Establishment of Ship Allocation Model by Using Marine Logistics Database (MLDB)

Mohammad Danil Arifin[1], Fanny Octaviani[2]
(Received: 20 April 2022 / Revised: 3 June 2022 / Accepted: 4 June 2022)

***Abstract*— recently, marine big data are significantly increased. If the data are effectively analyzed, it can give an advantage, and we can harness the data that is useful for a decision-maker in maritime industries. The Marine Logistics Database (MLDB) was successfully developed in the previous studies. It was developed by integrating big data into a relational database. By utilizing the extracted data from the developed database (DB), the model of ship allocation will be established. In this study, the main purpose is to develop a ship allocation model that matches the results with the existing ship allocation. In this study, the effectiveness of the allocation model was examined by checking the port constraints, ship specification, and allocation process itself. Moreover, some simulations were executed and discussed to develop new ship allocation and analyze the effective ship specification.**

Keywords—big data, marine logistics database, ship allocation.

I. Introduction

In recent years, big data (BD) has gathered enormous attention from academic researchers, governments in all aspects of information, and research institutes. In the case of the maritime fields, maritime data grow exponentially with the forming of the diversity of maritime data acquisition techniques, which formed as maritime big data. Maritime big data contain a vast value and embodies an enormous academic interest that can be converted into a pack of information for everyone to learn, explore, and preserve the maritime-related field, i.e., ecology, climate, disaster, and shipping industries.

For example, by observing the seismic and faulting activity data, the tsunami and undersea earthquakes can be successfully forecasted [1, 2]. The other example, by analyzing Argo data, the earth seeks to intensify the global hydrological cycle [3]. In addition, analyzing the acoustic remote sensing data of groups and species distribution can be robust scientific supporting confirmation to maintain the balance of maritime ecological [4]. In such a case, it can be realized that marine big data support warning potential problem and forecasting to help decision making.

Maritime big data is used for some applications in the shipping industry, i.e., by using the BDO (Big DataOcean). It can be used to semantically enrich and link data about the maintenance schedule to identify the maintenance schedules of vessels in the data lake and visualize the impact of equipment maintenance on the vessels [5]. The system dynamic model can be established by analyzing the big data correlation such as economic growth, sea cargo movement, ship bottom, ship order, ship construction, and scrapping ship to conduct the demand-forecasting [6]. Another example is the visualizing or monitoring system, which is important for ship construction and the visualization system of the cutting and subassembly processes [7].

In addition, the data handling frameworks with various types of data analytics are proposed by collecting the various onboard internet of things, sensors, and acquisition systems data based on the ship performance and navigation data [8]. Using these frameworks can enhance the quality and decrease the quantity of navigation information and ship performance, and it can enhance the standard and visualize the information appropriately. In the case of the exhaust gas emission evaluation, by utilizing the big data combination of geographic information system (GIS) and automatic identification system (AIS). The ship exhaust gas emission distribution can be predicted and calculated easily [9-10]. So, the green shipping environment could be shortly achieved through these data analyses.

In the previous study, to manage and utilize the availability of marine big data, such as route, port, ship, trade, and AIS data, the MLDB (Marine Logistics Database) was developed [11]. Moreover, using the extracted data from the developed database, the support system of ship basic planning also can be developed [12-15]. In summary, marine big data is also supported in shipping industries focused on ship operation, ship construction, ship maintenance, demand forecasting, ship emission, etc. However, there is limited research that is focused on ship allocation. Therefore, in this research, we focus on establishing the ship allocation model by using MLDB.

By comparing the estimation result of cargo volume with the trade volume amount from the international trade statistic database, the developed MLDB has been validated. The comparison result is shown in Figure 3. It shows that the cargo volume coverage of iron ore from Australia to Japan is 95%. At the same time, the cargo volume coverage of coal from Australia to Japan is reaching 91%. In summary, it is shown that the data extracted from the MLDB is reliable to be used.

Mohammad Danil Arifin is with Marine Engineering Department, Darma Persada University, Jakarta Timur, 13450, Indonesia. E-mail: danilarifin.mohammad@gmail.com
Fanny Octaviani is with Department of Naval Architecture, Darma Persada University, Jakarta, 13450, Indonesia. E-mail: fanny_octaviani@yahoo.com

A. Review of MLDB

MLDB is defined as marine big data integration structured into a database. Input data of marine logistics DB consists of five BD. For example, AIS data, ship data, port data, route data, and trade data. The concept of the MLDB is shown in Figure 1, and the structure of the MLDB is shown in Figure 2.

