27th International Congress on Waterborne Transportation, Shipbuilding and Offshore Constructions

Rio de Janeiro/RJ, 23rd to 25th October 2018

Estimating possible near miss collisions based on AIS big data for the Port of Rio de Janeiro

Maricruz A. F. S. Cepeda, UFRJ/COPPE, Rio de Janeiro/Brazil, maricruzcepeda@oceanica.ufrj.br
Gabriel Premoli Monteiro, UFRJ, Rio de Janeiro/Brazil, gabriel_premolimonteiro@poli.ufrj.br
João Vitor Marques de Oliveira Moita, UFRJ/COPPE, Rio de Janeiro/Brazil, joaov@oceanica.ufrj.br
Jean-David Caprace, UFRJ/COPPE, Rio de Janeiro/Brazil, jdcaprace@oceanica.ufrj.br

Abstract

Automatic Identification System (AIS) data records huge quantity of information regarding the safety of ships and port facilities in the international maritime transport sector. However, this big database is not only useful for the security of ships operations and port facilities. It can also be helpful for other important functions in maritime traffic such as estimating possible near miss ship collisions during a long period. This study develops an analytical approach to estimate possible near miss ship collisions in the Guanabara Bay of Rio de Janeiro (Brazil) using AIS database. The model is applied to rank the severity of an encounter between two vessels based on vessel conflict ranking operator (VCRO). The vessel size and the Minimum Distance to Collision (MDTC) concept are considered in the model. The results show that the proposed methodology is adequate for ranking and prioritizing encounters between ships. We suggest that ranking the possible near miss ship collisions around areas of high maritime traffic is important for decision makers and for the maritime authorities to make statements of maritime safety in relation to collision accidents.

Keywords: Ship, Near Miss Collision, Automatic Identification System, Big Data, Marine Traffic

1. Introduction

Maritime transportation is the main transportation mode for domestic and international trade. Over 80% of global trade by volume and more than 70% of its value being carried on board ships and handled by seaports worldwide, (UNCTAD, 2017).

Due to this, the growth in maritime traffic and the increase in the size of ships is inevitable. Consequently, it is essential to evaluate safety of marine traffic for the purpose of improvement of efficiency and safety of marine traffic, especially in areas where there is a high-density marine transport of passengers and/or cargo.

Therefore, it is important to develop and implement a method to know the behaviour of collisions and/or near miss collisions.

In general, Big Data has been receiving considerable attention in the specialized literature for the last 15 years. It is popular in shipping, especially on information collected about on-time ship characteristics to improve logistics, emissions, energy consumption and maintenance.

There are numerous studies focusing on the analysis of ship traffic. AIS data provides valuable input parameters in ship traffic simulation models for maritime risk analysis and for the prevention of shipping accidents. The article published by (Xiao, et al., 2015) reports the detailed comparisons of AIS data analysis between a Dutch case and a Chinese case. Another Chinese study published by (Zhang, et al., 2017) develops a tangible analytical approach to analyse ship traffic demand and the spatial–temporal dynamics of ship traffic in Singapore port waters using big AIS data. They found that the origin-to-destination pairs and navigation routes in the Singapore port waters maintain stability over time. Furthermore, they identify several hotspot areas in the Singapore Strait where ship sailing speeds are relatively high and ship sailing speeds in a few water areas vary greatly. (Zhang, et al., 2017) explored the relationship of hotspot areas with relation to the spatial distribution of ship accidents.

(Miyake, et al., 2015) developed and implemented an algorithm of collision avoidance that corresponded to a target area in the marine traffic simulation because actual events for collision avoidance depended on circumstances where ships were sailing. The authors of this article developed an automated marine traffic simulation system with AIS data and they proposed a series of systematic procedures for marine traffic simulation, including analysis for collision avoidance behaviours using AIS data.

