Impacts of wind and current on ship behavior in ports and waterways

A quantitative analysis based on AIS data

Zhou, Yang; Daamen, Winnie; Vellinga, Tiedo; Hoogendoorn, Serge P.

DOI
10.1016/j.oceaneng.2020.107774

Publication date
2020

Document Version
Final published version

Published in
Ocean Engineering

Citation (APA)
Zhou, Y., Daamen, W., Vellinga, T., & Hoogendoorn, S. P. (2020). Impacts of wind and current on ship behavior in ports and waterways: A quantitative analysis based on AIS data. *Ocean Engineering*, *213*, Article 107774. https://doi.org/10.1016/j.oceaneng.2020.107774

Ocean Engineering 213 (2020) 107774

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Impacts of wind and current on ship behavior in ports and waterways: A quantitative analysis based on AIS data

Yang Zhou [a,*], Winnie Daamen [b], Tiedo Vellinga [a], Serge P. Hoogendoorn [b]

[a] Department of Hydraulic Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Netherlands
[b] Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Netherlands

ARTICLE INFO

Keywords:
AIS data
Ship behavior
External disturbance
Wind impact
Current impact
Ports and waterways

ABSTRACT

In ports and waterways, the impacts of external navigational factors may lead to serious incidents due to limited space for ship maneuvering. Using nautical traffic models, these incidents can be predicted in advance. In current studies of nautical traffic models, the impacts of wind and current on ship behavior are seldom considered when modeling the ship behavior in a port area. The numerical maneuvering models simulate the individual ship behavior under such impacts by calculating the hydrodynamic forces working on the ship's hull. However, the input, maneuvering particulars of individual ships, are not available in ports. In order to fill the knowledge gap of estimating ship behavior under external impacts without detailed ship maneuvering information, the impacts of wind and current on the observed dynamic ship behavior (speed over ground and leeway and drift angle) in ports and waterways have been investigated by analyzing Automatic Identification System data (showing ship paths over time) and the meteorological and hydrological data collected from the port of Rotterdam. The relation between unhindered speed variation and ship size is revealed. The regression analysis results on ships with similar size indicate the differences between wind and current impacts. Especially for small ships, the current impact on speed over ground outweighs the wind, while the wind influences the leeway and drift angle more than the current. Based on the quantified impact variation over ship size, the proposed impact mechanism explains the variance of speed over ground and leeway and drift angle. Some conventional sailing habits based on good seamanship, such as a series of small-angle alterations rather than direct turning at waypoints, are also revealed by the statistical analyses. Considering the variation of wind and current conditions in the study area, the analysis result provides generic quantitative insights into the wind and current impacts on the individual behavior of ships of different sizes. These mathematical formulations can be adopted in a microscopic nautical traffic model to include the impacts of external conditions.

1. Introduction

Seaborne transport has been an important means of international freight transport, which accounted for over 80 percent of the global trade by volume and more than 70 percent by value until 2017 and grew by another 4 percent in 2018 (United Nations Conference on Trade and Development, 2018). According to this forecasting of UNCTAD, the maritime trade is projected to expand at an annual growth rate of 3.8 percent between 2018 and 2023. Due to the large amount of cargo carried by individual ships and the high frequency of ships visiting the hub ports, nautical traffic safety in ports has been an important and sensitive issue for nautical traffic management and port authorities. Unlike the large space for ship maneuvering at sea, the maneuverability is restricted under different external conditions, such as strong wind and current, in ports and inland waterways. In such areas, the impacts of external navigational factors may lead to more serious consequences, such as grounding or collision with vast loss of life and property and damage to the environment and local infrastructure. Thus, the understanding of external navigational impacts on ship behavior in real-life situations will benefit the effective management of nautical traffic considering the external conditions in the seaports and inland waterways.

To analyze and simulate the maritime traffic considering individual ship behavior in an area, various models have been developed, which are compared by Zhou et al. (2019a) from the ship behavior modeling perspective. However, external conditions, such as wind and current,

* Corresponding author.
E-mail addresses: y.zhou_navi@outlook.com, Y.Zhou-5@tudelft.nl (Y. Zhou).

https://doi.org/10.1016/j.oceaneng.2020.107774
Received 10 December 2019; Received in revised form 20 May 2020; Accepted 6 July 2020
Available online 21 July 2020

have seldom been considered in the models, even though it has been proven that the external factors do influence the ship behavior (Kepaptsoglou et al., 2015; Shu et al., 2017; Zhou et al., 2017). In the numerical models to simulate individual ship behavior considering the specific maneuverability, the effects of wind, tidal current and waves on moving ships are significant due to the hydrodynamic forces and moments working on the ship's hull (Chen et al., 2015; Soda et al., 2012). Using the detailed data of individual ship maneuverability, the wind, wave and current forces are estimated, which provides an accurate prediction of the ship behavior under the external conditions. In the nautical traffic models considering external environmental factors, two methods have been adopted to indicate such impacts. The simplified method is to introduce random variables (Qi et al., 2017; Qu and Meng, 2012) or generic rules (Almaz et al., 2006; Camci et al., 2009; Merrick et al., 2003). It shows a generic random variation of ship movement under external impacts, which is feasible when describing the traffic flows at an aggregated level. However, the corresponding mechanism of such impacts is not included when investigating the behavior of individual ships. On the contrary, the other method is to consider the maneuverability of each individual ship under specific wind and current conditions to model the corresponding behavior (Beschnidt and Gilles, 2005; Leguit, 1999; Sarıöz et al., 2002). This method specifies the hydrodynamic processes and requires maneuvering particulars for specific ships for model calibration, which cannot be used for simulation of generic nautical traffic in an area. Therefore, in the field of nautical traffic modeling considering individual ship behavior, the research gap is that neither method can be applied to model the wind and current impacts for different ships in an area where the maneuverability particulars of individual ships are unavailable.

In order to investigate ship behavior, Automatic Identification System (AIS) data have proven to be a valuable source (Yang et al., 2019). To analyze macroscopic navigation patterns or the nautical traffic characteristics in an area, AIS data are widely used due to its detailed record of behavior for almost all passing ships (Altan and Otay, 2017; Gao et al., 2017; Gunnar Aarsæther and Moan, 2009; Silveira et al., 2013). To analyze the safety distance during collision avoidance considering ship drift, the external forces can be considered by the difference between heading and course in AIS data (Altan, 2018). Since AIS equipment has been mandatory for most of the ships, the data can be obtained by all port authorities, which can be utilized to analyze the behavior in a port or other area. In this research, AIS data is collected to describe the ship behavior under different external conditions (including wind, current and visibility). Thus, the meteorological and hydrological data are also collected. By comparing the average behavior in hindered and unhindered behavior, the impacts of external factors including wind, current, visibility and encounters are found (Shu et al., 2017). Combining AIS data with meteorological and hydrological data, the impacts of wind and current on ship behavior have been qualitatively analyzed and presented (Zhou et al., 2017). It shows that both ship speed and lateral position in a waterway are affected by wind and current, where the wind and current directions are categorized into four directions, being from the bow, the stern, the port side and the starboard side. The impacts found are different for ships of different sizes. However, the qualitative analysis results cannot be used to estimate the behavior of different ships under different wind and current conditions in an area.

The aim of this paper is to quantitatively analyze and estimate the impacts of external conditions (wind and current) on ship behavior in ports and waterways, where the actual hydrodynamic forces cannot be calculated due to the unavailability of individual ship particulars. To focus on the wind and current, the impacts of visibility and ship encounter are eliminated by filtering the external navigational conditions. Based on the previous qualitative analysis results and the theory in dead reckoning to estimate ship position, a generic modeling paradigm of wind and current impacts on ship behavior is introduced. Using the AIS data and the meteorological and hydrological data in the same period, a regression analysis is performed to quantify the external navigational impacts on ship behavior (expressed by speed over ground, leeway and drift angle) as a function of the ship's own size and the wind and current conditions. The originality of this research is to reveal the mathematical formulations of ship behavior under the wind and current impacts without detailed ship particulars. These mathematical formulations can thus be used in a microscopic nautical traffic model to include the impacts of external conditions. It also provides the port authority with an insight into relations between ship behavior and external factors.

Based on this aim, the following research questions are proposed:

- Research question 1: What is the mathematical relation between the variation in speed over ground and ship size in unhindered situation?
- Research question 2: What is the mathematical formulation of ship behavior under the wind and current impacts considering ship size differences without individual ship maneuvering particulars?

In this paper, the research area and the collected data set are introduced in Section 2. Section 3 explains the behavior variables and the proposed research approach. The analysis results for wind and current impacts are presented in Sections 4. Section 5 concludes the paper with discussion and recommendations for further research.

2. Study area and data description

In this section, the study area is introduced in section 2.1, followed by the description of the data, including AIS data in section 2.2 and meteorological and hydrological data in section 2.3. These data have been collected for the whole year of 2014 by the port authority of Rotterdam. The AIS data reveal the ship behavior in the study area. Regarding external conditions, the meteorological data describe the condition of visibility and wind, and the hydrological condition is represented by current velocity.

2.1. Study area

The study area is a nearly straight waterway, Nieuwe Waterweg, located at the entrance of the port of Rotterdam, the Netherlands, as shown in Fig. 1. The reason for choosing an almost straight waterway for external impacts analysis is to eliminate the impact of a specific waterway layout on ship behavior. In a curved waterway, besides the impact of more complex current conditions due to the curve, the bridge team on board also needs to hold the ship position to follow the direction of the waterway. It leads to a large variation of ship behavior due to the maneuvering habits of individual officers when passing a curve. Thus, the impacts of wind and current are hardly separated from the resulting trajectories in a curved waterway, and a straight waterway is preferable to focus on such impacts. However, the study area is not exactly straight with parallel banks on both sides. The impacts of such slight bending waterway layout may still exist in the analysis results, but are considered to be negligible. The total direction changes of the waterway stretch in the study area is about 2°. The length of the study area is 2300 m, and its width is about 650 m. The changes the bridge team has to make to follow the waterway layout are therefore assumed to be negligible, and all changes visible in the trajectory are attributed to the external conditions. The traffic in the Maasgeul channel (see Fig. 1) splits into Nieuwe Waterweg and Calandkanaal, which are physically separated by a slightly bent mole, named the Splitsingsdam.

2.2. AIS data

In this research, AIS data is used to describe the ship behavior under different external conditions. The Automatic Identification System is an automated tracking system onboard ships to automatically transmit information about the ship to other ships and coastal authorities. In 2000,

