Title *Efficient, Safe and Sustainable Traffic at Sea*
Acronym *EfficienSea*

Contract No. 013

Document No. D_WP6_2_03
Document Access: Public

Analysis of the near-collisions using AIS data for the selected locations in the Baltic Sea.

Deliverable No. D_ WP6_2_03

Date: 24.01.2012

Part-financed by the European Union

DOCUMENT STATUS

Authors

Name	*Organisation*
Floris Goerlandt	Aalto University, Dept of Applied Mechanics
Jakub Montewka	Aalto University, Dept of Applied Mechanics
Erik Sonne Ravn	The Danish Maritime Safety Administration
Maria Hänninen	Aalto University, Dept of Applied Mechanics
Arsham Mazaheri	Aalto University, Dept of Applied Mechanics

Reviewing/Approval of report

Name	*Organisation*	*Signature*	*Date*
Tommi Arola	Finnish Transport Agency		
Marina Ladan	Aalto University		

Document History

Revision	*Date*	*Organisation*	*Initials*	*Revised pages*	*Short description of changes*

Table of contents

1. INTRODUCTION

Ship traffic poses risks in terms of possible accidents resulting in damage to vessels, personal injuries and loss of life. Environmental pollution, especially oil pollution, is of special concern for the Baltic Cea because of the vulnerability of the aquatic ecosystem.

Various analysis methods are available for assessing the safety level of maritime traffic, such as accident analysis, analytical and simulation models, Bayesian networks and fault tree analysis. A few examples are introduced below.

Kujala et al. (2009) analyzed accident data in the Gulf of Finland and concluded that groundings and ship-ship collisions are the most frequent accident types in the Gulf of Finland. Kristiansen (2010) combined accident data in a Bayesian network through a semi-automatic learning algorithm, focusing on powered groundings. Goerlandt et al. (2011) developed a simulation model for collision probability, which has been used to evaluate the risk of tanker collisions in the Gulf of Finland. Montewka et al. (2010, 2012) proposed an analytical model to assess the likelihood of ship collisions, using ship hydrodynamics to evaluate the so-called minimum distance to collision. Martins et al. (2010) used the formality of the Formal Safety Assessment method, applying fault tree analysis to assess the probability of collision and grounding, focusing on incorporation of human error in the analysis.

As pointed out by Inoue et al. (2007), near accidents in complex systems are much more common than actual accident. These events can give a statistically more significant evaluation of maritime safety than accident analysis, as they are much more common. Following this idea, this report presents a methodology to evaluate the occurrence of near ship-ship collisions in an open sea area, based on AIS data. The method is rooted in the well-established concept of a ship domain, which is first briefly introduced. An overview of the near collision detection method is then given and applied to the summer traffic in the Baltic Sea with respect to two test bench areas named: the Gulf of Finland and the Danish Straits. Challenges in the methodology are discussed and results are given.

2. NEAR-COLLISION CRITERION

In the analysis presented we consider a near-collision as a situation in which two ships come close to each other to a certain distance. This distance is called near-collision criterion and can be defined in two ways, either by using a ship domain concept or by applying dynamic safety ellipse.

A ship domain is usually defined as the area around the vessel which the navigator would like to keep free of other vessels, for safety reasons, see Goodwin (1975), Fujii (1971). Since the first introduction of the ship domain concept various researchers have attempted to quantify the size of this domain, either by statistical analysis of distances between vessel encounters or by expert elicitation using fuzzy methods. Wang et al. (2009) gives an overview of the different proposed domains, while proposing a unified analytical framework for the domains. Even though the ship domain is a well-established concept, certain problems with the application can be identified (Jingsong et al. 1993).

Domains can be classified by their shape: circular, elliptical and polygonal domains have been proposed. A distinction can also be made between fuzzy domains, where different domain sizes represent various levels of navigational safety, and crisp domains, for which an encounter is either safe or unsafe, without any gradation.

Fuzzy domains such as that proposed by Pietrzykowski et al. (2008) seem preferable in terms of safety analysis of marine traffic, but are at present still under development. Crisp domains use a simple classification of a situation between safe or unsafe, which evidently is a simplification. Moreover, the sizes of the domains proposed in the literature vary quite significantly (Wang et al. 2009).