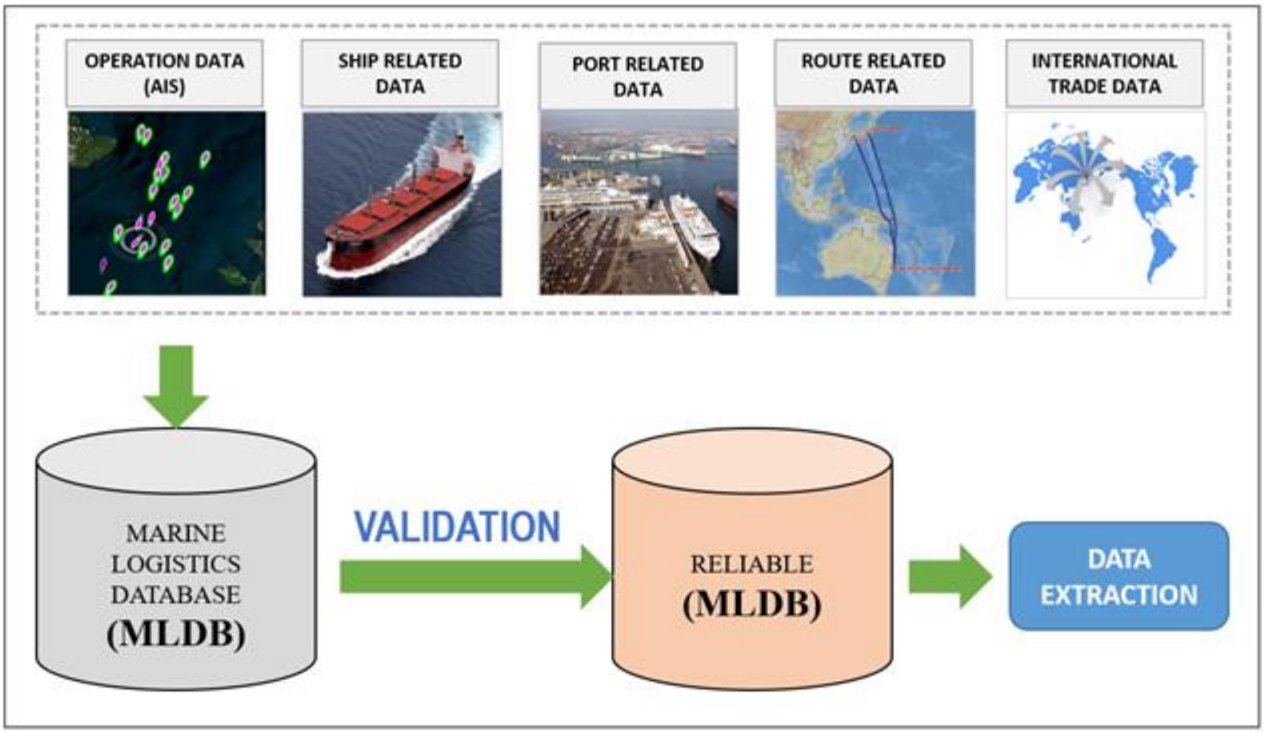


Figure 1. The basic concept of MLDB

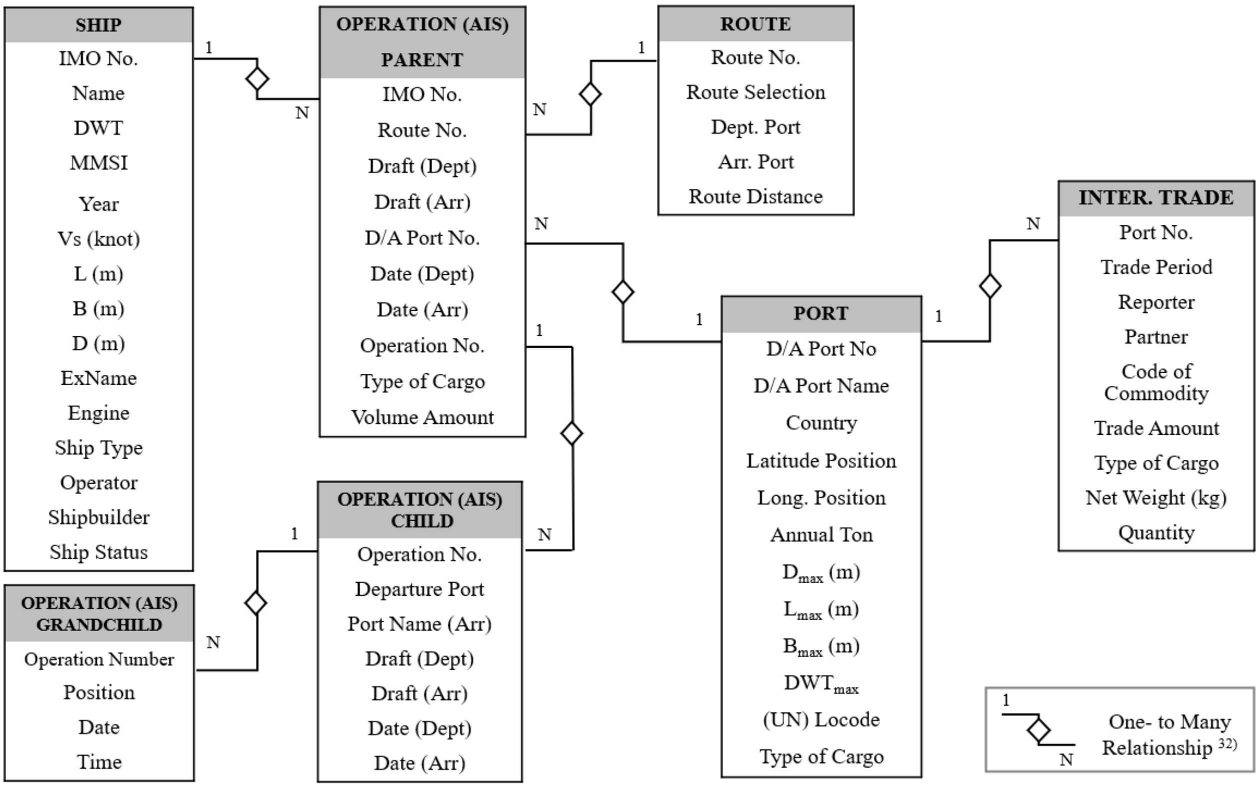


Figure 2. The structure of MLDB

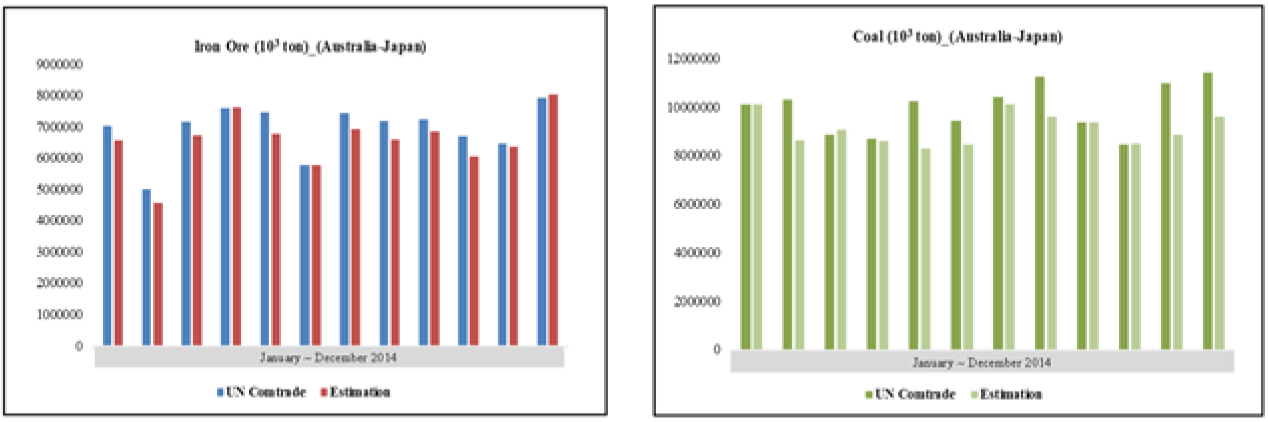
By comparing the cargo volume of estimation result with the amount of cargo volume from the international al trade data, the developed MLDB has been validated. The comparison result is shown in **Figure 3**. It shows that the cargo volume coverage of iron ore from Australia to Japan is 95%. At the same time, the cargo volume coverage of coal from Australia to Japan is reaching 91%. In summary, it is shown that the data extracted from the MLDB is reliable to be used.

B. Establishment of Allocation Model of Ship

The allocation model of the ship in this research is defined as a model that can reproduce the actual ship allocation using the extracted data from MLDB.

As an input to establish the ship allocation model following information is used: (1) canal and port information, i.e., DWT_{max}, LOA_{max}, B_{max}, D_{max}, etc.; (2) trade volume information and fuel price; (3) ship specification, i.e., DWT, LOA (m), B (m), service speed (knot), horsepower (hp), etc.

In this study, the allocation model of the ship is developed by the three selection models: the model of shipper, the model of shipowner model, and the model of the operator. Output to be achieved in this study is the ship allocation model's reproducibility that matches the ship operation and the port constraint. The principal concept of the proposed model is expressed in **Figure 4.**



(a) (b)

