Big Data Research 43 (2026) 100575

Contents lists available at ScienceDirect

Big Data Research

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

Techniques for interactive visual examination of vessel performance

Natalia Andrienko [a,b], Gennady Andrienko [a,b,*], Dimitris Zissis [c], Alexandros Troupiotis-Kapeliaris [c], Giannis Spiliopoulos [c]

[a] Fraunhofer Institute IAIS, Schloss Birlinghoven, Sankt Augustin, 53757, Germany
[b] City St George's, University of London, London, EC1V 0HB, UK
[c] Intelligent Transportation Systems Lab, University of the Aegean, Syros, Greece

A R T I C L E I N F O

Keywords:
Visual analytics
Trajectories
Vessel movement

A B S T R A C T

The development and evaluation of autonomous maritime vessels rely heavily on data-driven insights from iterative testing and analysis. While initial analyses are often conducted on small experimental datasets to explore key system characteristics, scaling these analyses to large datasets presents significant challenges. In this study, we extend our prior work on visual exploration of small-scale test bed data by proposing approaches to scaling the visual analytics techniques to large datasets. Using AIS data from ferry boats as a proxy for extensive maritime drone operations, we address the challenges of large-scale data exploration over eight days of repetitive ferry movements across a busy strait, simulating conditions suitable for autonomous vessels. Our approach investigates movement patterns, operational stability during repeated trips, and potential collision scenarios. To support such analyses, we propose a general, reusable workflow and a set of practical guidelines for applying visual analytics techniques to large maritime movement datasets. The findings highlight the scalability and adaptability of visual analytics methods, providing valuable tools for analyzing complex maritime datasets and advancing autonomous vessel technologies.

1. Introduction

In recent years, autonomous and semi-autonomous maritime vessels have seen growing adoption in diverse fields, ranging from marine research and environmental monitoring to maritime logistics [1,2]. These systems, commonly referred to as sea drones, are often deployed as test beds for advanced technologies, enabling iterative development and testing [3]. The testing and development of autonomous vessels rely on detailed examinations of movement trajectories and multi-dimensional sensor measurements. Effective visualization and interactive exploration of these data are essential for diagnosing malfunctions, understanding environmental influences, and refining performance. In addition to the inherent complexities of the trajectories and sensor data, maritime environments present unique challenges, including the influence of weather, external forces, and unconstrained two-dimensional movement, which complicate the analysis of vessel trajectories.

While small-scale datasets from experimental operations provide valuable insights, scaling the analysis to larger datasets is crucial for addressing challenges associated with real-world operations. This paper explores how the same analytical tasks can be effectively performed on both small and large datasets, demonstrating the adaptability and scalability of visual analytics techniques.

Another set of problems and tasks arises when exploring the potential for replacing human-operated vessels engaged in regular operations with autonomous ones. Autonomous vessels must not only be programmed to perform standard procedures, but also be capable of recognizing deviations from expected conditions and adapting their behavior accordingly. Achieving this requires the analysis of large volumes of historical data that capture both routine operations in the relevant domain and the irregular actions taken by human pilots in response to non-standard situations. In this paper, we demonstrate how visual analytics techniques can support the analysis of such historical movement data.

Building upon our prior work [4], which focused on the visual exploration of a small dataset from experimental maritime drones, we extend our approach to large data. As an example, we use a real-world set of AIS (Automatic Identification System, see [5]) data describing the operations of human-piloted ferry boats, which perform repetitive trips across a strait. These ferry operations represent scenarios that could realistically be handled by autonomous systems. To make this happen,

* Corresponding author.
E-mail addresses: natalia.andrienko@iais.fraunhofer.de (N. Andrienko), gennady.andrienko@iais.fraunhofer.de (G. Andrienko), dzissis@aegean.gr (D. Zissis), alextroupi@aegean.gr (A. Troupiotis-Kapeliaris), gspiliopoulos@aegean.gr (G. Spiliopoulos).

https://doi.org/10.1016/j.bdr.2025.100575
Received 29 January 2025; Received in revised form 23 April 2025; Accepted 23 November 2025
Available online 1 December 2025

analysts need to use the historical data to explore the potential behaviors, operational stability, and situational responses that future autonomous vessels might require. At the same time, this dataset can also serve as a meaningful proxy for large-scale maritime drone data.

We propose and implement interactive visual techniques designed to address the following analytical tasks:

- Detailed investigation of movement characteristics and sensor measurement recordings, focusing on detecting anomalies and unwanted behaviors.
- Assessment of operational stability during repeated movements, including consistency and deviations.
- Exploration of collective vessel movements to identify and analyze (a) potential collision situations or (b) maneuvers undertaken to avoid such situations.

We begin by applying these tasks to a small experimental dataset to explore the capabilities of maritime drones in controlled scenarios. This case study serves, first, to demonstrate the contents and properties of vessel movement data, and, second, to introduce the interactive visual interfaces enabling exploration and analysis of such data. We then describe another case study demonstrating how the same tasks are adapted to the larger AIS dataset. For this purpose, we rely on computational processing combined with visual exploration to uncover patterns, highlight deviations, and provide actionable insights for developers and operators of autonomous maritime systems. We then propose a general, reusable analysis workflow and a set of guidelines for scalable application of visual analytics techniques to large maritime traffic datasets.

The remainder of this paper is organized as follows. We first review related work on maritime data visualization and visual analytics for spatio-temporal data. Next, we describe both datasets, explain how we process them, and demonstrate how they can be analyzed using our proposed techniques. Finally, we summarize our findings, discuss the scalability of the methods, and outline future directions for supporting the analysis of autonomous vessel operations across varying data volumes.

2. Related work

The work in [6] highlights a critical gap in the domain of drone technology and robotics, emphasising the absence of visual analytics tools for effective analysis of multidimensional spatio-temporal data. In essence, this deficiency poses significant challenges to users seeking to monitor, comprehend, and control the behaviors of individual drones and drone fleets. Analysis tools should facilitate exploration and analysis of drone telemetry, trajectory data, environmental variables, and other kinds of information, thereby enabling users to gain actionable insights into drone functioning.

Drones are produced by various companies, each employing proprietary tools and data formats that lack compatibility, making data sharing challenging. Consequently, competitions such as autonomous boat race [7,8] serve as valuable platforms for gathering real-world datasets due to the limited availability of such data from proprietary sources.

Drone data consist of sequences of time-stamped geographic positions in 2D or 3D, annotated with measured attributes such as speed and direction, as well as characteristics of the moving object (e.g. weight or fuel consumption) and characteristics of the environment (e.g. wind speed and direction, water current attributes etc.) Such data are typical for mobility data science [9] and visual analytics of movement [10], with variety of analysis methods proposed in the literature. A framework for assessment of movement data quality was proposed in [11] and implemented as a protocol in a form of a Python library by [12]. The protocol addresses missing data, precision, consistency, and accuracy problems in respect to spatial, temporal, and attributive data components on the level of elementary data records, intermediate segments of trajectories, and overall trajectories and sets of them.

A large body of literature propose methods specific for maritime traffic (for example, [13]) and, more generally, in transportation studies [14]. The number of visual analytics papers proposing various approaches for analyzing movement data is very large and continues growing. Some of them deal specifically with data describing movements of vessels. Variants of dynamic density maps combined with specialized computations and techniques for interaction [15–17] support exploration of the density as well as other characteristics of maritime traffic. Kernel density estimation can be used to compute a volume of the traffic density in space and time [18], which can be visualized in a space-time cube [19] with two dimensions representing the geographical space and one dimension the time. [20] propose special glyphs for visualizing maritime data. [21] employ visual and interactive techniques for analyzing vessel trajectories together with weather data. [22] use vessel movement data to demonstrate the work of an interactive query tool called TimeMask that selects subsets of time intervals in which specified conditions are fulfilled. This technique is especially suited for analyzing movements depending on temporally varying contexts.

Similarly to drone and ferry boats, it is common for public transport to follow repeatedly the same route. Such settings allow visually-driven analysis of dynamics along the route and varying frequency of departures from origins towards destinations. [23] apply a 3D view to show similar trajectories as bands stacked on top of a map background. The bands consist of colored segments representing variation of dynamic attributes along the routes. Similarly, Itoh et al. [24] use stacked bands for displaying attributes along selected subway lines. [25] applied a complementary approach, building a kind of 2D matrix, with rows corresponding to stops along a bus line, columns representing minutes of the day, and colored cells reflecting aggregated attributes of movement such as average delay.

The literature study suggests that the following instruments are necessary for analyzing real-world vessel traffic data:

- Methods for pre-processing the data, including cleaning, stop detection, and division into trips;
- Clustering of trips by similarity for identifying repeatedly used routes and separating them from others;
- Representation of trips and attributes along them on maps (in space), space-time cube (in space and time) and along the time line;
- Interactions for focusing on particular aspects, including filtering by area, time, identities, and attributes;
- Transformation of times in trajectories (aligning them to common starts or ends) for their comparison;
- Detection of interactions between trajectories.

These techniques proved to be useful in our analyses.

3. Datasets used in this study

For this study we have utilized vessel trajectory data from distinct use cases. The first includes trajectories from small surface vessels that operate autonomously, as part of specific missions related to a University-level race. The other refers to movements of larger ferry vessels, traveling regularly between two ports. In this section we provide more details regarding each dataset.

3.1. The aegean race data

The 1st Aegean [7] Race (Autonomous Robotic Vessels Competition) took place in the island of Syros (Greece) in July 2022. The university-level competition was organised by the Intelligent Transportation Systems laboratory of the University of the Aegean and aimed to promote innovative ideas for smart shipping technologies. The student teams designed and developed autonomous robotic vessels on their own [26]. They competed, under real sea conditions, in speed, endurance and obstacle avoidance challenges, where their vessels had to operate completely autonomously without any interference by the users. Similar

Fig. 1. Analysis of the regularity of the position recording. The lengths of the time intervals between the recorded positions are represented by proportional sizes of circle symbols. The largest circle correspond to a time gap of 23 seconds, whereas the regular interval length is 1 second.

to a sailing regatta, the first challenge had vessels to perform a single round trip, bounded by three buoys, thus testing their speed capabilities on short, predefined trips. The second, collision avoidance, aimed to demonstrate the ability of the vehicles to detect, and effectively avoid, obstacles on their path, including other moving vessels or static objects (scattered buoys). Finally, the third challenge focused on the endurance of the vessel and its systems for voyages of longer duration. For this purpose, a round trip between two buoys was followed, with vessels performing as many laps as possible in the extended time frame, without stopping. The resulting data set consists of positional and mobility data of 3 vessels during all 3 challenges.

The data set has high temporal precision, with positions recorded almost every second (Fig. 1 shows an example trajectory), resulting in over 6900 positional reports. However, using GPS coordinates with only 7 decimal points lacks the precision needed to accurately track movement in small areas when recording data at a temporal resolution of about 1 second. This limitation can lead to distortions on maps, such as checkerboard-like patterns, and sudden fluctuations in derived movement metrics like speed, acceleration, direction, and turns.

3.2. Trajectories of ferry boats and surrounding traffic

For the second case study, we use a set of AIS data from 577 vessels that appeared in the part of Saronic Gulf near Athens from August 1, to August 9, of 2024 (however, the data for August 8, are missing). The study area, shown in Fig. 2, includes the Salamis Strait, a busy waterway used for ferry routes and other vessel traffic connecting Athens to the Salamis Island and beyond. From this dataset, we focus our analysis on the trajectories of 23 ferry boats circulating across the Salamis Strait. The data from the other vessels are used in analyzing the spatio-temporal context of the movements of the ferry boats.