Lastly, the study by (Szlapczynski & Szlapczynska, 2017) created a systematic and critical review of ship safety domain models, where they discussed multiple differences in approaches to the ship domain concept from definitions and safety criteria. In the literature review, we observed there are numerous studies focusing on marine traffic mainly because the transport risk assessment and accident investigation are more frequent with the increase of vessels size, speed and volume, which make it difficult to manoeuvre in small shipping lanes. The methods of determining ship domains have evolved with time.

(Fujii & Tanaka, 1971) introduced the concept of Ship Domain with the aim of defining the maximum capacity of Japanese fairways. They inferred a relation between the dimension of the ship domain and the length of the ship, or gross tonnage, on the basis of RADAR-observed traffic in the Tokyo Bay. They suggested the ship domain is an elliptical shape around the own ship. (Goodwin, et al., 1983) improved the concept of ship domain with the aim of marine traffic engineering and traffic control. In this reference, the dimension of the domain is deduced from statistical analysis of both simulator observations and observed traffic in the Thames Estuary. From this study, the authors suggest three sectors around the ship, each with a radius depending on sea area, traffic density and ship's size, i.e., length over all and Gross Tonnage.

Early models usually have been based on statistically processed radar data, but AIS has replaced radar as a data source and applied more advanced statistical methods to data processing, (Szlapczynski & Szlapczynska, 2017).

Currently, AIS data that collects navigation information on ships are becoming more readily available. However, it is shown that this data is actually largely underused, or the data is used partially for isolated studies.

The collisions and near miss collisions studies in the literature are mostly located in Europe and Asia, creating opportunities for studies in other areas such as South America.

The Guanabara Bay of Rio de Janeiro is an area where there is a massive transport of passengers and cargoes that may contaminate the environment. There are no studies about collisions and/or near miss collisions or that indicates that some areas of the bay are more susceptible to accidents.

Various methods have been proposed for analysing the accidental risk and safety of maritime transportation. The first category concerns *risk model and risk analysis*. Such methods aim to identify the areas of high accident risk and sometimes include mechanisms to assess the effect of incorporating risk control options. The second category focuses on *statistical analysis* of maritime accidents. Such methods aim to identify relations between variables to increase the understanding under which conditions maritime accidents occur. A third category utilises *non-accident information* to obtain insight in the safety of maritime transportation. Such methods take certain measures to indicate safety performance of the maritime transportation system, (Zhang, et al., 2017).

This paper aims to contribute to the last category of approaches, by proposing a method to identify potential zones of collisions and near miss collisions from AIS data of the bay of Rio de Janeiro. The authors developed a model to rank the severity of an encounter between two vessels based on vessel conflict ranking operator (VCRO). The vessel size and the Minimum Distance to Collision (MDTC) concept are considered in the model.

The purpose of the paper is ranking and prioritizing the encounters between ships inside the bay of Rio de Janeiro.

2. Methodology and data

In this paper, AIS data is used to develop an analytical approach to estimate possible near miss ship collisions. In this section, the area of the study, the description of the AIS data, the ship information data and the model are introduced.

2.1. Area of the study

Guanabara Bay (GB) is an oceanic bay, located on the Southeast Brazil in the state of Rio de Janeiro, between 22°40'S and 23°00'S latitude and between

043°00’W and 043°18’W longitude. The bay is known as the second largest bay in area in Brazil (after the Santos Bay). It has an area of approximately 384 km^2, including islands. On its western shore lies the city of Rio de Janeiro and fifteen other municipalities.

The populated region around the area of the study is composed by the following 16 municipalities, see Figure 1. There were 12 million people in 2017 based on Brazilian Institute of Geography and Statistics data (IBGE).

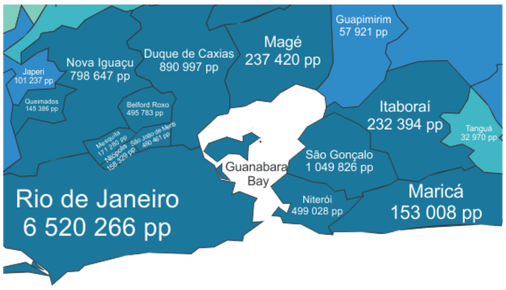

Figure 1 – Map of the regions surrounding the bay with total habitants in 2017 based on Brazilian Institute of Geography and Statistics data

The main port area in the Guanabara Bay is the Port of Rio de Janeiro, located in downtown of homonymous city at 23°45’S and 44°45’W, see Figure 2.

