Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira, Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria
 Published by Research Publishing, Singapore.
doi: 10.3850/ESREL2026061419_esrel26-p26517-cd

Ship collision avoidance behaviour identification and analysis method

Boyuan Zhang

Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK. E-mail: b.zhang@2024.ljmu.ac.uk

Hao Rong

Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK. E-mail: h.rong@ljmu.ac.uk

Zaili Yang

Liverpool Logistics, Offshore and Marine (LOOM) Research Institute, Liverpool John Moores University, Liverpool, UK. E-mail: z.yang@ljmu.ac.uk

In the field of intelligent ships, enabling ships to navigate with human-like decision-making capabilities has long been a research focus. However, few studies have objectively demonstrated or systematically mined the existence of ship collision avoidance behaviour (CAB) from real-world data. This study proposes a spatial-temporal characterisation framework for ship CAB based on extracting and analysing near-collision scenarios from historical Automatic Identification System (AIS) data. Firstly, collision avoidance manoeuvres are identified using a Change Point Detection (CPD) technique, in which a cost function is formulated to distinguish the manoeuvring behaviour from normal sailing behaviour, and critical encounter scenarios are extracted through encounter situation analysis. Second, contributions of collision avoidance characteristic indicators are determined through Principal Component Analysis (PCA), taking into account the requirements of the International Regulations for Preventing Collisions at Sea (COLREGs). The proposed framework is applied to the Øresund waterway between Denmark and Sweden, where the spatial–temporal distribution of critical encounters and behavioural patterns is examined. Results of the case study reveal the importance of relative bearing and phase in determining the timing of collision avoidance for both give-way and stand-on ships. This study provides a novel perspective for understanding ship navigation behaviour, enhancing situational awareness, and supporting the quantitative characterisation of collision risk.

Keywords: Ship Collision Avoidance Behaviour, Change Point Detection, Principal Component Analysis

1. Introduction

Ship collisions are a typical marine accident that results in significant casualties, substantial economic losses, and significant environmental pollution (Guedes Soares & Teixeira, 2001). To reduce ship collision accidents, the International Regulations for Preventing Collisions at Sea (COLREGs) stipulate when and how evasive actions should be taken under different encounter scenarios and collision avoidance responsibilities (IMO, 1972). However, some key rules of COLREGs are formulated in qualitative, non-deterministic language, which makes it difficult for human navigators to directly apply them to make collision avoidance decisions. Instead, navigators are required to interpret the COLREGs based on their professional knowledge and experience.

In simple encounter situations, competent navigators are generally able to formulate effective collision avoidance strategies. However, as navigators' mental workload increases with situational complexity (Montewka et al., 2022), accurate situation assessment and timely decision-making become significantly challenging in complex multi-ship encounters within restricted waterways. To address this challenge, a wide range of decision-support methods has been proposed to assist navigators in determining appropriate collision avoidance timing.

Distance at Closest Point of Approach (DCPA) and Time to Closest Point of Approach (TCPA) are widely used indicators for supporting navigators in determining collision avoidance timing. To jointly consider the DCPA and TCPA, the Collision Risk Index (CRI) was first proposed by Kearon (1977), which is a polynomial equation involving DCPA and TCPA. Zhao et al. (2016) considered DCPA, TCPA, relative distance, relative bearing, and relative speed, and then employed ER theory to assess the CRI to support a collision avoidance system. Hu et al. (2020) developed a CRI-based collision avoidance method, demonstrating that the timing of evasive actions is closely related to the CRI value. Furthermore, through a series of comparative experiments, they further identified a recommended range of CRI thresholds for initiating collision avoidance actions.

Except for CRI, many other methods have been proposed to support the timing of initiating evasive actions. Gil et al. (2022) determined the empirical values of Bow Crossing Range (BRC) during routine operations of merchant ships in ship encounters based on a ten-year AIS dataset. Zhang et al. (2015) developed a Vessel Conflict Ranking Operator (VCRO) method to evaluate near-miss ship collision risk, which considers the relative distance, phase, and speed. To assess the near-miss collision risk more accurately, Liu et al. (2023) proposed a Kinematics Feature-based Vessel Conflict Ranking Operator (KF-VCRO) method, which introduces ship specifications to the VCRO. Wang (2010) established a Fuzzy Quaternion Ship Domain (FQSD) incorporating ship length, speed, manoeuvrability, and COLREGs, thereby further estimating the spatial collision risk. Montewka et al. (2010) proposed a Minimal Distance to Collision (MDTC), which represents the last moment for taking avoidance actions under current manoeuvrability and state.