The trajectories of the ferries include in total 72,837 position records. The temporal resolution is predominantly 60 seconds (in about 80 % of the records). In 99.6 % of the records, the time interval to the next record does not exceed 3 minutes, but there are also large time gaps indicating missing data.

Additionally to the vessel coordinates and time stamps, the data include recorded values of the speed over ground (SOG), course over ground (COG), and heading. However, in 93.2 % of the position records, the heading attribute has a special value 511, which indicates absence of data. Among the values of SOG, there are obvious outliers, such as 102.3 knots, while the values in 96.2 % of the records range from 0 to 11.8 knots.

For the analysis, we needed to divide the continuous trajectories spanning over several days into discrete trips from one to the other side of the strait. For this purpose, we used the boundaries of the stopping areas (in particular, ports) identified earlier at the stage of data preprocessing described in Section 4. We divided the full trajectories into segments starting and ending in the stopping areas located on the western and eastern sides of the strait (Fig. 3, left). After skipping several segments consisting of only a few points with large time gaps in between, i.e., missing most of the points along the routes, we obtained 1536 trips, including 768 eastward and 766 westward trips shown in Fig. 3 using purple and green colors, respectively. Besides, there we have got two anomalous trip trajectories. One is starting in the middle of the route; evidently, the remaining positions from this trip were missing in the original data. In the other anomalous trip, the boat apparently turned back without entering the destination port on the eastern side.

4. Preprocessing of big data

In order to analyze the vessel voyages data, the input trajectories first need to be properly formatted and prepared. This would entail identifying instances with faulty feature values, removing unnecessary data, like duplicates, and annotating parts of the movement. Overall, this preprocessing step pertain to both individual positional records and full trajectory entries, while we assume that we have continuous sequences of messages from a very large number of vessels. The data records consumed by this mechanism may be historical (collected earlier) or come in real time.

For processing historical AIS data, a distributed map-reduce approach (e.g. Apache Spark[1]) is typically selected to extract routes from AIS positional reports [27,28]. Different processing tasks upon large-data can be materialized through this method, including trajectory compression [29], anomaly detection [30] or even clustering [31]. Popular processing frameworks, such as Apache Spark, transform traditional data processing pipelines into a sequence of map-reduce tasks. Operations such as filtering and de-duplication, as well as pairwise calculations, are supported through mapping, sorting and redistribution of data in distributed nodes is happening through the shuffling process, while aggregation and collection of results is happening in the reduce phase. Common operations such as selection, filtering, application of simple mathematical transformations and window operations are integral to the apache Spark application interface. In addition, complex and custom transformations are also supported through the definition of custom user defined functions (UDF), i.e custom functions that are applied at record level in the distributed dataset. In this context, the proposed pre-processing methodology can be easily transformed into a big data pipeline by executing four major steps.

[1] https://spark.apache.org/

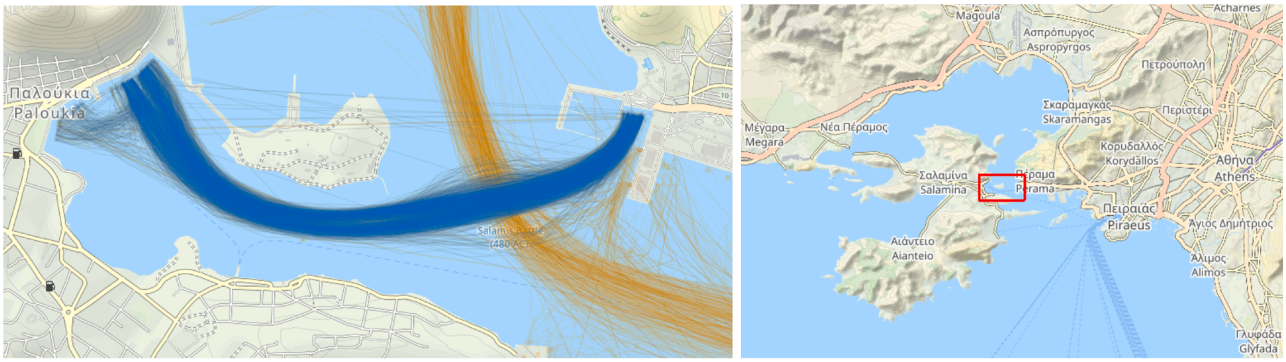

Fig. 2. Left: Salamis Strait with trajectories of ferries and other vessel traffic. Right: the position of the study area in a larger geographic context.

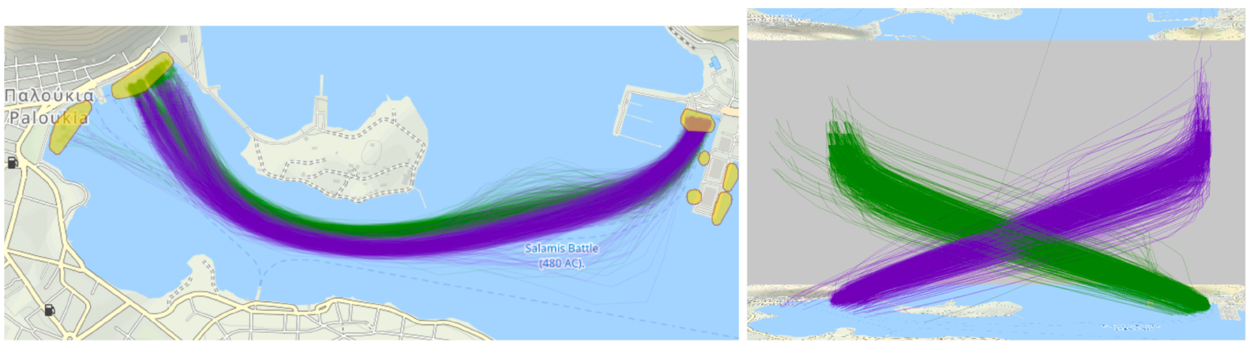

Fig. 3. Left:The trajectories of the ferry boats have been divided into trips between port areas. Right: The trajectories representing the trips have been aligned by their starting points and visualized in a space-time cube display, where the vertical dimension represents time.

First, we filter out all records that have erroneous values in regard to position (incomplete coordinates, latitude near poles, or zero are typical values indicated erroneous GPS reception from AIS transceiver), time (difference in timezones), identifier (non-AIS compliant MMSI specifications), as suggested by [11,12]. Provided that coordinates and timestamps can be trusted and the dataset is sorted in ascending time, errors in speed (extreme values of speed for vessels typically indicate errors) and direction (missing values or default) can be reconstructed. Specifically we can combine window operations and UDFs to calculate speed from the distance and time interval between two sequential positional records, while heading or course correction can be tackled by solving the inverse geodetic problem inside a UDF, for the same pair of records.

Secondly, to detect locations that vessels stop, we use a speed threshold. A slight movement of a vessel is reported to AIS even when at anchor due to GPS inconsistencies and the slight movement vessels perform around their anchor point. The selected speed threshold (e.g. less than 0.5 knots) indicates that the vessel is not underway using its engines. Later, this filtering process is combined with geo-fencing techniques to ensure that the detected stops are located within a port or anchorage area. If the geometries of such areas are not available in the form of an external dataset, they can be extracted directly from the data [32]. To achieve that, a separate dataset that contains only stopped vessels' records is extracted. This extracted dataset is much smaller than the original dataset, thus a density clustering approach (DBSCAN [33]) can be used to group together stopped vessel records and then form solid areas that vessels stop. The effectiveness of DBSCAN-based approaches free space movement data has been taken advantage in a multitude of other works regarding vessel traffic [34–36].

As a third step, we identify the sequences of vessels positions that indicate a complete trip. This is achieved through the respective stopping areas geometries, like ports. More precisely, all records are first organized in partitions based on the vessel identifier (i.e. MMSI) and sorted by time. Then, we join the records dataset with the geometries dataset so that all records that spatially intersect with a port/anchorage area are annotated with its identifier. For each vessel, all messages that are chronologically in between two stops are annotated using window functions with the same trip identifier.

Finally, similar trips are organized in groups, using as group identifier the vessel type, provided by AIS, origin and destination location of each trip. All trip trajectories within each group are sorted based on their length and duration. The representative or baseline trajectory for each group can be selected as the one with either median duration or length, or even apply more sophisticated methods that include sparsity into the selection as presented in [37].

5. Investigation of the experimental dataset

5.1. *Detailed exploration of individual trajectories*

Possible objectives of a detailed exploration of individual trajectories include inspection of the characteristics of the movement, position recording, and measurements. The most common visualization of trajectories is by lines on a map, as in Fig. 4, left. An animated map can show the progress of the movement over time but not the overall shape of the trajectory. A space-time cube [38], as in Fig. 4, right, shows the relative times of different segments of the trajectory, as well as the movement directions and speeds. The speed in a trajectory segment is indicated by the inclination of the corresponding line: the smaller the inclination, the higher the speed.

To explore the details of the position recording and sensor measurements, it is useful to combine the representation of the trajectory by a line with representing the recorded points by symbols, such as dots. Sizes and/or colours of the symbols can encode recorded measurements, as, for example, the speed in Fig. 5, left, or computed variables, such as the time to the next point in Fig. 1. The positions of the point symbols on a trajectory line indicate gaps in measurements and reveal line segments resulting from interpolation between known positions. Such estimations may significantly differ from the unknown actual path.

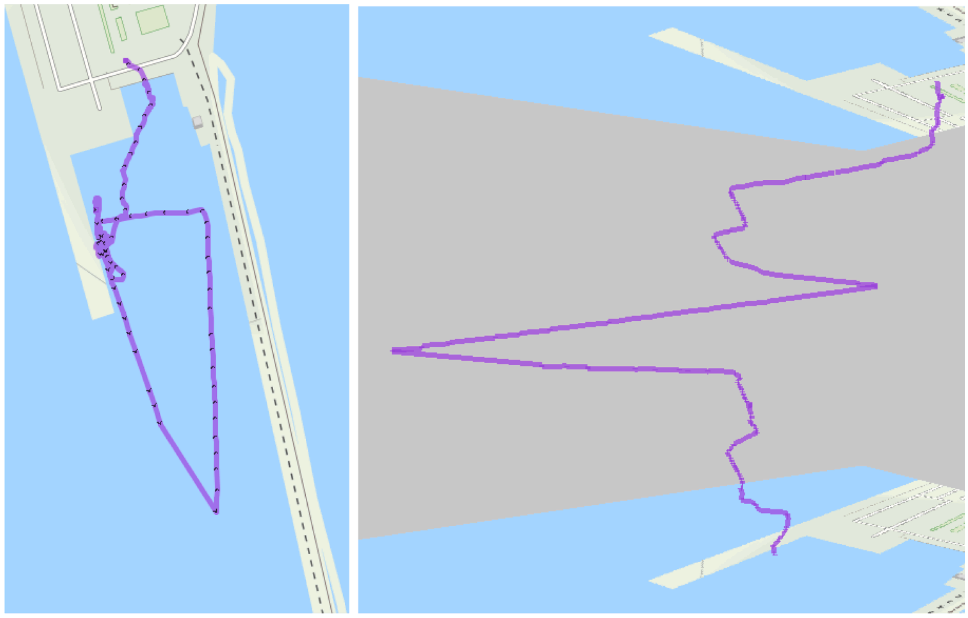

Fig. 4. A single trajectory represented on a map (left) and in a space-time cube (right). The time axis in the space-time cube is oriented upwards.