Figure 3. (a) Comparison results of iron ore, and (b) comparison results of coal

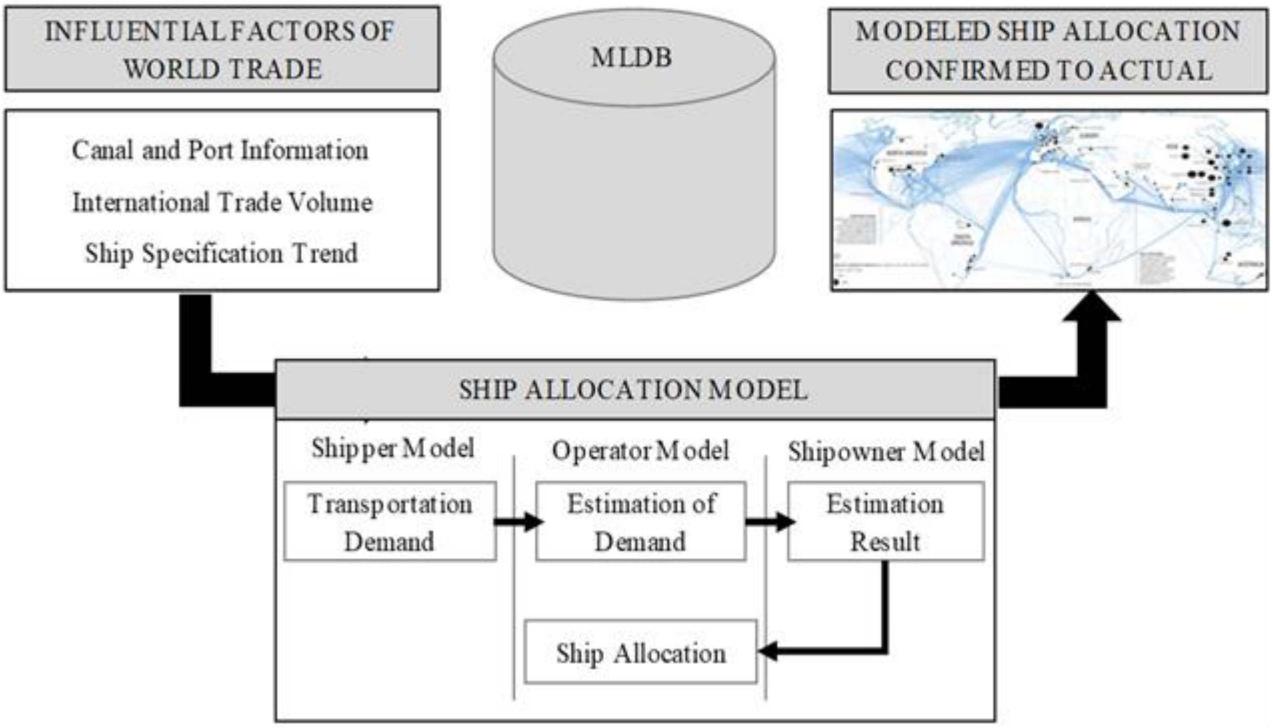


Figure 4. (a) The structure of MLDB

II. METHOD

As explained in the previous section, the ship allocation model was established by three distinct models, i.e., shipper, shipowner, and operator. A detailed explanation of each model is described as follows:

A. Shipper Model

The shipper model has defined a cargo transportation request between the shipper and consignee. The shipper model is generated by performing cluster analysis. Cluster analysis enabled defining the similitude of data to group similar data and collectively organize them into several clusters. The following steps were executed to develop this model:

1) Ship operation data extraction from MLDB: after the data was extracted, the arrival data was managed into matrix form; an example of ship operation is shown in **Table 1.**

TABLE 1.
MATRIX OF SHIP ARRIVAL DATA

	Port 1	Port 2	Port 3	Port 4	Port 5
Ship 1	0	0	0	5	2
Ship 2	0	0	0	1	0
Ship 3	1	1	0	0	0
Ship 4	0	0	1	0	0
Ship 5	2	3	0	0	0

2) Standardization calculation of ship arrival data: the standardized results are shown in **Table 2**.

TABLE 2.
STANDARDIZED SHIP ARRIVAL DATA

	Port 1	Port 2	Port 3	Port 4	Port 5
Ship 1	-0.7	-0.9	-0.4	2.19	2.24
Ship 2	-0.7	-0.9	-0.4	0	-0.4
Ship 3	0.65	0	-0.4	-0.5	-0.4
Ship 4	-0.7	-0.9	2.24	-0.5	-0.4
Ship 5	1.96	1.73	-0.4	-0.5	-0.4

3) Euclidean distance calculation of ship arrival data: the Euclidean distance was calculated using the standardized or normalized arrival data using equation (1), where x_i and y_i are the numbers of calls after standardizing ship i at port x and y, respectively. The results it is shown in **Table 3**.

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \quad (1)$$

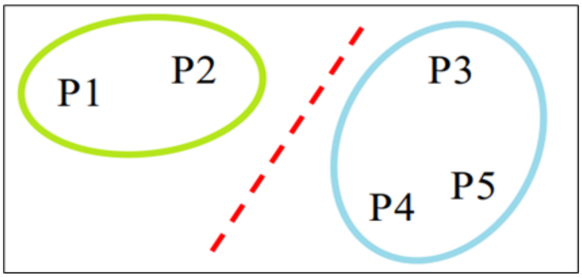
TABLE 3.
EUCLIDEAN DISTANCE

	Port 1	Port 2	Port 3	Port 4	Port 5
Port 1		1.71	3.94	4.04	3.94
Port 2	1.71		4.08	4.21	4.08
Port 3	3.94	4.08		3.87	3.79
Port 4	4.04	4.21	3.87		0.49
Port 5	3.94	4.08	3.79	0.49	

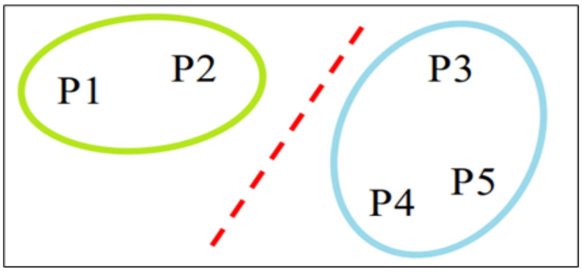
4) Defined port cluster using average linkage method and define a dendrogram: the hierarchical cluster is defined using the equation (2). The clustering result is illustrated in **Figure 5(a)**; otherwise, the dendrogram is illustrated in **Figure 5(b).**

$$d(C_1, C_2) = \frac{1}{|C_1||C_2|} \sum_{x_1 \in C_1} \sum_{x_2 \in C_2} d(x_1, x_2) \quad (2)$$

C_n is defined as a cluster, x_n is indicated as a port, and $d(C_1, C_2)$ is the Euclidean distance between C_1 and C_2.