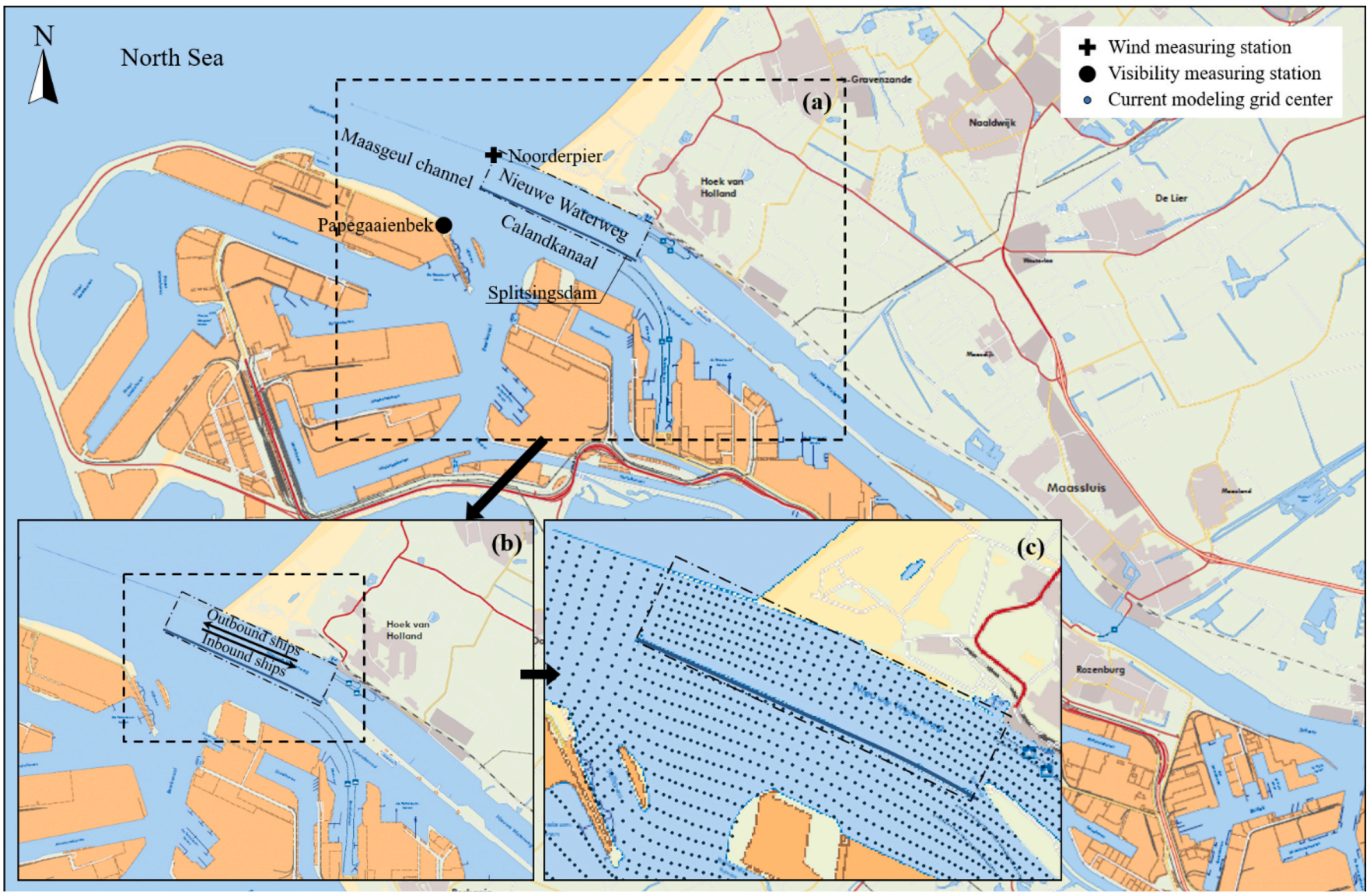

Fig. 1. (a) Location of the study area in the port of Rotterdam with the meteorological data measuring stations; (b) Zoom-in view of the study area with the sailing direction of the ships; (c) Zoom-in view of the hydrological data modeling grid map in the study area.

the International Maritime Organization (IMO) issued an amendment adopting a new requirement regarding the introduction of AIS system in the International Convention for the Safety of Life at Sea (International Maritime Organization, 1974). By the end of 2004, the AIS system was mandatory for all ships of 300 Gross Tonnage (GT) and more engaged on international voyages, cargo ships of 500 GT and more not engaged in international voyages and all passenger ships irrespective of size. Inland ships, both commercial and recreational, and sailing vessels longer than 20 m are mandatory to use AIS since December 1st, 2014 according to the resolution by Central Commission for the Navigation of the Rhine (2013). The resolution applies to most of the inland vessels in the Netherlands. In the study area, all seagoing ships including the ships below the GT limit of IMO regulation have installed AIS equipment and used it all the time as required by the local port authority. Since the year 2014 is a transition period, the majority of the collected AIS data of 2014 in the study area are seagoing ships. The collected AIS data in the study area contain 415,121 messages (inbound 4300 ship trajectories by 215, 926 messages, outbound 4732 ship trajectories by 199,195 messages). However, the exact number of missed inland ship trajectories in AIS data can hardly be estimated. There could be some inland ships without AIS equipment sailing in the area without record in the data, which may affect the analyzed ship behavior. The focus of this analysis remains to be seagoing ships recorded in the collected AIS data. One of the possible reasons that the data set contains less AIS messages for outbound ships, while there are more outbound ships is the different reporting interval of ships at different speed. Part of the outbound ships will take a left turn directing to Calandkanaal, and the speed will be low with longer reporting intervals and thus less AIS messages compared to other ships.

According to the guidelines by International Maritime Organization (2003), the AIS data contain three types of information: (1) static information (Maritime Mobile Service Identity number, IMO number, ship name, radio call sign, ship type, overall length, beam, etc.); (2) dynamic information (UTC time, ship position, speed over ground (SOG), course over ground (COG), heading, navigational status, etc.); and (3) voyage-related information (draught, destination, etc.).

The collected data set in text file has been processed from the original messages by an institute authorized by the local port authority. The data processing includes the data formatting and the combination of AIS data, radar data and the ship information in the system of IVS (Informative Verwerkend System in Dutch, Information Processing System in English). The document of data processing between the institute and the port of Rotterdam is not released in public. Thus, no reference can be cited due to confidentiality agreement. The format of the collected data after the official processing are listed in Table 1.

The static information is entered into the AIS system by the equipment provider when the equipment is initially installed or after a major change of the ship structure. According to the study on AIS data reliability by Harati-Mokhtari et al. (2007), MMSI number, ship name and call sign are fully correct for all ships. To ensure the reliability of the ship identity information, for the collected data set, the MMSI number and the IMO number of the ships have been checked with the identification in the system of IVS. Besides, when a ship enters the port, a temporary track number for the voyage is marked by the local authority. Together with this track number, the trajectories are uniquely identified in the data set. However, the information inconsistency problem of vessel type occurs in most of the ships, while the information of length and beam is mostly reliable.

The dynamic information is automatically updated based on the sensor data. In the collected data set, the x-position, y-position and heading values from the sources of Radar and AIS are checked, while

Table 1
Fields and format of the collected AIS data set processed by the authorized institute.

Field	Format/Unit	Range	Remarks
UTC-time	string with format 'dd-mm-yyyy hh:mm:ss'	–	
sensor type	–	Radar only, AIS only, Combined	
AIS IMO number	string	arbitrary number of digits	Sensor type ≠ 'R'
AIS MMSI number	string with 9 digits	000000000–999999999	Sensor type ≠ 'R'
AIS vessel type	string	arbitrary number of printable characters	Sensor type ≠ 'R'
AIS length	decimeter	0–10000	Sensor type ≠ 'R'
AIS beam	decimeter	0–10000	Sensor type ≠ 'R'
AIS x-position	meter (Rijksdriehoek)	−14000–113999	Sensor type ≠ 'R'
AIS y-position	meter (Rijksdriehoek)	381000–508999	Sensor type ≠ 'R'
AIS COG	degree	0–359	Sensor type ≠ 'R'
AIS SOG	centimeter/second	−2500–2500	Sensor type ≠ 'R'
AIS heading	degree	0–359	Sensor type ≠ 'R'
AIS draught	decimeter	0–1000	Sensor type ≠ 'R'
AIS nav. status	string	arbitrary number of printable characters	Sensor type ≠ 'R'
track number	–	1–10000	marked by local authority
length in IVS	decimeter	1–10000	fully identified in IVS
beam in IVS	decimeter	1–10000	fully identified in IVS
draught in IVS	decimeter	1–1000	fully identified in IVS
x-position	meter (Rijksdriehoek)	−14000–113999	
y-position	meter (Rijksdriehoek)	381000–508999	
speed	centimeter/second	−2500–2500	
heading	degree	0–359	

COG only derives from AIS data. According to the technical recommendations by International Telecommunication Union (2014), the precision of COG is 0.1°, while 1° for heading. However, as can be seen from Table 1, the precision of both COG and heading is 1° in the data set. In this research, for each message, the value of heading is deemed reliable when the data are consistent from the two sources, while the value COG is adopted when the other dynamic information are all consistent. As indicated by the International Telecommunication Union (2014), the reporting interval depends on the ship speed and course alteration. For the ships in the study area with small course changes, the time interval is 6 s when speed is larger than 14 knots (7.2 m/s), and 10 s for ships at speed lower than the value. The dynamic trajectories of inbound and outbound ships in the study area are illustrated in Fig. 2. Besides the layout sketch, the bouys in the area are also marked. It can be observed that all ships sail within the boundaries marked by the buoys.

The voyage-related information should be manually updated to the real-time situation by the officers on board. The actual draught may indicate the loading condition of the ship, which affects the ship's maneuverability. However, in the collected data set, errors are found in the draught information. The draught of most ships in the data set is not updated on each voyage. For some ships, the value of draught equals the molded draught in the registration. Other ships are recorded with a draught of 0 m or an unreasonable small value in the data set. Bailey et al. (2008) also show that 31% of the investigated messages have obvious errors in draught information. It implies that the data of ship draught are not reliable, thus these are not included in the analyses of this paper. Reliable ship draught data would have indicated the water depth that a ship is involved in. Since the current direction and speed can be different over the water depth in tidal waterway, the impact of current actually working on ship's hull can be analyzed with the draught information.

2.3. Meteorological and hydrological data

To analyze the impacts of wind and current on ship behavior, the wind and current conditions during the sailing of the ships are needed, i.e. the velocity of both wind and current. Thus, the meteorological and hydrological data in the study area are collected.

The meteorological condition refers to wind and visibility. Both are measured at different stations in the study area (see Fig. 1). The wind velocity is measured at an interval of 5 min, while the visibility is measured every minute. In non-extreme weather conditions, there is no sudden change of wind within 5 min. Thus, the measuring frequency of the data is sufficient in presenting the external conditions. The wind and visibility can be deemed to be homogeneous for the whole area. In the study area, there are some artificial dunes and the storage tanks for LNG

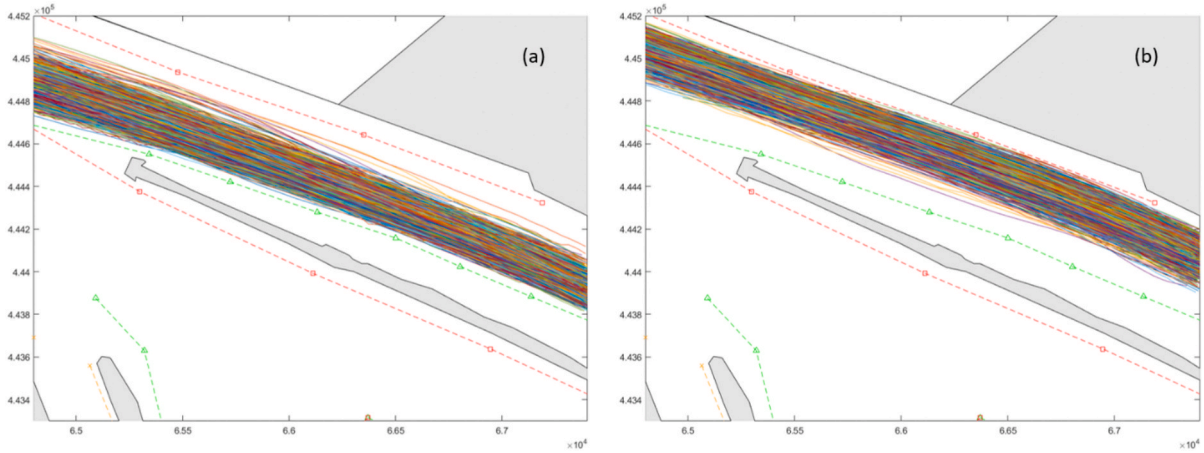

Fig. 2. (a) AIS trajectories of inbound ships; (b) AIS trajectories of outbound ships.

on the south side of the waterway, but without any high-rise buildings on land. Considering the scale and distance to the waterway, it is assumed that there is no impact on the wind and visibility in the area. Thus, the measuring data at a single station represents the wind and visibility conditions in the whole area. The wind direction probability in the study area is visualized by the wind rose diagram in Fig. 3. It can be observed that the wind direction is changing over the year, which is seldom parallel to the direction of the waterway (WNW/ESE). It means for ships sailing in the study area, besides the wind from the bow or stern direction, there is lateral forces on the ship hull by the crosswind for most of the time, which causes leeway in the observed ship behavior. The complexity of wind conditions is sufficient in the area to analyze the wind impact on ship behavior.