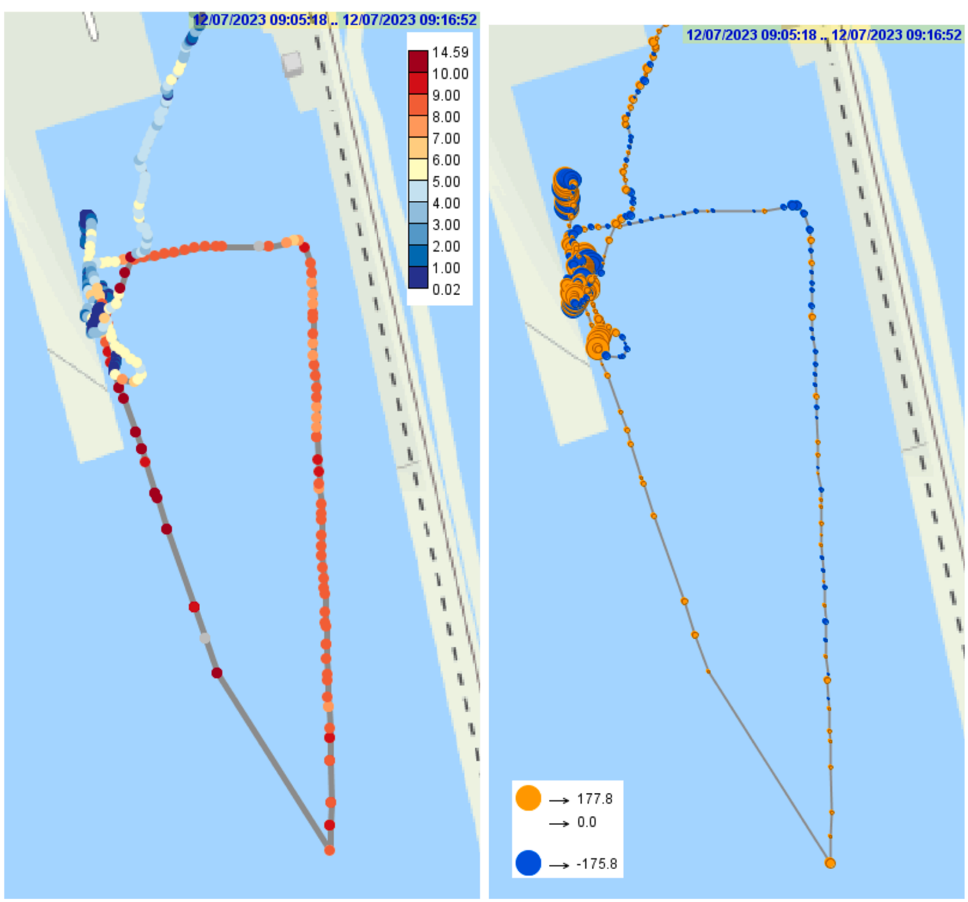

Fig. 5. Exploration of movement characteristics. Left: speed measurements are represented by point colouring. Right: deviations of the movement direction (computed from consecutive positions) from the vessel heading (recorded during the movement) are represented by proportional sizes and colours of circle symbols. Orange symbolises deviations to the right and blue to the left.

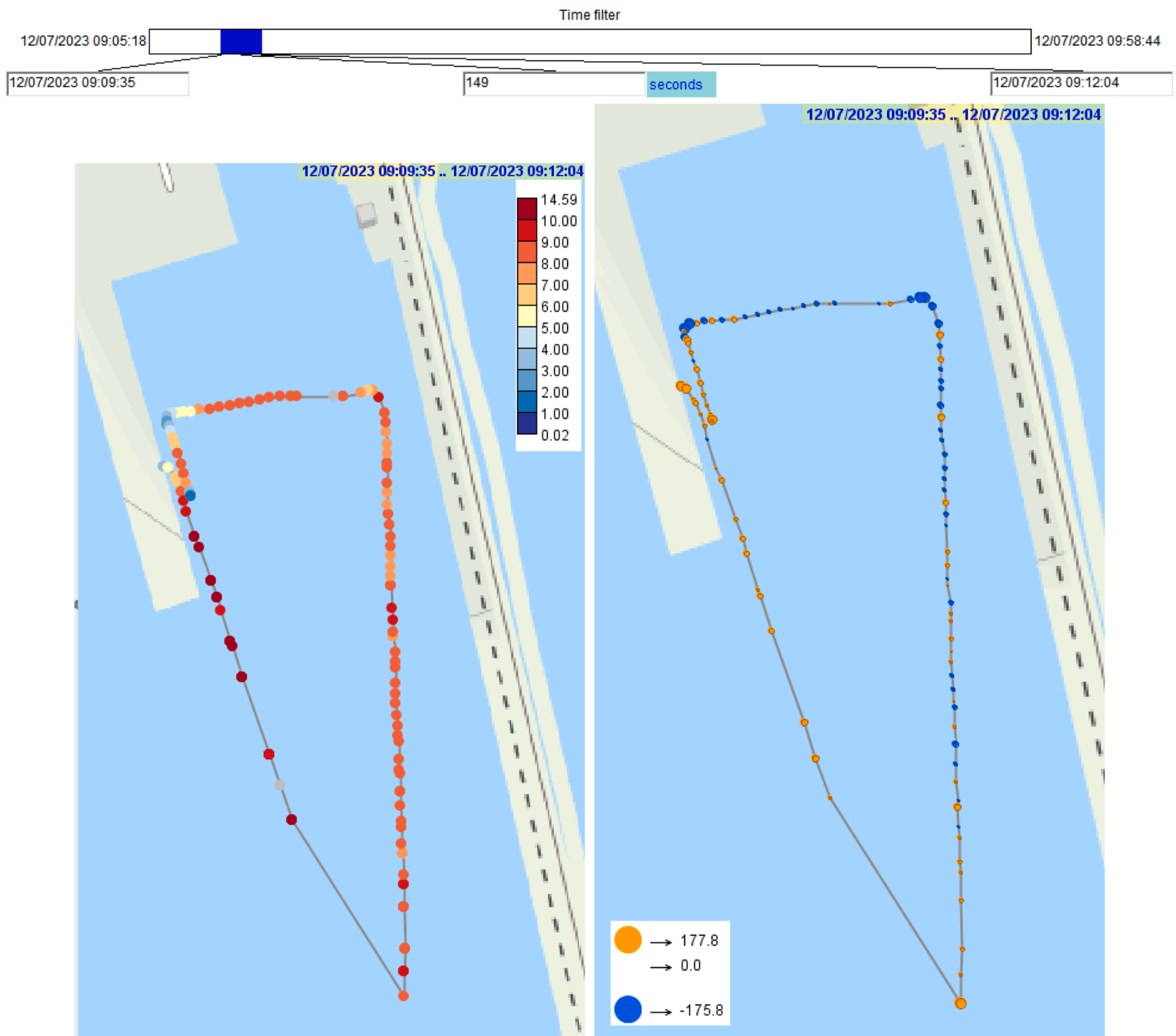

Fig. 6. Selection of a relevant part of the trajectory by means of temporal filtering.

Visualisation of speed and course data along a trajectory can also give a hint about the impacts of wind and waves on the vessel movement. Thus, we see on the map on the left of Fig. 5 that the speed of the southward movement was notably higher than in the movement to the north, which shows the impact of the wind blowing from the north and northeast. The impact of the wind on the vessel course can be explored by calculating and visualising the differences between the recorded vessel heading and the movement course computed from consecutive vessel positions. On the right of Fig. 5, the deviations are represented by dot symbols with the colour (blue or orange) encoding the direction of the deviation (left or right of the heading) and size proportional to the amount of the deviation, in degrees.

A necessary tool for interactive exploration of trajectory data is time filter allowing selection of time intervals for viewing only data generated in these intervals while the remaining data are hidden. The work of a time filter is illustrated in Fig. 6, where it was used to hide irrelevant parts of the trajectory that reflect the vessel movements before and after the race. The filter was applied to the data presented in Fig. 5. We see that the speed during the race was mostly quite high and the deviations of the course from the heading were low compared to the hidden parts of the trajectory that were visible in Fig. 5. Still, the orange and blue colours of the dot symbols signify the impact of the northern and north-eastern wind: the course slightly deviated to the right of the heading during the southward movement and to the left during the northward and westward movements.

In a similar way, one can explore any sensor measurements taken by the vessel along the route. To summarise, basic techniques for visual exploration of individual trajectories and associated point-based measurements include representation of the trajectories by lines on a map and in a space-time cube, using point symbols for showing the locations of the recorded trajectory points and any attributes associated with the points, and time filter for selection of time intervals and corresponding trajectory parts to focus on.

5.2. Exploration of repeated movements

During development and testing, drones often need to perform repeated tasks, following the same pre-defined route. Some variations of the route may occur due to changes in context such as weather conditions, activities of the drone itself, other events that happened nearby (e.g. proximity to stationary obstacles or other moving objects) etc. Examples of such data have been collected during the so-called endurance race, see Fig. 7. Similarly to Fig. 4, the space-time cube in the middle shows dynamics of the 3 trajectories. In the bottom, trajectories are divided into repeated fragments and their starting times are aligned.

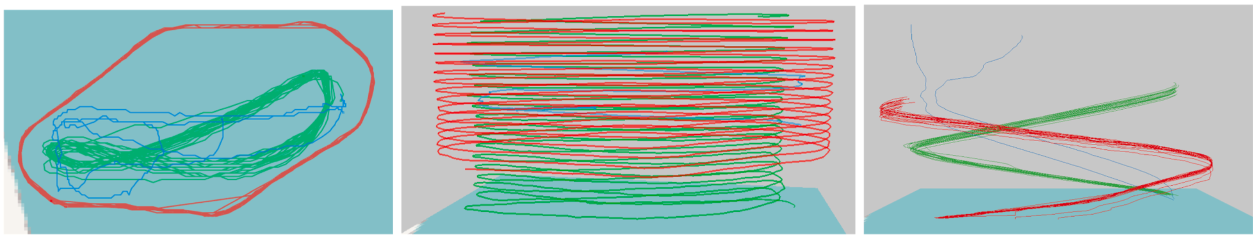

Fig. 7. Endurance race: trajectories of the 3 drones on the map (left) and map space-time cube (middle and right).

