



Calculation of the number of ship collision candidates using mesh-based estimation method for ship traffic data

Sonoko Kawashima¹ · Hiroko Itoh¹ · Yasumi Kawamura²

Received: 28 September 2018 / Accepted: 28 August 2022 / Published online: 9 September 2022
© The Author(s) 2022

Abstract

Predicting and assessing the ship–ship collision possibility in waters are important for discussions on ship traffic safety. The geometric number of collision candidates is one of the most commonly used indexes for representing the frequency of multi-ship encounters that have potential collisions. It has often been estimated for ship traffic in a delimited area based on existing concepts. However, to discuss ship traffic safety in wide-congested waters where ships navigate along various courses and various ship encounters occur, a comprehensive understanding of potential collisions corresponding to all encounter characteristics, such as the encounter angle and location, is necessary. This paper proposes a method, called a “mesh-based estimation method,” to calculate the geometric number of collision candidates. It can deal with various encounter angles by introducing a computational mesh to comprehensively assess potential collisions in wide-congested waters. The validation of the method is conducted by applying it to simple ship traffic and comparing the calculated result with a result calculated based on a conventional approach. In addition, the method is applied to traffic data of AIS-equipped ships navigating in Tokyo Bay in Japan to show locations that have potential collisions based on the encounter angles.

Keywords Ship traffic safety · The geometric number of collision candidates · Mesh-based estimation method · Encounter angle · Congested waters

1 Introductions

To discuss ship traffic safety, it is important to consider measures to predict and assess the possibility of ship–ship collisions, which are frequent maritime accidents. To assess the frequency and probability of collisions, Fujii et al. [1, 2] and Macduff [3] proposed a collision frequency model during the 1970s, which is most commonly used and represented as the product of the geometric number of collision candidates and the collision causation probability. Thereafter, some researchers (e.g., Kristiansen [4] and Montewka [5]) estimated the collision frequency as the product of the geometric number of collision candidates if no evasive

maneuvers are made and the probability of failing to avoid a collision. According to the model, it is important to first decrease the geometric number of collision candidates that have potential collisions to reduce the possibility of ship–ship collisions.

The geometric number of collision candidates is estimated statistically and quantitatively based on traffic data, such as the traffic volume, length, and velocity of two ship groups on a collision course. Fujii et al. [1, 2, 6, 7] initially proposed the concept and formulations to estimate the geometric number of collision candidates based on the encounters, such as head-on and overtaking situations, without considering the precise encounter angle between two ship courses. Following this concept, Pedersen [8], Kaneko [9, 10], and Friis-Hansen [11] proposed new concepts and formulations for estimating the geometric number of collision candidates when considering the encounter angles.

The geometric number of collision candidates estimated by the formulations is an important index for investigating and understanding potential collisions. It has often been estimated for ship traffic in a delimited area, such as traffic separation routes and narrow waters, where some ship groups navigate

✉ Sonoko Kawashima
kawashima-s@m.mpat.go.jp

¹ National Maritime Research Institute (NMRI), National Institute of Maritime, Port and Aviation Technology (MPAT), 6-38-1 Shinkawa, Mitaka-shi, Tokyo 181-0004, Japan

² Yokohama National University, 79-5 Tokiwadai, Hodogaya-ku, Yokohama-shi, Kanagawa 240-8501, Japan

along relatively single courses (e.g., [9, 11–15]). By contrast, it was difficult to investigate potential collisions in wide-congested waters in which ship traffic is not separated, because the understanding of long-term ship behavior was a laborious task before the introduction of an automatic identification system (AIS). In coastal waters connecting to principal ports and traffic routes, however, ships navigate along various courses, and collisions occur, resulting in casualties (e.g., [16]). It has also become easier to observe ship behavior for the longer term and in a wider area owing to the recent introduction of an AIS, and ship behaviors can be comprehended even in smaller areas. Thus, ship traffic in waters is receiving increased attention for the prevention of ship potential collisions through safety measures [17–19].

Conventional formulations can be easily and freely applied to every encounter angle. However, the range of an assessed sea area needs to be delimited in conventional concepts, so that encounter angles and patterns need to be limited to the entire sea area. As encounter angles and navigating positions differ by location in wide-congested waters, it is difficult to estimate locations that have potential collisions at a detailed level based on conventional concepts. Furthermore, it is necessary to estimate the geometric number of collision candidates per unit time and area to compare and discuss safety in the waterways [19]. Therefore, a method is required for estimating the geometric number of collision candidates in every location by dividing a target sea area into smaller areas.

The objective of this study is to develop a method, which the authors call a “mesh-based method,” to calculate the geometric number of collision candidates that can deal with various encounter angles introducing a computational mesh. This enables the geometric number of collision candidates considering encounter angles per evaluation time to be estimated for each evaluation area, which is small enough to apply a uniform distribution regardless of the probability distributions; thus, the geometric number of collision candidates in wide-congested waters can be assessed comprehensively. In this paper, conventional concepts used to estimate the geometric number of collision candidates are described, and a new method is proposed based on conventional concepts. Then, the proposed method is validated by applying it to simple ship traffic and by comparing it with a result calculated based on one of the conventional concepts. Finally, the proposed method is applied to traffic data of AIS-equipped ships navigating in Tokyo Bay in Japan to show locations that have potential collisions based on every encounter angle.

2 Ship collision frequency models

The ship collision frequency and probability model was initially proposed by Fujii et al. [1, 6] and Macduff [3] during the 1970s. According to Fujii’s concept, the ship collision frequency can be generally estimated as follows [1, 2, 6, 7]:

$$F = N_G \times P_C, \quad (1)$$

where F is the ship collision frequency, N_G is the geometric number of ship collision candidates, and P_C is the ship collision causation probability. In addition, N_G indicates the potential number of ship collisions if two ships encounter each other and no evasive maneuvers are made to avoid a collision. An estimation method for N_G has been developed by some researchers. It is based on ship geometric parameters that represent ship traffic conditions, such as ship traffic volume, length, and velocity. In addition, P_C is the probability that ships on a collision course will fail to avoid a collision and depends on a variety of factors, such as human errors, equipment failures, maneuverability, and sea conditions. It can be estimated based on historical accident data for each area [9, 20]. Another approach for estimating P_C is the application to analytical models such as a Bayesian network [21], which represents probabilistic relations among the factors and results by setting up a graphical model for the former. From Eq. 1, it is possible to say that F is the collision frequency in cases in which ships with potential collision conditions (N_G) have actual collision accidents based on certain causations (P_C).

The value of N_G is important for estimating the collision frequency based on the reality of the ship traffic. Fujii et al. proposed a concept and method for calculating N_G based on ship traffic conditions, and N_G for overtaking, crossing, and head-on situations was estimated based on ship traffic observations in waterways in the 1970s and the 1980s [1, 2, 6, 7, 12, 13]. Owing to this method, N_G can be represented by rough geometric parameters, such as the ship traffic volume, length, and velocity. Thereafter, some researchers followed Fujii’s method, and renewed methods estimating N_G have been proposed [5, 8–11, 20, 22, 23]. In the 1990s, Pedersen developed a method to calculate N_G [8] following Fujii’s idea. This model is more realistic in that ship encounter conditions can be represented by the use of conventional geometric parameters and new parameters such as the encounter angles between two ship courses and lateral traffic distributions based on the ship positions.

In this chapter, the concepts and calculations of representative models, i.e., Fujii’s model and Pedersen’s model, are presented to understand their characteristics. In addition, disadvantages when these models are applied to ships in congested waters are also discussed.

2.1 Fujii’s model

2.1.1 Concept of Fujii’s model

Fujii et al. proposed a model to estimate N_G assuming that two ship groups have an encounter per unit time [1, 2, 6, 7,

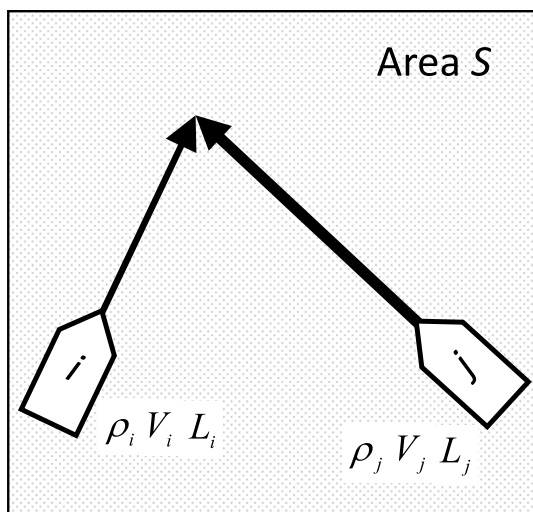


Fig. 1 Concept of ship collision in Fujii’s model

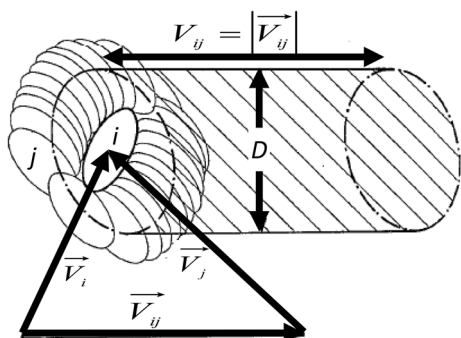


Fig. 2 Geometry of ship collision in Fujii’s model (created based on [6])

[12, 13]. That is, N_G indicates the number of ship encounters during a collision course. Figure 1 shows the concept of the collision model, and Fig. 2 shows the geometry of collisions, in which a ship belonging to group i with velocity V_i collides with the other ship belonging to group j with velocity V_j within the shaded area shown in Fig. 2 per unit time. The number of collisions for ship j can be estimated as $\rho_i \cdot D \cdot V_{ij} \cdot S \cdot T$, where ρ_i is the traffic density of group j , D is the geometric collision diameter shown in Fig. 2, V_{ij} is the relative velocity, S is the area of the target waterway, and T is time. Therefore, the geometric number of collision candidates N_G in which ship j collides with ship i with traffic density ρ_i is calculated using Eq. 2

$$N_G = \rho_i \cdot \rho_j \cdot D \cdot V_{ij} \cdot S \cdot T. \tag{2}$$

2.1.2 Calculation procedure and an example using Fujii’s model

Fujii et al. published some studies regarding the calculation of N_G in some Japanese straits, most of which were written in Japanese [7, 12, 13]. In this section, the study in [12] is introduced, and N_G in a waterway is calculated based on their model with past ship traffic observation data.

