J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 7063--7082

Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Development of Costing and Budget Control Strategy for Shipbuilding Based on Machine Learning

Yuan Jin[1,✉], Chenyu Zhang[1]

[1] *School of Economics and Management, Shanghai Maritime University, Shanghai, 201306, China*

ABSTRACT
Due to the complexity of the ship product structure and process, long production cycle and other factors, ship enterprises are plagued by the problem of profitability. Strengthening cost prediction and budget control is a very important means for ship enterprises to improve their profit margins. By analyzing the cost structure of shipbuilding, this paper proposes a rolling forecast model of shipbuilding cost based on long and short-term memory neural network (LSTM) as the estimation method of shipbuilding cost. Meanwhile, the traditional earned value method and target cost method are combined to sort out the shipbuilding cost control process and prepare the cost control plan as the control strategy of shipbuilding cost. Then we take the manufacturing data of a shipyard as the experimental object, use this paper's model for data mining, compare the data performance of this paper's model with similar algorithms, and verify the feasibility of this paper's model. Finally, the model of this paper is applied to real cases. In the comparison of the estimation results between this paper's model and the commonly used algorithms, the average error of cost estimation of this paper's model is ±4.95%, which is better than the average error of the commonly used algorithms. The superior accuracy of this paper's model in shipbuilding cost estimation is verified.

Keywords: LSTM model, shipbuilding cost, earned value method, target cost method

1. Introduction

For a relatively long period of time, the world shipping market continued to be in the doldrums, with the bottoming out of the freight indices of the three major shipping markets [1-2]. Although the delayed delivery of newbuildings, withdrawal of orders and massive dismantling of old ships have effectively contributed to the net growth of capacity, people are still cautious about whether this means the beginning of a full recovery of the shipping market due to the low demand [3-4]. In addition, the shipping industry itself is a high-risk industry, affected by many external factors such as the market, legal policy, political and economic environment, which will lead to the shipping

✉ Corresponding author.
E-mail address: yjin@shmtu.edu.cn (Y. Jin).
Received 05 February 2024; Revised 12 April 2024; Accepted 05 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-393

enterprises in the financing process will face a number of risks with complexity and unpredictability. Moreover, as most shipowners operate below the break-even point for a long time, the capital risk of the shipping industry is gradually erupting in full force, and it is common to see news of shipping enterprises suffering serious losses, being delisted, and declaring bankruptcy [5-6]. Most of the ship enterprises are faced with a lack of business start, the delivery of ships is increasingly difficult, hand-held orders to increase the risk of a substantial decline in corporate profitability, the industry is not stable development of many challenges [7-8].

However, with the advancement of the globalization process, international trade shows a high-speed growth trend, the rapid development of the shipping industry, which makes the global demand for ships continue to rise [9]. Ships as the infrastructure of international trade, its capacity and efficiency enhancement is crucial for the development of trade. The global ship market is expected to maintain stable growth in the coming years. Moreover, the Asian region is very important for the pull of ship demand due to the world's largest manufacturing base and consumer market [10]. In recent years, the shipbuilding industry in China, South Korea, Japan and other Asian countries has developed rapidly, becoming an important base for global shipbuilding and repair [11-13]. With the continuous expansion and upgrading of the Asian ship market, the market prospect of the ship industry in the region is even broader. At the same time, with the continuous progress of science and technology, the replacement cycle of ships is gradually shortened, and the application of new technologies and materials not only improves the safety and energy-saving of ships, but also improves the transportation efficiency of ships, and at the same time has a significant effect on the safety and efficiency of the shipbuilding process [14-17]. Therefore, the demand of ship enterprises for updating and upgrading existing ships to maintain competitiveness is becoming more and more urgent.

As a kind of industrialized production body, the process of ship construction involves a very complex technical process and the synergistic cooperation of multiple combined components, including design, material procurement, manufacturing, testing and other links, of which costing and budget control is a crucial link [18]. Costing and budget control of shipbuilding is not only related to the economic benefits of the enterprise, but also related to the progress and quality of manufacturing [19-20]. Shipbuilding industry is highly competitive, and enterprises must have lower cost and price competitiveness in order to be invincible in the market. Effective costing and budget control can help enterprises to reduce production costs and improve the price competitiveness of products, thus improving the competitiveness of enterprises [21].

In general ship construction costing, construction materials, workers, and construction equipment are the basic cost data. And Lin and Shaw [22], in addition to the consideration of the base cost data, also include the characteristics of the ship construction specifications in the final cost calculation, for this reason, they proposed the feature-cost estimation method, specifically the construction specifications of the characteristics of the construction data and the base cost data to build a relational topology, as a way to estimate the pre-cost of the shipbuilding. However, the defects of this method are also obvious, it is difficult to accurately calculate the similarity between the detailed design information of the target ship and the materials required for the construction, which leads to the low accuracy of the final costing. The operation costing method (ABC) is a costing method with high precision and accuracy, which is widely used in various fields. And ABC can categorize costs by identifying data such as demand triggered by the job process to determine direct and indirect costs and estimate the total cost and its various costs in the job process [23].Choir and Baroroh [24] used ABC to explore the costing method for the production of aluminum ships, comparing with the traditional method, the calculation method used by ABC allows companies to calculate the apportioned manufacturing costs of different products and the calculated production costs are higher because ABC considers more comprehensive factors. In addition, Azhar and Kristiyono [25] used linear regression to calculate the cost data of ships and their constituent parts in order to estimate the cost of new ship construction.Araujo et al [26] designed a decision support system in order to scientifically and objectively evaluate the cost budgets of the Brazilian Merchant Marine Fund (MBF) for the shipyards, which manages the data of the ship construction, analyzes the factors affecting the construction cost, and analyzes the cost of the ship construction data, and analyzes the factors affecting the construction cost. That analyzes the factors affecting the construction costs and evaluates the total cost of shipbuilding projects in shipyards, and then makes the construction budget. However, with the development and transition of energy sources, the ship construction process and materials are transformed, requiring a new approach to costing and budget control. Machine learning is widely used in the field of forecasting to cope with dynamic changes in manufacturing technology, construction materials and equipment, etc., and to realize scientific and objective cost management and financial expenditure and budget control in enterprises [27].

In the estimation method of shipbuilding cost, this paper analyzes the cost composition characteristics and estimation method of shipbuilding, and establishes a rolling prediction model of shipbuilding cost using LSTM model. In terms of shipbuilding budget control strategy, this paper introduces the target cost method and uses the traditional earned value method to make up for its lack of refinement, and establishes the shipbuilding cost control plan by combining the shipbuilding production mode and construction process. Given that shipbuilding cost estimation requires a certain amount of historical data to support, this paper collects the manufacturing data of a shipyard in the past ten years, uses the model in this paper to carry out data mining, and compares the estimation results of the model in this paper with those of similar algorithms. The practical application of the shipbuilding cost budget control strategy based on this paper's model is carried out to analyze the practicality of this paper's model and strategy.

