precisely the second missing piece in an approach relying solely on the use of AIS data. The third is that already in ice conditions of minor difficulty the majority of the ships follow frequently navigated ice channels. For the channels not only the ice resistance but also the characteristics of dynamic and thermodynamic evolution are different from the surrounding intact ice cover.

We sought to minimize these problems by selecting the subclass of independently navigating 1A Super ships. These are able to make steady progress through Bay of Bothnia ice cover with typical degree of deformation and do not change their power setting much. The number of ship types was about 10 and the class generally shows much less performance variation between ship types than the next class or 1A. The 1A Super ships also often select routes outside the main channels in order to take over convoys. Although the progress in the channels is faster, they still display thickness variation of the surrounding ice cover to some extent as channel typically is filled with rubble of different thicknesses. Especially the thicker rubble sections from broken ridges have intervals and thicknesses related to the density and size of the original ridges. In the 500 m resolution data the ice channels are not visible and a pixel with a channel becomes classified about similarly to a pixel comprising only the ambient intact ice.

The uncertainties from channel navigation, power setting changes and variation of performance capabilities could be further reduced by targeting more to the speed variation. This has characteristics forced by the ridging statistics which component should be about similar for all ships and navigation conditions. The speed variation in itself is a useful descriptor of navigability as, apart from characterizing the degree of deformation, it can be related to besetting probability. It is precisely for this reason ships with ice class 1A or lower often increase speed in thicker ice, that is, the power setting is selected in order to keep speed variation amplitude within acceptable bounds and not in

order to maintain some average speed. To include lower ice classes would greatly increase the data amounts. The present results intend to demonstrate the soundness of the approach using a limited dataset. It scratches the surface of the data comprising billions of AIS position reports from ten years and covering most variable ice conditions. Apart from increasing the data amounts there are several possible lines of research that are addresses in the next section.

6.2. Future prospects of the approach

As far as we know ours is the first attempt to generate ice navigation information from direct classification of SAR imagery with ship data. Similar methods have great application potential. Approaches seeking to generate navigationally relevant information by linking SAR signatures with ship performance can be divided into three types: data driven, classification oriented, and hybrid.

In the Baltic more than 60 million position reports from icegoing ships are obtained during every month of the ice season and it is perfectly feasible to generate maps of navigational difficulty from this data only. A primitive data driven approach would simply identify the SAR locations where several ships have difficulty and transfer this cumulating information to the next image in the series. The time gap between ship observations and images could be handled with ice drift models and SAR based methods but as well by expert estimates. This would fit to the daily practice of the FIS ice charting as it already requires the estimating of how the boundaries between ice types change due to ice drift. This is helped by the fact that most traffic is constrained to ice channels which can be rather easily identified from the high-resolution SAR images (resolution of 100 m or finer). They also make the problem essentially one-dimensional as along channel conditions are sufficient. Although navigation outside the main channels is common, see Fig. 2, providing SAR based navigability indices for them would already greatly improve the predictability of the winter navigation system.

On the other hand, our approach is classification oriented. Such methods seek to extrapolate the information gained from ship response to certain signatures and texture parameters to the whole image area. The extrapolated parameter can be speed, as in the present case, some numeral for navigational difficulty, or the load level experienced by the ship. There are also simpler options to advanced methods and automated classification of the whole image. In the Baltic the areas with difficult ice conditions are often discernible bandlike formations created by historic stormy periods. The difficulty experienced during some transect across the band is likely to be close to the truth when extrapolated to the whole band. This is especially so for brash barriers (windrows) that usually have very homogeneous properties in along band direction. In the Baltic operative practice this could combine the ice expert's estimates with automated classifications. The latter, on the other hand, provide new possibilities for ice routing where the objective is to optimize the route of independently navigating ships with respect to travel time or fuel consumption. The bottleneck has been that ice information has lacked the accuracy and detail that engineering models calculating ship speed from ice conditions require. Our approach in fact indirectly demonstrates that routing is feasible in the Baltic. From the ship data set about a quarter of the measurements was used for training and the rest for validation. As well the ship data collected during the first quarter of some time period could be used to generate speed map used in routing during a certain time period to follow. The data from the routed ships can then be used to update the speed maps for the ships arriving later by rerunning the regression procedure using the data set augmented with new observations. The applied regression method works the better the more versatile and comprehensive its training data set is. So the incremental increase of the training data set should improve the quality of the regression estimates.

In hybrid approach the AIS based SAR classification is combined with other classification methods, especially those seeking to retrieve ice thickness, ice concentration, and ridging parameters. Our approach applied the thickness and concentration from ice charts but it could have used as well SAR based products that enhance the chart information or are used in FIS to assist ice charting [18]. As there are established models on how ship performance depends on ice conditions these two types of variables can in principle be theoretically connected in SAR classification methods. Also, it is in

principle possible to derive ice parameters from ship speed records. A promising future direction of research is integrating AIS based SAR classification with methods seeking to retrieve ridging for SAR as ridging induces ship speed variation that is related to the ridge density and ridge height. In [10] we introduced a method to retrieve the ridging classes. The chart of the ridging classes can be combined to build a hybrid ridging/speed retrieving method in several ways, e.g., by adding the ridging class as one categorical variable to the current regression model. Further developments could also utilize ice models and kinematic algorithms for ice drift retrieval.

