**Ocean Systems Engineering**, *Vol. 9, No. 1 (2019) 79-95*
DOI: https:// doi.org/10.12989/ose.2019.9.1.079 

# Big data-based piping material analysis framework in offshore structure for contract design

Min-Jae Oh[1], Myung-Il Roh*[2], Sung-Woo Park[3],
Do-Hyun Chun[4] and Sehyun Myung[5]

*[1]Research Institute of Marine Systems Engineering, Seoul National University, Seoul, Republic of Korea*
*[2]Department of Naval Architecture and Ocean Engineering, and Research Institute of Marine Systems Engineering, Seoul National University, Seoul, Republic of Korea*
*[3]Interdisciplinary Program in Offshore Plant Engineering, Seoul National University, Seoul, Republic of Korea*
*[4]Department of Naval Architecture and Ocean Engineering, Seoul National University, Republic of Korea*
*[5]School of Automotive & Mechanical Design Engineering, Youngsan University, Yangsan, Republic of Korea*

*(Received December 10, 2018, Revised January 20, 2019, Accepted January 24, 2019)*

**Abstract.** The material analysis of an offshore structure is generally conducted in the contract design phase for the price quotation of a new offshore project. This analysis is conducted manually by an engineer, which is time-consuming and can lead to inaccurate results, because the data size from previous projects is too large, and there are so many materials to consider. In this study, the piping materials in an offshore structure are analyzed for contract design using a big data framework. The big data technologies used include HDFS (Hadoop Distributed File System) for data saving, Hive and HBase for the database to handle the saved data, Spark and Kylin for data processing, and Zeppelin for user interface and visualization. The analyzed results show that the proposed big data framework can reduce the efforts put toward contract design in the estimation of the piping material cost.

**Keywords:** piping material; big data analysis; offshore structure; contract design

## 1. Introduction

The material analysis of offshore structure is generally conducted in the contract design phase for the price quotation of a new offshore project. For contract design, engineers analyze previous projects and use this analysis to estimate the construction costs for new projects. This work is very important because the initial cost determines the success of the project. The analysis is conducted manually by the engineer, which is time-consuming and can lead to inaccurate estimation. The engineers collect the related data from the previous projects and use these data to estimate the cost. Unfortunately, there is often not enough time to precisely estimate the cost of a new project. It is very difficult to estimate precise costs in a short amount of time because there is lots of information from the previous projects that must be analyzed. In this study, a big data framework

*Corresponding author, Professor, E-mail: miroh@snu.ac.kr


http://www.techno-press.org/?journal=ose&subpage=7 ISSN: 2093-6702 (Print), 2093-677X (Online)

for material analysis is suggested for faster and easier analysis, and the piping materials are analyzed in order to verify the applicability of the framework.

The most important technologies that have become the driving force of the fourth Industrial Revolution in recent years are artificial intelligence, the Internet of Things, and big data. In the manufacturing industry in particular, there has been an active movement toward deriving meaningful results from the accumulated data. Therefore, the importance of big data technology is increasing. Recently, such a move has also been seen in the shipbuilding industry, as shipbuilding companies are operating big data teams and analyzing the accumulated big data. In doing so, they are trying to increase the efficiency of their business process and reduce their production costs by analyzing big data. They are also going to provide customized services by analyzing the needs of customers using big data analysis. Ship owners and operators are also interested in big data analytics. Huge amounts of data are gathered during the operation, and it is possible to analyze this data due to the development of the big data technologies. Ship owners and operators are attempting to reduce the Capital Expenditure (CAPEX) and Operational Expenditure (OPEX) from the big data analytics. Class societies have also started new intelligence services from big data analytics. For example, DNV GL started a new business from Automatic Identification System (AIS) data analysis (DNV GL, 2018). The suggested examples using AIS data analysis are emission monitoring, delay management in container shipping, and voyage management. There is plenty of interest in the maritime industry regarding big data analytics, but it is not easy to find suitable big data technologies for maritime industry because there are so many technologies that are often changing. In addition, new technologies are being developed very quickly. Therefore, in this study, we introduce some of the recent breakthroughs in big data technology and propose a big data framework to support designers in contract design.

The main contribution of this study is the proposal of the use of big data-based piping material analysis for contract design. It is possible to perform interactive analysis using query language in the big data framework, so that the amount of manual work necessary can be minimized. In addition, the big data framework can handle large amounts of data very quickly, so it does not take a long time to analyze large amounts of data from previous projects.

The organization of this paper is as follows: Section 2 introduces related works in big data analytics in the maritime industry. In Section 3, the big data framework to analyze the big maritime data is proposed. In order to show the applicability of the big data framework, the piping material analysis is conducted in Section 4. Section 5 concludes the study and gives directions for future work.

## 2. Related works

In general, big data refers to data with the characteristics of the 3Vs of volume, velocity, and variety, as suggested by Gartner (Pettey and Goasduff 2011), a global research and advisory firm. Gandomi and Haider (2015) defined the concept of big data and introduced the methods of analysis for unstructured big data. Chen *et al*. (2014) reviewed the background and state-of-the-art in 2014. The authors explained the definitions of big data, technologies, Internet of Things, and analysis of data. They defined big data as the datasets that could not be perceived, acquired, managed, and processed by traditional IT and software/hardware tools within tolerable time. Mauro *et al.* (2015) reviewed previous works for big data regarding information, technology, methods, and impact. For the general overview of the data analytics, Kelleher and Tierney (2018)

Table 1 Related works on big data analysis in the maritime industry

| Research | Purpose | Application | Big data for analysis | Big data technology |
|---|---|---|---|---|
| Kim *et al.* (2013) | Estimated of external force acting on the ship | Ship operation | AIS, External force | N/A |
| Feblowitz (2013) | Proposed the potential of big data analytics | Oil and gas industry | Data from the oil and gas project | Hadoop |
| Wang *et al.* (2015) | Proposed big data analytic-Internet of Things framework | Maritime industry | Data from sensors | N/A |
| Lee (2017) | Proposed the big data analysis guideline | Shipbuilding industry | Data from the construction | N/A |
| Perera and Mo (2017) | Proposed the big data framework for ship energy efficiency | Ship operation | AIS | N/A |
| Lytra *et al.* (2017) | Proposed the big data architecture | Maritime industry | Ocean data, Maritime data | SPARQL |
| Kim *et al.* (2017) | Estimated the weight of FPSO topside | Offshore structure design | FPSO weight data | Hadoop, RHadoop |
| Oh *et al.* (2018) | Estimated the material requirement of piping materials | Offshore structure design | Piping material, schedule | Hadoop, Spark |
| Yoo (2018) | Identified the near-miss areas and density | Ship operation | AIS | N/A |
| Park *et al.* (2018) | Analyzed the association of piping materials | Offshore structure design | Piping material | Hadoop, Spark |
| This study | Analyzed the usage of piping materials for contract design | Offshore structure design | Piping material | Hadoop, Spark, Hive, Kylin, Zeppelin |

presented the overview of the data science and machine learning, and O'Neil and Schutt (2013) introduced the definition of the data science and the machin learning algorithm for the big data processing.