International shipping associated with the development of the country and petroleum industry increased the marine traffic through the Bay. There are five type of facilities distributed throughout the bay that are the heart of industry and the mass transit of people. These type of facilities are: dry cargo terminal, passenger terminals, petroleum terminals, shipyards, navy facilities and yacht clubs, see Figure 2. The study examines the marine traffic that may potentially affect up to 12 million people.

Figure 2 – View of Guanabara Bay includes six types of industrial/commercial facilities in the area

2.2. Automatic Identification System (AIS) data and ship information data

The Automatic Identification System (AIS) is a mandatory collision avoidance system required to be installed on ships by the International Maritime Organization (IMO) and the Maritime Safety Administration in several countries. The AIS system makes it possible to locate the grand majority of vessels throughout the world.

International voyaging ships with a Gross Tonnage (GT) of 300 or more, passenger ships of all sizes, domestic vessels with a GT of 200 or more traveling in coastal waters, and inland ships with a GT of 100 or more, are all required to be equipped with AIS. Special purpose vessels such as military ships, fishery ships, sports ships, and public service ships are exceptions, see Figure 3, (Chen, et al., 2018), (IMO, 2003).

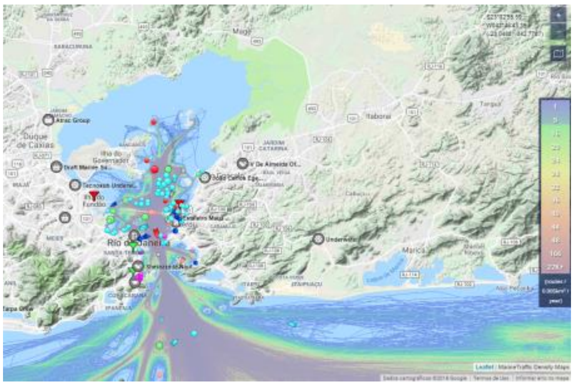

Figure 3 – View of Guanabara Bay included the Rio de Janeiro Port with AIS data, source: Marine Traffic 072018

There are, in fact, 2 classes of AIS, (Kerbiriou, et al., 2017): Class A: transponders are mandatory on board merchant ships exceeding 300 tonnages and all passenger ships meeting SOLAS standards (merchant navy, ferries, etc.). Class B: transponders concern small ships that are not required to comply with SOLAS conventions (recreational vessels, fishing vessels of less than 15 meters, etc.), to enable them to adapt voluntarily to the AIS system.

The objectives of IMO of implementing AIS are to enhance the safety and efficiency of navigation, safety of life at sea, and protection of the maritime environment. AIS facilitates communication between vessels and assist vessel traffic control functions in congested ports, locks and waterways, (Kerbiriou, et al., 2017).

The reported AIS data can be divided into static, dynamic, and voyage-related data categories. Static information includes ship name, ship type, length, breadth, etc. Dynamic data includes ship speed over ground, navigational status (operating mode), heading, rate of turn, position, etc. Voyage-related data includes current draught, description of cargo, and destination, (IMO, 2003). Besides ship information reported by AIS, detailed data for ship type, ship size, date of construction, design speed, gross tonnage and power of the engines can be obtained from the other databases as Marine traffic or IHS Maritime.

Our base station name is UFRJ-COPPE located at Technology Centre of Federal University of Rio de Janeiro. The hardware consists of one omnidirectional Sirio GP6E antenna of $2 \times 5/8\lambda$ (162 Mhz), one AIS receiver COMAR SLR350N and a Raspberry Pi 3 to provide Ethernet connectivity and to host a NMEA multiplexer server. The data warehouse configuration is composed of a NMEA message decoder and a Microsoft SQL server. The main table in the database contains 196 different fields extracted from the messages. The average range of the signal reception is 11 NM with a maximum of 74.5 NM. The location of the system is plotted in Figure 2 with a black antenna symbol.