7. Conclusion

We have presented an approach that uses position and speed data from the AIS messages of icegoing ship to translate SAR images over ice cover to ship speed maps. The methods were demonstrated for Baltic SAR mosaics for March 2013 and for a ship class with good ice performance. When developed further the approach has potential to answer to various needs of the ice navigation community, including improved ice information and functioning route optimization, and has also relevance for the maritime engineering research seeking to model ship ice performance. The method relies on random forest regression and the applicability of this advanced and recently much used method becomes demonstrated in the present context. The results were rather good, the RSR was able to generate useful speed maps from a limited training set although there are several uncertainties involved in the use of AIS data. Several possible lines of future research are immediate.

Data availability. All Matlab programs to analyze the data can be obtained by contacting the first author. The access to the AIS data is restricted. It requires the acceptance by the Finnish Traffic Administration. The SAR data can be delivered only for research purposes and requires the approval by FMI.

Author contributions. Formal analysis, Markku Similä and Mikko Lensu; Funding acquisition, Mikko Lensu; Investigation, Markku Similä and Mikko Lensu; Methodology, Markku Similä; Software, Markku Similä; Writing – original draft, Markku Similä and Mikko Lensu; Writing – review & editing, Markku Similä and Mikko Lensu.

Competing interests. The authors declare that they have no competing interest.

Acknowledgements We thank Alexandru Gegiuc and Juha Karvonen for the processing of the SAR mosaics. We also thank Eero Rinne for proofreading the paper and making valuable comments. Markku Similä and Mikko Lensu were supported by the BONUS STORMWINDS project funded jointly by the EU and the Academy of Finland (contract 291683).

1. Arguedas, V.F., Pallotta, G. and Vespe, M., Maritime Traffic Networks: From Historical Positioning Data to Unsupervised Maritime Traffic Monitoring. IEEE Transactions on Intelligent Transportation Systems, 19(3), pp.722-732, 2018.
2. Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer-Verlag, Berlin, 2006.
3. Breiman, L.: Random forests, Machine Learning, 45, 5–32, 2001.
4. Calström, A., Ulander, L. M. H.: Validation of backscatter models for level and deformed sea ice in ERS-1 SAR images, Int. J. Remote Sens., 16, 3245–3266, 1995.
5. Dierking, W., Pettersson, M. I., Askne, J.: Multifrequency scatterometer measurements of Baltic Sea ice during EMAC-95, Int. J. Remote Sens., 20, 349–372, 1999.
6. Dierking, W., Dall, J., Sea-Ice Deformation State From Synthetic Aperture Radar Imagery—Part I: Comparison of C- and L-Band and Different Polarization IEEE Transactions on Geoscience and Remote Sensing 54(11):3610–3622, 2006. DOI10.1109/TGRS.2007.903711
7. Dierking, W., Lang, O. and Busche, T., Sea ice local surface topography from single-pass satellite InSAR measurements: a feasibility study. The Cryosphere, 11(4), pp.1967-1985, 2017
8. Efron, B. and Tibshirani, R.J. An Introduction to the Bootstrap, Chapman & Hall/CRC Monographs on Statistics & Applied Probability, CRC Press, 1994.

