The Journal of Navigation (2024), 1–19
doi:10.1017/S0373463324000109

RESEARCH ARTICLE

Exploration of COLREG-relevant parameters from historical AIS-data

Inger B. Hagen, Karen S. Knutsen, Tor Arne Johansen,* and Edmund Brekke

Department of Engineering Cybernetic, Norwegian University of Technology and Science, Trondheim, Norway.
*Corresponding author: Tor Arne Johansen; Email: tor.arne.johansen@ntnu.no

Received: 30 June 2022; **Accepted:** 12 March 2024

Keywords: automatic identification system; COLREGs; ship behaviour

Abstract

Reliable anti-collision control algorithms conforming with the rules regulating traffic at sea, the International Regulations for Preventing Collisions at Sea (COLREG), are essential for the deployment of autonomous vessels in waters shared with other ships. The development of such methods is an active field of research. However, little attention has been given to how these rules are interpreted by experienced mariners, and how such information can be parametrised for use in automatic control systems and autonomous ships. This paper presents a method for exploiting historical automatic identification system (AIS) data to characterise parameters indicating the prevalent practices at sea in encounters with high collision risk. The method has been tested on data gathered in areas off the Norwegian coast over several years. Statistics on relevant parameters from the resulting dataset and the relation between them is presented. The results indicate that the strongest influence on vessel behaviour is the type of situation, and the amount of land and grounding hazards in the vessel's proximity.

1. Introduction

Maritime autonomous surface ships (MASS) and autonomous surface vehicles (ASVs) sharing waters with conventional ships is an inevitable consequence of the shift towards autonomy within the maritime sector. For this to be a viable and safe prospect with respect to collision avoidance, vessel interactions must be regulated.

The COLREG (IMO), which regulates the behaviour of vessels during encounters at sea, arose in a world where seafaring vessels were controlled by experienced navigators and captains. The rules were therefore left purposely vague, trusting the sailors' expertise (good seamanship) in interpreting how the rules should be applied in different situations. This intentional vagueness poses a challenge when developers attempt to implement the COLREG within the control algorithms of autonomous vessels. The question at hand is: How to transform the tacit knowledge of experienced sailors into documented information, suitable for implementation in computer programs?

One attempt at such a transformation is the concept of the ship domain. As remarked by Fujii and Tanaka (1971), there exists an area around ships underway that navigators tend to avoid. Originally denoted the effective domain, the term ship domain, introduced by Goodwin (1975), has since become the more commonplace expression, defined as the area around a ship that the officers on watch (OOW) would like to maintain empty of other vessels. While Fujii and Tanaka (1971), Goodwin (1975) and other earlier works were limited to information gathered by radar, the introduction of AIS provided researchers with more detailed information concerning vessels and their movements. This was exploited by Gucma and Marcjan (2012), where AIS data recorded over a one year period in the Gulf of Pomerania were employed in the creation of probabilistic models for the ship domains of different vessel categories (tankers, passenger and cargo ships) and different encounter types (head-on, crossing and overtaking). Similar studies into ship domain properties that also make use of AIS data are Pietrzykowski and Magaj (2017) and Procee et al. (2018), which focus on areas with high traffic density, and Hansen et al. (2013), where the limits for comfortable traffic flow in narrow channels are explored. In the latter, the domain is measured in ship lengths, removing any correlations to vessel size.

While the above mentioned works aim to determine the ship domain, the present paper investigates the characteristics of actions taken to avoid the violation of this area. Whether a small course change at a large distance is preferable to a large course change at a smaller distance, or whether the size of the course change is dependent on the size or the speed of the vessels, are examples of questions that must be answered. The answers to such questions are central, both for collision avoidance algorithms that aim to adhere to COLREG in given situations, such as those raised by Hagen et al. (2018), Zhao and Roh (2019), Kim et al. (2021), and for the evaluation/verification of such methods, exemplified by Woerner et al. (2019), Pedersen et al. (2019), Torben et al. (2022) and Hagen et al. (2023). As a step towards the safe coexistence of MASSs and conventional ships, this work attempts to: (a) identify parameters that can be used in characterising COLREG compliance; (b) ascertain acceptable values or intervals for COLREG parameters; and (c) identify correlations between parameters and external factors.

As in the previous works on ship domain, our source of information is historical AIS data. These originate from vessels in normal operation, thereby giving a realistic impression of vessel behaviour. Data gathered over several years from three different areas off the coast of Norway were studied to obtain a sufficient number of COLREG-related encounters for the observations regarding customary COLREG interpretation to be reliable. A database was constructed, containing encounters where at least one of the involved vessels performed an evasive manoeuvre. In addition to trajectory data and vessel information, supplementary parameters characterising the encounter and the vessels' behaviour was extracted from each situation and added to the dataset, see Section 4. These data were then examined to identify acceptable values for parameters that we consider as important, such as distance at closest point of approach (DCPA) and time to closest point of approach (TCPA) at the time when an evasive manoeuvre is initiated. Attempts to identify determining factors for the vessel's behaviour were also made, with special focus on the amount of land in the area and the encounter type. Preliminary results were given by Vassbotn (2022) and Knudsen (2022).

The remainder of this paper is structured as follows. Section 2 explains the choice of parameters for the study with regards to the COLREG. A brief overview of AIS is given in Section 3, which also contains a presentation of the raw datasets and pre-processing methods employed. Techniques used for extracting relevant parameters are explained in Section 4 followed by selected results in Section 5. The paper concludes with a brief discussion around the results, along with some concluding remarks in Sections 6 and 7, respectively.

2. COLREG parameters

Previous work on collision avoidance methods aiming for COLREG compliance have focused on a subset of the steering and sailing rules concerning the conduct of vessels in sight of one another (Vagale et al., 2021). The most commonly considered rules are rules 13–15, which describe the desired behaviour in overtaking, head-on and crossing situations. The actions required in these situations are further specified by rules 16, 17 and 8. However, several points within these rules can be regarded as open to interpretation. The following points are considered by the authors as being the most pertinent:

Rule 8(a) *'Any action to avoid collision shall be . . . made in* **ample time** *. . . '*;

Rule 8(b) *'Any alteration of course and/or speed to avoid collision shall . . . be* **large enough** *to be readily apparent to another vessel observing visually or by radar'*;

Rule 8(d) *'Action taken to avoid collision with another vessel shall be such as to result in passing at a* **safe distance***'*.

Interpreting these phrases is equal to identifying acceptable values for: (a) when an avoidance manoeuvre should be initiated ($TCPA_{man}$) (b) the necessary size of a course change ($\Delta\chi$) or speed change (ΔU) and (c) the appropriate DCPA. However, these values may differ depending on the type of situation and the vessels involved. Moreover, external factors may also affect these values. The factors focused on in this work are the amount of land in the area where the situation occurs, the speed and size of the involved vessels, and the number of vessels in the vicinity.

3. AIS data

3.1. The automatic identification system

The AIS is an automatic tracking system and navigational aid that allow vessels to broadcast both static and dynamic information about themselves via digital very high frequency (VHF) radio transmission, and simultaneously receive the same data from vessels nearby. The data studied in this work were gathered by the national AIS network, AIS Norway, which consists of shore-based facilities covering the area from the baseline to 40–60 nautical miles from the coast, along with satellites covering offshore areas, and is operated by the Norwegian Coastal Administration (NCA).