2. Shipbuilding budget control based on LSTM modeling

In this chapter, by analyzing the main cost components and cost estimation methods of shipbuilding, the LSTM model is chosen as the calculation and prediction model of shipbuilding cost. It also combines the earned value method and the target cost method to launch the research on the method of shipbuilding budget control.

2.1. Components and estimation of ship costs

Although due to the different functions of various types of ships, resulting in various types of ship production cost composition is different, but in general, the cost of the ship can be roughly divided into the cost of marine equipment, raw material costs and labor costs of three major parts, the ratio of the three is about 2:1:1, of which the cost of marine equipment is an important part of the cost of shipbuilding, the proportion of the total cost of about half, especially in the main engine of the ship and auxiliary power equipment, deck machinery and equipment, etc., the proportion is mainly half of the total cost of marine equipment. The proportion of deck machinery and equipment is dominant, accounting for half of the total cost of ship equipment. In general, shipbuilding costs mainly include raw material costs, ancillary equipment costs, self-product costs, generics costs, labor costs, production of special fees and taxes.

2.1.1. Cost of raw materials. Raw materials required for ship production include ship steel, welding materials, paints, auxiliary materials, etc., which are calculated according to the following methods.

In the process of ship construction, steel is the largest consumption of raw materials, in the calculation of the design amount of steel, usually using the main scale of the hull of the estimation method, the estimation formula as formula (1):

$$W = k1 * (L * B * H) \tag{1}$$

where W is the total weight of steel used in the production process of the ship in tons. $k1$ is the steel consumption coefficient of the ship as shown in Table 1. L is the length of the ship type in meters. B is the width of the ship type in meters. H is the hull depth in meters.

Table 1. Value of steel coefficient k1

Ship type and type	K1 Value			
	Below 10,000 tons	10,000 to 20,000 tons	20,000 to 50,000 tons	More than 50,000 tons
Oil carrier	0.144~0.147	0.143~0.146	0.142~0.144	0.139~0.143
Bulk cargo ship	0.125~0.131	0.123~0.131	0.123~0.131	0.119~0.126
Multipurpose cargo ship	0.123~0.125	0.121~0.123	0.119~0.127	0.119~0.126
Container ship	0.124~0.126	0.123~0.125	0.122~0.123	0.119~0.121

In the actual construction of the ship building process, due to the specific requirements of the various process links are not the same, the size of the steel structural components are different as well as the supply of steel of different sizes and other issues, the production process of the ship steel is difficult to achieve the full utilization, so in the actual calculation of the ship construction of steel costs, but also to take into account the actual steel utilization of the various types of ships such as formula (2):

$$G = g / m \tag{2}$$

where G is the gross weight of steel used in the production process of the ship in tons. g is the net weight of steel used in the ship's hull. m is the steel utilization rate. Related research shows that the total utilization rate of steel in the ship manufacturing process is 85%~90%.

Welding materials are composed of electrodes, wires, fluxes and brazing materials, etc. The cost of welding materials is a key part of the total cost of ship production. The most important reasons affecting the consumption of welding consumables are the content of molten metal of welding consumables and various process losses during welding. The calculation of ship welding consumptions is based on the overall steel plate consumption, which is expressed as equation (3):

$$F = G * k2 * n \tag{3}$$

where F is the cost of welding material, $k2$ is the welding consumptive factor, G is the total weight of steel (tons), and n is the unit price of welding material. The steel consumption factor is shown in Table 2.

Table 2. Steel consumption coefficient

Ship type	Total steel consumption (tons)	Welding material consumption coefficient $k2$
River cargo ships and oil tankers	Below 50	0.026
	50~200	0.025
	200~500	0.024
	500~1000	0.023
	More than 1000	0.022
Coastal cargo ships and oil tankers	Below 100	0.027
	100~500	0.026
	500~1000	0.025
	1000~5000	0.024
	More than 5000	0.023
Ocean cargo ships and cruise ships, container ships	Below 5000	0.0241
	5000~10000	0.0236
	10000~15000	0.0231
	15000~20000	0.0226
	More than 20000	0.0201

Painting materials mainly refer to paints and other anticorrosive agents, the cost of which is also an important component of shipbuilding costs, and whose net and actual consumption is determined by equation (4):

$$W = L * (B + H) * k3 \tag{4}$$

where W is the net weight of paint (tons) and $k3$ is the weight of paint per square meter.

The actual consumption of paint during ship production is shown in equation (5):

$$w = W * (i / k4) \tag{5}$$

where w for the actual consumption of paint in the ship production process, i for the corresponding paint dilution coefficient, combined with the characteristics of the painting process of different ship types, the value of the coefficient is generally 1.15-1.2, $k4$ for the paint process consumption coefficients, the value of Table 3.

Table 3. Paint weight per square meter for different ship types

L*(B+H)	$k3$	L*(B+H)	$k4$
<4000	0.0096	14000~16000	0.0085
4000~6000	0.0094	16000~17000	0.0084
6000~8000	0.0091	17000~18000	0.0083
8000~10000	0.0089	More than 18000	0.0082
10000~14000	0.0087		

The cost of auxiliary materials and other materials also accounts for a certain proportion of the cost of hull materials. Auxiliary materials are substances that are not directly involved in the composition of the hull in the ship manufacturing process, such as carbon dioxide, oxygen, acetylene, high-pressure gas, high-pressure water, etc., and the cost of these substances can be estimated based on the total mass of ship steel combined with the corresponding empirical judgment of formula (6):

$$F = W * P * k5 \tag{6}$$

where F represents the cost of auxiliary materials, W represents the total weight of steel on board, P represents the average unit price of steel, and $k5$ represents the consumption factor of auxiliary materials, which, according to industry experience, can be in the range of 0.04-0.05. The cost of other materials, including timber and non-ferrous metals, etc., can be valued at 2-3% of the total material cost.

2.1.2. Equipment costs. Ship equipment consists of three main types: engine equipment, electrical equipment, and outfitting equipment. There are many types and quantities of these equipment, and 20% of the main equipment accounts for almost 80% of the total value of the equipment. Since the cost of equipment has little to do with the production process, the cost of equipment can be determined by requesting quotations, or it can be valued according to industry experience.

2.1.3. Labor costs. Labor costs are determined based on the number of shipbuilding man-hours see equation (7):

$$F = H * n \tag{7}$$

where F is the cost of labor, H is the number of shipbuilding man-hours, and n is the unit price of man-hours.

2.1.4. Production-specific costs. Shipbuilding special costs belong to the fixed components of ship production costs, mainly design costs, related policy costs, equipment costs and other necessary costs related to production, this part of the cost due to the ship production process is not very relevant to the link with a large number of items and accounted for a small proportion of the cost of shipbuilding production costs are usually taken into account in the process of calculating the cost of ship production estimates of the fixed costs.