9. Eriksson, P.B., Majaniemi, S., Lensu, M. and Karvonen, J., February. Met-Ocean Services by FMI for Ice Management and Shipping in Ice. In OTC Arctic Technology Conference. Offshore Technology Conference, 2014.
10. Gegiuc, A., Similä, M., Karvonen, J., Lensu, M., Mäkynen, M., and Vainio, J.: Estimation of degree of sea ice ridging based on dual-polarized C-band SAR data, *The Cryosphere*, 12, 343-364, 2018.
11. Goerlandt, F., Goite, H., Banda, O.A.V., Höglund, A., Ahonen-Rainio, P. and Lensu, M., An analysis of wintertime navigational accidents in the Northern Baltic Sea. *Safety science*, 92, 66-84, 2017.
12. Goerlandt, F., Montewka, J., Zhang, W. and Kujala, P., An analysis of ship escort and convoy operations in ice conditions. *Safety science*, 95, pp.198-209, 2017.
13. Hastie, T., Tibshirani, R., Friedman, J. , *The Elements of Statistical Learning*, 2nd ed., Springer, New York, 2011.
14. Hu, S.; Liu, H.; Zhao, W.; Shi, T.; Hu, Z.; Li, Q.; Wu, G. Comparison of Machine Learning Techniques in Inferring Phytoplankton Size Classes. *Remote Sens.*, 10, 191, 2018.
15. Hui, F., Zhao, T., Li, X., Shokr, M., Heil, P., Zhao, J., Zhang, L. and Cheng, X., Satellite-Based Sea Ice Navigation for Prydz Bay, East Antarctica. *Remote Sensing*, 9(6), p.518, 2017
16. Johansson, R. , Askne, J., Modelling of radar backscattering from low-salinity ice with ice ridges, *Int. J. Remote Sens.* 8, (11), 1667-1677, 1987.
17. Johannessen, O.M., Alexandrov, V.Y., Frolov, I.Y., Sandven, S., Pettersson, L.H., Bobylev, L.P., Kloster, K., Smirnov, V.G., Mironov, Y.U. and Babich, N.G., Application of SAR for ice navigation in the Northern Sea Route. *Remote Sensing of Sea Ice in the Northern Sea Route: Studies and Applications*, 323-396, Springer, 2007.
18. Karvonen, J, Similä, M., Heiler, I., Ice Thickness Estimation Using SAR Data and Ice Thickness History, *Proc. of the IEEE Int. Geoscience and Remote Sensing Symposium 2003 (IGARSS'03)*, Toulouse, France, 2003.
19. Karvonen, J. , Similä, M. , Haapala, J. , Haas, C. and Mäkynen, M., Comparison of SAR Data and Operational Sea Ice Products to EM Ice Thickness Measurements in the Baltic Sea , *Proc. of the IEEE International Geoscience and Remote Sensing Symposium 2004 (IGARSS'04)*, Anchorage, Alaska, 2004.
20. Karvonen, J., Evaluation of the operational SAR based Baltic Sea ice concentration products, *Advances in Space Research*, 56(1), 119-132, 2015.
21. Karvonen, J., Shi, L., Cheng, B., Similä, M., Mäkynen, M., Vihma, T. Bohai Sea Ice Parameter Estimation Based on Thermodynamic Ice Model and Earth Observation Data. *Remote Sens.*, 9, 234, 2017.
22. Kullback, S. , *Information Theory and Statistics*, John Wiley & Sons, New York, 1959.
23. Kotovirta, V., Jalonen, R., Axell, L. Riska, K., Berglund, R., A system for route optimization in ice-covered waters, *Cold Regions Science and Technology*, 55(1), 52-62, 2009.
24. Kuuliala, L., Kujala, P., Suominen, M. and Montewka, J., Estimating operability of ships in ridged ice fields. *Cold Regions Science and Technology*, 135, 51-61, 2017.
25. Lensu, M., The evolution of ridged ice fields. A dissertation for the degree of Doctor of Philosophy. Helsinki University of Technology, Ship Laboratory Report Series M-280, 140 p, 2003.
26. Lensu, M., Assessing the ice performance of ships in terms of AIS data. In *Proceedings of the International Conference on Port and Ocean Engineering Under Arctic Conditions*, 2015.
27. Leppäranta, M., *The drift of sea ice*, 2nd Edn., Springer-Praxis, Heidelberg, Germany, 2011.
28. Lindner, C., Bromiley, P.M., Ionita, M.C., Cootes, T.F., Robust and Accurate Shape Model Matching Using Random Forest Regression-Voting, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 37(9), 1862-1874, 2015.
29. Lundhaug, M., ERS SAR studies of sea ice signatures in the Pechora Sea and Kara Sea region. *Can. J. Remote Sens.*, 28:2, 114-127, 2002.
30. Manninen, A., Surface morphology and backscattering of ice-ridge sails in the Baltic Sea. *Journal of Glaciology*, 42(140), 141-156, 1996. doi:10.3189/S0022143000030604
31. MATLAB and Statistics Toolbox Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States.
32. McElreath, R., *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, Boca Raton, 2016.
33. MDA (MacDonald, Dettwiler and Associates, Ltd.), Radarsat-2 product description, Tech. Rep. RN-SP-52-1238, Issue 1/11, MDA, BC, Canada, pp. 84., May 5, 2014.
34. Montewka, J., Goerlandt, F., Kujala, P. and Lensu, M., Towards probabilistic models for the prediction of a ship performance in dynamic ice. *Cold Regions Science and Technology*, 112, 14-28, 2015.

35. Montewka, J., Goerlandt, F., Lensu, M., Kuuliala, L. and Guinness, R., 2018. Toward a hybrid model of ship performance in ice suitable for route planning purpose. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, p.1748006X18764511, 2018
36. Mäkynen, M., Manninen, A., Similä, M., Karvonen, J., Hallikainen, M.: Incidence Angle dependence of the statistical properties of the C-Band HH-polarization backscattering signatures of the Baltic sea ice, *IEEE Trans. Geoscience and Remote Sensing*, 40, 2593–2605, 2002.
37. Mäkynen, M., Hallikainen, M.: Investigation of C- and X-band backscattering signatures of the Baltic Sea ice., *Int. J. Remote Sens.*, 25, 2061–2086, 2004.
38. Ruescas, A.B.; Hieronymi, M.; Mateo-Garcia, G.; Koponen, S.; Kallio, K.; Camps-Valls, G. Machine Learning Regression Approaches for Colored Dissolved Organic Matter (CDOM) Retrieval with S2-MSI and S3-OLCI Simulated Data. *Remote Sens.* 2018, 10, 786.
39. Similä, M., Leppäranta, M., Granberg, H. B., Lewis, J. E., The relation between SAR imagery and regional sea ice ridging characteristics from BEPERS-88, *Int. J. of Remote Sens.*, 13:13, 2415-2432, 1992. DOI: 10.1080/01431169208904279
40. Similä, M.: SAR image segmentation by a two-scale contextual classifier, *Proc. SPIE Conf. Image and Signal Proc. for Remote Sensing*, v. 2315, Ed. J. Desachy, 434–443, 1994.
41. Similä, M., Mäkynen, M., Heiler, I.: Comparison between C band synthetic aperture radar and 3-D laser scanner statistics for the Baltic Sea ice, *J. Geophys. Res. Oceans*, 115, C10056, 2010.
42. Zhang, W., Goerlandt, F., Montewka, J. and Kujala, P., A method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering*, 107, 60-69, 2015.