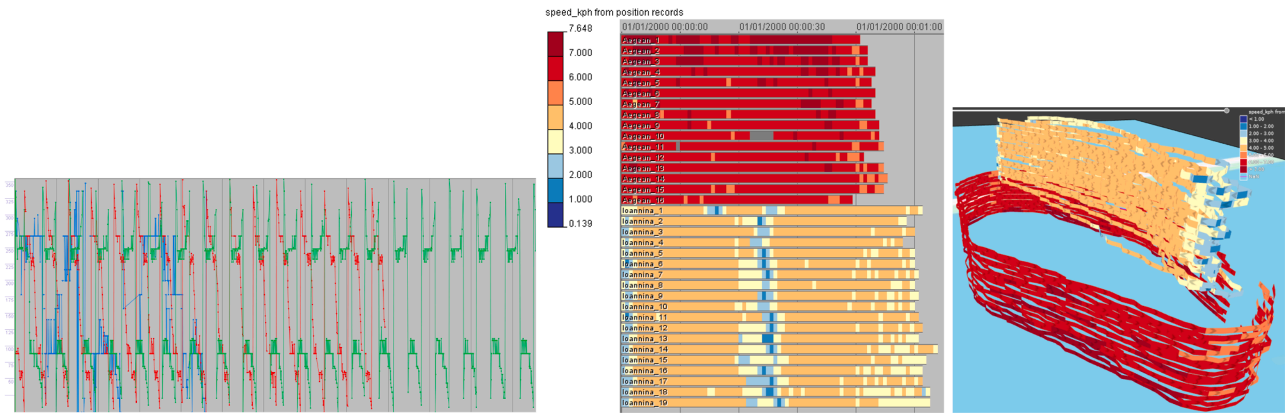

Fig. 8. Endurance race: dynamics of speeds over multiple loops. The display in the left shows dynamics over time; segmented time bars in the middle align starts of all loops; trajectory wall display [39] in the right shows speeds in their spatial context.

Analysis of repeated movement is not limited to purely spatial and spatio-temporal shape matching. In addition, it is necessary to study the dynamics of attributes for the whole trajectories and their dynamics within the trajectories. Thus, by computing average speeds over multiple fragments we observed gradual speed decrease over the sequence of loops for each drone, indicating their degrading performance. More detailed analysis can be done using time series displays, as shown in Fig. 8. Such displays are suitable for understanding the overall dynamics of movement in the repeated fragments and for identifying times and locations of speed changes, as well as sporadic fluctuations. In further analysis, these patterns can be matched to context data (e.g. weather attributes) or events of proximity to stationary obstacles or other vessels.

5.3. Exploration of interactions

Here we focus on the task of detecting and exploring events of close approach of vessels to other static or moving objects; we shall call such events *interactions* [40]. Interactions can be detected by computing the minimal distance from each point of a vessel trajectory to the boundary or location of another object at the time of attaining this point.

In computing the distance to a moving object, it is necessary to take into account the possible differences between the time moments when the locations of the given vessel and the other object were sampled. Thus, for a vessel position measured and recorded at time moment t there may be no position in the trajectory of another moving object having exactly the same time reference. Therefore, in computing the distances, it is necessary to take a temporal buffer $[t-\epsilon, t+\epsilon]$ around each position of the vessel trajectory, find the points from the other trajectory where times fit in this time interval, and compute the distances to all these points. The temporal threshold ϵ is chosen based on the coarsest temporal resolution of the position recording among all trajectories involved in the calculation.

To detect close approaches, it is also necessary to define what distance between objects can be treated as a close approach, i.e., to set a distance threshold δ. It is chosen depending on the sizes of the vessel and the objects that can be approached during the vessel movement.

As an example, we show results of detecting interactions between two autonomous vessels during a race using the threshold settings $\epsilon = 5$ seconds and $\delta = 1$ m.

Exploration of detected interactions requires them to be represented visually on a map, as, for example, in the middle of Fig. 9. The points of close approach are marked by dot symbols and connected to the corresponding points from the other trajectory by lines. Another visual representation is a space-time cube, as in the lower part of Fig. 9. It shows the approximate relative times of different interactions.

However, occlusions and line intersections in both the map and the cube complicate the examination of the details of the interactions. This problem can be solved using time filtering, as illustrated in Fig. 10. For convenience, a time interval containing one interaction can be selected using a mouse operation within the map display.

6. Investigation of the performance of ferry boats

In this case study, we discuss how to scale the analysis to large trajectory datasets, where detailed examination of individual trajectories becomes impractical. Nevertheless, it is important to retain the use of interactive visual interfaces, which enable interpretation of the information based on domain expert knowledge and support the detection of unforeseen situations and patterns that are difficult to identify algorithmically.

As mentioned earlier, analyzing historical data reflecting the behavior of human-piloted vessels can provide essential insights for developing autonomous alternatives. By understanding how routine operations are performed and how human operators respond to irregular situations, developers can better define the expected capabilities of autonomous systems. Such analysis is particularly valuable when large-scale data from autonomous vessels is not yet available.

Using the AIS data described in Section 3.2, we focus on ferry boat trips across the Salamis Strait. Our approach combines data aggregation and computational methods to detect and extract anomalies or disruptions in regular movements, enabling targeted, detailed investigations.

id	Start time	End time	Duration	Spatial extent
1	12/07/2023 11:28:36	12/07/2023 11:28:43	7	1.440
2	12/07/2023 11:30:24	12/07/2023 11:30:29	5	0.987
3	12/07/2023 11:30:48	12/07/2023 11:30:53	5	0.660
4	12/07/2023 11:35:55	12/07/2023 11:36:03	8	2.066
5	12/07/2023 11:37:01	12/07/2023 11:37:06	5	0.987

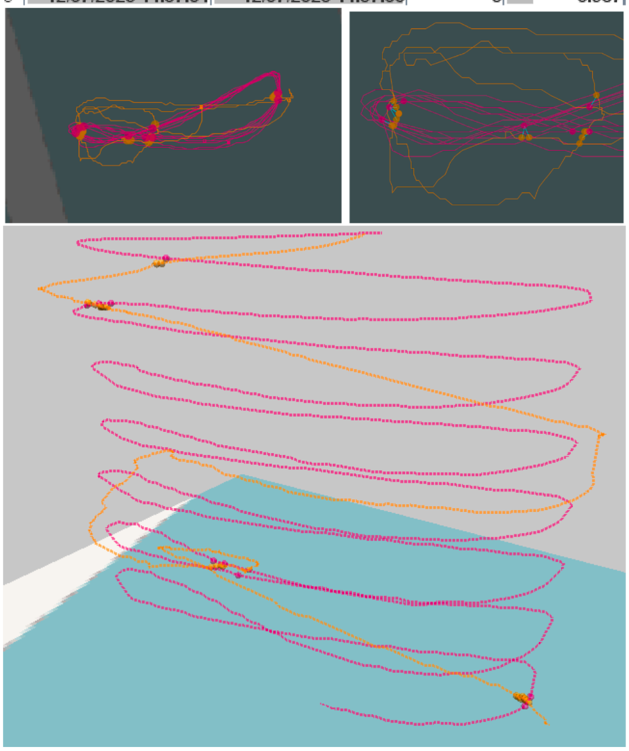

Fig. 9. Interactions between two autonomous vessels during a race. Top: a table describing the detected interactions. Middle: points of close approach are marked on a map. An enlarged map fragment is shown on the right. Bottom: the trajectories and points of close approach are displayed in a space-time cube.

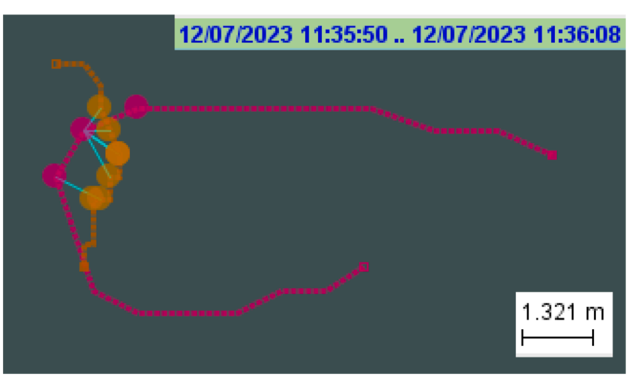

Fig. 10. One interaction has been selected for inspection by means of time filtering.

6.1. Detection and exploration of anomalous behaviors

A straightforward method for detecting anomalies in large trajectory datasets is through histograms that represent the distributions of relevant numeric attributes, such as trip duration and path length. This aggregated representation scales well to large datasets and effectively highlights outliers. For example, in Fig. 11, the histogram of trip durations (left) reveals a single extreme outlier of 61 minutes, while most trips last between 9 and 19 minutes. Similarly, the histogram of path lengths (right) identifies one unusually short and one unusually long trip.

The very short path corresponds to an incomplete trajectory, as mentioned in Section 3.2. The trip with the longest path also exhibits the longest duration. To investigate this outlier, we queried and visualized its trajectory on both a map and a space-time cube (Fig. 12). The map suggests that the ferry was approaching the eastern port but abruptly turned back without entering the port. However, the space-time cube reveals a significant time gap between two consecutive positions, represented as a steep diagonal line. A closer look highlights a 41-minute gap, marked on the map by a yellow circle. This indicates missing position records, suggesting the vessel likely entered the port, paused, and resumed its journey after the stop. Thus, these anomalies are artifacts caused by incomplete data rather than actual operational disruptions.

Apart from easily detectable outliers, there may be anomalies that are hard to detect in visualizations of aggregated data. Thus, in a bunch of trajectories connecting two ports, the shapes of some trajectories may deviate from the usually followed routes. Such trajectories can be noticed on a map and in a space-time cube, as in Fig. 3; however, this requires visual representation of each individual trajectory, which may not be feasible when the trajectories are very numerous. Besides, visual discrimination of typical routes from fluctuations requires high transparency in representing the trajectories. The typical routes become prominent on the map when many trajectories following these routes are drawn one on top of another. However, this approach makes deviating trajectories hard to notice due to their high transparency.

The problem of detecting unusual shapes or other kinds of anomalies that cannot be revealed using aggregated data displays can be solved by means of density-based clustering with an appropriate distance function. Thus, for distinguishing typical and unusual route shapes, one needs a distance measure assessing the spatial distance between trajectories [41,42]. It is also important to use a suitable clustering algorithm, such as OPTICS [43], which not only separates clusters from noise (i.e., objects dissimilar to others) but also associates clustered objects with two values: core distance and reachability distance. Objects with low core distances represent the most typical, frequently occurring cases. High reachability distances signify deviations from the cluster cores, i.e., from the typical cases. Hence, in addition to the noise, cluster members with high reachability distances may correspond to anomalous behaviors.

Applying OPTICS to the ferry trips dataset, we used the “route similarity” distance function ([10]) with a spatial threshold of 100 m and a minimum of 20 neighbors. Fig. 13 (top) shows the results. The standard routes are visualized in Fig. 13 (middle) and consist of trajectories with low core distances. Deviating routes, identified by high reachability distances or classified as noise, are shown in Fig. 13 (bottom). These trajectories are drawn with higher opacity for better visibility.

While density-based clustering is effective, its scalability is limited for very large datasets. To address this, clustering can be combined with classification [44]. In this hybrid approach, clustering is applied to a manageable random sample of data. Dense clusters representing frequent, typical behaviors are then used as training classes. The remaining data are processed using a classification algorithm, such as kNN, to assign instances to these classes or mark them as unclassified if they are dissimilar. Unclassified instances are flagged for further investigation as potential anomalies. This approach can also be applied to streaming data.

By combining aggregation, clustering, and classification, this methodology provides an efficient framework for exploring large trajectory datasets while ensuring that critical anomalies and disruptions are not overlooked.