(a)



(b)

Figure 5 (a-b). Illustrated hierarchical cluster and dendrogram

B. Shipowner Model

In this section, the shipowner model is described. This model is defined as an estimation of cargo transport volume and cost. The draught rate, sailing speed, and port staying time are estimated by generating deep learning to realize the shipowner model. Deep learning is a machine learning technique with a process for inputting the data for training into an early stage that is a learning machine. The next step is generating a learning model, and the last is processing data using the learning model.

The following step was executed to establish the shipowner model: (a) data extraction from MLDB, (b) creating of the deep learning model, and (c) calculating the shipment days, cargo transport volume, and operational transportation cost. The process of deep learning estimation methods is illustrated in **Figure 6.**

C. Operator Model

The selection of the most acceptable ship to be allocated on the selected route based on the shipowner model estimation result is defined as the operator model. The primary step to determine ship allocation was performed as follows:

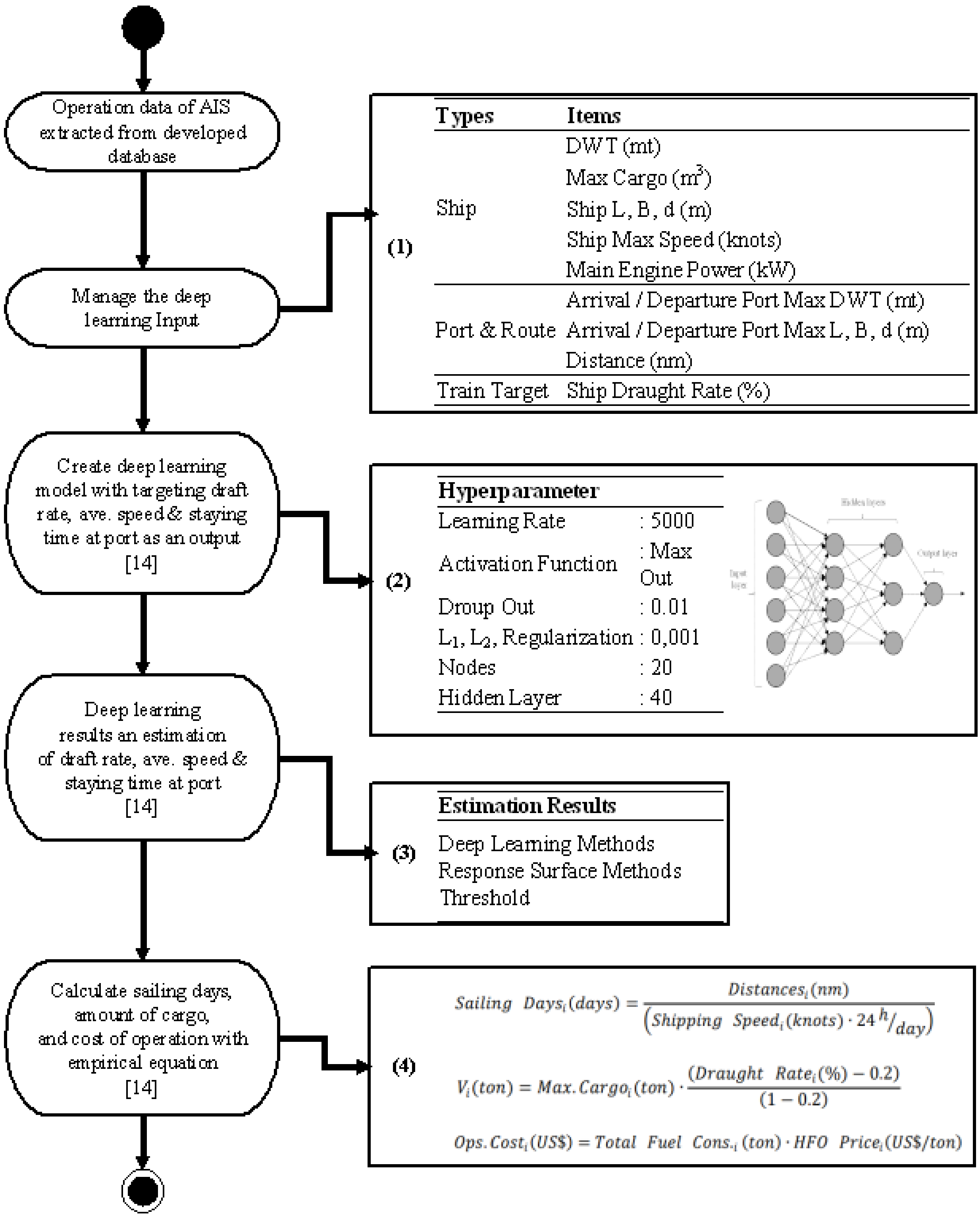


Figure 6. Process of deep learning estimation methods

1) Calculation of the total cost and amount of cargo volume: as an example, shipowners proposed for all shipment requests (ship A to ship D) based on the selected route (route A to route B).

   By examining the total amount of operation cost and the total amount of transportation volume *t*, the cost per unit of transport volume was developed. Each shipper's calculated the amount of operation cost and the amount of cargo volume are managed in matrix form for each selected route, as shown in **Table 4.**

TABLE 4.
TOTAL COST AMOUNT AND CARGO VOLUME

Shipper	Route	Cargo Amount (t)	Ship A ($/t)	Ship B ($/t)	Ship C ($/t)	Ship D ($/t)
A	A1	$3.5{\times}10^6$	14.8	14.1	16.9	19.9
	A2	$2.0{\times}10^6$	14.7	13.9	16.4	19.4
B	B1	$4.7{\times}10^6$	13.6	13	15.1	18.3
	B2	$6.0{\times}10^6$	13.1	12.6	14.5	18.2

2) Calculation of the standard deviation value: The deviation is defined as an index to decide which ship is suitable for delivering a given cargo type on a specific route.