Unlike existing collision avoidance timing determination studies that are based on the risk or conflict indices, this study focuses on analysing which indicators dominate the initiation of observed evasive actions in real AIS-based encounters. The main contribution is an integrated framework that first identifies collision avoidance manoeuvres through CPD, then extracts critical encounter scenarios by the intention of collision avoidance, and finally analyses the dominant indicators at the moment of action initiation using PCA.

This study first identifies manoeuvring behaviours and then extracts the corresponding encounter scenarios. Subsequently, Principal Component Analysis (PCA) is applied to identify the key indicators of collision avoidance characteristics, thereby supporting the development of effective decision support methods. The remainder of this paper is organised as follows. In Section 2, the methodology is developed and introduced, including ship behaviour identification, encounter extraction, and indicator contribution analysis. Then, the results analysis is presented in Section 3. Section 4 concludes this paper.

2. Methodology

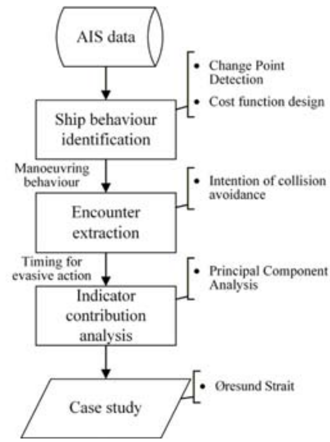

Fig. 1. Framework of collision avoidance behaviour identification and analysis method.

Collision avoidance characteristic indicators are essential for representing encounter situations and determining the timing of evasive actions. To quantify the contributions of these indicators to actual collision avoidance decisions, this study proposes a new methodology that consists of three parts: ship behaviour identification, encounter extraction, and indicator contribution analysis (Fig. 1). First, a Change Point Detection (CPD) method is employed to identify manoeuvring behaviours. Then, an encounter scenario extraction method is proposed to determine the timing of collision avoidance. Finally, a Principal Component Analysis (PCA) is applied to evaluate the relative importance of collision avoidance characteristic indicators for determining the collision avoidance timing. The considered indicators include commonly used ship relative features (e.g., DCPA, TCPA, relative distance, bearing, course, speed, and phase), as well as evasive action-related indicators, such as the maximum course alteration and the minimum ship-to-ship distance during avoidance.

2.1. Ship Behaviour Identification

In signal processing, CPD is the task of identifying changes in the underlying model of a signal or time series (Truong et al., 2020). The aim of CPD methods is to determine the optimal change points to minimise the cost. Subsequently, establishing an appropriate cost function is the most important step in the CPD method.

The overall cost function is shown in Eq. (1).

$$F(k,h \mid \mathcal{T},\mathcal{R}) = \lambda(m)\sum_{j=1}^{m-1} C_{h_j}\left(S_j,\mathcal{R}_j\right) \tag{1}$$

where k denotes the ordered set of the behavioural change points, $\mathcal{T}$ and $\mathcal{R}$ denote the trajectory and planned route, $C_{\text{Sail}}(\cdot)$ and $C_{\text{Mano}}(\cdot)$ are the sailing/manoeuvring segment costs, λ is a change point penalty, S_j and $\mathcal{R}_j$ are the trajectory and planned route segments in between change points k_j and k_{j+1}.

Given that collision avoidance behaviours exhibit a deviation–recovery evolution pattern (Fig. 2), this evolution pattern is considered in $C_{\text{Sail}}(\cdot)$ and $C_{\text{Mano}}(\cdot)$, instead of short-term variations in kinematic features, which are sensitive to noise. The $C_{\text{Mano}}(\cdot)$ is expressed in Eq. (2).

$$\begin{cases} C_{\text{Mano}}(S,\mathcal{R}) = C_d + C_\theta \\ C_d = C_{\text{d}}^{\text{Peak}} + C_{\text{d}}^{\text{Amp}} + C_{\text{d}}^{\text{Prop}} + g(S,\mathcal{R}) \\ C_\theta = C_\theta^{\text{Peak}} + C_\theta^{\text{Amp}} + C_\theta^{\text{Prop}} + C_\theta^{\text{ROT}} \end{cases} \tag{2}$$

where C_d and C_θ represent the cost in distance and relative course.