The advantages of AIS are, (Kerbiriou, et al., 2017): Efficiency to assist ships in collisions avoidance and the port and maritime authorities in traffic monitoring and ensuring better surveillance of the sea. It is relevant tool for the protection of the marine environment.

In this study, four months of AIS data, from January until April 2018, were utilized to estimate near miss collisions over the area of the study. Figure 4 shows the distribution of the type of ships of the 2681 vessels recorded during the period of the study.

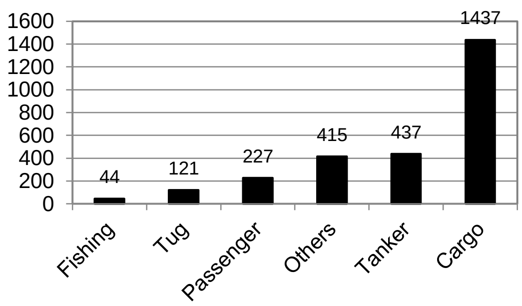

Figure 4 – Distribution of the type of ships

2.3. Workflow of the study

Figure 5 shown the workflow of the method developed in this paper

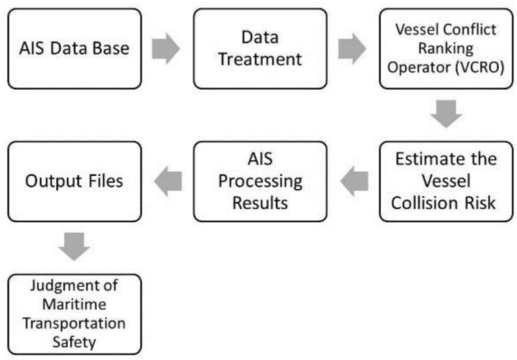

Figure 5 – Framework for research on near miss collisions

2.4. Model

The model and risk analysis is based on possible near miss collisions. Such methods aim to identify of areas of high accident risk. The AIS data enable investigation of the spatial and temporal relationship between two vessels, (Wu, et al., 2016).

This study proposes the identification and the evaluation of the risk areas.

The VCRO is constructed using a mathematical model based on the generic characteristics of ship–ship encounters. Based on the expert interviews, following factors are included in the model, (Zhang, et al., 2015):

a) The distance between the two ships.
b) The rate of change of the distance in the course of the encounter, determined by the relative speed of the two ships.
c) The relative orientation of the two ships, determined by the difference between their headings.

In this study, theoretically, the distance at time between two vessels can be calculated as the Equation
(1 based on Euclidian Distance.

$$D_{ij}^{t} = \sqrt{\left(x_i^t - x_j^t\right)^2 + \left(y_i^t - y_j^t\right)^2} \quad (1)$$

The invasion of ellipse is calculated by the Equation (2.

$$\frac{\left(\left(x_i^t - x_j^t\right) * \cos\alpha_i^t + \left(y_i^t - y_j^t\right) * \sin\alpha_i^t\right)^2}{(1.6 * L_i)^2} + \frac{\left(\left(x_i^t - x_j^t\right) * \sin\alpha_i^t - \left(y_i^t - y_j^t\right) * \cos\alpha_i^t\right)^2}{(4 * L_i)^2} \leq 1 \quad (2)$$

The approach presented in this paper to estimate the vessel collision risk based on AIS data is based on the domain theory presented by (Weng, et al., Proceedings of the 2014 Transportation Research

Board Annual Meeting) and (Weng , et al., 2012) proposed. At first, the relative speed of vessel i over vessel j denoted as $|\vec{v}_{ij}^{t}|$can be determined according to Equation (3, (Wu, et al., 2016).