2.2. Rolling forecast model of shipbuilding costs

Rolling forecasting is a problem of solving regression using time series data, this problem usually uses autoregressive sliding average model (ARIMA), but ARIMA model is usually used for the prediction of a single time series, while the shipbuilding material cost prediction uses multiple sets of data on the material consumption of the shipbuilding process, so it is not applicable to the study in this paper.LSTM is an excellent deep learning model, which is able to automatically LSTM is an excellent deep learning model, which can automatically capture the relationship between complex time series data, and is not only applicable to various types of time series data, but also can effectively deal with long-term dependency relationships.

Therefore, this paper chooses to use LSTM to construct a rolling prediction model for shipbuilding cost according to the characteristics of shipbuilding cost data with multiple groups and time series.

LSTM is a recurrent neural network (RNN) architecture for processing sequence data, which allows the model to selectively forget or store information, and its basic cell of the network is shown in Fig. 1.This cell state is continuously passed throughout the sequence, thus allowing the model to decide which information should be remembered or forgotten according to different parts of the sequence.

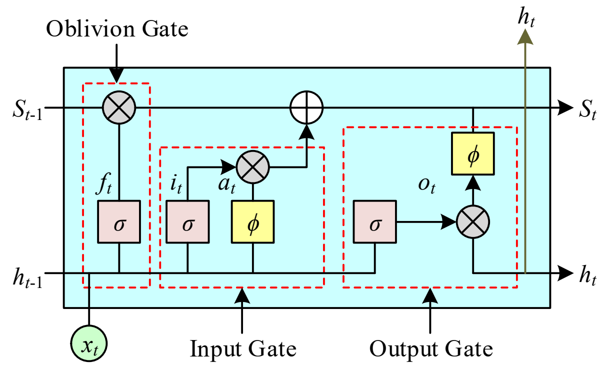

Fig. 1. LSTM Network Basic Unit

The basic unit of LSTM network contains three gates: forgetting gate, input gate and output gate. Among them, the forgetting gate can automatically migrate the current acquired knowledge to the next new moment, through the input vector x_t, the state memory unit S_{t-1} and the intermediate output h_{t-l} work together to determine the part of the forgetting gate that needs to be forgotten. The input gate filters the input vectors to remove noise and extract useful information. Input vector x_t and intermediate output h_{t-l} are transformed by the *sigmoid* and tanh functions, which together determine the portion of the input gate that needs to be retained. The output gate combines the information calculated from the previous two gates to obtain the intermediate output h_t and update the state memory cell S_t.

The formulas involved are as in Eqs. (8)-(13):

$$f_t = \sigma\left(W_{fx}x_t + W_{f_h}h_{t-1} + b_f\right) \tag{8}$$

$$i_t = \sigma\left(W_{it}x_t + W_{ih}h_{t-1} + b_i)\right) \tag{9}$$

$$g_t = \varphi\left(W_{gx}x_t + W_{gh}h_{t-1} + b_g\right) \tag{10}$$

$$o_t = \sigma\left(W_{ox}x_t + W_{oh}h_{t-1} + b_o\right) \tag{11}$$

$$S_t = g_t \odot i_t + S_{t-1} \odot f_t \tag{12}$$

$$h_t = \varphi\left(S_t\right) \odot o_t \tag{13}$$

where: f_t, i_t, g_t, o_t, h_t and S_t are the states of the oblivion gate, the input gate, the output gate, the input node, the output node, the output gate, the intermediate outputs, and the state unit, respectively. where the elements have different weights on the eigenvectors corresponding to each subspace, $W_{fx}, W_{fh}, W_{ix}, W_{ih}, W_{gx}, W_{gh}, W_{ax}$, and W_{oh} are the matrix weights of the corresponding gates multiplied by input x_t and intermediate output h_{t-l}, respectively. b_f, b_i, b_g, and b_o are the bias terms of the corresponding gates, respectively. $\odot$ denotes the multiplication of the elements of the vector by bits. σ denotes the sigmoid function variation, and ϕ denotes the tanh function variation.

In order to evaluate the performance of the model, three indexes, root mean square error (RMSE), mean absolute error (MAE), and prediction accuracy (FA), are used, and the equations for solving the three indexes are shown in Eqs. (14)-(16):

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(X_{pred,t} - X_{act,t}\right)^2} \tag{14}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|X_{pred,t} - X_{act,t}\right| \tag{15}$$

$$FA = \left(1 - \frac{\left|X_{act,t} - X_{pred,t}\right|}{X_{act,t}}\right) \times 100\% \tag{16}$$

2.3. Cost Control Strategies for Shipbuilding

The Earned Value Method is a method of project performance management, the core purpose of which is to compare the difference between the actual cost of a project and the plan, focusing mainly on the differences between projects and tasks in terms of cost, schedule, quality, etc., and to scientifically predict and control the remaining unfinished portion of the project based on the available data. The Earned Value Method has three basic indicators, Planned Cost (PV), Actual Cost (AC) and Earned Value (EV), which are used to measure the actual and planned amount of the project.

The target cost method is a consensus pricing method among suppliers and purchasers that determines the manufacturing cost of a product based on a given competitive price to ensure that the product profit is expected to be managed.

The Earned Value Method and Target Cost Method have their own strengths in controlling project costs, which are compatible with the cost control needs of shipbuilding, and this subsection combines the two methods to develop a process and plan for cost control in shipbuilding.

2.3.1. Establishment of project cost control processes. In this paper, the content of the task package is quantitatively decomposed by using the earned value method, firstly, the ship's manufacturing process and the corresponding parts of the task package are analyzed by the earned value, and the order of the veins is clarified. The progress of each part is analyzed and summarized in detail step by step: then the duration of the whole project is predicted according to the summarized situation. This can ensure that when monitoring the project in real time, the potential hidden problems affecting the project process can be accurately found and solved in time, and the calculation steps of the calculation method are shown in Table 4.

Table 4. Application flow of earned value method based on task package

Step	Concrete content
1	Set up WBS work breakdown structure and work schedule
2	Set up CBS cost breakdown structure and make cost schedule
3	Set up the cost-schedule plan and determine the detection time node
4	Refine the work plan of the task package, determine the work package on the critical path, and clarify the logic between the task packages Collect the BCWP and ACWP data of the task package of the time node and create the task Three index analysis system of packet
5	Based on the three index data of each task package of the detection time node, the project is accumulated layer by layer and the earned value analysis and prediction are carried out to find the reasons for differences, make adjustments and corrections, and finally obtain the estimated completion time and cost of the whole project
6	Draw the cost-schedule trend chart or update the project completion plan based on the analysis of the task package index system

Adopting the target cost control process of shipbuilding project combined with the earned value method and target cost method, the above process can be mainly and roughly divided into the following four steps.