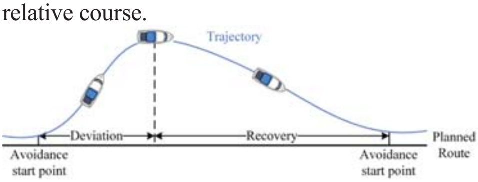

Fig. 2. Evolution pattern in collision avoidance behaviour.

2.1.1. Peak Cost

COLREGs state that the evasive action should be large enough to be readily apparent to another ship. Therefore, the peak values in the deviation and recovery stages are introduced to the cost function as the extent of evasive action, as shown in Eq. (3).

$$\begin{cases} C_{\mathcal{X}}^{\text{Peak}} = f_H\left(P_{\mathcal{X}},\tau_{\mathcal{X}}^{\text{Peak}}\right), \mathcal{X} \in \{d,\theta\} \\ f_H(a,\tau) = \begin{cases} \dfrac{1}{a+\varepsilon}, 0 \le a \le \tau \\ 0, \quad a > \tau \end{cases} \end{cases} \tag{3}$$

where $C_{\mathcal{X}}^{\text{Peak}}$ represents the peak cost associated with feature $\mathcal{X}$, $P_{\mathcal{X}}$ denotes the minimum peak of feature $\mathcal{X}$ in deviation and recovery stages, $\tau_{\mathcal{X}}^{\text{Peak}}$ is the threshold, $f_H(\cdot)$ denotes the inverse hinge function.

2.1.2. Amplitude Cost

Due to environmental constraints, ships may sail parallel to the planned route, resulting in a large peak in distance. Consequently, the amplitude of evolution patterns is introduced to represent the change in kinematic features, as expressed in Eq. (4).

$$C_{\mathcal{X}}^{\text{Amp}} = f_H\left(A_{\mathcal{X}},\tau_{\mathcal{X}}^{\text{Amp}}\right) \tag{4}$$

where $C_{\mathcal{X}}^{\text{Amp}}$ represents the amplitude cost associated with feature $\mathcal{X}$, $A_{\mathcal{X}}$ denotes the minimum amplitude of feature $\mathcal{X}$ in deviation and recovery stages, $\tau_{\mathcal{X}}^{\text{Amp}}$ represents the threshold.

2.1.3. Proportion Cost

Compared to stable route-following behaviour, collision avoidance behaviour exhibits a clear deviation-recovery pattern, with significant changes in kinematic features. Therefore, the proportion of kinematic variations is introduced to capture this difference, as shown in Eq. (5).

$$C_{\mathcal{X}}^{\text{Prop}} = f_H\left(Prop_{\mathcal{X}},\tau_{\mathcal{X}}^{\text{Prop}}\right) \tag{5}$$

where $C_{\mathcal{X}}^{\text{Prop}}$ represents the proportion cost associated with feature $\mathcal{X}$, $Prop_{\mathcal{X}}$ denotes the minimum proportion of variation of feature $\mathcal{X}$ in deviation and recovery stages, $\tau_{\mathcal{X}}^{\text{Prop}}$ is the threshold.

2.1.4. Rate of Turn Cost

The Rate of Turn (ROT) reflects the rate of change in ship course. Since COLREGs require the ship to avoid a succession of small alterations of course in collision avoidance, the average of ROT is considered, as shown in Eq. (6).

$$C_\theta^{\text{ROT}} = f_H\left(\overline{ROT},\tau_{ROT}\right) \tag{6}$$

where $\overline{ROT}$ denotes the average of $|ROT|$, τ_{ROT} is the threshold.

2.1.5. Sailing Penalty

To accurately identify the change points between sailing and manoeuvring behaviours, a sailing penalty is introduced to reduce the misclassification of sailing behaviours as manoeuvring behaviours before and after manoeuvring. The sailing penalty is shown in Eq. (7).

$$g(S,\mathcal{R}) = \sum_{j=1}^{0.2n-1} f_H(d_{j+1} - d_j, 0) + \sum_{j=0.8n}^{n-1} f_H(d_{j+1} - d_j, 0) \quad (7)$$

where $g(\cdot)$ denotes the sailing penalty, d is the distance between the trajectory segment S and planned route $\mathcal{R}$, n is the length of S.