$$|\vec{v}_{ij}^{t}| = \sqrt{\left(|\vec{v}_i^t|\cos\alpha_i^t - |\vec{v}_j^t|\cos\alpha_j^t\right)^2 + \left(|\vec{v}_j^t|\sin\alpha_i^t - |\vec{v}_j^t|\sin\alpha_j^t\right)^2} \quad (3)$$

And then the relative angle from $\vec{v}_i^t$ to $\vec{v}_i^t$ is in Equation (4.

$$\theta_{ij}^t = \cos^{-1}\left[1 - \frac{\left(\frac{|\vec{v}_i^t|\sin\alpha_i^t - |\vec{v}_j^t|\sin\alpha_j^t}{|\vec{v}_{ij}^t|} - \frac{x_j^t - x_i^t}{D_{ij}^t}\right)^2 + \left(\frac{|\vec{v}_i^t|\cos\alpha_i^t - |\vec{v}_j^t|\cos\alpha_j^t}{|\vec{v}_{ij}^t|} - \frac{y_j^t - y_i^t}{D_{ij}^t}\right)^2}{2}\right] \quad (4)$$

Then the following criteria can be used to categorize the type of vessel conflict:

- **Overtaking conflict** if $\sin\alpha_i^t * \sin\alpha_j^t > 0 \; and \; (\theta_{ij}^t \leq 10° \; or \; \theta_{ij}^t \geq 170°)$
- **Head-on conflict** if $\sin\alpha_i^t * \sin\alpha_j^t < 0 \; and \; (\theta_{ij}^t \leq 10° \; or \; \theta_{ij}^t \geq 170°)$
- **Crossing conflict** if $\theta_{ij}^t > 10° \; and \; \theta_{ij}^t < 170°$

Where α_i^t and α_j^t are courses of vessels i and j at time t, respectively, and θ_{ij}^t is the relative angle at time t. If one vessel is already within the domain of another at time t, a vessel conflict does happen.
The above rules can be also applied to elliptical domains.

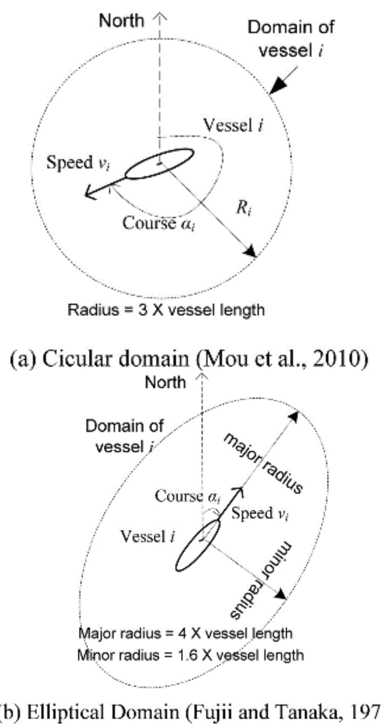

Figure 6 –Demonstration of a vessel's domain, (Wu, et al., 2016)

Figure 6 shows these two types of domains. It also demonstrates the speed, course, and domain of vessel. It seems that elliptical domains, focus more on conflicts between vessels which share similar or opposite courses, (Min Mou, et al., 2010).

3. Results

The AIS data captured by the antenna described in section 2.2 allowed plotting a high-resolution geographical characterization of traffic from January to April 2018, see Figure 7.

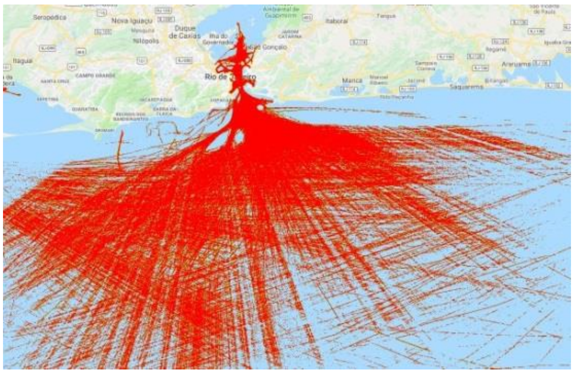

Figure 7 - Distribution of marine traffic around Rio de Janeiro (Guanabara Bay)

The distribution of marine traffic around Rio de Janeiro shown in Figure 7 check the large number of ships that sail in this area and the large amount of data that can be analysed for different purposes.