Table 1. The 10 most important features

Statistics	Polarisation	Scale [km]	Component	Importance
maximum	HH	2.5	13	3.3
h_i			2	3.3
minimum	HH	2.5	16	2.3
std	HV	3.5	11	2.6
minimum	HH	2.5	16	2.3
mean	HV	3.5	8	2.3
minimum	HV	2.5	16	2.3
mean	HV	2.5	9	2.1
IC			1	2.1
mean	HH	3.5	5	2.0

Table 2. Estimation results for March. Symbols for the SAR update: 2=totally updated SAR coverage, 1=partly updated, 0=no updates.

day	speed (M)	speed diff. (M-E)	MSE	MAPE [%]	KL-d.	estimates [%]	SAR update
2	7.703	0.348	1.901	15	0.779	45	2
3	7.406	0.039	1.088	12	0.539	66	2
4	7.494	0.238	0.975	11	1.387	74	0
5	7.593	0.160	1.224	12	0.836	64	2
6	7.734	0.322	1.412	14	1.181	57	2
7	7.849	0.271	1.548	16	3.251	59	1
8	7.053	-0.266	0.951	12	1.242	75	2
9	8.118	0.113	0.440	6	0.329	87	2
10	7.274	-0.050	1.343	15	2.692	64	1
11	8.037	0.692	1.969	16	5.347	41	0
12	7.405	0.311	1.590	16	1.361	59	2
14	7.512	-0.038	1.182	13	1.034	67	2
15	7.979	0.231	1.217	11	1.333	67	2
16	8.271	0.410	1.208	12	2.991	58	2
17	7.403	-0.262	1.063	12	1.704	74	1
18	7.252	-0.080	1.368	14	0.538	62	0
19	6.943	-0.294	1.535	17	0.946	65	2
20	7.591	-0.079	2.690	22	3.554	52	2
21	6.900	-0.480	1.804	19	0.933	52	0
22	6.751	-0.638	1.779	18	0.620	66	2
23	7.868	-0.099	1.571	16	3.260	63	2
24	7.115	-0.034	2.266	22	4.049	50	0
25	7.637	-0.221	1.640	17	1.476	71	1
26	6.794	-0.755	2.938	24	1.759	49	1
27	7.222	0.055	0.867	11	0.626	76	0
28	7.961	0.106	1.352	13	2.130	61	0
29	7.356	-0.120	1.267	13	0.716	71	2
30	8.120	-0.066	0.707	8	1.179	85	0
31	7.683	-0.126	1.286	13	1.040	64	1

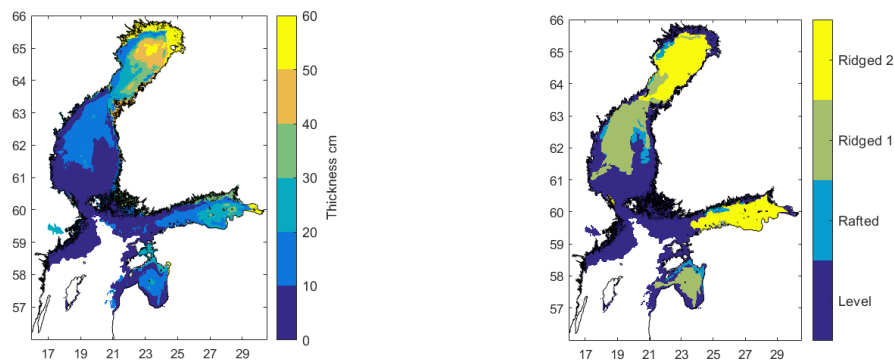


Figure 1. Mean ice thickness and degree of ridging March 2013 (left,right).

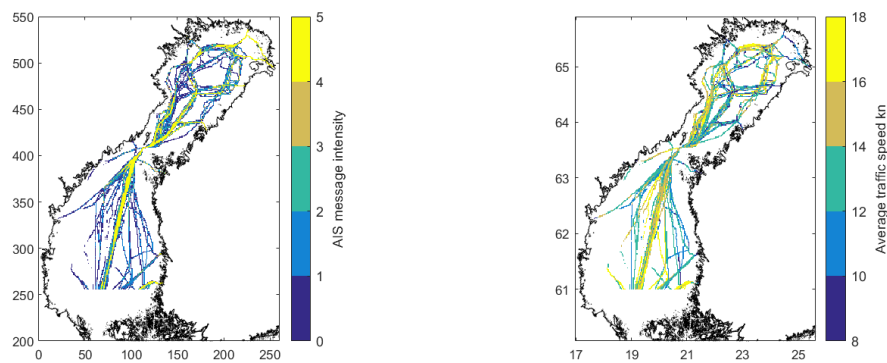


Figure 2. Intensity of AIS messages and average measured speeds of 1A Super ships in March 2013 (left,right).