Of the information contained in the AIS-messages, the fields most relevant to this work are the identification (maritime mobile service identity (MMSI)) and position of the vessel, the position time stamp in coordinated universal time (UTC), speed over ground (SOG), course over ground (COG) and navigational status. The MMSI is static and entered into the device upon installation on the vessel. The remaining fields mentioned are dynamic and automatically updated from ship sensors, except for the navigational status, which must be changed manually by the OOW. Static data are transmitted every 6 min, or upon request, while the update rate for dynamic information depends on the vessel’s speed and course alterations.

The information available is, however, limited to vessels equipped with AIS transponders, which is required on ships of 300 gross tonnage and above engaged on international voyages, cargo vessels of 500 gross tonnage and above not engaged on international voyages, and all passenger ships, according to International Convention for the Safety of Life at Sea (SOLAS). The AIS referred to by the SOLAS is commonly known as AIS Class A, but less expensive units, termed AIS Class B, intended for non-SOLAS vessels such as domestic commercial vessels and pleasure crafts, are also available. Class B units communicate and operate in conjunction with Class A units, but have less functionality. When studying encounters collected from AIS data, one must therefore be aware of the possibility of the situation being influenced by vessels not recorded in the data.

3.2. Datasets

The AIS data that form the basis for this research were gathered from the areas shown in Figure 1. These areas were chosen to capture encounters occurring in both open waters and in coastal areas where grounding hazards may restrict the vessels’ movements. The northern dataset consists of data from the coast of mid-Norway, collected from 01.01.2018 to 31.07.21, excluding data from 01.06.2019 to 31.07.2019. The southern dataset contains data from Skagerak, the waters between Norway and Denmark, dated from 01.01.2018 to 31.12.2019. The western dataset is from the area around Bergen and contains data from 01.01.2019 to 31.12.2020.

3.3. Pre-processing

The pre-processing procedure is based on the work presented by Vassbotn (2022). The following paragraphs describe the steps taken to extract slices from the original dataset containing at least one encounter, these slices are henceforth called cases.

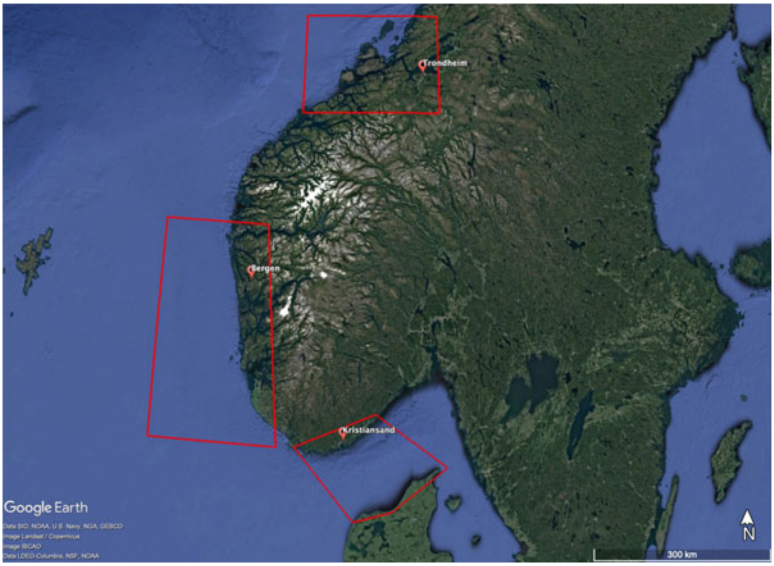

Figure 1. *Areas where AIS data were collected, outlined in red.*

3.3.1. Masking docking sites

Locations where multiple vessels gather, such as ports, fish farms and ship-yards, form clusters in the datasets. Data transmitted from such locations originate from vessels that are either docked or, if in transit, their conduct is likely to be governed by local rules or conventions and should therefore be excluded from the datasets. The density-based spatial clustering of applications with noise (DBSCAN) algorithm of Ester et al. (1996) was employed to identify possible docking locations from the positional information of vessels with their navigational status set to 'at anchor'. These were then confirmed by manually checking for actual docking facilities at the site. A mask was placed over the site, covering an area with a radius of five kilometres (approximately $2 \cdot 7$ nautical mile (NM)) around the identified dock, excluding data transmitted within that area.

3.3.2. Down sampling and case identification

When possible, marine vessels tend to travel in straight lines, making linear approximation a suitable method for down sampling. Disregarding data points where positional data can be linearly interpolated with an accuracy of ten metres[1] significantly reduces the amount of data to be processed in the case identification.

The first step in the case construction is to locate two vessels with an approximate DCPA of less than five kilometres (approximately $2 \cdot 7$ NM). Around this approximate closest point of approach (CPA), a time-frame for the case is constructed, defined as the period where the vessels are within 15 km of each other. The geographical extent of the case is then found by taking the rectangular area enclosing these trajectory segments plus a margin of 10 km. Data transmitted within these spatial and temporal limits are then extracted from the dataset and combined into a case. A plot including all positional data from one example case is shown in Figure 2.

3.4. Interpolation

The relevant data have now been separated into smaller cases containing one or more encounters. For each case, the sample times of the AIS messages containing dynamic information are synchronised using

---

[1] The approximate accuracy of AIS positional data as specified by the International Maritime Organization (IMO) Organization (2015).

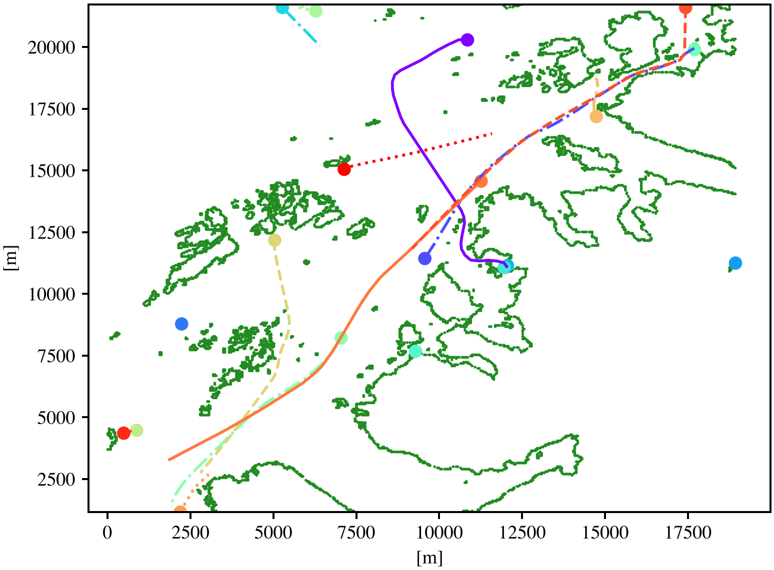

Figure 2. *Example plot of vessel trajectories contained in one case.*

linear interpolation with a sample interval (Δt) of one minute. Relevant fields from messages containing static information, such as navigational status, ship type and length, are also added for each vessel.

4. Parameter extraction

4.1. Cases and encounters

While the static parameters can be directly extracted from the case data, other information, such as situation type, must be calculated based on the trajectory data from each vessel. This is done separately for each two-vessel encounter. As in the case construction, an encounter is defined as two vessels with an approximated DCPA less than five kilometres. A vessel may therefore be involved in several encounters within the same case, each encounter resulting in a separate set of parameters. As a consequence, multi-vessel encounters will be divided into the appropriate number of two-vessel encounters.

An overview of the most relevant parameters can be found in Table 1. The remainder of this section explains how parameters are extracted from case data, and specifies the conditions that must be met for an encounter to be included in the final dataset, based on Vassbotn (2022).