   In this study, the standard deviation of some ships was calculated for each route due to the amount of cost per unit transport volume from the previous step. The estimation of the standard deviation is shown in **Table 5.**

TABLE 5.
STANDARD DEVIATION VALUE

Shipper	Route	Cargo Amount (t)	Deviation Value			
			Ship A	Ship B	Ship C	Ship D
A	A1	$3.5{\times}10^6$	57.2	60.31	47.9	34.59
	A2	$2.0{\times}10^6$	56.64	**60.43**	45.58	34.35
B	B1	$4.7{\times}10^6$	56.81	59.74	49.51	33.92
	B2	$6.0{\times}10^6$	56.84	59.12	50.45	33.58

3) Ship assignment: is defined as deciding the acceptable ship to be operated regularly by considering the value of standard deviation.

   The highest standard deviation values of the ship are assigned to the shipment operation on the selected route. As expressed in **Table 6,** the B ship was selected to operate in route A2.

TABLE 6.
THE SHIP ASSIGNMENT

Shipper	Route	Cargo Volume (t)	Ship A ($/t)	Ship B ($/t)	Ship C ($/t)	Ship D ($/t)
A	A1	$3.5{\times}10^6$	14.8	—	16.9	19.9
	A2	$\mathbf{0.6{\times}10^6}$	14.7	—	16.4	19.4
B	B1	$4.7{\times}10^6$	13.6	—	15.1	18.3
	B2	$6.0{\times}10^6$	13.1	—	14.5	18.2

III. RESULTS AND DISCUSSION

This study discussed the ship's operation from Australia to Japan or Japan to Australia extracted from MLDB. Based on the proposed methods, the result of the ship allocation model can be discussed as follows:

A. Shipper Model

In this research, the model was intended for iron ore from Australia (Shipper) to Japan (Consignee). As an evaluation of the shipper model, the clustering analysis of ship operations in 2014 was examined. The cluster analysis was done and evaluated due to the consignee point of view and shipper point of view, as shown as follows:

- Cluster result based on the point of view of the shipper: based on it, the shipper was grouped into three clusters as follows:
  1) Cluster A1: Port Hedland
  2) Cluster A2: Port Walcott and Dampier
  3) Cluster A3: Esperance and Parker Point
- Cluster result based on the point of view of the consignee: the consignee was grouped into four clusters as follows:
  1) Cluster J1: Chiba, Kawasaki, Fukuyama, Mizushima
  2) Cluster J2: Kashima, Oita, Kisarazu
  3) Cluster J3: Nagoya, Wakayama, Tobata
  4) Cluster J4: Hachinohe, Tomakomai, Saganoseki, Kure, Himeji, Higashi-Harima

To confirm the clustering result, the clustering result of the operation from Australia-Japan 2014 is compared to the clustering result of the operation from Australia-Japan 2017. The clustering result of its operation is compared in **Table 7**.

Based on the comparison result above, it can be seen that the clustering result of 2014 and 2017, both from shipper and consignee points of view, are consistent. It can be concluded that the results are confirmed.

Through the extracted data from MLDB and clustering results, some essential information, such as characteristics of shippers based on shipper and consignee point of view, could be identified, as presented in **Table 8**.

TABLE 7.
COMPARISON OF CLUSTERING RESULTS

2014	**2017**
Shipper point of view	Shipper point of view
Cluster A1: Port Hedland	Cluster A1: Port Hedland
Cluster A2: Port Walcott and Dampier	Cluster A2: Port Walcott and Dampier
Cluster A3: Esperance and Parker Point	Cluster A3: Esperance and Parker Point
Consignee point of view	Consignee point of view
Cluster J1: Chiba, Kawasaki, Mizushima, Fukuyama,	Cluster J1: Chiba, Kawasaki, Mizushima, Fukuyama
Cluster J2: Kashima, Oita, Kisarazu	Cluster J2: Kashima, Oita, Kisarazu
Cluster J3: Nagoya, Wakayama, Tobata	Cluster J3: Nagoya, Wakayama, Tobata
Cluster J4: Hachinohe, Tomakomai, Saganoseki, Kure, Himeji, Higashi-Harima	Cluster J4: Muroran, Saganoseki, Kure, Higashi-Harima

TABLE 8.
THE SHIPPER POINT OF VIEWS CHARACTERISTICS

Cluster	**Port Name**	**Port Constraints**			
		DWT (10^4)	**L (m)**	**B (m)**	**d (m)**
	Chiba	22	300	50	18
J1	Fukuyama	22	300	50	18
JFE Steel	Kawasaki	26	340	50	18
	Mizushima	22	340	50	18
J2	Kashima	30	340	60	19
Nippon Steel & Sumitomo	Oita	40	450	60	25
Metal Terminal (NSSMT)	Kisarazu	30	330	60	19
J3	Tobata	16	327	43	16
Nippon Steel & Sumitomo	Wakayama	16	300	43	14
Metal Terminal (NSSMT)	Nagoya	11	300	43	16
J4	Kure	27.6	360	45	18
KOBELCO-Nisshin	Higashi-Harima	18	330	47	17