Mourouzis (2018) suggested how big data analytics will transform the shipping industry. Application areas were suggested as fuel consumption, route and supply-chain optimization, operational efficiency, maintenance prediction, and so on. Kim *et al.* (2018) addressed the necessity of a big data platform in offshore applications. In spite of the increasing interest in big data in the maritime industry, there have not been many studies that applied big data technology to maritime applications so far. The previous studies on big data analytics in the maritime industry are summarized in Table 1. Kim *et al.* (2013) proposed the assessment method of an external force acting on the ship in maritime traffic using big data analytics. The Automatic Identification System (AIS) data and the external force data from sensors were defined as big data. Feblowitz (2013) proposed the potential of big data in the oil and gas industry. The author introduced the definition of big data and pointed out that the oil and gas companies were not familiar with big data and analytics in 2013. The author also introduced big data technologies such as Hadoop. Wang *et al*. (2015) suggested the analysis of big data with an Internet of Things (IoT) framework in the maritime industry. The authors pointed out that the new big data analytics and IoT technologies are emerging quickly, but most of them are for land-based applications. They found that cloud-based big data analytics for Offshore Support Vessels is not a suitable method. Therefore, the vessel-based big data analytic layer and the land-based big data analytic layer were proposed. Lee (2017) provided the big data analysis guideline for the shipbuilding industry. The author analyzed the shipbuilding process and suggested the reference model, which is composed of four layers: the construction phase, the category of analysis, analysis method, and detailed algorithm. In addition, the big data applications were introduced per construction phase. Perera and Mo (2017) developed the onboard big data handling framework for ship energy efficiency. The authors proposed the data pre-processing section on board for sensor fault detection, data classification, and so on, and the data post-processing section on shore for more time-consuming data processing. Lytra *et al.* (2017) proposed the big data architecture for ocean data and maritime applications. The authors defined what big marine data is, and suggested four applications: fault prediction, mare protection, anomaly detection, and wave power identification. Yoo (2018) presented a method for identifying ship near-miss areas and relative density maps using AIS data. Kim *et al.* (2017) proposed a big data platform based on Hadoop and estimated the weight of offshore structure topside using RHadoop package. Oh *et al.* (2018) used the big data technology to estimate the material requirement of piping materials in an offshore structure. In this study, based on Kim *et al.* (2017) and Oh *et al.* (2018)'s studies, we introduce new big data technologies that can be applied to analyzing and visualizing big data, and propose a new big data framework using each technology. We analyze the piping materials of offshore structures using the framework for contract design.

## 3. Big data framework for the analysis

The big data framework that is newly suggested in this study by improving the previous work by the authors (Kim *et al.* 2017) is shown in Fig. 1. The big data framework consists of data storage, database management system, data mining, and processing system, as well as a distributed analysis processing engine and visualization notebook for the user interface. It consists of HDFS, Hive, Hbase, Spark, Kylin, and Zeppelin.

The big data framework is based on the Hadoop framework (Apache Hadoop 2018), which is the big data environment used. Hadoop is a Java-based open source framework for distributing and processing large amounts of data. It handles large amounts of data using MapReduce. MapReduce is a software framework for processing large amounts of data in distributed parallel computing. Hadoop includes the Hadoop Distributed File System (HDFS), which is a file distribution repository big enough to distribute large amounts of data across multiple servers. Since HDFS stores data in blocks of a specific size and is stored in a distributed server, it has merit in that it is a storage system that can be configured using lower-cost hardware than that required by a conventional large-capacity file system. The data in HDFS can be transformed as a relational database, so that users can access and process the data. A relational database is a dataset in which data is organized into columns and rows, and data is identified through a primary key. This relational database is stored in Hive (Apache Hive 2018), which acts as a big data DBMS (Database Management System), which summarizes, queries, and analyzes data. Because Hive supports built-in HiveQL queries or MapReduce functionalities, users can send queries to analyze data. Using the queries, it is convenient to extract, transform, and load the data for analysis. The data stored in Hive is processed using the distributed processing framework Spark. Previously (Kim *et al.* 2017), Hadoop MapReduce was used for data distribution processing, but it has a disadvantage in that it is difficult to process data in real time, and it is also difficult to process complex data types because it cannot use query language like SQL. Spark not only improves real-time data processing speed in memory, but also supports various query languages such as Spark SQL, so that query execution as well as immediate feedback can be obtained through an interactive shell. Hive stores HDFS data in a relational database through several attribute items. Each attribute item can be created as a dimension, and each attribute item can be collected and converted into the form of multidimensional data. In order to carry out this collecting and converting process, the data analysis engine, Kylin (Apache Kylin 2018), is used to analyze the multidimensional data in a variety of ways using various criteria. This interactive analytic is called On-Line Analytical Processing (OLAP). Kylin collects only the parts it wants to analyze from the data stored in a relational database format on Hive and defines it as a multidimensional database (or data cube). As a result, it allows for fast and interactive analysis. The generated multidimensional database is stored in a database called HBase (Apache HBase 2018), and can be quickly analyzed by queries using the built-in SQL language of Kylin. Queries for multidimensional data analysis can be made using Kylin's native windows environment or Zeppelin (2018), a web-based notebook-style environment. All of the data processing tasks described so far are performed using a notebook capable of multi-purpose data processing, retrieval, analysis, and visualization called Zeppelin. Zeppelin can send a query and analyze data by accessing various big data analysis tools and databases such as Spark, Kylin, and Hive according to needs and usage. Zeppelin also supports data visualization and concurrent editing, so it is possible to visualize the analysis results on the Zeppelin window. The big data framework is built on Linux, and the states of all of the components are monitored through Ambari (Apache Ambari 2018).