4.2. Situation type

To investigate how the extracted parameters are affected by the type of COLREG situation, each encounter was classified as either overtaking, crossing or head-on. For overtaking and crossing situations, the vessels involved were also assigned a role as either the stand-on or give-way vessel, cf. Rules 16 and 17. The classification method is based on an algorithm (Woerner et al., 2019, Alg. 3) categorising encounters according to the vessels' relative poses, see Figure 3. We note that since AIS data do not contain heading information, we use COG instead. Since we only consider vessels that have a substantial speed, the difference between COG and heading is expected to be limited to a few degrees. In addition, we removed from the dataset the situations involving ferries with stand-on rights according to local rules, and vessels where Rule 18 could change their role or responsibility, e.g. vessels engaged in fishing activities (navigational status set to 7) as well as other types of vessels mentioned in this rule. Since we also have validated each case by manual inspection, we do not consider this to be a major error source when classifying situation types.

Table 1. *Selection of parameters included in the final dataset.*

Information type	Parameter
Situation	Originating dataset ID Originating case ID Date Situation type Number of vessels included in case DCPA
Vessels	ID Length Average speed during encounter
Manoeuvre	Total course angle change Total speed change TCPA at manoeuvre start TCPA at manoeuvre end

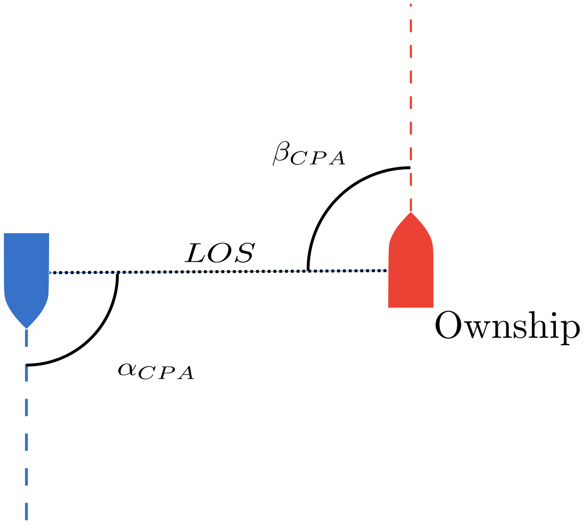

Figure 3. *Relative pose between vessels: contact angle,* $\alpha \in [-180°, 180°)$*, and relative bearing,* $\beta \in [-180°, 180°)$*. The red ship is the ownship, while the blue ship is the target ship.*

4.3. Complete/non-complete situations

Due to the spatial and temporal limits of the case data, some of the encounters included in the case may be incomplete and should therefore not be included in the final dataset. Encounters are marked as incomplete if:

- the vessels are in an overtaking, crossing or head-on situation according to COLREG at the first and/or last time step included in the case, indicating an ongoing or unresolved situation;
- DCPA is less than 50 m, indicating a collision or deliberate interaction;
- the duration of the situation is less than five minutes, since such situations might not adhere to the COLREG requirement for early action;
- either vessel moves less than 100 m.

Table 2. *Parameter values employed in the manoeuvre detection procedure.*

Parameter	Value	Parameter	Value
$\epsilon_{\dot{\chi}}$	$0 \cdot 01$ rad/s	$\epsilon_{\dddot{\chi}}$	$0 \cdot 005$ rad/s^3
$\epsilon_{\ddot{\chi}}$	$0 \cdot 01$ rad/s^2	$\epsilon_{\dot{U}}$	$0 \cdot 8$ m/s

In addition, situation specific requirements on the vessels' relative poses, given by α and β (see Figure 3), must be fulfilled for an encounter to be labelled as complete.

Head on: The vessels must have passed each other; $|\alpha|_{max} \geq 90^\circ \wedge |\beta|_{max} \geq 90^\circ$.

Overtake: The vessels must be close to parallel at CPA; $30^\circ \leq |\alpha_{cpa}| \leq 90^\circ \wedge 30^\circ \leq |\beta_{cpa}| \leq 90^\circ$.

Stand-on vessel: Must have passed give-way vessel; $|\alpha|_{min} \leq 90^\circ \wedge |\beta|_{max} \geq 90^\circ$.

Give-way vessel: Must have been passed by stand-on vessel; $|\beta|_{min} \leq 90^\circ \wedge |\alpha|_{max} \geq 90^\circ$.

Crossing: One of the vessels must cross the other's path; there exists a time step k within the encounter such that $\operatorname{sign} \alpha_k \neq \operatorname{sign} \alpha_{k+1} \vee \operatorname{sign} \beta_k \neq \operatorname{sign} \beta_{k+1}$.

4.4. Manoeuvre detection

Manoeuvres are manifested by a change in course angle (χ) and/or speed over ground (U). Due to differences in the nature of the signals, the detection process studies each component separately making use of the variables' derivatives, found by finite central differences, based on Vassbotn (2022).

4.4.1. Course change

A signal change detection method (Basseville and Nikiforov, 1993) is employed to identify course changes in the data. It is based on derivatives that are smoothed by Gaussian convolution, and is similar to edge detection in image processing (Canny, 1986). A manoeuvre is detected at the kth sample if the following conditions are fulfilled:

$$\begin{aligned} &|\dot{\chi}_k| \geq \epsilon_{\dot{\chi}}, \\ &|\ddot{\chi}_k| \leq \epsilon_{\ddot{\chi}}, \quad \operatorname{sign} \ddot{\chi}_k \neq \operatorname{sign} \ddot{\chi}_{k-1}, \\ &|\dddot{\chi}_k| \geq \epsilon_{\dddot{\chi}}, \quad \operatorname{sign} \dddot{\chi}_k \neq \operatorname{sign} \dot{\chi}_{k-1}, \end{aligned} \tag{1}$$

where $\epsilon_{\dot{\chi}}$, $\epsilon_{\ddot{\chi}}$ and $\epsilon_{\dddot{\chi}}$ are adjustable parameters, their values can be found in Table 2. The start and end indices (k_{man} and k_{stop}) of the manoeuvre are defined as the time of the nearest third derivative zero, giving the following expression for the total course change of the manoeuvre:

$$\Delta\chi = \chi(k_{stop}) - \chi(k_{man}) = \sum_{k_{man}}^{k_{stop}} \dot{\chi}\Delta t. \tag{2}$$

4.4.2. Speed change

The speed changes of marine vessels tend to be small and well defined, permitting the use of a simpler detection method where only the first derivative is used. A manoeuvre is in progress at the kth sample if the following condition is fulfilled:

$$|\dot{U}_k| \geq \epsilon_{\dot{U}}, \tag{3}$$

where $\epsilon_{\dot{U}}$ is an adjustable parameter, see Table 2. The total speed change of the manoeuvre is given by

$$\Delta U = U(k_{stop}) - U(k_{man}) = \sum_{k_{man}}^{k_{stop}} \dot{U}\Delta t. \tag{4}$$

4.4.3. Evasive and non-evasive manoeuvres

To exclude encounters where vessel behaviour is directed by factors other than the COLREG, only encounters where the give-way vessel performs an evasive manoeuvre are retained. This will exclude multi-vessel encounters where the manoeuvres made are restricted by, or intended to avoid, vessels not considered in the two-vessel encounter. It will also exclude port manoeuvres since they may have been agreed using radio communication or other information not available through AIS. A manoeuvre is tentatively marked as evasive if the following is true:

- the DCPA is predicted at the start and end points of the manoeuvre using a constant velocity (CV) model for both vessels and the manoeuvre causes an increase in the predicted DCPA;
- for head-on situations and for give-way vessels in crossing situations, any course change must be towards the starboard side.

