# Port call extraction from vessel location data for characterising harbour traffic

Clément Iphar, Iwan Le Berre, Éric Foulquier, Aldo Napoli

Ocean Engineering 293 (2024) 116771

Contents lists available at ScienceDirect

Ocean Engineering

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

Research paper

# Port call extraction from vessel location data for characterising harbour traffic

Clément Iphar [a,*], Iwan Le Berre [a], Éric Foulquier [a], Aldo Napoli [b]

[a] LETG-Brest GEOMER, UMR 6554 CNRS, IUEM-Université de Bretagne Occidentale, Rue Dumont D'Urville, F-29280 Plouzané, France
[b] CRC, Mines Paris - PSL, Rue Claude Daunesse, F-06560 Sophia Antipolis, France

ARTICLE INFO

Dataset link: https://zenodo.org/records/10380638

*Keywords:*
Port activity
Port calls
AIS
Lesser Antilles
Marine transportation

ABSTRACT

Ports are essential facilities for the functioning of our globalised economy, as well as for the territories they supply. Monitoring port activity is therefore a relevant way of assessing the economic intensity they generate, and enabling regional analyses of their relative weight, their areas of activity and their development trajectories. However, comparative analysis of port activity is hampered by the variable quality and uneven distribution of the data produced by port authorities, which is far from being homogeneous, which makes comparative approaches difficult. In this paper, using vessel positioning data obtained through the Automatic Identification System (AIS), we propose a method for the extraction, quantification and qualification of port calls. We apply this method to a dataset in the Caribbean, covering the whole of 2019. Our method was able to describe the activity of 101 ports (in the sense of port areas) in the Lesser Antilles, by identifying 171,626 calls and their duration, by 5,907 ships of various type, size, age or flag. A comparison with the statistics published by a number of reference ports authorities and by the UNCTAD (United Nations Conference on Trade and Development) shows that the median difference of computed port calls is about 11%, thus validating the proposed approach. By qualifying the type of vessel that passes through port waters, according to its characteristics, this approach based on the number of calls makes it possible to observe the capabilities and the infrastructural quality deployed in the ports in terms of vessels, and, stemming from it, to develop the hypothesis of port vulnerability, that is a perspective for our future work.

## 1. Introduction

In an era marked by the great importance of maritime transportation in the international trade landscape, ports play an key role in the global economic structure. The assessment of port activities, however, presents a complex task, given the diversity of functions they cater to, from goods handling and passenger transit to bunkering, ship repair, and safety inspections, among others (Rodrigue, 2020).

Assessments of port activities are critically dependent on the type and detail of information provided by Port and Customs Authorities. This typically includes their annual activity reports and the nation-specific statistical systems. However, the data derived from these sources poses several challenges. Primarily, they are often consolidated and presented annually, broken down into broad categories such as the number of containers and passengers, volumes of dry and liquid bulk, and volumes of break bulk cargo. This annual approach of data prevents from considering port activity under the light of seasonality, or year-round activity fluctuations, which is detrimental in the scope of studying congestion effects and wait time. Secondarily, such data is not universally available, especially in countries with the weakest administrative systems, thereby hindering effective comparisons between ports (Jia et al., 2019; Yan et al., 2022). Thirdly, these datasets are typically designed to cater to a volume-based traffic analysis, although alternative indicators might provide a more comprehensive understanding of port activities. Port calls is an example of such an indicator. This approach to port calls also allows for the identification of characteristics of the infrastructural capacity deployed by shipowners on each service, in terms of size, type, and even age of the ships observed in each port. This opens up perspectives for observing disparities in the logistics of supplying territories. Due to the scarcity of detailed ship movement data, port call analysis across various ports or specific regions has been historically challenging. However, vessel monitoring data, particularly from the Automatic Identification System (AIS), opens up novel perspectives for in-depth port activity analysis.

* Corresponding author.
*E-mail addresses:* clement.iphar@univ-brest.fr (C. Iphar), iwan.leberre@univ-brest.fr (I. Le Berre), eric.foulquier@univ-brest.fr (É. Foulquier), aldo.napoli@minesparis.psl.eu (A. Napoli).

https://doi.org/10.1016/j.oceaneng.2024.116771
Received 31 July 2023; Received in revised form 14 November 2023; Accepted 13 January 2024

This paper proposes a method for extracting port calls and their characteristics from AIS data describing the vessel trajectories, in order to characterise port traffic on a regional scale. To validate this approach, our results are compared against statistical data published by several reference port authorities and international organisation. This validation enables us to extend the analysis to the regional scale of the Caribbean, in order to build a complete map of port calls, and systematically compare their characteristics depending on the kind of vessel (*e.g.* by type, size, age, flag). This validated AIS-based computation of vessel port calls forms the originality of this article, which is structured as follows. In Section 2, the background around the subject is presented. In Section 3, the computation method is proposed and its results subsequently enriched in Section 4. Results and validation by comparison with ground truth data is shown in Section 5, before some conclusive remarks on the analytical prospects thus opened up.

## 2. Background

### 2.1. The AIS as a global monitoring system

AIS data has revolutionised the analysis of shipping activities. Thanks to the recent development of satellites harvesting vessel signals on a large scale, possibly combined to coastal systems such as radars, and the Internet as a worldwide platform for data sharing, vessels no longer disappear beyond the horizon line (Wahl et al., 2005). In the maritime domain, such legally-enforced system has been put in place in the 2000's by the International Maritime Organisation (IMO), and called the Automatic Identification System (AIS). Originally deployed on-board vessels to prevent collision risk, it has become the largest provider of vessel information at sea, and enables an accurate global modelling of vessel behaviour and maritime traffic (Pallotta et al., 2013; Rong et al., 2022).

This system is widely used for maritime situational awareness, and is a large source of maritime navigation data (Fournier et al., 2018). AIS data are exploited for a wide range of applications which make use of the modelling of the maritime traffic at sea: economical, environmental and societal impacts (Prochazka et al., 2019): exhaust of greenhouse gases from cruise and ferry operations (Tichavska and Tovar, 2015) or commercial shipping (Styhre et al., 2017), impact on populations (Toscano et al., 2021), voyage planning (Cai et al., 2021) and port activity (Liu et al., 2020).

AIS high rate of transmission (from one position message every three minutes at minimal speeds to one position message every three seconds at maximal speed), and the vast network of receiving antennas and satellites enables a precise tracking of vessels, on both short and large temporal and geographical scales (Liu et al., 2023, 2020). Albeit not being perfect, as erroneous data and falsification data have been demonstrated (Iphar et al., 2020b), AIS data enable in particular estimating the spatiotemporal maritime traffic organisation as well as identifying the main maritime areas such as routes, ports and anchorage areas (Iphar and Jousselme, 2023).

Since the system is equally operating when the vessel is underway or stopped (although with various reporting rates), the AIS can be a powerful tool to model maritime traffic, not only during cruising along maritime routes, but also within the area of the port (Jia et al., 2019; Yan et al., 2022; Wu et al., 2020).

Indeed, most research conducted about the applications of the AIS involve the kinematics of the motion of the vessel or the notion of maritime route, which means the vessel is mainly considered as a mobile between two ports, and studies are mainly based on this mobility capability. In this paper, we would rather focus on the port, and more particularly find out to which extent it is possible to extract port call information from AIS data.

### 2.2. Various approaches around port calls

In the literature, most AIS-related studies are linked to the ship motion, and traffic modelling helps generate and build on maritime routes, or predict the upcoming port calls (Sakan et al., 2018), including the estimated time of arrival at the next port of call (El Mekkaoui et al., 2021) in the frame of optimisation, and the prediction of the destination using methods such as Markov models (Chen, 2022) or uncertain reasoning (Wu et al., 2020). On the contrary, there is less interest about vessels once stopped, at the port, which is what we want to focus on in this article.

The main use of AIS with vessels located within the area of a port is for pollution purposes (Tzannatos, 2010), in order to evaluate the tally of vessel emissions in port waters, for visiting vessels, local service vessels such as tugs (Chen et al., 2021), and vessels that are unidentified, and for which their characteristics must be inferred from the others (Zhang et al., 2019).