Table 1 shows the ship traffic data observed in the East waterway in the Bisan–Seto Strait from 1970 to 1981, which were obtained by eye and radar observations [12]. It should be noted that the East waterway in the Bisan–Seto Strait is one of the waterways with the heaviest traffic volume in Japan. The data in Table 1 include the traffic volume, average velocity, and average ship length. First, the traffic volume Q , which is the average daily number of ships, was obtained visually for 3 days each year and was counted for each category of gross tonnage (GT): “mini” (less than 100 GT), “small” (100–500 GT), “medium” (500–3000

Table 1 Ship traffic data per day on average in east waterway in the Bisan–Seto Strait from 1970 to 1981 [12]

Year	Traffic volume (daily average): Q					Total Q (num)	Velocity (knot)	Length (m)
	Mini	Small	Medium	Large	Huge			
1970	102	408	96	12	1	619	–	40
1971	42	324	189	22	2	589	–	50
1972	55	307	214	51	5	632	9.37	55
1973	26	286	254	48	9	623	11.09	60
1974	28	323	161	36	7	555	11.84	54
1975	25	445	122	27	7	626	–	48
1976	33	337	158	46	7	581	11.46	55
1977	12	243	143	39	5	442	–	58
1978	22	302	154	55	7	540	–	58
1979	32	389	126	51	9	607	10.30	54
1980	21	346	137	45	7	556	–	54
1981	31	377	122	60	5	595	11.25	54

GT), “large” (3000–20,000 GT), and “huge” (over 20,000 GT). “Total Q ” is total value of traffic volume of all GT categories for 4 years. There is no detailed description of the method used to count the traffic volume for each GT category in [12]. However, Fujii (1981) indicated that visual observations were usually conducted using binoculars or a telescope at the port entrance or cape with a clear view and unobstructed by buildings, and the number of ships passing through one location can be counted [7]. The ship length and gross tonnage were determined by reading the name of the observed ship and searching for information on the ship using a “Ship registry.” As another approach, telescopes will also be helpful, having a fractional scale for measuring the viewing angle. If the horizontal viewing angle and distance between an observed ship and an observation position were determined, the ship length can be estimated. The gross tonnage (GT) was estimated using the reduction relation with the ship length (L), i.e., $GT = L^3/250$ [7]. It should be noted that the suitable distance between an observed ship and an observation position when using a telescope is 2–3 km. Second, the “Velocity” was the average value for each GT category. The ship velocity when navigating in waterways was obtained based on radar observations, and the velocity distribution was created. The “Velocity” was calculated based on the distribution. Third, the “Length” was the average value estimated based on the concept of the equivalent length in [12]. Traditionally, the degree of congestion for different ship sizes is estimated by the conversion factor, C_{conv} , as shown in Table 2. Then, the converted traffic volume Q_L for the small-sized ships [$C_{conv} < 1.0$ (ship size of “medium” is assumed as a standard)] is calculated as $Q_L = C_{conv} \times Q$, where Q is the actual traffic volume. This means that the degree of congestion can be considered to be smaller if the ship size is smaller when the number of ships in a certain area is constant. This Q_L concept is often used as an assessment index of the ship density and congestion in a conventional marine traffic engineering [12, 24–27]. In [12], Fujii et al. assumed that the equivalent length, L , for each category of gross tonnage can be estimated using $L = C_{conv} \times 70.0$ (m), where the standard size (“medium”) of ships is regarded as 70.0 m. The length shown in the right column in Table 1 was calculated by this concept using C_{conv} in Table 2, as detailed in [12]. In the following calculations, the data in Table 1 were used.

The concept and procedures for calculating N_G by Fujii et al. [7, 12, 13] are as follows:

Table 2 Conversion factors for ship traffic volume [12]

Ship size	Mini	Small	Medium	Large	Huge
Conversion factor (C_{conv})	0.25	0.50	1.0	2.0	4.0

- (1) The traffic density, ρ , is obtained by Q (the number of ships per unit time (s)), V (m/s), and the width of the waterway, W (m), as follows:

$$\rho = \frac{Q}{V \cdot W} \tag{3}$$

If Q is the total traffic volume in two ships of groups i and j in a waterway, and there is little difference in traffic volume between the groups, ρ_i and ρ_j are assumed as $\rho_i = \rho_j = \rho/2$.

- (2) In addition, D (m) is the geometric collision diameter, which is the width of the possible collision area between the ships of groups i and j and is taken perpendicular to the direction of V_{ij} (see Fig. 2). Although a method for calculating D was not mentioned in Fujii’s study [1, 2, 6, 7, 12, 13], D takes a value of $(L_i + L_j)/6$ to $L_i + L_j$ depending on the ship velocity and encounter angle between two ship courses. In most of their work, D is assumed to be $(L_i + L_j)/2$, which is the average ship length.
- (3) Next, V_{ij} is calculated using V_i and V_j , which are the average velocities of ship groups i and j , respectively. In Fujii’s study, the average velocity of both ship groups is assumed to have the same value ($V = V_i = V_j$). Then, V_{ij} is $2V$ in the case of a head-on encounter, $\sqrt{2}V$ in a crossing encounter, and $\alpha \cdot V$ in an overtaking encounter. Although α should be calculated based on the velocity distributions and encounter situations, α is usually assumed to be approximately 0.3 [7]. In [12], α was assumed to be 0.25, and this value was used in the following calculation.
- (4) In addition, S (square meter) is calculated by multiplying the length and width of the waterways.
- (5) Finally, T (s) is the target time. Then, N_G for a certain period can be obtained using Eq. 2.

In this paper, N_G is calculated for the East waterway in the Bisan–Seto Strait based on the above procedure from (1) to (5), where V , ρ_i , ρ_j , D , and V_{ij} are calculated according to the traffic data in Table 1. Because the method for calculating V in blank spaces in Table 1 was not mentioned in [12], V is assumed as the average of the observed data between 1970 and 1981 in this paper. As the East waterway in the Bisan–Seto Strait is almost a parallel waterway and most ships meet in head-on and overtaking cases, V_{ij} is calculated under such situations. In addition, W is 1400 m and S is 518,000,000 m² (see the area enclosed by the line in Fig. 3).

Here, the results of the calculation example are shown. Figure 4 shows N_G calculated by Fujii et al. in [12] (“ N_G in the reference*”) and our calculation results (“ N_G by our calculation”) in head-on and overtaking cases every

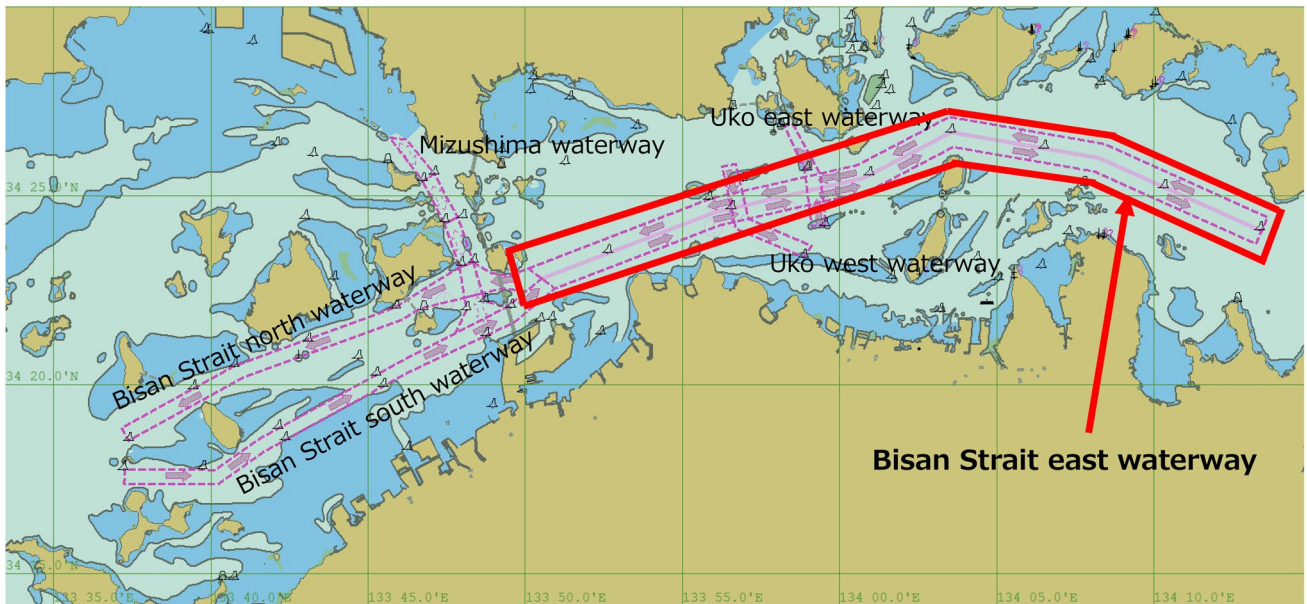
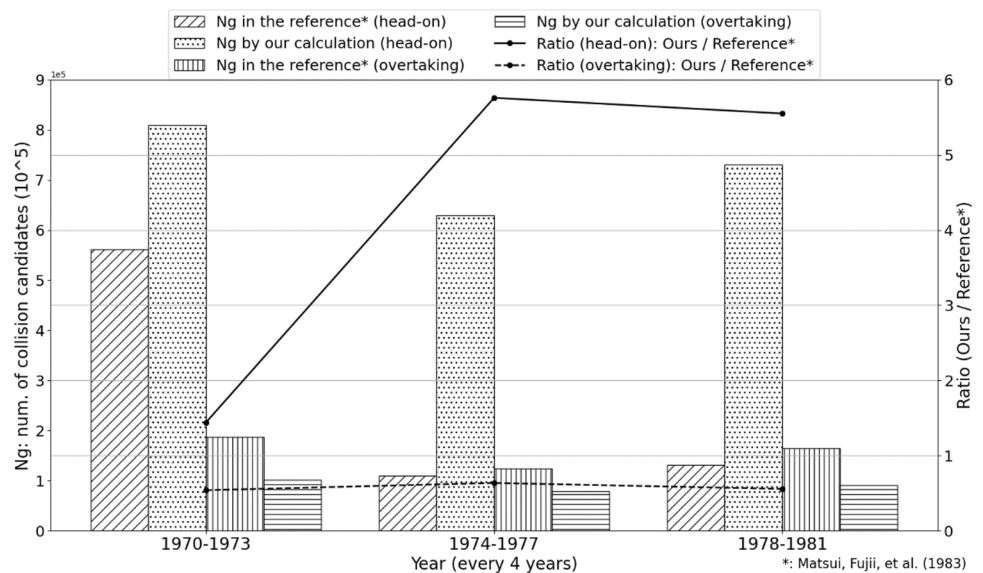


Fig. 3 The Bisan–Seto Strait in Japan

4 years. The ratios of N_G by our calculation to N_G in [12] are also shown in Fig. 4. Figure 4 shows that N_G in the head-on situation from 1970 to 1973 does not differ much between the N_G in [12] and N_G in our calculation. By contrast, N_G based on our calculation of the head-on situation from 1974 to 1977 and from 1978 to 1981 is much larger than that in [12]. This may be because the Maritime Traffic Safety Act was enforced in 1973 and traffic separation is applied in the strait, and thus, the ship traffic density distribution by traffic separation may have an effect on N_G . Because Eq. 2 does not consider the density distribution, our calculation does not include the effect of the

distribution. However, in the calculation by Fujii et al. [12], the density distribution for traffic separation was considered, although the method for calculating the distribution was not described in detail. Therefore, N_G of the head-on case based on our calculation is much larger than N_G in [12]. In addition, the value of N_G for an overtaking situation through our calculation is smaller than N_G in [12], but they do not differ much between the calculation results compared with the head-on situation. This means that the effect of traffic separation on N_G under an overtaking situation is smaller, even if the density distribution for a traffic separation is not considered in our calculation.

Fig. 4 Geometric number of ship collision candidates per 4 years in east waterway by Fujii’s model



Although a method for calculating the density distribution is not described in their previous paper, the distribution effects on N_G per year are shown in [12]. According to [12], the distribution effect on N_G increased in the East waterway from 1973, which coincides with the start year of traffic separation enforcement. In a head-on situation, the value of the distribution effect on N_G from 1973 was estimated to be approximately 0.2 [12]. If N_G by our calculation from 1974 to 1977 and from 1978 to 1981 are multiplied by 0.2, respectively, it is almost equal to N_G in [12]. In the overtaking case in 1973, the distribution effect on N_G was estimated to be approximately 2.0 [12]. This is also almost the same value between the results of our calculations and [12].

It is concluded that N_G in our calculation does not differ much from that in [12], and the density distributions influence the results of the calculated N_G , particularly in waterways to which traffic separation is applied.

