Big maritime data for the Baltic Sea with a focus on the winter navigation system

Published in: MARINE POLICY

DOI: 10.1016/j.marpol.2019.02.038

Published: 01/06/2019

Please cite the original version:
Lensu, M., & Goerlandt, F. (2019). Big maritime data for the Baltic Sea with a focus on the winter navigation system. MARINE POLICY, 104, 53-65. https://doi.org/10.1016/j.marpol.2019.02.038
Big maritime data for the Baltic Sea with a focus on the winter navigation system

Mikko Lensu\textsuperscript{a,}\textsuperscript{*}, Floris Goerlandt\textsuperscript{b,1}

\textsuperscript{a} Finnish Meteorological Institute, P.O. Box 503, FI-00101, Helsinki, Finland
\textsuperscript{b} Aalto University, School of Engineering, Department of Mechanical Engineering, Marine Technology, P.O. Box 15300, FI-00076, Aalto, Finland

\textbf{ARTICLE INFO}

\textbf{Keywords:}
AIS data
Databases
Data integration
Marine information
Marine models
Ice conditions
Ice navigation
Baltic Sea

\textbf{ABSTRACT}

The automatic identification system (AIS) has become a key element in maritime domains of inquiry and the number of related articles has increased rapidly. The systematic integration of AIS data with other datatypes has received less attention and has mostly resulted in application-specific datasets that are small relative to the available AIS data. This work presents an accumulating multi-purpose database for the northern Baltic Sea that combines nine years of AIS data with marine environmental data. The main application is winter navigation research, for which purpose the environmental data is from ice charts and ice drift models. The AIS data is from terrestrial stations and amounts to 6 billion messages. It has a full update rate which is also required for the analysis of ice navigation as this involves close encounters, icebreaker assistance, convoy operations, and rapid speed changes. To identify and study such traffic features, distances between ships that are close to each other are included in the database. Application examples are given for spatial traffic statistics, reduction of ship speed with increasing ice thickness, and for icebreaker assistance.

1. Introduction

The automatic identification system (AIS) is a system for enhancing the safety and efficiency of navigation. It is designed to automatically provide information about a ship to other ships and to coastal authorities. Regulation 19 in SOLAS (Safety of Life At Sea) Chapter V (Carriage requirements for shipborne navigational systems and equipment)\textsuperscript{[1]}, supplemented by the European Union Directive 2002/59/EC (known as ‘the Vessel Traffic Monitoring Directive’)\textsuperscript{[2]}, lays down the regulatory demands of vessels to be equipped with AIS. These are supplemented with guidelines for onboard installation\textsuperscript{[3]} and guidelines for the operational use\textsuperscript{[4]} of AIS. The technical details of the AIS and especially of the STDMA (self-organized time-division multiple access method) are found in Ref.\textsuperscript{[5]}. The STDMA method uses the precise timing of GPS signals to synchronise multiple data transmissions from numerous users on a single narrowband channel, in effect making AIS possible.

The main intention of the mandatory carriage of AIS is to support navigation and assist in collision avoidance decision-making onboard, as well as to improve maritime situational awareness for vessel traffic services (VTS)\textsuperscript{[6]}. While this remains the main practical use of this equipment, recorded AIS data has also been a rich resource for maritime authorities to gain insight into the spatial and temporal characteristics of maritime transportation activities in particular sea areas\textsuperscript{[7–9]}.

In academic research, AIS is rapidly becoming a key element in various maritime-related domains of inquiry. Fig. 1 gives an overview of the number of articles where AIS data has been used in the 2001–2017 period. A separation is made between articles focusing on the operational use of AIS, e.g. for maritime surveillance for prevention\textsuperscript{[10]} or response\textsuperscript{[11]}, and articles where AIS is used to support policy decisions, e.g. for ship emission estimations\textsuperscript{[12]}, or fishery management\textsuperscript{[13, 18, 49]}. The data used in this analysis is obtained from a search in Scopus, a major abstract and citation database of peer-reviewed literature. The search was conducted on 31 May 2018, using the keywords “AIS data” and “ship”. An additional screening was made to remove irrelevant articles from the results based on the information provided in the title, abstract, and keywords. Where these provided insufficient information, the article itself was inspected. The top and bottom diagrams of Fig. 1 show articles with operational and policy-oriented emphasis respectively. In both diagrams the articles are divided into two groups: in the first group, only AIS data is in the focus, whereas in the latter group, AIS data is used in conjunction with other data sources.

\textsuperscript{*}Corresponding author.

E-mail address: mikko.lensu@fmi.fi (M. Lensu).

\textsuperscript{1}Present affiliation: Dalhousie University, Department of Industrial Engineering, Halifax, Nova Scotia, B3H 4R2, Canada.

https://doi.org/10.1016/j.marpol.2019.02.038

Received 14 July 2018; Received in revised form 18 January 2019; Accepted 21 February 2019

0308-597X/ © 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).
From the figure, it can be seen that the number of articles using AIS data has increased dramatically over the last two decades. The earliest publications focused exclusively on the operational use, but the policy-oriented uses have gained increased interest in the last decade. It can also be seen that AIS data is relatively often combined with other data sources, first in operations-related research, and more recently also more in policy-oriented work. The literature search also revealed the wide spectrum of research areas where AIS data is used, including transport engineering and management [14–16], maritime economics [17], fisheries [13, 18, 49], maritime risk and safety analysis [19–21], ship performance analysis [21], and regarding the ecological impacts of shipping [12, 22, 23].

In the research literature it is usually reported how the applied AIS data is organised for access, managing and updating and how it is possibly linked with other data types. Any such system can be called a database. Often geographic information system (GIS) software is used. More recently several authors have presented the development of integrated databases in a more technical sense, including selected applications of these databases [24–27]. These often apply some established database or data warehouse software. This is an important issue to facilitate future research efforts, because, as outlined in Section 2, most research efforts utilize relatively small, purpose-built datasets for specific research purposes, which are usually not easily re usable for other purposes. Furthermore, the existing literature on AIS database development only includes AIS data, whereas recent research efforts show the great research potential of linking this with other (marine) data, as outlined in Section 2. Although the integration of marine data is quite straightforward in itself, advanced database solutions like data warehouse could offer more flexibility by allowing incorporation of less structured data. This applies especially to maritime risk analyses that often involve a very large number of risk factors of different origin. The applicability of well-arranged AIS data to support maritime policy issues has been highlighted as well [28, 29].

In the Baltic Sea area, the seasonal presence of sea ice is a significant obstacle to ship navigation and has important implications for the navigational practices of maritime transportation in the area [30]. As in other sea areas, it is meaningful to develop an integrated AIS database, and especially to link this data with other data sources, including but not limited to data related to sea ice.

In light of this, the objectives of this paper are two-fold. First, the development of an integrated, extendable, and searchable database linking AIS data with other marine-related data is presented, with a focus on winter navigation. The navigation has a systemic nature due to icebreaker assistance and being restricted to ice channels. Operations must be coordinated and unexpected events for an individual ship may propagate to affect the whole traffic situation. With integration it is understood that both the state of the ship and state of the marine environment around the ship are linked, which enables research on their interaction. Second, a number of analysis results of this integrated database are shown to demonstrate its applicability to support marine policy issues, with a focus on winter conditions.

The remainder of this article is organised as follows. Section 2 provides background about the research use of AIS data, in particular its integration with other data sources, and the database in the context of Big Data. In Section 3, the integrated database is presented, with application examples shown in Section 4. A discussion is provided in Section 5. Section 6 concludes the paper.