6.2. Assessment of operational stability during repeated movements

After detecting abnormal or irregular behaviors, the next step is to investigate their locations and frequency. Not all deviations from standard routes are equally significant. For example, Fig. 13 (bottom) shows some irregular trajectories deviating from standard routes within or near port areas. These deviations typically comply with traffic regulations in

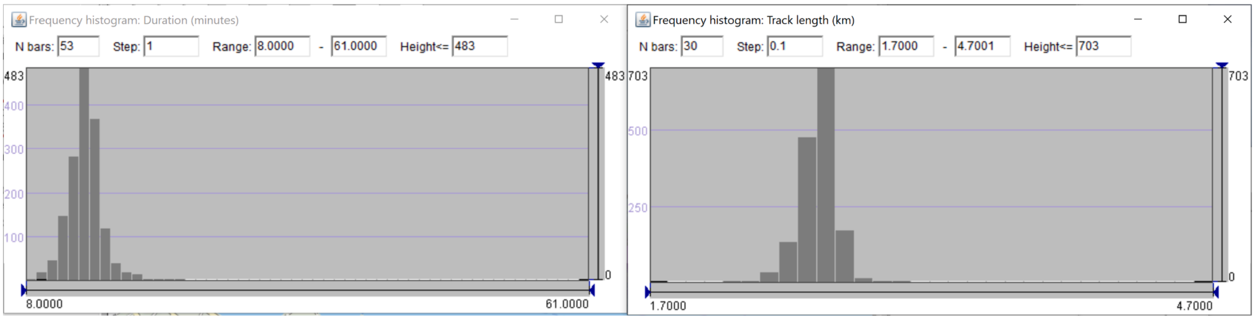

Fig. 11. Histograms representing the distributions of the trip duration and length.

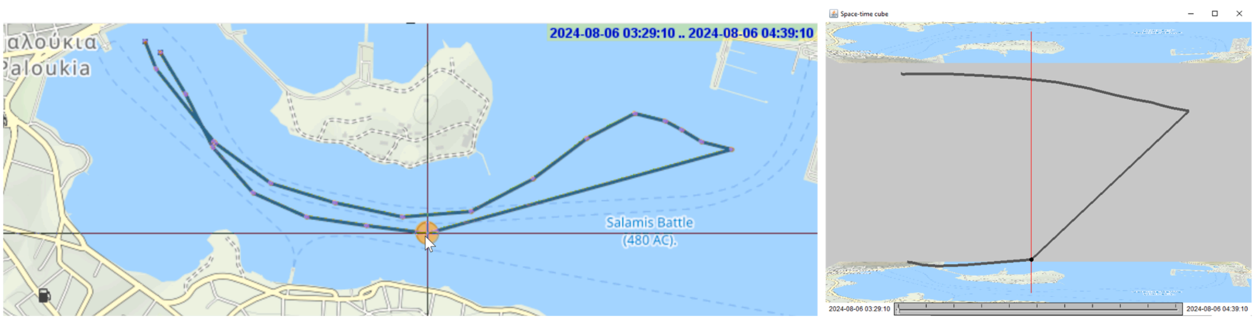

Fig. 12. Visual investigation of a detected anomalous trajectory on a map and in a space-time cube.

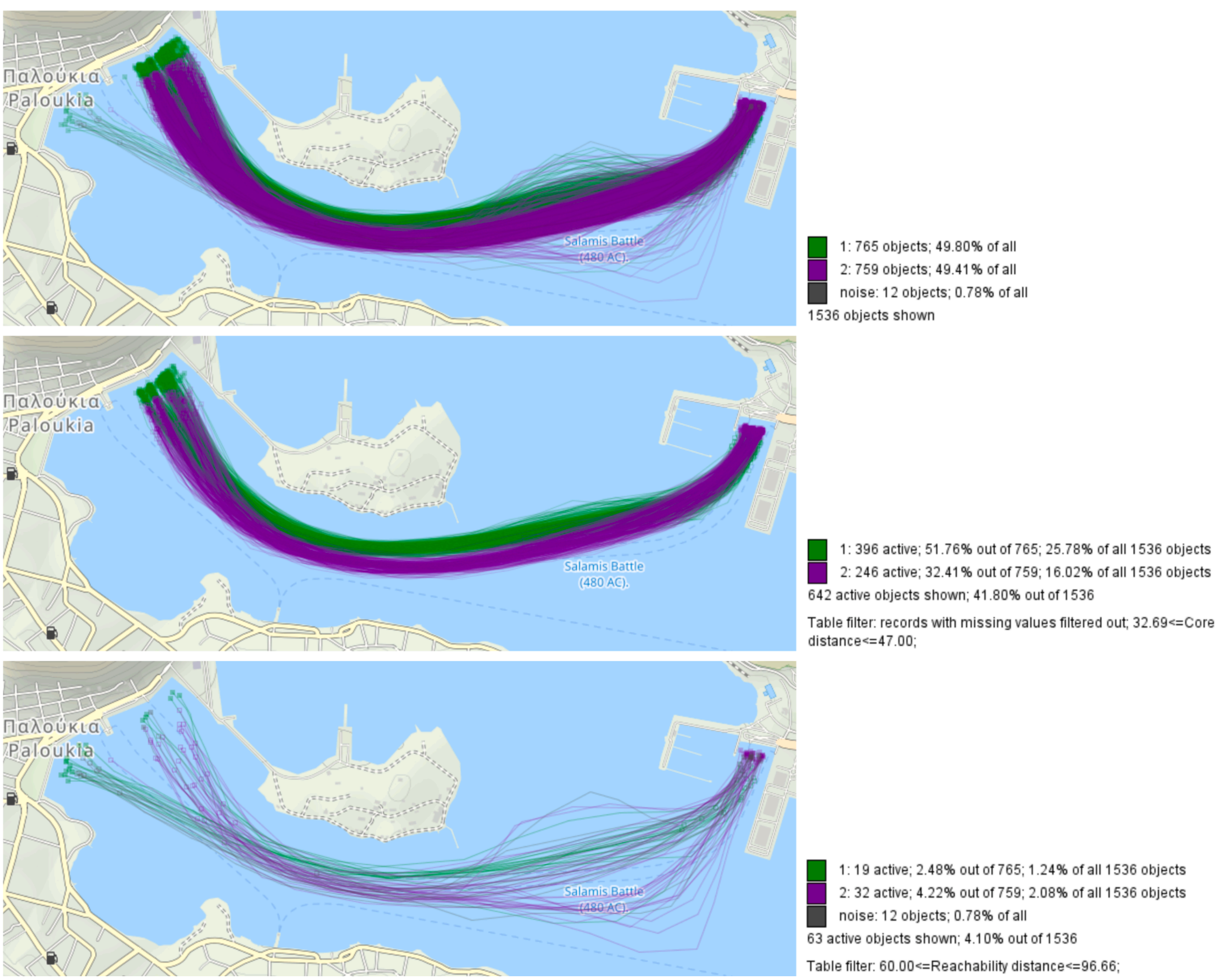

Fig. 13. Detection of regular and irregular routes by means of clustering using OPTICS.

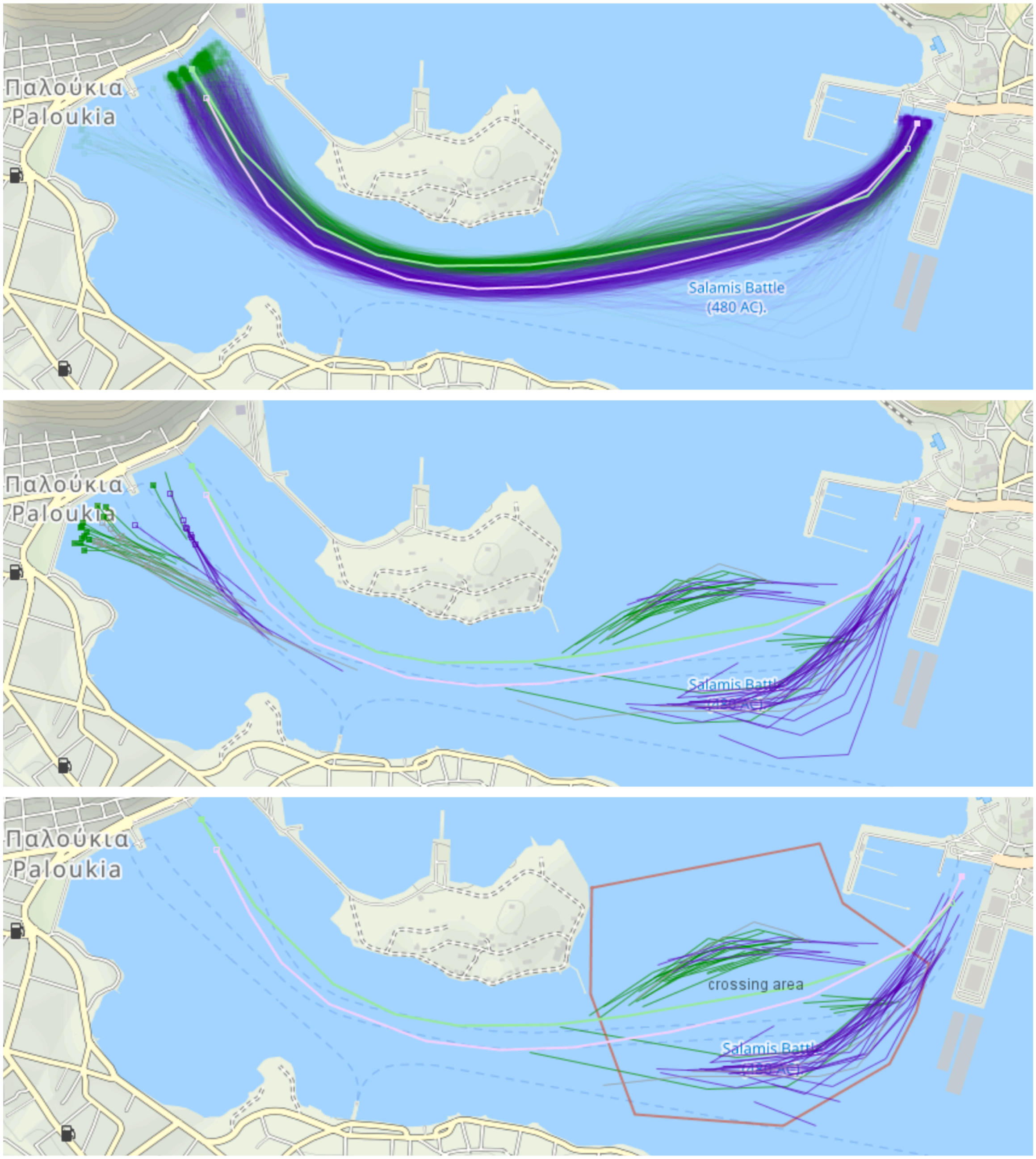

Fig. 14. Top: Central trajectories of the trip clusters constructed from averaged corresponding positions of the core cluster members. Middle: Trip segments deviating from the central trajectories by 100 or more meters. Bottom: Spatial filter to extract the deviating segments occurring in the crossing area.

ports and can be considered normal. However, deviations occurring in the middle of the strait, where traffic flows intersect (Fig. 2, left), may deserve special attention. Targeted analysis of such deviations requires their detection and extraction through computational techniques.