2.1.6. Change Point Penalty

Since the number of change points is unknown, the optimal segmentation is obtained by searching for the set of change points that minimises the total cost under different numbers of change points. To prevent the number of change points from increasing indefinitely, a change point penalty is introduced in Eq. (8).

$$\lambda(m) = 1 + m\tau_{Decline} \quad (8)$$

where m indicates the number of change points, $\tau_{Decline}$ is the threshold.

2.1.7. Sailing Cost

The most typical pattern of sailing behaviour is to sail along the planned route. In an ideal condition, when a ship sails along the planned route, the distance and course relative to the planned route are near zero. Therefore, the sailing cost is shown as Eq. (9).

$$C_{\text{Sail}}(S,\mathcal{R}) = \sum_{j=1}^{n} |d_j| + |\theta_j| \quad (9)$$

where the C_{sail} represents the sailing cost function, θ denotes the course relative to the planned route.

Based on the sailing and manoeuvring cost functions, dynamic programming (DP) is employed to search for the optimal change points that minimise the total segmentation cost for a given number of change points, n. The number of change points is then increased iteratively until the total cost no longer decreases.

2.2. Ship Encounter Extraction

When a ship takes evasive action, there are often multiple other ships surrounding it. To extract encounters involving ship-to-ship interactions, a new method is proposed. The key idea of this method is to compare the real trajectory with a hypothetical route-following trajectory, to examine whether the manoeuvring action is motivated by collision avoidance and which ships are avoided by the evasive action. By analysing differences in ship interactions between the real and reference trajectories, the avoidance intention can be inferred, and the corresponding critical encounter scenarios can be accurately identified. Consequently, crossing encounters are further subdivided into four types by comparing the passing condition and distance in the reference trajectory and real manoeuvring trajectory, as shown in Table 1.

Table 1. Subdivided crossing encounter types.

Crossing encounter types	Reference trajectory		Real manoeuvring trajectory	
	Con	Dis	Con	Dis
$Cross_{bow}$	Bow	d	Bow	$\geq \tau d$
$Cross_{bow2stern}$	Bow	-	Stern	-
$Cross_{stern}$	Stern	d	Stern	$\geq \tau d$
$Cross_{Stern2bow}$	Stern	-	Bow	-

(Con and Dis denote the passing condition and relative distance between two ships, d is the passing distance between ships, - is not applicable, τ is the threshold)

2.3. PCA-based Indicator Contribution Analysis

Evaluating the relative contributions of collision avoidance characteristic indicators at the moment of avoidance action initiation is essential for understanding the dominant factors underlying collision avoidance behaviour. To achieve this, PCA is employed to quantify the contribution of different indicators when evasive actions are initiated.

2.3.1. Collision Avoidance Characteristic Indicators

Closest Point of Approach (CPA): CPA represents the point at which the distance between own ship and target ship is minimal under the assumption that both own ship and target ship maintain their speed and course. The distance at CPA is DCPA, while the time to CPA is TCPA. Let points $p_o = (lon_o, lat_o, v_o, \psi_o)$ and $p_t = (lon_t, lat_t, v_t, \psi_t)$ represent the position of own ship and target ship. The relative distance and relative speed are expressed as Eqs. (10-14).

$$D_r = 2R_e \sin^{-1}\left(\sqrt{\sin^2\left(\frac{lat_o - lat_t}{2}\right) + \cos lat_o \cos lat_t \sin^2\left(\frac{lon_o - lon_t}{2}\right)}\right) \quad (10)$$

$$v_r = \sqrt{(v_o \cos\psi_o - v_t \cos\psi_t)^2 + (v_o \sin\psi_o - v_t \sin\psi_t)^2} \quad (11)$$

where D_r represents the relative distance between own and target ships, v_r denotes the relative velocity, R_e is the radius of Earth.

$$\psi_r = \cos^{-1}\left(\frac{v_o - v_t \cos(\psi_o - \psi_t)}{v_r}\right) \quad (12)$$

$$DCPA = D_r \sin(\psi_r - \alpha_t - \pi) \quad (13)$$

$$TCPA = \frac{D_r \cos(\psi_r - \alpha_t - \pi)}{v_r} \quad (14)$$

where ψ_r denotes the relative course, α_t represents the bearing of the target ship relative to the own ship.