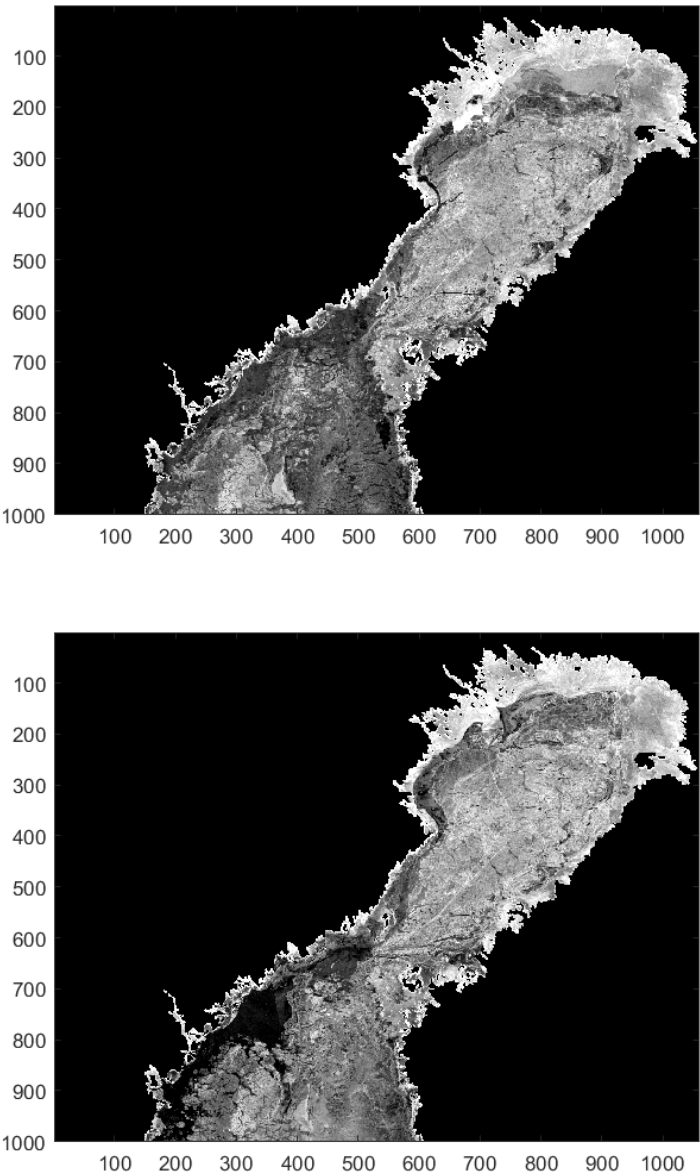


Figure 3. Ice movement in two successive days, March 28 and March 29 2013 (upper panel, lower panel).©MDA

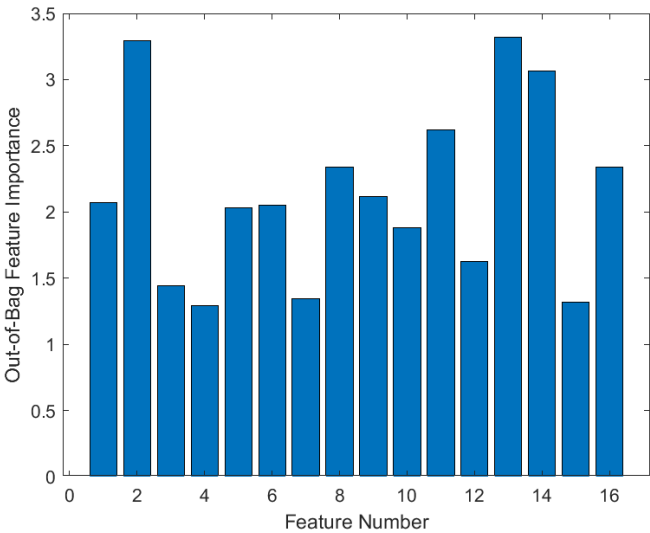


Figure 4. The importance of different statistics in a feature vector in the regression analysis. See Table 1 for the explanation of the feature components.

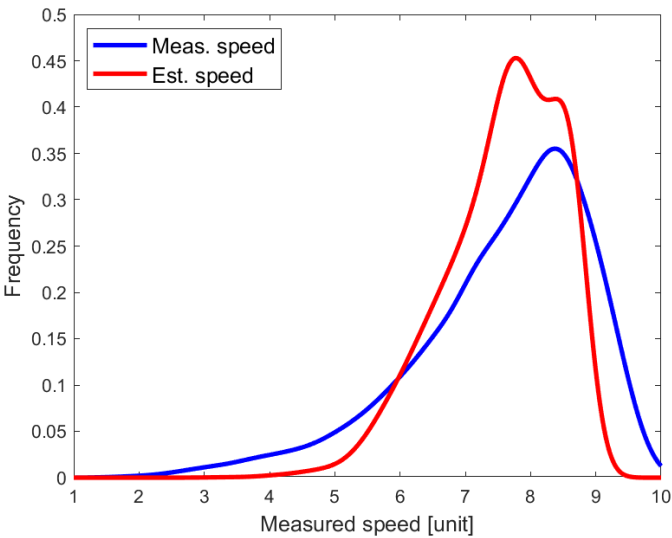


Figure 5. The measured (blue line) and estimated (red line) speed distributions.