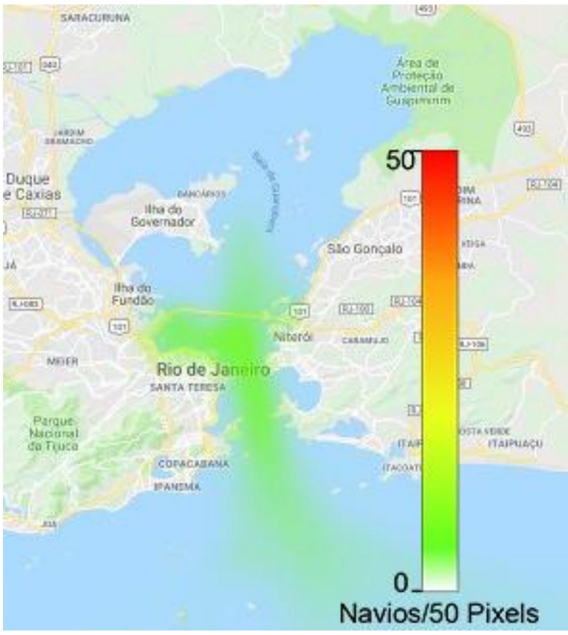

Figure 8 - Density of ships in the Guanabara Bay (relative distance of 50 pixels)

Figure 8 shows the Density of ships in the Guanabara Bay in this case each pixel calculates in its direct area of influence (circle of 50 pixels) the number of ships that is in that interval.

Nevertheless, the results show the existence of possible areas of risk, see Figure 8, as the follows:

- The transport of passengers in the route Rio de Janeiro – Niteroi with high speed passenger vessels
- The routes of logistic support and transportation of goods, tools, equipment and personnel to and from offshore oil platforms and other offshore structures by Platform supply vessels (PSVs)
- The potential risk of ship graveyards
- The entrance of the bay

The analysis of the near missed collision of the vessels using the previous mentioned methodology is plotted on Figure 9.

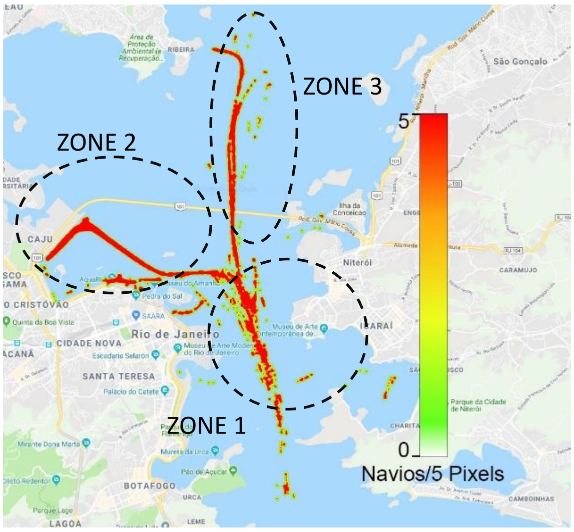

Figure 9 – Near missed collision in Guanabara Bay

In Figure 9, we identify three zones of conflict. The zone 1 representing the access to the Guanabara bay in the central zone. The zone 2 representing the access to the Rio the Janeiro Port (dry bulk, logistic and container terminals). Finally, the zone 3 that represent the access to the petroleum facilities passing under the bridge between Rio de Janeiro and Niteroi city.

After a deeper analysis, it has been discovered that the near missed collision plotted in the Figure 9, was taking into account the operation of the tug boats entering in the safety ellipse of the ships. It is obviously not a situation that it should be considered at risk. Therefore, the tug boats have been erased in order to plot the near missed collision of the ships in Figure 10.

