IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

A Big Data Approach on Predicting and Optimizing Shipping Management – A Systematic Literature Review and Research

Dr B.Swaminathan
Associate Professor,
School of Maritime Management,
Indian Maritime University, Chennai Campus
bswaminathan@imu.ac.in

Abstract

In the shipping industry, where a lot of data is gathered to better understand and optimize logistics, emissions, energy use, and maintenance, big data is becoming more and more popular. The cost and quality of on-board sensors and data collecting systems, satellite connection, data ownership, and technical challenges to efficient collection and use of big data are some of the constraints on its utilization. The process of gathering and organizing data, notably in the field of e-navigation, may be made simpler by new protocol standards. This essay provides a summary of some of these problems and potential fixes.

Keywords: Industry 4.0, Cyber Shipping, Voyage Data Recorder, Automatic Identification System, Cyber Physical system, Big Data, IoT

Introduction

The fourth technological revolution (Industry 4.0), commonly referred to as "shipping 4.0" or "cyber-shipping," is now underway in the shipping industry (Lasi, 2014). The switch from sail to steam about 1800, the switch from steam to diesel engines around 1910, and the advent of automation and computerized systems around 1970 were all revolutions (Babicz, 2015). According to Hermann et al., the present revolution in maritime operations revolves around the use of digital data in all facets of ship operations (Hermann, 2015). Cyber-Physical Systems (CPS), the Internet of Things (IoT), and the Internet of Services are all heavily utilized in shipping 4.0. With the use of this technology, on-board equipment may be made more intelligent and equipped with integrated computers that can access a wealth of new data and information as well as new shore services to put the data to use (Shafique, 2020). With the use of this technology, on-board equipment may be made more intelligent and equipped with integrated computers that can access a wealth of new data and information as well as new shore services to put the data to use (Zhen, 2020).

That this is indeed "big data" is not immediately apparent. Big data is sometimes described as having large volume, high speed, and/or great diversity, rendering "traditional" data processing approaches insufficient to effectively use the data for analytics, decision-making, or control. (Christos, 2020)

As per Yun et,al one can scarcely assert that shipping nowadays is at this level of processing and computing capability. While handling massive amounts of data from ships may be difficult or impossible with some simple technologies, big data issues are not truly present for today's computer systems. But in order for consumers to grasp the information effectively, the amount and complexity of the data will necessitate new approaches and tools, which may be seen as a type of a big data challenge (Yuen, 2020).

Although the fourth shipping revolution opens up a tone of new opportunities for off-line analytics and more sophisticated online control, there are a number of other issues beyond the sheer bulk of the data sets (Wiegmans, 2020). This analytical study outlines some of the problems that the researcher had come up in his work and giving some solution to that.

2. Growth in data and its importance

There are a number of developments that help make more information on ships and ship systems, such as navigation and automation systems, more readily available. Despite the fact that each of these developments is related to shipping 4.0, they all manifest in very distinct ways and have very different effects on how simple it is to utilize the data (Cui, 2020). A summary of some of the most important data sources currently in use will be provided in this section.

2.1 Bridge data network

The IEC 61162 set of standards, IEC 61162, and digital interfaces are typically used to link the ship bridge equipment (IEC, 2007-2015) . Accessing measurements from the navigation sensors and equipment is made rather simple by this. The amount of data accessible from the bridge is growing gradually but steadily as a result of the flag states' and the IMO's transport requirements. (Guo, 2020)

Additionally, specific equipment pertinent to their operations may be needed for some particular ships. Wave radars, oil spill detectors, very accurate inertial navigation sensors, etc. may fall within this category. Sometimes these data are collected using the Voyage Data Recorder (VDR), but one should be aware of the VDR's restrictions regarding the amount of data points it can gather and how often it can register the data (Liu, 2020).

2.2 Traditional Automation System

The majority of ships include automation systems that gather a lot of data, albeit of varying quality. There might be hundreds of input/output points even on an older bulk carrier. On more advanced and complicated ships, there may be many tens of thousands of data points. (Erős, 2019).

According to Dahl et al.these kinds of data have two key issues: Since the data is typically only accessible within closed, vendor-specific systems, the first step is to get access. The second is the data's level of quality. The researcher also says the sensors may not be very precise since the data is used in closed loop control or in alarm limit surveillance. The presence of defective or unconnected sensors may also be a concern for older automation systems. The user must determine whether the measurements are reasonable and whether values are stuck or change excessively because automation systems seldom offer quality features on the data (Dahl, 2019).