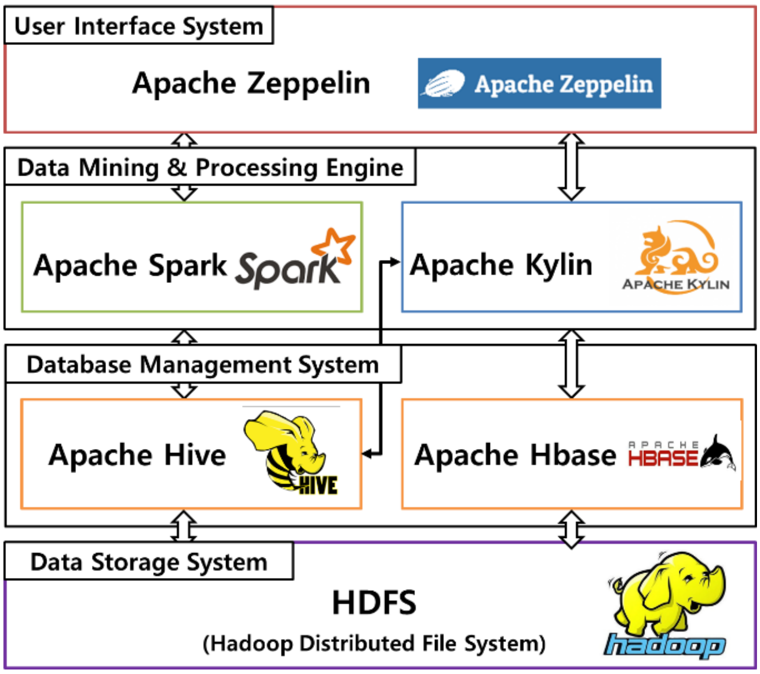


Fig. 1 Big data framework for the analysis of piping materials

The proposed big data framework can be changed according to the needs of the applications. For example, if the data is not too large, or the real-time processing is not necessary, then Kylin, which can create a data cube for fast and interactive analysis, is not necessary to be included in the big data framework. In addition, new big data technologies are being developed now; those can be added to the proposed big data framework. The proposed big data framework can be built in cloud services such as Amazon Web Service (AWS) (Amazon Web Services Inc., 2018) and Google Cloud (Google 2018), or in a local server. In this study, we developed the big data framework on a local server. Using the framework, we analyze the usage of piping materials in an offshore structure for contract design. The details will follow in the next section.

## 4. Analysis of the piping material usage for contract design

### *4.1 Piping materials of the offshore structure*

In this study, the entire piping material list for an offshore structure is used for analysis. There are many kinds of piping materials such as pipe, elbow, tee, flange, gasket, clamp shoe, and so on. Additionally, each material has various kinds of properties depending on the materials such as GRE (Glass Fiber Reinforced Epoxy), stainless steel, carbon steel, titanium, and copper-nickel. The piping material data is stored in the shipbuilding company's internal database system of a 3D CAD (Computer Aided Design) system, ERP (Enterprise Resource Planning) system, or PLM (Product Lifecycle Management) system, depending on the applications. The stored piping material data can be analyzed using a relational database management system. However, since the types and materials of piping materials are very diverse, it is difficult to calculate accurate and fast statistics, and it is also difficult to perform comparative analysis on several ships or offshore structures. This is the reason why the exact estimation of the price quotation for a bidding is difficult and time-consuming in the contract design phase. In this study, we analyze the usage of the piping materials for an offshore structure and verify the applicability of the proposed big data framework. The analyzed offshore structure consists of 110 modules, of which 90 (module number 1 to 90) are comprised of process and utility modules topside, and the remaining 20 (module number 91 to 110) are composed of living quarters.

### *4.2 Analysis results*

Analysis of the piping materials is performed according to the properties and quantities of the materials. The most used piping material is a pipe, so we performed the analysis for the pipe and the other materials separately. Each piping material can have one of the material properties, and it is necessary to check how many piping materials are used with the properties as the cost can differ according to the material properties.

#### *4.2.1 Material analysis*

The materials that are used in the offshore structure are GRE, stainless steel, carbon steel, duplex, and so on. Each piping material can have one of the material properties, and it is necessary to check how many piping materials are used according to the properties because the cost can be different from the material properties. For example, the duplex pipe is much more expensive than the carbon steel pipe. Therefore, if the length of the pipe with duplex material used in an offshore structure can be reduced, then the construction cost will be down. This is why the piping material analysis is necessary to find the information from big data.

Fig. 2 shows the rate of the pipe length per material used for the offshore structure. Table 2 shows the length value of the entire pipe material in Fig. 2. The offshore structure used in the analysis is the most used with 33.51% of GRE material for the pipe. GRE is relatively inexpensive compared to carbon steel or stainless steel, but because of the low pressure it can tolerate, it is often used for constructing an open drain that does not have to withstand high pressures.

Fig. 3 shows the ratios of the materials used for the offshore structures, excluding pipe, and Table 3 shows the details. The piping materials except for pipe are elbow, flange, gasket, coupling, U-bolt, bracing, valve, connection, and so on. According to the results shown in Fig. 2 and Table 2, the GRE material is the most commonly used (32%), followed by the stainless steel material. The same material must be used for both pipe and its connectors such as elbows and flanges. Dissimilar metals can induce electrochemical potential and hence could trigger corrosion. Depending on the topside processing design or the other piping design of an offshore structure, if a specific material is used a lot, the other materials of the same material are also increased accordingly.

Table 2 Used materials of the pipe

| Material | Quantity (m) | Percentage (%) |
|---|---|---|
| GRE | 232,729 | 33.51 |
| Stainless Steel | 195,502 | 28.15 |
| Carbon Steel | 174,054 | 25.06 |
| Duplex | 45,253 | 6.52 |
| Copper | 37,483 | 5.40 |
| Cu-Ni | 9,306 | 1.34 |
| Titanium | 263 | 0.04 |
| Sum | 694,590 | 100.00 |

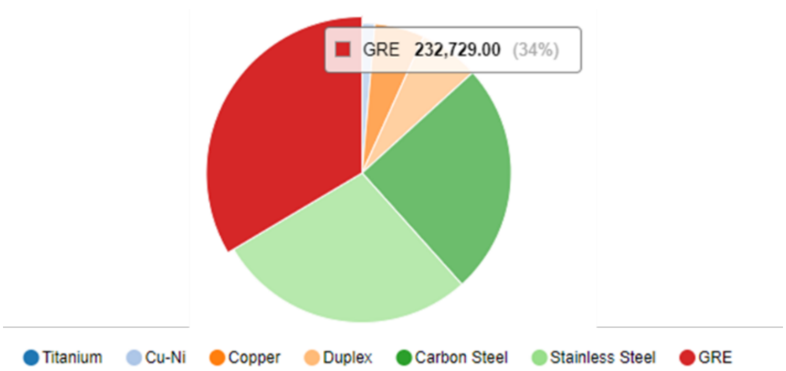


Fig. 2 Rate of used materials of the pipe

Table 3 Used materials except pipe

| Material | Quantity (EA) | Percentage (%) |
|---|---|---|
| GRE | 47,291 | 32.01 |
| Stainless Steel | 42,836 | 29.00 |
| Carbon Steel | 37,292 | 25.24 |
| Duplex | 11,998 | 8.12 |
| Copper | 7,028 | 4.76 |
| Cu-Ni | 1,127 | 0.76 |
| Titanium | 149 | 0.10 |
| Sum | 147,721 | 100.00 |