In this report two criteria are used depending on the sea area. In the case of the Gulf of Finland, considered an open sea area, the smallest ship domain found in the literature is applied. This means the domain proposed by Fujii (1971), defined as an ellipse with the major axis along the ship's length and the minor axis perpendicular to the ship's beam, as illustrated in Fig. 1. The half-length of the major axis is taken as 4L, while the half-length of the minor axis is taken as 1.6L, with L the ship length, see Fig 1.

In the case of Danish Strait which is considered a restricted area a dynamic safety ellipse (DSE) is adopted, defined as an area around a ship with the length of 4L, unequally divided. The ellipse size is as follows: 0.5L behind the ship and 2.5L in front, wherease the width of the ellipse is 5B, where B stands for ship breadth. Moreover, this initial ellipse is subject to scale up or down by a continuous factor, depending on the relative courses and speeds of the ships.

If the two ships sail on the same course with almost the same speed (small relative speed), the length is decreased to a a factor not smaller than 0.4. This implies that the total length of the ellipse is not smaller than 0.4*3L=1.2L and the total width of the ellipse is 0.4*5B=2.0B. If they are heading toward each other the length is increased up to a factor not larger than 1.2. The near-miss candidates are then manually evaluated.

The criteria adopted seem justifiable, since the aim of the method proposed is finding the most critical encounters between ships.

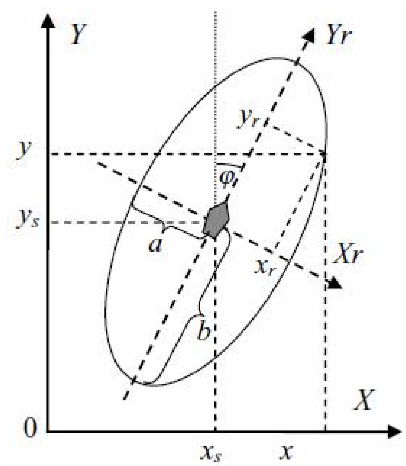

Fig. 1: Ship domain according to Fujii et al. (1971). Figure: Wang et al. (2009). a = 1.6L, b = 4L

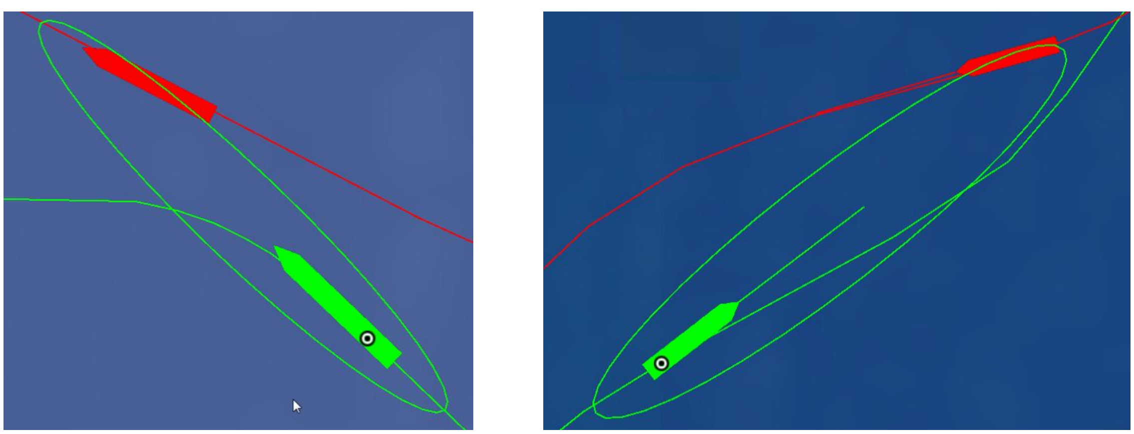

Fig.2 A DSE concept, for small relative velocities on the left (3.5L, 4B) and otherwise on the right (4.5L, 5B).