For encounters containing multiple evasive manoeuvres, only the manoeuvre causing the largest increase in DCPA is included in the set of parameters.

However, these measures do not guarantee that the primary purpose of the manoeuvre is collision avoidance. Other possible reasons include grounding hazards and vessels not included in the situation (with or without AIS transponders). To increase the likelihood that the intention behind the manoeuvre is truly collision avoidance, the trajectories from each encounter were manually inspected by the authors before its parameters are added to the final dataset. If there was some doubt, the manoeuvre was excluded from further analysis.

4.5. Missing and erroneous messages

For some vessels, the AIS messages do not contain information on the vessel size and the information had to be retrieved manually from an online resource, Marine Traffic, using the vessels' MMSI.

5. Results

5.1. Encounters extracted from the data

The procedure detailed in the previous sections produces a dataset containing parameters extracted from encounters where the DCPA is less than 5 km, and an evasive manoeuvre has been performed in accordance with the qualitative behaviour prescribed by the COLREG. From the 2,974 cases extracted from the raw AIS data, 28,421 encounters were identified. Of these, 782 encounters were considered complete, COLREG compliant, and containing both an evasive manoeuvre and the necessary information. The used dataset thus consists of 782 entries, whereof 110 crossing, 230 overtaking and 442 head-on situations. The rest of this section presents statistics inferred from this dataset and shows the relations between selected parameters.

5.2. External parameters

One of the questions that this work set out to answer is whether external parameters affect vessel behaviour. Considered as particularly interesting is vessel size, the type of area in which the encounter occurs and traffic density in the area.

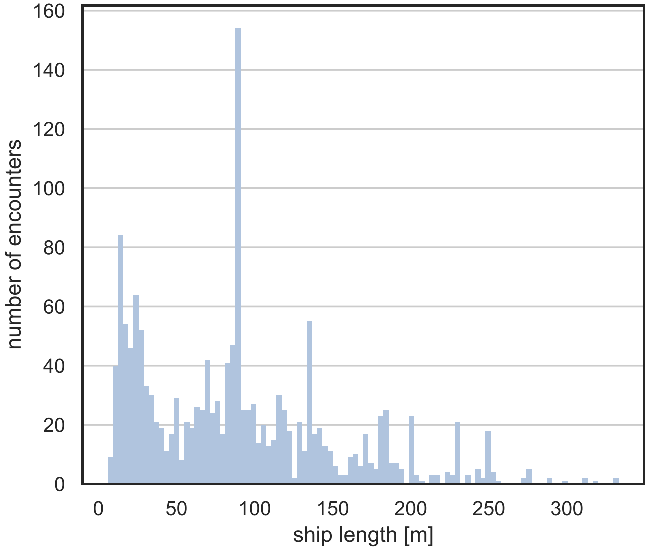

Figure 4. *Distribution of vessel lengths in the dataset, the lengths of both ownship and obstacle ship are included.*

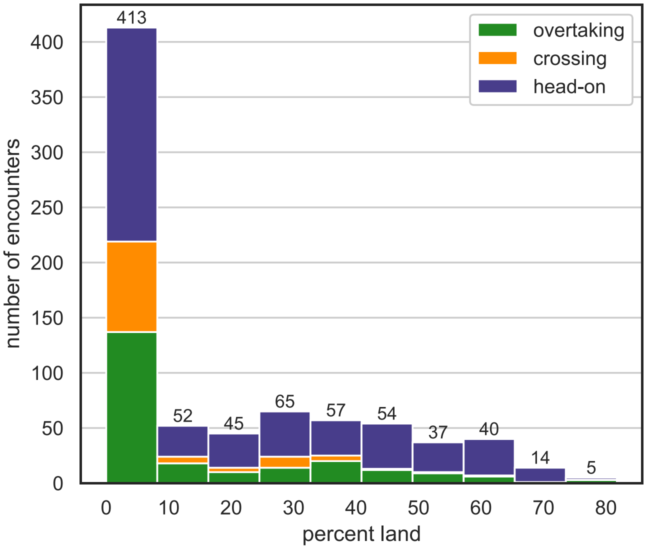

Figure 5. *Distribution of encounters versus the percentage of land in a 14 by 14 km area surrounding the encounter, according to situation type.*

The length of a vessel is used as a representation of its size. This was chosen because of its significant effect on both ship dynamics and ship domain, while the recorded lengths in the parameter set range from 6 to 333 m, see Figure 4, most (approximately 73%) of the vessels are between 15 and 150 m long.

Instead of attempting to classify encounter locations into area types, such as open water, archipelago or fjord, it was chosen to calculate the percentage of land in a 14 by 14 km area around the encounter, for simplicity, this number will from now on be referred to as land coverage. While this classification method may seem simplistic, it is used to indicate the likelihood of a vessels' behaviour being affected by grounding hazards and is sufficient for our purposes. The distribution of encounters according to percentage of land is shown in Figure 5. Of the 782 encounters, almost half of them occur in areas with less than 10% land coverage.

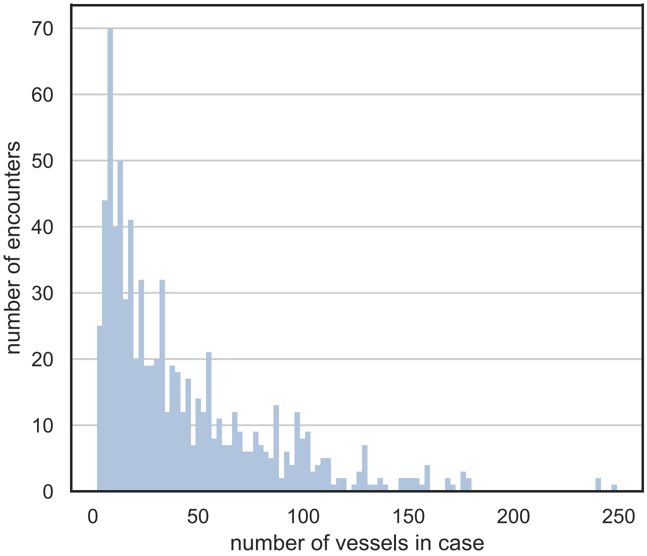

Figure 6. *Distribution of the number of vessels in the originating case for each situation.*

The number of vessels included in the originating case was used as a measure for traffic density, the distribution is shown in Figure 6. Again, this is a simple and straightforward approach but may reveal whether traffic density is an important factor in vessel behaviour.

5.3. DCPA

The average DCPA with regards to situation type and land coverage is shown in Figure 7. When disregarding the land coverage, overtaking situations have the highest average DCPA at 1,007 m, followed by crossing situations at 886 m and head-on situations at 597 m. The difference between the highest and lowest is 410 m, which is more than four ship lengths for 64% of the recorded vessels. For all situation types, the average DCPA decreases when the amount of land in the area increases.[2]

The relation between DCPA and land coverage, according to situation type, is also illustrated in Figure 8. Notable in this plot is the 121 overtaking situations that occurred in areas with zero land coverage, where the DCPA ranges from 60 to 4,700 m. The large spread in the recorded values makes it difficult to recommend a general DCPA value suitable for this type of situation. For crossing situations in areas with more than 60% land coverage, only one situation falls within the category and no recommendations can be given. For the remaining categories, the mean values and typical ranges, e.g. represented as 5% and 95% percentiles as given in Table 3, appears to be a reasonable estimate for the preferred DCPA as practised at sea.