In many applications, port calls are used as a proxy for the port activity, and more precisely as a proxy for the demand Bai et al. (2022), with the ability to capture the changes on the shipping with time (Michail and Melas, 2020). Port activity (Xiao et al., 2015) and port performance evaluation (Yang et al., 2019) are also performed with port call data, and global maps of sea lines generated (Heiland and Ulltveit-Moe, 2020). The number of port calls has otherwise be deemed as a factor in the estimation of the accident rates by ports (Bye and Almklov, 2019), identifying at-risk waterways or vessels involved (Mou et al., 2010; Luong et al., 2021), in order to model and help avoid potential collisions.

Within the port area, AIS is also used for port call optimisation measures, for instance to measure fuel efficiency in shipping (Merkel et al., 2022), berthing velocity (Roubos et al., 2017), which is the speed at which the vessels arrived along the pier, wharf or quay, that differs with the tonnage and type of the vessel. As a whole, AIS has a prominent capacity to provide information in the prospect of smart ports (Iphar et al., 2021).

As for the determination of the location of vessel stops, Kernel Density Estimation has been used (Zocholl et al., 2021; Millefiori et al., 2016), along with clustering with DBScan algorithm (Fuentes, 2021), enabling a mapping of the local berthing areas. However, this purely spatial approach fails to consider the time spent at the port, which is important from an economic point of view (Slack et al., 2018) and valuable to compute. For instance, Merkel et al. (2022) uses the time at anchor in their study, to optimise $CO_2$ emissions, but do not explain the method used.

The approach that is closest to ours is Yan et al. (2022), with a somewhat similar method to extract port calls from AIS data that leads to the determination of the location of vessel berthing areas. We extend this reflection by adding the duration of the port call (which is an ambiguous notion we will discuss in Section 5), while discriminating vessels by their type, and validating the computed number of port calls with actual data from either reference datasets or port information.

## 3. Method

The analysis of trajectories are a main feature of AIS data. In Yan et al. (2008) a trajectory is defined as "*a record of the evolution of the position (perceived as a point) of an object that is moving in space during a given time interval in order to achieve a given goal*". The high frequency of AIS data allows us to identify vessels whose trajectory arrives in a port, stops there and then leaves, thus characterising a port call. Since the terms are similar, we have chosen to distinguish between a **port of call**, which is a port in which a vessel makes a stop, and a **port call**, which is the fact for a vessel to stop at a given port. The notion of port of call addresses a network issue, while the port call deals with "port attendance" issues.

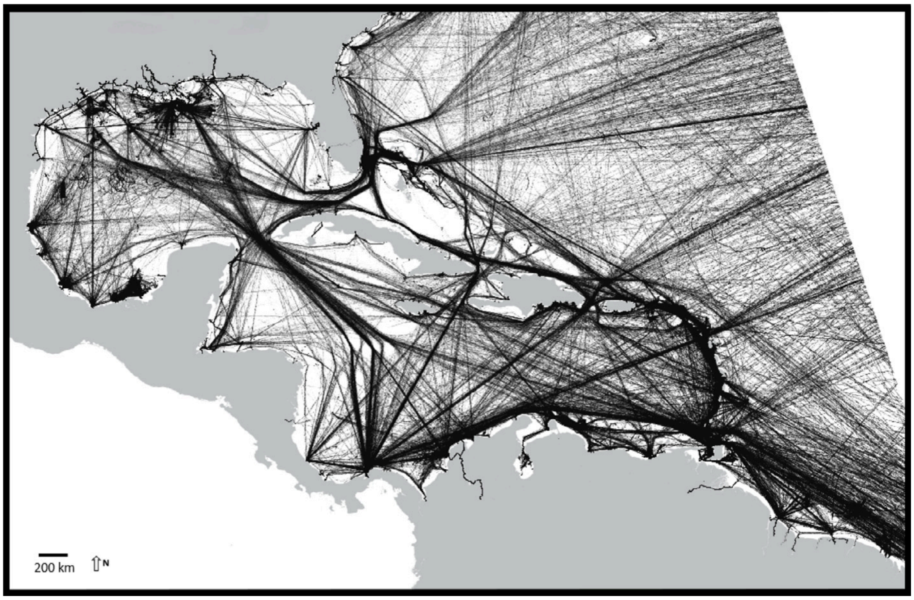


**Fig. 1.** AIS contacts of exactEarth dataset, for January 2019.

### 3.1. AIS dataset

The vessel position dataset has been puchased from the firm exactEarth . Its timespan covers the whole of 2019, and its spatial extent spans from 3.86 degrees North to 34.05 degrees North in terms of latitude, and from 98.02 degrees West to 51.17 degrees West in terms of longitude. It covers the entire Caribbean, including the Gulf of Mexico.

The whole dataset encompasses a grand total of 641,709,724 data contacts, which represents a mean value of 1.76 million AIS messages per day. This considerable amount of data requires carefulness when processed in order to maintain reasonable computational times.

Fig. 1 represents all data contacts for the sole month of January 2019 (46 million messages), that clearly shows the bounding box of our dataset.

This aggregated AIS dataset does not have distinct dynamic and static messages, but rather aggregated messages, with the most recent static information received being added to dynamic messages to complement them, with data including, but not limited to, length, width and destination. Dynamic messages also include a navigational status field, that is supposed to indicate which is the current state of the vessel. Those data fields, however, are filled in manually by the vessel crew, and are prone to massive reliability issues (Ben Abdallah et al., 2019). This is why in this paper we do not rely on them for assessing port calls and vessels characteristics, but rather rely on AIS dynamic data analysis and vessel register queries. The AIS system, therefore, is not a system that faithfully and perfectly depicts the maritime environment (Iphar et al., 2020a). Those imperfections, such as possible errors, or even falsifications, have been taken into consideration in the crafting of the method presented in this paper.

### 3.2. Study area

The Caribbean Sea is an important region for international trade and home to major shipping routes, connecting the Atlantic Ocean with the Gulf of Mexico and the Pacific Ocean, especially through the Panama Canal. The region is also a global hub for the cruise industry, and is home to intensive maritime traffic : our 2019 AIS dataset shows the presence of nearly 12,000 commercial ships, equivalent to a quarter of the world fleet.

Moreover, in this region, which is both continental and archipelagic, maritime transport is vital for many island or micro-island territories. Because of these geographical characteristics, the inter-island service is particularly developed. In addition, the Caribean is divided into a multitude of territories with varying sovereignty status, with a mix of some developed countries such as the United States, a large collection of small developing insular nations of various sizes, and some dependencies of other nations, such as France, the Netherlands, the United Kingdom or the United States. Indeed, this region is home to a large number of ships registered under flags of convenience.

In this paper, we focus on the Lesser Antilles, encompassing a group of diverse countries and territories spanning from the Virgin Islands to Trinidad and Tobago, including Barbados and the Dutch islands off the Venezuelan coast (Aruba, Bonaire and Curaçao).

### 3.3. Port dataset

Although global port databases do exist, they cover only the most important ports, and are neither exhaustive nor very precise (Wu et al., 2020). For example, in the whole of the Caribbean, the World Port Index (WPI) identifies 280 ports, while the IHS Markit database describes 301 ports, whereas we found 394 ports. The ports described are not always precisely located, as they are only located by a point: they are not delimited and their terminals are not mapped.

To have a cartographic reference frame for calculating port calls, a dataset describing the infrastructure – quays and wharves – of all ports of interest in the Caribbean has been produced using GIS photo-interpretation. In order to provide an overview of all port terminals capable of receiving merchant vessels, digitisation was carried out at a scale of 1:5000, using *Google Earth* and *Bing Map* satellite imagery, and the *OpenStreetMap* cartographic repository.

In the Lesser Antilles, a total of 101 ports, spanning across 19 territories (sovereign nations or dependencies) on 46 different islands (Table 1) have been identified and labelled (Fig. 2).