We propose an approach for extracting deviating trip segments, illustrated in Fig. 14. First, the central routes of the trip clusters are constructed using the core cluster members (Fig. 13, middle). The central route construction algorithm [45] groups spatially close points from different trajectories and calculates the spatial centers of these groups to form the central trajectory. In Fig. 14, top, the central routes for the green and purple clusters are shown as thick lines colored in light green and pink, respectively. This step is performed once after identifying the regular trajectory shapes, and the central routes are stored for subsequent use.

Next, for each point on a trip, its distance to the nearest segment of the central route is calculated. This operation can be applied to both historical and streaming data. Trip points and segments with distances exceeding a defined threshold (e.g., 100 m, as in Fig. 14, middle) are flagged as deviations. Finally, only segments falling within a predefined area of interest (e.g., the traffic crossing area in our case study) are extracted for further analysis, as illustrated in Fig. 14, bottom. This spatial filtering step is computationally efficient and scalable, making it suitable for streaming data analysis.

Once the deviating trajectory segments are extracted, they are subjected to detailed inspection. Since these anomalies are expected to be relatively infrequent, they can be effectively analyzed using visual and interactive techniques, supplemented by computational methods as needed.

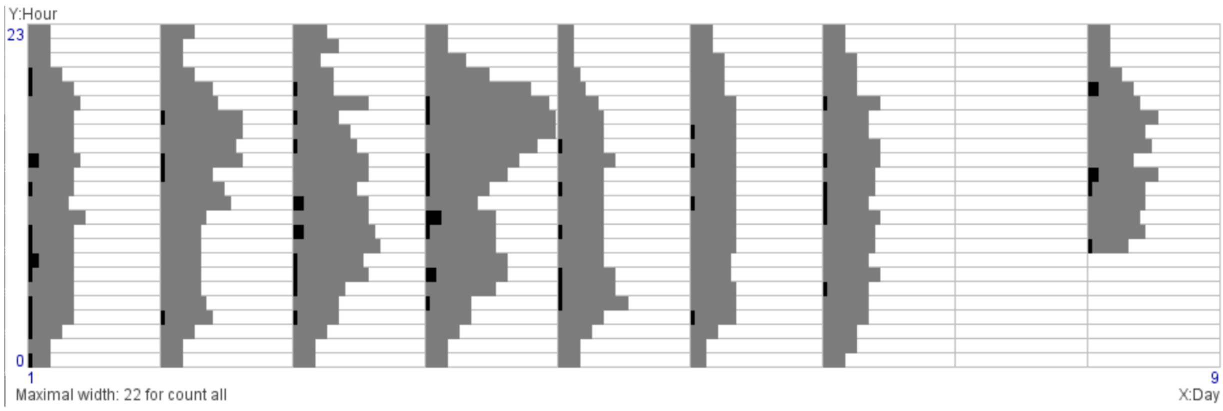

Fig. 15. 2D time histogram shows the distribution of the starts of the ferry trips by days (horizontal dimension) and hours (vertical dimension). The bar lengths are proportional to the counts of the trip starts in the 1-hour time intervals. The black segments represent the counts of the trips deviating from the regular routes in the traffic crossing area.

6.3. Examination of spatio-temporal contexts of normal and abnormal movements

The temporal distribution of deviating trips among all ferry trips can be explored using temporal histograms, such as the 2D time histogram shown in Fig. 15. In this visualization, the horizontal and vertical axes represent days and hours of the day, respectively. The length of the gray bars indicates the total counts of ferry trips starting within 1-hour intervals, while the black segments within the bars represent the counts of trips deviating from the standard routes within the traffic crossing area. These deviations correspond to trips containing the previously extracted abnormal segments. As observed in Fig. 15, no clear temporal patterns emerge in the occurrence of deviations, nor do these occurrences show any correlation with the overall number of ferry trips across the strait.

Deviations from standard routes could potentially be influenced by specific weather conditions, such as strong side winds. With access to relevant weather data, it is technically possible to correlate weather parameters with the timing of deviations. However, in this case, no continuous periods of heightened deviation frequency are evident, suggesting that weather is unlikely to be a significant factor.

The more probable cause of these deviations is the traffic conditions in the strait, as ferry boats may alter their routes to avoid collisions with other vessels. To investigate this hypothesis, we analyzed the relationship between deviations and overall traffic density in the crossing area. Using AIS data from all vessels in the study area, we computed hourly time series representing the number of vessels present in the crossing area and compared this with the time series of deviation occurrences.

Fig. 16 illustrates the use of 2D matrix displays for these time series. Each row corresponds to a day, and each column represents an hour of the day. The left panel shows the hourly counts of deviation events, while the right panel visualizes the z-scores of total vessel counts in the crossing area. Positive z-scores are displayed as gray squares, and negative z-scores as white squares.

Our analysis reveals no apparent similarity between the temporal patterns of the two time series. This lack of correlation suggests that the deviations are not driven by overall traffic intensity. Instead, these anomalies likely result from specific configurations of vessels and their movements within the spatio-temporal vicinity of the ferry boats. Investigating such localized interactions requires detailed examination of individual cases. Given that these situations are typically infrequent, interactive visual analysis is a feasible approach.

The maritime traffic domain knowledge includes the fact that, under specific vessel configurations, a vessel may need to deviate from their planned route in accord with the International Regulations for Preventing Collisions at Sea (COLREGs). The latter consists of a set of rules and mitigation actions that apply worldwide to avoid collisions at sea. The most relevant rule for the area we examine relates to crossing vessels, as shown in Fig. 3. The prescribed mitigation action for two powerboats (vessels using engines) crossing at a risk of a collision is that the vessel that approaches from the left side the other vessel must deviate giving priority to the other vessel.

In our case study, 59 deviations occurred over eight days, with a maximum of 14 deviations per day and two per hour. In a real-time monitoring scenario, anomalies can be examined as soon as they are detected, ensuring timely insights into potential navigation issues.

A visual examination of three distinct cases of ferries deviating from their standard routes is presented in Fig. 17. The upper images display map fragments, providing spatial contexts for each deviation. On top of each route a vessel marker is depicted. The colored segments in the front of the markers indicate the left and the right side of each vessel. In a crossing situation, the vessel facing the red side of another vessel must give priority and move behind the other vessel (this movement is indicated with an arrow in top left and center figure in Fig. 17). These maps show the trajectories of vessels moving within the crossing area at the times the deviations occurred. The screenshots of space-time cubes below the maps further illustrate the temporal progression of these situations.

In the first case, a ferry heading toward the eastern port performed a maneuver to avoid a collision with a vessel crossing its route and moving northward. In the second case, a ferry bound for the western port deviated northward to give way to a vessel traveling from the north. The space-time cube reveals that the positions of the two vessels were in close proximity at one point. However, this apparent closeness could be an artifact caused by the coarse temporal resolution of the AIS data and the linear interpolation applied between recorded positions, which were sampled at intervals of approximately one minute. Both deviations appear to be in-line with the mitigation actions COLREGs define.

The third case depicts a more complex scenario involving multiple ferry boats traveling in parallel. This situation forced one of the ferries not only to deviate from its standard route but also to reduce its speed to avoid potential conflicts. Again, in this situation according to COLREGs, when vessels have their maneuverability restricted (often due to coastline, bathymetry or other vessels nearby), they must coordinate with nearby vessels to ensure a safe passage.

The visual analysis of the ferries deviations showed that most deviations result from the application of the navigational rules for the area rather than an emergency scenario. Generally, for a deviation to comply with COLREGs in this area, we expect the vessel to deviate north when it moves towards west and south when it moves towards east. The map representation of the deviating segments in Fig. 14 shows that it is typically the case. Still, there are exceptions, which may deserve additional investigation using interactive visualizations.

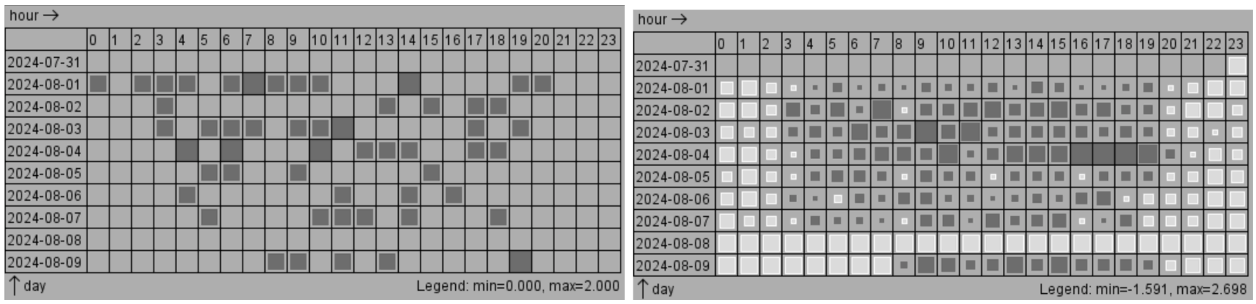

Fig. 16. 2D representation of the hourly time series of the counts of the deviations (left) and the overall intensity of traffic in the crossing area (right). The display on the right represents the normalized deviations of the vessel counts from the mean number of vessels in the crossing area. The positive deviations are shown in gray and negative in white.

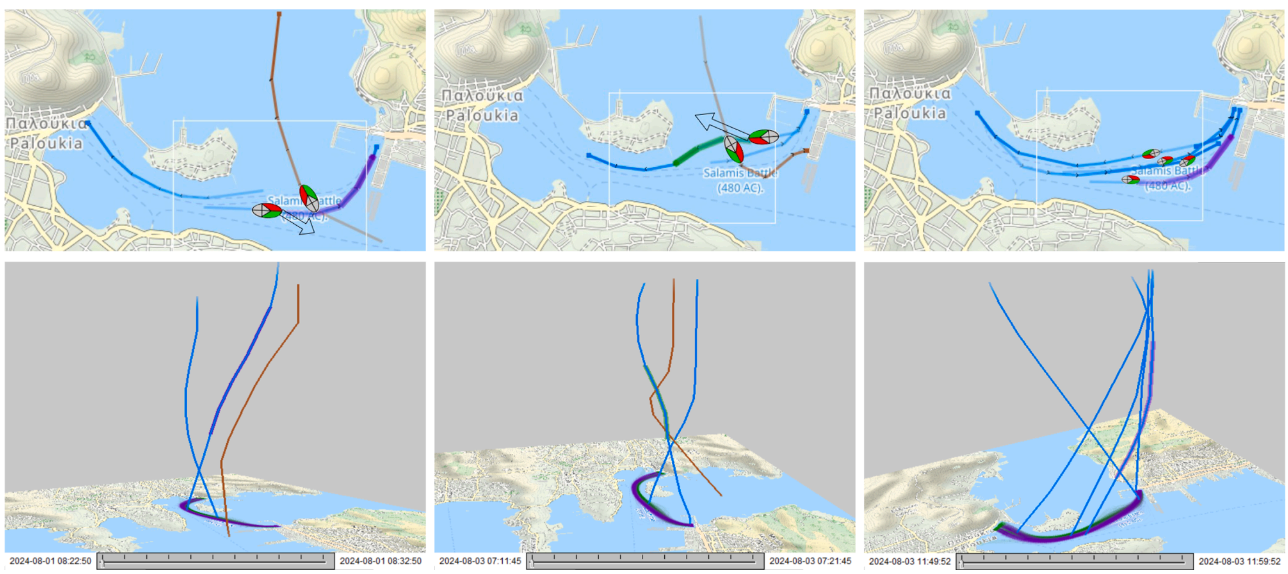

Fig. 17. Visual exploration of details of individual deviations from standard routes. The upper images are map fragments showing the spatial contexts of the deviations. Below them, the spatio-temporal contexts are represented in a space-time cube. The ellipsoid markers at the origin of each deviation on the map indicate each vessel's direction. The colored segments indicate sailing navigational lights that are located in the front part of each vessel, according to the International Regulations for Preventing Collisions at Sea (COLREGs) [46]. The vessels with arrows in the top-center and top-left figures indicate clear application of the COLREGs for powerboats crossing each other's routes. In crossing situations, a vessel approaching another vessel's left/red side (port side in nautical terms) must avoid crossing ahead of her, while the other vessel has to maintain her course. In the top-right figure, multiple vessels approach the port area simultaneously restricting significantly their maneuverability, forcing vessels to reduce speed or even stop to avoid collision. Vessels in restricted maneuverability typically coordinate their actions to ensure safety at sea.