The relationship between DCPA and the average speed of the ownship, again according to situation type, is shown in Figure 9. The plot shows that vessels overtaking others, not surprisingly, tend to have higher average speeds than vessels in other situations. It also appears like the lowest DCPA seems to increase slightly with the speed, but no strong correlations are evident in the data. With regards to vessel length, there does again seem to be a slight correlation between the length of the ownship and the DCPA, which can be seen in Figure 10.

---

[2] The exception is crossing situations where the average DCPA for the highest percentages of land is above the situation mean. However, the dataset only contains one entry in this category which has been excluded from the plots to avoid misrepresenting the results since a single observation is not statistically significant.

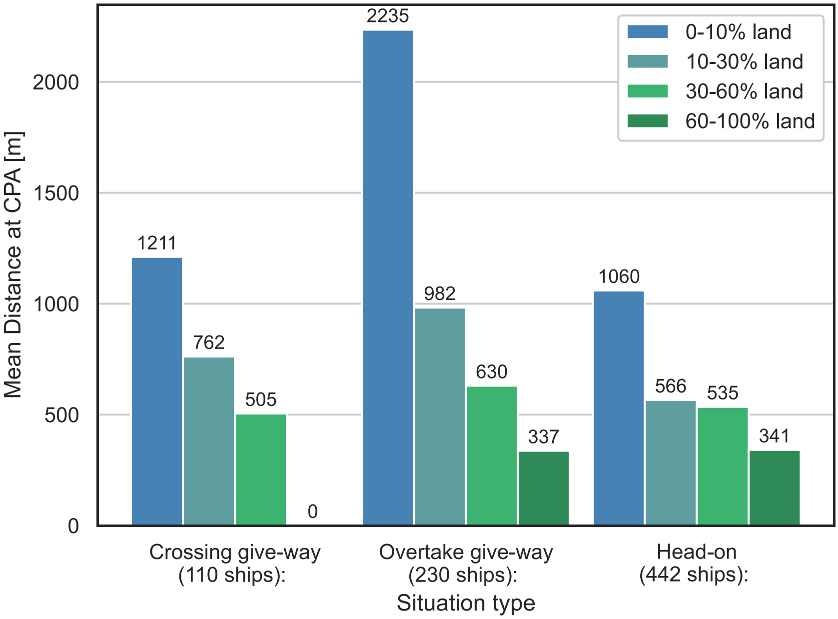

Figure 7. *Mean DCPA for each situation type, according to situation type and the area's land coverage.*

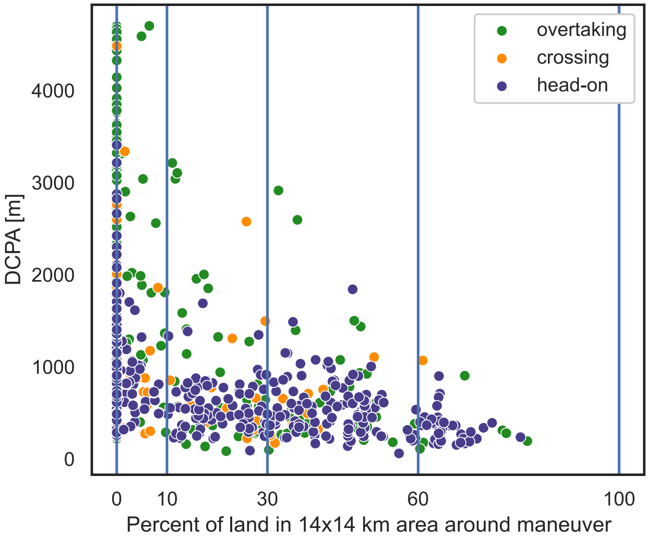

Figure 8. *DCPA versus land coverage, according to situation type.*

5.4. TCPA

The TCPA is the predicted time to CPA, assuming both vessels keep constant course and speed, and is often made available to the OOW through navigational aids. The TCPA when an evasive manoeuvre is initiated ($TCPA_{man}$) is therefore a parameter of interest. The means of the $TCPA_{man}$ according to land coverage, Figure 11, show that land coverage is a determining factor. Independent of land coverage, the mean value is 2,454 s (40 · 9 min) for crossing situations, 2,338 s (40 · 0 min) for overtaking situations and 1,658 s (27 · 6 min) for head-on situations. Encounters occurring in open waters or with little land coverage display significantly larger values for $TCPA_{man}$ than encounters with higher land coverage. When disregarding areas with less than 10% land coverage, the mean $TCPA_{man}$ values are 1,101 s

***Table 3.** Statistics extracted from the data shown in Figure 8.*

Land cover	Mean	Std. dev.	5% percentile	95% percentile
Distance at CPA in crossing situations				
0–10%	1,210 m	739 m	404 m	2,736 m
10–30%	761 m	596 m	285 m	1,820 m
30–60%	504 m	282 m	187 m	946 m
Distance at CPA in overtaking situations				
0–10%	2,234 m	1,220 m	607 m	4,523 m
10–30%	982 m	853 m	150 m	3,066 m
30–60%	630 m	551 m	191 m	1,484 m
60–100%	336 m	221 m	137 m	712 m
Distance at CPA in head-on situations				
0–10%	1,059 m	583 m	313 m	2,235 m
10–30%	565 m	278 m	241 m	958 m
30–60%	534 m	279 m	220 m	1,035 m
60–100%	340 m	155 m	168 m	669 m

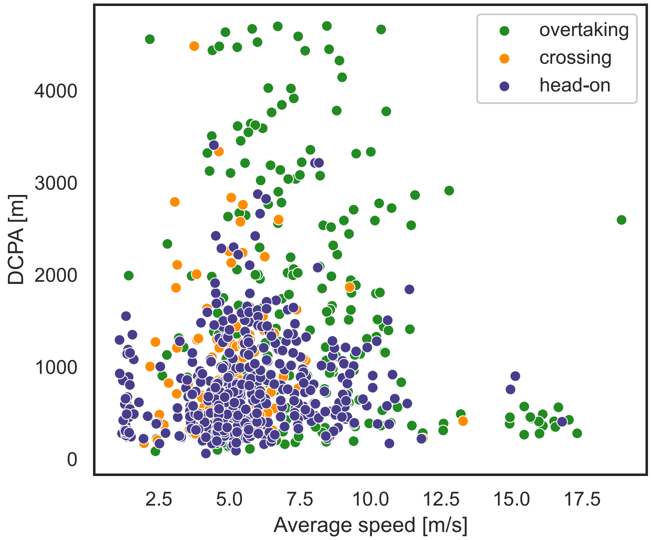

***Figure 9.** DCPA versus own ship's average speed during the encounter, according to situation type.*

(18 · 3 min) for crossing situations, 825 s (13 · 8 min) for overtaking situations and 861 s (14 · 4 min) for head-on situations. Again, due to a lack of data, the mean is not shown for crossing situations in areas with more than 60% land coverage.

To put this into perspective; the difference in mean $TCPA_{man}$ between crossing and overtaking situations (1,101 and 825 s, respectively) equals a distance of approximately 1,380 m for a vessel travelling at a relative velocity difference of 10 knots, or approximately 5 metres per second. This represents more than four ship lengths for all the recorded vessels and more than 8 ship lengths for 88% of them.