2. Background and related work

In the Baltic Sea area, AIS data has been integrated with ice parameters for case studies of ship performance in ice [31], accident and oil spill risk modelling [32], analysis of icebreaker escort and convoy operations [33], and for the analysis of navigational accidents [34]. These studies integrate AIS and ice forecast data for selected ships and voyages. A more comprehensive integration of all traffic in the northernmost part of the Sea of Bothnia during January–April 2011 is made in Ref. [35] and applied to study ship speed reduction in ice. This links ice data from the Swedish ice forecast high resolution operational model for the Baltic Sea (HIROMB) with 14,000 AIS reports with reduced update rates. In the Arctic Ocean, comprehensive AIS-based analyses of traffic statistics [36] and emissions [37] exist, but only small-scale case studies have used AIS data to study the effects of sea ice on Arctic ship operations [38].

In other marine environmental data, oil spill backtracking poses the problem of optimally combining AIS data and drift modelling [39]. AIS data also provides source data for acoustic propagation models and noise pollution inventories [40]. In emission modelling, integration with wave model data may improve results [41]. The general potential of integrated AIS and marine environmental data is recognized in Ref. [42] where also the inherent problem of spatiotemporal match is discussed. The same problem is central in the fusion of satellite AIS data with SAR imagery [43] or marine radar data, and with electronic chart display and information system (ECDIS) data [44]. The aims of data fusion approaches are usually different from the data integration discussed in this paper, but the methods of spatio-temporal interpolation and matching are transferable. Also their user interface solutions with layers and pop-up data windows are potential superstructures for an integrated AIS database. The same comments apply to the incorporation of AIS data in GIS-based systems for spatial maritime planning [45].

Searches in the scholarly literature for ‘big data’ result in hundreds of thousands hits. The term is applied in contexts where the data amounts are very large, or where the data is too complex and heterogeneous for traditional statistical methods. Common big data methods include, amongst others, machine learning, advanced regression, clustering algorithms and network analysis. Method-oriented work often develops and demonstrates the tools with a restricted dataset. This is also the case for most applications of big data analytics in AIS contexts, e.g. Ref. [46]. Relating to the potential amounts of data the term is warranted. In 2015, all of the steaming in the world for the year was
estimated to be close to 3 billion nautical miles using global AIS data [47]. This results in about 200 million unique messages per day from steaming ships, assuming an average speed of 15 knots and a conservative estimate of a 10-s update rate. Global data providers such as MarineTraffic [46] and Orbcorn [47] report the reception of about 0.5 billion messages per day, including duplicates. The data aggregated in repositories contains both terrestrial data and intermittent satellite data. The latter comprises about 20 million messages per day. Terrestrial data is received from ships closer than 50 NM to the coastline and covers a large part of the most congested sea areas. Full update rate is usually unnecessarily high for most applications and resampling is needed for better comparability of the mixed data. For most global analyses, the variations in the full update rate data from interpolated resampled data are negligible. Examples include data on general shipping statistics [48], ship emission inventories [12], and sustainability of fisheries [49]. These utilize datasets with up to 20 billion messages. In similar studies on coastal seas the terrestrial datasets can have comparable magnitudes [14,50]. Only a few approaches utilizing tools common in big data contexts use large or global datasets [51]. In most studies with billions of messages, the data has been arranged for a specific purpose and the other potential uses are not considered. Database oriented approaches are more common in ship and navigation research, but the datasets are not very large [33–35,52]. One exception is [27] which reports 1.5 billion messages and a daily reception of 66 million messages.

3. The integrated AIS database

3.1. Database arrangement

The structure of the database is shown in Fig. 2. In consists of the core data, sources of external data, and functionalities to search, filter, process, analyze and present the data. In its present setup the database is used in studies for which archive data is sufficient and it has been updated after each ice season. However, the plan is to develop it into a real-time resource updated with daily data products (Section 5.2) and used also to generate information products for the maritime community. The database uses MATLAB software.

The core data is persistent and accumulating the content of the database. The basic data integrates the AIS position reports with gridded marine environmental data. A number of environmental parameters are connected to each report. The environmental data is sourced from daily operational products. In addition, each reporting ship is linked to a separate database on ship particulars.

The basic data incorporates the data as it is in the AIS reports and environmental products, while derived big datasets are processed from the basic data and have the same extent. The most important derived datatype is proximity data related to the distance between ships. Supporting datasets can be linked to the ships or spatiotemporal locations of the basic data. There are not specific formats and the number of different supporting data types is not limited. Navigationally important sets include the static and voyage data extracted from AIS messages, information on loading conditions and port visits, piloting details, information on traffic restrictions, and data on the changing routing network for icebreaker operations.

The core data components and functionalities are described in more detail in Sections 3.2–3.4, and selected applications are presented in Section 4. The core part is complemented by external data sources characterized below.

In present use copies of the database are accessible without a network connection on external drives. The addition of new data types to the basic data is limited by the practicable size of the database, which for the moment is 2TB. Linkable big datasets consist of data products that could be incorporated with the basic data but are left out due to the data magnitude or some other issue. Regular and comprehensive operational data products are preferred in the basic data, while satellite data especially is often spatially patchy and temporally irregular. The spectrum and volume of linkable data is limitless and its use is often project specific. The required parts are incorporated with the same routines as used for the basic data, in which case the database functionalities also apply to them. The linkable datasets are from marine environmental models, satellite images, and coastal radars. The model data may consist of less important variables, and can have a higher spatiotemporal resolution, or lack spatial or temporal coverage, or come from alternative models. The satellite data includes images mostly from radar satellites (Synthetic Aperture Radar, SAR), or products derived from them. Coastal radar data is in principle similar to SAR but has a very high local spatiotemporal resolution.

Case specific datasets are sets of observational data that can be used to complement the Core Data and validate it. For winter navigation research, the most useful data is from ship ice transits and consists of propulsion data, data on ice loads against the ship hull, and observations and measurements of ice parameters. Propulsion data is the key missing element from the basic data and is currently available only for a small fraction of all navigation. When stored on board, as it is in regular practice, it is usually not accessible or is proprietary. Important non-shipborne datasets include helicopter-borne ice profile data, from which ice thickness and ice ridging parameters can be obtained. Although much can be gained with the basic data, for research on ice-going ships the case specific datasets are instrumental and can greatly improve large statistical analyses.

3.2. AIS data

AIS data has 27 types of messages used for different purposes [5]. The receiving station adds a header to the messages which is not standardized but usually includes the date and time. The remaining part has a fixed structure. First this contains the message metadata, including NMEA (National Marine Electronics Association) message types, channel information, and structure identifiers needed in case the information is divided between several messages. Next comes the message payload and finally the checksum. The payload has binary coding which is explained in the AIS manuals. All messages contain a Maritime Mobility Service Identity (MMSI) number reported in the IMO (International Maritime Organization) number reported in messages of Type 5. For most maritime analyses only messages of Type 1–3 and 5 are utilized. Types 1–3 are the Class A Position Reports that include the basic navigational data. They have essentially identical content but serve somewhat different operative purposes. There are also slightly different Class B Position Reports for light commercial ships and leisure boats. Type 5 is static and voyage related data where the static data contains ship identifiers and particulars, and the voyage data describes the destinations and estimated arrival times (ETA). The voyage data needs to be updated manually by the ship’s crew. Apart
from a few navigation status numerals, the Type 1–3 data include the following: the Maritime Mobile Service Identity (MMSI) number that identifies the ship radio station, latitude, longitude, speed over ground (SOG), course over ground (COG), heading (HDG) and rate of turn (ROT). The time is reported as seconds of coordinated universal time (UTC) and the receiving station provides the remaining part of the time tag.