### 3.4. Determination of the occurrence of a port call

The computation of port calls is performed in three steps. First, AIS data is retrieved and the string of data points that display the behaviour of a port call are individuated into Raw Computed Port Calls (RCPCs), for which characteristics such as the length and the location are computed. Then, two consecutive operations, namely the concatenation and the merge, are performed, generating the set of Actual Port Calls

**Table 1**
Number of ports by territory in our port database for the Lesser Antilles.

| Territory | Code | Ports |
|---|---|---|
| Anguilla | AIA | 1 |
| Antigua and Barbuda | ATG | 4 |
| Bonaire, Sint Eustatius and Saba | BES | 5 |
| Barbados | BRB | 3 |
| Curacao | CUW | 5 |
| Dominica | DMA | 5 |
| Guadeloupe | GLP | 11 |
| Grenada | GRD | 6 |
| Saint Kitts and Nevis | KNA | 4 |
| Saint Lucia | LCA | 4 |
| Saint Martin | MAF | 2 |
| Montserrat | MSR | 2 |
| Martinique | MTQ | 7 |
| Sint Marteen | SXM | 3 |
| Trinidad and Tobago | TTO | 11 |
| Saint Vincent and the Grenadines | VCT | 8 |
| British Virgin Islands | VGB | 11 |
| United States Virgin Islands | VIR | 6 |

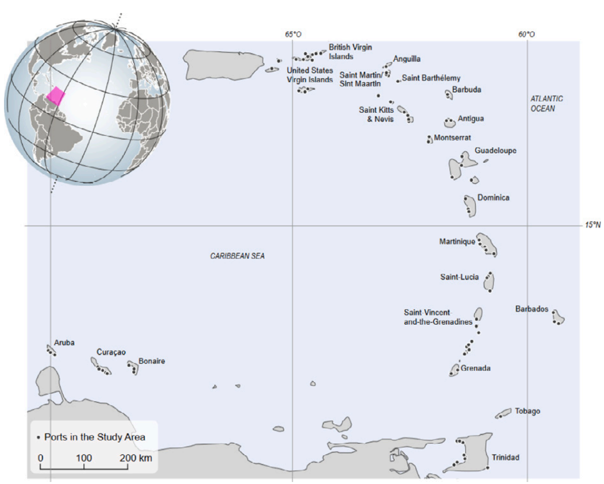


**Fig. 2.** Location of all 101 ports in our dataset for the Lesser Antilles.

(APCs) that we will use for the remainder of this article. The first step is shown in Section 3.4.1 and both concatenation and merge operations are shown in Section 3.4.2. One major statistical element pertaining to port calls, which is the assessment of their duration, will later be addressed in Section 3.4.3 (see Fig. 3).

### 3.4.1. Raw computed port calls

The first step of port call extraction presented in this section is the characterisation of RCPCs. Procedure 1 shows the algorithmic workflow, taking as input variables the AIS dataset, the set of ports $P$, a cut value for port proximity denoted $\epsilon$ and a cut value for the vessel speed that characterises the beginning and the end of a port call, denoted $\sigma$. In our study, $\epsilon$ has been set in accordance with literature on proximity to port infrastructure to half a maritime mile (ca. 925 m) (Klovning, 2020), and $\sigma$ has been set in accordance with literature about the standard speed value threshold between mooring and manoeuvring of 0.5 knots (Yan et al., 2022). The outputs of the procedures are the number of port calls and the port calls themselves, with their characteristics: $M_i$ is the vessel name, $P_k$ the port of call, $t_\alpha$ the arrival time and $t_\beta$ the departure time. Some elements of the AIS messages are used within the procedure, namely the timestamp $t_j$, speed over ground $S_j$ and localisation $L_j$ for the $j$th message.

The general principle of the algorithm is to deal with one vessel at a time, restrict the AIS messages to those within a close proximity to a port, and determine from the speed profile of the vessel the beginning,

**Procedure 1** Extraction of raw port calls

**Input:** $AIS, P, \epsilon, \sigma$
**Output:** $S, k$
  $k \leftarrow 0$ {Counter initialised}
  $S \leftarrow \emptyset$ {Port call repository initialised}
  $M \leftarrow List(AIS_{MMSI})$ {List all vessels within dataset}
  **for** $i \in 1 \cdots Card(M)$ **do**
    $D_i \leftarrow AIS_{M_i}^{d(P,AIS)<\epsilon}$ {All points from vessel $M_i$ near ports, ordered by time}
    $j = 1$ {Initialise counter, then scan all messages until last is reached}
    **while** $j \leq Card(D_i)$ **do**
      **if** $S_j > \sigma$ **then**
        $j = j + 1$ {Vessel is underway, message counter incremented}
      **else**
        $j = j + 1$ {Vessel is under speed threshold, message counter incremented}
        $t_\alpha = t_j$ {Port call start time}
        $C \leftarrow \bot$ {Flag initialised}
        **while** $C = \bot$ **do**
          **if** $S_j > \sigma$ **then**
            $C = \top$ {Vessel is underway again, port call has ended, flag is raised}
            $t_\beta = t_j$ {Port call end time}
            $P_k = min(n \in 1 \cdots Card(P))d(P_n, L_j)$ {Port of call determination}
            $k = k + 1$ {Port call counter is incremented}
            $S \leftarrow (k, M_i, P_k, t_\alpha, t_\beta)$ {New port call enters repository}
          **end if**
          $j = j + 1$ {Message counter incremented}
        **end while**
      **end if**
    **end while**
  **end for**

the end and the location of the port call. As shown in Fig. 4, we assume that the port call begins at the first data contact for which the speed is below the threshold value $\sigma$ upon the arrival of the vessel at the port, and it ends at the last data contact for which the speed is below the threshold value $\sigma$ as the vessel accelerates and leaves the port once the call is completed.

### 3.4.2. Merging computed elements into a single port call

Two consecutive operations need to be completed in order to generate the final table of all port calls: the first is a concatenation of RCPCs, resulting in Concatenated Port Calls (CPCs), the second is a merge of CPCs when the gap between them is small enough.

The concatenation is a consequence of the computational method. Indeed, RCPCs are computed one vessel at a time. However, vessel that are present in the port area for a long time will report a consequent number of messages. In order to avoid handling considerably large data arrays and save computational time, data from a single vessel has been divided in a number of data blocks, that number being directly proportional to the total number of messages received from that vessel.

Therefore, if a data block ends during a port call, the next data block will begin during this port call and display two separate RCPCs after the computation. One port call, if long enough, could even span across more than two data blocks. Since, in our computational workflow, the first message preceding and following a RCPC are retrieved, it is possible to merge those RCPCs using the timestamps and therefore recompute the characteristics of the newly generated CPC, as it actually took place, from the characteristics of every single RCPC that compose it.

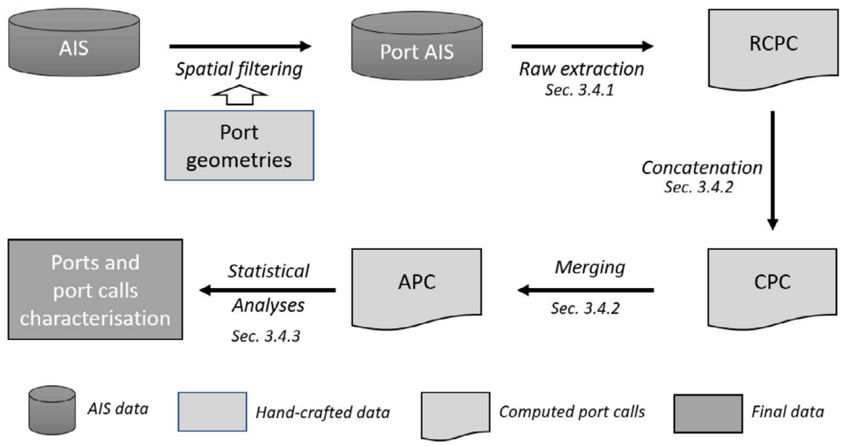


**Fig. 3.** Workflow from raw AIS data to port calls and associated statistical elements.

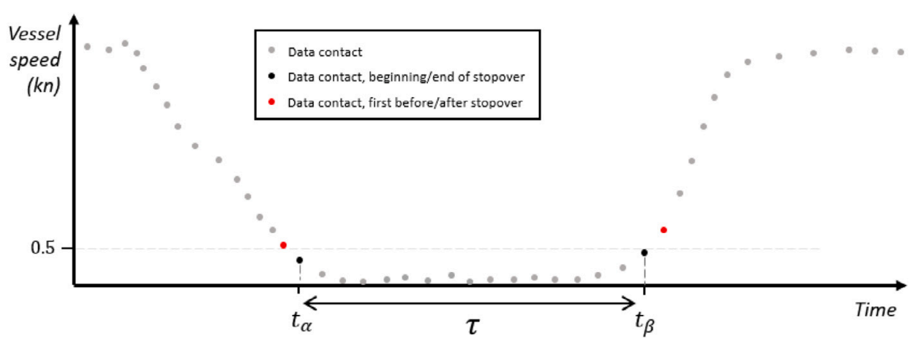


**Fig. 4.** Schematic speed profile of a vessel, characterising a port call.

The merge of CPCs occurs when two CPCs for the same vessel in the same port take place in succession. The merging of two CPCs is shown in Fig. 5. There is a gap in between those CPCs as, for a certain amount of time, the vessel has reached the threshold speed and therefore has computationally ended a CPC. This interval period in-between two CPCs can be due to several reasons: small movements of the vessel because of the current, movement from one quay to another one, movement after a short stop upon arrival, GPS inaccuracies that generate a false movement, amongst others.