According to the collected data, the time with visibility distance less than 1000 m holds 0.52% of the year, while the frequency of visibility less than 2000 m is 4.87%. When the visibility is less than 2000 m, specific restriction measures are applied by the local port authority (Port of Rotterdam, 2014), e.g. entry restriction, specific traffic guidance by Vessel Traffic Service center, etc. Thus, the number of ship trajectories in restricted visibility is limited, and the reflected ship behavior involves the effects of local restrictions. We have removed the trajectories of the ships sailing in restricted visibility, as this paper focuses on the impacts of wind and current. The data of visibility will be used to filter the sailing situation under restricted visibility, i.e. to exclude the impact of visibility on ship behavior (speed) as revealed by other researchers (Shu et al., 2017).

The hydrological condition is represented by the velocity of the current in the waterway. The ship behavior is influenced via the hydrodynamic forces and moments working on the ship's hull under different current conditions. Unlike wind and visibility, the measured current velocity at a specific measuring station is not representative for the whole area, due to the propagation of flow and the velocity variation over the water-depth. Thus, the current velocity field is calculated by the port authority using the SIMONA model (Vollebregt et al., 2003) using the measured water level from eight stations around the port as input. The modeled velocity has been validated by comparing it to the measured velocity at one station in the area. The collected data describe the current velocity in 41 × 7 orthogonal curvilinear grids with a resolution of about 85 m (see Fig. 1). The current velocity in each grid cell is presented by 10 layers with the same depth averaged by the water depth of the grid at an interval of 15 min. For most of the ships, the length is larger than 85 m, so the grid resolution is sufficiently accurate. During each movement of the ships, the current velocity is instantly updated. The current velocity varies among grids and over water depth. The studied waterway links the inland waterway and the sea with natural physical boundaries on both sides (see Fig. 1), which is a tidal reach.

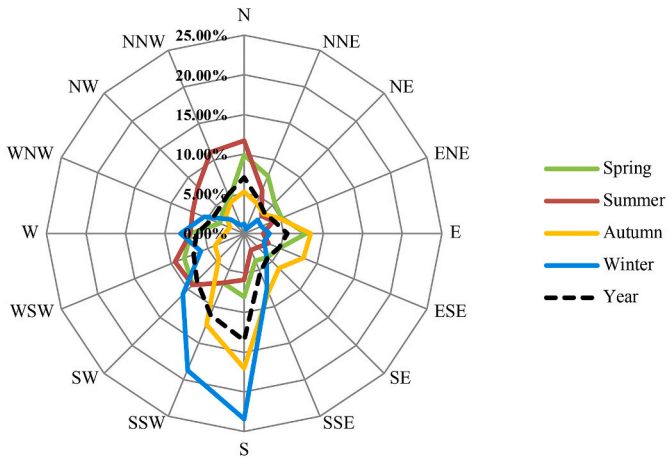

Fig. 3. Wind rose diagram of the study area.

Through the ebb and flood of the tides, the current directions in all grids at different water-depth do not always follow the sailing direction of the ships or the direction of the waterway.

3. Research approach

This research uses AIS data to statistically investigate the impact of wind and current on ship behavior via a regression analysis approach. In this section, the behavior variables in the AIS data and the wind and current directions are illustrated in the coordinate system. With an introduction of the underlying assumptions of the research, the data analysis method is explained in steps, including the data preparation and the approach to answer the two research questions proposed in Section 1.

3.1. Behavior variables and their coordinate system

The coordinate system to present dynamic ship motion is shown in Fig. 4. It consists of the space-fixed coordinate system $o_0 - x_0y_0$ and the moving ship-fixed coordinate system $o - xy$. Compared to the geographical coordinate system, the x direction points to the true North. The ship heading ψ is defined as the angle between x and x_0 axes.

The behavior variables discussed in this paper are the resulting behavior of all factors (see Fig. 4), rather than the ship maneuvering variables which are not known (e.g. rudder angle, and engine rate). Among the presented behavior variables, v_{SOG}, ψ and ϕ are directly collected from AIS data. The difference between ψ and ϕ is defined as γ, which is the leeway and drift angle indicating the angular deviation due to the external impacts. When the ship moves into the heading direction (i.e. $\phi = \psi$), γ equals to zero, which can happen in two situations. One situation is that the external conditions do not affect the ship behavior at

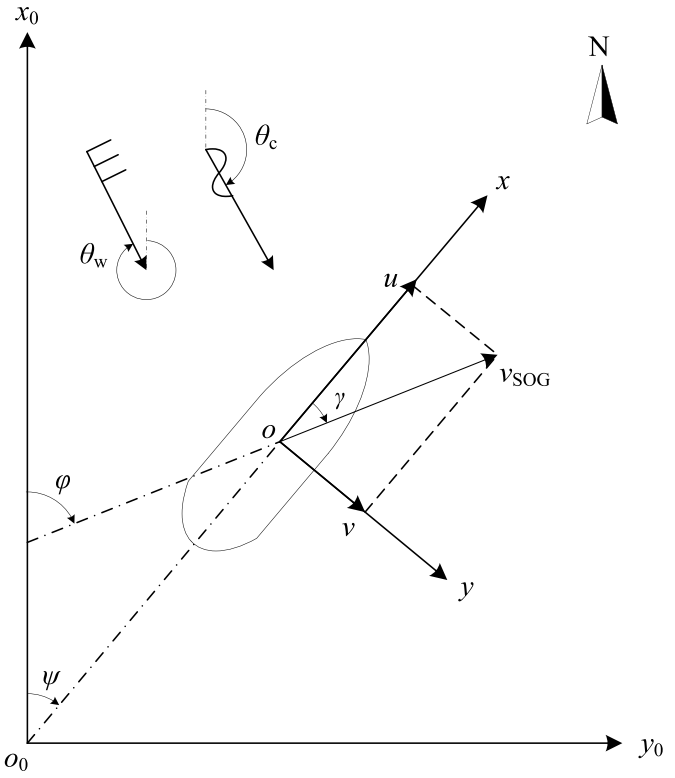

Fig. 4. Illustration of ship behavior variables and wind and current directions in the coordinate systems (v_{SOG}: speed over ground; ϕ: course over ground; ψ: heading; γ: leeway and drift angle; u: longitudinal speed; v: lateral speed; θ_w: wind direction in degrees as indicated by the arrow; θ_c: current direction in degrees as indicated by the arrow).

all. The other situation is that the different external impacts on ship behavior compensate each other, so the sum of directional impacts is zero. This way, the combination of v_{SOG} and γ can present the dynamic motion of a ship during sailing. Similarly, to directly represent the ship motion in longitudinal and lateral directions, the velocity components of v_{SOG} in x and y directions, namely u and v, can be calculated. These two variables in the ship-fixed coordinate system $o-xy$ directly describe the ship motion of surge and sway. During the data analysis, two sets of behavior variables (v_{SOG}, γ and u, v) have been tested. Both sets basically describe the same phenomenon of the ship motion, in which one is described in the space-fixed coordinate system, and the other in the ship-fixed coordinate system. The results are similar, and part of the corresponding results shown in Appendix. Thus, in this paper, only the results for v_{SOG} and γ, which are derived directly from AIS data, are explained in detail.

Besides the ship behavior variables, the directions of wind and current are illustrated in the coordinate system. According to common practice, the direction of wind θ_w describes which direction the wind is from, while the direction of current θ_c indicates the direction into which the water flows. The visibility is indicated by the visibility distance without specific direction indicated.

3.2. Assumptions and generic expression of the wind and current impact

In this paper, the following assumptions are applied to simplify the process.

- Besides the wind and current impacts analyzed in detail in this paper, the ship maneuvering in confined waterways and the human factors of the bridge team will affect the ship behavior variables described in Fig. 4. However, the impacts of human factors are not investigated in this research.
- It is assumed that the waterway layout and sailing direction (approach to or departure from a port) affect the ship behavior. The ships slightly change course to follow the waterway, and the inbound ships decelerate when approaching to the terminal. Thus, the behavior of inbound and outbound ships is separately investigated. However, these two factors are not quantitatively analyzed in this paper due to a lack of data on individual terminals of departure and destination.
- In unhindered situations, ships of similar size are assumed to maintain similar behavior considering the inertia of ships. The bridge teams onboard ships of similar size are assumed to take similar maneuvering decisions. Under the impacts of wind and current, the resulting behavior of such ships is assumed to be similar. Thus, ship size is the only internal factor to distinguish ships in this paper, irrespective of the maneuverability differences among individual ships.
- Without an encounter with other ships, the behavior of a ship in good visibility is assumed to be affected by the external factors of wind, current. The bridge teams are expected to take action based on the information of ship size, wind, and current, in line with good seamanship.

When considering the impacts of wind and current on ship speed, the linear combination form has been widely accepted for ship behavior modeling when considering ship as an integral rigid body and using the maneuvering particulars of individual ships (Beschnidt and Gilles, 2005; Yasukawa and Yoshimura, 2015). In their models considering such impacts, the mass or the weights regarding the under- and above-water parts of the hull for individual ships are needed to estimate the wind and current forces on the hull. The method of dead reckoning to estimate ship position is used to calculate the difference between heading and COG as the addition of leeway angle caused by wind and drift angle caused by current (Ni et al., 2010). Combining the above assumptions, a generic expression of speed over ground and leeway and drift angle under the impacts of wind and current can be formed as follows, while the detailed elaboration of each impact is given in Section 3.3.3.

$$v_{SOG} = f_{SOG,\ size}(s_s) + f_{SOG,\ wind}(v_w, \theta_w, \phi, s_s) + f_{SOG,\ current}(v_c, \theta_c, \phi, s_s) + \varepsilon_{SOG} \quad (1)$$

$$\gamma = f_{\alpha,\ wind}(v_w, \theta_w, v_{SOG}, \psi, s_s) + f_{\beta,\ current}(v_c, \theta_c, v_{SOG}, \psi, s_s) + \varepsilon_{\gamma} \quad (2)$$

where s_s denotes the size of a ship, the functions $f_{\text{behavior variable, factor}}$ explain the detailed impact mechanisms of each factor, γ is the sum of the leeway angle α and the drift angle β. ε_{SOG} and ε_{γ} are included in the equations to represent the behavior variation of individual ship due to the bridge team onboard. The bank effects on ship behavior or the proactive deceleration/acceleration when approaching/departing the terminals, are not considered either.

3.3. Data analysis method

The flow diagram in Fig. 5 illustrates the steps of the research approach, which are further explained in this section. The collected data are first processed to generate the data sets of ship behavior analyzed in this paper. Then, two phases of data analysis are developed to answer the research questions proposed in Section 1, respectively. The data set of unhindered behavior is used to explain the speed variation of ships due to the size differences. Based on this result, the impacts of wind and current are investigated using the whole data set of ship behavior.

3.3.1. Data preparation

Since the port authority of Rotterdam only stores the mandatory fields in AIS data, the ship size characterization is limited to length, beam, and draught. However, the information on draught is not reliable since too many errors are found. Thus, in this paper, only length and beam are adopted as the proxy for ship size.

During the data processing (the first step of the data preparation), the raw AIS data are filtered. This so-called data cleaning is performed using two steps. Firstly, the messages with sensor type marked as radar only are filtered, since there are no AIS information for these ships. The second step is to filter the messages with inconsistent information from Radar and AIS. During this step, the values of dynamic information are checked. The clean AIS data are linked with the meteorological and hydrological data based on the time and ship position in each AIS message. The drift angle γ is calculated for each AIS message. the resulting data set contains ship behavior in all external conditions.

As stated in Section 1, to focus on the impacts of wind and current, the impacts of visibility and ship encounter should be eliminated. In line with the preliminary analysis result, visibility $\geq$2000 m is defined as good visibility to avoid the impact of restricted visibility on ship behavior (Zhou et al., 2017). In order to exclude the impact of ship encounters, the processed data set has filtered the trajectories of ships with any encounter with another ship in the study area. The three types of encounters identified in the International Regulations for Preventing Collisions at Sea (COLREGs) are considered, namely head-on, overtaking, and crossing situations. If more than two ships are involved, the situation is deemed as a combination of several two-ships encounters of the above-mentioned types. The crossing situation can be easily distinguished by the relative position between ships, while the overtaking situation is characterized by the speed differences and position changes in between over time. A head-on situation at sea is defined when one ship is coming towards the other one roughly within 6° on either side of the heading. Considering the length and width of the waterway in the study area, the head-on situation is identified and filtered, when one ship encountering the other from the opposite direction in the study area. So far, the resulting data set includes all ship behavior in good visibility and without any encounter of another ship. The wind and current conditions are not used to filter any ship behavior data.