2.3 Cyber physical system

Advanced equipment with built-in sensors and control systems is used aboard modern ships. This comprises power plants and engines, torque-controlled winches, sophisticated dynamic position systems, new navigational sensor systems, and so on (Mohanty, 2020). Computers and sensors are used in the systems for closed-loop control, condition monitoring, and alert situations (Jamwal, 2020). One of the key elements of Industry 4.0: Cyber-Physical Systems is the integration of physical systems with computer control (CPS) (Colombo, 2020).

2.4 Data of Ship Performance

Ship instrumentation for performance monitoring and optimization experienced a significant growth in later years, notably during the high oil price period (Bui, 2020). This frequently featured trim measures, enhanced environmental sensors, fuel mass-flow meters, shaft torque meters, etc (Kim, 2020). This equipment was specifically installed to give data for better ship performance, therefore depending on how and by whom the instruments were installed, one may anticipate that both access to and the quality of the data are high (Dalheim, 2020). In other circumstances, third parties may offer the tools and services, which might limit access to the data.

2.5 Ship Reporting

Several operational and administrative reports are being sent from ships to land, various required port and port state reports, midday at sea reports, technical maintenance reports, etc. are included in this. Many of these represent operational choices made by the crew and can be quite helpful as data analysis input. These reports are frequently created manually, which renders them prone to data entry mistakes. Some reports also affect the ship's financial performance; therefore data may be changed to avoid fines or additional expenses.

2.6 Weather data

Depending on the amount of information needed, forecasted and historical weather data is either free or fee-based. Although the free information is of a low resolution, it can provide useful high-level information regarding the impact of the environment on the ship. In places where meteorological events are impacted by smaller scale geographic characteristics, such as close to islands or inside narrow ocean currents, the data is less valuable along the shore or in such locations (Vettor, 2020).

2.7 Port Call Data

On ship movements in and around ports, port and ship agents are gathering important data. Most of this information is operational and may provide extremely precise information on port delays, loading times, and discharge times, among other things. The information is frequently sent to the ship and ship owner in the form of a statement of facts or a port record. Since the statistics are frequently used to determine port fees, performance claims, and other crucial business information, they tend to be reliable (Conca, 2018).

3. Challenges for collecting ship data on board

An overview of frequently utilized data sources on or off the ship, as well as some of the difficulties experienced while attempting to access and use the data, are provided in the preceding section.

3.1 Contextual data quality

All measurements that are collected from one system and is utilized in another having, the same sense, any issue due to the particular collection context because of sensor data (Gao, 2018). One illustration is several sensors of comparable sorts, their locations, and their connections: For instance, a ship often includes a number of position sensors, each of which has a unique local location reference and unique quality characteristics that might not be known to systems other than the bridge. Using the navigation system's raw position data may result in accuracy or consistency issues.

According to Li et al.; sensors employed in closed loop control or as alarm limit sensors experience a similar issue. Basic sensor characteristics like linearity, offsets, accuracy, and stability will constantly be modified for the work at hand, and the fitting of a particular sensor involves making trade-offs between cost, robustness, and suitability for the task (Li, 2016). The exact quality of the measurement might not be as important for control and alarm applications since sensor artifacts can be reduced in the control and monitoring software. The usage of these measures outside of their systems quickly becomes troublesome unless the adjustment variables are known to other systems (RØDSETH, 2016)

3.2 Security and Integrity

Connections to the bridge or other data networks that perform duties linked to safety or security may pose a risk to those factors (Zaman). Any physical entry point into networks or systems can serve as a vehicle for malicious assaults or the spread of errors. Thus, flag state or class authorities frequently forbid such links. Bridge networks have benefited from advancements in firewall technology and safer Ethernet networks for connecting to external networks (Vujičić, 2020).

3.3 Entry Errors in Non-Automated data entry

The manual entering of data into report forms, computer systems, or AIS (Automatic Identification System) transceivers has proven to be a significant cause of inaccuracies in the past (Mitchell, 2014). This will show up specifically in journey destination, ship draught, and sailing mode for AIS, but it can also occasionally appear as unlawful or incorrect ship identification.