2.1.3 Points in Fujii's model for more adequate calculation

From the above calculation and discussion, some points in Fujii's model for a more adequate calculation can be summarized as follows:

- The collision diameter D is not clearly described, because the equation for calculating D based on the encounter conditions was not established.
- The detailed relative velocity V_{ij} in such crossing situations was not calculated.
- Although a rough estimation of N_G for head-on and overtaking situations is possible based on Fujii's model, it is difficult to calculate N_G accurately for a crossing situation.
- Equation 2 is unsuitable for traffic separation waterways, because it is difficult to consider the density functions of separated ship traffic in all waterways.

As shown in the above points, it is necessary to calculate N_G when considering the encounter conditions, such as the density distribution, collision diameter, and relative velocity, as well as the encounter angles between two ship courses, and thus, Fujii's model can only roughly estimate N_G under certain situations. However, the basic idea of this model is to calculate the geometric number of collision candidates in an area by focusing on the density of ships.

2.2 Pedersen's model

2.2.1 Concept of Pedersen's model

Pedersen published a study on the calculation of N_G [8], following the concept of Fujii's model. He presented a formulation for estimating N_G when waterways 1 and 2 cross with angle θ_{ij} , as shown in Fig. 5. In waterway 1 in Fig. 5, $Q_i^{(1)}$ is the number of ships per unit time in class i , $V_i^{(1)}$ is the average velocity of the ships in class i , $f_i^{(1)}(z_i^{(1)})$ is the lateral traffic distribution of ships in class i , and $z_i^{(1)}$ is the distance from the centerline of waterway 1. In the same way, $Q_j^{(2)}$, $V_j^{(2)}$, $f_j^{(2)}(z_j^{(2)})$, and $z_j^{(2)}$ are defined for waterway 2. It should be noted that superscripts (1) and (2) indicate each waterway, which is the direction of ship traffic, and subscripts (i, j) represent the ship class assuming that the ships navigating in the waterway are classified by the ship type and ship size.

Then, Pedersen's formulation to calculate N_G becomes the following [8]:

$$N_G = \sum_i \sum_j \iint_A \frac{Q_i^{(1)} \cdot Q_j^{(2)}}{V_i^{(1)} \cdot V_j^{(2)}} \cdot f_i^{(1)}(z_i^{(1)}) \cdot f_j^{(2)}(z_j^{(2)}) \cdot V_{ij} \cdot D_{ij} \cdot dA \cdot \Delta t, \quad (4)$$

where A is the target sea area for the calculation, V_{ij} is the relative velocity between the ships in waterways 1 and 2, D_{ij} is the collision diameter, and Δt is the time for the assessment.

The derivation of Eq. 4 can be explained as follows: First, the number of ships per unit time in a small segment, dz_j , of waterway 2 can be represented by $Q_j^{(2)} \cdot f_j^{(2)}(z_j^{(2)}) \cdot dz_j$, which can be regarded as the number of ships in the area, $dz_j \cdot V_j^{(2)}$, because the movement distance of a ship in waterway 2 per unit time is $V_j^{(2)}$. The density of ships (the

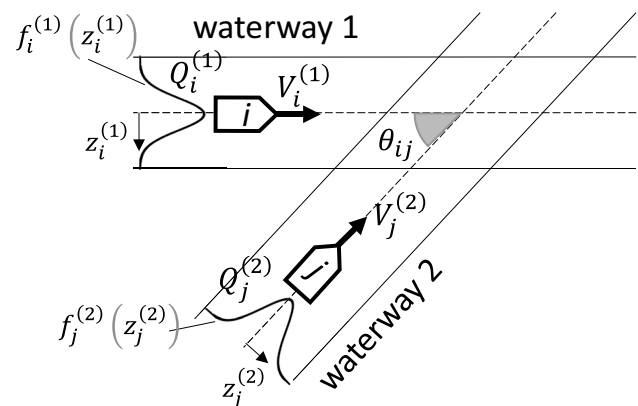


Fig. 5 Concept of ship collision in Pedersen's model (created based on [8])

number of ships per unit area) in waterway 2 can then be obtained by dividing the number by the area as follows:

$$\frac{Q_j^{(2)}}{V_j^{(2)}} \cdot f_j^{(2)}(z_j^{(2)}) \tag{5}$$

Then, the number of ships belonging to class j in waterway 2 on a collision course with one ship belonging to class i in waterway 1 per unit time is defined by Eq. 6, because the area where a ship in waterway 2 is on a collision course per unit time can be approximately represented by $V_{ij} \cdot D_{ij}$, as shown in Fig. 6, assuming that the probability distribution of $f_j^{(2)}(z_j^{(2)})$ is constant within the area

$$\frac{Q_j^{(2)}}{V_j^{(2)}} \cdot f_j^{(2)}(z_j^{(2)}) \cdot V_{ij} \cdot D_{ij} \tag{6}$$

Because Eq. 6 is the number of ships along the collision course for only one ship of class i in waterway 1, and the total number of collisions in area A at time Δt can be computed by multiplying the density of ships of class i in waterway 1, $(Q_i^{(1)}/V_i^{(1)}) \cdot f_i^{(1)}(z_i^{(1)})$, with Eq. 6, and integrating over the considered area and summing over ship classes i and j , as shown in Eq. 4.

Although the concept of Pedersen’s formulation is similar to that of Fujii’s model, the improvements and differences between the two models are as follows:

(1) Collision diameter D_{ij} is defined by Eq. 7 derived from the conceptual figure shown in Fig. 6 [8], where V_{ij} is the relative velocity computed using Eq. 8

$$D_{ij} = \frac{L_i^{(1)}V_j^{(2)} + L_j^{(2)}V_i^{(1)}}{V_{ij}} \sin\theta_{ij} + B_j^{(2)} \left\{ 1 - \left(\sin\theta_{ij} \frac{V_i^{(1)}}{V_{ij}} \right)^2 \right\}^{\frac{1}{2}} + B_i^{(1)} \left\{ 1 - \left(\sin\theta_{ij} \frac{V_j^{(2)}}{V_{ij}} \right)^2 \right\}^{\frac{1}{2}} \tag{7}$$

$$V_{ij} = \sqrt{(V_i^{(1)})^2 + (V_j^{(2)})^2 - 2 \cdot V_i^{(1)} \cdot V_j^{(2)} \cdot \cos\theta_{ij}} \tag{8}$$

In the formulation above, $L_i^{(1)}$ and $B_i^{(1)}$ ($L_j^{(2)}$ and $B_j^{(2)}$) represent the average length and breadth of the ships in class i (or j) in waterway 1 (or 2). It should be noted that the angle θ_{ij} between two ship courses is explicitly considered, and the relative velocity, V_{ij} , is also calculated according to the angle between courses. It should be noted that Kaneko [9] pointed out that Pedersen’s formulation does not consider the probability distribution of the appearance of ships in the collision area ($D_{ij} \cdot V_{ij}$) and proposed a precise formulation to evaluate N_G .

(2) The lateral distribution of ship traffic in the waterway ($f_i^{(1)}(z_i^{(1)})$ and $f_j^{(2)}(z_j^{(2)})$) is considered. Any distribution can be applied if the spatial distribution is understood [28]. Because the collision diameter, D_{ij} , for two ship courses with arbitrary angles is explicitly defined, and the traffic distribution in the waterway ($f_i^{(1)}(z_i^{(1)})$ and $f_j^{(2)}(z_j^{(2)})$) is also clearly defined, it is possible to say that the number of ship collision candidates can be estimated more accurately by Pedersen’s model than by Fujii’s model.

It should be noted that Eq. 4 might be rather an approximate formulation when compared with more accurate formulation [10]. However, it is thought that this equation is very simple to use and can be regarded as an effective evaluation formulation.

2.2.2 Calculation example of Pedersen’s model

To show the characteristics of Pedersen’s model and compare it with those of Fujii’s model, a calculation example is described in this section. Here, the ship traffic data of the East waterway in the Bisan–Seto Strait in Table 1 is applied to the following equation introduced by Pedersen, which calculates N_G for a parallel waterway in a head-on situation. This is derived from the integral of Eq. 4 with a simplification, where $f_i^{(1)}(z_i^{(1)})$ and $f_j^{(2)}(z_j^{(2)})$ are normally distributed [8]

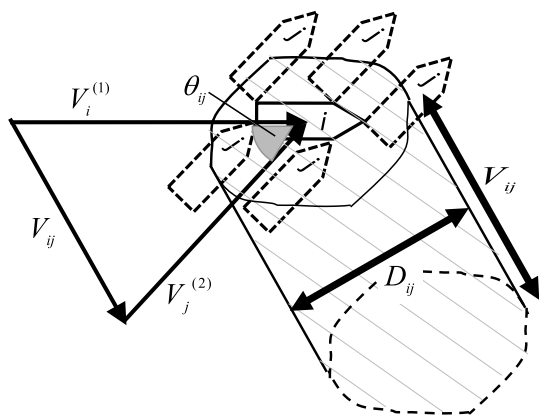


Fig. 6 Geometric collision diameter in Pedersen’s model (created based on [8])

$$N_G = \frac{1}{2\sqrt{\pi}} \cdot L_w \sum_i \sum_j \frac{Q_i^{(1)} \cdot Q_j^{(2)}}{V_i^{(1)} \cdot V_j^{(2)}} \cdot (V_i^{(1)} + V_j^{(2)}) \cdot (B_i^{(1)} + B_j^{(2)}) \cdot \frac{1}{\sigma_{ij}} \cdot \exp\left(\frac{-\mu^2}{4(\sigma_{ij})^2}\right) \cdot c \cdot \Delta t, \tag{9}$$

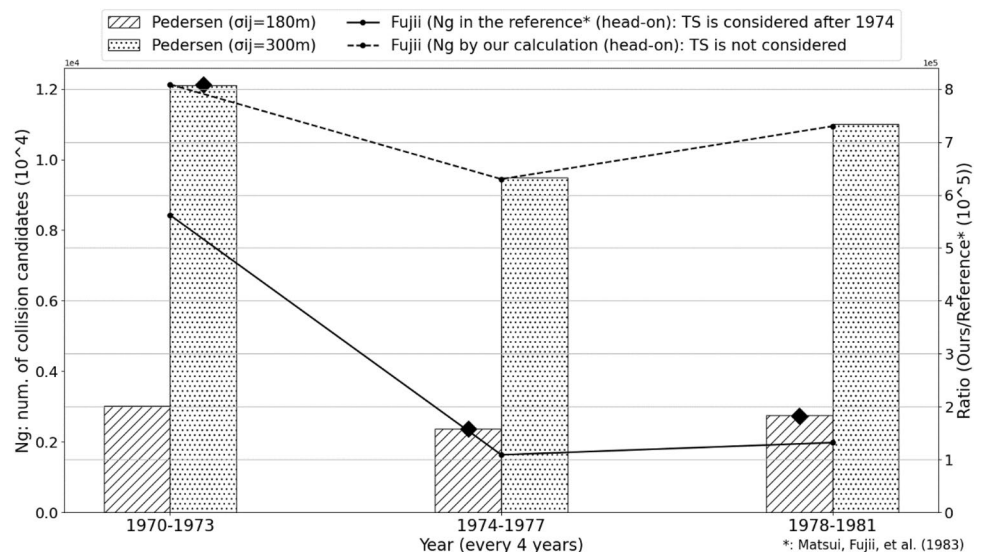
where L_w is the length of the parallel waterway, μ is the distance between the average positions of the two ship groups, and σ_{ij} is the standard deviation of $z_i^{(1)}$ and $z_j^{(2)}$. Note that the standard deviations of $z_i^{(1)}$ and $z_j^{(2)}$, which are considered to be $\sigma_i^{(1)}$ and $\sigma_j^{(2)}$, are assumed to be the same ($\sigma_{ij} = \sigma_i^{(1)} = \sigma_j^{(2)}$). In this calculation, L_w is the length of the East waterway, i.e., 3700 m (see Fig. 3), and Q ($Q_i^{(1)}$, $Q_j^{(2)}$ based on the number of ships per second) and V ($V_i^{(1)}$, $V_j^{(2)}$, m/s) are determined as indicated in Table 1. The average breadth of ships the, B ($B_i^{(1)}$ and $B_j^{(2)}$, m), is calculated using the approximation equation $L/B=6.0$ in this paper, because B was not observed. According to Inoue [26], μ was estimated to be approximately 620 m and σ_{ij} was estimated to be approximately 180 m in the East waterway in 1976. Although σ_{ij} in 1976 can be defined as shown above, σ_{ij} prior to 1973 may be much larger than in 1976, because the traffic separation rule was not applied before 1973. It is difficult to define an accurate value of σ_{ij} , because there were no statistical data for the ship traffic before 1973. For this reason, N_G is calculated for different values of σ_{ij} in this paper, that is, σ_{ij} is set as 180 and 300 m, whereas μ is fixed at 620 m for both cases.