To further elaborate the behavior variation due to ship size, a data set

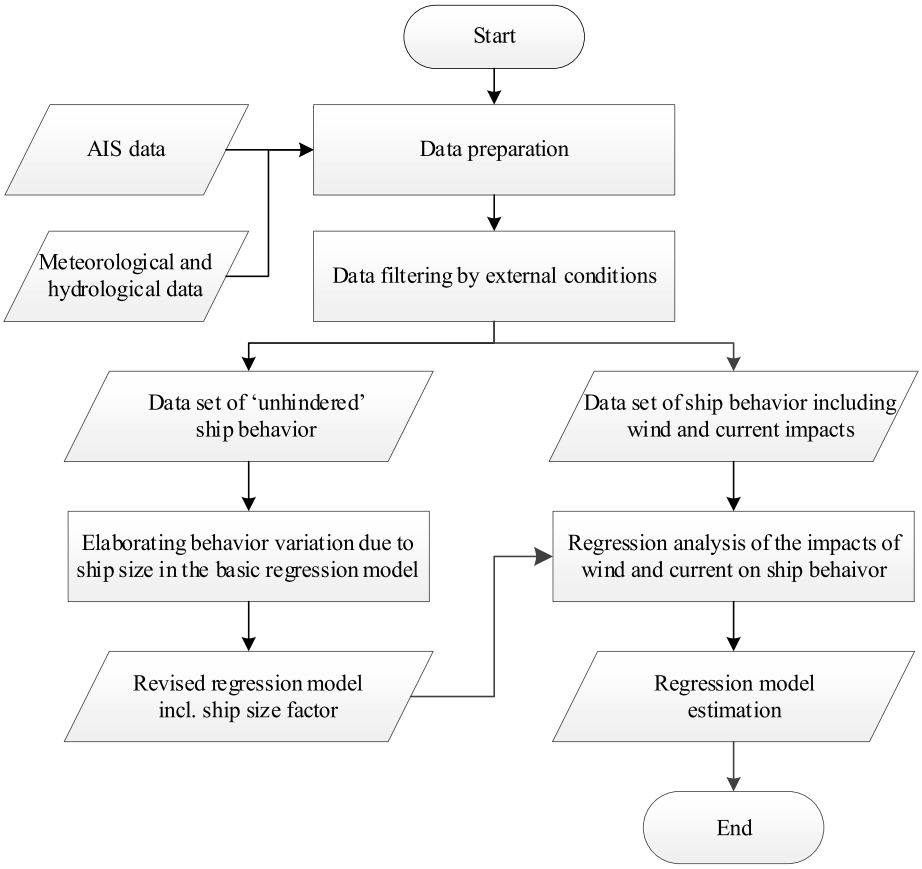

Fig. 5. Flow diagram of data preparation and data analysis.

in the unhindered situation is prepared. The thresholds have been previously analyzed, which characterize the situation by visibility ≥2000 m, wind speed <8 m/s (15.55 knots), and current speed < 0.37 m/s (0.72 knots) (Zhou et al., 2017). However, it should be noticed that the weak impacts of wind and current still exist in such a situation. Thus, when analyzing the impacts of wind and current, the whole data set including both hindered and unhindered situations will be used.

Since the waterway is not exactly straight with parallel banks on both sides, the impacts of the slightly bending waterway layout may affect the ship behavior, as stated in the assumptions. Besides, the sailing direction may influence the speed of a ship in the unhindered situation. Thus, the data set of inbound and outbound ships in the study area is separated and analyzed independently. By comparing the analysis results of these two data sets, it can prove whether the impacts of these two factors can be qualitatively proved.

3.3.2. *Variation of unhindered speed due to ship size*

In respect of the unhindered ship behavior, the speed variations among ships of different sizes have been observed (Shu et al., 2013; Zhou et al., 2019b, 2015). However, the direct relationship between ship size and SOG is still not revealed, which is represented by $f_{SOG,\ size}(s_s)$ in the generic expression of the impact mechanism. It is expected to find an isolate function to appropriately describe such a relationship using the data set of ship behavior in the unhindered situation. Thus, during the quantitative analysis of the wind and current impacts on ships of different sizes, the unhindered SOG of such different ships can be first estimated. The detailed steps to analyze speed variation due to ship size are present in Fig. 6.

Firstly, a correlation analysis between unhindered SOG and ship size (length and beam) is performed. It is to indicate the strength of correlation relationship in between and identify which size criterion better characterizes the ship behavior variation.

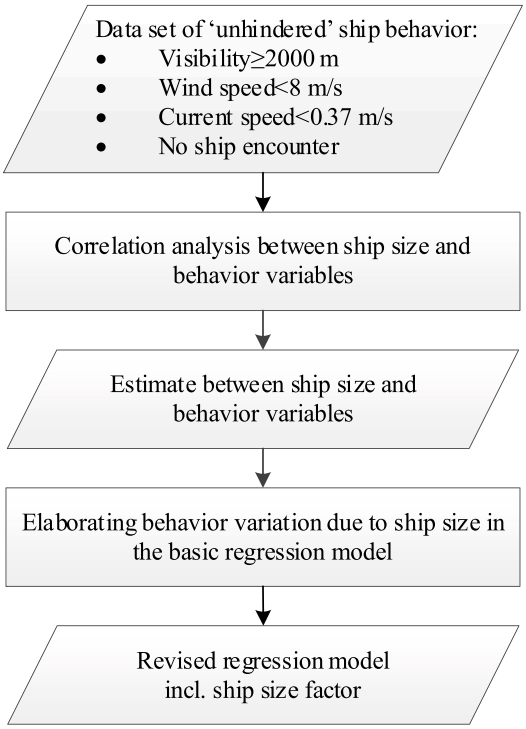

Fig. 6. Steps to analyze speed variation due to ship size.

Using the selected ship size criterion, the function to estimate the relationship with ship behavior is tested with monotonic elementary function types, which is considering the findings of behavior variation over ship size by Shu et al. (2013). The function type yielding the highest estimate result is adopted to describe the variation of unhindered ship behavior due to ship size. In case of the speed variation that is not monotonic as found in the preliminary analysis (Zhou et al., 2017), a piecewise function will be adopted for ships divided by the size threshold where the variation pattern changes.

The variation of SOG in the unhindered situation due to ship size can be described by $f_{SOG,\ size}(s_s)$ in Equation (1) using the determined function form. In the following section analyzing the wind and current impact, the variation due to ship size will also be considered.

3.3.3. Impacts of wind and current

To quantify the impacts of wind and current on different ships, three steps, in general, will be taken as shown in Fig. 7. Firstly, the functions to describe the mechanism of wind and current impacts in Equations (1) and (2) need to be determined, which specifies the form of regression models. Secondly, the regression analysis will be performed using the subsets of ship behavior with a similar ship size. The estimated results of all subsets will indicate whether the wind and current impacts vary among different sizes of ships. Finally, the overall functions to describe the wind and current impacts considering the variation pattern over ship size will be specified with coefficients estimated directly using the whole data set of ship behavior. In the following, each step will be elaborated upon.

Step 1. specifying the wind and current impact mechanism

As explained in the generic expression of the impact mechanism, the impacts of wind and current are assumed to be linear. This assumption could be tested by the calculation of hydrodynamic forces and moments working on the ship's hull. However, as the detailed ship particulars of each individual ship cannot be collected from AIS data, the wind and current impact mechanism are expressed using the generic ship size information and the wind and current velocity, as shown in Equations (3) and (4). Comparing to the estimate using specific information, such generic ship particulars may lead to a less accurate estimate result. But the method can be applied to estimate the ship behavior in a port, where the specific particulars for all visiting ships are unknown.

$$v_{SOG} = c_{SOG,\ s} \cdot f_{SOG,\ s}(s_s) + c_{SOG,\ w} \cdot f_{SOG,\ w}(s_s) \cdot (- v_w \cos(\phi - \theta_w)) + c_{SOG,\ c} \cdot f_{SOG,\ c}(s_s) \cdot (v_c \cos(\theta_c - \phi)) + c_{SOG} \tag{3}$$

$$\gamma = c_{\alpha,\ w} \cdot f_{\alpha,\ w}(s_s) \cdot \left(\frac{v_w}{v_{SOG}}\right)^2 \sin(\psi - \theta_w) + c_{\beta,\ c} \cdot f_{\beta,\ c}(s_s) \cdot \arcsin\left(\frac{v_c}{v_{SOG}} \sin(\theta_c - \psi)\right) + c_\gamma \tag{4}$$

where $f_{SOG,\ s}(s_s)$ denotes the variation of unhindered speed due to ship size, v_w, v_c and θ_w, θ_c describe the speed and direction of wind and current, the functions $f_{\text{behavior variable, factor}}(s_s)$ explain the variation of external impacts for different size of ships, the coefficients $c_{\text{behavior variable, factor}}$ will be estimated by the regression analysis. Since not all factors affecting ship behavior have been analyzed in the model, a constant c_{SOG} and c_γ has been added to each model to represent the impacts due to other unexplained factors.

The unhindered speed that a ship maintains when no effects of wind and current, is assumed to be affected by the ship size, which is analyzed in section 3.3.2. Regarding the external impact on ship's SOG, it is represented by the projection of wind/current velocity on the direction of v_{SOG} in Equation (3). The direction of velocity vector has been considered in the projection calculation.

Based on the theory of dead reckoning to estimate the ship position, the drift angle γ is the sum of the leeway angle α due to the wind and the

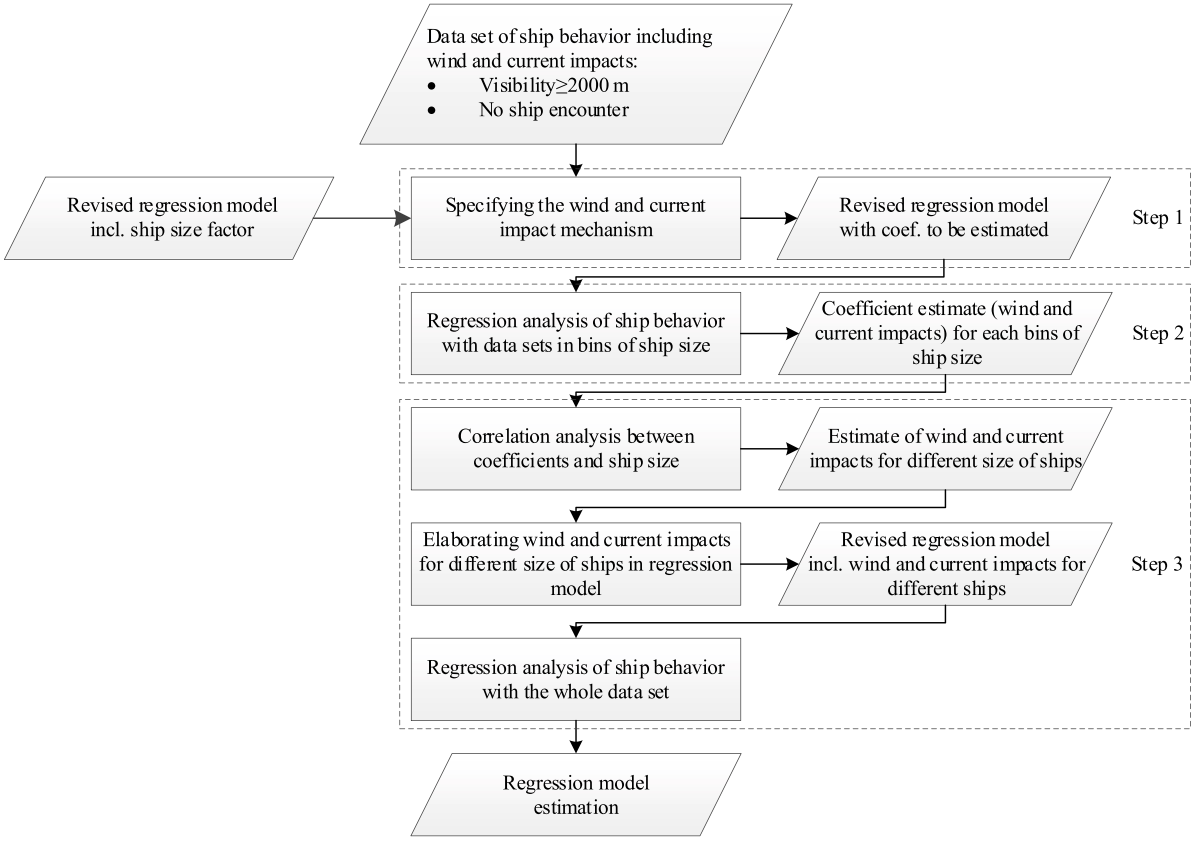

Fig. 7. Steps to estimate the wind and current impacts.

drift angle β for the current in Equation (4) (Bowditch, 2017). The leeway angle is calculated according to the empirical equation for water surface leeway analysis (Richardson, 1997). However, the coefficients are achieved by field experiments for specific physical objects, which can be only applied for specific circumstances. Thus, using AIS data combining the meteorological and hydrological data, the coefficients will be estimated by regression analysis. The obtained results can be applied to predict such impacts on the ship behavior in the area. The drift angle is calculated in the current triangle adopting the law of sines, using the angle between the current direction and heading, current speed, and SOG.