Concerns with the transfer of navigational data, such as rate of turn and true heading, which should be obtained from sensors outside of the AIS, also exist. Many AIS transmitters give invalid data or data that is internally generated from the position since they are not attached to such sensors (Alizadeh, 2021).

3.4 Low quality instrumentation – less quality attributes

Sometimes, the quality of the sensor data on older ships and sensor data that is not frequently used operationally is questionable. This may be the result of a sensor failure, disconnection, discomfort, or any number of other issues. This issue with data from the automation and alarm systems is very typical (Ribeiro, 2018).

Most automation systems won't do quality checks on the data or provide the measurements any additional quality criteria. If data is exported from the system, it will be challenging to determine whether or not the data can be trusted (Kalra, 2020).

3.5 Ownership of special or derived data

Specific limitations on the usage of system-internal data may be included in the delivery of some systems (Mizumoto, 2020). The ship owner may occasionally get condition-based maintenance and general monitoring as a service even if the owner has no ownership or access to the underlying data. Similar outside services are occasionally employed for a variety of ship or fleet monitoring or optimization tasks. This could also indicate limitations on data ownership.

3.6 Challenges in Cyber security

Although cyber security is not thought to be a significant problem in the gathering of general data on-board, one should be aware that some data may be blocked or faked by cyber-attacks (Heering, 2020). Even GPS position signals may be fooled, according to examples. Every piece of information that depends on wireless broadcasts from other ships or the shoreline is vulnerable to spoofing or jamming. This applies, for instance, to AIS or radar targets that the ship gathers and transmits to land. Additionally, it will apply to AIS signal data gathered by land-based or satellite-based systems (Tusher, 2020).

4. Application Maritime Big data and its applications

As the preceding sections have demonstrated, there are several data sources, potentially large amounts of data, and therefore numerous opportunities for inaccuracy aboard the ship. Data management must be carried out along three axes both on board and ashore, as demonstrated in Fig.1.

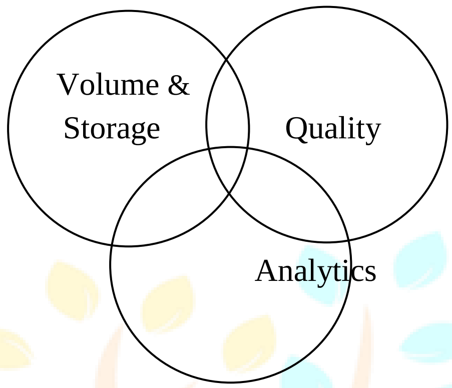

Fig1: Three dimensions of big data management

Large amounts of data are managed in an organized collection by volume and storage management on board and/or ashore (Lv, 2020). The expense and feasibility of transporting substantial amounts of data between ship and shore will also need to be taken into account when managing volume (Schwatke, 2020). Data analytics is the branch of science that deals with analyzing data to derive meanings (Mohammadpoor, 2020). For the accuracy of the analysis results, quality control of the data must be taken into account by either party or as a separate operation.

4.1 Volume and Storage

Ship performance and operational data are gathered in vast amounts by sensors and data gathering systems (Zhang, 2021). Some systems, especially those that monitor bridges and performance, can be fitted with a number of sensors that often gather characteristics of the same kind (Rawson, 2021). Similar circumstances apply to AIS (Automatic Identification System) observations, when many base stations might simultaneously receive data from a single ship. The time resolution of data collected by many sensors, including AIS, is frequently substantially better than that needed for ship performance analysis (Ignatov, 2021).

In many cases, measured data might be redundant in terms of functionality or time, and such instances should be found throughout the volume and storage management process. By eliminating duplication, most data sets can often be shrunk down to a considerably smaller data set. This greatly enhances the processing and exchange of data (L.P. & MO, 2016).