It would be natural to assume that vessel speed could influence the choice of when to make a manoeuvre. Figure 12 shows that the $TCPA_{man}$ has no significant dependence on the ownship's average speed, except for the high-speed vessels travelling faster than 10 m/s, and that are typically smaller and

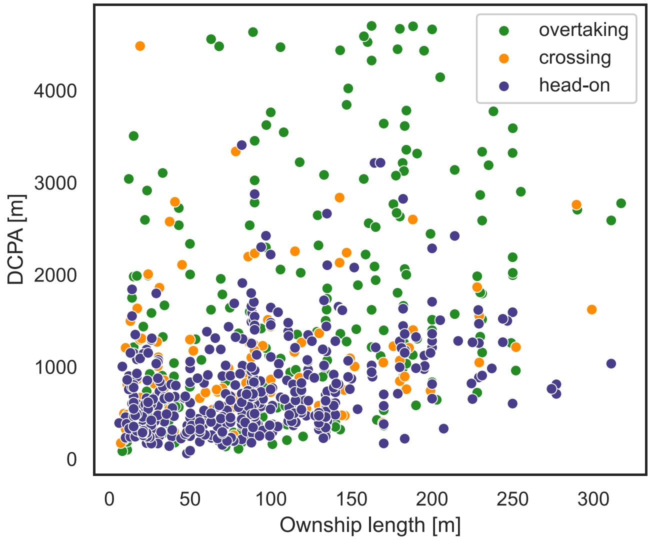

Figure 10. *DCPA versus ownship's length during the encounter, according to situation type.*

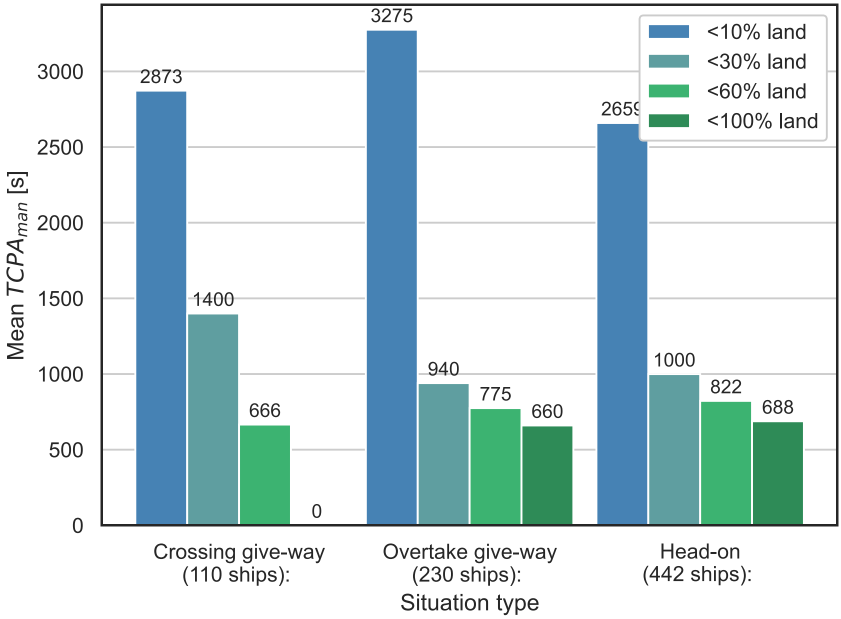

Figure 11. *Mean predicted TCPA at manoeuvre start, according to situation type and the area's land coverage.*

highly manoeuvrable. While for larger and slower ships, there is no significant difference in terms of TCPA, it is clear that manoeuvres initiated at similar TCPA translates to manoeuvres being initiated at larger distances between vessels when vessels proceed at higher speeds.

5.5. Course change

A change in course angle is more readily apparent to other vessels than a change in speed and, especially in open waters, it is the preferred action to avoid collision according to the COLREG. It is thus a very

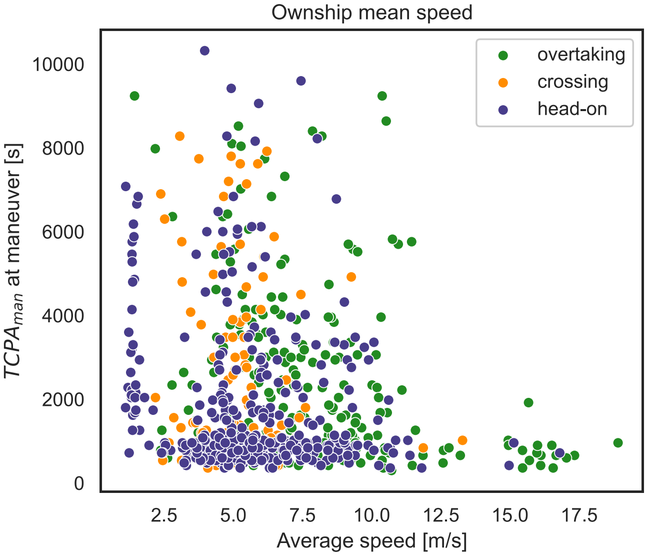

Figure 12. *Predicted TCPA at manoeuvre start versus ownship's average speed, according to situation type.*

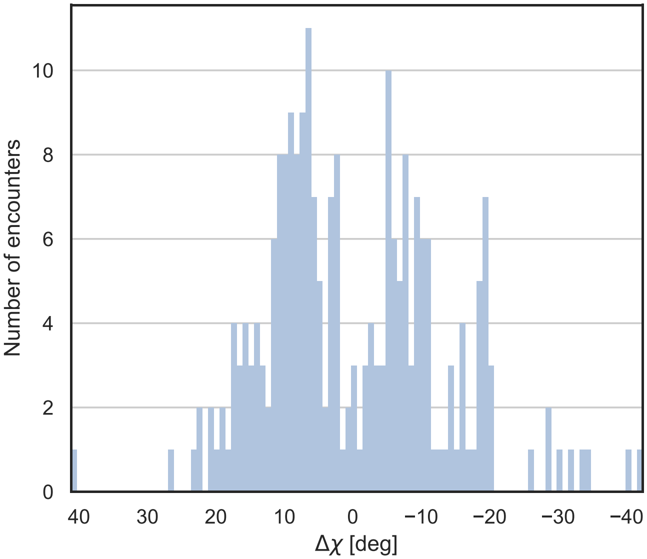

Figure 13. *Distribution of change in course angle for the 230 overtaking manoeuvres.*

relevant parameter for both collision avoidance and evaluation algorithms. In both head-on and crossing situations, the COLREG prescribe that any change in course angle should normally be towards starboard, while in overtaking situations, both port and starboard manoeuvres are allowed. In this section's plots, a negative change in course angle signifies a starboard turn, while positive values indicate port turns.

The distribution of course angle changes in overtaking situations is shown in Figure 13. In 102 (45%) of the 230 situations, the give-way vessel make a starboard turn, and in 125 (55%), a port turn. In absolute values, the change in course angle ranges from 0 to 42°, with 76% of the manoeuvres in the 5 to 25° range.