Relative bearing: relative bearing refers to the angle between the own ship course and the location of target ship. According to the COLREGs, the encounter types are partially determined based on relative bearing, influencing the timing and measure of collision avoidance.

Relative velocity: refers to the velocity vector of the target ship with respect to the own ship, including both its magnitude (relative speed) and direction (relative course). The relative speed reflects the potential severity of collision consequences, while the relative course reflects the possible pattern of relative motion.

Relative distance: refers to the distance between own ship and target ship. It is one of the most widely used indicators in collision avoidance decision-making, as many methods trigger evasive actions when the distance falls below a certain threshold.

Phase: refers to the angular difference between the courses of the own ship and the target ship, with a range of $[-\pi, \pi]$ (see (Zhang et al., 2015)). A positive value indicates that the ships are approaching, whereas a negative value denotes that the ships are receding from each other.

Magnitude of evasive action: characterises the extent of manoeuvring action during collision avoidance. Collision avoidance decisions consist of two coupled components: the timing and the magnitude of evasive actions. To account for this coupling, two indicators are considered: the maximum deviation course and the minimum distance between ships. The maximum deviation course represents the largest course alteration relative to the planned route, reflecting the intensity of manoeuvring, while the minimum distance describes the achieved safety distance during avoidance. The maximum deviation course and minimum distance are expressed in Eqs. (15) and (16).

$$\theta_{max} = \max(\theta_{dev}) \tag{15}$$

$$d_{min} = \min(D_r) \tag{16}$$

where θ_{max} is the maximum deviation course, θ_{dev} represents the course relative to the planned route in the deviation stage, d_{min} denotes the minimum distance.

2.3.2 Principal Component Analysis

PCA is a classical multivariate statistical technique for dimensionality reduction and feature importance analysis. Its objective is to transform a set of correlated variables into a new set of orthogonal variables, referred to as principal components, which successively maximise the variance of the data.

Let $\boldsymbol{M} \in \mathbb{R}^{m \times n}$ denote a feature matrix with m samples and n features. The matrix is first standardised to remove scale effects by Eq. (17).

$$\widehat{\boldsymbol{M}}_j = \frac{(\boldsymbol{M}_j - \mu_j)}{\sigma_j} \tag{17}$$

where $\boldsymbol{M}_j$ denotes the matrix of j^{th} feature, $\widehat{\boldsymbol{M}}_j$ is its standardised form, μ_j and σ_j are the corresponding mean and standard deviation.

The covariance matrix $\boldsymbol{C}$ of the standardised $\widehat{\boldsymbol{M}}$ and its eigenvalues λ_i and eigenvectors $\mathrm{v_i}$ are then obtained by Eqs. (18) and (19)

$$\boldsymbol{C} = \frac{1}{m-1} \widehat{\boldsymbol{M}} \widehat{\boldsymbol{M}}^T \tag{18}$$

$$\boldsymbol{C}\mathbf{v_i} = \lambda_i \mathbf{v_i} \tag{19}$$

Each eigenvector defines a principal component direction, while the corresponding eigenvalue represents the amount of variance explained. The principal components are ranked by descending eigenvalues, and the first k components explaining most of the variance are retained. The contribution of feature j is computed as the squared loading normalised by the sum of squared loadings over all features, providing an objective measure of the relative importance of different indicators.

3. Case studies and Results

The case AIS data cover the Øresund Strait, including the Helsingør–Helsingborg (HH) ferry route connecting Helsingør, Denmark, and Helsingborg, Sweden. They are collected from 1 January 2024 to 31 December 2024. Before analysis, a pre-processing procedure is applied to ensure data quality. First, trajectories are segmented using a 900s time-gap threshold. Then, each segment is linearly interpolated at 10 s. Through the pre-processing, 44530 validated segments across 7799 ships are obtained. Owing to the dense traffic and frequent ferry crossings in the study area, the dataset provides a representative test bed for restricted-water ship encounters, although it does not cover all navigational environments.

3.1. Ship Behaviour Identification Results

To determine the thresholds in the cost function, over 1200 trajectories are randomly sampled and manually labelled to evaluate the performance of different threshold sets. However, exhaustively exploring all combinations of the eight threshold parameters results in prohibitively high computational complexity. Therefore, an orthogonal experimental design is adopted to systematically explore the parameter space.