Step 2. estimating the impacts on ship behavior with similar size

In the previous analysis, the impacts on different sizes of ships are observed to be different as well (Zhou et al., 2017). But it is still unknown whether the cause of such differences is occasional fluctuation or due to the relationship with ship size. To answer this question, a quantified analysis is performed for ships in bins, which groups the ships with the same or similar size. The analysis results are compared to identify whether the impacts vary along with the change of ship size.

The whole data set is split into subsets of ship behavior according to the ship size. The variation of unhindered ship behavior due to ship size $f_{SOG,\ s}(s_s)$ has been revealed in section 3.3.1. The regression analysis will be performed based on Equations (5) and (6) for each subset of ship behavior data.

$$v_{SOG} = c_{SOG,\ s} \cdot f_{SOG,\ s}(s_s) + c_{SOG,\ w} \cdot (- v_w \cos(\phi - \theta_w)) + c_{SOG,\ c} \cdot (v_c \cos(\theta_c - \phi)) + c_{SOG} \quad (5)$$

$$\gamma = c_{\alpha,\ w} \cdot \left(\frac{v_w}{v_{SOG}}\right)^2 \sin(\psi - \theta_w) + c_{\beta,\ c} \cdot \arcsin\left(\frac{v_c}{v_{SOG}} \sin(\theta_c - \psi)\right) + c_\gamma \quad (6)$$

Compared to Equations (3) and (4), the functions to represent impact variation for different ships have been removed, since these models will be applied for ships with the same or similar size. *F*-test and *t*-test are used to determine the significance of the estimated models and the coefficients (a 0.05 significance level is adopted). The models are estimated using standardized scores (Z-scores) to obtain the standardized coefficients. The results of standardized coefficients of wind and current impacts can be compared within each subset. The comparison results present the weights of wind and current impacts on ship behavior for this size of ships.

Step 3. estimating the impacts on ship behavior for ships of different sizes

To identify whether the external impacts change along with the ship size or not, the correlation analysis is performed using the estimated coefficients of wind/current impacts with the average ship size of each bin. If the correlation is significant at the level of 0.01 (p-value), the impact of the external factor on the behavior variable is deemed as related to the ship size. If the correlation is not significant, it implies that the wind/current impact is not strongly correlated to the ship size. It can also be because that the local wind/current speed are quite small, or the ship size range is limited, the correlation in between cannot be revealed based on the data set. The function to estimate the relationship between external impact and ship size should be removed in Equations (3) and (4). For the impacts significantly correlated to ship size, the function type to describe the variation is determined by selecting the function yielding the highest estimate result.

With the estimated functions $f_{SOG,\ w}(s_s)$, $f_{SOG,\ c}(s_s)$, $f_{\alpha,\ w}(s_s)$, $f_{\beta,\ c}(s_s)$, the generic regression models of each behavior variable for all ships with different size are determined. The models will be estimated using the whole data set of ship behavior. The final estimated regression model explains the quantitative impacts of wind and current on ship behavior.

4. Results and discussions

Since the ship speed is influenced by the proactive maneuvering for approaching/departing a port (i.e. the inbound ships mostly decelerate, while the outbound ships accelerate) and the course is influenced by the waterway layout, the behavior of inbound and outbound ships is independently analyzed. During the process to determine the form of the regression model, the results are similar for both inbound and outbound ships. Only the results for inbound ships are presented and explained in detail. For the estimation of the final regression model, both results will be shown and the reasons for the differences are discussed.

4.1. Variation of speed due to ship size

In order to intuitively estimate the relationship between ship size and speed, the speed over ground v_{SOG} in the unhindered situation is visualized as a function of ship size in Fig. 8. In the boxplot, the distribution of ship speed within each bin is shown. It can be found for the first several groups of small ships and the last couple of large ship groups, the difference between the 25 and 75 percentile is rather small or rather large. This is because the number of ships within such groups are small, which leads to a large variation of the observations due to individual behavior differences. However, by comparing the median value of the bins, the overall variation pattern can be observed. For small ships, the speed increases when the ship size grows to maintain the maneuverability in the narrow waterway. For large ships, the value gradually decreases to a certain stable state when the size becomes larger, since the large ships cannot sail too fast in case of emergent maneuvering with big inertia. Thus, to use a single function describe the variation pattern is not feasible. In the previous sensitivity analysis for qualitative analysis, the threshold to distinguish small or large ships is 150 m for length and 23 m for beam, which also holds for the pattern shown in Fig. 8 (Zhou et al., 2017). It means the length or beam from AIS data can be used to categorize the ships as small or large ones. To further identify which of the ship size (length or beam) best describes the speed over ground, the correlation analysis between unhindered speed and ship size has been performed, as shown in Table 2.

In the data set, the ship length ranges between 24 m and 333 m, while the beam varies between 8 m and 60 m. According to the correlation analysis results, the ship beam is expected to better describe the relationship between v_{SOG} and ship size than ship length. Thus, in the remaining part of this paper, the beam is selected as the proxy for ship size during the quantitative analysis of wind and current impacts. To estimate the relationship between v_{SOG} and ship beam, four types of monotonic elementary functions have been tested for small and large ships, respectively. The estimated results are presented in Table 3. It should be noticed that the low R^2 value of the estimate result is due to ignoring other factors affecting ship behavior. Even in unhindered situations, there is still wind and current influencing ship behavior, and other unexplained factors as well.

Ideally, the function $f_{SOG,\ s}(s_s)$ to explain the relation between SOG and beam adopts the function type yielding the highest estimate result. However, it leads to different types of functions for small and large ships (logarithmic function for small ships and exponential function for large ships), which will result in different forms of the regression model for different ship sizes. The aim of this paper is to find a generic model form for all ships. Comparing the overall performance of four function types to different ships, the result of logarithmic function ranks the best for small ships and the second-best for large ships. When estimating the behavior for large ships, the difference of R^2 to the best function type (exponential function) is 0.003, which is acceptable. Thus, the logarithmic function is adopted to describe the relationship between v_{SOG} and the ship beam. The function $f_{SOG,\ s}(s_s)$ is included as $\log(B)$ in Equation (3).

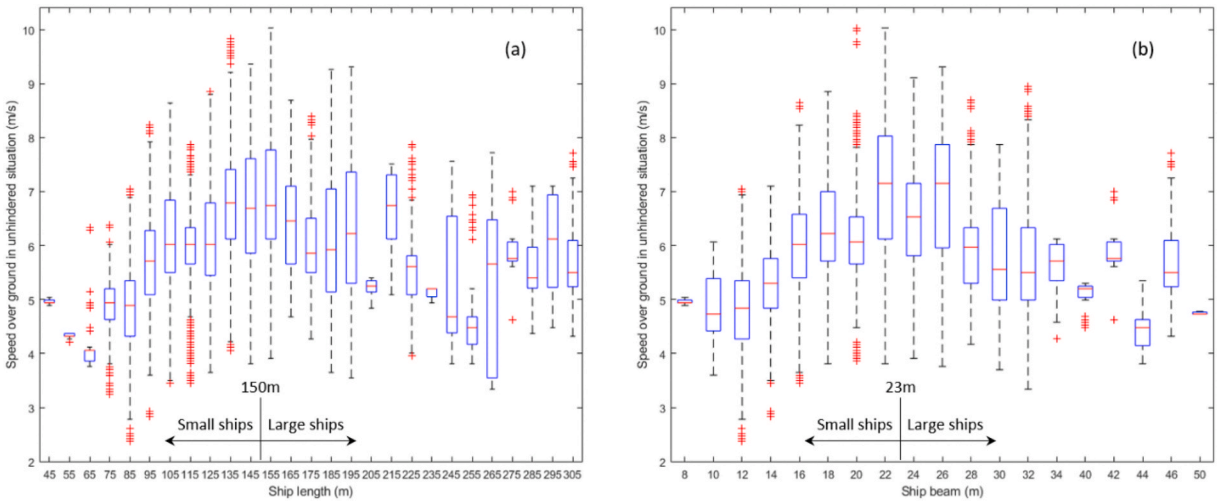

Fig. 8. Boxplot of speed over ground v_{SOG} in the unhindered situation as a function of ship length (a) and beam (b). (The line in the middle of each bin indicates the median value. The 25 and 75 percentiles of the values form the lower and upper boundaries of the box. The crosses outside the bin represent the statistical outliers.)

Table 2
Correlation analysis between unhindered speed over ground and ship size.

Behavior variable	Small ships		Large ships	
	Length (<150 m)	Beam (<23 m)	Length (≥150 m)	Beam (≥23 m)
v_{SOG}	0.580	0.660	−0.234	−0.404

** All correlations are significant at the 0.01 level of p-value (2-tailed).

Table 3
Comparison of estimated results through R^2 between unhindered speed over ground (v_{SOG}) and ship beam.

Type of function	Small ships (Beam < 23 m)	Large ships (Beam ≥ 23 m)
Linear: $v_{SOG} = c_{SOG,\ B} \cdot B + c_{SOG}$	0.435	0.163
Logarithmic: $v_{SOG} = c_{SOG,\ B} \cdot \log(B) + c_{SOG}$	0.437	0.167
Inverse: $v_{SOG} = c_{SOG,\ B} \cdot B^{-1} + c_{SOG}$	0.427	0.165
Exponential: $v_{SOG} = c_{SOG,\ B} \cdot \exp(B) + c_{SOG}$	0.423	0.170

* All results are significant at the 0.01 level of p-value.

4.2. Quantification of wind and current impacts on ship behavior

In this section, the regression analysis results of wind and current impacts on subsets of ships with similar size are firstly explained. The results quantitatively compare the wind and current impacts on similar ships and prove whether the impacts vary with ship size. Then the analysis results considering the external impact variation among the different sizes of ships are presented.

4.2.1. Wind and current impacts on similar-sized ships

The whole data set of ship behavior is split into subsets with the same or similar ship beam. The bin size is mostly set as 1 m, while for beams smaller than 10 m or larger than 32 m, the bin size is set as 5 m to include sufficient data (more than 30 ships) in each subset.

The regression models in Equations (5) and (6) for v_{SOG} and γ are estimated for each subset of ship behavior with similar beams. Since the speed of wind is much larger than current in measured values, the estimated unstandardized coefficients are not directly comparable. However, the standardized coefficients are estimated from the standardized regression analysis where the variances of variables are 1. They explain which of wind and current impacts have a greater influence on ship behavior in this multiple regression model. These coefficients of wind and current for the two behavior variables in each subset are presented in Fig. 9 and Fig. 10. As an example, similar results for surge and sway speed (u and v) are shown in Appendix.

Two comparisons are taken to interpret the estimated standardized coefficients. When looking at the standardized coefficients for similar size within the same subset, the weights of wind and current impacts can be compared. For small ships, the impact of current on v_{SOG} is dominant, compared to the impact of wind. But for large ships, the impact differences are becoming smaller. The speed of a ship is mainly provided by the propeller, which is underwater and affected by the current. For large ships with high superstructures, the wind area is also large. The wind impact may outweigh the current, while the difference is small. However, for the impact on leeway and drift angle α and β, the impact of wind is mostly larger than current. It means, in the port area, when the officers onboard change heading to prepare leeway and drift angle, the wind direction would be the primary factor for their decision. For large ships, the wind and current impacts are comparable, probably due to the large draught underwater.