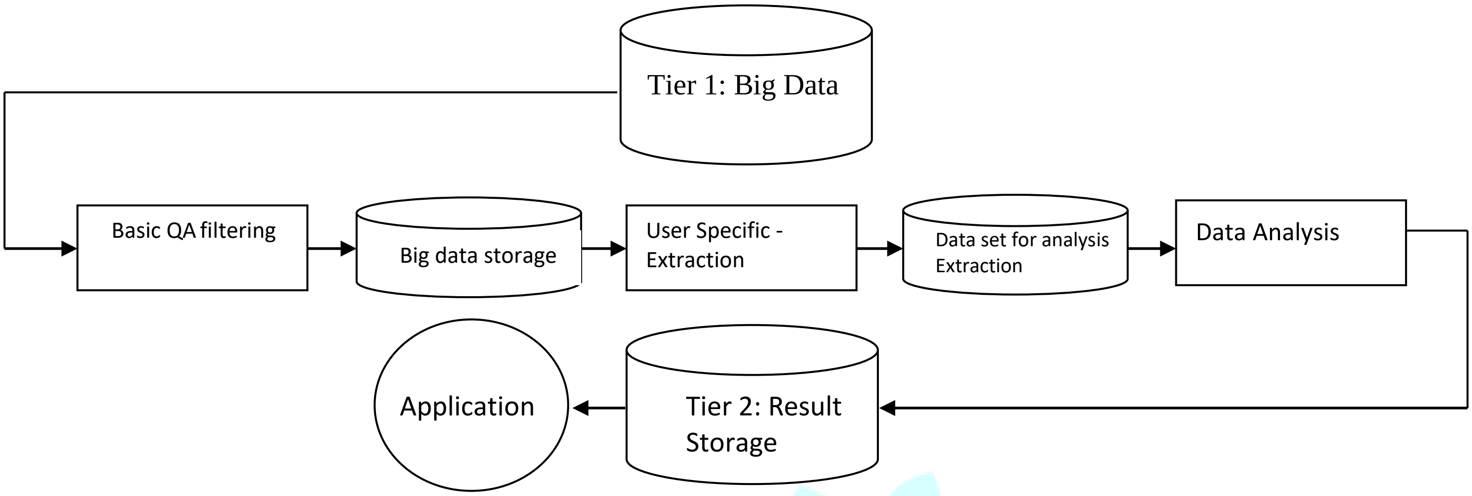

Fig 2: Proposed Data Analysis System for Ship Analysis

Data storage and organization are important factors to take into account while managing data, especially when handling huge data volumes effectively. We are looking into a two-tier storage system, as shown in Fig.2, to address the tension between having consistent and simple to manage data sets and preserving details in the data sets that could be useful later. After early poor data removal and thorough data reduction, the lowest tier will leverage big data technologies to keep as much data as possible. The second tier will apply certain extracts, structures, and methods to the current analytical issue.

5. Ship big data application in operational perspective

In many various time scales, from online decision support tools to long-term fleet statistics analysis, the new big data applications may be deployed (Lee, 2022). Within this paradigm, loops may be utilized to update control parameters in online decision support using the findings of long-term study (Gil, 2022).

5.1. Ship decision support by OLAP

Early in the 1990s, concepts like data warehouses and online analytical processing (OLAP) were proposed; by the mid-1990s, they had become popular. While OLAP concentrates on data analysis in data warehouse and turns it into decision-making information, data warehouse focuses on storing and maintaining subject-oriented data (Zhu, 2020).

5.2. Optimization of Ship Performance

Other applications will analyze data over longer periods to determine more general optimization possibilities (Farag, 2020). In addition to best practice operational guidance, this can also include technical condition monitoring and maintenance planning (Pennino, 2020). One example of such a service is technical condition monitoring which uses monthly controlled measurements to assess the technical condition of ship machinery (Chen, 2020).

5.3. Fleet Optimization

The performance of similar ships over extended periods of time may be compared using even bigger data sets (Cannarile, 2015). This is often used to identify statistical operational restrictions for specific routes or to provide more general recommendations for best practices throughout the fleet. Various technical and operational methods, such as maintenance plans, voyage optimization, hull coating performance, etc., can also be examined for their technical performance (Choi, 2022).

5.4 Predictive analysis

Using historical data to identify new ship features or operating principles is another intriguing possibility that big data brings us (Nguyen, 279-299). One intriguing use of this is the virtual development and simulation of new ship designs under known and historically validated operating situations (Zhong, 2022).

Conclusion

Big data processing for ships and shipping has significant promise for creating more effective technological solutions and operating regimes for these industries. It can access a variety of previously underutilized data sources. However, there are several mistake sources that might impede the analytical procedures and result in false findings. To deal with such circumstances, a variety of online and offline tools and strategies must be created. This is crucial for systems that run online and might be included as a component of contemporary integrated bridge systems. Systems employing data collected over longer time periods have a greater chance of filtering out bogus poor data points by using bigger historical data sets and averaging over longer time periods. However, there are new techniques that are developing that might offer improved sensor and acquisition fault situations detection even for online and real-time data handling. The quality of the data sets may be raised, allowing for more accurate and detailed observations of immediate vessel performance.