Table 4 Used materials of pipe topside

| Material | Quantity (m) | Percentage (%) |
|---|---|---|
| GRE | 217,222 | 37.71 |
| Carbon Steel | 160,713 | 27.90 |
| Stainless Steel | 142,657 | 24.77 |
| Duplex | 45,253 | 7.86 |
| Copper | 9,780 | 1.70 |
| Titanium | 263 | 0.05 |
| Cu-Ni | 101 | 0.02 |
| Sum | 575,989 | 100.00 |

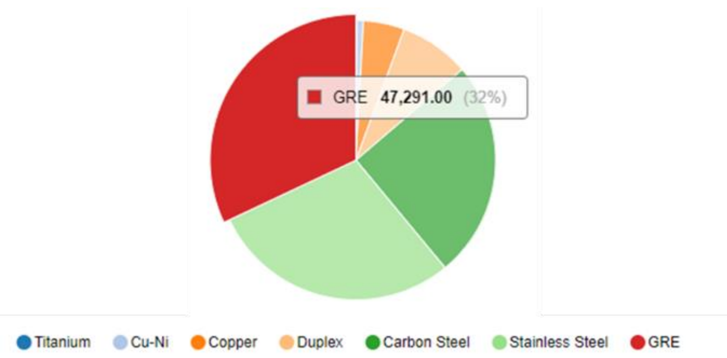


Fig. 3 Rate of used materials except pipe

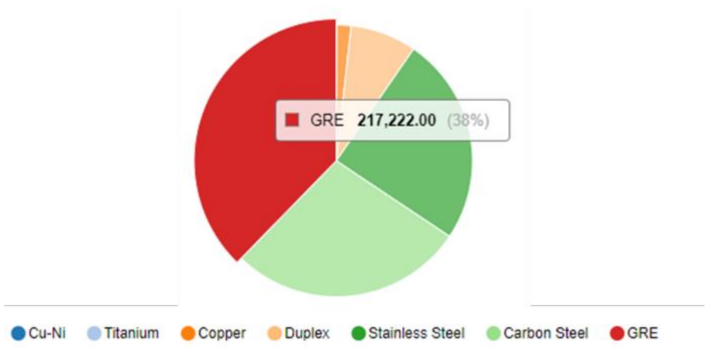


Fig. 4 Rate of materials of pipe at topside

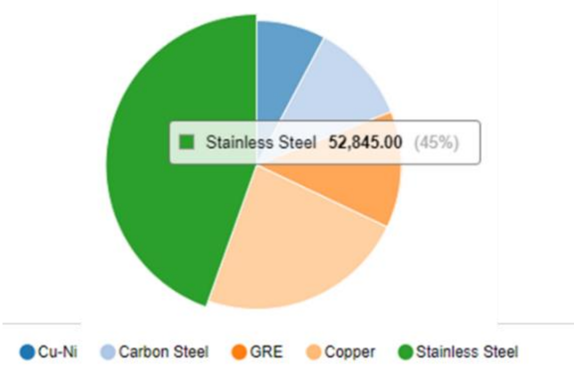


Fig. 5 Rate of materials of pipe at living quarter

Fig. 4 shows the ratio of the length of the pipe per material used at the topside, and the values are shown in Table 4. The most used material is GRE, but the second most used is the carbon steel, and the third most used is stainless steel. From Table 3, it was found that the stainless steel is the second most used for the entire offshore structure, but this changed when the analysis is limited to the topside. This means that stainless steel material is more used in living quarters. One important aspect is the portion of the duplex pipe. When the analysis is limited to the topside, the ratio of the duplex material is increased from 6.52% in Table 2 to 7.86%. The reason is that the duplex pipe is only used in the topside. Another peculiarity is that the ratios of copper and Cu-Ni alloy materials are significantly reduced. These observations will be discussed further in Fig. 5 and Table 5.

Fig. 5 shows the ratio of the length of the pipe per material used in the living quarters, and the values are shown in Table 5. It is found that the proportion of stainless steel is high and that the ratio of copper and Cu-Ni alloy is higher than that of the topside. As the copper and Cu-Ni alloy materials do not corrode, the materials are mainly used in the fresh water line, so the proportion of the materials in the living quarter is increased. Stainless steel piping lines can be delivered in an assembled form, so stainless steel material has been used to shorten the working time in a relatively narrow space of living quarters in the offshore structure. It is also found that the duplex material is not used in the living quarters. The duplex pipe is used for the chemical processing, and the cost is relatively higher than that of the other materias. In the same way, the other materials except for pipe can be analyzed.

Fig. 6 and Table 6 show the materials that are used topside except for pipe. The composition is similar to the ratio of the pipe material on topside because the material with the same material of the pipe should be connected to the pipe.

Table 5 Used materials of pipe at the living quarter

| Material | Quantity (m) | Percentage (%) |
|---|---:|---:|
| Stainless Steel | 52,845 | 44.56 |
| Copper | 27,703 | 23.36 |
| GRE | 15,507 | 13.07 |
| Carbon Steel | 13,341 | 11.25 |
| Cu-Ni | 9,205 | 7.76 |
| Sum | 118,601 | 100.00 |

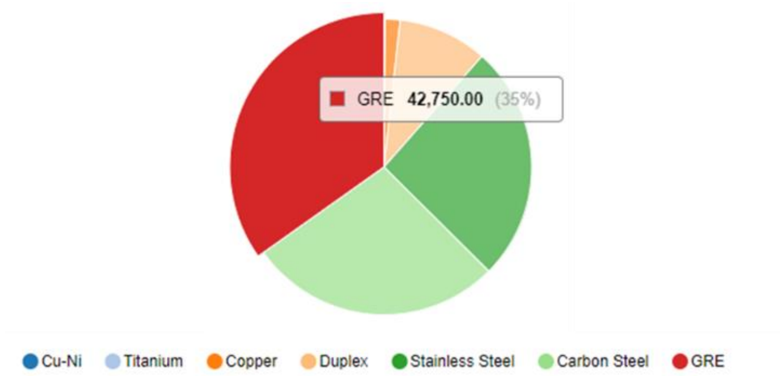


Fig. 6 Rate of materials except pipe at topside

Table 6 Used materials except pipe at topside

| Material | Quantity (EA) | Percentage (%) |
|---|---|---|
| GRE | 42,750 | 34.82 |
| Carbon Steel | 34,043 | 27.73 |
| Stainless Steel | 31,828 | 25.92 |
| Duplex | 11,998 | 9.77 |
| Copper | 2,005 | 1.63 |
| Titanium | 149 | 0.12 |
| Cu-Ni | 10 | 0.01 |
| Sum | 122783 | 100.00 |

Table 7 Used materials except pipe at living quarter

| Material | Quantity (EA) | Percentage (%) |
|---|---|---|
| Cu-Ni | 1,117 | 4.48 |
| Carbon Steel | 3,249 | 13.03 |
| GRE | 4,541 | 18.21 |
| Copper | 5,023 | 20.14 |
| Stainless Steel | 11,008 | 44.14 |
| Sum | 24,938 | 100.00 |

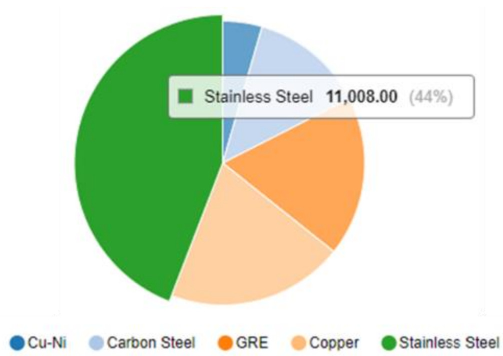


Fig. 7 Rate of materials except pipe at living quarter

In the same way, the materials can be checked for the living quarters from Fig. 7 and Table 7. It is also found that the composition of the materials is similar to that of the piping materials used in the living quarters.