1) After the overall target cost is determined, the project leader needs to use WBS to decompose the project links and finalize the critical path of the whole project execution. Secondly, strengthen the use of BOM and project management database, timely analysis of material inputs and machine parts information is centralized, and then allocated to each manufacturing process, the original non-manpower cost breakdown to the different segments.

2) Using the cost and schedule decomposition of the project after the previous step, the target cost method can not be subdivided downward, at this time, the smaller level of the task package of the indicators using the Earned Value Method to the smallest level of the target cost method to calculate the cumulative level, the level to achieve the convergence of the two, for the subsequent comparison of the difference between the analysis, back to the adjustment to lay the foundation.

3) Compare the target cost of a task and the cumulative value of earned value, calculate the difference between the two, compare this difference with the predetermined cost difference allowable range, if the difference is greater than the allowable range, then the larger cost value for retrospective analysis, find the reasons for the difference, and try to make the two costs close to each other, so as to take the final cost as the target cost of the project.

4) In the decision-making level to complete the target cost plan, strengthen the use of the earned value method of dynamic monitoring of the target cost, timely feedback information can be the first time to find deviations, if there is overspending, need to be revised in a timely manner.

2.3.2. Cost containment plan preparation. Target cost method is based on the target cost = target selling price - target profit formula to develop profits, usually with the two sides of the agreement price and market price as a benchmark, the cost target is determined, according to the WBS work decomposition structure will be the target cost decomposition to the various smaller level of the task, the process is mainly based on previous empirical data or historical databases, combined with the characteristics of the various tasks of each task for the decomposition of the operation. However, in the actual operation of the process due to resource constraints or technical limitations, the target cost in the decomposition of the decomposition is often only decomposed to the intermediate level of the task, and continue to subdivided activities are more cumbersome and complex, in order to facilitate the management and calculation, can only stay in a more macro level of operation. Therefore, only using the target cost method, the management is still sloppy, difficult to meet the control needs of shipbuilding enterprises, so it must be refined and decomposed into smaller tasks target cost.

To address the shortcomings of the target cost method, the earned value method can make up for them in terms of refined classification. The cost decomposition plan of a shipbuilding enterprise is shown in Figure 2. When the target cost method is decomposed from top to bottom to N layer, the Earned Value Method should be utilized to decompose the manufacturing tasks according to the process and manufacturing schedule, and then continue to decompose the manufacturing tasks to the smallest activity unit, that is, the task package.

First of all, the personnel involved in cost control should work together with the production staff to classify the classification level of the task package, marking the smallest task package as level X, and then start calculating the earned value data from level X according to the unit cost of the planned and actual work volume, and determine the cost and progress data by referring to the enterprise cost database: Finally, the data from all levels of the task package should be added up and deduced all the way up to level N, realizing the confluence of the Earned Value Method Decomposition and the Target Cost Method Decomposition. The Earned Value decomposition and Target Cost decomposition are converged to provide more refined data support for the subsequent analysis and assessment.

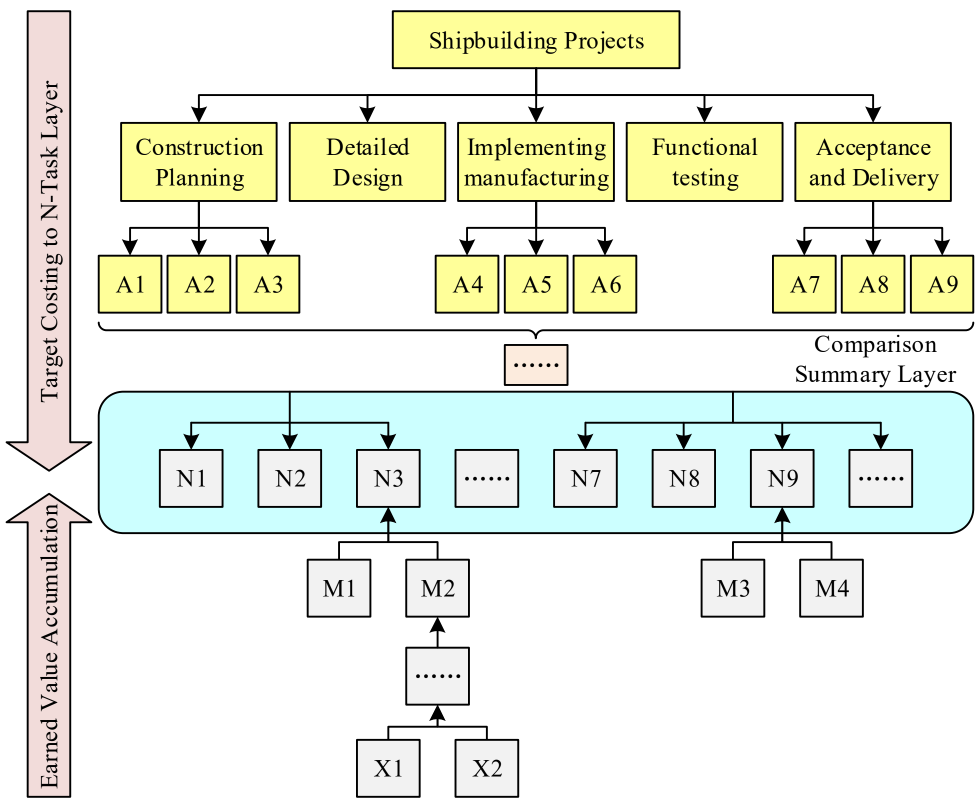

Fig. 2. Target cost of shipbuilding enterprise under earned value method

In this chapter, based on the cost composition of shipbuilding, LSTM model is selected as the cost estimation method for shipbuilding. Meanwhile, the cost control process of shipbuilding is established by combining the earned value method and target cost method as the budget control strategy of shipbuilding. By combining the cost estimation method and budget control strategy, the cost estimation and budget control method of ship manufacturing based on LSTM model is designed.

3. Data representation of the model

In this chapter, after the data preprocessing operation, based on the association rules, using this paper's ship manufacturing cost rolling prediction model respectively a shipbuilding enterprise's ship manufacturing data, combined with the knowledge of the shipbuilding industry to explain and analyze the data mining results. On the basis of data mining, the effect of ship manufacturing cost estimation is compared between this paper's model and the commonly used topable instance GA-BP network model.

3.1. Data preparation

Association rules are one of the most active methods in the field of data mining, however, the real shipbuilding data cannot be directly used for association rule mining, or the results are not satisfactory. In order to improve the effect of data mining, data preprocessing technology was born in response to the situation, which preprocesses the data before data mining to ensure the quality of the data and improve the reliability of the data mining results.

The accuracy of the data used for data mining will affect the reliability of the mining results, so it is necessary to determine the source of the original data, and the object of data mining in this paper is the ship manufacturing data of a shipyard and other related influencing factors data.