Since this can occur not only with two CPCs, but several times in a row, only one new APC will be created during the computation, replacing all merged ones, and characteristics will be recomputed. For the sake of simplicity, in this section, we consider the case where there are only two CPCs to be merged.

In Fig. 5, the two CPCs are separated by $\tau'$, last respectively $\tau_1$ and $\tau_2$, and begin and end respectively at $t_{\alpha_1}$, $t_{\beta_1}$, and $t_{\alpha_2}$, $t_{\beta_2}$. The maximum gap period between CPCs to be merged, denoted $\tau_{Max}$, has been set to 30 min, in accordance with literature (Yan et al., 2022) and with the additional advice of a maritime navigation expert. The characteristics of the new ACP are as such: $t_\alpha = t_{\alpha_1}$, $t_\beta = t_{\beta_2}$, and $\tau = \tau_1 + \tau' + \tau_2$.

*3.4.3. Validation of computed port call duration against missing data*

In addition to occurrence, another interesting variable is the duration of port call, which can be used as an indicator of port performance, in particular by comparing the port calls of the same kind of vessels in different ports.

But, the determination of the duration of a port call is not straightforward, as it highly depends on available data quality. Indeed, while it is possible to know whether a ship has made a port call even with missing data, for assessing the correct duration of this port call a series of conditions must be met: the time of arrival and the time of departure must be known, and time series should show no large temporal gap.

Fig. 6 shows four cases in which it is possible to tell that a port call has taken place, but any temporal measure of the duration thereof would be unsure. In the case 1, data of vessel arrival to port is missing, therefore the computed port call time is shorter than the actual port call time. In the case number 2, the situation is similar except that the missing data is for vessel departure. Case number 3 encompasses both cases 1 and 2, and vessel data is missing both upon arrival and departure of the vessel, thus the computed port call time is presumably much shorter than reality. The case number 4 is different, as data on arrival and on departure are correctly reported, but there is a reporting gap in between. Therefore, the computed port call time could be correct, but it is unsure whether or not the vessel has moved during this gap, and that this event reported as one port call could actually be two distinct calls of various duration.

## 4. Enrichment

### *4.1. Vessel categorisation*

In order to obtain a precise understanding of the possible differentiation between the vessels, a categorisation based on the vessel type has been performed. For this purpose, the IHS Markit database classification, otherwise purchased, has been used. Within all categories of vessels, the selected types are bulk carriers, container ships, cruise vessels, general cargo vessels, inter-island ships, service vessels and tankers. A last category encompasses all vessels that have not been classified under any of those seven classes. Note that those classes only cover commercial fleets and service vessels, as fishing and recreational navigation is not of interest for the port call approach that we have.

It must be noted that the classification is not perfect, as some vessels that do fall under one of those categories is not present within the IHS Markit database, or is present under a different class. Also,

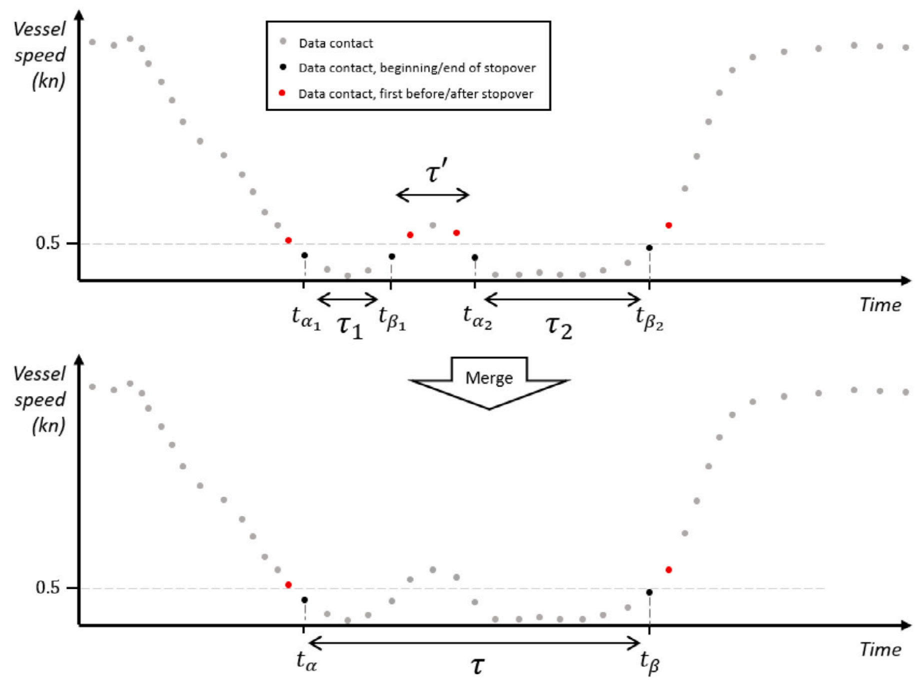


**Fig. 5.** The merging of two port calls separated by a small amount of time.

**Table 2**
Labels for vessel type classification.

| Vessel type | Abbr. | Database | Data field | Entry | Card. |
|---|---|---|---|---|---|
| Bulk Carriers | *bulk* | IHS Markit | shiptypelevel2 | “Bulk Carriers” | 4328 |
| Container Ships | *cont* | IHS Markit | shiptypelevel3 | “Container” | 1373 |
| Cruise Vessels | *crui* | IHS Markit | shiptypelevel4 | “Passenger (Cruise) Ship” | 243 |
| General Cargo Vessels | *gcar* | IHS Markit | shiptypelevel3 | “General Cargo” | 1154 |
| Inter-island Ships | *iisl* | IHS Markit | shiptypelevel4 | “Passenger Ship” | 65 |
| Service Vessels | *serv* | IHS Markit | shiptypelevel2 | “Miscellaneous” | 1271 |
| Tankers | *tank* | IHS Markit | shiptypelevel2 | “Tankers” | 4190 |

classes in the IHS Markit classification are somewhat blur and several levels of detail exist (*i.e.* different data fields offer various levels of granularity when describing the vessel type). As a consequence, in order to properly distinguish the vessel types of interest, some elements have to be retrieved in various data fields, which increase the risk of misclassification. However, to the best of our understanding, those classes are mutually exclusive. Table 2 shows the data fields and the corresponding data field label that has been selected for each of the seven vessel types of interest. The values shown in the “Card” column stand for the number of unique vessels of each type that are listed within the vessel register.

The *Miscellaneous* category includes most service vessels, and a few other kind of vessels (such as research vessels or pollution control vessels for instance), which offers a satisfactory compromise between simplicity and speed of classification of vessels, and their optimisation. In the IHS Markit database, out of the 1271 vessels classified under the *Miscellaneous* category, 1197 (94%) are service vessels, hence the compromise.

### *4.2. Generation of a berths dataset*

There is some interest in being able to determine the docking location of a vessel, particularly in the scope of vessel type classification. In this respect, for each port call, an unweighted barycenter of vessel locations during the port call is computed for all points for which the speed value is null.

Since vessel stops are clustered, a manual delimitation of berths has been performed, based on a visual assessment of clustering locations. In Pointe-à-Pitre, 20 of such areas, hereafter named berths, have been drawn.