The thresholds are selected by jointly evaluating representative parameter sets using accuracy, precision, recall, and F1-score. The adopted set achieves the highest recall (0.8253) and F1-score (0.8235), while maintaining high accuracy (0.9356) and precision (0.8217). This result demonstrates that the joint cost function defined in Eqs. (1)-(9) effectively distinguishes manoeuvres from route-following behaviours by capturing their distinct behavioural patterns. The specific parameter settings are shown in Table 2.

Table 2. Parameter settings.

Parameter	Value	Parameter	Value
$\tau_{dist}^{\mathrm{Peak}}$	1.000	$\tau_{\Delta COG}^{\mathrm{Peak}}$	10°/180°
$\tau_{dist}^{\mathrm{Amp}}$	0.600	$\tau_{\Delta COG}^{\mathrm{Amp}}$	0.200
$\tau_{dist}^{\mathrm{Prop}}$	0.200	$\tau_{\Delta COG}^{\mathrm{Prop}}$	0.200
τ_{ROT}	10°/min	$\tau_{Decline}$	0.100

3.2. Ship Encounter Extraction Results

The distribution of encounter types extracted using the proposed method is summarised in Table 3. Routes 1 and 2 represent the ferry route from Helsingborg to Helsingør and Helsingør to Helsingborg, respectively. The extracted encounters are assumed to be in good visibility conditions, and therefore Rule 19 does not apply. Furthermore, this study focuses on encounters involving cargo and passenger ships, both of which are power-driven vessels with the same priority level under Rule 18. Consequently, the analysis of extracted encounters focuses on Rules 13–15.

Table 3. Distribution of encounter types.

Encounter types	Route 1	Route 2
$Cross_{bow}$	384 (9.98%)	131 (5.76%)
$Cross_{bow2stern}$	787 (20.45%)	815 (35.81%)
$Cross_{stern}$	830 (21.56%)	921 (40.47%)
$Cross_{Stern2bow}$	16 (0.42%)	11 (0.48%)
Head-on	1647 (42.79%)	318 (13.97%)
Overtaking	93 (2.42%)	46 (2.02%)
Overtaken	92 (2.39%)	34 (1.49%)

In the crossing encounter types, $Cross_{stern}$ and $Cross_{bow2stern}$, which correspond to give-way ships passing astern of the target ships, account for 1617 and 1736 cases on Routes 1 and 2, respectively. The number of these encounter types is significantly higher than that of bow-passing encounters, which are 400 and 142, respectively.

This observed distribution indicates a clear preference for passing astern of the target ships, consistent with the requirements of Rule 15 of the COLREGs. The consistency between the extracted encounter patterns and COLREGs demonstrates the effectiveness of the proposed encounter extraction method.

3.3. Indicator Contribution Results

According to the COLREGs, give-way ships are required to take early and substantial action, whereas stand-on ships are expected to maintain their course and speed and only take evasive action when a collision cannot be avoided by the action of the give-way vessel alone. Given this fundamental difference in the timing of initiating collision avoidance manoeuvres, the contributions of collision avoidance characteristic indicators are analysed separately for give-way and stand-on ships.

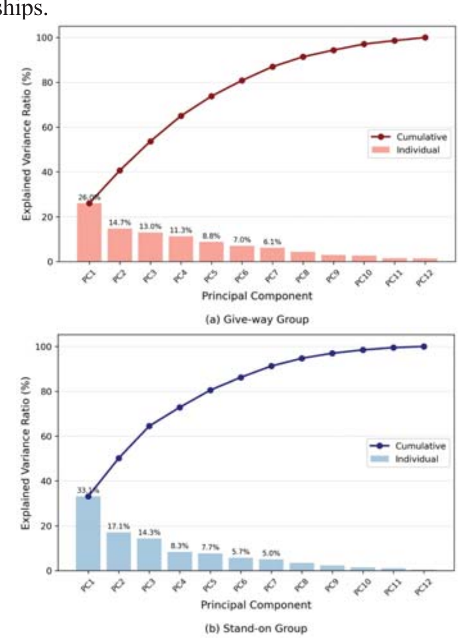

Fig. 3. Explained variance ratio of principal components.