The AIS data is transferred to the Finnish Meteorological Institute (FMI) in real time from the Finnish Traffic Administration (FTA). The FTA receives the data from terrestrial stations shown in Ref. [52]. Messages are error checked and made available by the FTA for vessel traffic service (VTS) centers and other users. The end user must add a time stamp. The data is complete, containing both Class A and Class B messages with full update rates. The rate is for Class A position reports from steaming ships about once every 10 s or better [5]. This also resolves speed variations resulting from the ship’s response to local ice conditions. Researchers must manage both the transfer and the archiving themselves and cannot publish results for identifiable ships without permission from both the FTA and the shipowner. The FTA also provides a web interface with a real-time AIS stream with full update rates for research institutions and at a reduced rate for application developers and the general public. This is intended for traffic situation displays.

The total area of the Baltic Sea is 377,000 km². The range over which the FTA stations provide full coverage is typically 100 km, after which the reception is intermittent and depends on conditions. The range of applicable data is typically 150 km, which is sufficient to cover the Bay of Bothnia (36,000 km²), the eastern part of the Sea of Bothnia (total area 79,000 km²), the Northern Baltic proper, and the Gulf of Finland (30,000 km²) (see Fig. 3a). More intermittent data is obtained even further. The fading of message numbers with distance can be seen in Fig. 3b. Close to the Swedish coast of the Bay of Bothnia slight intermittency may be found, but the difference to the Finnish side is mostly due to the larger number of visits to Finnish ports [53]. In the western part of the Sea of Bothnia and in the Gulf of Riga the coverage is intermittent and dependent on radio weather. As a rough estimate, from 1 to 10 percent of all AIS messages are included from the Sea of Bothnia that lies west of 19 °E and is shown in blue in Fig. 3b (fewer than 1,000 messages per square nautical mile). In the Baltic proper the intermittent data reaches south of 57° N. The range of reception must be taken into account in the winter navigation analyses. During mild ice winters the data covers all pack ice navigation well as only the Bay of Bothnia and Gulf of Finland have pack ice. During average ice winters the ice navigation data from the western half of the frozen Sea of Bothnia is incomplete, and during severe winters when the Baltic proper freezes it is missing from the Southern Baltic.

The FTA data can be completed with reduced update rate data which is available from commercial providers or the from intergovernmental body the Baltic Marine Environment Protection Commission (HELCOM). The compiling of an AIS data archive by HELCOM is sustained and based on agreements between all Baltic rim states. The data has been collected since 2005 and covers the whole Baltic with few temporal gaps. Although the update interval is only 4–6 min the amount of data in 2014 was 1.54 billion messages mostly from the Southern Baltic [54]. As the focus of the present database setup has been winter navigation analyses involving speed variations and details of convoy operations or encounters, the HELCOM data has not been mixed with the full FTA update rate data but will be included as a separate datatype.

The FMI receives and stores complete AIS from FTA. Transmission interruptions have not been infrequent and the daily AIS dataset may be incomplete or missing. The Class A position reports are decoded and stored as daily matrices typically containing 2–3 million messages during the ice season. Due to the winter navigation focus the annual data is arranged into navigation years extending from the beginning of July to the end of June next year. The navigation years are referred to in periods such as 2010–2011. The numbers of AIS position reports are shown in Table 1 where the incomplete season 2015–2016 ends in March, and the spatial distribution is shown in Fig. 3b. The data amounts to 5.7 billion messages overall and 2.5 billion messages for ice transits. The loss of temporal coverage of AIS data is due to interruptions to the data transfer from FTA to FMI. Although the total data loss is considerable, the ice season is less affected as the missing data is mostly from the ice-free season. The data cannot be fully recovered as the FTA has not stored the data and HELCOM has stored it at a reduced rate. The data loss may affect studies of winter navigation characteristics that require AIS data at the full update rate, especially the response to ice, wave and wind conditions and the detailed analysis of wintertime navigational situations and convoy operations.

3.3. Marine environmental and ship data

3.3.1. AIS data combined with environmental data

AIS position report data is stored as daily matrices in the database. A number of marine environmental parameters are attached to each position report. These are gridded environmental data types and the values have a time tag and pertain to the cells of a fixed grid. Each AIS position report is assigned the parameters pertaining to the grid cell where the ship is navigating at the time of broadcasting its position report. The time difference between AIS and environmental data is minimized. Daily operative products are used as this ensures the reliability, consistency and continuation of the data. The environmental

Fig. 3. a) Northern Baltic Sea and the average annual number of ice days. b) The number of database messages per square nautical mile.
data is arranged in daily matrices with the same number of rows as the AIS matrices. In addition, selected data grids are made accessible to serve as general synoptic reference of the environmental conditions.

The main source of inaccuracy is the time gap between the AIS position report and the environmental data. This mostly concerns ice-covered areas and is discussed for each data type below. Apart from ice data, the AIS reports have been integrated with wave model data that includes, among other variables, significant wave heights and wave directions. This has navigational relevance, but, respecting the focus of this study, a detailed description is given for the ice data only.

### 3.3.2. Ice chart data

Ice charts are published daily by the Finnish Meteorological Institute Ice Service [55]. During the very first and last phases of the ice season the charts may be bi-weekly. In the charts, the ice cover is divided into polygons characterized by a set of parameters related to the ice thickness, ice concentration, and the degree of ice ridging (Table 2). The main information sources are radar satellite (SAR) images and observations. The ice thickness values are based on observations made by icebreakers and, for the fast ice zone, at fixed stations. The station values are from drilling, while the icebreaker values are estimates from upturning floes and occasional drillings. The thickness values refer to ice types with a flat surface that can be level or rafted ice and they seek to characterize the regional conditions. The concentration values and the degree of ice ridging are based on SAR images and observations made by icebreakers. No rules to estimate the total thickness from the ice chart thicknesses and the degree of ice ridging have been consolidated yet although a clear correlation exists [56].

Apart from the graphic charts, gridded versions with different resolutions are prepared for various purposes. The database incorporates gridded ice chart data matching the grid of the FMI operative ice forecast model (Helsinki Multicategory Sea Ice Model, HELMI). The model is described in the next section. The grid covers the Baltic from the latitude 56.74°N northward. The resolution is 1/60 degrees in the N-S direction and 1/30 degrees in E-W direction, which corresponds to a 1 × 1 NM grid cell size at 60°N. The ice chart variables are attached to an AIS position message broadcasted from a certain grid cell for the same grid cell and from the same day. The complete set of gridded ice chart data is also made available as a separate datatype. This set can be used to characterize the ice conditions and their variation on daily to interannual time scales, and to generate background charts for the database graphics.

The SAR images used for the daily ice charts can be from 2 to 24 h old when the chart is ready at around noon UTC. Thus the time gap between an AIS report and the utilized SAR data can range from 10 h in advance to a 36-h delay. In the Baltic, as a rule of thumb, the ice drift speed after drift onset is 1–2% from the wind speed and speeds exceeding 50 cm/s may be observed [57]. Thus the ice may drift as much as 2 km per hour although 100–300 m per hour is typical. In areas of thinner mobile ice near the ice margin the inaccuracy can be considerable although the ice chart polygons, over which the chart data is constant, are quite large. For case studies, the match can be improved to some extent by modelling the ice drift.