To preserve, search, and retrieve the necessary data sets, which allow for a detailed analysis of the performance of the whole fleet of vessels while using offline data handling, it may be necessary to use extra data management tools. As a result, the growth of data management and analytics will be crucial in the area of big data in transportation.

For certain ships, connectivity to land is still an issue, but as communication technology advances and bandwidth costs fall, this issue is becoming less of a concern. There are a lot of intriguing big data applications for ship and fleet analysis, notwithstanding connection difficulties.

The absence of technological standards for linking to various data sources is a key cost-increasing problem for big data applications on ships. This necessitates the creation of expensive, ad hoc interface solutions for each ship. Additionally, it will be expensive to set up databases, develop system models, and connect the various measurements to the pertinent events. Although effort is being done in this area, owners, yards, system makers, and scientists must provide their assistance more.

Understanding the potential and constraints of the various data sources and the technologies available to address the various issues is another essential component for the effective implementation of big data applications in the marine industry.

Reference

Alizadeh, D. A. (2021). Vessel trajectory prediction using historical automatic identification system data. *The Journal of Navigation* , 156-174.

Babicz, J. (2015). *Encyclopedia of ship technology*. Wärtsilä Corporation.

Bui, K. Q. (2020). Advanced data analytics for ship performance monitoring under localized operational conditions. *Ocean Engineering* , 235.

Cannarile, F. C. (2015). Handling reliability big data: A similarity-based approach for clustering a large fleet of assets. In Safety and Reliability of Complex Engineered Systems-Proceedings of the 25th European Safety and Reliability Conference. *ESREL* .

Chen, H. Z. (2020). Optimization of sizing and frequency control in battery/supercapacitor hybrid energy storage system for fuel cell ship. *Energy* .

Choi, Y. N.-T. (2022). Optimization of haulage-truck system performance for ore production in open-pit mines using big data and machine learning-based methods. *Resources Policy* .

Christos, S. C. (2020). Combined multi-layered big data and responsible AI techniques for enhanced decision support in Shipping. *International Conference on Decision Aid Sciences and Application (DASA)* (pp. 669-673). IEEE.

Colombo, A. W. (2020). Learning industrial cyber-physical systems and industry 4.0-compliant solutions. *IEEE Conference on Industrial Cyberphysical Systems (ICPS)* (pp. 384-390). IEEE.

Conca, A. D. (2018). Automation in freight port call process: Real time data sharing to improve the stowage planning. *Transportation research procedia* , 70-79.

Cui, Y. K. (2020). Manufacturing big data ecosystem: A systematic literature review. *Robotics and computer-integrated Manufacturing* , 62.

Dahl, M. E. (2019). Control components for collaborative and intelligent automation systems. *IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)* , 378-384.

Dalheim, Ø. Ø. (2020). A computationally efficient method for identification of steady state in time series data from ship monitoring. . *Journal of Ocean Engineering and Science* , 333-345.

Erős, E. D. (2019). A ROS2 based communication architecture for control in collaborative and intelligent automation systems. *Procedia Manufacturing* , 349-357.

Farag, Y. B. (2020). The development of a ship performance model in varying operating conditions based on ANN and regression techniques. . *Ocean Engineering* .

Gao, M. S. (2018). Online prediction of ship behavior with automatic identification system sensor data using bidirectional long short-term memory recurrent neural network. . *Sensors* , 4211.

Gil, M. K. (2022). Know your safety indicator–A determination of merchant vessels Bow Crossing Range based on big data analytics. *Reliability Engineering & System Safety* .

Guo, J. &. (2020). Dynamic response analysis of ship-bridge collisions experiment. *Journal of Zhejiang University-SCIENCE A* .

Heering, D. E. (2020). suring cybersecurity in shipping: Reference to Estonian shipowners. . *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation* .

Hermann, P. T. (2015). *Design principles for Industrie 4.0 scenarios: a literature review*. TU Dortmund.

IEC. (2007-2015). *Maritime navigation and radiocommunication equipment and systems- Digital interfaces,*. IEC 61162.