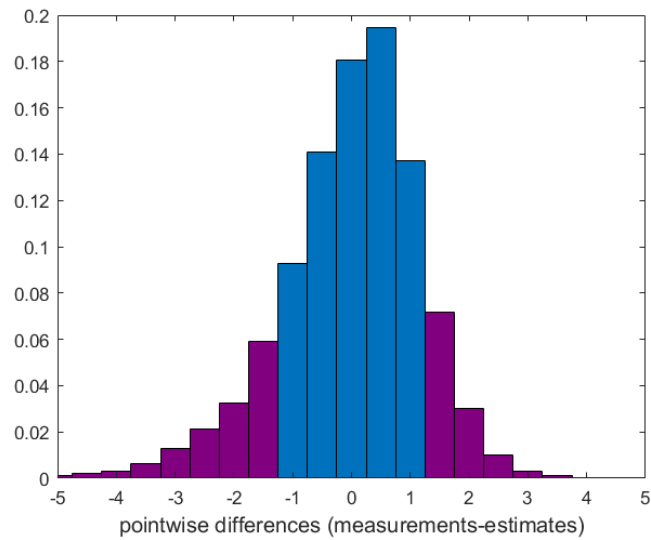


Figure 6. The pointwise differences in scale 0.5 km between AIS measurement and estimates for the whole test data set.

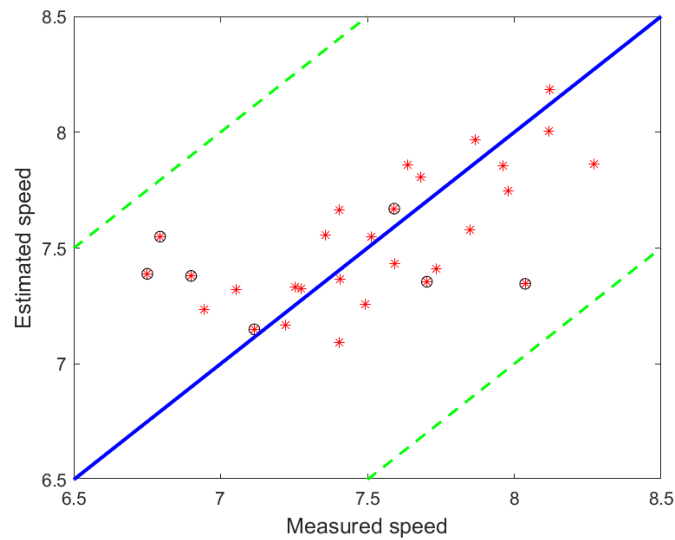


Figure 7. The measured and estimated mean speeds for every route.

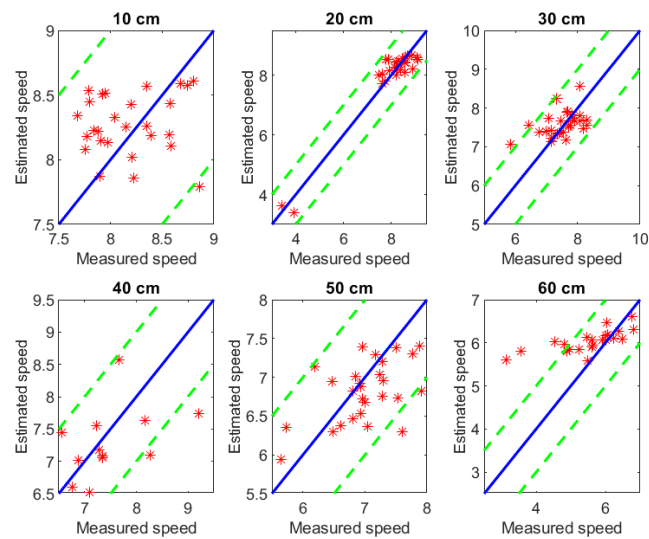


Figure 8. The measured and estimated mean speeds for each ice class.

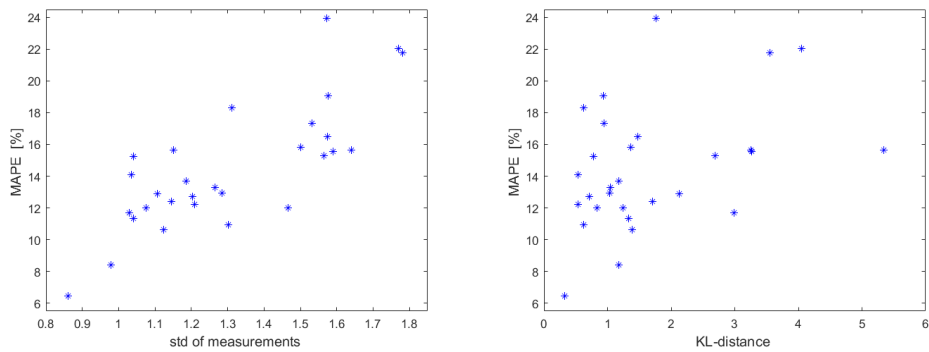


Figure 9. The routewise std of measurements and MAPE (left). The KL-distance and MAPE (right)