Fig. 3 shows that the first six principal components account for more than 80% of the total variance for both give-way and stand-on ships, indicating that collision avoidance is difficult to characterise by a single or two indicators, but rather by a series of indicators.

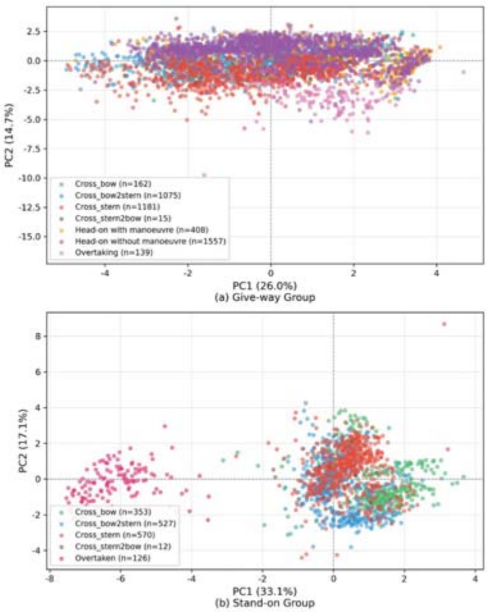

Fig. 4. PCA of collision avoidance indicators for give-way and stand-on ships.

Fig. 4 illustrates the projection of collision avoidance characteristic indicators onto the first two principal components (PC1 and PC2). Each point represents a set of indicators at the moment of initiating evasive action in an extracted encounter scenario, coloured by encounter type. Encounters of the same type exhibit a noticeable clustering tendency, indicating that the PC1–PC2 space captures consistent characteristics within each encounter type.

In the give-way group, different encounter types exhibit a relatively compact and overlapping distribution in the PC1–PC2 space. This suggests that, despite differences in encounter situations, the dominant factors influencing the timing of collision avoidance actions for give-way ships do not differ significantly across encounter types. From a regulatory perspective, this observation is consistent with the COLREGs, which require give-way ships to take early and substantial action, regardless of specific encounter configuration. As a result, the collision avoidance characteristic indicators appear to be more homogeneous for give-way ships.

In contrast, for the stand-on group, overtaken encounters form a clearly separated cluster from crossing encounters in the PC1–PC2 space. This indicates that the factors influencing the timing of avoidance actions for stand-on ships differ substantially between overtaken and crossing situations.

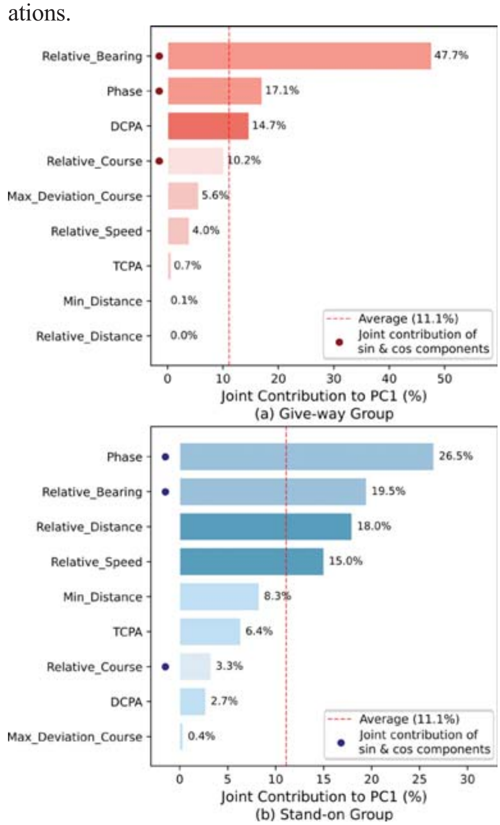

Fig. 5. Contribution of collision avoidance characteristic indicators for give-way and stand-on ships.

Fig. 5 illustrates the joint contribution of collision avoidance characteristic indicators to the first principal component, which represents the direction of maximum variance obtained by projecting the original variables onto the leading eigenvector of the covariance matrix. Relative bearing exhibits the highest contribution in the give-way group (47.7%) and the second-highest contribution in the stand-on group (19.5%), while the phase indicator shows the opposite trend. This indicates that relative bearing and phase jointly dominate the principal variation associated with collision avoidance decision timing. Such a dominant pattern is consistent with the encounter classification in COLREGs, where encounter types are primarily determined by relative bearing. Moreover, the phase indicator essentially integrates relative bearing with the target ship's course, thereby capturing key aspects of encounter characteristics. The high contributions of these indicators suggest the importance of these encounter-related characteristics for both give-way and stand-on ships.