TABLE 9.
THE CONSIGNEE POINT OF VIEWS CHARACTERISTICS

Cluster	**Port Name**	**Port Constraints**			
		DWT (10^4)	**L (m)**	**B (m)**	**d (m)**
A1					
BHP Billiton	Port Hedland	34	335	60	19.5
A2	Dampier	26	330	55	19.2
RIO Tinto A	Port Walcott	25	340	55	19.5
A3	Esperance	22	300	50	18.5
RIO Tinto B	Parker Point	22	300	50	18.0

B. *Shipowner Model*

The draft rate, average service speed (knot), and staying time in port (days) are estimated by generating the deep learning analysis as explained in the previous section. As an example, the results of the shipowner model are expressed in **Table 10**. Hence, the deep learning errors and their comparison are compared in **Table 11**.

TABLE 10.
RESULT OF THE PREDICTION BY DL

Operation Route	Ship	Average Speed (V_s Loaded)	Average Speed (V_s Unloaded)
Port	A	11.6 knot	11.7 knot
Hedland-	B	11.3 knot	12.4 knot
Chiba	C	11.0 knot	11.6 knot
Shipping Route	**Ship**	**Staying Time (Loaded)**	**Staying Time (Unloaded)**
Port	A	3.1 days	1.9 days
Hedland-	B	4.1 days	1.9 days
Chiba	C	3.4 days	1.6 days

TABLE 11.
DIFFERENTIATION OF DEEP LEARNING (DL), RESPONSE SURFACE (RS), AND THRESHOLD (TH)

Method	Draft Rate (%)	Service Speed (knot)	Staying Time (days)
DL	3.4	0.2	0.9
RS	5.9	-	-
Th	3.5	0.9	1.2

As illustrated in Table 10, the draft rate of error average using the DL method is 3.4% for the draft rate, the average of service speed is 0.2 knots, and the port staying time results is 0.9 days. As an evaluation of the shipowner model, the estimation results of the deep learning method are compared to other methods, i.e., response surface and threshold. The DL estimation result is better than the Th, although the RS method is the worse.

C. *Allocation Model of Ship*

The evaluation of the developed allocation model that is proposed in this study is examined by clarifying the applicability of the proposed model. The ship allocation model is simulated to evaluate the reproducibility. In this simulation, the number of the ship, the allowed ship specification, and the port limitation or constraint were set as inputs. The actual ship allocation was compared to the result of the developed allocation model. Furthermore, all information that is used to simulate the developed ship allocation was taken from the database.

As a result, the shipper is grouped into four clusters (J1-J4), as shown in **Figure 7**. The vertical axis is defined as the operation's number, and then the horizontal axis is defined as ship size (10^4 DWT). Based on the result, the cluster was clustered into six groups: In this study, to examine the applicability/reproducibility of the proposed model, the following process was executed:

- Port constraint
  Step 1: Extracting the port constraint from port data in the database. 1st, the port limitation was acquired from port data. However, some limitations were not available or did not match with the actual or real conditions.
  Step 2: Modify the port constraint using operation data. The port constraint modification was conducted based on the ship specification when the two data did not match, or the constraint was unavailable. Some port constraint is shown in **Table 8**. The white color reflects the data from step 1, and the grey reflects the modification data based on the actual operation.
- Ship specification
  The ship specification of the actual operation was extracted from MLDB. The specification of the typical ship for each ship size is shown in **Table 12**.

TABLE 12.
TYPICAL SHIP'S SPECIFICATIONS

DWT	L (m)	B (m)	D (m)	d (m)	HP
106,507	255	43	19	13	16,680
177,855	292	45	25	18	22,920
210,036	300	50	25	18	21,808
229,013	320	54	24	18	30,499
250,813	330	57	25	18	29,789
297,736	325	55	29	21	30,808

- Allocation process
  Figure 7 should be started from the ship in J3 because the port limitations were the most severe. Hence, breadth ≤ 43 m and draft ≤ 14 m became a functional limitation, and ships with 10 (10^4 DWT) (breadth ≤ 43 m) were selected. After that, the J4 was selected to be allocated where the breadth should be lower than 45m. After that, the Shipper J1 should be allocated because its port limitation was tougher than the limitation of shipper J2. The last remaining ships should be allocated to shipper J2.
  As represented in **Figure 7**, the results for the simulation of all shippers (J1-J4) mostly coincident with the real or actual condition. Moreover, it shows that shipper J1 generally used 21 (10^4 DWT) ships for their operation. Shipper J2 used several types of ships from 17-30 (10^4 DWT).

Since the ship allocation simulation results coincide with the real or actual ship allocation and the ship specification generally coincides with port limitations, the reproducibility and applicability of the proposed model are confirmed.