The data for ship manufacturing cost association rule mining in this chapter comes from the ship manufacturing data of 135 ships with poor cost control in a shipyard in the past ten years (from 2007 to 2016) after removing unpredictable factors such as material price, labor price, exchange rate and other factors such as marginal contribution pricing and other factors of business strategy, and then determining the raw data used for association rule mining according to the principles of influencing factor selection and the reasons for the selection, i.e. Ship's start date, shipowner, ship type, classification society, project manager, supervision company, outsourcing team, production plant and production saturation of the data contained in the nine influencing factors, such data in the existing shipbuilding enterprises in the relational database, the use of retrieval technology will be the shipbuilding data selection summary, and then selected from the influencing factors involved in the data, some of the contents of the cleanup , remove the content unrelated to mining, and carry out manufacturing database reconstruction. Part of the ship information is shown in Table 5.

Table 5. Part of data of shipbuilding

Hull number	1	2	3	...	135
Date of commence	2013-7.28	2015-3-12	2013-7-10	...	2013-9-28
Shipowner	Singapore	Germany	China	...	Brazil
Ship type	chemical tanker	Oil tanker	chemical tanker	...	bulk cargo ship
Classification society	LR	CCS	CCS	...	ABS
Project manager	F	B	J	...	E
Supervision company	d	c	b	...	e
A team of contractors	three-four-five	One-two-four	Two-five-four	...	Two-four-five
Production area	Third factory area	Second factory area	First factory area	...	Third factory area
Production saturation	0.68	0.99	0.7	...	1.06

3.2. Analysis of data mining results

In this data mining experiment, the minimum support is taken as 0.1 and the minimum confidence level is taken as 0.65.The complete experimental result output of the model in this paper is shown in Table 6.

Table 6. Association rules of impact factors of shipbuilding cost

Rule ID	Former term	Final term	Living example	Degree of confidence
1	Bulk carrier, Classification Society LR	First factory area	15	0.89
2	Production saturation [0.62,0.71]	First factory area	15	0.83
3	Detail unit B, oil tanker	First factory area	17	0.72
4	Classification Society LR	First factory area	16	0.71
5	[December, February], bulk carriers	First factory area	14	0.70
6	[June, August], bulk carriers	First factory area	16	0.69
7	Detailed Unit B, [December, February]	Second factory area	16	0.69
8	[June, August], Southeast Asian shipowners	Contractors three, four, five	15	0.66
9	Production saturation [1.02,1.11]	First factory area	15	0.66
10	[December, February], detail unit B	Contractors three, four, five	14	0.66

The association rule algorithm is applied to mine the business data of ship manufacturing process to realize the mining of the association relationship between the influencing factors. The association rules are explained as follows:

1) The shipyard is prone to cost overruns on bulk carriers. The reason is that the ship type is simple and the design and production difficulty is low, resulting in high competition in the industry and low ship price, which leads to the setting of lower target cost, thus triggering cost overruns. The shipyard should give full play to its competitive advantages in design and production by accepting and favoring high value-added ship types such as oil tankers and offshore support vessels, and selectively abandoning low value-added and low-difficulty ship types such as bulk carriers.

2) Under the premise of low production saturation, the cost of a single ship constructed in the first factory is easy to overspend. The reason is that in recent years, the first plant has been carrying out equipment upgrading and technological transformation, although the degree of automation and production efficiency has been greatly improved, but the fixed cost is large, and when the production saturation is low, it cannot give full play to its high efficiency advantage, instead, it makes the fixed cost sharing insufficient, which leads to cost overruns. The shipyard should rationally arrange the factory areas according to the production plan, giving priority to the first factory area when the production is full, and arranging to other factory areas when it is not full.

3) The construction cost of bulk carriers starting in December-February and June-August is prone to cost overruns. The reason is that the plant's mature bulk carrier start-up superseded by a 140-day shipboard cycle and a 50-day shipboard cycle, and the shipboard cycle of bulk carriers started in December-February straddles the rainy and typhoon seasons, with poor operating conditions and low efficiency, which is prone to cost overruns. The slipway cycle of bulk carriers starting in June-August straddles the Chinese New Year holiday, and the main outsourced engineering team has low attendance for 15 days each before and after the Chinese New Year, which is very likely to cause slipway delays and cost overruns. This shipyard should reasonably plan the start time according to the production rhythm characteristics of each ship type, and avoid arranging the ship types with compact slipway cycle to carry out slipway operations during the rainy season and Chinese New Year.

4) The ability and experience of the project manager should have an impact on cost control, and at the same time, there should also be an association between the project manager and the ship type, shipowner, supervisor and other factors, but the data mining did not produce the corresponding results in this regard. The reason for this is that the 135 ships collected in this sample in 10 years involve many project managers, which does not reach the minimum support and minimum confidence required by the model algorithm, and it needs to be further analyzed with the help of manual assistance to improve the correlation rules. Similarly, the outsourced labor team also has a great impact on the shipbuilding cost, but since the same ship is usually constructed by multiple

outsourced labor teams, it is difficult to dig out the impact of a single outsourced labor team on the cost.

The above results can be used as a guiding suggestion for ship enterprises before, during and after the production process, as well as a reflection of the correlation relationship between potential ship cost influencing factors in the mining results.

3.3. Performance comparison with similar algorithms

3.3.1. Comparison and analysis of cost estimation results. Comparing the cost estimation results of the model in this paper and the GA-BP network model of the topable example, the comparison of the cost estimation results of the two models is shown in Table 7 (only some typical data are listed).

Table 7. Compare the results of the two estimation methods

Customer quotation data		Extension examples and neural networks		BP neural network algorithm	
Data number	Cost / 10,000 yuan	Forecast cost/10,000 yuan	Relative error/%	Forecast cost/10,000 yuan	Relative error/%
001	380	389.14	2.41	381.1	0.29
002	165	158.63	-3.86	165.22	0.13
003	512	512.32	0.06	513.8	0.35
004	324	343.45	6	324.02	0.01
005	196	201.46	2.79	195.97	-0.02
006	1024	1055.63	3.09	1025.34	0.13

In order to react to the prediction error, the performance of each model in this paper is measured by the average absolute error as a measurement comparison parameter, and the average error of the model in this paper is ±4.95% for each 30 sets of validation data, and the estimated value of 15 sets of data is within 1.0% of the actual error rate, and the estimated value of the error of 30 sets of data is within 10% of the actual error rate. The average results of the topable example GA-BP network model gave an average error ratio of ±6.65%, with 25 sets of data within 3.5% of the actual error rate and 30 sets of data within 15% of the estimated error value. It is clear that the model in this paper not only has less localized errors in improving the feature parameters, but also improves the prediction accuracy. As can be seen from the table, the model cost estimation method of this paper is better than the cost estimation method of the topable instance GA-BP network model with less error, which can be used as the basis for ship manufacturing cost decision-making.