### 3.3.3. Ice model and weather data

The FMI operative ice forecast model HELMI is a multi-category ice model that resolves ice drift, five categories of level ice, the category of rafted ice, and the category of ridged ice. The variables are listed in Table 2 and the model grid was specified in the preceding section. The level ice categories enable the description of successive freeze-up events. Ice dynamics is governed by a momentum conservation equation with a viscous-plastic rheology for the ice-ice interaction effects [58]. The ice mass balance is governed by continuity equations for concentration and thickness. The HELMI data incorporated in the database is obtained from reanalyses run after each ice season. The re-analyses are forced by the FMI weather forecast model (High Resolution Local Area Model, HIRLAM) re-analyses, which are in essence the best dynamic interpolations of meteorological observations. The HELMI re-analysis output also includes wind speed, air temperature, and sea surface temperature data of the HIRLAM forced fields. The HELMI model also calculates the snow thickness for each of the five level ice categories from the

### Table 1

Temporal coverage and the number of AIS position reports (millions of reports).

<table>
<thead>
<tr>
<th>Year</th>
<th>Data coverage %</th>
<th>Reports, navigation year</th>
<th>Reports, ice season</th>
<th>Max ice extent 1000 km²</th>
<th>Ice winter characterization</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007–2008</td>
<td>62</td>
<td>477</td>
<td>211</td>
<td>49</td>
<td>Extremely mild</td>
</tr>
<tr>
<td>2008–2009</td>
<td>81</td>
<td>663</td>
<td>274</td>
<td>110</td>
<td>Mild</td>
</tr>
<tr>
<td>2009–2010</td>
<td>81</td>
<td>637</td>
<td>262</td>
<td>244</td>
<td>Average - severe</td>
</tr>
<tr>
<td>2010–2011</td>
<td>79</td>
<td>722</td>
<td>385</td>
<td>309</td>
<td>Severe</td>
</tr>
<tr>
<td>2011–2012</td>
<td>75</td>
<td>706</td>
<td>262</td>
<td>179</td>
<td>Average</td>
</tr>
<tr>
<td>2012–2013</td>
<td>76</td>
<td>718</td>
<td>376</td>
<td>177</td>
<td>Average</td>
</tr>
<tr>
<td>2013–2014</td>
<td>78</td>
<td>797</td>
<td>350</td>
<td>100</td>
<td>Mild</td>
</tr>
<tr>
<td>2014–2015</td>
<td>66</td>
<td>629</td>
<td>267</td>
<td>51</td>
<td>Extremely mild</td>
</tr>
<tr>
<td>2015–2016</td>
<td>45</td>
<td>413</td>
<td>147</td>
<td>110</td>
<td>Mild</td>
</tr>
<tr>
<td>Overall</td>
<td>73</td>
<td>5,762</td>
<td>2,534</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

The ice chart and ice model data.

<table>
<thead>
<tr>
<th>Ice chart concentrations</th>
<th>HELMI level ice categories</th>
<th>HELMI dynamical variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pack ice concentration 0–99%</td>
<td>Category 1 thickness, concentration and snow thickness on ice</td>
<td>Ice velocity u-component</td>
</tr>
<tr>
<td>Fast ice (concentration 100%)</td>
<td>Category 2 thickness, concentration and snow thickness on ice</td>
<td>Ice velocity v-component</td>
</tr>
<tr>
<td>Minimum thickness cm</td>
<td>Category 3 thickness, concentration and snow thickness on ice</td>
<td>Ice pressure parameter x</td>
</tr>
<tr>
<td>Maximum thickness cm</td>
<td>Category 4 thickness, concentration and snow thickness on ice</td>
<td>Ice pressure parameter y</td>
</tr>
<tr>
<td>Average thickness cm</td>
<td>Category 5 thickness, concentration and snow thickness on ice</td>
<td>HELMI forcing data</td>
</tr>
<tr>
<td>Ice chart degree of ridging</td>
<td>HELMI deformed ice categories</td>
<td>Wind speed u-component</td>
</tr>
<tr>
<td>0 Level ice</td>
<td>Rafted ice thickness</td>
<td>Wind speed v-component</td>
</tr>
<tr>
<td>1 Rafted ice</td>
<td>Rafted ice concentration</td>
<td>Air temperature</td>
</tr>
<tr>
<td>2 Slightly ridged ice</td>
<td>Rridged ice thickness</td>
<td>Sea surface temperature</td>
</tr>
<tr>
<td>3 Rigid ice</td>
<td>Rridged ice concentration</td>
<td>HELMI derived variables</td>
</tr>
<tr>
<td>4 Heavily ridged ice</td>
<td></td>
<td>Total concentration</td>
</tr>
<tr>
<td>5 Brash barrier (windrow)</td>
<td></td>
<td>Average thickness</td>
</tr>
</tbody>
</table>
HIRLAM precipitation. The re-analyzed HELMI fields are stored each hour and the model variables for the same grid cell and for the nearest hour are attached to each AIS position message broadcasted from a certain grid cell. The spatial accuracy of this match is of the order of one grid cell or better and the uncertainties come almost exclusively from the inaccuracy of the model itself.

For a synoptic characterization of the ice conditions the complete HELMI fields are included every 6 h for the following variables: total ice thickness and concentration, rafted ice thickness and concentration, ridged ice thickness and concentration. The complete set of hourly fields can be accessed as external data for specific purposes. The same applies to the ensemble results, for which HELMI is run several times with somewhat different sets of parameters and initial assumptions. Instead of deterministic values these provide ice conditions in terms of probability of occurrence.

### 3.3.4. Ship data

The ship particulars data is obtained from a separate particulars database with about 96,000 vessels. The type and ice class numerals are included as well. The basic identification and particulars data are obtained from the IHS Fairplay databases completed with data from various sources. The main data entries for all ships are shown in Table 3, and the ship types in Table 4. The ice classes include Finnish-Swedish classes (numerals 1–5) and two general characterizations (numerals 6 and 7).

As the number of different ships in the daily AIS data is not very large, usually no more than few hundred, the particulars data is not arranged in matrices matching the AIS matrices but is linked by the MMSI number. Unlike the International Maritime Organization (IMO) numbers, the MMSI may change, which usually occurs when the owner changes and can be cross-checked with the static AIS data reporting both numbers. Also the ice classes are sometimes incorrect in the IHS Fairplay data. The valid ice class can be found from the Finnish port traffic information system Portnet.

### 3.4. Database functionalities

#### 3.4.1. Selection and compilation of data

Data analyses usually commence by selecting a subset of the integrated AIS and environmental data in terms of time, geographical area, ship data, and environmental parameters. This can be done by directly specifying filter conditions for the data matrices that are easily retrieved from the database. A geographical area can be defined by a bounding box or as any subset of the $1 \times 1$ NM HELMI grid. The analysis can then proceed to the specific research topic at issue.