3. RESEARCH METHOD

3.1 Description of the AIS system

The determination of near ship-ship collisions is based on an extensive analysis of data obtained from the Automatic Information System (AIS), installed onboard merchant ships. The system transmits certain data about the navigational status of the vessel to other vessels and onshore receiving facilities. The International Maritime Organization (IMO) requires following ships to be equipped with AIS:

- All ships of 300 Gross Tonnes (GT) and up-ward, engaged on international voyages.
- Cargo ships of 500 GT and upwards, not en-gaged on international voyages.
- All passenger ships, irrespective of size.

It follows that not all traffic is recorded in AIS databases (pleasure crafts are only sparsely represented based on voluntary carriage), but most merchant traffic is recorded. The certain vessel types have not been taken into account into the analysis, like ice breakers and tugs, as these vessels are meant to operate in close vicinity of merchant vessels. The analyzed data for the Gulf of Finland is for the period 01.04.2007-31.10.2007, whereas in the case of Danish Straits the full year 2007 is analyzed. This implies that only summer traffic is evaluated. This restriction is needed, as the Gulf of Finland freezes during the winter, which has an important impact on the vessel traffic. The data type used in the context of this report is summarized in Table 1.

Table 1. Information in AIS transmission, used in this paper, (IMO 2001)

Data Type	Accuracy	Unit
Maritime Mobile Service Identity *	-	-
Timestamp : UTC time	1	s
Position: longitude & latitude †	1/10000	min
Ship type ‡	-	-
Ship length	1	m

* MMSI: a 9-digit code uniquely identifying a vessel, † WGS-84 reference system, ‡ Cargo ship, tanker, passenger, HSLC or other vessel

It has been recognized that AIS data contains various types of error sources, such as data corruption, faulty position reports and erroneous MMSI numbers, see Graveson (2004), Norris (2007). These have been filtered out before analyzing the data. The transmission rate of AIS data ranges from every 2 s to every 3 min, depending on the ship speed and the rate of turn. This results in an enormous data set. The analyzed data for 2007 unfortunately does not contain all the data points. For a given vessel, the sample rate is about 5 minutes on average. This is sufficient to get a rough estimate of the number of the close encounters, but is insufficient to evaluate the actions of the navigators in these close encounters.

3.2 Data analysis

Raw AIS data consists of millions of data points, containing information given in Table 2. In order to analyze the maritime traffic this data had to be grouped into routes. Routes are defined here as a set of trajectories between a departure and arrival harbor.

The AIS data was first gathered per ship, based on the MMSI number. After sorting this data chronologically, the data per ship was further split up to form individual ship trajectories, using a methodology described by Aarsæther et al. (2009). These trajectories were then further processed and grouped per route.

As the sample rate of these vessel trajectories is about 5 minutes on average, the trajectory data was artificially enhanced to contain data for each second in order to enable a comparison between vessel positions at the exact same point in time. The extrapolation for the vessel position was performed using an algorithm suitable for data in the WGS-84 reference frame, see Vincenty (1975). The ship speed has been linearly interpolated between known values.

3.3. An algorithm detecting near collisions

The near collision detection algorithm is shown in Fig. 2. The basic idea is to scan the database for events where the ship contour of one vessel (i.e. the ship area in terms of ship length and width) enters the ship domain of another vessel. If the domain is violated, the event is labeled as a near collision and relevant details such as time of occurrence, location, encounter type, ship types and ship flags are stored for further analysis. The near collision detection algorithm was coded in Matlab.

The algorithm starts with evaluating whether or not the trajectories of the two considered vessels occur in an overlapping timeframe. If so, the closest distance between vessel positions for contemporary time instances is computed. If this closest distance is smaller than the extreme value of the ship domain, the actual vessel contours in terms of length and width are constructed for the smallest vessel and the ship domain is constructed for the largest vessel, for each second. Concurrent ship domains and vessel contours are evaluated to overlap or not. If there is an overlap, the relevant situational data is stored. If there is no overlap, the next case is investigated.