### *4.2.2 Material quantity analysis*

In this section, the number of piping materials will be discussed. By analyzing the piping material list, it is possible to determine how many materials are used per material types, such as the elbow, flange, gasket, etc.

Fig. 8 shows the most used piping materials of the offshore structure. Elbow has the largest quantity of 35,549, followed by flange 17,323, and gasket 16,698. Based on these results, it is possible to calculate how many total materials are used in one project.

Fig. 9 is the result of analyzing the piping materials only for the GRE material except for pipe. It was found that the GRE material is the most used in the offshore structure from Table 3, and Fig. 9 shows how many quantities are used per piping material. Elbow with GRE material was used the most at 9,884, and flange and gasket materials with GRE material were also used often. In particular, the quantity of clamp shoe is 4,777, and it can be analyzed that the clamp shoe with GRE material is used to support GRE pipe.

Among the piping materials of the offshore structure, the total usage of pipe materials with GRE material can be analyzed by each module, as shown in Fig. 10. The largest total length of GRE pipe is 13,733 m, and this is found in the module 32. Likewise, the other materials with GRE except for pipe are analyzed as shown in Fig. 11. It is found that the distribution is similar to that of pipe material because the piping line is assembled with the same material.

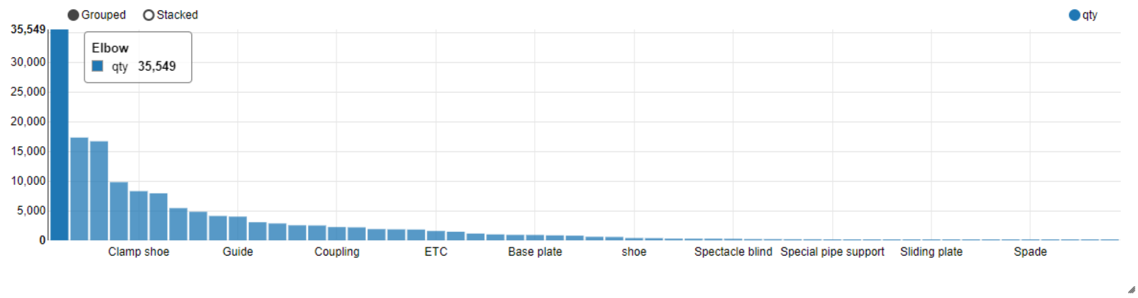


Fig. 8 Most used materials

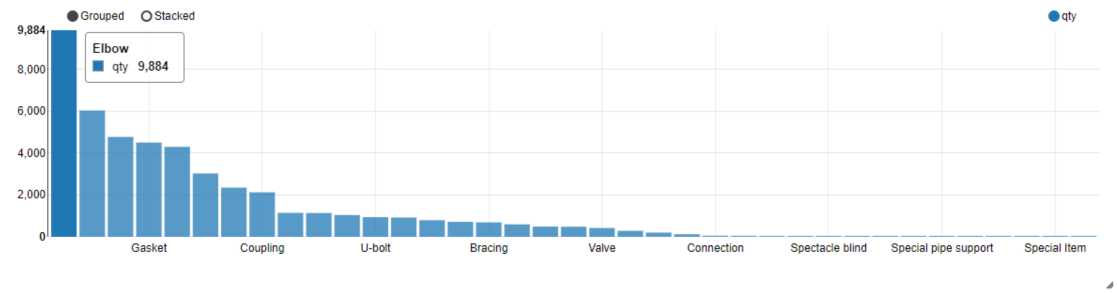


Fig. 9 Most used materials with GRE material

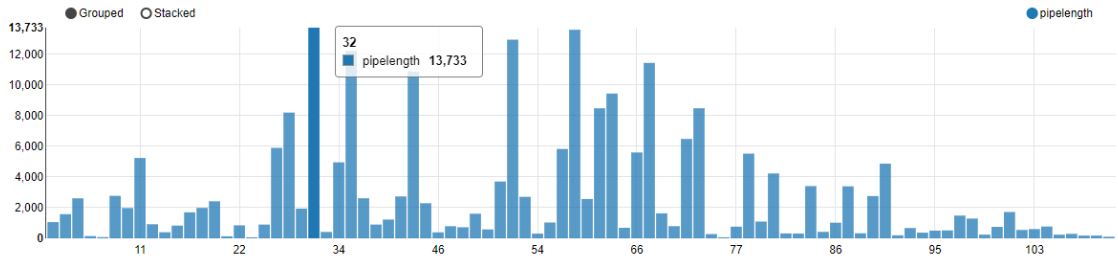


Fig. 10 Pipe length of GRE material per module

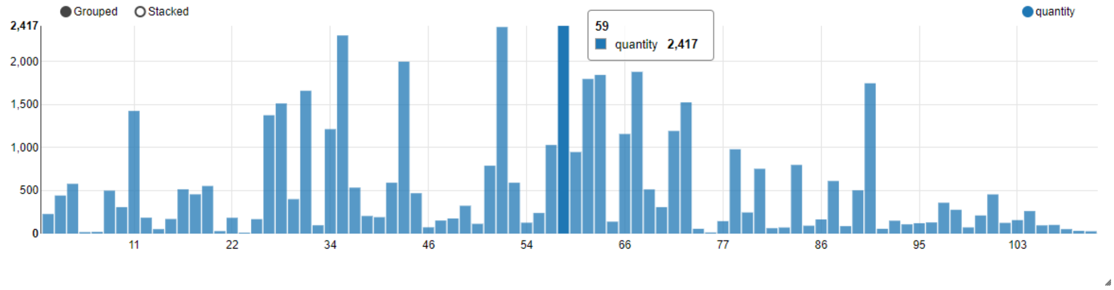


Fig. 11 Quantity of materials with GRE material per module

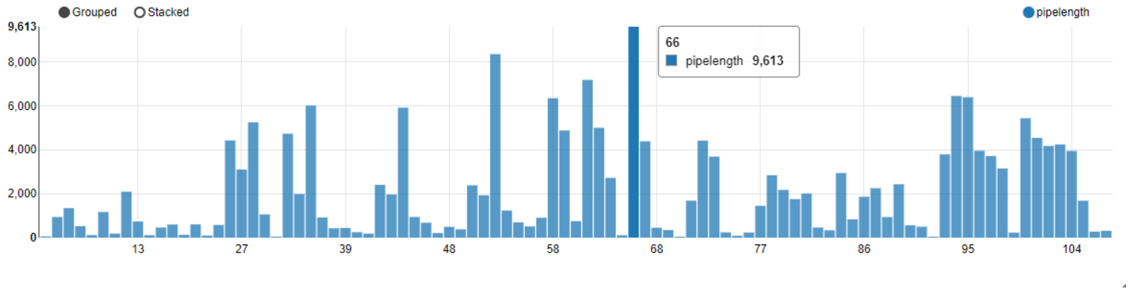


Fig. 12 Pipe length of stainless steel material per module

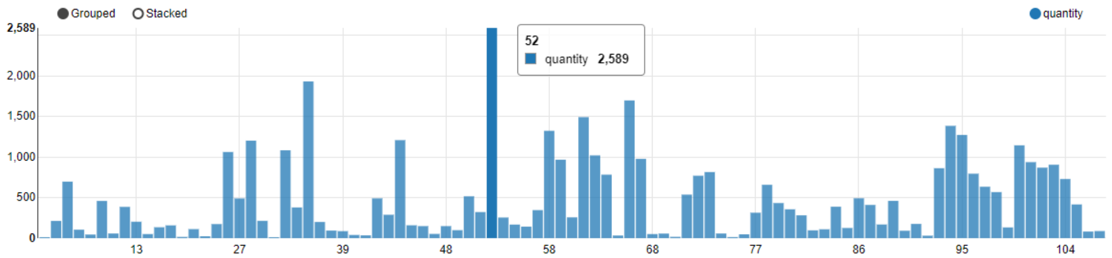


Fig. 13 Quantity of materials with stainless steel material per module except pipe

The same analysis is performed on stainless steel, and the results are shown in Fig. 12 and Fig. 13; Fig. 12 shows the result of usage for pipe material with stainless steel per module and Fig. 13 is the result of usage for the remaining materials with stainless steel per module. In particular, it is found that the use of stainless steel material in the living quarters is relatively high. The reason for this is that there are many clean water piping lines for the people on board, and stainless steel is used because the material is good for corrosion. In addition, considering the small installation area and short installation time, the assembly lines with stainless steel are used. This can also be seen in Fig. 13.

Unlike other steel pipe material such as carbon steel or stainless steel, duplex material indicates improved resistance to crevice corrosion and stress corrosion cracking. Because of this kind of advantage, it is often used in harsh environments, such as high-pressure processes or the transport of fluids containing sulfuric acid. Figs. 14 and 15 show the length of the duplex pipe and quantity of other materials with the duplex material, respectively. Module 28 contains the most duplex material, and we can predict that the process of the module will require a relatively harsh environment.

As previously mentioned, Cu-Ni material has a high corrosion resistance, and it is usually used for lines carrying drinking water. For the offshore structure, most of drinking water line is arranged at galley which is inside of the living quarters. The galley of the analyzed offshore structure is located at module 95 as shown in Figs. 16 and 17 and the analysis shows this fact precisely.

Copper has properties similar to those of Cu-Ni alloys but is used in slightly different applications. Copper material is also used to carry fresh water, but it also carries cooling water or refrigerant of air conditioners or cooling device of process equipment. For this reason, copper has been placed in a variety of modules instead of Cu-Ni alloys, and this can all be confirmed by Figs. 18 and 19.