3.3.2. Comparison of calculation results of different methods. Using the model of this paper, regression model, price software for a company to produce 50 ship composite parts of the manufacturing cost of measurement, and its error with the actual cost of the statistical situation is shown in Figure 3. It can be seen that the model of this paper on the existing data has a good fitting effect. Meanwhile, in the training process, the overfitting situation is prevented to a certain extent by the grouping method.

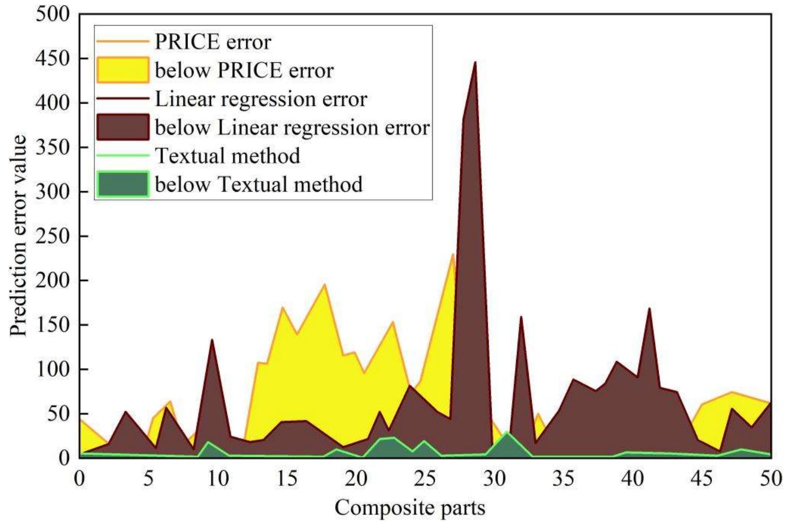

Fig. 3. Error comparison of three models

4. Practical application of shipbuilding budget control strategies

This chapter cites an example of a 61KBC ship construction project, and by using the LSTM model-based shipbuilding cost control method proposed in the previous section, decomposes the cost plan of this project, tracks and counts the actual costs at all reported time points, and launches the application analysis.

4.1. Statistical project data

From the beginning of the signing of the 61KBC shipbuilding contract to the final delivery and acceptance of the ship by Company D, the collection and analysis of data on the shipbuilding progress, cost and expense of each node are carried out throughout the process, and the fast, efficient and accurate acquisition of the required cost and progress data at each node is the prerequisite and guarantee for applying the model and strategy of this paper to the 61KBC shipbuilding project. Only by obtaining relatively accurate and reliable data, and analyzing and summarizing them, can we effectively assist the decision makers of the shipbuilding project to make the correct decisions, and then lay the foundation for the efficient operation of ship production and construction. In this paper, we set up a unified and standardized calculation method to design the BN segment construction node plan according to the WBS work decomposition of the 61KBC shipbuilding project in the previous paper, which includes the starting and stopping time of each node, the work cycle and other information. There are nine major nodes, respectively, for the start of construction (A), keel laying (B), dock shifting (C), main engine mounting (D), launching (E), re-checking (F), sea trial (G), delivery of the ship (H), the preparation of the cost of each plan corresponding to the detection of nodes is shown in Table 8, and in summary, the calculation of the cost of each plan corresponding to the cost of the nodes of the details, for the subsequent analysis of the earned value of the Y company to provide data to support (the source of the data by the procurement department), Human Resources Department, Finance Department, and Manufacturing Headquarters to provide support).

Table 8. Cost of each node

Macronode	Time	As a percentage of total costs(%)	Cost completed (10,000Yuan)
A	2021.01.22	0%	0
B	2021.05.09	21%	1350
C	2021.08.04	76%	4940
D	2021.08.15	81%	5265
E	2021.09.16	88%	5790
F	2021.11.09	97.5%	6274
G	2021.12.06	98%	6400
H	2022.01.05	100%	6500

Note: The above data comes from the manufacturing department and the sales Department

Applying the model and strategy of this paper to the 61KBC shipbuilding project of Company Y, with earned value as the main evaluation index, the following section focuses on the calculation of the required indexes for the Earned Value Management of this project. First of all, in order to fulfill the customer's requirements, it is planned to spend 65 million yuan to build the 61KBC shipbuilding project, in which the total construction period is from January 2022 to December 2022 for a total of 11 months. This project is based on natural time points and important process points to set up data dynamic monitoring time points, the beginning of the month as the data collection and analysis nodes, starting from the keel construction of each node data in order to summarize and calculate the planned cost of Table 9, the actual cost of Table 10, the amount of completion of Table 11, the earned value of Table 12. The table shows that the planned cost was 65.17 million yuan, the actual cost was 62.57 million yuan, the completion was 100%, and the earned value was 65.1 million yuan.

Table 9. Ship building program costs(10,000Yuan)

Reporting time		1	2	3	4	5	6	7	8	9
Node	Planned cost	2022.5	2022.6	2022.7	2022.8	2022.9	2022.10	2022.11	2022.12	2023.1
B	1300	450	400	450						
C	3650				1990	1660				
D	325				110	215				
E	525					480	45			
F	489						104	355	30	
G	128							25	103	
H	100								45	55
Total	6517	450	400	450	2100	2355	149	380	178	55
Accumulate		450	850	1300	3400	5755	5904	6284	6462	6517

Table 10. Actual cost of ship construction(10,000Yuan)

Reporting time		1	2	3	4	5	6	7	8	9
Node	Planned cost	2022.5	2022.6	2022.7	2022.8	2022.9	2022.10	2022.11	2022.12	2023.1
B	1285	450	438	397						
C	3500				1975	1525				
D	303				99	204				
E	495					460	35			
F	476						101	345	30	
G	115							15	100	
H	83								36	47
Total	6257	450	438	5369	2074	2189	136	360	166	47
Accumulate		450	888	1285	3359	5548	5684	6044	6210	6257

Table 11. Shipbuilding completions

Reporting time		1	2	3	4	5	6	7	8	9
Node	Proportion	2022.5	2022.6	2022.7	2022.8	2022.9	2022.10	2022.11	2022.12	2023.1
B	19.6%	7.3%	7.1%	5.2%						
C	57.5%				31.5%	26%				
D	4%				1.7%	2.3%				
E	7%					6.5%	0.5%			
F	6.5%						1.3%	4.1%	1.1%	
G	3.2%							0.5%	2.7%	
H	2.2%								0.9%	1.3%
Total	100%	7.3%	7.1%	5.2%	33.2%	34.8%	1.8%	4.6%	4.7%	1.3%

Table 12. Shipbuilding earned value(10,000Yuan)