Figure 10 shows the near missed collisions without tugboats. It can be observed that the majority of near missed collision in zone 2 disappeared. It was mainly due to the operation of the operation of the tugboats in this region. A similar behaviour is observed in zone 3. The density of near missed collisions observed in zone one remain constant that indicate that this area seems to be at a higher risk level. For this reason, the authors suggest that for a further investigation this area should be studied in priority.

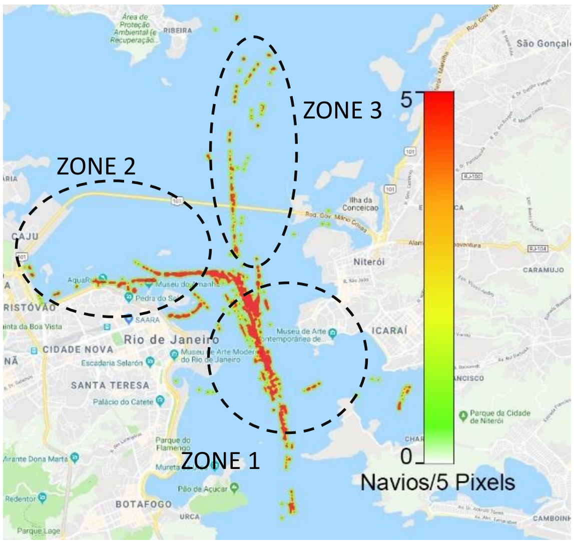

Figure 10 - Near missed collision in Guanabara Bay without the tugboats

During the four months of AIS data collection from January until April 2018, 2553 unique ships passed the area (without tugboats - 121 ships and pilot-boats – 7 ships). For any discrete moment of time, e.g., a whole minute, the traffic situation was analyzed and the distance and relative bearing from any vessel to the nearest target was calculated.

The density plot was created using the smallest distance between ships, i.e., for each register of the database the Euclidian distance between the reference ship and all the surrounding ships was calculated, see Equation(1). Only the registers with speed above 4 knots were considered in order to avoid interference in the graph due to barges and anchored vessels. Then, the smallest distance is selected and plotted within a radius of 2 kilometers, see Figure 11. It represents a total of 185239 points. The ship safety domain (ellipse) can be observed and present a 400 m length and 100 m width.

The type of vessel conflict, presented in Table 1 represents the quantity of ships in overtaking, crossing and head-on conflict.

Table 1 shows that the main type of conflict happens in Overtaking, the less in the Crossing category.

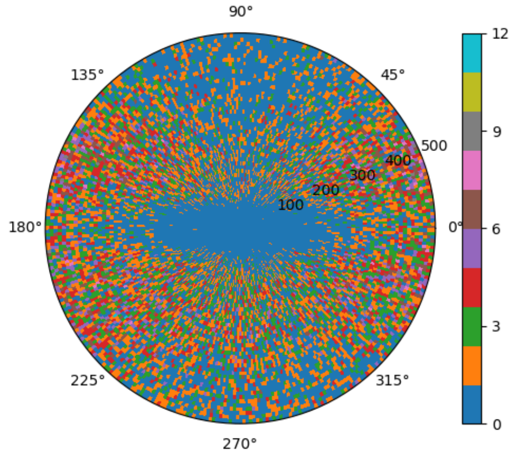

Figure 11 - Relative bearing and distance to all targets in Guanabara Bay without the tugboats and pilot boats.

Table 1 – The type of vessel conflict in Guanabara Bay without the tugboats and pilot boats

Classification of conflicts	Quantity
Total of conflicts	21128
Overtaking conflict	18881
Crossing conflict	922
Head-on conflict	1325

4. Conclusions and recommendations

The Rio de Janeiro Guanabara Bay, one of the busiest ports of Brazil, has a great environmental and socio-economic importance for the region of the study.