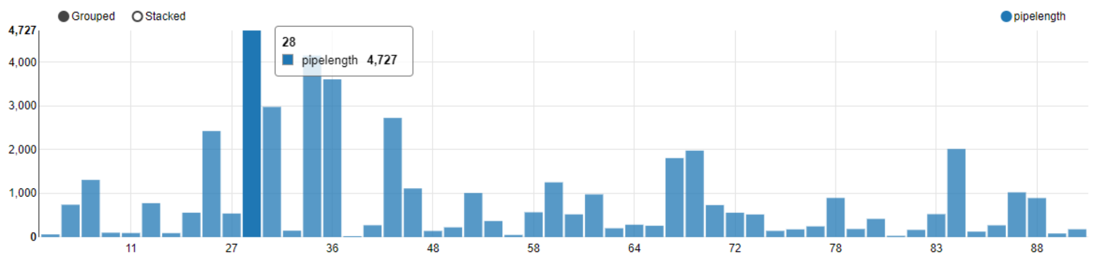


Fig. 14 Pipe length of duplex material per module

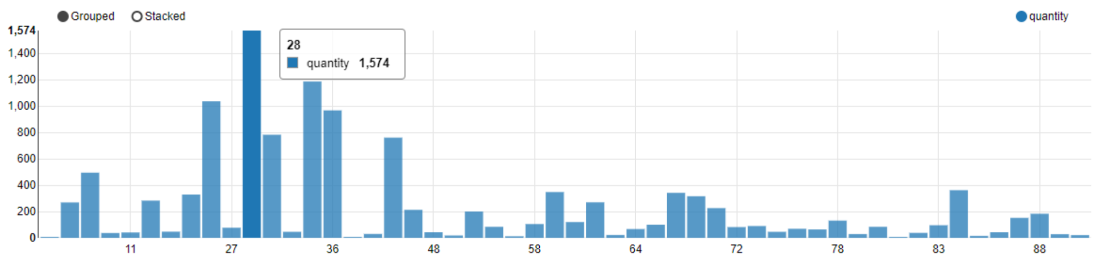


Fig. 15 Quantity of materials with duplex material per module except pipe

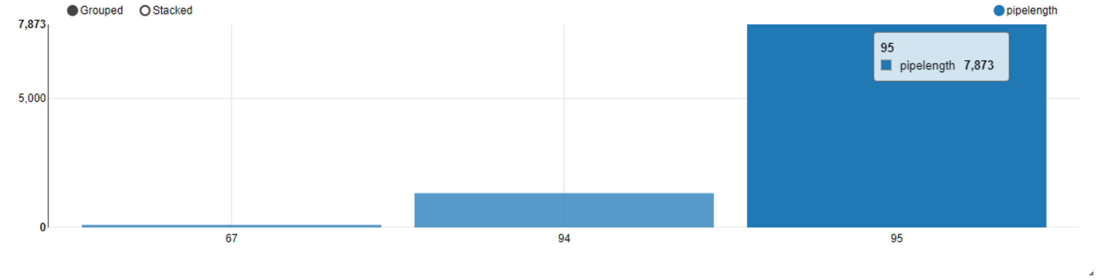


Fig. 16 Pipe length of Cu-Ni material per module

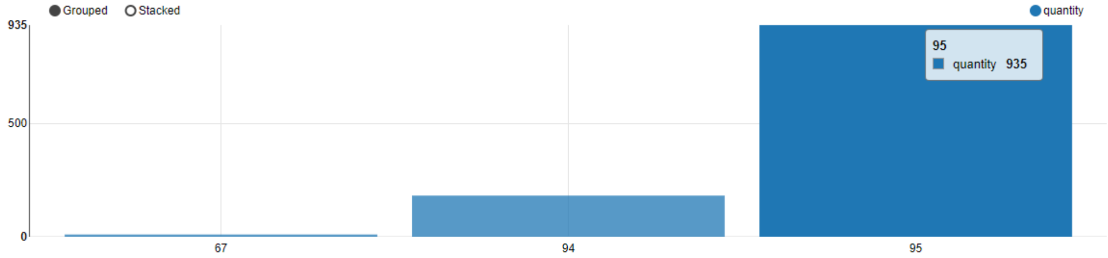


Fig. 17 Quantity of materials with Cu-Ni material per module except pipe

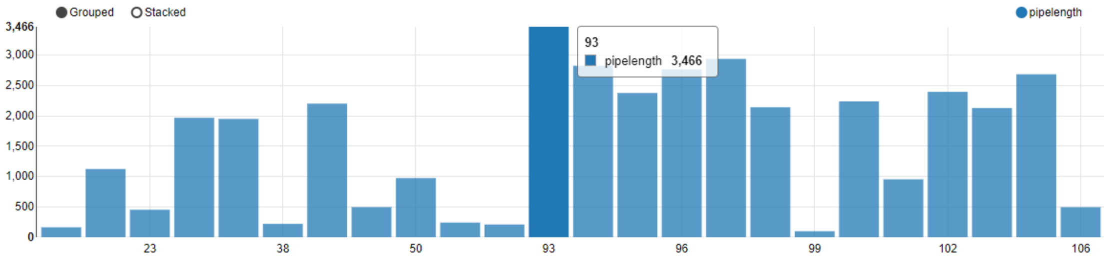


Fig. 18 Pipe length of copper material per module

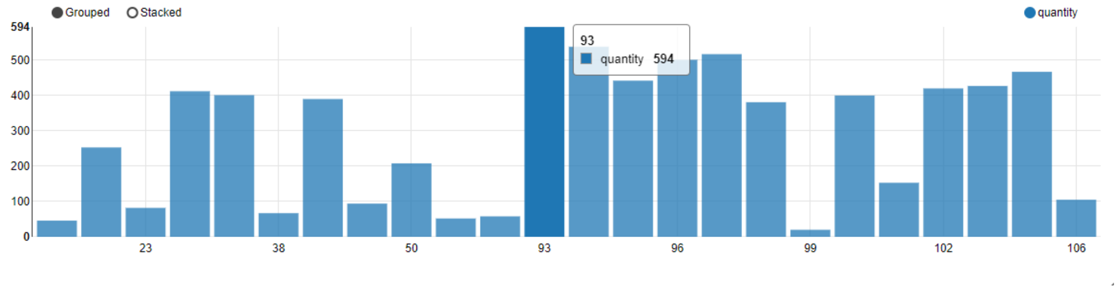


Fig. 19 Quantity of materials with copper material per module

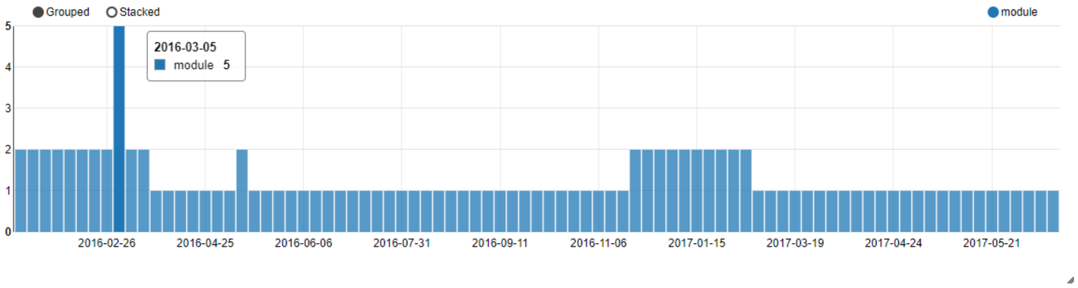


Fig. 20 Number of modules that start steel cutting per schedule

Fig. 20 shows the number of modules that started steel cutting per schedule. Through this analysis, it is possible to predict the workload of the production according to the progress of the process. According to the analysis result, the constructions of one or two modules are started on the overall schedule, and the workload is then evenly distributed. However, at a specific point in time, the steel cutting process for the five modules was carried out, indicating that the workload was high at that particular time.

### *4.3 Discussion*

The proposed big data framework is verified by analyzing the piping material of the offshore structure. The piping material list is saved in HDFS. The saved data is loaded to Hive for interactive analysis by sending queries. Spark can access Hive and process the loaded piping materials using queries. The queries can be sent using Zeppelin as a user interface. The data cube generation function of Kylin is used to analyze the piping materials used by defining the analysis range as either topside or living quarters. Kylin can be used when the data size is so