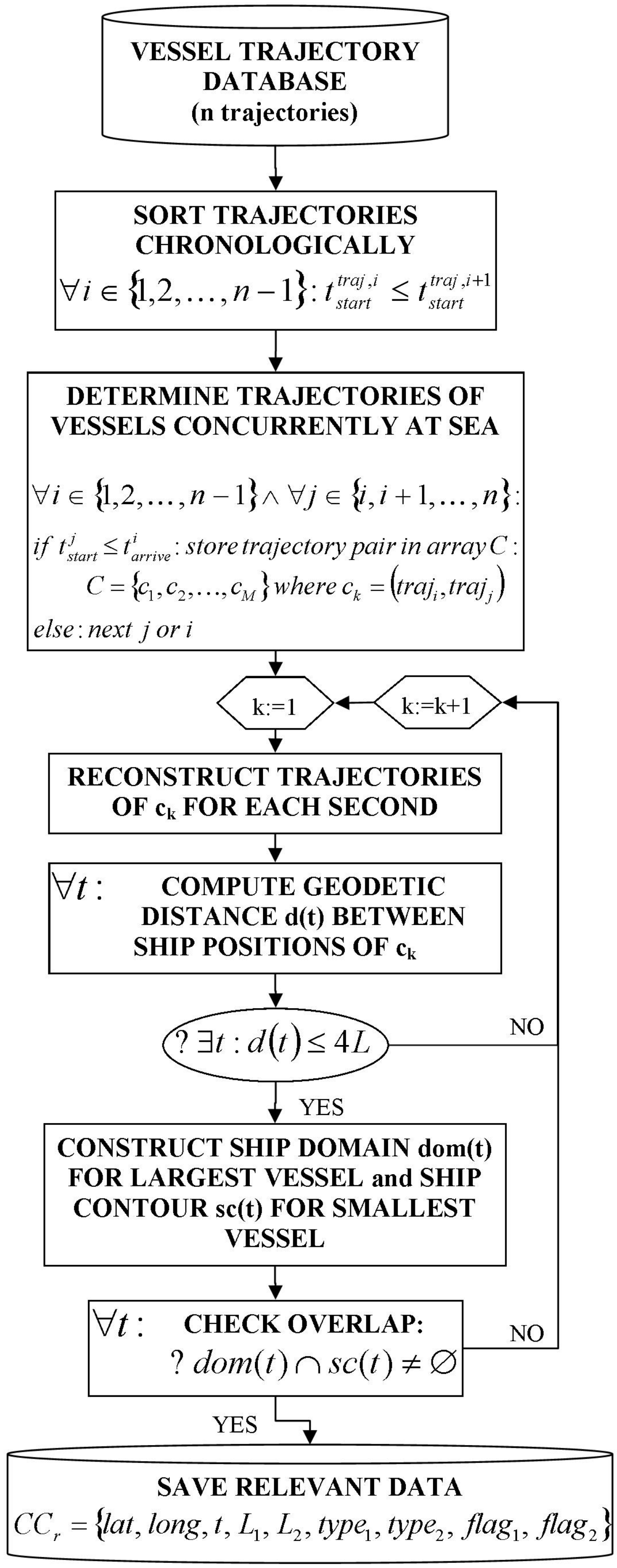

Figure 2. Near collision detection algorithm.

4. RESULTS

The proposed method is applied to two sea areas. Firstly it is run for the open sea area in the Gulf of Finland, which extends eastwards from 21°38" E and northwards from 59°54" N. Secondly the restricted sea areas of the Danish Straits are analyzed, with the bounds depicted in Fig.3, moreover the near-collisions outside the Danish territorial water are discarded.