Fig. 7 shows the spatial distribution of all computed barycenters for the port of Pointe-à-Pitre, in the French archipelago of Guadeloupe, along with the location of such clusters, in Brown, in Pointe-à-Pitre harbour.

### *4.3. Dataset enhancement*

Once the berths have been determined, the stopping locations within the port are individuated, and emerging patterns of life can be studied, in order to have a more comprehensive understanding of the dynamics of the port, and the specialisation of berths and wharves. In addition, as berths contain both known and unknown vessels, they can be used to qualify unknown vessels, based on the reasonable assumption that berths are specialised and that similar vessels will tend to berth at the same locations.

To illustrate the matter, we propose to pursue the example of the port of Pointe-à-Pitre, in the French island of Guadeloupe. Table 3 shows the distribution of vessel types for a selection of berths.

In Table 3 are shown only 11 berths, which are those for which at least one vessel is known (so that a counting of vessel types can take place) and for which at least one vessel is unknown (so that an assignment can be performed and data effectively enhanced). Other berths have no vessels of unknown type, or all their vessels are of unknown type. In both cases, data enhancement based on vessel type is impossible.

Out of the 11 berths shown in Table 3, one has too few cases of port calls done by a vessel for which the type is known: it is berth number

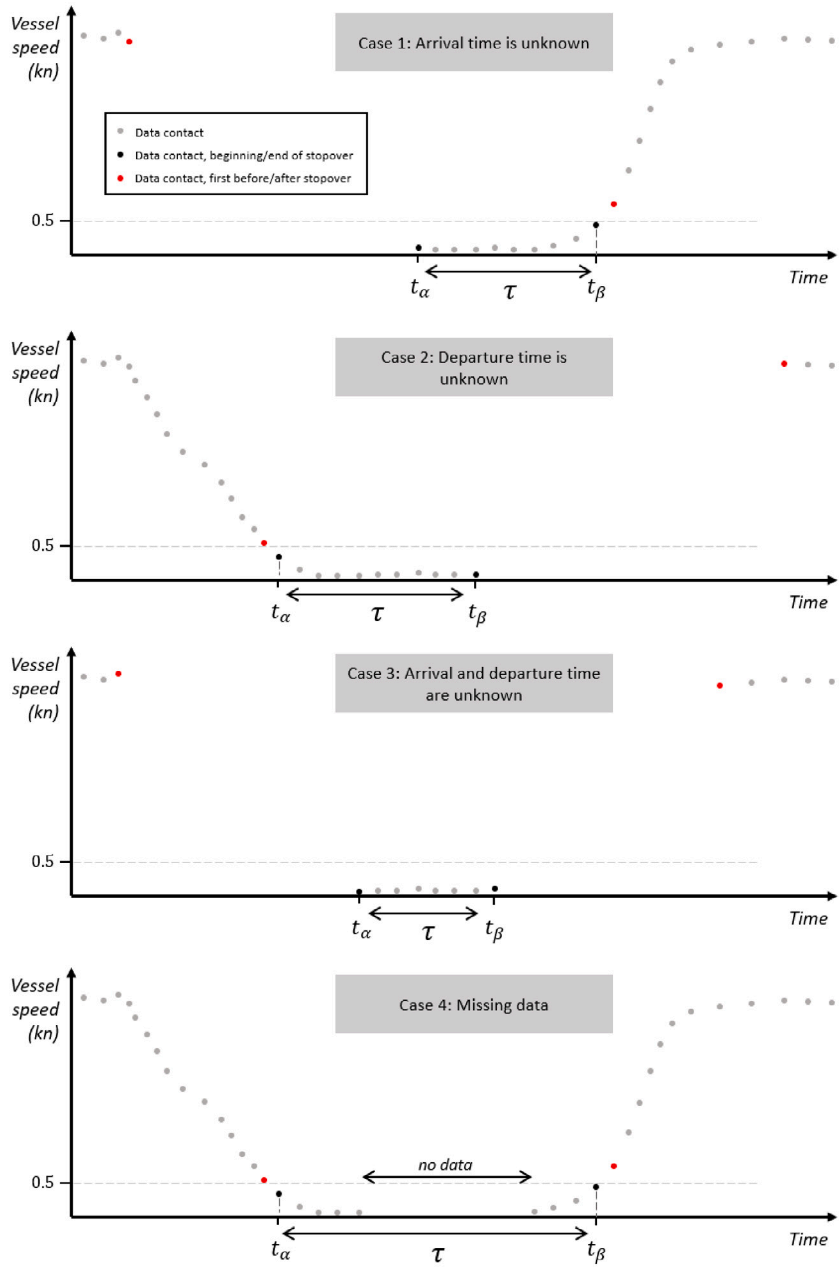


**Fig. 6.** Cases in which the determination of the port call duration is unsure.

**Table 3**
Determination of the main vessel class for all berths from Pointe-à-Pitre port for which at least one vessel is known and at least one vessel is unknown.

| Berth | Total stops | Known stops | Assign | First class | | Second class | | Third class | |
|---|---|---|---|---|---|---|---|---|---|
| **4** | 61 | 30 | **iisl** | iisl | 30 | – | – | – | – |
| **6** | 173 | 172 | **iisl** | iisl | 172 | – | – | – | – |
| **7** | 21 | 2 | - | serv | 2 | – | – | – | – |
| **8** | 31 | 30 | **serv** | serv | 30 | – | – | – | – |
| **9** | 372 | 371 | **serv** | serv | 371 | – | – | – | – |
| **12** | 696 | 695 | **serv** | serv | 695 | – | – | – | – |
| **13** | 192 | 87 | **cont** | cont | 80 | gcar | 5 | serv | 2 |
| **16** | 69 | 41 | **serv** | serv | 41 | – | – | – | – |
| **17** | 66 | 50 | **serv** | serv | 50 | – | – | – | – |
| **18** | 43 | 42 | **tank** | tank | 38 | serv | 4 | – | – |
| **20** | 473 | 472 | **serv** | serv | 438 | cont | 34 | – | – |

**Table 4**
Number of unique vessels and port calls for both the set of vessels for which the type is known and the set of vessels for which the type is inferred.

| Vessel type | Number of vessels | | Number of port calls from | |
|---|---|---|---|---|
| | Type already known | Type inferred | Type already known | Type inferred |
| bulk | 73 | 9 | 946 | 151 |
| cont | 116 | 24 | 3176 | 313 |
| crui | 147 | 175 | 4118 | 2228 |
| gcar | 199 | 56 | 4075 | 3748 |
| iisl | 30 | 310 | 16 625 | 11 030 |
| serv | 128 | 204 | 24 471 | 8527 |
| tank | 677 | 65 | 4235 | 629 |

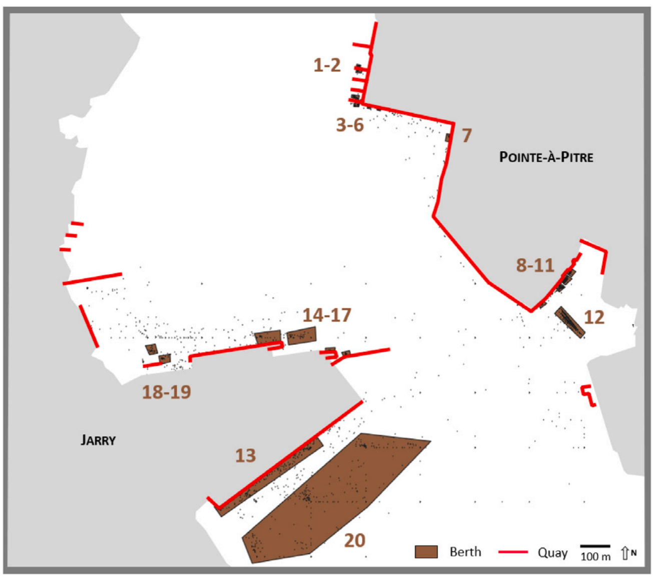


**Fig. 7.** Location of quays, stopping locations and the areas in which they cluster (mainly berths) in Pointe-à-Pitre harbour.