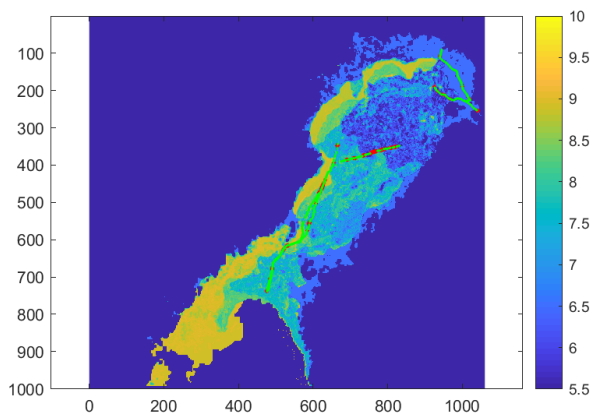
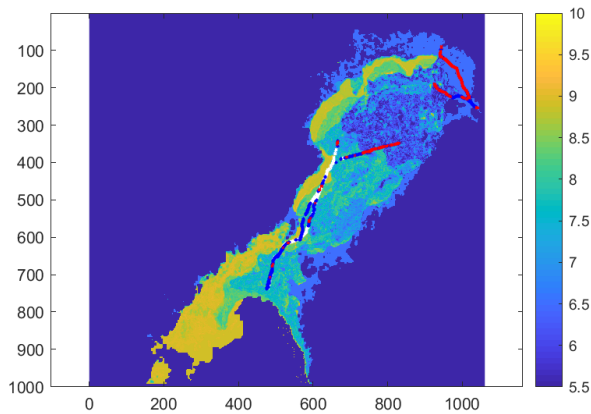
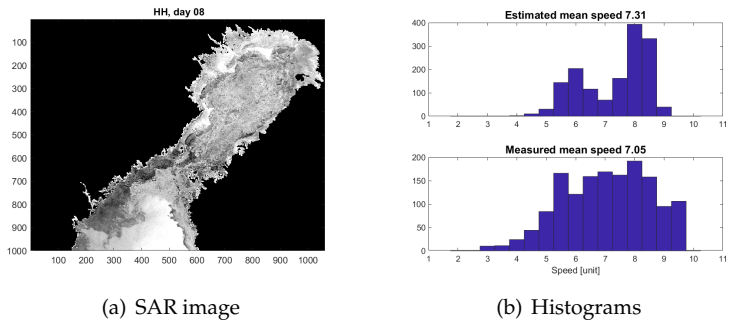


Figure 10. March 8 results

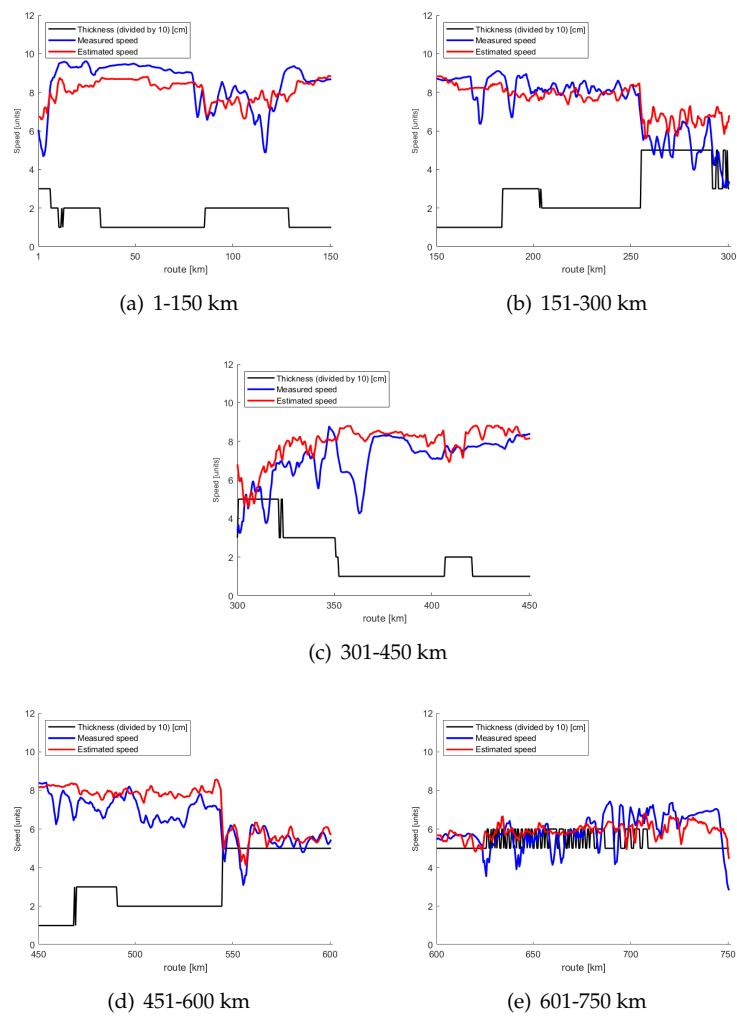
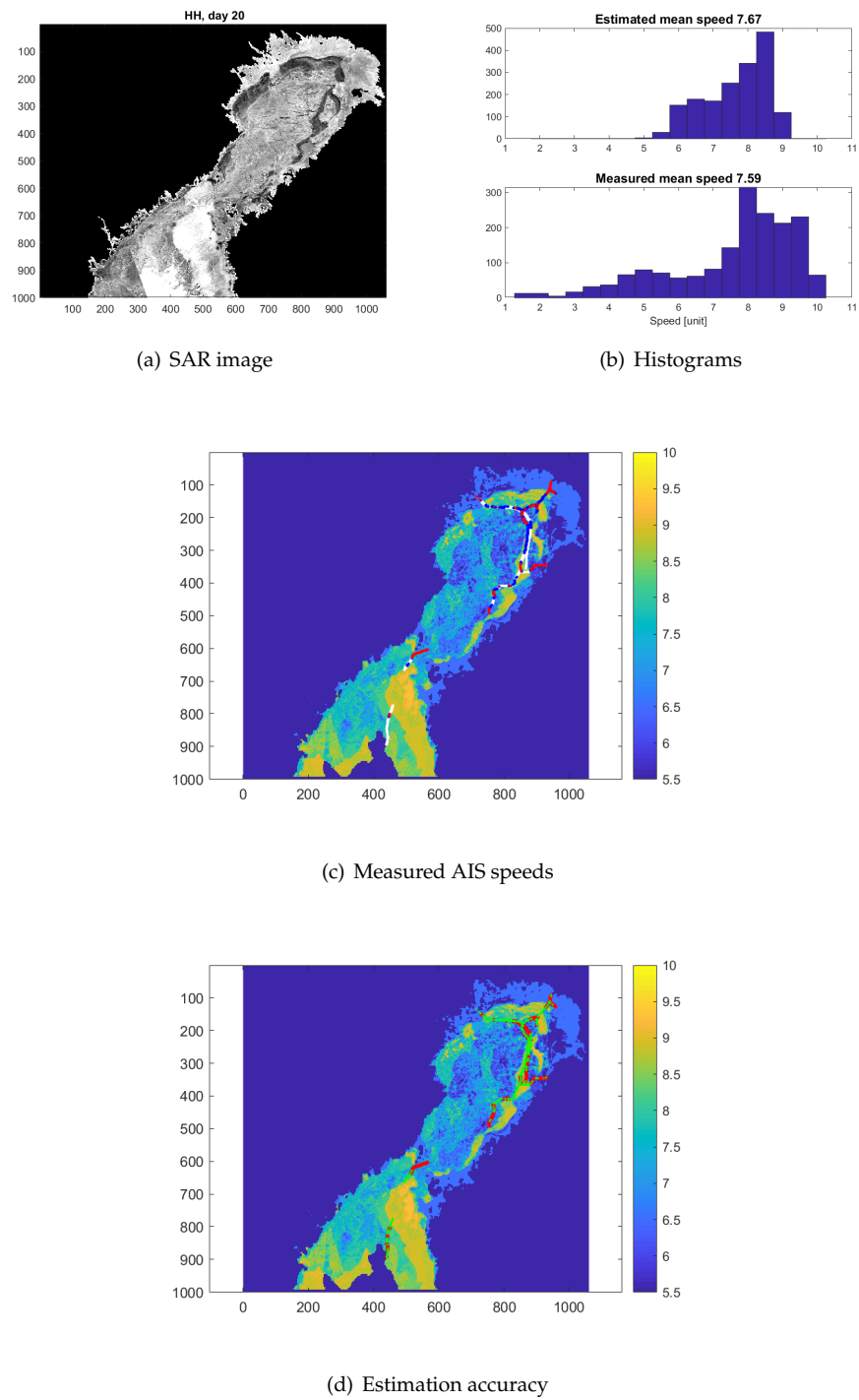