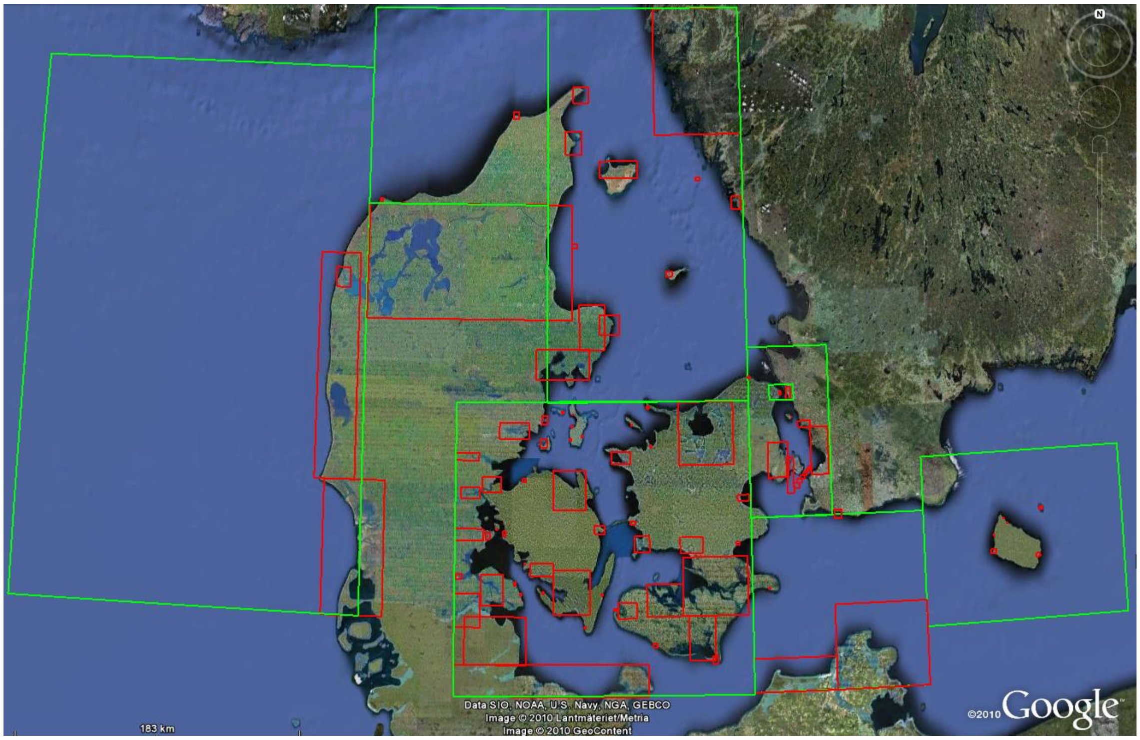

Fig. 3 The Danish Straits - the green rectangles are where the near-miss calculations are carried out. The red rectangles are omitted.

4.1. Results of the analysis for the Gulf of Finland

In the analysis of the locations of the near collisions, a distinction is made between three different encounter situations, as defined in the Collision Regulations (2002). Thus, crossing, head-on and overtaking encounters are distinguished in Fig. 4 to Fig. 6. In these figures, the near collisions are assigned a color code according to the closest distance between the vessels, normalized with respect to the length of the largest vessel.

The elliptical Fujii domain leads to 178 ship domain violations for head-on encounters and 170 for crossing encounters. However, 2253 cases were identified for overtaking encounters. This is due to the fact that the Fujii domain does not take the regulation of traffic in terms of traffic separation schemes into account. In order to get more meaningful results, a heuristic solution for this is proposed, by requiring that the number of domain violations for overtaking is equal to the average number of critical encounters for head-on and crossing. To this effect, the Fujii domain is evaluated with a width of 1.25L for overtaking encounters (as opposed to the original 1.6L), where L is the length of the largest vessel in the en-counter, see Table 2. In appreciating the results, it is evident that different regions in the Gulf of Finland are more likely than others to see near collisions between vessels. Moreover, the encounter type is also strongly location-dependent.

Concerning critical crossing encounters, the crossing area between Helsinki and Tallinn and the junction near Primorsk are seen to be dangerous locations. This corresponds to the marking of these areas on nautical charts for the area. However, other locations may be identified as well: the junctions on the main traffic scheme towards Sköldvik and Kotka also show an increased number of critical encounters. Finally, two areas in the mouth of the Gulf of Finland between Paldiski and Hanko show up. This is where traffic bound for or coming from Sweden crosses the traffic separation scheme.

As for critical head-on encounters, the approaches to Primorsk and St. Petersburg seem risky, but it should be noted that navigation in these areas occurs in a rather narrow waterway, which is marked with a series of buoys. This means that the Fujii-domain may not be an appropriate tool to assess the navigational risk in these locations. Head-on encounters are seen to occur frequently at the start of the traffic separation scheme at the western edge of the studied area and in the mouth of the Gulf of Finland, between Paldiski and Hanko.