The other comparison is between the coefficients for different groups of ships, which indicates the variation of the external impacts among ships of different sizes. It can be observed that for small ships, both impacts of wind and current on SOG decrease when the ships get larger. For small ships, the smaller size usually comes with smaller inertia to maneuver in emergent circumstances. Thus, those ships do not consume extra fuel to compensate the influences of wind and current. But the larger ships in this group need to keep their speed either for basic maneuvering requirements or for emergent maneuvering. However, both impacts on the drift angle slightly increase with the increase of ship size, which is due to the larger wind and current forces on ship hull and propeller. For large ships, the variation of wind and current impacts are not always the same. When the ships get larger, the impact of wind on SOG gradually increases, but the impact on the leeway angle fluctuates with a decrease. Because for very large ships, the wind forces on superstructures are large. Once the maneuvering requirement can be fulfilled, the ships will not spend extra effort (consumption of more fuel) to compensate such impact on speed. But for leeway angle, such large ships need to avoid collision with banks under the wind forces. Thus, the resulting impact on behavior seems smaller. Meanwhile, for the impact of current, the opposite relationship is presented. The impact of current on SOG for larger ships is smaller, and the impact on the drift angle is larger. The reasons are the same as for the impact of current on the group

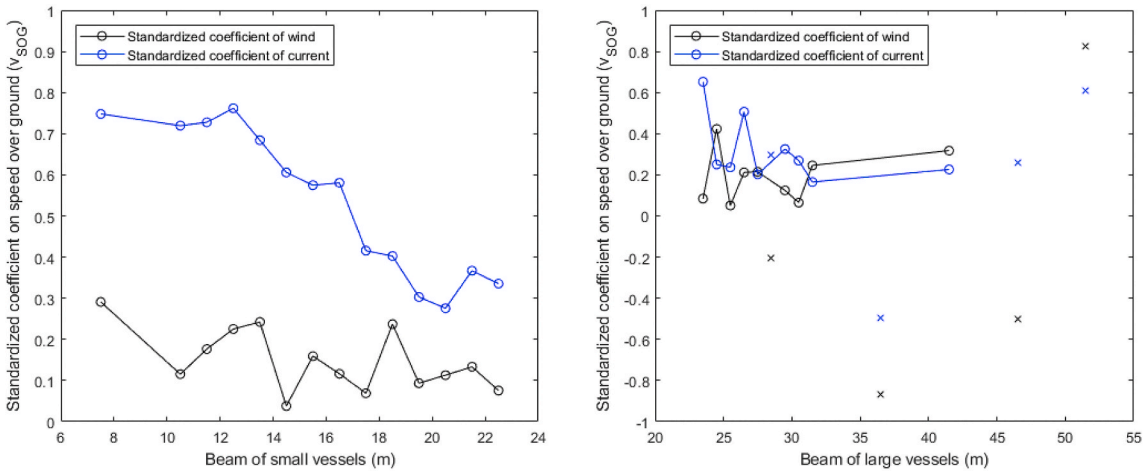

Fig. 9. Standardized coefficients of wind and current impact on speed over ground $c_{SOG,\ w}$ (in black) and $c_{SOG,\ c}$ (in blue) as a function of ship beam (The subsets containing less than 30 trajectories are marked as crosses in the right figure). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

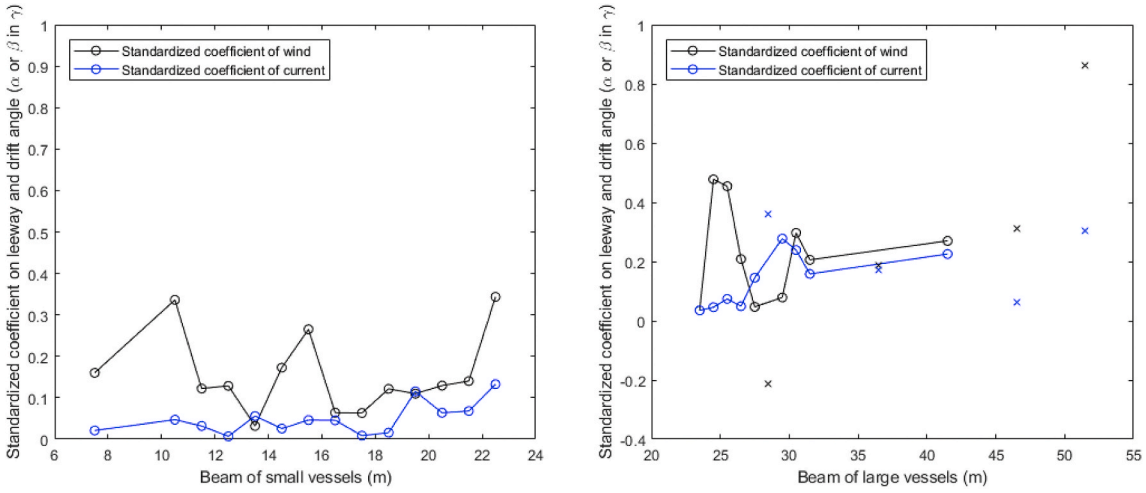

Fig. 10. Standardized coefficients of wind and current impact on leeway angle α and drift angle β, $c_{\alpha,\ w}$ (in black) and $c_{\beta,\ c}$ (in blue) as a function of ship beam (The subsets containing less than 30 trajectories are marked as crosses in the right figure). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

of small ships.

The regression analysis test on ship behavior with similar size proves the external impact variation, which is in line with the preliminary qualitative analysis result and follows our expectation (Zhou et al., 2017). The detailed quantification of such variation along with the size change will be determined by statistical analysis in the following section.

4.2.2. Wind and current impacts considering ship size variation

To figure out the impact factors significantly varied among different sizes of ships, the relationship between the unstandardized coefficients for wind/current impact and ship beam are statistically tested by correlation analysis. The correlation coefficients are listed in Table 4. The p-value of 0.01 is taken as the threshold of significant correlation. The positive values indicate positive correlations in between, while the negative coefficient refers to the negative correlation.

Table 4
Correlation analysis between unstandardized wind/current impact coefficients for v_{SOG}, α, β and ship beam for inbound ships.

	$c_{SOG,\ w}$	$c_{SOG,\ c}$	$c_{\alpha,\ w}$	$c_{\beta,\ c}$
Small ships	−0.529	−0.934**	0.326	0.729**
Large ships	0.317	−0.819**	−0.184	0.377

**: Correlation is significant at the 0.01 level of p-value.

From the statistical test perspective, for all behavior variables of both small and large ships, the impact of wind varies without a strong correlation with ship size. Such variation can be caused by behavior differences among individual officers on board or other unexplained factors. This way, the corresponding functions in the regression model to indicate such correlation are removed from models in Equations (3) and (4), including $f_{SOG,\ w}(s_s)$, $f_{\alpha,\ w}(s_s)$. However, for the impacts of current, the coefficients are significantly correlated to the ship's size, except for the impact on the drift angle for large ships. Therefore, the functions presenting the relationship $f_{SOG,\ c}(s_s)$, $f_{\beta,\ c}(s_s)$ need to be elaborated to quantify such impact differences. For the three significant correlations, the same four types of monotonic elementary functions have been tested for each coefficient, as presented in Table 5.

The ideal situation is to adopt a generic function form for all ships, the same as the function to describe the variation of speed due to ship size. However, the correlation between the current impact on drift angle and the size of large ships is not significant, while the one for small ships

Table 5
Comparison of estimate results R^2 for different functions between unstandardized wind/current coefficients and beam for inbound ships.

Type of function	Beam <23 m		Beam ≥23 m
	$c_{SOG,\ c}$	$c_{\beta,\ c}$	$c_{SOG,\ c}$
Linear	0.873**	0.531**	0.672**
Logarithmic	0.858**	0.491**	0.684**
Inverse	0.792**	0.432*	0.688**
Exponential	0.845**	0.363*	0.685**

**: The result is significant at the 0.01 level of p-value. The function yielding the highest estimate result is marked as grey.
*: The result is significant at the 0.05 level of p-value.

is significant (see Table 4). It leads to different forms of functions for small and large ships. Thus, the functions best describing the relationship are adopted in the regression models to consider the impact variation due to ship size. The corresponding functions in Equations (3) and (4) are elaborated for small ships and large ships, respectively, as follows. In the linear functions, two coefficients need to be estimated.

For small ships (ship beam <23 m),

$$f_{SOG,\ c,\ s}(s_s) = c_{SOG,\ c,\ s1} \cdot B + c_{SOG,\ c,\ s2} \tag{7}$$

$$f_{\beta,\ c,\ s}(s_s) = c_{\beta,\ c,\ s1} \cdot B + c_{\beta,\ c,\ s2} \tag{8}$$

For large ships (ship beam ≥ 23 m),

$$f_{SOG,\ c,\ l}(s_s) = c_{SOG,\ c,\ l} \cdot B^{-1} \tag{9}$$

So far, the regression models for v_{SOG} and γ to quantitatively analyze the wind and current impacts are generated for small and large ships. The impact variation among the different sizes of ships is also included. In this paper, the models are estimated based on the data set of inbound and outbound ship behavior separately. The estimation results of the regression analysis are shown in Table 6 and Table 7.

According to the regression analysis results, about 70% of the variance in SOG of small ships can be explained by the ship size, wind and current (R^2 is 0.743 for inbound ships and 0.696 for outbound ships). Comparing the standardized coefficients, the choice of SOG in unhindered situation accounts for the major weight of the final sailing speed, which is mostly determined due to the size of a ship. It means the ships only adjust their speed in unhindered situation when there are the impacts of wind and current. Looking at the unstandardized coefficients of unhindered SOG, the estimation results for inbound and outbound ships are different, which indicates the sailing direction also affects the speed choice. The impact of current on SOG of small ships (both inbound and outbound ships) outweighs the wind impact, which is the same when performing the regression analysis for ships in bins with similar size (see Fig. 9). However, for large ships, the explained variance drops to around 40% (R^2 is 0.395 for inbound ships and 0.440 for outbound ships). Two reasons may explain this result. Firstly, there is a large variation of SOG for large ships in the unhindered situation, even between the ships with similar size (see Fig. 8). The variance of speed by individual ships cannot be precisely predicted by the generic model. It also holds when comparing the results of inbound and outbound ships, in which the model performs better for outbound ships than inbound ships. When sailing in the study area, most of the outbound ships try to reach their desired speed at sea. However, the inbound ships need to decelerate or keep their speed depending on the distance to their destination terminal, which is different for individual ships. Thus, there is more variation in the choice of speed for inbound ships. The other reason is that the detailed impacts of wind and current on the speed of large ships need to consider the specific above- and under-water ship hull, which is hard to achieve for ships in an area.

Compared to the explanation of SOG, the wind and current impacts only account for 25% of the variance in leeway and drift angle, which seems that the relationship between external factors and the drift angle is not very strong. The standardized coefficients also indicate similar results. According to the ordinary practice of seamen, the set of leeway and drift angle is based on the surrounding sailing situation, including wind, current and waterway layout. However, the instant decision

Table 6
Estimation results of the regression model in final forms for the whole data set of inbound ship behavior.