Reporting time		1	2	3	4	5	6	7	8	9
Node	Earned value	2022.5	2022.6	2022.7	2022.8	2022.9	2022.10	2022.11	2022.12	2023.1
B	1269	469	455	345						
C	3710				2015	1695				
D	330				115	215				
E	490					445	45			
F	385						92	270	23	
G	190							40	150	
H	136								54	82
Total	6510	469	455	345	2130	2355	137	310	227	82
Accumulate		469	924	1269	3399	5754	5891	6201	6428	6510

4.2. Determining the time point for testing and collecting data

The shorter the time between inspections, the better the control of the project, but this tends to result in a higher investment of resources, so when determining the time of inspection, the first element to consider is the entire duration of the work, and ultimately determine the time interval. At the same time, based on the consideration of the time interval, it is also necessary to combine the important construction time points and the end of the quarter to be determined. Combining the total duration of the 1503 shipbuilding project, important construction points, and payment points, the final test points determined in this case are “Segmental start (S1), start of piggybacking (S2), end of the first quarter of 2022 (S3), end of the third quarter of 2022 (S4), start of launching (S5), end of quay operation (S6), end of joint commissioning (S7), End of sea trial (S8)” totaling eight testing time points, and the specific values of the first three testing time points are given in Table 13.

Table 13. The parameter of earned value of the 1503 ship at the time of inspection (10,000Yuan)

Detection time point	Time	Planned Value(PV)	Earned Value(EV)	Actual Cost(AC)
S1	2021/3/25	158.00	158.00	163.62
S2	2021/10/24	2029.78	2024.08	2049.88
S3	2022/1/31	4189.61	4182.91	4251.25
S4	2022/9/30	8765.78		
S5	2022/12/29	10642.03		
S6	2023/4/18	12332.86		
S7	2023/10/3	13611.28		
S8	2024/2/16	14992.31		

4.3. Application analysis

4.3.1. Collection and computational analysis of basic parameters. Assuming that construction is currently underway to the third inspection point, Table 14 gives an analysis of the relevant statistical earned values for Bulk Carrier 1503 on January 31, 2022, based on the summary of relevant data collection.

Table 14. Earned value analysis of 1401 ship construction project(2016/1/31) (10,000Yuan)

Number	Mission name	Total planned value	Planned value	Earned Value	Actual Cost	Schedule variance	Cost variance
1	Construction planning	103.00	103.00	103.00	100.00	0.00	3.00
2	supplies purchasing	32.00	32.00	32.00	40.53	0.00	-8.53
3	Sectional construction	5341.15	2356.00	2353.11	2382.93	-3.87	-26.93
4	Bottom section	1716.85	1716.85	1714.85	1730.48	-2.00	-13.63
5	Section 902	641.00	641.00	641.00	777.00	-3.26	-136.00
6	Bulkhead section	694.15	615.77	613.85	625.84	-1.93	-10.07
7	Section 131	338.62	260.25	258.33	256.50	-1.92	3.75
8	Side section	1167.40	24.35	24.45	26.50	0.05	-2.15
9	Section 132	572.23	24.35	24.45	86.50	0.05	-62.15
10	Section 135	595.17					
11	deck panel	446.60					
12	three-dimensional unit	356.48					
13	Building section	587.78					
14	Bottom lifting	384.92	303.67	303.22	304.70	-0.45	-1.03
15	Equipment entry	3688.15	2370.96	2368.00	2400.00	-2.36	-29.04
16	Dock action	1691.83					
17	joint debugging	1278.44					
18	Trial trip	1381.041					
19	delivery of vessel	135.00					
20	Total	15025.32	4189.65	4182.95	4251.25	-6.69	-61.6

If the traditional earned value method is used, it only focuses on the project as a whole and can only know what the cost overruns were on January 31, 2016, when the overall project schedule for the 1503 was slightly delayed, i.e., the last row of the total in Table 13, are.

The control strategy using the fusion of the Earned Value Approach and the Target Cost Approach is to analyze the overall multi-tiered sub-projects to know which specific sub-projects are responsible for the project's eventual cost overruns. Firstly, the first level of the project is analyzed, from the data in the table, the delay in the schedule is due to the delay in the two phases of segment construction and dock loading, in terms of cost, the construction planning sub-project is a slight saving, but the material procurement overspent by 85,300 yuan, and the segment construction and dock loading overspent by more, which are 269,300 yuan and 328,900 yuan respectively. Secondly, analyzing the second level, taking the sub-project of segment construction as an example, the overall overrun is mainly due to the bottom segment and bulkhead segment second level projects. On the basis of the second level and then analyzing the third level projects, it is found that the main overspending of the sub-projects of segmental construction comes from 902 sub-section and bulkhead 132 sub-section, so it is necessary to focus on checking the reasons for overspending of these two sub-projects, pay attention to the cost control of the subsequent links of these two sub-projects, and strictly control the subsequent cost expenditures.

4.3.2. Cost estimation and analysis. In order to clearly compare the accuracy of the traditional earned value method and the control strategy estimation in this paper, Table 14 gives the results of the comparison of the Estimated Completion Cost (EAC) of the project completion cost estimated by the two methods. Since the change point is at week 15, the data from week 15 and onwards are selected for comparison. Table 15 contains mainly the EAC at each weekly point in time from week 15 to week 30, and the degree of deviation (PE(i), where i denotes the week) between the EAC and the actual project completion cost (CAC).

Table 15. Comparison of evaluation indicators of EAC estimation after improvement

Week	Traditional earned value method		Textual control strategy	
	EAC-i(ten thousand yuan)	PE-i	EAC-i(ten thousand yuan)	PE-i
15	1284.86	-19.7%	1387.45	-13.28%
16	1325.87	-17.13%	1391.8	-13.01%
17	1366.84	-14.57%	1406.85	-12.07%
18	1376.26	-13.98%	1408.25	-11.98%
19	1382.89	-13.57%	1417.89	-11.38%
20	1384.05	-13.5%	1419.04	-11.31%
21	1424.77	-10.95%	1420.92	-11.19%
22	1442.86	-9.82%	1470.66	-8.08%
23	1454	-9.13%	1480.79	-7.45%
24	1454.55	-9.09%	1482.21	-7.36%
25	1475.95	-7.75%	1494.91	-6.57%
26	1484.84	-7.2%	1500.27	-6.23%
27	1502.23	-6.11%	1530.07	-4.37%
28	1541.66	-3.65%	1532.04	-4.25%
29	1555.21	-2.8%	1577.84	-1.39%
30	1562.67	-2.33%	1595.96	-0.25%

From Table 15, the closer the estimation results of the traditional earned value method and the control strategy of this paper are to the final actual completion cost, the smaller the error with the actual cost is. The estimation deviation of the traditional Earned Value Method ranges from -2.33% to -19.7%, while the estimation deviation of this paper's control strategy ranges from -0.25% to -13.28%, which is better than the estimation results of the traditional Earned Value Method. The use of the control strategy in this paper improves the accuracy of the completion cost estimation, increases the credibility of the comparison with the target cost, and is conducive to the dynamic control of ship construction cost.