large that it takes a long processing time even though the data is processed by the big data framework. In this study, only the functionality of Kylin is checked because the data size is not so big. The analysis results are visualized using Zeppelin, as shown in the figures.

The piping materials and its properties are analyzed, and the statistics about the material type, quantity, module, etc. are obtained. In addition, by analyzing the number of piping materials used, it could be found that the elbow material is most used except for pipe material. The piping materials with GRE except pipe were also analyzed, and the amount of the materials used are in the descending order of elbow, flange, clamp shoe, and gasket, and the numbers used were also correctly calculated. It also found that where the duplex, Cu-Ni, and copper material properties are used, and the reasons why the material properties are used are discussed. In addition, by analyzing the steel cutting point of each module, it was possible to analyze at which point the work was overloaded. Based on these analyses, it is expected to estimate the precise cost for the offshore project with the other information, such as the material cost, labor cost, wasted materials, and so on.

## 5. Conclusions

In this study, we present a big data-based piping material analysis for contract design. We analyzed the piping materials and checked what kinds of materials were used. These result can be used for estimating the price quotation for a new offshore project. The engineers can analyze the piping meterial data using query language. This is also faster than the manual work for the analysis. We confirmed that the big data framework proposed in this study could be used as a data analysis tool for the contract phase with interactive and fast analysis of big data.