This study focus the assessment of the identification of collisions and near collisions zones due to marine traffic base on 4 months AIS data (January to April 2018).

The major findings of this study shows that the area that present crossed flow of ship entering in the bay with the high speed passenger ships operating between Rio de Janeiro and Niteroi present a higher risks of collision that the other parts of the area (zone 1).

This preliminary results need to be improved to better identify what are the root causes of that risk. We propose to improve the model by extending first the period of the study. Another important point consist to analyse the influence of certain type of specific operations such as Pilot boat interferences, PSVs that are in the anchorage areas, as well as the eventual seasonality (night/day) or possible correlations with the weather factors (fog, waves, wind).

The consequence of conflict scenario includes human life loss, property loss and environmental loss.

From the analyses, it appears that the ship domains seems to be elliptical shaped, about twice as long as it is wide, sometimes symmetrical around the own ship, sometimes shifted towards the bow of own ship.

The continuous monitoring of the bay trough AIS data is producing a consistent database that is enabling the future improvement of the model.

5. Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

6. References

Chen, D. et al., 2018. Contribution of ship emissions to the concentration of PM2.5: A comprehensive study using AIS data andWRF/Chemmodel in Bohai Rim Region, China. *Science of the Total Environment,* p. 1476–1486.

Fujii, Y. & Tanaka, K., 1971. Traffic Capacity. *The Journal of Navigation,* 24(4), pp. 543-552.

Goodwin, E. M., Lamb, W. G. P. & Kemp, J. F., 1983. Select Quantitative Measurements of Navigational Safety. *The Journal of Navigation,* 36(3), pp. 418-429.

IMO, 2003. *Guidelines for the Installation of a Shipborne Automatic Identification System (AIS),* London: IMO.

Kerbiriou, R., Lévêque, L., Rajabi, A. & Serry, A., 2017. *The Automatic Identification System (AIS) as a data source for studying maritime traffic.* Solin, Faculty of Maritime Studies, pp. 1-17.

Min Mou, J., van der Tak, C. & Ligteringen, H., 2010. Study on collision avoidance in busy waterways by using AIS data. *Ocean Engineering*, April, 37(5-6), pp. 483-490.

Miyake, R., Fukuto, J. & Hasegawa, K., 2015. Procedure for Marine Traffic Simulation with AIS Data. *the International Journalon Marine Navigation and Safety of Sea Transportation,* March, 9(1), pp. 59-66.

Szlapczynski, R. & Szlapczynska, J., 2017. Review of ship safety domains: Models and applications. *Ocean Engineering,* Volume 145, pp. 277-289.

UNCTAD, 2017. *Review of Maritime Transport.* Geneva: UNITED NATIONS.

Weng , J., Meng, Q. & Qu, X., 2012. Vessel Collision Frequency Estimation in the Singapore Strait. *The Journal of Navigation,* April, 65(2), pp. 207-221.

Weng, J., Meng, Q. & Li, S., Proceedings of the 2014 Transportation Research Board Annual Meeting. *Quantitative risk assessment model for ship collisions in the Singapore Strait.* Washington, DC, s.n.

Wu, X., Mehta, A. L., Zaloom, V. A. & Craig, B. N., 2016. Analysis of waterway transportation in Southeast Texas waterway based on AIS data. *Ocean Engineering,* July, Volume 121, pp. 196-209.

Xiao, F., Ligteringen, H., Gulijk, C. v. & Ale, B., 2015. Comparison study on AIS data of ship traffic behavior. *Ocean Engineering,* February, Volume 95, p. 84–93.

Zhang, L., Meng, Q. & Fang Fwa, T., 2017. Big AIS data based spatial-temporal analyses of ship traffic in Singapore port waters. *Transportation Research Part E: Logistics and Transportation Review.*

Zhang, W., Goerlandt, F., Montewa, J. & Kujala, P., 2015. A method for detecting possible near miss ship collisions from AIS data. *Ocean Engineering,* Volume 107, pp. 60-69.