For comprehensive analyses for a whole navigation year or several years data selection is done in terms of binary logical $v$ vectors defined for each navigation year. The length of the vector is the same as the number of AIS position reports for the whole navigation year. For a given condition the values 1 and 0 identify the subset for which the condition holds or does not hold. As binary data, the logical vectors are not larger than 0.8 GB and are well managed on regular laptops. The commonly applied ones are stored, and new conditions can be derived with operators AND, OR and NOT. A compiling function uses $f$ to create a single data matrix optionally restricted by further conditions and column selection for the navigation year. After the logical $f$ vector is set the compiling is fast. In winter navigation analysis, common conditions include: restriction to one of the main sea areas, independency of navigation, selecting between open water navigation and ice navigation, selecting between fast ice and pack ice zones, selecting the degree of ice ridging numeral, selecting ship ice classes, identification of icebreakers, and restricting the results to certain ship types.

The rationale of this approach is that the data needed in each specific application context need not be saved for future use but only the logical vectors for compiling it. The data is compiled when it is needed for the processing step at hand. This prevents the clogging of the database by the derived datatypes. For example if the aim is to determine Aframax tanker ice performance in the Gulf of Finland let $f$ select this subset, $f_1$ be independent navigation, and $f_2$ be ice class IA Super. Now if the topic is the relationship $v = v(h,c)$ on how speed $v$ depends on the ice thickness $h$ and concentration $c$, then the compiling of $(v, h, c)$ for each of the combinations $f \& f_1 \& f_2$ and $f \& f_1 \& \sim f_2$ and $f \& \sim f_1 \& f_2$ and $f \& \sim f_1 \& \sim f_2$ generates a four-field of relationships $v = v(h,c)$ in terms of independency and ice class.

#### 3.4.2. Derived datatypes

Derived datatypes are results calculated from the basic data for which the processing is time consuming or for which the saving is meaningful for other reasons. Examples of the former are the proximity and adjacency considered in next two sections. An example of the latter is the ship matrix datatype which is a spatial statistical datatype based on the HELMI grid. It essentially adds to the HELMI grid $(i,j)$ a third discrete dimension $k$ which is a ship identifier and refers to the row of the ship database matrix that presently has 96,000 rows or ships. This is a derivative of the daily AIS position report files and has the value $(i,j,k) = 1$ if ship $k$ has visited the grid cell $(i,j)$ during the day. Visitation statistics for periods of one day or longer are obtained for any subregion definable in terms of the HELMI grid and for any ship class definable in the terms of ship particulars in Table 2.

#### 3.4.3. Independent navigation and proximity

To take into account the distance between ships, first a simple descriptor of independent navigation is generated and then a binary relationship of adjacency is generated (described in the next section). These are derived datatypes calculated from the AIS position reports.

By independent navigation it is understood that the operation of a
ship is not affected by the presence of other ships. This is defined principally to quantify the effect of ice conditions on ship performance. This requires that changes of power setting are identified, or that their role is minimized. Especially the effect of other ships should be filtered out. This may be due to chance encounters, but in ice navigation it is often due to convoys and icebreaker assistance. Non-independency is identified by a filter for ship proximity, and a variable is attached to each AIS position report with value $n$ if the ship is non-independent and proximate to $n$ other ships and value $0$ if the navigation is independent. To make processing fast a filter is defined in terms of covering windows instead of the distance and time difference. Value $1$ is assigned to a position report from ship $A$ if there exists a report from exactly one ship $B$ such that both reports fit into the same spatial and time windows. If there are $n$ different ships $B$ for which this holds, the value is $n$. The standard setting is a $3 \times 3$ NM spatial window and a 10-min time window. The time window makes the detection of proximity possible also for the reduced data rate. It also has some extent takes into account that the relative ship speed affects the anticipation of the encounter. If ship $A$ follows a northbound course and $B$ approaches on a southbound course with a relative speed of 24 knots or 4 NM per 10 min the proximity changes to 1 when the mutual distance is 5 NM.

In practical calculations both the space and time are discretized with the resolution ($\Delta x$, $\Delta y$, $\Delta t$). The positions and times of the AIS reports are mapped to the closest nodes of the discrete spatial grid and discrete time axis. The standard values are $\Delta x = 1/60$ degrees of latitude, $\Delta y = 1/30$ degrees of longitude, and $\Delta t = 5$ min. At 60°N the spatial resolution is 1 NM in both degrees. The windows of the discrete filter are $3 \times 3$ spatial block and a $1 \times 2$ time block. Two messages from two ships are always proximate when the distance between their positions is less than 2 NM and the time difference is less than 5 min. The proximity is always 0 if the distance is more than 3/2 NM and the time difference more than 10 min. Between these all-inclusive and all-exclusive cases the proximity is fuzzy. The fuzzy margin can be reduced by finer discretization, and the window sizes can be increased, but as a general characterization of independent navigation the standard procedure is considered sufficient.

For each AIS position report from ship $A$, the unary proximity filter finds the number of proximate ships but does not specify these. This makes the processing fast. On the other hand, for specific analyses of situations involving proximity, such as encounters and convoys, a binary filter specifying all pairs of proximate ships $A$ and $B$ is needed. This is addressed in the next section.

3.4.4. Adjacency and connected groups of ships

Adjacency is a binary relation for a pair of identified proximate ships. The presentation is in terms of square symmetric matrices and the applied terminology conforms to the use of such matrices to describe connectedness in various contexts. The matrices pertain to the daily sets $S(m)$, $m = 1, \ldots, M$ of the AIS position reports and have the dimensions $M \times M$.

An adjacency matrix $A$ is first generated with proximity filtering and then used to select subsets for further analysis. Thus $A(m,n) = 1$ if the position reports $S(m)$ and $S(n)$ are from different ships and are proximate according to the adopted criterion, otherwise $A(m,n) = 0$. The calculation applies a similar discretized approach as described in the preceding section but it is slower because all possible pairs of ships are cross-checked. As most values are zero, $A$ has a manageable size as a daily sparse matrix. Now if $f$ is a metric for $S$ it defines another matrix $F(m,n)$ with the value $f(S(m),S(n))$ when $A(m,n) = 1$ and otherwise $F(m,n) = 0$. Such matrices include mutual distances $D$, mutual time differences $T$, mutual speed differences $V$, and mutual differences of HDG and COG.

If a selected ship identified by $j$ has $M_j$ daily messages, a $M_j \times M$ submatrix $A_j$ of $A$ comprises the adjacency data between the selected ship and the other ships. The distance, speed, HDG and COG differences calculated from $A_j$ can be transformed into a coordinate system where the selected ship has speed zero and sits in the origin with HDG upwards. This constitutes ship domain data which is sufficient for the analysis of binary encounters for the selected ship [59]. However, convoys and other situations with several pairs of adjacent ships are common in winter navigation. In these cases, the ship domain approach is not sufficient for their analysis, and the characterization of ships belonging to a group and of the manner the group moves as one unit is needed. This is done in terms of connected components.

The adjacency matrix $A$ has an equivalent representation as a graph $G$ where reports $S(m)$ correspond to the vertices (nodes) of $G$ and adjacency $A(m,n) = 1$ is assigned to an edge linking the vertices. The basic concept is the connected component $G_k$, which is a maximal subset of $G$ in which all vertices can be connected with a chain of edges. It is required that $G_k$ contains at least two vertices. Each $S(m)$ is then either from an independently navigating ship or belongs to precisely one $G_k$. The connected components correspond to situations, or sequences of situations, during which the ships are grouped by their adjacency to each other, for example in encounters and convoys. They can be complex, for example whenever two convoys meet they belong the one and same $G_k$. It is also to be noted that $G_k$ may connect ships that are not proximate, for example the first and last ship of a long convoy.