In head-on situations, the COLREG require that changes in course angle should be to starboard, i.e. negative. The distribution of course angle changes for this situation type is shown in Figure 14 and

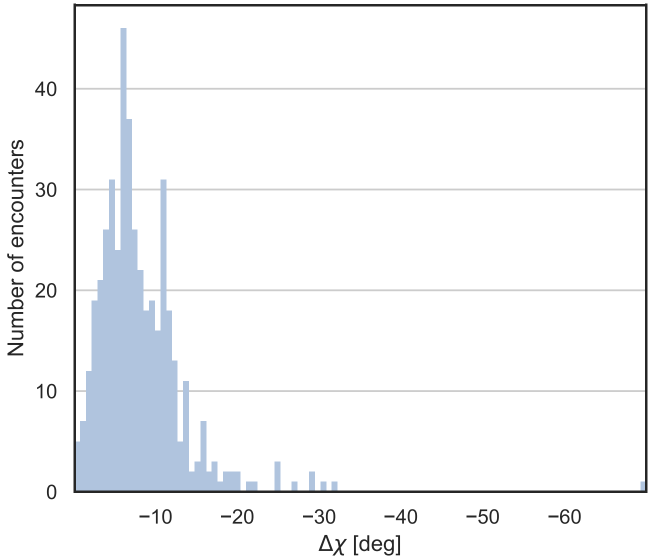

Figure 14. Distribution of change in course angle in 442 head-on situations.

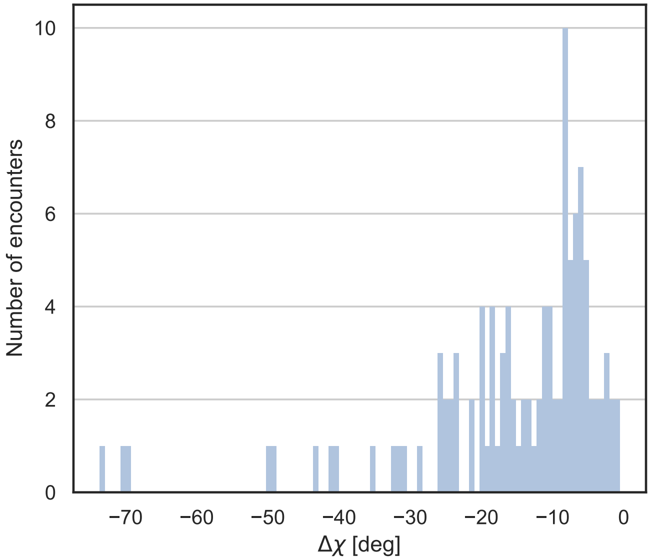

Figure 15. Distribution of change in course angle by give-way vessels in 110 crossing situations.

ranges from 0 to −70°. For this situation type, 99% of the manoeuvres are below 30° and 66% are in the range 5 to 15°.

The course angle changes in give-way manoeuvres in crossing situations, which should also be negative, fall within the range −1 to −74° Their distribution is shown in Figure 15, where 76% of the manoeuvres are within the −5 to −30° range.

The means of the course angle changes in the different situations according to land coverage (see Table 4) show that in head-on and crossing situations, there seems to be a tendency towards smaller course angle changes with increased land coverage. However, for overtaking situations there is no obvious correlation between course angle change and land coverage.

Table 4. *Mean course angle change according to situation and land coverage. For overtaking situations, the mean is shown for all manoeuvres (*^{abs}*), starboard turns only*$^{+}$ *and port turns only*$^{-}$*.*

	Mean course angle change [deg.]			
	Land coverage			
Situation type	0–9%	10–29%	30–59%	60–100%
Head-on	−9 · 6°	−8 · 0°	−7 · 5°	−6 · 1°
Crossing	−16 · 2°	−11 · 4°	−10 · 6°	–
Overtaking^{+}	9 · 8°	8 · 9°	10 · 3°	10 · 3°
Overtaking^{-}	−14 · 8°	−8 · 3°	−12 · 4°	−12 · 4°
Overtaking^{abs}	11 · 3°	8 · 6°	11 · 6°	11 · 6°

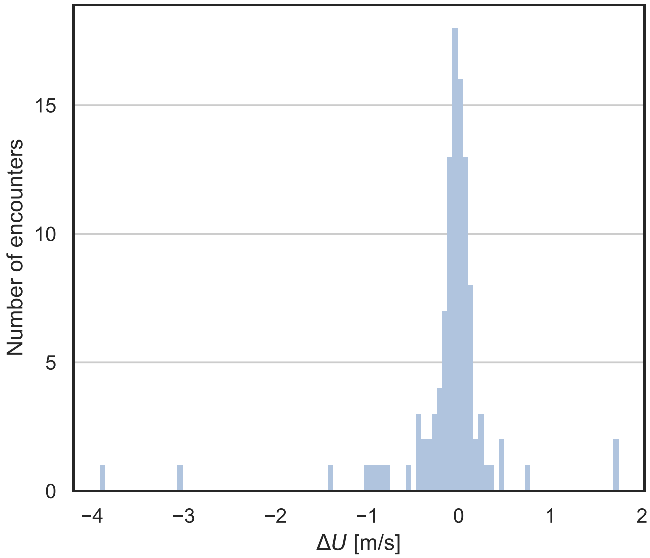

Figure 16. *Distribution of speed changes for crossing situations, excluding situations with no changes in speed.*

5.6. Speed changes

If changing the course angle to avoid collision is not practicable, the speed can be changed. This type of action is most suitable for crossing situations. However, as shown in Figure 16, the speed changes recorded in the dataset are, in general, too small to have a significant influence on the situation. In total, only 27 of the 782 situations have speed changes larger than 1 metre per second (approximately 2 knots), and it is evident that changing course is the preferred action in collision avoidance situations.

5.7. Traffic density

The effect of traffic density in the area with regards to the different parameters was also investigated. As an example, DCPA is plotted against the number of vessels in the encounter's originating case in Figure 17. As expected, a low number of vessels in the area allows for a large DCPA, while more traffic tends to restrict the DCPA to lower values. There is also a tendency towards higher traffic in areas with

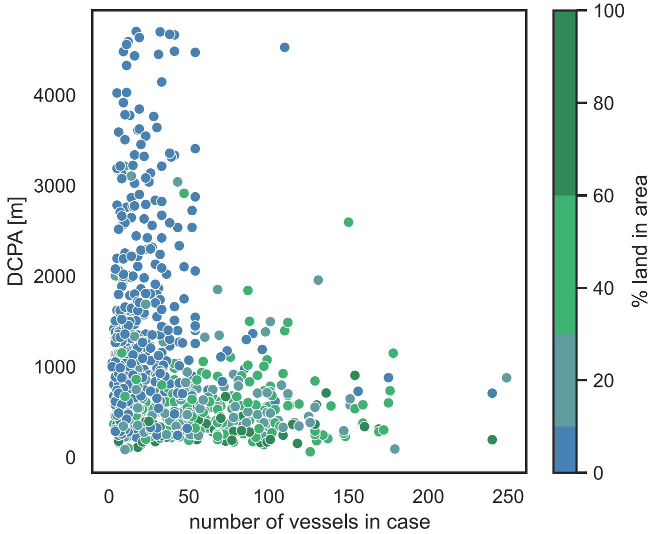

Figure 17. *DCPA versus number of vessels in originating case.*

more land coverage. This may be somewhat misleading with regards to the DCPA, as some of this traffic may be constrained by land, thereby not having any influence on the vessels' behaviour.

6. Discussion

Encounters between vessels are relatively rare and their behaviour may be directed by other factors than the COLREG. Filtering out irrelevant data and identifying encounters where the vessels behave according to the COLREG is therefore an important task in this work. While the drastic reduction in the number of situations after filtering may seem excessive, it is necessary to obtain reliable information about COLREG compliant behaviour. Nevertheless, obtaining a larger selection of encounters that can produce more statistically significant results is an important task that remains for future work.