Ignatov, A. R. (2021). Real-time video super-resolution on smartphones with deep learning, mobile ai 2021 challenge. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* , 2535-2544.

Jamwal, A. A. (2020). Development of cyber physical system based manufacturing system design for process optimization. *Materials Science and Engineering* (p. 012048). IOP Publishing.

Kalra, A. C. (2020). Machine learning for automation of radiology protocols for quality and efficiency improvement. *Journal of the American College of Radiology* , 1149-1158.

Kim, H. K. (2020). Artificial intelligence vision-based monitoring system for ship berthing. *Ieee Access* .

L.P., P., & MO, B. (2016). Data compression of ship performance and navigation information under deep learning. *35th Int. Conf. Ocean, Offshore and Arctic Engineering (OMAE 2016),*. Busan.

Lasi, H. F. (2014). *Industry 4.0* (Vol. 6(4)). Business & information systems engineering.

Lee, J. S. (2022). Extracting the maritime traffic route in Korea based on probabilistic approach using automatic identification system big data. *Applied Sciences* .

Li, G. Z. (2016). Analysis and modeling of sensor data for ship motion prediction. *In OCEANS 2016-Shanghai* , IEEE.

Liu, S. L. (2020). Rational processing of monitored ship voyage data for improved operation. *Applied Ocean Research* , 104.

Lv, S. (2020). Construction of marine ship automatic identification system data mining platform based on big data. *Journal of Intelligent & Fuzzy Systems* .

Mitchell, K. N. (2014). Waterway performance monitoring with automatic identification system data. *Transportation Research Record* , 20-26.

Mizumoto, K. K. (2020). Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan. *Eurosurveillance* .

Mohammadpoor, M. &. (2020). Big Data analytics in oil and gas industry: An emerging trend. *Petroleum* .

Mohanty, S. P. (2020). Healthcare cyber-physical system is more important than before. *IEEE Consumer Electronics Magazine* , 6-7.

Nguyen, S. C. (279-299). Container shipping operational risks: an overview of assessment and analysis. *Maritime Policy & Management* .

Pennino, S. G. (2020). Development of a new ship adaptive weather routing model based on seakeeping analysis and optimization. *Journal of Marine Science and Engineering,* .

Rawson, A. &. (2021). Developing contextually aware ship domains using machine learning. *The Journal of Navigation* .

Ribeiro, D. M.-P. (2018). Technical quality of root canal treatment performed by undergraduate students using hand instrumentation: a meta-analysis. *International endodontic journal* , 269-283.

RØDSETH, Ø. (2016). Integrating IEC and ISO information models into the S-100 Common Maritime Data Structure. *E-navigation Underway International* .

Schwatke, C. D. (2020). Volume variations of small inland water bodies from a combination of satellite altimetry and optical imagery. . *Remote Sensing* .

Shafique, K. K. (2020). Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. (pp. 23022-23040). Ieee Access.

Tusher, H. M. (2020). Cyber security risk assessment in autonomous shipping. . *Maritime Economics & Logistics,* , 1-20.

Vettor, R. S. (2020). Towards improving optimised ship weather routing. *Polish Maritime Research.*

Vujičić, S. H. (2020). Distributed ledger technology as a tool for environmental sustainability in the shipping industry. *Journal of Marine Science and Engineering* , 366.

Wiegmans, B. W. (2020). Big data of the past: Analysis of historical freight shipping corridor data in the period . *Research in Transportation Business & Management* , 1662–1855.

Yuen, K. F. (2020). Special issue on 'Artificial Intelligence & big data in shipping. *Maritime Policy & Management* , 575-576.

Zaman, I. P. (n.d.). Challenges and opportunities of big data analytics for upcoming regulations and future transformation of the shipping industry. *Procedia engineering* , 537-544.

Zhang, D. Z. (2021). Data mining approach for automatic ship-route design for coastal seas using AIS trajectory clustering analysis. *Ocean Engineering* .

Zhen, L. W. (2020). Green technology adoption for fleet deployment in a shipping network. *Transportation Research Part B: Methodological* , 388-410.

Zhong, Y. Z. (2022). I give discounts, I share information, I interact with viewers: a predictive analysis on factors enhancing college students' purchase intention in a live-streaming shopping environment. . *Young Consumers.*

Zhu, A. &. (2020). Research on decision support system of ship planning management based on data warehouse. *Journal of Physics* .