The final step is taken from the connectedness of the AIS reports, which usually have a different time tag, to the connectedness parameterized by common continuous time. In the notation this is indicated by an added time variable. After the final step, the connected components $G_k(t)$ are time dependent groups of ships where the list of ships may change with time. Without going into the detail it is outlined how this is attained. Each $G_k$ can also be expressed as an adjacency matrix $A_k$ for a list $S_k$ of position reports from at least two ships. The ordered time column of $S_k$ is adopted as a common time axis to which all navigational parameters in $S_k$ are interpolated. For each ship this is done only for gapless data sections as a gap indicates that the ship has for a while lost the connection with $G_k$ and then returned back. The interpolated data defines a denser adjacency matrix which is then reduced to $A(t)$ by requiring that all time differences between reports from different ships are zero. From $A(t)$ and from the interpolated and reduced report list $S(t)$, mutual distances and speed differences are obtained, as well as differences of HDG and COG as functions of common time. From $A(t)$ the time parameterized ship groups are identified as connected components $G_k(t)$. This information is included in $A(t)$ by assigning an identification number instead of 1 to each group and adding the group number column to the report list $S(t)$. One more column connects each interpolated report in $S(t)$ with an actual report in $S$ from the same ship and with a minimal time difference. This re-establishes the connection with the environmental data linked to $S$.

In practice, the ship group data $A(t)$ and $S(t)$ includes all the information that is needed for the analysis of encounters, convoys, and icebreaker operations. The data gaps must be noted, however. Also, if the data pertains to the daily lists of position reports, the possible continuation of operations over midnight needs consideration.

4. Application examples

4.1. Spatial traffic statistics

It has been asserted that the time range of AIS datasets can be limited in traffic analyses as the traffic patterns repeat themselves [16]. However, this limits the possibility to detect annual and interannual variations, cycles and trends, and the traffic patterns during the ice season are in many respects unique for each season. Also, the spatial smoothness and coverage provided by large AIS data sets has its methodological benefits. It can be seen from Fig. 3b that the nine-year traffic data covers all navigable waters of the Northern Baltic at a resolution of 1 NM. The figure can be taken as an approximation of a continuously varying field. In analogy with continuum mechanics, this field can be described either in terms trajectories and events for
individual ships (the particles of the Lagrangian picture) or in terms of the traffic flux and events at each location (the Eulerian picture). In particular, the grids generated for the whole time-range can be taken to represent a mean field and grids for shorter time periods can be presented as a deviation from the mean field.

The grid values may relate, among many other possibilities, to ship density, navigational status, and navigational events. As an example the spatial distribution is shown for the degree of non-independency, which is the ratio of non-independent navigation to all navigation (Fig. 4). This is calculated, applying the criterion in Section 3.4.3 over the whole database time-range both for ice-free conditions and ice navigation. The degree of non-independency is a dimensionless variable for ship crowding. This combines the effects of traffic intensity and proximities required by navigational operations and is relevant for the estimation of navigational risks. For ice-free conditions this increases mainly with the traffic intensity only. The traffic generally takes the fastest route to the destination and seeks to maintain a safe distance. Certain straits, reporting areas, and major width-limited fairways are identified as hot-spots. On the other hand, for ice navigation, additional features due to the winter navigation system emerge. In most sea areas, the degree of non-independency arises due to convoy operations and ice channel navigation. The traffic spreads as the fastest route to the destination varies with the ice conditions. New hotspots include icebreaker meeting points, and extensive roadstead stays in the Eastern Gulf of Finland where ships queue for port access.

Concerning the interpretation of the grids in Fig. 4, as a mean field this is likely to hold well for ice-free conditions. Further years of data may change the values, but the spatial patterns will remain similar. The trends and variances of non-independency can then be studied as differences from the mean field. On the other hand, for ice navigation this is uncertain. The ice seasons include five mild winters (mainly ice-free in the Sea of Bothnia), two average winters (Sea of Bothnia ice covered), and two from average to severe winters (Northern Baltic proper ice covered). Also between ice seasons with similar maximum ice extent a wide range of ice conditions is found. Especially the amount and location of ridged ice depends on the amount of wind and it is difficult to identify a pair of ice seasons with comparable navigational conditions. In Fig. 4 the signatures of individual ice seasons are visible and the grid can hardly be interpreted as a mean field for all sea areas. The assumption is more warranted in the Gulf of Finland where most traffic follows a traffic separation scheme whatever the ice conditions. The example illustrates the challenge of spatially modelling the winter navigation system and the need for comprehensive datasets.

4.2. Ice navigation speed

One basic aim of winter navigation research is to determine the effect of ice conditions on ship speed. This is instrumental for assessing the efficiency of the winter navigation system. The task may be to describe ice resistance that reduces the speed from the open water value if the propulsion power is kept constant, or increases the fuel consumption if the speed is maintained. On the other hand, the task may be to describe the actual realized operational speed in which case there are also indirect effects of ice conditions. These include waiting times due to port logistics and icebreaker assistance, the efficiency of assistance, and conforming to convoy speed limited by the less capable ships.

Here speed reduction due to ice resistance only is considered but the concepts apply as well to operation speed. For simplicity it is assumed that only the ice thickness is relevant. A ship’s transit is divided into three modes, continuous icebreaking, ramming and being beset. In ramming, difficult ice features such as ridges are penetrated by repeated attempts. The ship may become beset in a ramming attempt but also due to dynamic compression in the ice. When just continuous icebreaking is considered, the relationship \( v = v(h) \) between speed and thickness is often called the \( h-v \) curve. In a data driven approach this is determined from the ice thickness and the AIS-retrieved speed data for a period \( T \), the sea area \( R \), and a fleet \( F \) of ships. Depending on how \( T \), \( R \), and \( F \) are involved a range of definitions is obtained. There are two basic types. Definitions of the first class use the basic data of the database, and those of the second type use derived gridded data types.

The definitions of first type select a suitable set of linked speed and thickness data from the basic data for \( T \), \( R \) and \( F \). The first option is that thickness is grouped into bins \( h \) and the speed values linked with \( h \) have an average \( v \). This is called the mean average over \( T \) and \( R \). The obtained \( h-v \)-curves does not take into account the different capabilities and different numbers of messages for the ships. It characterizes the general effect of a speed reduction on the traffic flux and transportation efficiency. On the other hand, if \( F \) contains only one ship the \( h-v \)-curve is a ship average possibly related to the ship-specific physical speed reduction. If the \( h-v \)-curve is determined for each ship of a fleet \( F \), this family of curves can be further averaged to give the fleet average. As each ship enters the average with equal weight it better characterizes the typical ice performance than the mean average which is dominated by ships on many visits. Another set of \( h-v \)-curves is obtained by normalizing the speed with the open water speed. Especially normalizing the \( h-v \)-curves for each ship in \( F \) and taking a fleet average provides an \( h-v \)-curve of the average normalized speed reduction for the ships in \( F \).

Fig. 5 shows fleet averaged \( h-v \)-curves determined from the data for all Finnish-Swedish ice class IA Super ships. The ships in this class are able to navigate without assistance also in thick Baltic ice and often find their routes outside the main ice channels that are occupied by weaker ships. Only the independent navigation (Section 3.4.3) of steaming ships is considered. The speed threshold between steaming and idling is...
set to 0.5 knots. The IA Super ships are also less likely to change their power settings during ice transits \[60\]. Despite the independence, the winter navigation system affects thicker ice types as the traffic seeks to follow the main ice channels. On the other hand, by separating a subclass of frequently visiting ships, which are usually more powerful and faster, it can be seen that for normalized speeds the result does not change. This indicates that the normalized fleet average is a meaningful descriptor of the effect of the ice conditions on speed.