In Fig. 7, the bar graph shows the results of N_G calculated in this paper by Pedersen’s formulation (Eq. 9), and it was observed that N_G varies considerably according to the magnitude of σ_{ij} . Because σ_{ij} is an important parameter for representing the lateral traffic distribution ($f_i^{(1)}(z_i^{(1)})$

and $f_j^{(2)}(z_j^{(2)})$ in Pedersen’s formulation), it is possible to imagine that an accurate definition of $f_i^{(1)}(z_i^{(1)})$ and $f_j^{(2)}(z_j^{(2)})$ is important for accurately estimating N_G . In addition, Fig. 7 shows that N_G through Pedersen’s formulation is smaller than that by Fujii’s model in the head-on situation shown in Fig. 4. This is because the lateral traffic distribution is not considered properly assumed in Fujii’s model, whereas a practical distribution, which is a normal distribution, is assumed in Pedersen’s model. Another reason is the difference in collision diameter between the models (D in Fujii’s model and D_{ij} in Pedersen’s formulation). Here, D of the head-on situation in Fujii’s model is calculated by the simple assumption of the ship length, $(L_i + L_j)/2$. However, D_{ij} of the head-on situation in Pedersen’s formulation can be calculated by $(B_i + B_j)$, which is smaller than D in Fujii’s model. Therefore, N_G by Fujii’s model seems to be overestimated and becomes larger than that of Pedersen’s formulation.

To discuss the trend of the calculated N_G for each model, N_G of the head-on situation by Fujii’s model in Fig. 4, “ N_G in the reference* (head-on)” in which a traffic separation (TS) is considered after 1974, and “ N_G by our calculation (head-on)” in which a TS is not considered, are also plotted by line graphs in Fig. 7. It should be noted that both line graphs of N_G by Fujii’s model refer to the vertical axis of the right-hand side of Fig. 7, and the trend of change of N_G by year is similar between the results of Pedersen’s formulation using each σ_{ij} (180 and 300 m) and

Fig. 7 Number of ship collision candidates by the model developed by Pedersen and Fujii (head-on situation)



the result by Fujii's model when TS is not considered (N_G by our calculation (head-on)). In addition, when N_G from 1970 to 1973 is regarded as the value calculated by the larger σ_{ij} (300 m) and N_G after 1974 is regarded as a value calculated by a smaller σ_{ij} (180 m) (see \blacklozenge in Fig. 7), the result is similar to the trend of change of N_G by year by Fujii's model when TS is considered [N_G in the reference * (head-on)]. It can be stated that the lateral traffic distribution is important for accurately estimating N_G and should be considered based on the actual ship traffic conditions.

2.2.3 Points in Pedersen's model for more general and adequate calculation

Pedersen's model improves the points of Fujii's model, which was described in Sect. 2.1.3, and it is possible to calculate a more detailed calculation of N_G in an area. For a more general and adequate calculation, some points in Pedersen's model can be summarized based on the above calculation and discussion as follows:

- (a) Pedersen's concept can be applied only to ship traffic in a delimited area, and thus, encounter situations and patterns are limited. Estimating N_G in wide-congested waters where many ships navigate along various courses and encounter various angles is unsuitable.
- (b) The lateral distribution of ship traffic in wide-congested waters cannot be assumed by a simple distribution profile such as a normal distribution [28] and differs by location.

2.3 Points to be improved for discussing ship traffic safety in every location

The models of Fujii and Pedersen are important for calculating N_G in an area. Although both models are based on actual ship traffic, the model of Pedersen is more detailed than that of Fujii in that navigational situations such as the ship encounter angle and lateral traffic distribution are considered in a delimited area.

However, in recent years after introducing AIS, ship traffic safety in wide coastal waters around ship-congested bays such as Tokyo Bay and Ise Bay in Japan has received attention [29–31]. In these waters, where ship traffic is not separated, many ships navigate along various courses and encounter various angles in various locations. It is difficult to consider the encounter angles and lateral traffic distributions for each location in water using these concepts. To discuss ship traffic safety in waters, a new method for estimating N_G in every location should be developed by considering the concepts with certain modifications. The points to be improved are as follows.

- (a) Consideration of various encounter angles between two ships in water, including wide-congested waters, to investigate locations that have potential collisions in detail.
- (b) Estimation of lateral traffic distributions for each location in wide-congested waters.

3 Proposal of mesh-based estimation method of N_G

When Fujii and Pedersen proposed their concept to estimate N_G , the observation and comprehension of ship behavior was laborious and was difficult to observe for a lengthy period of time. Their concepts, which express the characteristics of the ship behavior as a probability distribution, are effective in estimating N_G in a delimited area based on a small amount of information. However, it has become easier to observe the ship behavior for the longer term and in a wider area owing to the recent introduction of an AIS. At present, when big data related to ship traffic can be used, it is not necessary to express the characteristics of the entire sea area as a probability distribution, and the ship behavior, such as the ship density, velocity, and size, can be comprehended even in smaller areas.

If the area is sufficiently small, the characteristic number of ship behaviors can be assumed to be constant. In this paper, a new method based on the concept of Fujii and Pedersen is proposed to calculate N_G in an area which is sufficiently small to assume a uniform distribution of the ship behavior by introducing a computational mesh, which we call a "mesh-based estimation method" [32–35]. This enables an improvement of the points discussed in Sect. 2.3 and deals with various encounter angles even in wide-congested waters. In this chapter, the concept, calculation process, and validation of the method are presented.

In addition, N_G , which is calculated using this method, is based on ship traffic data, such as the ship traffic volume, position, velocity, course, and length. Thus, the method can be applied to any type of ship, such as ships equipped with an AIS ("AIS-equipped ships") and those not equipped with an AIS, if the traffic data are obtained. In this paper, the target was AIS-equipped ships, and AIS data were used to calculate N_G .

3.1 Concept of mesh-based estimation method

First, to simplify the explanation of a "mesh-based estimation method" in this chapter, the formulation of the geometric number of collision candidates per unit area is simplified, as shown in Eq. 10, which is for the case in which the ships of course group i with direction θ_i and

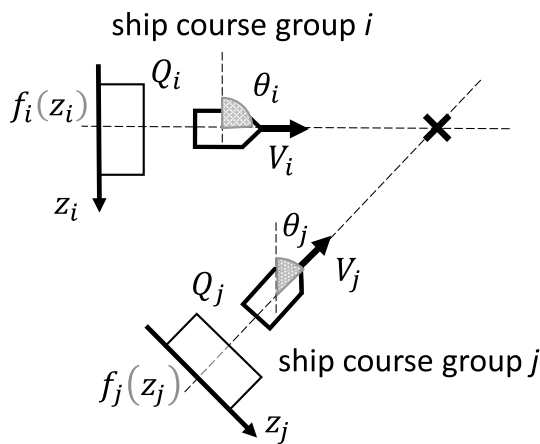


Fig. 8 Concept of ship encounter in mesh-based estimation method

course group j with direction θ_j have an encounter, as shown in Fig. 8. The method focuses on dealing with the ship density, which is the number of ships per unit area, as shown in $Q_i \cdot f_i(z_i)/V_i$ and $Q_j \cdot f_j(z_j)/V_j$ in Eq. 10, without considering the direction of integration. Equation 10 follows Eq. 4 of Pedersen’s model and is in the same form as Eq. 2 of Fujii’s model, which estimates the geometric number of collision candidates in an area by dealing with the ship density. It should be noted that one ship class for

each waterway in Eq. 4 is considered to be one ship course group in Eq. 10

$$N_{G_perunitarea} = \frac{Q_i}{V_i} \cdot f_i(z_i) \cdot \frac{Q_j}{V_j} \cdot f_j(z_j) \cdot V_{ij} \cdot D_{ij} \cdot \Delta t. \quad (10)$$

In Eq. 10, Q_i , V_i , and $f_i(z_i)$ represent the number of ships per unit time, the ship velocity, and the lateral traffic distribution of the ship course group i , respectively. In addition, V_{ij} and D_{ij} are the relative velocity and collision diameter, respectively, which can be estimated using Eqs. 7 and 8, respectively:

The prerequisite and calculation process of the method for applying Eq. 10 to the ship traffic is shown in Fig. 9 and described Sects. 3.1.1 and 3.1.2.

3.1.1 Prerequisite process

We developed software to obtain AIS data, such as the position, velocity, course, and length of the ships, by setting the virtual gate lines in a sea area and obtaining the AIS data of ships passing through each gate line [36]. In the prerequisite process of the method, virtual gate lines are set in a target sea area in the software to obtain AIS data of ships passing through each virtual gate line. We previously arranged

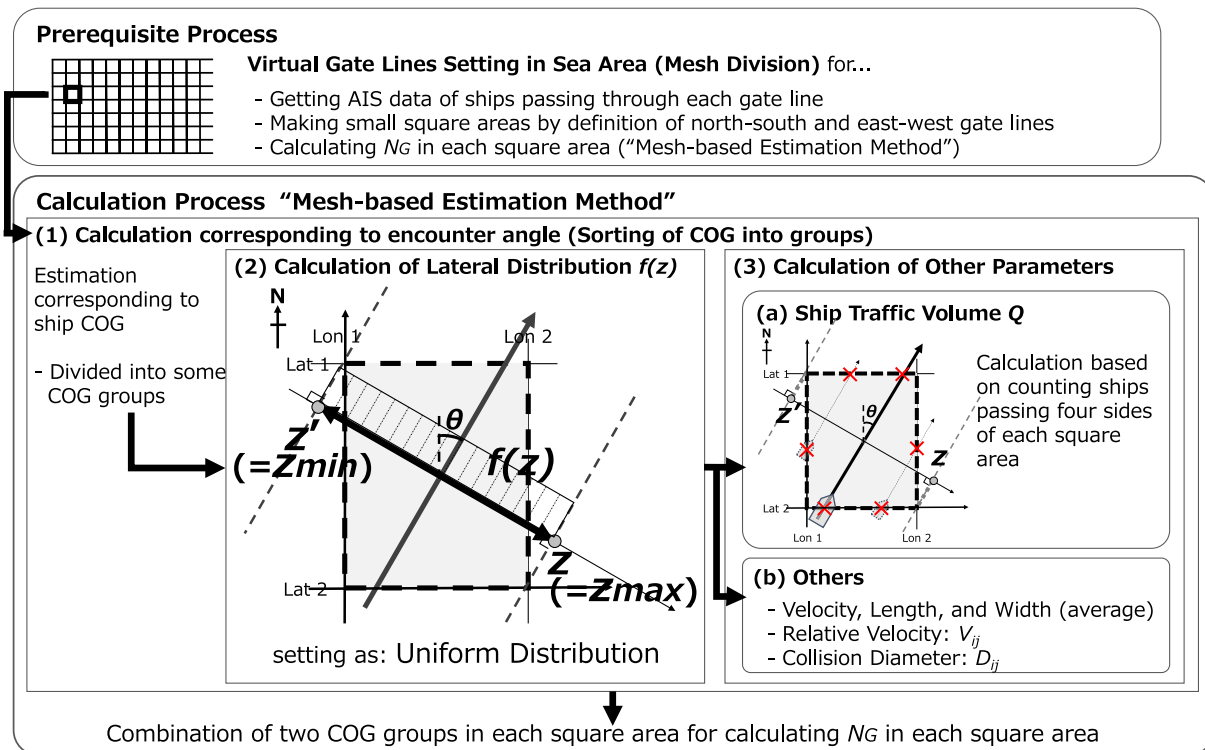


Fig. 9 Prerequisite and calculation processing of mesh-based estimation method

virtual gate lines almost perpendicular to the direction of ship traffic to make it easier to obtain AIS data [29, 31–38].