7. Discussion

In this work, we proposed and demonstrated a set of visual analytics techniques for analyzing maritime vessel movements. We investigated the applicability and scalability of these techniques for different analysis tasks using both small experimental datasets and larger AIS datasets as proxies for real-world scenarios involving autonomous vessels. The discussion focuses on evaluating the methodologies, highlighting the challenges encountered, and considering the implications of the findings.

Scalability and adaptability of the methods. The techniques demonstrated in this study effectively scaled from analyzing a small experimental dataset of sea drones to a large AIS dataset of ferry movements across the Salamis Strait. Tasks such as anomaly detection, stability assessment, and spatio-temporal exploration were successfully scaled to larger datasets through computational preprocessing, clustering, and filtering. The combination of interactive visualization and computational processing proved essential for handling the volume and complexity of AIS data, while still enabling detailed examination of individual cases when needed. Table 1 summarizes our experiences as a general reusable workflow.

Insights into vessel behavior and anomalies. Interactive visual inspection is crucial for domain experts and maritime traffic operators to understand events and assess the criticality of unusual and complex situations based on their domain knowledge. Our case studies demonstrated how the visual analytics approach revealed valuable insights into both normal and abnormal vessel behaviors. For example, deviations from standard routes were identified as artifacts of missing data, responses to specific vessel configurations, or compliance with maritime traffic rules such as COLREGs. The ability to discern such cases highlights the importance of combining domain knowledge with visual analytics for context-aware interpretation.

Challenges in data quality and real-time application. The study underscored several data quality challenges, including missing position records, coarse temporal resolution, and potential inaccuracies in sensor measurements. These issues limited the granularity of certain analyses, such as determining precise collision risks. While the preprocessing methods mitigated some of these challenges, their resolution remains critical for advancing real-time applications. Although we described how the analysis workflows could be adapted to real-time monitoring, this requires efficient implementation of the com-

Table 1
Workflow for scalable visual analytics of maritime trajectory data.

Step of workflow	Task	VA techniques	Adaptation to large data
Data preprocessing	Clean data	Visualize attribute distributions to set thresholds distinguishing valid values from errors	Aggregate data before visualizing and use suitable visualizations (e.g., histograms)
Data preprocessing	Extract stop places	Visually supported density-based clustering of stop positions; creating areas around clusters	Automatically extract stop positions using a speed threshold; divide the study area into smaller spatial regions to process in parallel or iteratively
Data preprocessing	Extract complete trips between stop places of interest	Interactively select relevant stop places; visualize and explore results of automatic extraction	Visualize summary statistics of extracted trips (e.g., path length, duration); highlight extreme cases for inspection; set thresholds to exclude invalid results
Model normal behaviors	Identify subsets of normal trips for each stop pair	Density-based clustering of trips by similarity in shapes and movement dynamics; visualization to tune parameters	Apply clustering and visualization to manageable trajectory samples; sample from different time periods to verify consistency
Model normal behaviors	Define normal behaviors	Interactively select core cluster members and summarize into model trips	Construct model trips from sampled trajectories, assuming frequent typical patterns and diversity of anomalies
Detect and analyze anomalies	Identify anomalies	Automatically quantify each trip's similarity to the respective model; visualize distance statistics; interactively inspect deviating trips	For very large sets of anomalies, aggregate deviations by grid cells and visualize spatial distribution; use spatial queries to extract outliers from selected areas
Detect and analyze anomalies	Analyze temporal distributions of anomalies	Visualize normal and abnormal behavior frequencies in a calendar-style aggregated display	Aggregate data on a server, then visualize time-based summaries
Detect and analyze anomalies	Examine details of selected trips	Select anomalous trips in key locations (e.g., traffic lanes); visually examine context and interactions	After inspecting selected trips, use queries to extract similar cases and compare contexts

putational and visualization pipelines to handle high-frequency data streams.

Applicability to autonomous vessel development. The findings demonstrate that these techniques provide valuable insights for developing, testing, and operating autonomous maritime systems and managing other types of maritime traffic. The methods support key analytical tasks, such as identifying performance deviations, ensuring operational stability, and understanding the spatio-temporal contexts of vessel movements. By applying these techniques, developers can better evaluate system performance under varying environmental and operational conditions.

Implications for future research and development. Future research should focus on enhancing the integration of visual analytics with real-time decision support systems for autonomous vessels. This includes:

- Improving methods for detecting and visualizing interactions between vessels;
- Extending clustering and classification approaches for more robust handling of large datasets;
- Incorporating additional data sources, such as weather or current conditions, to contextualize anomalous behaviors further;
- Developing open-source tools and libraries to make advanced visual analytics techniques more accessible to researchers and practitioners in the maritime domain.

Limitations and generalizability. While the case studies demonstrated the effectiveness of the methods for analyzing ferry movements, their generalizability to other vessel types or operational contexts requires further validation. For example, differences in vessel behavior, traffic density, and environmental influences may necessitate adaptations of the techniques.

Still, our experience gained in this study allows us to outline several general approaches that can support effective application of visual analytics techniques to large-scale spatio-temporal datasets (see also Table 1):

- **Aggregate before visualizing:** Summarize large volumes of data using aggregation techniques (e.g., histograms, calendar matrices, spatial grids) to reveal overall patterns without overwhelming the analyst.
- **Decompose spatially and temporally:** Divide the dataset into smaller spatial regions or time intervals to make clustering and other computations more manageable and interpretable.
- **Sample for modeling:** Use representative trajectory samples for clustering and defining normal behaviors, assuming that frequent, typical movement patterns will dominate in the data.
- **Preprocess on the server side:** Perform aggregation and filtering operations server-side (or in batch processes) to prepare summaries suitable for interactive, client-side visualization.
- **Preserve interactivity for discovery:** Retain interactive tools for filtering, brushing, and selecting subsets of interest. Human-in-the-loop exploration enables flexible refinement and contextual understanding of patterns and anomalies.
- **Use scalable anomaly detection workflows:** Quantify deviation from modeled behaviors, then aggregate and localize anomalies (e.g., by grid cells or time windows) to prioritize detailed inspection.

These strategies provide a foundation for adapting VA techniques to large, complex datasets in maritime and other movement-related domains. They balance computational efficiency with the interpretive power of human-guided analysis.

Practical relevance. The techniques presented in this paper are intended to support real-world decision-making, particularly, in the preparation for deploying autonomous vessels in routine transportation operations. The ability to analyze historical data at scale is essential for training, evaluating, and fine-tuning autonomous systems based on observed behaviors and anomalies in human-operated operations.

As the techniques presented in this paper are applied in an offline, retrospective context, a question can be raised about their applicability in real-time scenarios. Generally, such scenarios rely heavily on automation, whereas human involvement becomes necessary in exceptional or unforeseen situations. In such cases, analysts require targeted visualizations that not only present relevant details clearly but also help direct attention to anomalies that may require human judgment.

While comprehensive discussion of real-time use cases is beyond the scope of this paper, the visual analytics techniques presented here lay the groundwork for future monitoring systems that combine automated detection with interactive, human-in-the-loop decision support in time-sensitive scenarios.

8. Conclusion and future work

This study demonstrated the scalability and applicability of visual analytics techniques for analyzing maritime vessel movements, using both small experimental and large AIS datasets. The proposed methods effectively addressed key analytical tasks, such as anomaly detection, stability assessment, and spatio-temporal exploration, while enabling insights into both normal and abnormal vessel behaviors.

The findings emphasize the value of combining computational processing and interactive visualization to support the development and operation of autonomous maritime systems. Despite these advancements, challenges such as data quality issues, coarse temporal resolution, and the need for real-time analysis remain critical areas for improvement.

To build on this work, future research should focus on:

- Developing real-time pipelines for efficient analysis combining algorithmic methods with visualization.
- Enhancing techniques for detecting and analyzing vessel interactions under complex traffic conditions.
- Integrating additional data sources, such as weather and ocean currents, for richer contextualization.
- Testing and adapting the methods to diverse operational scenarios and vessel types.

Advancing visual analytics in maritime applications will improve situational awareness and operational safety, ultimately supporting the broader adoption of autonomous vessel technologies.

Data availability

The authors do not have permission to share data.

CRediT authorship contribution statement

Natalia Andrienko: Writing – review & editing, Writing – original draft, Visualization, Software, Data curation, Conceptualization; **Gennady Andrienko:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Data curation, Conceptualization; **Dimitris Zissis:** Writing – review & editing, Writing – original draft, Conceptualization; **Alexandros Troupiotis-Kapeliaris:** Writing – review & editing, Writing – original draft, Data curation, Conceptualization; **Giannis Spiliopoulos:** Writing – original draft, Data curation, Conceptualization.

Declaration of competing interest

The authors have no financial or personal relationships that may be perceived as influencing their work

Acknowledgment

This work was supported by Federal Ministry of Education and Research of Germany and the state of North-Rhine Westphalia as part of the *Lamarr Institute for Machine Learning and Artificial Intelligence* (Lamarr22B), and by EU in project *CrexData* (grant agreement no. 101092749).

References

[1] D. Madeo, A. Pozzebon, C. Mocenni, D. Bertoni, A low-cost unmanned surface vehicle for pervasive water quality monitoring, IEEE Trans. Instrum. Meas. 69 (4) (2020) 1433–1444.

[2] Y. Gu, J.C. Goez, M. Guajardo, S.W. Wallace, Autonomous vessels: state of the art and potential opportunities in logistics, Int. Trans. Oper. Res. 28 (4) (2021) 1706–1739.

[3] X. Bai, B. Li, X. Xu, Y. Xiao, A review of current research and advances in unmanned surface vehicles, J. Mar. Sci. Appl. 21 (2) (2022) 47–58.

[4] N. Andrienko, G. Andrienko, D. Zissis, A. Troupiotis-Kapeliaris, G. Spiliopoulos, Techniques for interactive visual examination of autonomous vessel performance, in: Proceedings of BigVis 2024: 7th International Workshop on Big Data Visual Exploration and Analytics, 2024. https://bigvis.imsi.athenarc.gr/bigvis2024/papers/BigVis2024_01.pdf.

[5] K. Bereta, K. Chatzikokolakis, D. Zissis, Maritime Reporting Systems, Springer International Publishing, Cham, 2021, pp. 3–30. https://doi.org/10.1007/978-3-030-61852-0_1

[6] F. Chen, Y. Yu, L. Ni, Z. Zhang, Q. Lu, DSTVis: toward better interactive visual analysis of Drones' spatio-temporal data, J. Visualization (2024). https://doi.org/10.1007/s12650-024-00982-2

[7] RoBoat, 1st Aegean RoBoat Race, 2023, https://smartmove.aegean.gr/events/1st-aegean-roboat-race/.