Critical overtaking encounters seem to be more frequent in the eastern Gulf of Finland, especially south of Suursaari south of Kotka and in the eastern stretch of the traffic separation scheme.

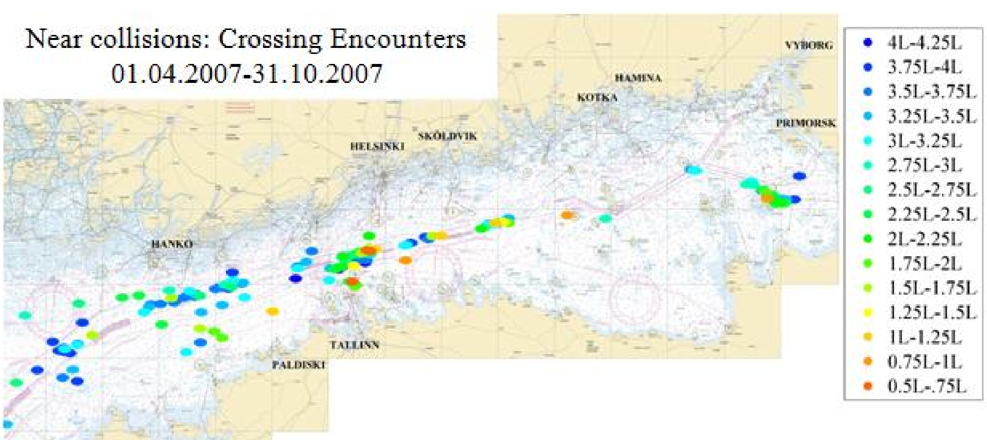

Figure 4. Near collision locations: crossing encounters, map: © Merenkulkulaitos lupa nro 1321 / 721 / 2008

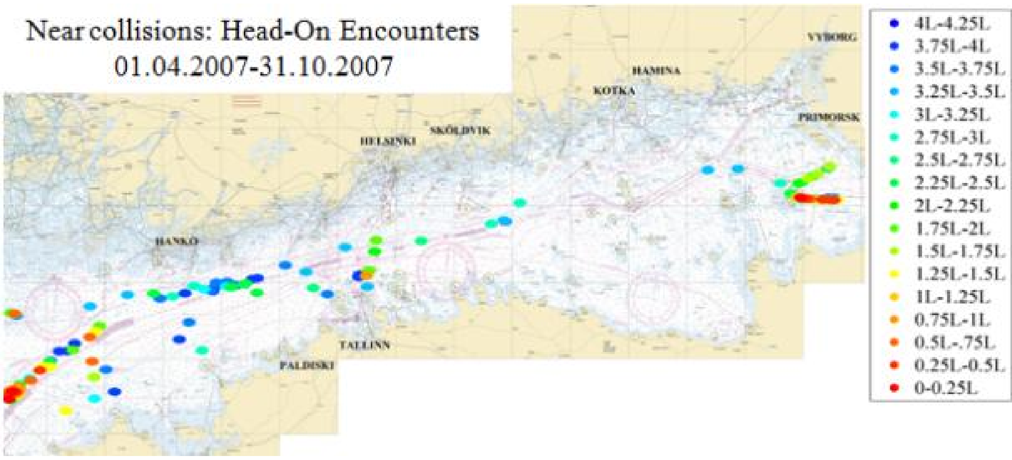

Figure 5. Near collision locations: head-on encounters, map: © Merenkulkulaitos lupa nro 1321 / 721 / 2008