In the prerequisite process, as shown in Fig. 9, virtual gate lines are set parallel to the latitude lines (WE gate lines) for detecting ships navigating in a north–south direction, and are set parallel to the longitude lines (“NS gate lines”) for detecting ships navigating in an east–west direction, or are set to any other lines, finely at regular intervals in a targeted sea area. This will make small square areas surrounded by four gate lines and enable the acquisition of AIS data for each square area.

3.1.2 Calculation process

During the calculation process, the three processes in (1)–(3) are conducted in each square area.

(1) Encounter angles between the two ship courses.

To estimate the N_G corresponding to the encounter angles between two ship courses, the course over ground (COG) θ is divided into groups (“COG group”) in each square area. Then, through a combination of two COG groups in each square area N_G can be calculated by the following procedure.

In the following calculation examples in Sect. 3.2, the number of COG groups (n_{cog}) is taken as 72 with intervals of 5° , which means that COG group 1 is $-2.5 \leq \theta_1 < 2.5$, COG group 2 is $2.5 \leq \theta_2 < 7.5, \dots$, and COG group 72 is $352.5 \leq \theta_{72} < 357.5$ (COG group i is $-2.5 + 5 \cdot (i - 1) \leq \theta_i < 2.5 + 5 \cdot (i - 1)$). It should be noted that we call each COG group “ $\theta_i = 0$,” “ $\theta_i = 5$,” ..., “ $\theta_i = 355$ ” (degree) in the following sections, because a medium value for each group is $0^\circ, 5^\circ, \dots, 355^\circ$, respectively.

(2) Lateral traffic distributions.

Figure 10 shows the calculation of the lateral distributions based on this method. The lateral distributions of ship traffic can be assumed to have a uniform distribution in each square area if the gate lines are set at small intervals and each square area becomes sufficiently small (see the left part of Fig. 10). To apply Eq. 10 to each square area, the lateral traffic distribution $f_i(z_i)$ for each COG group (heading angle) θ_i can be assumed as shown in the right part of Fig. 10, and can be represented by Eq. 11, where $z_{imax} - z_{imin}$ indicates the width of the square area when the width is considered perpendicular to the direction of θ_i . The lateral traffic distribution $f_j(z_j)$ for another heading angle θ_j can also be calculated in the same way as Eq. 11

$$f_i(z_i) = \frac{1}{z_{imax} - z_{imin}}. \tag{11}$$

(3) Other parameters.

The other parameters in Eqs. 10 are calculated for each square area using the following method:

(a) Ship traffic volume.

The ship traffic volume Q (Q_i and Q_j in Eq. 10) is assumed to be the number of ships with COG group θ (θ_i and θ_j in Eq. 10), navigating each square area per second. The number of ships can be obtained from the traffic data by counting ships passing through four sides of each square area at time T . It should be noted that one ship passes the sides of the square area twice, as shown in Fig. 11, and thus, Q is defined by all counts passing through four sides of each area, N , in time T , as shown in Eq. 12

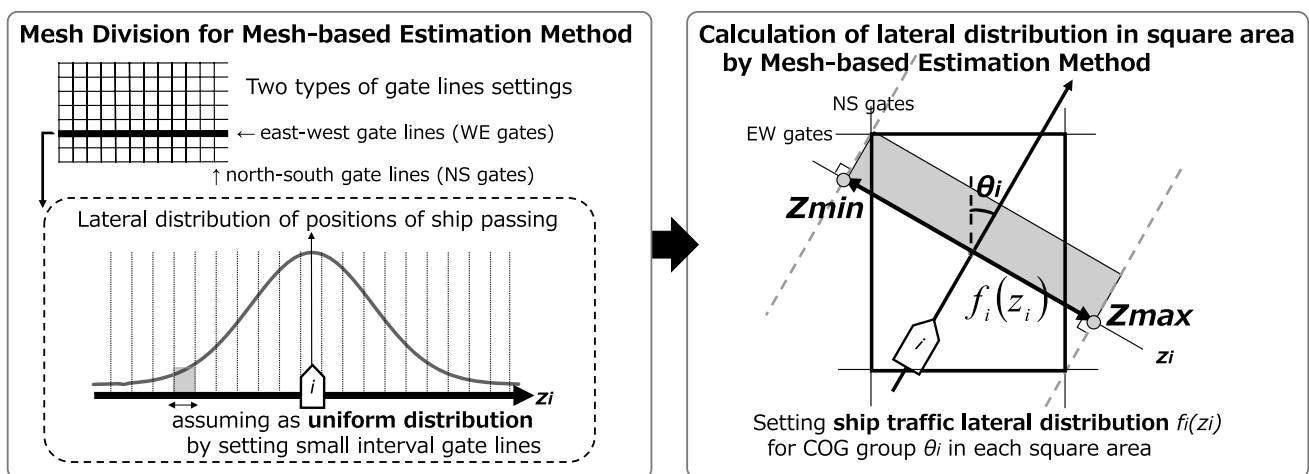
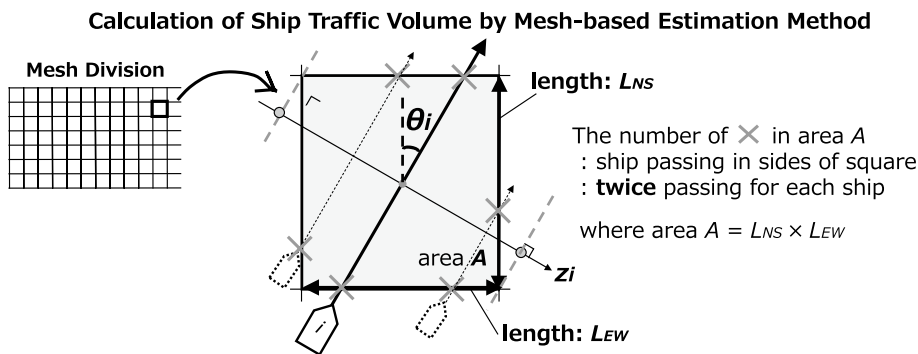


Fig. 10 Calculation of lateral traffic distribution using mesh-based estimation method

Fig. 11 Assumption of ship passing and calculation of ship traffic volume using mesh-based estimation method



$$Q = \frac{N}{2T} \tag{12}$$

(b) Other parameters.

The relative velocity V_{ij} and collision diameter D_{ij} can be calculated using Eqs. 7 and 8, respectively. The ship velocity V (V_i and V_j), ship length L (L_i and L_j), and ship width B (B_i and B_j) are assumed to be the average values in each COG group θ (θ_i and θ_j) in each square area obtained from the traffic data.

3.1.3 Calculation of N_G

With this method, the geometric number of collision candidates per unit area during the time, Δt , is estimated at the representative point of the square area by Eq. 10. Then, by multiplying area, A , of the square area, the geometric number of collision candidates in the area is calculated. It should be noted that Eq. 10 and the procedures shown in Sects. 3.1.1 and 3.1.2 are for one ship size class for each direction, and the geometric number of collision candidates is calculated by summing the calculation in each square area between two ship COG groups of θ_i and θ_j .

It should also be noted that Eq. 11 is used not to consider the direction in the integration but to consider the number of ships per unit area, as shown in Eq. 5, which is used only to calculate Eq. 12. This means that the number of ships per unit area can be calculated using Eqs. 11 and 12, because the number of ships navigating within the breadth of $(z_{\max}^{(1)} - z_{\min}^{(1)})$ is counted to calculate Q .

Based on the above, the geometric number of collision candidates per assessment time, Δt , and the assessment area, A , which we call “Encounter frequency E_f ,” is given by Eq. 13

$$E_f = \sum_{i=1}^{n_{cog}} \sum_{j=i+1}^{n_{cog}} \frac{Q_i^{(1)} \cdot Q_j^{(2)}}{V_i^{(1)} \cdot V_j^{(2)}} \cdot f_i^{(1)}(z_i) \cdot f_j^{(2)}(z_j) \cdot V_{ij} \cdot D_{ij} \cdot A \cdot \Delta t. \tag{13}$$

Here, A is the square area calculated as $L_{NS} \times L_{WE}$ by the length of the sides of the square shown in Fig. 11, and n_{cog} is the number of COG groups in the area. As shown in Eq. 4, Pedersen proposed the method to calculate N_G by dividing ships into many classes, because the difference of ship length, type, and velocity are important elements which affect the geometric number of collision candidates. In this chapter, only the difference of COG is considered as shown in Eq. 13 to show validation of the proposed method. However, the proposed method can deal with these elements in the future when calculating the number of collision candidates.

It should be noted that the geometric number of collision candidates N_G is the total sum of all values of E_f in the target sea area based on this concept.

3.2 Validation of mesh-based estimation method

In this section, the validation of the mesh-based estimation method (“the proposed method” in this section) is carried out by comparing the calculated result of N_G using the proposed method with the result obtained through a conventional method for a simple example problem with a normal distribution. The example problem is established based on the ship traffic data obtained from AIS data of ships passing through gate lines in Tokyo Bay, which is a congested sea area in Japan (see Fig. 12). It should be noted that the intervals of each gate line, which is the size of the mesh division, are taken as 0.2 min, where 1.0 min is $1/60^\circ$ in latitude or longitude. For validation, we only considered the head-on situation, i.e., northbound (N/B) ships of group $i=1$ at $\theta_i=0^\circ$ and south-bound (S/B) ships of group $j=37$ at $\theta_j=180^\circ$. The other information for the traffic data is as follows:

- Target period: 92 days (From August 1, 2017 to October 31, 2017).