		R^2	F-stat	Estimate	Std. error	t-stat	Std. estimate
Inbound Beam<23 m	$v_{SOG,\ s} = c_{SOG,\ B,\ s}\cdot\log(B) + c_{SOG,\ w,\ s}\cdot(-v_w\cos(\phi-\theta_w)) + (c_{SOG,\ c,\ s1}\cdot B + c_{SOG,\ c,\ s2})\cdot(v_c\cos(\theta_c-\phi)) + c_{SOG,\ s}$						
	Model	0.743	48099.002				
	$c_{SOG,\ B,\ s}$			3.383	0.009	368.112	0.625
	$c_{SOG,\ w,\ s}$			0.025	4.58e-04	54.063	0.092
	$c_{SOG,\ c,\ s1}$			−0.051	0.001	−42.471	−0.315
	$c_{SOG,\ c,\ s2}$			1.833	0.020	92.090	0.683
	$c_{SOG,\ s}$			−3.825	0.025	−151.407	
	$\gamma_s = c_{\alpha,\ w,\ s}\cdot\left(\frac{v_w}{v_{SOG}}\right)^2\sin(\psi-\theta_w) + (c_{\beta,\ c,\ s1}\cdot B + c_{\beta,\ c,\ s2})\cdot\arcsin\left(\frac{v_c}{v_{SOG}}\sin(\theta_c-\psi)\right) + c_{\gamma,\ s}$						
	Model	0.288	2860.903				
	$c_{\alpha,\ w,\ s}$			0.121	0.003	44.686	0.110
	$c_{\beta,\ c,\ s1}$			0.013	0.003	4.817	0.049
	$c_{\beta,\ c,\ s2}$			0.631	0.042	15.181	0.153
	$c_{\gamma,\ s}$			−0.136	0.007	−18.826	
Inbound Beam≥23 m	$v_{SOG,\ l} = c_{SOG,\ B,\ l}\cdot\log(B) + c_{SOG,\ w,\ l}\cdot(-v_w\cos(\phi-\theta_w)) + \frac{c_{SOG,\ c,\ l}}{B}\cdot(v_c\cos(\theta_c-\phi)) + c_{SOG,\ l}$						
	Model	0.395	3667.836				
	$c_{SOG,\ B,\ l}$			−2.832	0.031	−92.419	−0.349
	$c_{SOG,\ w,\ l}$			0.012	0.001	10.916	0.041
	$c_{SOG,\ c,\ l}$			14.861	0.271	57.742	0.207
	$c_{SOG,\ l}$			15.393	0.102	150.399	
	$\gamma_l = c_{\alpha,\ w,\ l}\cdot\left(\frac{v_w}{v_{SOG}}\right)^2\sin(\psi-\theta_w) + c_{\beta,\ c,\ l}\cdot\arcsin\left(\frac{v_c}{v_{SOG}}\sin(\theta_c-\psi)\right) + c_{\gamma,\ l}$						
	Model	0.254	2358.404				
	$c_{\alpha,\ w,\ l}$			0.121	0.003	37.017	0.149
	$c_{\beta,\ c,\ l}$			0.412	0.013	30.914	0.124
	$c_{\gamma,\ l}$			−0.561	0.008	−69.248	

* The estimated coefficients with different signs compared to outbound ships are marked as grey.

Table 7
Estimation results of the regression model in final forms for the whole data set of outbound ship behavior.

		R^2	F-stat	Estimate	Std. error	t-stat	Std. estimate
Outbound Beam<23 m	$v_{SOG,\ s} = c_{SOG,\ B,\ s}\cdot\log(B) + c_{SOG,\ w,\ s}\cdot(-v_w\cos(\phi-\theta_w)) + (c_{SOG,\ c,\ s1}\cdot B + c_{SOG,\ c,\ s2})\cdot(v_c\cos(\theta_c-\phi)) + c_{SOG,\ s}$						
	Model	0.696	33215.208				
	$c_{SOG,\ B,\ s}$			2.549	0.009	288.901	0.554
	$c_{SOG,\ w,\ s}$			0.033	4.46e-04	74.203	0.142
	$c_{SOG,\ c,\ s1}$			−0.032	0.001	−27.083	−0.233
	$c_{SOG,\ c,\ s2}$			1.502	0.020	75.423	0.649
	$c_{SOG,\ s}$			−0.525	0.024	−21.501	
	$\gamma_s = c_{\alpha,\ w,\ s}\cdot\left(\frac{v_w}{v_{SOG}}\right)^2\sin(\psi-\theta_w) + (c_{\beta,\ c,\ s1}\cdot B + c_{\beta,\ c,\ s2})\cdot\arcsin\left(\frac{v_c}{v_{SOG}}\sin(\theta_c-\psi)\right) + c_{\gamma,\ s}$						
	Model	0.216	1761.709				
	$c_{\alpha,\ w,\ s}$			0.173	0.014	12.093	0.132
	$c_{\beta,\ c,\ s1}$			0.017	0.008	1.291	0.034
	$c_{\beta,\ c,\ s2}$			0.610	0.072	3.017	0.133
	$c_{\gamma,\ s}$			0.109	0.017	3.989	
Outbound Beam≥23 m	$v_{SOG,\ l} = c_{SOG,\ B,\ l}\cdot\log(B) + c_{SOG,\ w,\ l}\cdot(-v_w\cos(\phi-\theta_w)) + \frac{c_{SOG,\ c,\ l}}{B}(v_c\cos(\theta_c-\phi)) + c_{SOG,\ l}$						
	Model	0.440	4629.630				
	$c_{SOG,\ B,\ l}$			−2.617	0.027	−95.268	0.207
	$c_{SOG,\ w,\ l}$			0.014	0.001	15.166	0.031
	$c_{SOG,\ c,\ l}$			16.228	0.247	65.666	0.247
	$c_{SOG,\ l}$			15.391	0.091	168.242	
	$\gamma_l = c_{\alpha,\ w,\ l}\cdot\left(\frac{v_w}{v_{SOG}}\right)^2\sin(\psi-\theta_w) + c_{\beta,\ c,\ l}\cdot\arcsin\left(\frac{v_c}{v_{SOG}}\sin(\theta_c-\psi)\right) + c_{\gamma,\ l}$						
	Model	0.236	2071.510				
	$c_{\alpha,\ w,\ l}$			0.379	0.016	22.953	0.125
	$c_{\beta,\ c,\ l}$			0.241	0.050	4.765	0.110
	$c_{\gamma,\ l}$			0.259	0.026	10.145	

* The estimated coefficients with different signs compared to inbound ships are marked as grey.

differs among individual officers onboard, considering their sailing habit and experience. In the same situation, some officers may take several degrees of drift angle, but the other officers may keep sailing without heading change. Even with the same wheel order, the observed results of leeway and drift angle (the difference between heading and COG) still depend on the rate of turn of individual ships. Besides, the precision of both heading and COG are the same as 1° in the collected data after the official data processing. It can happen that the values are the same while there exist a small difference in actual situation. Or the leeway and drift angle is calculated where the actual difference is quite small at a precision of 0.1°. Therefore, the explanation of the variance in γ is not as good as SOG.

The signs of the coefficients together with the functions in the regression model explain the relationship between behavior variables and the impact of ship size, wind and current. The results prove that the theoretical expression of the impact mechanism and the revealed impact variation over ship size by analyzing the subsets of data are correct, when analyzing the whole data set consisting of ships with different sizes.

Comparing the weights of wind and current impacts by standardized coefficients in Tables 6 and 7, the result for the whole data set is similar to the regression analysis using the subsets of inbound ship behavior with the same size (see Figs. 9 and 10), which follows our expectation. Regarding the impacts on SOG, the current impact outweighs the wind. However, the impact of current on drift angle is slightly larger than the wind on leeway angle. For the current impact on small ships, $c_{SOG,\ c2}$ and $c_{\beta,\ c2}$ represents the generic impacts, while $c_{SOG,\ c1}$ and $c_{\beta,\ c1}$ indicates the variation due to the size differences. Comparing their standardized estimates, the weight of the direct impact of the current itself is larger than the correction regressor for ship size.

The constant in the model of SOG plays a dual role. On the one hand, it corrects the unhindered speed due to ship size. Besides, it includes the other impacts of unexplained external factors. Regarding the constant for γ, it is expected to be zero in an ideal sailing situation without wind and current when sailing in the straight waterway. However, the estimated results are with different signs for inbound and outbound ships, which are marked as grey in Tables 6 and 7. The main reason is that the study area is not exactly straight with parallel banks (see Fig. 1). For inbound ships, the waterway slightly bends to the starboard side. The negative sign of c_γ indicates that the heading of a ship directs to the starboard side comparing to the $o-x$ direction of the ship-fixed coordinate system. It follows good seamanship that a ship will take a series of small-angle alteration to follow the designed route, rather than a sharp turning at the waypoint considering the ship maneuverability. For outbound ships, the positive sign represents the turning direction to the port side to follow the layout of the waterway. It also proves that the leeway and drift angle of a ship in inland waterways is affected by the bank besides the wind and current impacts. The bending direction of the waterway indicates the sign of the coefficient.

The estimated regression model provides a quantification of the wind and current impacts. Some behavior following good seamanship is also statistically revealed by the estimated regression model.

5. Conclusions

This paper proposes a regression model to quantitatively analyze the impact of wind and current on ship behavior (speed over ground and drift angle) derived from AIS data. The variations of ship behavior and the external impacts due to the size differences are also included during the analysis.

The variation of speed over ground in the unhindered situation due to ship size can be observed. The correlation analysis shows that the ship beam is better to indicate the relationship with v_{SOG} than length, which can be described through a logarithmic function.

The wind and current impact on ship behavior also vary for ships of different sizes. For small ships, both wind and current impacts on v_{SOG} decrease when the ships get larger. However, for large ships, the impact of wind on v_{SOG} gradually increases along with growing ship size, but the impact on the leeway angle fluctuates with a decrease. The current impact on v_{SOG} of larger ships is smaller, but the impact on the drift angle

is larger. For the coefficients that are significantly correlated to the ship's size by correlation analysis, the functions best estimate the relationship are adopted in the regression models.

According to the regression analysis results using the whole data set of ship behavior consisting of different sizes, about 70% of the variance in v_{SOG} of small ships can be explained by the factors of ship size, wind, and current. The choice of v_{SOG} in unhindered situation accounts for the major weight of the final sailing speed, which is mainly due to the ship size. However, for large ships, the explained variance drops to around 40%, possibly due to the large variation in the unhindered situation and the complex interaction of wind and current forces on ship hull. Compared to v_{SOG}, the wind and current impacts only account for 25% of the variance in leeway and drift angle, which is due to the instant decision differences between individual officers onboard and the maneuverability of individual ships. The results prove that the proposed theoretical expression of the impact mechanism and the revealed impact variation over ship size by analyzing the subsets of data are correct.

The estimated regression model provides the quantitative relationship of wind and current impacts on ship behavior considering ship size variation. Some conventional sailing habits of course alteration to follow the designed route in line with good seamanship are also statistically revealed by the estimate results. The analysis result could benefit both researchers and the port authority. For the researcher, a quantification of the impact mechanism of wind and current helps to further simulate ship behavior in such external conditions. For the port authority, the revealed insight into the relations between ship behavior and external factors will help the ship traffic management under different wind and current conditions and the corresponding risk control in port.

Within this paper, a nearly straight waterway is studied, which eliminates the impact of the waterway layout on ship behavior. As indicated by the estimated result, a port area with the more complex layout should be analyzed to identify such impact. According to the comparison of inbound and outbound ships, the distance to destination or the sailing direction of approaching or departing from a terminal also affects the speed choice, which can be further investigated. Based on a series of quantitative analyses looking into the relationship between the observed ship behavior and the external factors, a new nautical traffic model can be expected to predict the ship behavior under different conditions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yang Zhou: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing. **Winnie Daamen:** Methodology, Writing - review & editing, Supervision. **Tiedo Vellinga:** Writing - review & editing, Supervision, Funding acquisition. **Serge P. Hoogendoorn:** Methodology, Supervision.

Acknowledgment

This work is initiated by the project, Nautical traffic model based design and assessment of safe and efficient ports and waterways, under the Netherlands Organization for Scientific Research (NWO). The fellowship of Yang Zhou is supported by China Scholarship Council and Delft University of Technology. The support from SmartPort, both financially and by embedding the research in the practical context of the Port of Rotterdam, is highly appreciated. The authors would also like to thank the department of Data Management in the port of Rotterdam during the data collection, and appreciate Frank Cremer for accessing AIS data, Cor Mooiman for providing wind and visibility data, Bob van Hell and Lamber Hulsen for simulating current data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.oceaneng.2020.107774.