7, with 2 stops known out of 21. Out of the remaining 10, all have a clearly dominating vessel type, with various levels of share: 3 with a ratio $\in [0.9, 1[$ of known stops, and 7 with all 100% of known stops. In those 7 berths, it is possible to assign the majority class to all remaining vessels with a high level of confidence. The same applies to the 3 berths with share of the majority class over 90%.

Since the vessels will move across different ports and be located in various berths, it is necessary to check that a vessel has not been assigned two different classes across all berths. There is no such occurrence in our computation.

Out of the 101 ports of the Lesser Antilles, 1370 vessels had their type known initially, and the type of 843 vessels have been inferred with this method. More specifically, the separation by vessel types can be seen in Table 4.

## 5. Results and validation

The data set of all computed port calls, all port geometries, vessel trajectories, berths and vessel types, including enriched, which has been the topic of Sections 3 and 4, as well as some reference data that will be discussed in Section 5.1, is published. It is publicly available online (Iphar et al., 2023) and will be complemented with a data paper offering a comprehensive understanding of data generation.

### 5.1. Reference data and semantic alignment

The reference data we use to assess the quality of our results are twofold. On the one hand, we back our study with data from UNCTAD,[1] the United Nations Conference on Trade and Development, and on the other hand we use released data from ports authorities, either publicly or upon request.

Data from UNCTAD is plentiful but grouped by country or dependency, and not by port, even less at the level of single terminals or quays. In order to assess the port of call computation, we use a table that we extracted from UNCTAD website and rearranged within our database. This table shows, by territory, the number of vessel port calls across the entire 2019 year for the following vessel classes: liquid bulk, dry bulk, dry breakbulk, liquefied petroleum gas, liquefied natural gas, roll-on roll-off vessels, container ships and passenger vessels.

The second source of reference data are the port authorities themselves, with the major drawback that data are seldom publicly available. It implies that we will have only a handful of pieces of information at our disposal, because of the scarcity of information. Indeed, ports are reluctant to openly publish their data, as those are information of strategic value. In this paper, we will compare our computation against data from five ports, which are Fort-de-France, in the French overseas region of Martinique, Port-of-Spain (Trinidad and Tobago), Bridgetown in Barbados and both Willemstad and Oranjestad in the Dutch overseas territory of the Netherlands.

When comparing values coming from various sources, and with our own results, it is of great importance to ensure that a common vocabulary is used when it comes to vessel types. Indeed, various denominations may encompass different realities, and it is necessary to be as accurate as possible when comparing the results, as it pertains to the validation of our approach. UNCTAD data nomenclature, albeit peculiar, somewhat aligns with the data types of our study (see Table 5).

### 5.2. Results with data from ports

Table 6 allows a comparison between the computed number of port calls in 7 occurrences covering 5 ports and 3 types of vessels, and the corresponding values are reported by the relevant port authorities. Those results will be discussed in Section 5.4, where the ratio is computed as the number of ground truth calls over the number of computed calls.

### 5.3. Results with data from UNCTAD

Table 7 shows the comparison of the number of computed port calls with our method, with the number of port calls that individual nations declared to UNCTAD for the year 2019. It shows the containers, cruise vessels and tankers figures for countries and dependencies of the Lesser Antilles. Cases for which data is lacking (for instance tankers in Dominica) are not shown. In this Table, similarly to Table 6, the ratio is computed as the number of computed calls over the number of ground truth calls. Those results will be discussed in Section 5.4.

[1] https://unctad.org/

**Table 5**
All corresponding values.

| Source | Field name | Alignment |
|---|---|---|
| UNCTAD | Dry bulk | Bulk |
| UNCTAD | Liquid Bulk + liquefied petroleum gas, liquefied natural gas | Tankers |
| UNCTAD | Container ships | Containers |
| Fort-de-France Port Authority | Goods | Bulk + Containers + General Cargo + Tankers |

**Table 6**
Computed and declared number of port calls by port authorities.

| Port | Country | Type | Origin | Computed value | Declared value | Ratio |
|---|---|---|---|---|---|---|
| Oranjestad | Aruba | Cruise | Port Authority | 268 | 324 | 1.21 |
| Port-of-Spain | Trinidad and Tobago | Cruise | Port Authority | 24 | 26 | 1.08 |
| Fort-de-France | Martinique | Cruise | Port Authority | 165 | 169 | 1.02 |
| Fort-de-France | Martinique | Goods | Port Authority | 602 | 704 | 1.17 |
| Bridgetown | Barbados | Cruise | Port Authority | 408 | 422 | 1.03 |
| Bridgetown | Barbados | Tankers | Port Authority | 133 | 155 | 1.17 |
| Willemstad | Curacao | Cruise | Port Authority | 265 | 313 | 1.18 |

**Table 7**
Computed and declared values for cruise, container and tanker vessels by the UNCTAD.

| Country | Type | Computed value | Declared value | Ratio |
|---|---|---|---|---|
| Antigua and Barbuda | Cruise | 320 | 346 | 1.08 |
| Antigua and Barbuda | Container | 64 | 48 | 0.75 |
| Barbados | Cruise | 408 | 470 | 1.15 |
| Barbados | Container | 238 | 248 | 1.04 |
| Barbados | Tankers | 133 | 150 | 1.13 |
| Dominica | Cruise | 172 | 159 | 0.92 |
| Dominica | Container | 40 | 40 | 1 |
| Grenada | Cruise | 185 | 219 | 1.18 |
| Grenada | Container | 163 | 163 | 1 |
| Guadeloupe | Cruise | 130 | 218 | 1.68 |
| Guadeloupe | Container | 271 | 248 | 0.92 |
| Guadeloupe | Tankers | 157 | 157 | 1 |
| Martinique | Cruise | 169 | 168 | 0.99 |
| Martinique | Container | 204 | 181 | 0.89 |
| Martinique | Tankers | 182 | 178 | 0.98 |
| Saint Christopher and Nevis | Cruise | 349 | 379 | 1.09 |
| Saint Christopher and Nevis | Container | 61 | 67 | 1.10 |
| Saint Christopher and Nevis | Tankers | 85 | 59 | 0.69 |
| Saint Lucia | Cruise | 341 | 420 | 1.23 |
| Saint Lucia | Container | 130 | 152 | 1.17 |
| Saint Lucia | Tankers | 340 | 373 | 1.10 |
| Saint Vincent and the Grenadines | Cruise | 142 | 1995 | 14.0 |
| Saint Vincent and the Grenadines | Container | 134 | 178 | 1.33 |
| Saint Vincent and the Grenadines | Tankers | 58 | 19 | 0.33 |
| Trinidad and Tobago | Cruise | 53 | 1739 | 32.8 |
| Trinidad and Tobago | Container | 930 | 825 | 0.89 |
| Trinidad and Tobago | Tankers | 1629 | 1135 | 0.70 |

### *5.4. Discussion*

Out of the 34 lines of comparison, 3 have computed the exact same number of port calls, 15 have calculated a very similar number (*i.e.* with a ratio between 0.9 and 1.1), and 25 have computed a total of port call that is less the 20% off the declared ground truth (ratio values between 0.8 and 1.2). This high number of matching results encourages us to validate the method that leads to the individuation of port calls, and invites us to reflect on the circumstances that lead to mismatching cases.

For instance, two very clear outliers are the ground truth values of port calls for cruise vessels in Saint Vincent and the Grenadines and in Trinidad and Tobago, respectively at 1995 and 1739, while computed cruise calls are at a tally of 142 and 53, respectively. After further investigation, it turns out that we computed a total of 2390 and 2017 calls for inter-island vessels for Saint Vincent and the Grenadines and Trinidad and Tobago, respectively. Those figures do match the ones declared by the countries to the UNCTAD. Both cruise vessels and inter-island vessels are passenger vessels, and we set here a reasonable hypothesis that there has been a confusion on the type of vessel to report data. This particular example shows us that ground truth data must be handled carefully, as semantic alignment is of paramount importance in the validation of this study, as we previously presented in Section 5.1.