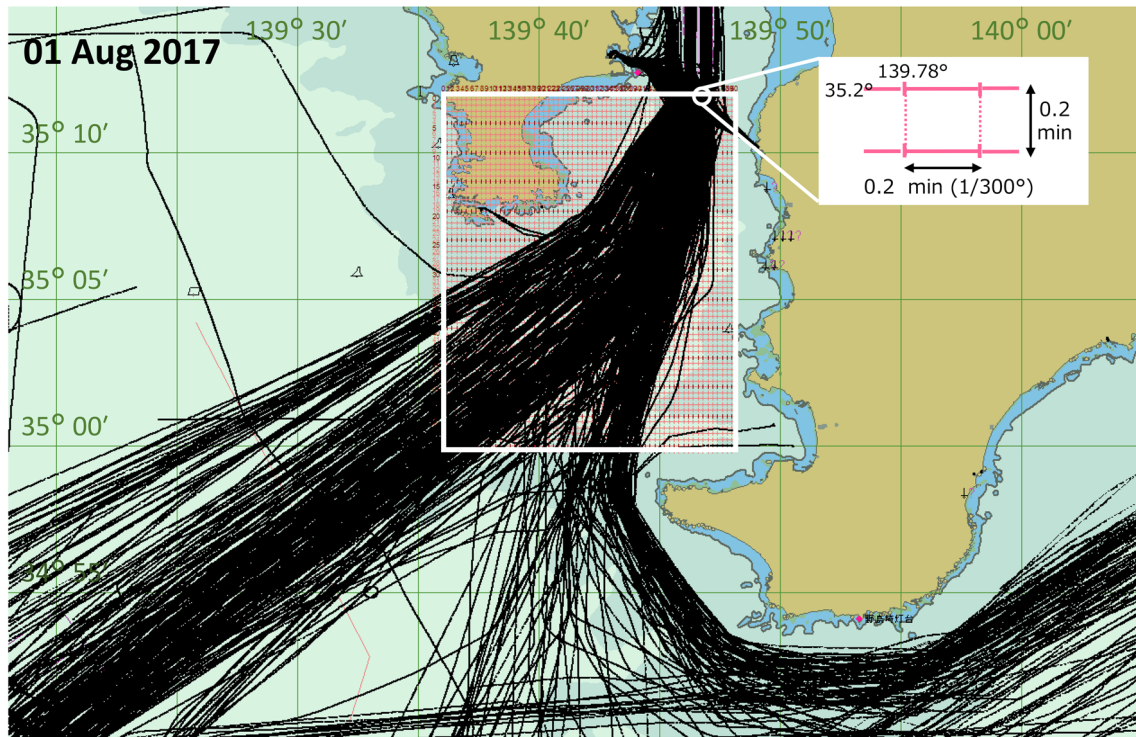


Fig. 12 Trajectories in Tokyo Bay in Japan. (Target sea area is surrounded by rectangular line)

- Target sea area: Tokyo Bay, Japan (Lat 35° 00' (35.0°)–35° 12' (35.2°) N, Lon 139° 36' (139.6°)–139° 48' (139.8°)).

Figure 12 shows the ship trajectories on August 1, 2017 in Tokyo Bay, and the target sea area is surrounded by a rectangular line. It should be noted that there are over 1000 navigating ships every day in Tokyo Bay, and most ships navigating in this bay are north- or south-bound ships.

For the calculation of N_G using the conventional method, the following equation for the head-on situation is used, which was introduced by Friis–Hansen and is a result of the integral of Eq. 4 without simplification [11]:

$$N_G = L_w \cdot \sum_{i,j} P_{Gi,j}^{head-on} \cdot \frac{V_{ij}}{V_i^{(1)} \cdot V_j^{(2)}} \cdot (Q_i^{(1)} \cdot Q_j^{(2)}), \quad (14)$$

where L_w is the length of the segment (target waterway), $P_{Gi,j}^{head-on}$ is the probability that two ships will collide in a head-on situation, and the other parameters are the same as in Eq. 13. It is possible for $P_{Gi,j}^{head-on}$ to be expressed as Eq. 15, when the lateral distribution is a normal distribution with distribution parameters $(\mu_i^{(1)}, \sigma_i^{(1)})$ and $(\mu_j^{(2)}, \sigma_j^{(2)})$ [11]

$$P_{Gi,j}^{head-on} = \Phi\left(\frac{B_{ij} - \mu_{ij}}{\sigma_{ij}}\right) - \Phi\left(-\frac{B_{ij} + \mu_{ij}}{\sigma_{ij}}\right), \quad (15)$$

where B_{ij} is the average ship breadth, $\Phi(x)$ is the standard normal distribution function, μ_{ij} is the mean sailing distance between two ships, which equals $\mu_i^{(1)} + \mu_j^{(2)}$, and σ_{ij} is the standard deviation of the joint distribution, which is equal to $\sqrt{(\sigma_i^{(1)})^2 + (\sigma_j^{(2)})^2}$. Using the above equation, N_G is calculated on the lines of constant latitude in the target sea area, assuming that, along each latitude line, the traffic distribution of the N/B and S/B ships is approximated based on a normal distribution. This means that the mean values and standard deviations of the passing positions for N/B ships $(\mu_i^{(1)}, \sigma_i^{(1)})$ and S/B ships $(\mu_j^{(2)}, \sigma_j^{(2)})$ are used to evaluate Eq. 14. It should be noted that N_G is calculated for all 59 lines of constant latitude from Lat 35.0° to 35.2° N.

By contrast, with the proposed method, N_G is calculated in each square area, which is surrounded by WE gate lines and NS gate lines with intervals of 0.2 min. The distributions of the ship traffic volume Q (Q_i and Q_j) in each square area are assumed to be uniform, which is determined from the approximated normal distributions used in the calculation with the conventional method. For the ship velocity V (V_i and V_j), the relative velocity V_{ij} and the collision diameter D_{ij} are calculated from the AIS data in each square area.

Figure 13 compares the calculation procedure of the conventional formulation shown in Eq. 14 with that of the

Fig. 13 Comparison of mesh-based estimation method (proposed method) with a conventional method

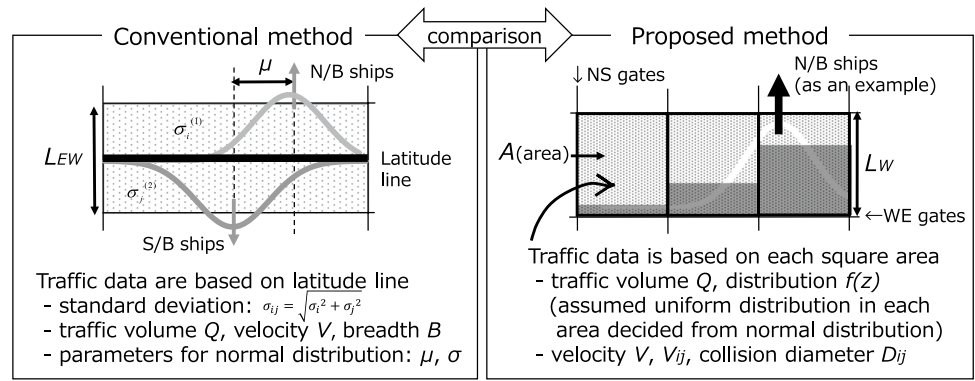
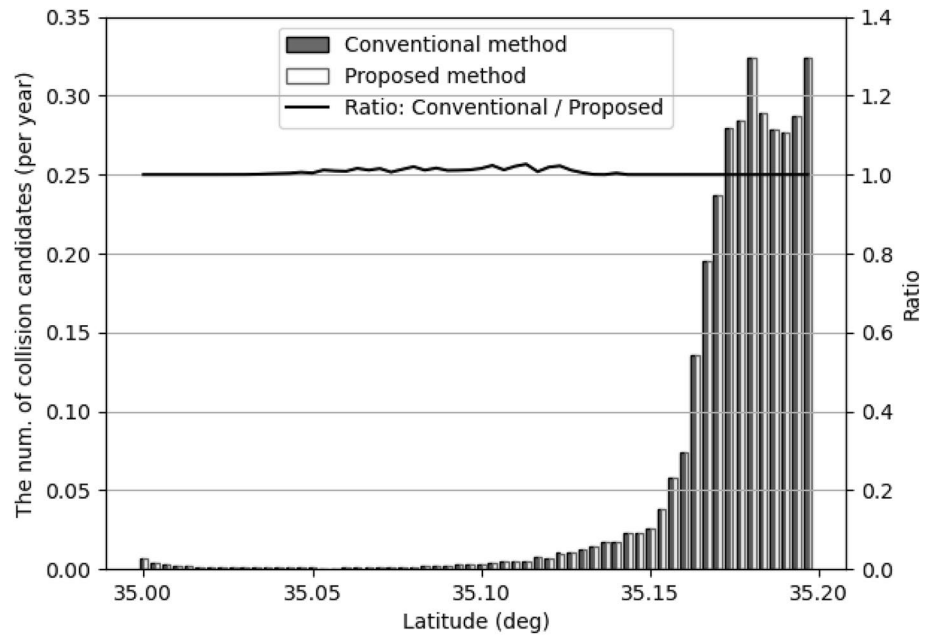


Fig. 14 N_G using the conventional method and the proposed method assuming $f(z)$ as a normal distribution



proposed method. In the procedure using Eq. 14, to calculate N_G , the parameters are calculated along a latitude line, and the length of target waterway L_{EW} is set in accord with the length of the side of the square area, $L_{EW} = 0.2$ min (the left part of Fig. 13). By contrast, with the proposed method, the square areas of the latitude line were considered in the calculation (the right part of Fig. 13).

The bar graphs in Fig. 14 show the values of N_G calculated using the proposed method and the conventional method (Eq. 14). It should be noted that the actual traffic distributions are occasionally different from a normal distribution. In such cases, it is reasonable to use a proper distribution to compute the actual N_G . However, in this section, to check the validity of the proposed method, we calculated N_G and E_f for a comparison with the calculation results by the conventional method using an approximated normal distribution. In addition, the ratio of N_G by the conventional method to the proposed method is shown as a line graph in Fig. 14, which refers to the vertical axis of

the right-hand side. The horizontal axis represents the latitude (unit, degree). It was found that the values calculated by both methods almost coincide, and the average ratio between them is 1.006 and the average standard deviation is 0.0076. It can be concluded that the calculation procedure and mesh size of the proposed method are validated. It should be noted that this section and Fig. 14 show the validation for head-on situation, but N_G for crossing situation whose encounter angle is 10° to 170° has been also validated by the authors by calculating N_G using the proposed method for virtually generated data of crossing traffic flow.

4 Characteristics of encounter situation in Tokyo Bay

This chapter shows the analysis results of E_f and N_G for Tokyo Bay in Japan using a mesh-based estimation method to discuss the characteristics of the types of encounters

in this bay. In this paper, the method was applied to AIS data during the first stage of the method development. The initial settings for the size of the mesh division, number of COG groups (n_{cog}), and target period are the same as those in Sect. 3.2.

- Size of mesh division: 0.2 min mesh.
- n_{cog} : 72 groups at intervals of 5°.
- Target period: 92 days (From August 1, 2017, to October 31, 2017).

It should be noted that all traffic data in all directions during the target period in Tokyo Bay were used in this analysis.

4.1 Calculation example for one square area using mesh-based estimation method

In this section, the difference in the calculated value of E_f between some COG groups in one square area using Eq. 13 is shown as a calculation example. The upper left coordinate point of the target square area is “Lon 139.78° E, Lat 35.2° N,” as shown in the left part of Fig. 12, which is the south entrance of Uraga Channel. The calculation condition is the encounter between the ship group $i=2$ ($\theta_i=5^\circ$) and the other COG group, θ_j , whose encounter angles with the group of $\theta_i=5^\circ$ are less than 10° ($j=1, 3, 4, 72$: $\theta_j=0, 10, 15, 355$), is calculated. Table 3 shows the values of the parameters of the ship group of $\theta_i=5$ in the target square area, and Table 4 shows the values of the parameters of the ship groups of $\theta_j=0, 10, 15, 355$, and the calculated E_f between ship group $\theta_i=5^\circ$ and each ship group θ_j within the area. It should be noted that E_f , whose encounter angle is less than 10° , can be regarded as an overtaking situation. In addition, it notes that E_f shown in Table 4 is calculated as the value of E_f per day within the square area. It can be seen from the table that E_f between $\theta_i=5^\circ$ and $\theta_j=10$ (group 2–3) under an

overtaking situation is the highest, which means that there are many ships entering the Uraga Channel with such COGs and encounter angles. It should be noted that, in the following sections, E_f is calculated in the way described above.