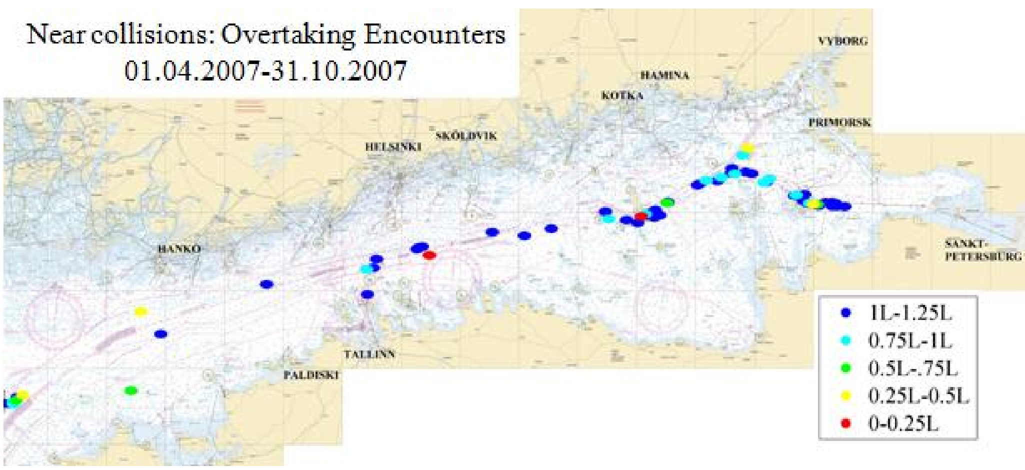

Figure 6. Near collision locations: overtaking encounters, map: © Merenkulkulaitos lupa nro 1321 / 721 / 2008

Table 2. Results of the near-collision analysis for three types of encounters for the Gulf of Finland and the Fujii's domain as a criterion.

Encounter type	Number of near-collisions within analysed time period	Annual number of near-collision	Adjusted number of near-collision using ratio from Table 3
Crossing	170	290	22
Overtaking adjusted	174	298	20
Head-on	178	305	23
SUM	522	893	65

4.2. Results of the analysis for the Danish Straits

The results obtained for the Danish Straits are depicted in Table 3, they yield the total number of near-collision, and no distinction is made for the types of encounter (*N_(n-c)*). However the post-analysis evaluation is made, and the encounters which are not compromising the safe passage of two ships are disregarded. The evaluation of the results is made with the presence of experts, who analyze each situation and label it as "normal operation" of "abnormal operation". Therefore the numbers of encounters in which two ships are in a critical situation are determined (*N_(n-c_corr))*. Finally the ratio of numbers of evaluated near-collisions to the number of such situations obtained from the algorithm is obtained. This ratio is then applied as a thinning factor for the results obtained in the Gulf of Finland, which has not been evaluated, see Table 2

Table 3. Results of the near-collision analysis for the Danish Straits with the DSE as a criterion.

Encounter type	Near-collisions N_(n-c)	Near-collisions evaluated N_(n-c_corr)	Ratio N_(n-c_corr)/N_(n-c)
All	432	31	0.072

5. DISCUSSION

In this report two methods to detect (possible) near collisions between two vessels, valid for open sea areas in ice-free conditions are presented. The methodology uses the concept of the ship domain to detect critical encounters between vessels from a AIS database. The proposed method has been successfully applied to the summer traffic in the Gulf of Finland and the Danish Straits, using AIS data from 2007.

It has been found, that the ship domain as proposed by Fujii (1971) is not well suited for navigation areas where traffic separation schemes are enforced, especially for overtaking encounters, as the Fujii domain does not make a distinction between the different encounter situations. Moreover, the Fujii-domain is symmetric with respect to the ship's sailing direction, which implies that ships passing in front of the bow present the same risk as ships passing at the stern. This is of course unrealistic. Other domains, such as the one proposed by Kijima et al. (2003) or Pietrzykowski (2008) take this into account. However, the Kijima domain is larger than the Fujii domain, which nullifies this benefit. The Pietrzykowski domain is a fuzzy domain, which takes the specific encounter situation between the vessels into account, and seems the most appropriate domain to be used in a near collision detection method. However, this domain is at present still under active development and cannot yet be applied to ships of any type or size.

Another point of reflection is the AIS data sampling rate. Unfortunately, not all transmitted AIS messages were retained in the TraFi database: the actual sample rate is about 5 minutes on average. Due to the relatively slow nature of maritime traffic, this is sufficient for a first rough assessment of near collisions. However, for a detailed analysis of the actions taken by vessels in these close encounters, a higher sampling rate is required. In addition, it should be acknowledged that this sampling rate may influence the obtained results somewhat.