The above definitions do not take into account spatio-temporal variations. These are more consistently included in the definitions of the second type that advance from a time stack of gridded thickness data. To each thickness grid \( h_{ij} \) a grid of average speeds \( v_{ij} \) is associated. The binning and averaging proceeds then spatially for the \( ij \)-grids and temporally through the stack of grids. Two cases deserve specific definition. In the first, the ice thickness and ship speed are averaged temporally and independently of each other. The \( h-v \)-curve is constructed from the two resulting grids by binning and averaging. When this is done over several ice seasons this is called the climatological average as it compares long-term averages of \( h \) and \( v \) as independent quantities. This is the usual approach in ice navigation research when linked data is not available. In the other case, the \( h-v \)-curve is first calculated as a temporal average for each grid cell and the final \( h-v \)-curve is the spatial average over this family. This grid average has the advantage that less visited grid cells that are more likely to be navigated outside channels have the same weight as cells from intensively navigated main ice channels.

Fig. 6 shows that also the most general approach with climatological averages captures the effect of ice on ship speed. The data comprises all AIS position reports from independently navigating IA Super ships steaming through ice-covered water. The speed normalized by the open water speed was calculated for each 8 × 8 NM cell as a fleet average. By the force of the law of large numbers, the speed map reproduces the essential features of the spatial patterns of the 8 × 8 NM ice thickness map. Although the scatterplot between speed and thickness values spreads out, the \( h-v \)-curve generated by binning and averaging has a consistent linear slope close to the values obtained in more detailed analyses \[60\].

4.3. Icebreaker assistance

All convoy and assistance operations can be identified with the methods described in Section 3.4.4 and can have remarkable complexity. For example, convoys led by icebreakers may meet, divide, merge or exchange ships. A simple but representative case where an icebreaker assists a single ship outside existing channels is show in Fig. 7. The data is from March 2, 2016, and from the northwestern quadrant of the Bay of Bothnia close to the Swedish coast. The ships head southwards against 12 m/s headwinds and the icebreaker breaks a channel in about 40 cm thick and somewhat ridged ice. The ridges induce speed variations, the amplitude of which is about 2 knots. The assisted tanker follows with less ice induced speed variation and seeks to maintain a mutual distance of about 4–6 cable lengths.

Three periods with difficulties are seen. Before 10 UTC the icebreaker speed slows down to 5 knots. It gains some of the speed back after which the speed is reduced close to zero. Although the icebreaker loses its momentum thereafter it gains its operation speed rapidly. The tanker appears to have no difficulty to adjust its speed and keep a safe distance from the icebreaker. It also accelerates well after the speed drop. Another similar event follows during the last quarter of the hour, after which the distance is reduced. The reason is probably the anticipation of difficulties from the ice conditions. This does not help as the tanker becomes beset at 11:10 and cannot continue. The icebreaker halts, makes a full turn and cuts the tanker free by passing by with a minimum distance of 0.4 cable lengths. Some AIS messages are lost when the icebreaker heading is less favorable for the transmission. After being freed the tanker gains speed less easily. There are no indications of ice induced difficulties in the icebreaker’s transit. Thus the likely reason is that the ice cover has been in a state of static wind induced compression which has changed to dynamically closing the channel and besetting the tanker.

This simple case illustrates the usefulness of the full update rate for AIS and proximity data in detailed investigations of winter navigation,
especially in the analyses of risks in convoy operations. In most cases there are several assisted ships, often more than ten, and complicated interactions arise when the ships react to the speed changes of the other ships.

5. Discussion

5.1. AIS as research resource

The original intention of the AIS was collision avoidance as well as providing situation awareness for VTS and for ships themselves. Several elements were considered important, among other things, automatic data exchange and ship identity information. However, the research applications were not foreseen. From research point of view, the AIS combines detail and comprehension with some fundamental shortcomings. These are due to both the characteristics of the AIS and the infrastructure for the transmission and managing of the AIS messages. There are problems with the data quality, archiving, intermittency of satellite AIS data, and the lack of ship propulsion data. These are felt especially in winter navigation research where the propulsion data plays a key role and ship dynamics vary over short time and length scales. Also, the winter navigation system needs be studied over long periods to include the interannual variation of ice season severity.

AIS data can be received by anyone and the opportunity has been seized by commercial providers. However, the AIS stream is vulnerable to manipulation and hacking [61] and may contain considerably erroneous data [62]. It should be cross checked against other information apart from just the checksum. For terrestrial data, the check is best done by traffic administrations who must provide accurate data for the VTS. On the other hand, it is usually not a specified task for authorities to archive the data. When they do so, the handing out of the data involves legal or other uneasy issues [63]. It can be used to study the ship performance and routes in detail, survey fisheries, and with satellite images used to detect dark ships, especially military vessels. In winter navigation the main issue is the reluctance of shipowners against the ice performance assessments of their ships. The authorities have somewhat facilitated the access as the data is known to be available anyway. The services opened in Finland provide real-time AIS. The archiving has been left to researchers, which is not a sustained solution as is seen from the data gaps (Table 1). A better and pan-Baltic solution would be to develop the HELCOM archive as the full data with 50 billion annual messages should not pose a problem. Possibilities for risk mitigation and cost savings from winter navigation system analyses are also evident. Terrestrial AIS is also hoarded at the pan-European level by, the European Maritime Safety Agency (EMSA), but the data is provided only for specially agreed uses, mainly for emission inventories [64]. Whatever the administrative solution, certain restrictions are expected to stay in effect.

In the Baltic full coverage can be attained by terrestrial stations, but in the world ocean only satellite reception is possible. The intermittency of satellite AIS is not a problem for most research because congested areas for which traffic details are important are covered by terrestrial data. For Polar ice navigation, especially along the Northern Sea Route and Northwest Passage, this is a shortcoming. Microsatellites for frequently updated SAR images are expected to improve Arctic ice routing. This would improve even more with added AIS reception or with specific AIS nanosatellites [65].

Research applying AIS puts a lot of effort into compensating for the lack of propulsion data. Ship emission modelling, one of the main uses of AIS, calculates the power from the speed, hull geometry and resistance models. The models are quite reliable for open water steaming but for ice transit the situation is different. The research on ship ice performance seeks to develop models connecting the triad of ice conditions, ship speed and propulsion power. Usually none of the components are well known. Ice information is usually inaccurate so that speed estimates are unreliable, even though the resistance and propulsion are known. The resistance models are validated only for certain ship types and for idealized ice conditions. With propulsion power AIS data could improve this. On the other hand, with good models the estimation of ice conditions from ship speed and power is a feasible undertaking. Access to propulsion power data would be a holy grail and allow the cross-validation of all components. If the ice conditions are well known, state of the art h-v-curves could be used for transit speed forecasting and ice routing. The results and the h-v-curves could then be validated and improved with AIS and propulsion data. On the other hand, if h-v-curves are well established it would be possible use them to derive ice conditions from speed and propulsion data and improve ice information. These undertakings constitute a feedback loop. However, as an IMO SOLAS agreement, AIS is not easily amended to cater for this need. Shipowners may release propulsion power data when needed in tailored ice routing services, and it is already obtained from a large number of ships in real time for remote control or weather routing purposes.