4.2 Color map of E_f

The color map of E_f obtained from the value of E_f for each square area is suitable for understanding the perspective of E_f in an entire sea area. Using a color map, it is also easy to see how high or low E_f is in each square area. Figure 11 shows a color map of E_f per day. To understand the characteristics around Tokyo Bay, the target sea area is set larger than the area in Sect. 3.2, i.e., Lat 34.8°–35.2° N, Lon 139.4°–140.0° E. Square areas in which E_f is above 0.01 (times/day) are shown by the maximum color.

Figure 18 shows that E_f in Tokyo Bay is higher than that outside the bay, and it was found that there is a difference in traffic volume between in and outside the bay. This is because Tokyo Bay has some principal ports and harbors, and many ships gather from various locations. Therefore, it can be stated that there might be a correlation between the traffic volume and E_f . In general, when E_f is estimated using the models of Fujii and Pedersen for a certain area, it is thought that E_f usually becomes high if the traffic volume (Q) is large. This can be seen from Table 1 and Figs. 4 and 7, in that the number of ships (traffic volume) from 1970 to 1973 is the largest among the other 4-year periods (1974–1977 and 1978–1981) in Table 1, and that E_f from 1970 to 1973 calculated by both models is the highest (Figs. 4 and 7). From this viewpoint, the calculated E_f by the mesh-based estimation method also shows the same trend, that is, there is a correlation between traffic volume and E_f .

By contrast, some wide fairways can be seen outside the bay in Fig. 18. It was found that the sea area around the bay

Table 3 Parameters of the ship group $\theta_i=5^\circ$ in a square area

Group no. (i or j)	COG θ_i or θ_j	Q (/s)	V (m/s)	L (m)	B (m)	$f(z)$
$i=2$	5°	$2.68 \cdot 10^{-4}$	6.06	125.9	20.1	$3.00 \cdot 10^{-3}$

Upper left coordinate point: “Lon 139.78° E, Lat 35.2° N”

Table 4 Parameters and calculated E_f between the ship group $\theta_i=5^\circ$ and other ship groups θ_j within a square area

Group no. (i or j)	COG θ_i or θ_j	Q (/s)	V (m/s)	L (m)	B (m)	$f(z)$	E_f (times/day) per square area	
$j=1$	0°	$7.55 \cdot 10^{-5}$	6.18	123.0	19.6	$3.31 \cdot 10^{-3}$	$7.13 \cdot 10^{-3}$	Group 2–1
$j=3$	10°	$1.44 \cdot 10^{-4}$	6.18	104.6	16.3	$2.77 \cdot 10^{-3}$	$1.05 \cdot 10^{-2}$	Group 2–3
$j=4$	15°	$4.45 \cdot 10^{-5}$	6.58	87.1	14.2	$2.58 \cdot 10^{-3}$	$5.65 \cdot 10^{-3}$	Group 2–4
$j=72$	355°	$2.57 \cdot 10^{-5}$	6.04	117.3	18.6	$3.00 \cdot 10^{-3}$	$4.25 \cdot 10^{-3}$	Group 2–72

Upper left coordinate point: “Lon 139.78° E, Lat 35.2° N”

is congested with ships entering or leaving the bay in various courses, and that there are some encounter routes, as shown in Fig. 18a–c. In the western sea area in Fig. 18, the encounter route is shown as (a) in the figure, and is wide in the north area which is close to the bay. In addition, the encounter route is as shown in (b) in the figure, whose value of E_f is smaller than E_f in other areas. It is thought that various

ship encounters with various angles occur in this area. In the eastern sea area shown in Fig. 18c, ship encounters are relatively high in the waters off Sunosaki and Nojima-saki, where ships are used as waypoints to alter the ship course.

Fig. 18 Color map of number of collision candidates per day using new method

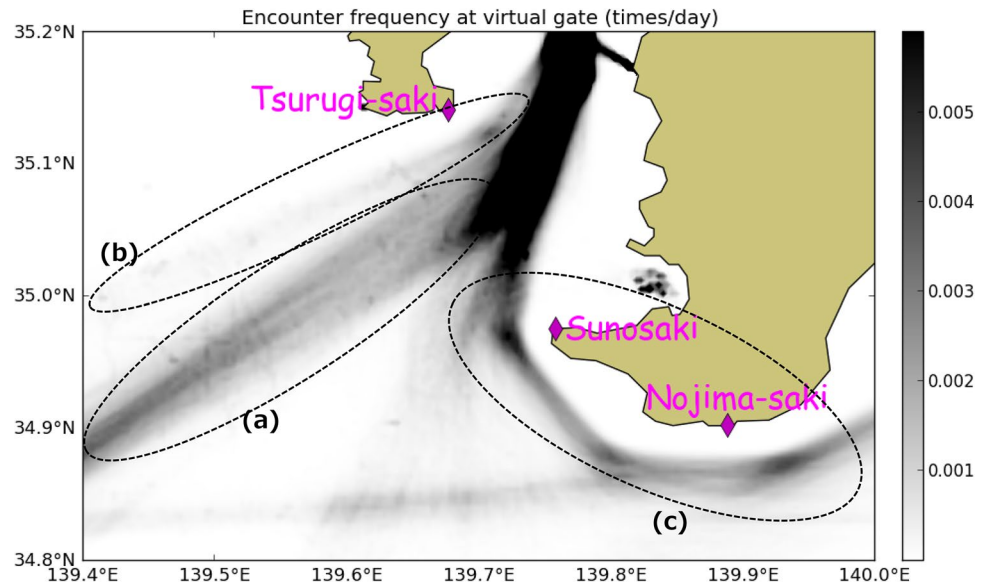


Table 5 Definition of encounter situations

Encounter situation	Definition of position relationship	Encounter angle θ_{ij}
Overtaking	A ship is in the sea area posterior to another ship from a direction more than 20° abaft her beam	$0 \leq \theta_{ij} < 70$
Head-on	Two ships meet along reciprocal or nearly reciprocal courses	$165 \leq \theta_{ij} < 180$
Crossing	Two ships cross each other	$70 \leq \theta_{ij} < 165$

Fig. 19 Number of collision candidates of overtaking encounters per day

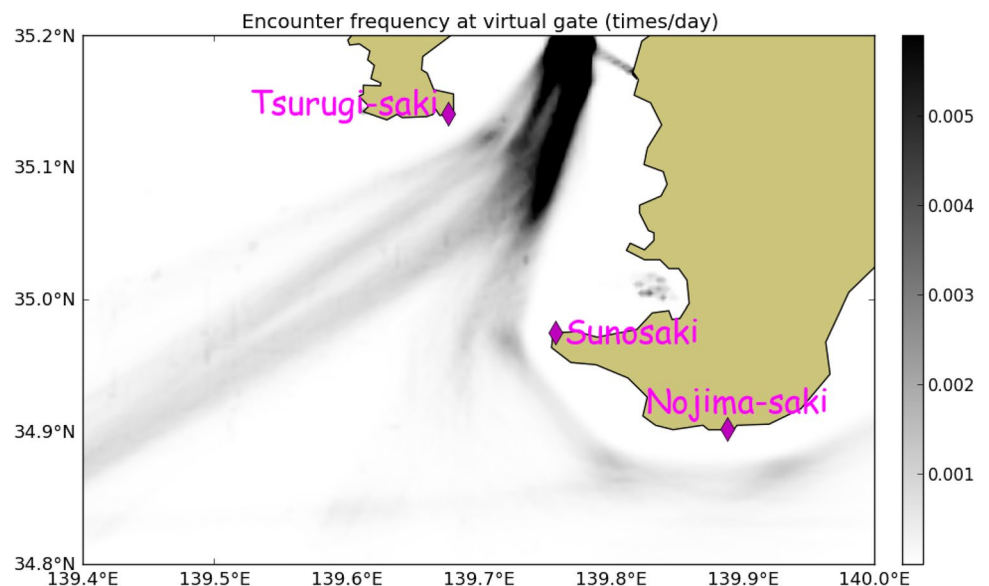


Fig. 20 Number of collision candidates of head-on encounters per day

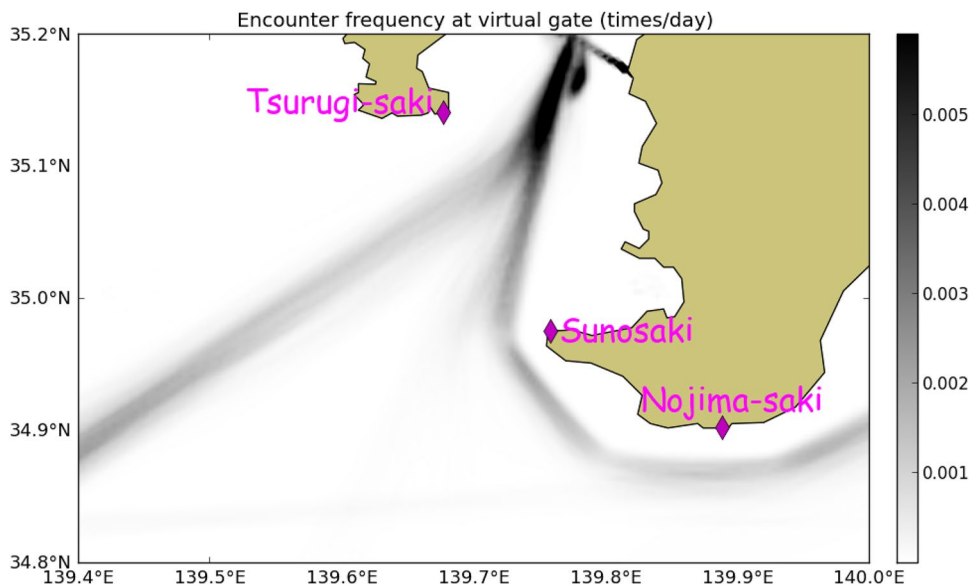
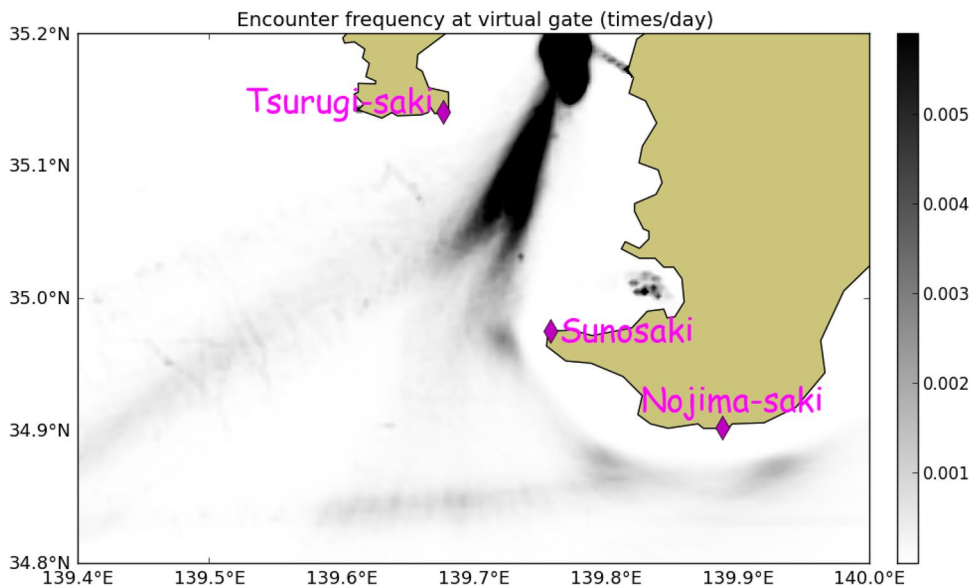