A point of caution should be raised concerning the evaluation of the near collisions. The basic assumption that a violation of a ship indicates a near collision, may be questioned. This is because the intention of the vessel movements should be taken into account in the question whether a certain encounter is dangerous or not. In order to assess this, the vessel trajectories of vessels within close quarters should be plotted, along with possible background traffic. The proposed method should be seen as a first filtering of the cases which need to be studied in more detail by means of trajectory plotting and expert judgment.

A final comment concerns the terminology used in this paper, and its implication. The International Maritime Organization defines a near miss as *"a sequence of events and/or conditions that could have resulted in loss. This loss was prevented only by a fortuitous break in the chain of events and/or conditions."* see IMO (2008). This implies that in an assessment of near misses, the question why the accident has been averted is paramount. However, based on analysis of AIS data, this question cannot be answered for near collisions. However, the conditions in which these events occur (i.e. where, what and how) can be evaluated.

6. REFERENCES

1. Aarsæther, K.G. and Moan, T. 2009. Estimating Navigation Patterns from AIS. The Journal of Navigation 62, 587-607.
2. Fujii Y, Tanaka K. 1971. Traffic capacity. The Journal of Navigation, 24, 543-552.
3. Goerlandt F, Kujala P. 2011. Traffic simulation based ship collision probability modeling. Reliability Engineering & System Safety, 96, 1. 91-107.
4. Goerlandt F, Montewka J, Kujala P. 2011. Risk analysis of tanker collisions in the Gulf of Finland. In: Weintrit A, editor. Marine navigation and safety of sea transportation. London, Taylor and Francis.
5. Goodwin EM. 1975. A statistical study of ship domains. The Journal of Navigation, 57, 329-341.
6. Graveson, A. 2004. AIS – an inexact science. The Journal of Navigation, 57, 339-343.

7. Inoue K, Kawase M. Innovative probabilistic prediction of accident occurrence. In: Weintrit A, editor. Marine navigation and safety of sea transportation. London, Taylor and Francis: p.31-34.
8. International Maritime Organization (IMO). 2001. Guidelines for the onboard operational use of shipborne Automatic Identification Systems (AIS), Resolution A.917(22).
9. International Maritime Organization (IMO). 2002. COLREG: Convention on the international regulations for preventing collisions at sea 1972, as amended.
10. International Maritime Organization (IMO). 2008. Guidance on near-miss reporting. MSC-MEPC.7/Circ.7 10 October 2008
11. Jingsong Z, Zhaolin W, Fengchen W. 1993. Comments on ship domains. The Journal of Navigation, 46, 422-436.
12. Kristiansen S. 2010. A BBN approach for analysis of maritime accident scenarios. In Ale, Papazoglou & Zio (Eds.), ESREL 2010, London. Taylor and Francis: 382-389.
13. Kijima K, Furukawa Y. 2003. Automatic collision avoidance system using the concept of blocking area. Proceedings of IFAC conference on maneuvering and control of marine craft, Girona, Spain.
14. Kujala P., Hänninen M., Arola T. and Ylitalo J. 2009. Analysis of the marine traffic safety in the Gulf of Finland. Reliability Engineering & System Safety 94, 8: 1349-1357.
15. Martins MR, Maturana MC. 2010. Human error contribution in collision and grounding of oil tankers. Risk Analysis; 30(4):674-698.
16. Montewka, J., Hinz, T., Kujala, P. and Matusiak, J. 2010. Probability modeling of vessel collision. Reliability Engineering & System Safety 95, 5. 573-589.
17. Montewka J., Goerlandt F., Kujala P. 2012. Determination of collision criteria and causation factors appropriate to a model for estimating the probability of maritime accidents. Ocean Engineering 40, pp. 50-61
18. Norris, A. 2007. AIS implementation – success or failure. The Journal of Navigation, 60, 373-389.
19. Pietrzykowski Z. 2008. Ship's fuzzy domain – a criterion for navigational safety in narrow fairways. The Journal of Navigation, 51, 499-514.
20. Vincenty T. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. Survey Review, 176, 88-93.
21. Wang N, Meng X, Xu Q and Wang Z. 2009. A unified analytical framework for ship domains. The Journal of Navigation, 62, 643-655.