5.2. Further development and utilization of the database

The database is an institute resource used for research, commissioned work, and information production. The research is progressing towards the characterization of the Baltic winter navigation system. During the midwinter the systemic nature of navigation becomes pronounced as the traffic gravitates to the main ice channels. In difficult conditions, powerful ships are less willing to overtake slower ships but conform to a system speed that settles around 8 knots [67]. The adaptation of road traffic models [66] suggests itself. The optimization problem concerns the whole system and not a collection of individual voyages, and the system decomposition should be in terms of vessel groups. A system analysis would be very useful from a societal perspective as it would make headway towards more reliable planning of icebreaker operations and the consequent cost reductions. Also conventional routing and ice performance analysis need to take the background system into account [67]. System-oriented projections of the future should include long-term scenarios of Baltic wintertime traffic, where climatic change should be taken into account, and short-term forecasting of the traffic situation.

Forecasting requires that the database is incorporated in the operative production system. Readiness for this exists but the reliability of the AIS data stream should be improved. A framework extending the present situation is sketched in Fig. 8. The integrates not only the...
marine environmental data but also its production with AIS. The historical part of the database shown in Fig. 2. A new component is a real-time database for core data only. This integrates the marine environmental and satellite data as soon as they come available with the AIS data and is used to generate products that amend and complement the usual environmental information. For ice navigation, the least challenging are displays of the traffic situation and ice performance together with nowcasted or forecasted fields of ice conditions. The next level would be to extrapolate observed ice resistance to the whole sea area, which would enable ice routing. Such a service could be ship-specific or be expressed as speed maps for ship classes. The forecasting could be extended to the traffic situation [68]. Learning algorithms and other big data methods can be applied to take into account the numerous non-physical factors. On a still higher level the forecasting and optimization could be made for the winter navigation system. This would finally answer the expressed wish of the winter navigation community.

An essential improvement made possible by the proposed framework is the integration of the AIS already with the production of environmental information. Restricted to ice conditions, this means the systematic utilization of the fact that ship responses mirror the difficulty of the ice conditions. This can be used to improve and validate ice information products that suffer from a chronic lack of observation data. A particular example is that of SAR images used in ice charting. Considerable effort has been put into extracting the ice thickness and ice ridging parameters from SAR but without conclusive results. However, in the Baltic it is possible to classify the images in terms of navigational difficulty by simply observing how the ship speed responds to the various SAR textures [69]. On the lowest level this can be done without any reference to ice conditions, however combining it with existing SAR methods for ice parameter retrieval further increases the promise of the approach. For Polar waters, combining the production of ice information with the AIS would be even more appropriate than in the Baltic.

5.3. Societal and policy aspects

AIS data has shown its usefulness in maritime research. Although big data and data integration are common buzzwords, for AIS there exists yet little that is both big, integrated and applying advanced mathematical methods. If the data is truly big, the methods are usually traditional, and advanced methods common in other big data contexts are applied only to small demonstration datasets. The comment on traditionality applies also to the database described here as the integration relies on temporal and georeferencing or exactly definable data links. This on the other hand creates a solid foundation on which other approaches using less structured data and advanced methods may be built. The described database suffers somewhat from data gaps following from the provisional transfer solution for the AIS data. A more sustained integration of the AIS data with marine environmental data and the distributing and archiving of the products would require agreements between the relevant actors. In Finland the actors would be the FTA, the Navy and the Frontier Guard, who have an established mechanism to share AIS data [62], and FMI. If the process is initiated, issues of data ownership and national security must be resolved. For the moment the integration and archiving relies on research groups.

On the societal and policy level, the most important application of AIS for the moment is ship emission inventorying done by research groups. If in the future emission-based fees and tighter emission controls emerge, this outsourcing is a bit of a problem as there may be a deficit of trust. It would then be better that the inventories are made by a maritime authority to whom the governance of AIS data is entrusted, for example the EMSA in the EU. The same applies to ocean noise and water discharge inventories. AIS data is relevant to the Energy Efficiency Design Index (EEDI) process that imposes ever tightening conditions for energy efficiency. In the ice navigation context, this is linked to the fact that EEDI ships tend to have less power and thinner shell plating and thereby inferior ice-going capabilities and higher damage risk [70]. This leads to the increased risk of oil spills in winter conditions. In addition, oil spill preparedness and response during the ice season is in need of more integration. The forecasting of oil spreading, providing ice information, controlling ship traffic, and managing the response fleet are currently divided between different authorities. There is presently little means to predict how well the response fleet is able to operate in difficult ice conditions. Systems integrating ice information, spreading modelling, AIS information and ship performance modelling could answer this need.

On European Union level maritime products generated from AIS data are made available by the European Marine Observation and Data Network (EMODnet) and, on the other hand by EMSA trough Copernicus Maritime Surveillance. Administrations also often realize the applicability of AIS to research topics that are under their mandate. However, national and transnational strategies to facilitate the use of AIS data as such for research purposes do not exist or are insufficient. Aspects of winter navigation are mostly untouched also in the general context not restricted to research. In Finland, governmental strategic white papers prepared during the last five years address open data, spatial data, and big data. The AIS data and also winter navigation are missing from these, and the focus is more on new business opportunities than on societal relevance. In Sweden the official statistics authority has prepared a white paper on the use of AIS data in maritime statistics [71], and HELCOM has recently updated its Maritime Assessment [52] that contains a detailed description of the HELCOM AIS database and its developments. The specific aspects of winter navigation are not addressed in these reports, although in particular the HELCOM mission is pollution prevention and maritime safety. HELCOM also continues to resample the data, incapacitating its use for several winter navigation research topics, although Moore’s law has made this historical choice obsolete.

6. Conclusion

The usefulness of AIS data integrated with ice information has been recognized in the winter navigation research. A comprehensive and accumulating database integrating nine years of Baltic full update rate AIS data with ice charts, ice forecasts and ship particulars data was presented. The integration links a string of variables describing the ice conditions to each AIS position report at the location and, on the other hand, a string of ship particulars. The database provides a good time resolution for the analysis of complicated ship configurations and a multi-annual span for resolving the climatological variation of ice conditions. It can be used to address a wide spectrum of research questions related to winter navigation. Especially relevant are...
functionalities that enable detection and analysis of convoys and other groups of proximate ships. These provide tools for analyzing and optimizing the winter navigation system in terms of convoys and assistance operations. Three application examples were presented: on the effect of ice on ship staticalistics, on the relationships of speed reduction and ice thickness, and on convoy analysis. Average speed reduction curves for IA Super ice class ships were presented. It was made evident that a full update rate is required for proper analysis of rapid speed changes that are typical for ice transits. The AIS data comes from Finnish territorial stations and their range is sufficient to cover the Gulf of Finland and the Bay of Bothnia completely. In the western half of the Sea of Bothnia the AIS reception is intermittent and the same applies to the northern Baltic proper which limits certain applications of the data in these sea areas. The aim is to improve the coverage of full update rate AIS data in cooperation with Baltic traffic administrations and HELCOM.

Conflicts of interest

There are no undisclosed relationships or undisclosed funding sources that may pose a conflict of interest.

Acknowledgements

This work resulted from the BONUS STORMWINDS project supported by BONUS (Art 185), funded jointly by the EU and the Academy of Finland (decision number 291683).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpol.2019.02.038.

References