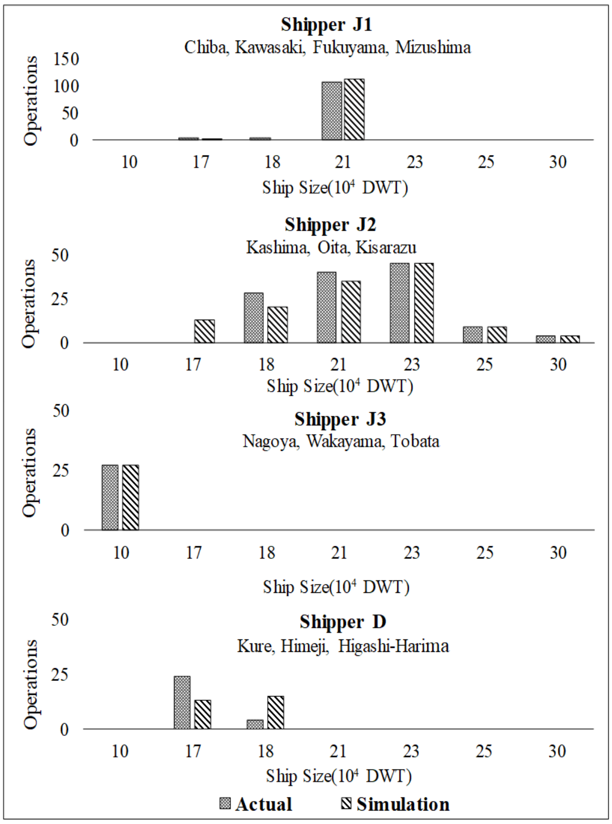


Figure 7. Comparison of actual allocation and simulation

IV. CONCLUSION

In this study, the developed database manages the available data from maritime BD extracted from MLDB to establish the model of ship allocation. The shipper, shipowner, and operator models established the ship allocation model. The shipper can be clustered easily, and each shipper's characteristics can be identified using the developed model. The reproducibility and applicability of the ship allocation model were confirmed using the proposed model. It was confirmed based on the evaluation by considering the three following processes: port constraints, ship specification, and allocation process. Based on the results, it can be concluded that the estimation results of ship allocation matched the actual ship allocation.

REFERENCES

[1] Rogers G C, Meldrum R D, Heesemann M, Mulder T L, Brillon C D, Cassidy J F, 2009 *Seismological Research Letters.* **81** no 2 pp 369

[2] Alexander B R, Richard E T, Isaac V F, 2013 *Pure and Applied Geophysics.* **170** no 9-10 pp 1529-1565

[3] Paul J D, Susan EW, Richard J M, 2012 *Science.* **336** pp 455-458

[4] Craig J B, Stephen J S, Peter L, John T A, 2011 *Estuarine, Coastal, and Shelf Science, Science Direct.* **95** pp 502-520

[5] Ioanna L, Maria E V, Fabrizio O, Judie A, 2017 *Proc. of International Conference on Engineering, Technology and Innovation.* pp 1216-1226

[6] Yujiro W, Kunihiro H, Noritaka H, Yui I, M.D Arifin, 2015 A study on system dynamics model for the demand forecasting of shipbuilding *Proc. of the 12th International Marine Design Conference* **2** pp 167-182

[7] Aoyama K, Yotsuzuka T, Tanaka Y, Tanabe Y, 2021 Monitoring platform of monitoring and visualizing system for shipyard: Application to Cutting and subassembly processes *Practical Design of Ships and Other Floating Structures* **65** pp 321-337

[8] Lokukaluge P P, 2017 Handling big data in ship performance and navigation monitoring *Proc. of Smart Ship Technology, The Royal Institution of Naval Architects Smart Ship Technology, The Royal Institution of Naval Architect* pp 89-97

[9] Artana KB, Pratiwi E, Dinariyana B, Pitana T, Masroeri A, Ariana M, M.D Arifin, Enhancement on methodology for estimating emission distribution at Madura Strait by integrating automatic identification system (AIS) and Geographic Identification system, 2014 *Proc. of 3rd International Symposium of Maritime Sciences* pp 219-224

[10] Hendra S, Adi M, Jaswar K, M.D Arifin, 2013 Estimation and distribution of exhaust ship emission from marine traffic in the Straits of Malacca and Singapore using Automatic Identification System Data *Proc. of the 8th Int. Conference on Numerical Analysis in Engineering* pp 131-142

[11] M.D Arifin, Kunihiro H, Hirata N, Ouchi K 2016 Development of marine logistic database using automatic identification system (AIS) and statistical data *Proc. of the 30th Asian-Pacific Technical Exchange and Advisory Meeting on Marine Structures* pp 395-402

[12] Kai I, Kunihiro H, Hirata N, M.D Arifin 2018 Development of the basic planning support system for tanker using marine logistics big data *Proc. of the Japan Society of Naval Architects and Ocean Engineers*

[13] Kunihiro H, Hirata N, Kai I, Dimas A F M, M.D Arifin 2019 Development of ship basic planning support system using marine logistics big data and its application to ship basic planning *Practical Design of Ships and Other Floating Structures: Proc. of the 14th International Symposium PRADS pp* 283-307

[14] M.D Arifin, Kunihiro H, Hirata N, Yuki K 2017 A study on the support system of ship basic planning by using marine logistics big data *Proc. of the International Conference on Computer Application in Shipbuilding* **1** pp 61-68

[15] M.D Arifin, Fanny O 2021 Exploiting marine BD to develop MLDB and its application to ship basic planning support *International Journal of Marine Engineering Innovation and Research* **6** no **4** pp 259-266.