The plots presented in Section 5 show that vessel behaviour is correlated with the percentage of land coverage in the area, but a more accurate method for classifying the area type may improve our understanding on this point. For instance, it is likely that vessels travelling in a fjord or narrow strait, where vessels often move in established lanes, will display behaviours different to those of vessels travelling in an archipelago where the vessels are restricted by multiple shallows and islands.

The presented results give an indication of customary behaviour in collision avoidance situations, which may be improved if even larger datasets are studied, as further division of the data according to vessel type and size may reveal hereto undiscovered correlations. Further work is also required to identify correlations between other parameters that may have an impact on vessel behaviour. This also applies to environmental influences such as wind, visibility and sea-state, which have not been considered in this work.

7. Conclusion

A procedure for identifying vessel encounters containing collision avoidance situations from recorded AIS data has been presented, along with a method for extracting parameters characterising the evasive manoeuvres performed. The technique was tested on AIS data gathered from vessels in three areas off the Norwegian coast over a period of several years, resulting in a dataset containing information on how the COLREG are currently practised. The data are presented graphically, showing the distributions of, and relationships between, relevant parameters such as type of situation, DCPA, TCPA when manoeuvres are initiated, magnitude of manoeuvres and speed changes, and how they correlate with the amount of land in the area and the ship's length and average speed. The data do not recommend specific values for use in methods for collision avoidance or the evaluation of these. However, for several parameters, the range of recorded values along with their average do point to what should be considered as safe and acceptable values, depending on the type of situation and the amount of land in the area. This contributes to a better understanding of what factors should be considered by collision avoidance algorithms and how these can be verified for use in autonomous vessels.

The natural continuation of this work is to obtain a larger dataset providing more statistically significant results, and further the investigation into determining factors for vessel behaviour. Moreover, multi-vessel situations where a ship is both stand-on and give-way have been excluded from the dataset. However, these are important data to gain experience from for autonomous ships, so they should be considered in future work.

Acknowledgements. We would like to thank The Norwegian Coastal Administration (Kystverket) for providing the data that form the foundation for this study. We would also like to thank Olav B. Vassbotn for his contributions to the pre-processing and data analysis software used in this paper. We are grateful for comments provided by employees at Kongsberg Maritime.

Financial statement. This research was primarily funded by Kongsberg Maritime. It was also funded by the Research Council of Norway grant 223254 to the Center of Excellence on Autonomous Marine Operations and Systems (NTNU-AMOS).

References

Basseville, M. and Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Application*. Vol. **104**. Englewood Cliffs: Prentice-Hall.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-8**(6), 679–698. doi: 10.1109/TPAMI.1986.4767851

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In: *KDD'96: Proceedings of the Second ACM International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon, 226–231.

Fujii, Y. and Tanaka, K. (1971). Traffic capacity. *Journal of Navigation*, **24**(4), 543–552. doi: 10.1017/S0373463300022384

Goodwin, E. M. (1975). A statistical study of ship domains. *Journal of Navigation*, **28**(3), 328–344. doi: 10.1017/S0373463300041230

Gucma, L. and Marcjan, K. (2012). Examination of ships passing distances distribution in the coastal waters in order to build a ship probabilistic domain. *Scientific Zeszyty Naukowe Akademii Morskiej w Szczecinie*, **32**(104), 34–40.

Hagen, I. B., Kufoalor, D. K. M., Brekke, E. F. and Johansen, T. A. (2018). MPC-based collision avoidance strategy for existing marine vessel guidance systems. In: *IEEE International Conference on Robotics and Automation*, Brisbane, QLD.

Hagen, I. B., Vassbotn, O., Skogvold, M., Johansen, T. A. and Brekke, E. (2023). Safety and COLREGS evaluation for marine collision avoidance algorithms. *Ocean Engineering*, **288**, 115991.

Hansen, M. G., Jensen, T. K., Lehn-Schiøler, T., Melchild, K., Rasmussen, F. G. and Ennemark, F. (2013). Empirical ship domain based on AIS data. *Journal of Navigation*, **66**(6), 931–940. doi: 10.1017/S0373463313000489

International Maritime Organization (IMO) (1972). Colregs - international regulations for preventing collisions at sea.

International Maritime Organization (2015). *Resolution a.1106(29), revised guidelines for the onboard operational use of shipborne automatic identification systems (AIS)*. Technical report, International Maritime Organization. https://edocs.imo.org/Final

Kim, H.-G., Yun, S.-J., Choi, Y.-H., Ryu, J.-K. and Suh, J.-H. (2021). Collision avoidance algorithm based on COLREGS for unmanned surface vehicle. *Journal of Marine Science and Engineering*, **9**(8), 863. doi: 10.3390/jmse9080863

Knudsen, K. S. (2022). *Analysis of AIS data for COLREGS compliance*. M.Sc. thesis, Norwegian University of Science and Technology, Trondheim, Norway.

Marine Traffic (2022). *Marine traffic*. https://www.marinetraffic.com

Pedersen, T. A., Glomsrud, J. A. and Haugen, O. I. (2019). Towards simulation-based verification of autonomous navigation systems. In: *Proceedings of the International Seminar on Safety and Security of Autonomous Vessels (ISSAV) and European STAMP Workshop and Conference (ESWC) 2019*, Aalto. doi: 10.2478/9788395669606-00

Pietrzykowski, Z. and Magaj, J. (2017). Ship domain as a safety criterion in a precautionary area of traffic separation scheme. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, **11**(1). doi: 10.12716/1001.11.01.10

Procee, S., Borst, C., van Paassen, R., Mulder, M. and Bertram, V. (2018). Using augmented reality to improve collision avoidance and resolution. In: *Proceedings of the 17th International Conference on Computer and IT Applications in the Maritime Industries*, Pavone, Italy.

Torben, T. R., Glomsrud, J. A., Pedersen, T. A., Utne, I. B. and Sørensen, A. J. (2022). Automatic simulation-based testing of autonomous ships using gaussian processes and temporal logic. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 1748006X2110692. doi: 10.1177/1748006X211069277

Vagale, A., Oucheikh, R., Bye, R. T., Osen, O. L. and Fossen, T. I. (2021). Path planning and collision avoidance for autonomous surface vehicles I: a review. *Journal of Marine Science and Technology*, **26**(4), 1292–1306. doi: 10.1007/s00773-020-00787-6

Vassbotn, O. B.. 2022). *Analysis of ship collision risk encounters and COLREG behaviours using machine learning and AIS data*. M.Sc. thesis, Norwegian University of Science and Technology, Trondheim, Norway.

Woerner, K., Benjamin, M. R., Novitzky, M. and Leonard, J. J. (2019). Quantifying protocol evaluation for autonomous collision avoidance: toward establishing COLREGS compliance metrics. *Autonomous Robots*, **43**(4), 967–991. doi: 10.1007/s10514–018-9765-y

Zhao, L. and Roh, M.-I. (2019). COLREGs–compliant multiship collision avoidance based on deep reinforcement learning. *Ocean Engineering*, **191**, 106436. doi: 10.1016/J.OCEANENG.2019.106436

Cite this article: Hagen I.B, Knutsen K.S, Johansen T.A, Brekke E (2024). Exploration of COLREG-relevant parameters from historical AIS-data. *The Journal of Navigation* 1–19. https://doi.org/10.1017/S0373463324000109