In addition, relative distance and relative speed present notable contributions in the stand-on group, accounting for 18.0% and 15.9%, respectively, while their contributions in the give-way group are negligible. This suggests that stand-on ships rely more on distance- and speed-related information to assess encounter evolution.

4. Conclusions

This study investigated the contribution of collision avoidance characteristic indicators to the timing of evasive action by integrating CPD-based ship behaviour identification, encounter extraction and PCA-based contribution analysis. Using AIS data from the Øresund Strait, a total of 6125 encounter scenarios were identified and analysed. The results show that relative bearing and phase exhibit dominant contributions for both give-way and stand-on ships, indicating that encounter geometry leads the principal variation associated with collision avoidance timing. Moreover, the distinct contribution patterns between give-way and stand-on ships reflect differences in their perceptual priorities.

Although PCA provides an effective means of quantifying the relative contributions of indicators, the underlying mechanisms behind these contribution patterns still require further investigation to improve interpretability. In addition, PCA is more suitable for continuous variables and has limited capability in analysing discrete indicators, such as ship domain violation. Since the focus of this study is to establish an integrated analysis framework, the comparison with existing methods remains limited. Moreover, the framework depends on the robustness of behaviour identification and encounter extraction, and its applicability in more heterogeneous real-world scenarios still requires further validation. Finally, due to the frequent ferry crossings in the study area, the results may partly reflect the navigational characteristics of this specific area.

Acknowledgement

This research was funded by a European Research Council project under the European Union's Horizon 2020 research and innovation programme (TRUST CoG 2019 864724).

References

Gil, M., Kozioł, P., Wróbel, K., & Montewka, J. (2022). Know your safety indicator–A determination of merchant vessels Bow Crossing Range based on big data analytics. *Reliability Engineering & System Safety, 220*, 108311.

Guedes Soares, C., & Teixeira, A. (2001). Risk assessment in maritime transportation. *Reliability Engineering & System Safety, 74*(3), 299–309.

Hu, Y., Zhang, A., Tian, W., Zhang, J., & Hou, Z. (2020). Multi-Ship Collision Avoidance Decision-Making Based on Collision Risk Index. *Journal of Marine Science and Engineering, 8*(9), 640. doi:10.3390/jmse8090640

IMO. (1972). Convention on the International Regulations for Preventing Collisions at Sea, 1972 (COLREGs). In: International Maritime Organization London, UK.

Kearon, J. (1977). *Computer program for collision avoidance and track keeping.* Paper presented at the Proc. Conf. Math. Aspects Mar. Traffic.

Liu, Z., Zhang, B., Zhang, M., Wang, H., & Fu, X. (2023). A quantitative method for the analysis of ship collision risk using AIS data. *Ocean Engineering, 272*, 113906. doi:10.1016/j.oceaneng.2023.113906

Montewka, J., Hinz, T., Kujala, P., & Matusiak, J. (2010). Probability modelling of vessel collisions. *Reliability Engineering & System Safety, 95*(5), 573–589. doi:10.1016/j.ress.2010.01.009

Montewka, J., Manderbacka, T., Ruponen, P., Tompuri, M., Gil, M., & Hirdaris, S. (2022). Accident susceptibility index for a passenger ship-a framework and case study. *Reliability Engineering & System Safety, 218*, 108145. doi:10.1016/j.ress.2021.108145

Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing, 167*. doi:10.1016/j.sigpro.2019.107299

Wang, N. (2010). An Intelligent Spatial Collision Risk Based on the Quaternion Ship Domain. *Journal of Navigation, 63*(4), 733–749. doi:10.1017/s0373463310000202

Zhang, W., Goerlandt, F., Montewka, J., & Kujala, P. (2015). A method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering, 107*, 60–69. doi:10.1016/j.oceaneng.2015.07.046

Zhao, Y., Li, W., & Shi, P. (2016). A real-time collision avoidance learning system for Unmanned Surface Vessels. *Neurocomputing, 182*, 255–266.