[8] D. Zissis, A. Troupiotis-Kapeliaris, E. Xidias, T. Kogias, G. Spiliopoulos, L. Millefiori, P. Braca, The aegean ro-boat race: a regatta to accelerate innovation in marine robotics and autonomous vessels, IEEE Aerosp. Electron. Syst. Mag. 39 (9) (2024) 36–45. https://doi.org/10.1109/MAES.2024.3425317

[9] M. Mokbel, M. Sakr, L. Xiong, A. Züfle, J. Almeida, T. Anderson, W. Aref, G. Andrienko, N. Andrienko, Y. Cao, S. Chawla, R. Cheng, P. Chrysanthis, X. Fei, G. Ghinita, A. Graser, D. Gunopulos, C. Jensen, J.-S. Kim, K.-S. Kim, P. Kröger, J. Krumm, J. Lauer, A. Magdy, M. Nascimento, S. Ravada, M. Renz, D. Sacharidis, C. Shahabi, F. Salim, M. Sarwat, M. Schoemans, B. Speckmann, E. Tanin, Y. Theodoridis, K. Torp, G. Trajcevski, M. van Kreveld, C. Wenk, M. Werner, R. Wong, S. Wu, J. Xu, M. Youssef, D. Zeinalipour, M. Zhang, E. Zimányi, Mobility data science (dagstuhl seminar 22021), Dagstuhl Rep. 12 (1) (2022) 1–34. https://doi.org/10.4230/DagRep.12.1.1

[10] G. Andrienko, N. Andrienko, P. Bak, D. Keim, S. Wrobel, Visual analytics of movement, Springer Science & Business Media, 2013.

[11] G. Andrienko, N. Andrienko, G. Fuchs, Understanding movement data quality, J. Locat. Based Serv. 10 (1) (2016) 31–46. https://doi.org/10.1080/17489725.2016.1169322

[12] A. Graser, An exploratory data analysis protocol for identifying problems in continuous movement data, J. Locat. Based Serv. 15 (2) (2021) 89–117. https://doi.org/10.1080/17489725.2021.1900612

[13] N. Andrienko, G. Andrienko, Visual Analytics of Vessel Movement, Springer International Publishing, Cham, 2021, pp. 149–170. https://doi.org/10.1007/978-3-030-61852-0_5

[14] G. Andrienko, N. Andrienko, W. Chen, R. Maciejewski, Y. Zhao, Visual analytics of mobility and transportation: state of the art and further research directions, IEEE Trans. Intell. Transp. Syst. 18 (8) (2017) 2232–2249. https://doi.org/10.1109/TITS.2017.2683539

[15] O.D. Lampe, H. Hauser, Interactive visualization of streaming data with kernel density estimation, in: IEEE Pacific Visualization Symposium, PacificVis 2011, Hong Kong, China, 1–4 March, 2011, 2011, pp. 171–178. https://doi.org/10.1109/PacificVis.2011.5742387

[16] R. Scheepens, N. Willems, H. van de Wetering, G.L. Andrienko, N.V. Andrienko, J.J. van Wijk, Composite density maps for multivariate trajectories, IEEE Trans. Vis. Comput. Graph. 17 (12) (2011) 2518–2527. https://doi.org/10.1109/TVCG.2011.181

[17] N. Willems, H. van de Wetering, J.J. van Wijk, Visualization of vessel movements, Comput. Graph. Forum 28 (3) (2009) 959–966. https://doi.org/10.1111/j.1467-8659.2009.01440.x

[18] U. Demšar, K. Virrantaus, Space–time density of trajectories: exploring spatio-temporal patterns in movement data, Int. J. Geograph. Inf. Sci. 24 (10) (2010) 1527–1542. https://doi.org/10.1080/13658816.2010.511223

[19] M.-J. Kraak, The space-time cube revisited from a geovisualization perspective, Proceedings of the 21st International Cartographic Conference (2003) 1988–1996.

[20] R. Scheepens, H. van de Wetering, J.J. van Wijk, Non-overlapping aggregated multivariate glyphs for moving objects, in: IEEE Pacific Visualization Symposium, PacificVis 2014, Yokohama, Japan, March 4–7, 2014, 2014, pp. 17–24. https://doi.org/10.1109/PacificVis.2014.13

[21] P. Lundblad, O. Eurenius, T. Heldring, Interactive visualization of weather and ship data, in: Proceedings of the 13th International Conference on Information Visualization IV2009, IEEE Computer Society, 2009, pp. 379–386.

[22] N. Andrienko, G. Andrienko, E. Camossi, C. Claramunt, J.M.C. Garcia, G. Fuchs, M. Hadzagic, A.-L. Jousselme, C. Ray, D. Scarlatti, G. Vouros, Visual exploration of movement and event data with interactive time masks, Visual Inf. 1 (1) (2017) 25–39. https://doi.org/10.1016/j.visinf.2017.01.004

[23] C. Tominski, H. Schumann, G. Andrienko, N. Andrienko, Stacking-based visualization of trajectory attribute data, IEEE Trans. Vis. Comput. Graph. 18 (12) (2012) 2565–2574. https://doi.org/10.1109/TVCG.2012.265

[24] M. Itoh, D. Yokoyama, M. Toyoda, Y. Tomita, S. Kawamura, M. Kitsuregawa, Visual exploration of changes in passenger flows and tweets on mega-city metro network, IEEE Trans. Big Data 2 (1) (2016) 85–99. https://doi.org/10.1109/TBDATA.2016.2546301

[25] C. Palomo, Z. Guo, C.T. Silva, J. Freire, Visually exploring transportation schedules, IEEE Trans. Vis. Comput. Graph. 22 (1) (2016) 170–179. https://doi.org/10.1109/TVCG.2015.2467592

[26] A. Troupiotis-Kapeliaris, N. Gavalakis, K. Koutis, D. Lamparas, G. Melissourgos, G. Nikolaidis, N. Sapountzis, L.-V. Skouras, M. Stampoulou, G. Voudimos, et al., Building an autonomous boat: a multidisciplinary design engineering approach, in: 2023 IEEE 10th International Workshop on Metrology for AeroSpace (MetroAeroSpace), IEEE, 2023, pp. 556–561.

[27] D. Zissis, K. Chatzikokolakis, G. Spiliopoulos, M. Vodas, A distributed spatial method for modeling maritime routes, IEEE Access 8 (2020) 47556–47568. https://doi.org/10.1109/ACCESS.2020.2979612

[28] G. Spiliopoulos, M. Vodas, G. Grigoropoulos, K. Bereta, D. Zissis, Patterns of life: global inventory for maritime mobility patterns, 27, 2024, p. 770 - 777. https://doi.org/10.48786/edbt.2024.67

[29] C. Liu, J. Li, B. Li, Z. Kan, Vessel trajectory compression algorithm based on MapReduce, in: Seventh International Conference on Electromechanical Control Technology and Transportation (ICECTT 2022), 12302, SPIE, 2022, pp. 1022–1027.

[30] X. Wang, X. Liu, B. Liu, E.N. de Souza, S. Matwin, Vessel route anomaly detection with hadoop mapreduce, in: 2014 IEEE International Conference on Big Data (Big Data), IEEE, 2014, pp. 25–30.

[31] T.-A. Pham, X.-K. Dang, N.-S. Vo, Optimising maritime big data by K-means clustering with mapreduce model, in: International Conference on Industrial Networks and Intelligent Systems, Springer, 2022, pp. 136–151.

[32] G. Spiliopoulos, D. Zissis, J. de La Cueva, I. Kontopoulos, Modelling and simulating vessel emissions in real time based on terrestrial AIS data, in: Global Oceans 2020: Singapore - U.S. Gulf Coast, 2020, pp. 1–4. https://doi.org/10.1109/IEEECONF38699.2020.9389295

[33] E. Schubert, J. Sander, M. Ester, H.P. Kriegel, X. Xu, DBSCAN revisited, revisited: why and how you should (still) use DBSCAN, ACM Trans. Database Syst. (TODS) 42 (3) (2017) 1–21. https://doi.org/10.1145/3068335

[34] Z. Chen, J. Guo, Q. Liu, DBSCAN algorithm clustering for massive AIS data based on the hadoop platform, in: 2017 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), IEEE, 2017, pp. 25–28.

[35] B. Lei, A DBSCAN based algorithm for ship spot area detection in AIS trajectory data, in: MATEC Web of Conferences, 291, EDP Sciences, 2019, p. 01008.

[36] N. Newaliya, Y. Singh, Multivariate hierarchical DBSCAN model for enhanced maritime data analytics, Data & Knowl. Eng. 150 (2024) 102282. https://doi.org/10.1016/j.datak.2024.102282

[37] N. Zygouras, G. Spiliopoulos, D. Zissis, Detecting representative trajectories from global AIS datasets, in: 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), 2021, pp. 2278–2285. https://doi.org/10.1109/ITSC48978.2021.9564657

[38] M.-J. Kraak, The space-time cube revisited from a geovisualization perspective, in: Proc. 21st International Cartographic Conference, ICA, 2003, pp. 1988–1996.

[39] G. Andrienko, N. Andrienko, H. Schumann, C. Tominski, Visualization of Trajectory Attributes in Space–Time Cube and Trajectory Wall, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, pp. 157–163. https://doi.org/10.1007/978-3-642-32618-9_11

[40] D. Orellana, G. Andrienko, N. Andrienko, M. Wachowicz, Uncovering interaction patterns in mobile outdoor gaming, in: Advanced Geographic Information Systems & Web Services, International Conference on, IEEE Computer Society, Los Alamitos, CA, USA, 2009, pp. 177–182. https://doi.org/10.1109/GEOWS.2009.13

[41] H. Su, S. Liu, B. Zheng, X. Zhou, K. Zheng, A survey of trajectory distance measures and performance evaluation, VLDB J. 29 (2020) 3–32. https://doi.org/10.1007/s00778-019-00574-9

[42] K. Toohey, M. Duckham, Trajectory similarity measures, Sigspatial Special 7 (1) (2015) 43–50.

[43] M. Ankerst, M.M. Breunig, H.-P. Kriegel, J. Sander, OPTICS: Ordering points to identify the clustering structure, ACM SIGMOD Record 28 (2) (1999) 49–60.

[44] G. Andrienko, N. Andrienko, S. Rinzivillo, M. Nanni, D. Pedreschi, F. Giannotti, Interactive visual clustering of large collections of trajectories, in: 2009 IEEE Symposium on Visual Analytics Science and Technology, 2009, pp. 3–10. https://doi.org/10.1109/VAST.2009.5332584

[45] G. Andrienko, N. Andrienko, G. Fuchs, J.M.C. Garcia, Clustering trajectories by relevant parts for air traffic analysis, IEEE Trans. Vis. Comput. Graph. 24 (1) (2018) 34–44. https://doi.org/10.1109/TVCG.2017.2744322

[46] International Maritime Organization, Convention on the International Regulations for Preventing Collisions at Sea, 1972, 1972, International Maritime Organization (IMO), London, https://www.imo.org/en/About/Conventions/Pages/COLREG.aspx.