Appendix

As stated in Section 3.1, the two sets of behavior variables (v_{SOG}, γ and u, v) basically describe the same phenomenon of the ship motion. To show similar results as an example, the standardized coefficients of wind and current impact on surge and sway speed (u and v) are shown in Fig. 11 and Fig. 12. The corresponding results for speed over ground and leeway and drift angle are presented in Figs. 9 and 10 in Section 4.2.1. It can be observed that the estimated results for these two sets of behavior variables are the same in trend and close in value. Thus, in this paper, only the results for the variables directly derived from AIS data are explained in detail.

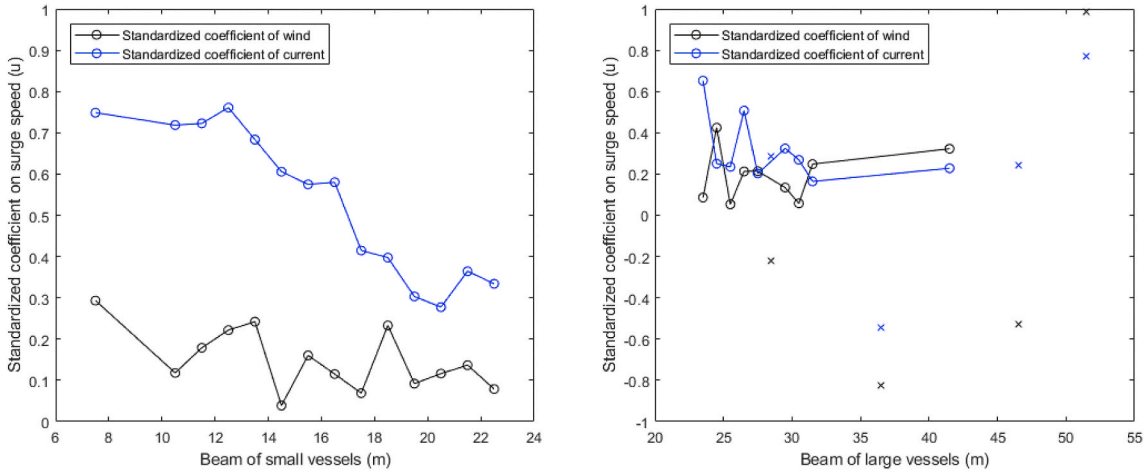

Fig. 11. Standardized coefficients of wind and current impact on surge speed u (wind in black and current in blue) as a function of ship beam (The subsets containing less than 30 trajectories are marked as crosses in the right figure)

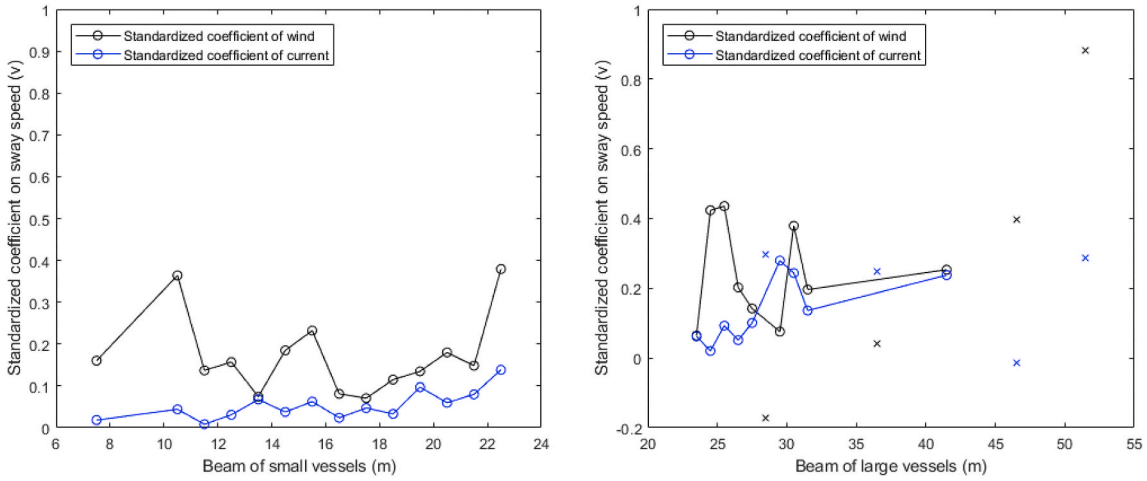

Fig. 12. Standardized coefficients of wind and current impact on sway speed v(wind in black and current in blue) as a function of ship beam (The subsets containing less than 30 trajectories are marked as crosses in the right figure).

References

Almaz, O.A., Or, İ., Özbaş, B., 2006. Investigation of transit maritime traffic in the strait of Istanbul through simulation modeling and scenario analysis. Int. J. Simulat. Syst. Sci. Technol. 7, 1–9.

Altan, Y.C., 2018. Collision diameter for maritime accidents considering the drifting of vessels. Ocean Eng. under-revi 106158. https://doi.org/10.1016/j.oceaneng.2019.106158.

Altan, Y.C., Otay, E.N., 2017. Maritime traffic analysis of the strait of istanbul based on AIS data. J. Navig. 70, 1367–1382. https://doi.org/10.1017/S0373463317000431.

Bailey, N., Ellis, N., Sampson, H., 2008. Training and Technology Onboard Ship : How Seafarers Learned to Use the Shipboard Automatic Identification System (AIS). Lloyd's Register Educational Trust Research Unit. Seafarers International Research Centre, and Cardiff University, Cardiff.

Beschnidt, J., Gilles, E.D., 2005. " Virtual Waterway " – a traffic simulation environment for inland and coastal waterways. WIT Trans. Built Environ. 79, 351–360.

Bowditch, N., 2017. American Practical Navigator: an Epitome of Navigation, 54th ed. National Geospatial-Intelligence Agency, Virginia.

Camci, F., Eldemir, F., Uysal, O., Ustun, I., 2009. Istanbul strait marine traffic simulation using multiple serially connected machinery concept. In: Proceedings of the 2009 Summer Computer Simulation Conference. Society for Modeling & Simulation International, pp. 424–429.

Central Commission for the Navigation of the Rhine, 2013. Resolution CC/R 2013 II.

Chen, C., Shiotani, S., Sasa, K., 2015. Effect of ocean currents on ship navigation in the east China sea. Ocean. Eng. 104, 283–293. https://doi.org/10.1016/j.oceaneng.2015.04.062.

Gao, X., Makino, H., Furusho, M., 2017. Analysis of ship drifting in a narrow channel using Automatic Identification System (AIS) data. WMU J. Marit. Aff. 16, 351–363. https://doi.org/10.1007/s13437-016-0115-7.

Gunnar Aarsæther, K., Moan, T., 2009. Estimating navigation patterns from AIS. J. Navig. 62, 587–607. https://doi.org/10.1017/S0373463309990129.

Harati-Mokhtari, A., Wall, A., Brooks, P., Wang, J., 2007. Automatic identification system (AIS): data reliability and human error implications. J. Navig. 60, 373–389. https://doi.org/10.1017/S0373463307004298.

International Maritime Organization, 2003. SN Circular 227 Guidelines for the Installation of a Shipborne Automatic Identification System (AIS).

International Maritime Organization, 1974. International Convention for the Safety of Life at Sea.

International Telecommunication Union, 2014. Technical characteristics for an automatic identification system (AIS) using time division multiple access in the VHF maritime mobile frequency band. Recomm. ITU-R.

Kepaptsoglou, K., Fountas, G., Karlaftis, M.G., 2015. Weather impact on containership routing in closed seas: a chance-constraint optimization approach. Transport. Res. C Emerg. Technol. 55, 139–155. https://doi.org/10.1016/j.trc.2015.01.027.

Leguit, D., 1999. A Possible VTS-Operator Support System Based on Vessel Traffic Simulation. Delft University of Technology.

Merrick, J.R.W., Van Dorp, J.R., Blackford, J.P., Shaw, G.L., Harrald, J., Mazzuchi, T.A., 2003. A traffic density analysis of proposed ferry service expansion in San Francisco bay using a maritime simulation model. Reliab. Eng. Syst. Saf. 81, 119–132. https://doi.org/10.1016/S0951-8320(03)00054-1.

Ni, Z., Qiu, Z., Su, T.C., 2010. On predicting boat drift for search and rescue. Ocean. Eng. 37, 1169–1179. https://doi.org/10.1016/j.oceaneng.2010.05.009.

Port of Rotterdam, 2014. Port Information Guide.

Qi, L., Zheng, Z., Gang, L., 2017. Marine traffic model based on cellular automaton: considering the change of the ship's velocity under the influence of the weather and sea. Phys. A Stat. Mech. its Appl. 483, 480–494. https://doi.org/10.1016/j.physa.2017.04.125.

Qu, X., Meng, Q., 2012. Development and applications of a simulation model for vessels in the Singapore Straits. Expert Syst. Appl. 39, 8430–8438. https://doi.org/10.1016/j.eswa.2012.01.176.

Richardson, P.L., 1997. Drifting in the wind: leeway error in shipdrift data. Deep. Res. Part I Oceanogr. Res. Pap. 44, 1877–1903. https://doi.org/10.1016/S0967-0637(97)00059-9.

Sarıöz, K., Kükner, A., Narlı, E., 2002. A real-time ship manoeuvring simulation study for the strait of istanbul (Bosporus). J. Navig. 52, 394–410. https://doi.org/10.1017/s0373463399008498.

Shu, Y., Daamen, W., Ligteringen, H., Hoogendoorn, S., 2013. Vessel speed, course, and path analysis in the Botlek area of the port of Rotterdam, Netherlands. Transp. Res. Rec. J. Transp. Res. Board 2330, 63–72. https://doi.org/10.3141/2330-09.

Shu, Y., Daamen, W., Ligteringen, H., Hoogendoorn, S.P., 2017. Influence of external conditions and vessel encounters on vessel behavior in ports and waterways using Automatic Identification System data. Ocean. Eng. 131, 1–14. https://doi.org/10.1016/j.oceaneng.2016.12.027.

Silveira, P.A.M., Teixeira, A.P., Soares, C.G., 2013. Use of AIS data to characterise marine traffic patterns and ship collision risk off the coast of Portugal. J. Navig. 66, 879–898. https://doi.org/10.1017/S0373463313000519.

Soda, T., Shiotani, S., Makino, H., Shimada, Y., 2012. Research on ship navigation in numerical simulation of weather and ocean in a bay. TransNav. Int. J. Mar. Navig. Saf. Sea Transp. 6, 19–25. https://doi.org/10.1201/b11343-24.

United Nations Conference on Trade and Development, 2018. Review of maritime transport. https://doi.org/10.18356/a9b345e7-en.

Vollebregt, E.A.H., Roest, M.R.T., Lander, J.W.M., 2003. Large scale computing at Rijkswaterstaat. Parallel Comput. 29, 1–20. https://doi.org/10.1016/S0167-8191(02)00217-X.

Yang, D., Wu, L., Wang, S., Jia, H., Li, K.X., 2019. How big data enriches maritime research – a critical review of Automatic Identification System (AIS) data applications. Transport Rev. 39, 755–773. https://doi.org/10.1080/01441647.2019.1649315.

Yasukawa, H., Yoshimura, Y., 2015. Introduction of MMG standard method for ship maneuvering predictions. J. Mar. Sci. Technol. 20, 37–52. https://doi.org/10.1007/s00773-014-0293-y.

Zhou, Y., Daamen, W., Vellinga, T., Hoogendoorn, S., 2017. AIS data analysis for the impacts of wind and current on ship behavior in straight waterways. Proc. 17th Int. Marit. Assoc. Mediterannean 265–272.

Zhou, Y., Daamen, W., Vellinga, T., Hoogendoorn, S., 2015. Vessel classification method based on vessel behavior in the port of Rotterdam. Sci. Journals Marit. Univ. Szczecin 42, 86–92.

Zhou, Y., Daamen, W., Vellinga, T., Hoogendoorn, S.P., 2019a. Review of maritime traffic models from vessel behavior modeling perspective. Transport. Res. C Emerg. Technol. 105, 323–345. https://doi.org/10.1016/j.trc.2019.06.004.

Zhou, Y., Daamen, W., Vellinga, T., Hoogendoorn, S.P., 2019b. Ship classification based on ship behavior clustering from AIS data. Ocean. Eng. 175, 176–187. https://doi.org/10.1016/j.oceaneng.2019.02.005.