The data management systems such as PLM or ERP, which have been used in the maritime industry, are based on the relational database. However, because these systems are not based on the big data technologies, it takes a lot of time to analyze many kinds of materials, and it is especially difficult to apply them to a comparative analysis of various projects. Therefore, if the data management system is developed based on the big data framework proposed in this study, it is expected that the big data can be processed efficiently. As a result, much more meaningful information can be obtained, and this results in the decreased cost of contract, production, and operation.

In this study, the piping material analysis for a single offshore structure was performed. It is possible to verify the function of big data framework because of various kinds of material and quantity. However, it is difficult to define it as a big data accurately based on the amount of data. The big data handling framework is not required to process the presented data because it can be analyzed using Microsoft Excel. However, the real data in the ship and offshore industries is much larger, and it is difficult to process the scaled data using the traditional software and hardware at once. The ship and offshore industries already have been accumulating such data, so it could be possible to estimate the initial material cost using the proposed big data framework. It is sure that the proposed framework can process big data as it is based on the big data technologies, but its efficiency should be checked for real applications. Therefore, future studies should investigate the data processing, analysis, and comparison of several offshore projects at once.

## Acknowledgements

This work was partially supported by (a) MSIP (Ministry of Science, ICT and Future Planning), Republic of Korea, under Development of Ship Design Standard PLM Platform based on Big Data (NIPA-2016-S1106-16-1025) supervised by the NIPA (National IT Industry Promotion Agency) and (b) Research Institute of Marine Systems Engineering of Seoul National University, Republic of Korea.

## References

Amazon Web Services Inc. (2018), Amazon Web Services (AWS) - Cloud Computing Services [WWW Document]. URL https://aws.amazon.com/ (accessed 7.9.18).

Apache Ambari [WWW Document] (2018), URL http://ambari.apache.org/ (accessed 7.11.18).

Apache Hadoop [WWW Document] (2018), URL http://hadoop.apache.org/ (accessed 7.11.18).

Apache HBase [WWW Document] (2018), URL http://hbase.apache.org/ (accessed 7.11.18).

Apache Hive [WWW Document] (2018), URL https://hive.apache.org/ (accessed 7.11.18).

Apache Kylin [WWW Document] (2018), URL http://kylin.apache.org/ (accessed 7.11.18).

Apache Zeppelin [WWW Document] (2018), URL http://zeppelin.apache.org/ (accessed 7.11.18).

Chen, M., Mao, S. and Liu, Y. (2014), "Big data: A survey", *Mob. Networks Appl.*, **19**, 171-209. https://doi.org/10.1007/s11036-013-0489-0

DNV-GL (2018), AIS – Meant for navigational safety, used for business intelligence - DNV GL [WWW Document]. URL https://www.dnvgl.com/maritime/energy-efficiency/automatic-identification-system-data-insights.html (accessed 7.9.18).

Feblowitz, J. (2013), Analytics in Oil and Gas: The Big Deal About Big Data, in: SPE Digital Energy Conference, 5-7 March, The Woodlands, Texas, USA. p. 6. https://doi.org/http://dx.doi.org/10.2118/163717-MS

Gandomi, A. and Haider, M. (2015), “Beyond the hype: Big data concepts, methods, and analytics”, *Int. J. Inf. Manage*., **35**, 137-144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Google (2018), Google Cloud [WWW Document]. URL https://cloud.google.com/ (accessed 7.9.18).

Kelleher, J.D. and Tierney, B. (2018), Data Science, MIT Press.

Kim, K.I., Jeong, J.S. and Park, G.K. (2013), “Assessment of external force acting on ship using big data in Maritime traffic”, *J. Korean Inst. Intell. Syst*., **23**, 379-384. https://doi.org/10.5391/JKIIS.2013.23.5.379

Kim, K.S., Roh, M.I., Lee, S.M., Kim, H.S. and Ahn, H. (2018), “Weight-estimation method of FPSO topsides considering the work breakdown structure”, *J. Offshore Mech. Arct*., **140**, 1-10. https://doi.org/10.1115/1.4037828

Kim, S., Roh, M., Kim, K. and Oh, M. (2017), “Big data platform based on hadoop and application to weight estimation of FPSO topside”, *J. Adv. Res. Ocean Eng*., **3**, 32-40.

Lee, Y. (2017), A Reference Model for Big Data Analysis in Shipbuilding Industry, Ulsan National Institute of Science and Technology.

Lytra, I., Vidal, M.E., Orlandi, F. and Attard, J. (2017), “A big data architecture for managing oceans of data and maritime applications”, *Proceedings of the 2017 International Conference on Engineering, Technology and Innovation (ICE/ITMC). IEEE*, 1216-1226. https://doi.org/10.1109/ICE.2017.8280019

Mauro, A.D., Greco, M. and Grimaldi, M. (2015), “What is big data? A consensual definition and a review of key research topics”, *Proceedings of the AIP Conference*, 97-104. https://doi.org/10.1063/1.4907823

Mourouzis, T. (2018), How Big Data Analytics Will Transform the Shipping Industry – Cyprus International Institute of Management [WWW Document]. URL https://www.ciim.ac.cy/how-big-data-analytics-will-transform-the-shipping-industry/ (accessed 7.8.18).

Oh, M.J., Roh, M.I., Park, S.W. and Kim, S.H. (2018), “Estimation of material requirement of piping materials in an offshore structure using big data analysis”, *J. Soc. Nav. Archit. Korea*, **55**, 243-251. https://doi.org/10.3744/SNAK.2018.55.3.243

O'Neil, C. and Schutt, R. (2013), Doing Data Science: Straight Talk from the Frontline, O'Reilly Media.

Park, S.W., Roh, M.I., Oh, M.J. and Kim, S.H. (2018), “Association analysis of piping materials of an offshore structure using big data technology”, *J. Sh. Prod. Des*., https://doi.org/10.5957/JSPD.170058

Perera, L.P. and Mo, B. (2017), “Machine intelligence based data handling framework for ship energy efficiency”, *IEEE T. Veh. Technol*., **66**, 1-8. https://doi.org/10.1109/TVT.2017.2701501

Pettey, C. and Goasduff, L. (2011), Gartner Says Solving “Big Data” Challenge Involves More Than Just Managing Volumes of Data [WWW Document]. Gartner. URL https://www.gartner.com/newsroom/id/1731916 (accessed 7.6.18).

Wang, H., Osen, O.L., Li, G., Li, W., Dai, H.N. and Zeng, W. (2015), “Big data and industrial Internet of Things for the maritime industry in Northwestern Norway”, *Proceedings of the 2015 IEEE Reg. 10 Conf.*, https://doi.org/10.1109/TENCON.2015.7372918

Yoo, S.L. (2018), “Near-miss density map for safe navigation of ships”, *Ocean Eng*., **163**, 15-21. https://doi.org/10.1016/j.oceaneng.2018.05.065

MK