In addition to those outliers, the UNCTAD data are aggregated over the whole administrative area, whether it is a country or a territory, making the computation more complex, as one has to ensure that no port is missing and that the ports that are taken into consideration for their countrywide computation are the exact same than those that we consider for computation. Indeed, despite our meticulous work of surveying every single commercial pier within ports, it may be that some do miss. And it is also possible that the data deemed as official do not encompass every single seaport in the nation (which may be particularly valid for large nations). Those discrepancies make the UNCTAD data somewhat less trustworthy than the port authority reports, as those only consider one single seaport, for which the authority is aware of all used piers. This might be the reason why our results better fit port authority data, with computed values consistently within 20% of computed values.

Another feature that may be a cause of mismatching results is the very notion of port call, that can have a variety of definitions. In this

**Table 8**
Possible cases of discrepancies in the comparison of the number of port calls.

| Possible cause | Nature | Port data | UNCTAD data | Our method |
|---|---|---|---|---|
| Erroneous report by ports | Qualitative and quantitative | X | X | |
| Erroneous aggregation at national level | Qualitative and quantitative | | X | |
| Missing ports | Quantitative | | X | X |
| Uncertainty about vessel arrival or departure | Qualitative | | | X |
| Vague definition of a port call | Quantitative | X | X | X |
| Vague definition of small movements | Quantitative | X | X | X |
| Confusion on vessel types | Qualitative | X | X | X |

paper, we stuck to definitions and threshold values found in literature, but there is no indication whatsoever about the port authorities or UNCTAD notions of a port call, which we expect to be declaration-based rather than data-based, and relying on harbour masters and official reports. Those possible discrepancies can only be an additional hurdle in a proper comparison of computed and ground truth port calls.

In addition, the enrichment that we propose may be incomplete, for instance in the unlikely case where all vessels docking in a specific pier are not present within the vessel information dataset we use. Although unfortunate, this possibility adds an uncertainty on the accuracy of the total number of computed port calls for each seaport (see Table 8).

Overall, given all possible causes of discrepancies between computed and reported number of port calls, the results are satisfactory to provide an accurate valuation of activity in each port of our study area, for each kind of vessel. The breakdown of traffic by vessel type and the use of an external reference database (*i.e.* IHS Markit vessel database) enable us to generate a set of statistics about the age of the vessels, their tonnage, or the duration of the port calls. This enable us to extract meaningful information about the pool of vessel mooring at a given port, or even specifically at a given pier, therefore possibly producing data that might be of interest for port authorities in a competitive market, enabling them to compare their performances with the ones of their competitors. In doing so, when considering the port call durations, only those for which the arrival time, the departure time, and with frequent position reporting while in the port shall be taken into consideration, as shown before in Section 3.4.3, so that comparisons are meaningful.

## 6. Conclusions

The work presented in this paper is part of the research in the fields of maritime transportation and port activity estimation.

We proposed a method for the extraction of port calls from AIS data, and satisfactorily applied it to the Lesser Antilles, for the 2019 year. From the original set of AIS messages, each extended period of time for which a vessel is stopped, or has a very reduced speed, has been characterised and the algorithmic workflow provided. The set of computed raw port calls then went concatenated and merged, to account for small movements within the port that do not constitute a separate port call.

We focus on merchant vessels, distinguishing between passengers (cruise and inter-island vessels) and goods transportation (tankers, bulk, general cargo and container vessels). We assume that such a wide range of ship types entails different port traffic patterns and port call practices, particularly in a region like the Caribbean, where political, legal and socio-economic contrasts are significant.

To qualify these different categories of vessels, we used an external database (IHS Markit vessel database), that describes ship characteristics such as tonnage or age. For vessels in our dataset that are not described in this database, we developed an enrichment method based on the assumption that vessels of the same type tend to frequent specific terminals. We were then able to produce a separate analysis of port calls, using their duration as an indicator of specific uses of port terminals.

Once the total number of calls by ship type has been calculated for all ports in the Lesser Antilles, our approach is validated using both port authority and UNCTAD data. This validation shows that the median error is about 11%. Causes for the discrepancies observed for some ports and some vessel types have been discussed.

With respect to existing literature, although the general method for extracting port calls are somewhat similar, our strength lies in our consideration for vessel and port call characteristics, which enables a deeper understanding of seaport activity. On the one hand, we extract the vessel characteristics from a register, thus allowing for a differentiated approach of maritime traffic. On the other hand, the temporal dimension is taken into consideration through the call duration, which gives hints about port activity, and the spatial dimension is handled via the computation of the exact location of berthing, enabling a characterisation of ports down to the level of the berth.

Overall, our method allows for a reliable estimation of the number of port calls, by territory, by port or by terminal, and generates a set of statistics useful for their description: duration, number of vessels, age. The use of AIS data and the protocol we have developed thus provide a more comprehensive view of port calls than can be achieved using port authority data, which is not systematically available. It is also more detailed than UNCTAD data, which is only available at national or island territory level. The database thus constituted enables analyses and comparisons to be made at different scales, from local to regional, between terminals, ports and territories in the Caribbean; the integration of datasets covering years other than 2019 would enable a temporal analysis to be envisaged, for example, to track the trajectories of ports and terminals. This new knowledge of ports of call is of interest, not only in the scope of further research in the field of human geography, but also is likely to be of interest to shipping stakeholders.

Indeed, our future work will aim to complete our set of indicators by identifying shipping lines from the extraction of consecutive port calls, to study the connectivity of ports within their local and global environment, for the whole of the Caribbean.

In addition, the concept of port calls can be used to characterise the level of exposure to navigation-related hazards of port areas, since ships are risk carriers, regardless of their position in the operating chain. As the Port Authority does not control the qualities of the ships that use it, most of which are operated by private shipowners, it can be assumed that preventive control of ship-related risk is partly beyond its control. By qualifying the type of vessel that passes through port waters, according to its size, its freight, its emissions, its age, and even its registration number, the approach based on the number of calls makes it possible to develop the hypothesis of port vulnerability. Exploring this notion is also a perspective for our future work.

## CRediT authorship contribution statement

**Clément Iphar:** Conceptualization, Data curation, Funding acquisition, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **Iwan Le Berre:** Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Validation, Writing – review & editing. **Éric Foulquier:** Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Writing – review & editing. **Aldo Napoli:** Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Writing – review & editing.

## 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.

## Data availability

Data is available online: https://zenodo.org/records/10380638.

## Acknowledgements

This work has benefited from several state fundings managed by the Agence Nationale de la Recherche under the "Investissements d'avenir" program: the Labex DRIIHM ANR-11-LABX-0010, and the ISblue project "Interdisciplinary graduate school for the blue planet" ANR-17-EURE-0015. It has also received funding from the European Union's Horizon 2020 research and programme under the Marie Sklodowska-Curie grant agreement No 899546.

## References

Bai, X., Xu, M., Han, T., Yang, D., 2022. Quantifying the impact of pandemic lockdown policies on global port calls. Transp. Res. Part A: Policy Pract. http://dx.doi.org/10.1016/j.tra.2022.08.002.

Ben Abdallah, N., Iphar, C., Arcieri, G., Jousselme, A.-L., 2019. Fixing errors in the AIS destination field. In: Proceedings of the OCEANS 2019 Marseille Conference.

Bye, R.J., Almklov, P.G., 2019. Normalization of maritime accident data using AIS. Mar. Policy http://dx.doi.org/10.1016/j.marpol.2019.103675.

Cai, J., Chen, G., Lutzen, M., Gorm Maly Rytter, N., 2021. A practical AIS-based route library for voyage planning at the pre-fixture stage. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2021.109478.

Chen, D., 2022. The Destination Port Prediction for Tramp Ships Based on AIS Trajectory Data Mining: A Case Study of VLCC (Dissertation). The Hong Kong Polytechnic University - Department of Logistics and Maritime Studies.

Chen, S., Meng, Q., Jia, P., Kuang, H., 2021. An operational-mode-based method for estimating ship emissions in port waters. Transp. Res. Part D: Transp. Environ. http://dx.doi.org/10.1016/j.trd.2021.103080.

El Mekkaoui, S., Benabbou, L., Berrado, A., 2021. Predicting ships estimated time of arrival based on AIS data. In: Proceedings of the SITA'20 Conference. http://dx.doi.org/10.1145/3419604.3419768.