Fig. 21 Number of collision candidates of crossing encounters per day



4.3 E_f based on encounter situations

In Table 5, three encounter situations are defined based on the relative encounter angle θ_{ij} , shown in Eq. 16. Then, E_f for each situation was calculated. Each encounter situation is called an overtaking, crossing, and head-on situation. Figure 19 shows the value of E_f for an overtaking situation, Fig. 20 shows E_f for a head-on situation, and Fig. 21 shows E_f for a crossing situation

$$\begin{aligned} |\theta_{ij}| &= |\theta_i - \theta_j| \text{ (in case } |\theta_i - \theta_j| \leq 180\text{degree)} \\ |\theta_{ij}| &= 360 - |\theta_i - \theta_j| \text{ (in case } |\theta_i - \theta_j| > 180\text{degree)}. \end{aligned} \tag{16}$$

Figures 19, 20, 21 show that the calculated value of E_f for each encounter situation has its own characteristics. In the overtaking and head-on situations shown in Figs. 19 and 20, some encounters can be seen along some principal fairways outside the bay, as shown in Fig. 18a–c, which is the area in the western sea area in the figure and in the waters off Sunosaki and Nojima-saki. In the western sea area shown in Figs. 19 and 20 almost the same encounter routes are present. However, the encounter sea area for a head-on situation is narrow, and the encounter sea area for an overtaking has certain routes and is relatively wide. On the other hand, Fig. 21 shows that there are a large number of encounters for crossing situations at the entrance of the bay, because ships enter and leave the bay crossing in this area. Using the color

maps in Figs. 19, 21, and 22, the encounter areas of each encounter situation can be detected.

4.4 Calculation of E_f using mesh-based estimation method

It was concluded from the above results that it is possible to estimate E_f for congested waters by calculating E_f for each divided small square area based on the data of actual ship traffic, which was AIS data in this paper, through a mesh-based estimation method. The mesh-based estimation method considers the encounter angle between two ship courses and assumes the lateral traffic distributions as a uniform distribution if a square area divided by the mesh is sufficiently small; thus, E_f can be calculated regardless of the sea area, the encounter angle, and the lateral traffic distribution. Using this method, high and low values of the actual E_f in each square area can be observed. In addition, the comparison of E_f for different encounter situations is shown, and the characteristics of each situation can be discussed. Using a mesh-based estimation method, it might become easier to forecast hazardous areas for each encounter situation.

In addition, when collision frequency F in Eq. 1 is estimated for a wide sea area such as Tokyo Bay, the total E_f in a wide area should be calculated. To calculate the total E_f in a wide target area, the calculated E_f values for all small square areas in the area are summed. By applying this calculation, it is possible to quantitatively assess the collision occurrences for arbitrary large sea areas.

5 Conclusion

In this paper, the authors proposed a new idea and method for calculating the geometric number of collision candidates, which can deal with various encounter angles by location in waters including wide-congested waters. The models of Fujii and Pedersen were presented to understand its characteristics, and the proposed method was based on these models. We reached the following conclusions:

- (1) It is difficult to calculate the geometric number of collision candidates using the conventional models when arbitrarily wide-congested waters are considered, where various encounter angles between two ship courses occur and various distributions of positions of ship passing should be considered based on actual ship traffic. In this paper, we proposed a new method by introducing a computational mesh, which estimates the geometric number of collision candidates corresponding to the encounter angles in each divided area when considering each lateral traffic distribution. This

is called the “encounter frequency,” and the method is called a “mesh-based estimation method.”

- (2) The method was validated by applying simple ship traffic with a normal distribution. The results calculated using the proposed method mostly coincide with the results calculated through the conventional model, and thus, the proposed method was validated.
- (3) The proposed method was applied to traffic data of AIS-equipped ships navigating in and around Tokyo Bay, which is a congested water area in Japan. Using this calculation, high and low encounter frequencies were estimated in each location, where the entire target sea area was divided into sections, and some characteristics of the encounter frequency were demonstrated through encounter situations such as head-on, overtaking, and crossing situations.

It can be concluded that the proposed method can be used to discuss the ship traffic safety regardless of the sea area, encounter angle, and lateral traffic distributions, and can be used as a tool for navigational support that can show hazardous areas to navigating ships in advance.

Acknowledgements This work was partially supported by Innovative Science and Technology Initiative for Security Grant Number JPJ004596, ATLA, Japan.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Fujii Y et al (1974) Some factors affecting the frequency of accidents in marine traffic. *J Navig* 27(2):235–252
2. Fujii Y (1983) Integrated study on marine traffic accidents. *IABSE Colloquium on Ship Collision with Bridges and Offshore Structures* 42:91–98
3. Macduff T (1974) The probability of vessel collisions. *Ocean Ind* 9(9):144–148
4. Kristiansen S (2005) *Maritime transportation: safety management and risk analysis*. Routledge
5. Montewka J, Goerlandt F, Hanninen M, Ylitalo J, Seppala T (2011) Algorithm development and documentation. *Efficient, Safe and Sustainable Traffic at Sea (EfficienSea)*, pp 1–124
6. Fujii Y, Shiobara R (1971) The analysis of traffic accidents. *J Navig* 24(4):534–543

7. Fujii Y, Makishima T, Hara K (1981) Marine traffic engineering, Kaibun-do. (in Japanese)
8. Pedersen PT (1995) Collision and grounding mechanics. In: Proceedings of WEMT95, vol 1, pp 125–157
9. Kaneko F, Hara D (2007) Estimation of dangerous encounters' number from observed ship trajectories. In: Proceedings of the fourth International Conference on Collision and Grounding of Ships (ICCGS 2007), pp 187–193
10. Kaneko F (2013) An improvement on a method for estimating number of collision candidates between ships. In: Proceedings of the sixth International Conference on Collision and Grounding of Ships and Offshore Structures (ICCGS 2013), pp 27–37
11. Friis-Hansen P, Ravn ES, Engberg PC (2008) Basic modelling principles for prediction of collision and grounding frequencies: The BaSSy ToolBox Baltic Sea Safety. Technical University of Denmark, pp 1–59
12. Matsui T, Fujii Y, Yamanouchi H (1983) Investigation on Marine Traffic in the Bisan Seto – No.1 Probability of Vessel Collision and Grounding. Electronic Navigation Research Institute Papers No. 43, pp 1–19 (in Japanese)
13. Matsui T, Fujii Y, Yamanouchi H (1985) The probability and the risk of marine traffic accidents. J Jpn Inst Navig 73:75–86 (in Japanese)
14. Kujala P, Hanninen M, Arola T, Ylitalo J (2009) Analysis of the marine traffic safety in the Gulf of Finland. Reliab Eng Syst Saf 94(8):1349–1357
15. Silveira PAM (2013) Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal. J Navig 66(6):879–898
16. Japan Transport Safety Board (2015) Marine Accident Investigation Report, MA2015-2
17. Miyake R, Itoh H, Nishizaki C, Fukuto J (2016) Method of safety assessment for establishing ship routing system with marine traffic simulation. In: Proceedings of the seventh International Conference on Collision and Grounding of Ships (ICCGS 2016), pp 161–166
18. Miyake R, Itoh H, Nishizaki C, Fukuto J (2017) Safety assessment for establishing ships' routing -recommended route off the Western Coast of Izu O Shima Island. In: Proceedings of the fourth Asian Conference on Defense Technology (ACDT)
19. Itoh H, Miyake R (2019) Research on change of traffic safety accompanying the implementation. In: Proceedings of the eighth International Conference on Collision and Grounding of Ships (ICCGS 2019), pp 247–254
20. Kaneko F (2004) Effectiveness of Separation scheme for prevention of collision by diminishing ships' encounter probability. In: Proceedings of the third International Conference on Collision and Grounding of Ships (ICCGS 2004), pp 211–220
21. Itoh H, Kaneko F, Mitomo N, Tamura K (2007) A probabilistic model for the consequences of collision casualties. In: Proceedings of the fourth International Conference on Collision and Grounding of Ships (ICCGS 2007), pp 201–206
22. Søfartsstyrelsen (2008) Risk analysis for sea traffic in the area around Bornholm. COWI A/S Report No.: P-65775–002 (0), pp 1–113
23. Li S, Meng Q, Qu X (2012) An overview of maritime waterway quantitative risk assessment models. Risk Anal 32(3):496–512
24. Fujii Y, Tanaka K (1971) Traffic Capacity. The Journal of Navigation 24(4):543–552
25. Tanaka K, Yamada K (1970) On the equivalent number of vessels of various size in the marine traffic. J Jpn Inst Navig 44:67–72 (in Japanese)
26. Inoue K (1980) Lane width required at congested route. J Jpn Inst Navig 62:67–76 (in Japanese)
27. Gao X, Makino H, Furusho M (2014) Analysis of the waiting activity in entering port using AIS data. J Jpn Soc Civil Eng 70(2):I-948-I–53 (in Japanese)
28. Gluver H, Olsen D (2001) Survey of ship tracks in Fehmarn Belt. In: Proceedings of the 2nd International Conference on Collision and Grounding of Ships (ICCGS), pp 13–22
29. Kawashima S, Kawamura Y, Itoh H, Fukuto J (2015) Generation of ship traffic flow based on principal component analysis of AIS data and its application to ship traffic simulations for evaluation of encounter probability. In: Proceedings of Asia Navigation Conference 2015 (ANC 2015), pp 255–264
30. Itoh H, Isimura E, Yanagi Y, Mori Y (2012) Cognitive Model of maritime navigation and its use for collision accidents analysis. In: 2012 Fifth International Conference on Emerging Trends in Engineering and Technology, pp 93–99
31. Itoh H, Isimura E, Kudou J, Mori Y (2013) Estimation of ship encounter frequency at coastal areas using AIS data. In: Conference Proceedings of the Japan Society of Naval Architects and Ocean Engineers No.16 pp 309–312 (in Japanese)
32. Kawashima S, Kawamura Y, Itoh H, Fukuto J (2018) Development of estimation method of ship encounter frequency in congested sea areas. In: Conference Proceedings of the Japan Society of Naval Architects and Ocean Engineers No. 26, pp 195–199 (in Japanese)
33. Kawashima S, Itoh H (2019) Assessment of ship encounter and collision in congested sea areas. In: Proceedings of the 8th International Conference on Collision and Grounding of Ships and Offshore Structures (ICCGS 2019), pp 247–254
34. Kawashima S, Itoh H, Kawamura Y (2021) Estimation of collision causation probability based on collision frequency model. J Jpn Inst Navig 144:32–41 (in Japanese)
35. Itoh H (2022) Method for prediction of ship traffic behavior and encounter frequency. J Navig 75(1):106–123
36. Itoh H, Yakabe F (2014) Modeling ship traffic distributions in coastal areas. J Jpn Soc Naval Arch Ocean Eng 19:235–244 (in Japanese)
37. Kawashima S, Itoh H, Kimura A (2017) Collision frequency to offshore floating installations based on analysis of ship traffic flow. J Jpn Inst Navig 136:80–87 (in Japanese)
38. Seshita A, Kawamura Y, Fukuto J, Itoh H (2016) Proposal of traffic flow tube model for collision risk assessment method of congested sea area. In: Proceedings of Asia Navigation Conference 2016 (ANC 2016), pp 212–221

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.