Figure 11. March 8 routes. For explanation of the used coloring see text.



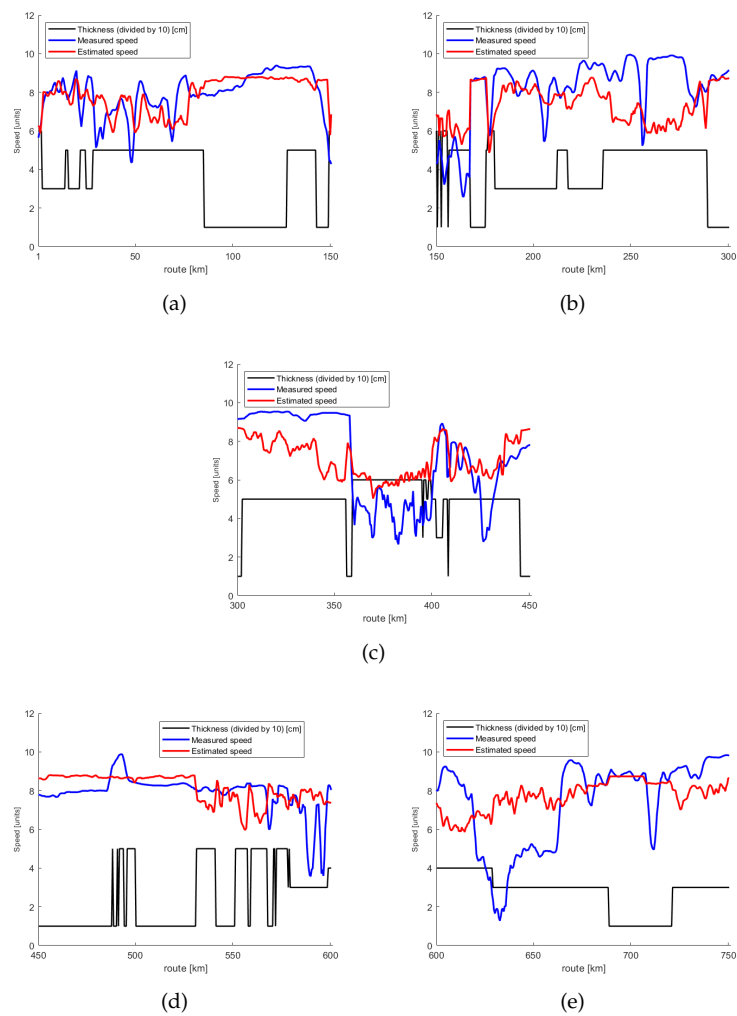


Figure 13. March 20 routes. The coloring and lines as in Fig. 11.