Fournier, M., Casey Hilliard, R., Rezaee, S., Pelot, R., 2018. Past, present, and future of the satellite-based automatic identification system: areas of applications (2004–2016). WMU J. Marit. Affairs 17 (3), 311–345. http://dx.doi.org/10.1007/s13437-018-0151-6.

Fuentes, G., 2021. Generating bunkering statistics from AIS data: A machine learning approach. Transp. Res. Part E: Logist. Transp. Rev. http://dx.doi.org/10.1016/j.tre.2021.102495.

Heiland, I., Ulltveit-Moe, K.H., 2020. An unintended crisis in sea transportation due to COVID-19 restrictions. In: Baldwin, R., Evenett, S.J. (Eds.), COVID-19 and Trade Policy: Why Turning Inward Won'T Work. CEPR Press, pp. 151–164.

Iphar, C., Jousselme, A.-L., 2023. A geometry-based fuzzy approach for long-term association of vessels to maritime routes. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2023.114755.

Iphar, C., Le Berre, I., Sahuquet, M., Napoli, A., Foulquier, E., 2023. Port calls and vessel trajectory dataset in the caribbean with accurate port quays survey. http://dx.doi.org/10.5281/zenodo.10091946, Data set. Licence CC-BY-NC-SA-4.0 and CC-BY-3.0-IGO. Zenodo.

Iphar, C., Napoli, A., Ray, C., 2020a. An expert-based method for the risk assessment of anomalous maritime transportation data. Appl. Ocean Res. http://dx.doi.org/10.1016/j.apor.2020.102337.

Iphar, C., Ray, C., Napoli, A., 2020b. Data integrity assessment for maritime anomaly detection. Expert Syst. Appl. 147, http://dx.doi.org/10.1016/j.eswa.2020.113219.

Iphar, C., Zocholl, M., Jousselme, A.-L., 2021. Semantics of maritime routes: Conciliating complementary views. In: Proceedings of the OCEANS 2021 San Diego Conference. http://dx.doi.org/10.23919/OCEANS44145.2021.9705934.

Jia, H., Prakash, V., Smith, T., 2019. Estimating vessel payloads in bulk shipping using AIS data. Int. J. Shipp. Transp. Logist. (1), 25–40.

Klovning, E., 2020. Wind affecting berthing operations. Transnav, Int. J. Mar. Navig. Saf. Sea Transp. (3), 721–725. http://dx.doi.org/10.12716/1001.14.03.26.

Liu, C., Liu, J., Zhou, X., Zhao, Z., Wan, C., Liu, Z., 2020. AIS data-driven approach to estimate navigable capacity of busy waterways focusing on ships entering and leaving port. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2020.108215.

Liu, L., Shibasaki, R., Zhang, Y., Kosuge, N., Zhang, M., Hu, Y., 2023. Data-driven framework for extracting global maritime shipping networks by machine learning. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2022.113494.

Luong, T.N., Hwang, S., Im, N., 2021. Harbour traffic hazard map for real-time assessing waterway risk using marine traffic hazard index. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2021.109884.

Merkel, A., Kalantari, J., Mubder, A., 2022. Port call optimization and CO2-emissions savings - estimating feasible potential in tramp shipping. Marit. Transp. Res. http://dx.doi.org/10.1016/j.martra.2022.100054.

Michail, N.A., Melas, K.D., 2020. Shipping markets in turmoil: An analysis of the Covid-19 outbreak and its implications. Transp. Res. Interdiscip. Perspect. http://dx.doi.org/10.1016/j.trip.2020.100178.

Millefiori, L.M., Cazzanti, L., Zissis, D., Arcieri, G., 2016. Scalable estimation of port areas from AIS data. In: Vespe, M., Mazzarella, F. (Eds.), Proceedings of the Maritime Knowledge Discovery and Anomaly Detection Workshop. In: JRC Conference and Workshop Reports, pp. 48–51.

Mou, J.M., Tak, C.v.d., Ligteringen, H., 2010. Study on collision avoidance in busy waterways by using AIS data. Ocean Eng. 37 (5–6), 483–490. http://dx.doi.org/10.1016/j.oceaneng.2010.01.012.

Pallotta, G., Vespe, M., Bryan, K., 2013. Vessel pattern knowledge discovery from AIS data: A framework for anomaly detection and route prediction. Entropy 15 (6), 2218–2245. http://dx.doi.org/10.3390/e15062218.

Prochazka, V., Adland, R., Wolff, F.-C., 2019. Contracting decisions in the crude oil transportation market: Evidence from fixtures matched with AIS data. Transp. Res. Part A: Policy Pract. 130, 37–53. http://dx.doi.org/10.1016/j.tra.2019.09.009.

Rodrigue, J.-P., 2020. The Geography of Transport Systems, Fifth Edition Routledge, http://dx.doi.org/10.4324/9780429346323.

Rong, H., Teixeira, A., Guedes Soares, C., 2022. Ship collision avoidance behaviour recognition and analysis based on AIS data. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2021.110479.

Roubos, A., Groenewegen, L., Peters, D.J., 2017. Berthing velocity of large seagoing vessels in the port of rotterdam. Mar. Struct. 202–219. http://dx.doi.org/10.1016/j.marstruc.2016.10.011.

Sakan, D., Rudan, I., Zuskin, S., Brcic, D., 2018. Near real-time S-AIS: Recent developments and implementation possibilities for global maritime stakeholders. Sci. J. Marit. Res. 211–218. http://dx.doi.org/10.31217/p.32.2.6.

Slack, B., Comtois, C., Wiegmans, B., Witte, P., 2018. Ships time in port. Int. J. Shipp. Transp. Logist. (1), 45–62.

Styhre, L., Winnes, H., Black, J., Lee, J., Le-Griffin, H., 2017. Greenhouse gas emissions from ships in ports – case studies in four continents. Transp. Res. Part D: Transp. Environ. 54, 212–224. http://dx.doi.org/10.1016/j.trd.2017.04.033.

Tichavska, M., Tovar, B., 2015. Port-city exhaust emission model: An application to cruise and ferry operations in las palmas port. Transp. Res. Part A: Policy Pract. 78, 347–360. http://dx.doi.org/10.1016/j.tra.2015.05.021.

Toscano, D., Murena, F., Quaranta, F., Mocerino, L., 2021. Assessment of the impact of ship emissions on air quality based on a complete annual emission inventory using AIS data for the port of Naples. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2021.109166.

Tzannatos, E., 2010. Ship emissions and their externalities for the port of piraeus - Greece. Atmos. Environ. 400–407. http://dx.doi.org/10.1016/j.atmosenv.2009.10.024.

Wahl, T., Høye, G.K., Lyngvi, A., Narheim, B.T., 2005. New possible roles of small satellites in maritime surveillance. Acta Astronaut. 56 (1–2), 273–277. http://dx.doi.org/10.1016/j.actaastro.2004.09.025.

Wu, L., Xu, Y., Wang, F., 2020. Identifying port calls of ships by uncertain reasoning with trajectory data. Int. J. Geo-Inf. http://dx.doi.org/10.3390/ijgi9120756.

Xiao, F., Ligteringen, H., van Gulijk, C., Ale, B., 2015. Comparison study on AIS data of ship traffic behavior. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2014.11.020.

Yan, Z., Cheng, L., He, R., Yang, H., 2022. Extracting ship stopping information from AIS data. Ocean Eng. http://dx.doi.org/10.1016/j.oceaneng.2022.111004.

Yan, Z., Macedo, J., Parent, C., Spaccapietra, S., 2008. Trajectory ontologies and queries. Trans. GIS 12, 75–91. http://dx.doi.org/10.1111/j.1467-9671.2008.01137.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. Transp. Rev. 6, 755–773. http://dx.doi.org/10.1080/01441647.2019.1649315.

Zhang, Y., Fung, J.C., Chan, J.W., Lau, A.K., 2019. The significance of incorporating unidentified vessels into AIS-based ship emission inventory. Atmos. Environ. 102–113. http://dx.doi.org/10.1016/j.atmosenv.2018.12.055.

Zocholl, M., Iphar, C., Jousselme, A.-L., Ray, C., 2021. Ontology-based approach for vessel activity recognition. In: Proceedings of the OCEANS 2021 San Diego Conference. http://dx.doi.org/10.23919/OCEANS44145.2021.9705824.