5. Conclusion

This paper is based on the whole process of shipbuilding, combining the cost characteristics of shipbuilding, designing shipbuilding cost prediction model and shipbuilding budget control strategy. Specific research conclusions are as follows:

1) Aiming at the complexity of shipbuilding cost prediction, this paper selects the LSTM model, which can automatically capture complex time series data and mine hidden information, to build a ship cost prediction model. In the performance comparison experiment of model data processing, the average error of this paper's model for each 30 groups of validation data is ±4.95%, which is smaller than the average error of commonly used algorithms, verifying the data generalization ability of this paper's model and the accuracy of shipbuilding cost prediction.

2) For the budget control of shipbuilding cost, this paper is based on the traditional earned value method and target cost method, combing,

Compile the shipbuilding cost control process and plan. In the actual case application of this paper's model and budget control strategy, the deviation of cost estimation is within -0.25% to -13.28%, with high accuracy. It shows that the prediction model and budget control strategy based on machine learning (LSTM) in this paper can help shipbuilding enterprises control costs more scientifically and effectively improve the market competitiveness of shipbuilding enterprises.

References

[1] QI, B., DU, L., WANG, Y. Y., PAN, F., & SHEN, S. W. (2020). Review of World Shipping Market in 2019 and Prospects in 2020. Ship & Boat, 31(01), 1.

[2] Kim, J. H. (2019). Studies on supply and demand paradox in shipping market. The Journal of Industrial Distribution & Business, 10(1), 19-27.

[3] Laurent, D., Lee, C., & Spieth, J. (2021). Shipbuilding policy and market developments in selected economies. OECD Science, Technology and Industry Policy Papers.

[4] Sibilia, E. A. (2019). Oceanic accumulation: Geographies of speculation, overproduction, and crisis in the global shipping economy. Environment and Planning A: Economy and Space, 51(2), 467-486.

[5] Alexandridis, G., Kavussanos, M. G., Kim, C. Y., Tsouknidis, D. A., & Visvikis, I. D. (2018). A survey of shipping finance research: Setting the future research agenda. Transportation Research Part E: Logistics and Transportation Review, 115, 164-212.

[6] Syriopoulos, T., Tsatsaronis, M., XYNOU, H., & VAGIANOS, E. (2018, April). Capital Structure Determinants According to Risk Periods in Shipping Companies. In International Conference on Business & Economics of the Hellenic Open University (Vol. 17, p. 2023).

[7] Park, S., & Kim, P. (2021). Operational performance evaluation of Korean ship parts manufacturing industry using dynamic network SBM model. Sustainability, 13(23), 13127.

[8] Guan, H., Wang, Y., Dong, L., & Zhao, A. (2022). Efficiency Decomposition Analysis of the Marine Ship Industry Chain Based on Three-Stage Super-Efficiency SBM Model—Evidence from Chinese A-Share-Listed Companies. Sustainability, 14(19), 12155.

[9] Naghash, H., Schott, D., & Pruyn, J. (2024). Shifting waves of shipping: a review on global shipping projections and methodologies. Journal of Shipping and Trade, 9(1), 1-43.

[10] Felipe, J. (2018). Asia's industrial transformation: the role of manufacturing and global value chains (Part 1). Asian Development Bank Economics Working Paper Series, (549).

[11] Hossain, K. A., Zakaria, N. M. G., & Sarkar, M. A. R. (2017). SWOT analysis of China shipbuilding industry by third eyes. Procedia Engineering, 194, 241-246.

[12] Brun, L., & Frederick, S. (2017). Korea and the Shipbuilding Global Value Chain. Korea in Global Value Chains: Pathways for Industrial Transformation, Durham, NC: Duke GVC Center, 4-1.

[13] Totten, G. O. (2019). The reconstruction of the Japanese shipbuilding industry. In Japan and the New Ocean Regime (pp. 130-172). Routledge.

[14] Huh, J. H., Koh, T., & Seo, K. (2018). Design of a shipboard outside communication network and its testbed using PLC: for safety management during the ship building process. Processes, 6(6), 67.

[15] Bhardwaj, S., Bhattacharya, S., Tang, L., & Howell, K. E. (2019). Technology introduction on ships: The tension between safety and economic rationality. Safety science, 115, 329-338.

[16] Hüffmeier, J., & Johanson, M. (2021). State-of-the-art methods to improve energy efficiency of ships. Journal of Marine Science and Engineering, 9(4), 447.

[17] Jebbor, I., Benmamoun, Z., & Hachimi, H. (2023). Optimizing Manufacturing Cycles to Improve Production: Application in the Traditional Shipyard Industry. Processes, 11(11), 3136.

[18] Chen, G., Fang, K., Zhang, X., & Ru, H. (2023). Research on Knowledge Management Practices for the Whole Process of Ship Design and Construction. International Core Journal of Engineering, 9(12), 45-53.

[19] Rum, M., & Nurjalal, A. M. (2024, December). Savings on the Cost of Production of Phinisi Ships in After-Sales Condition Through Quality Cost Budgeting. In IECON: International Economics and Business Conference (Vol. 2, No. 1, pp. 380-390).

[20] Bruce, G., & Bruce, G. (2021). Ship Project Strategy. Shipbuilding Management, 63-74.

[21] Lee, J. S. (2017). A Study on International Business Competitiveness in Ship Industry. The Journal of Economics, Marketing and Management, 5(3), 30-38.

[22] Lin, C. K., & Shaw, H. J. (2017). Feature-based estimation of preliminary costs in shipbuilding. Ocean Engineering, 144, 305-319.

[23] Wahidi, S. I., Virmansyah, V. M., & Pribadi, T. W. (2021). Study on implementation of activity-based costing (ABC) system on determination of indirect costs in ship production. Kapal: Jurnal Ilmu Pengetahuan dan Teknologi Kelautan, 18(1), 1-7.

[24] Choir, I., & Baroroh, I. (2023). Activity-Based Costing Method Application Activity in Determining the Production Cost of Aluminum Ships in the Small Shipyard Industry. BAITA Engineering: Journal of Naval Architecture and Marine Engineering, 1(1), 28-35.

[25] Azhar, A., & Kristiyono, T. A. (2022). Method for Estimating Cost of New Ship Building with Linear Regression. In IOP Conference Series: Earth and Environmental Science (Vol. 972, No. 1, p. 012014). IOP Publishing.

[26] Araujo, A. O., Rezende Filho, M., & Pires, F. C. (2017). A ship construction cost data management system. Marine systems & ocean technology, 3(12), 93-103.

[27] Pham, T. Q. D., Le-Hong, T., & Tran, X. V. (2023). Efficient estimation and optimization of building costs using machine learning. International Journal of Construction Management, 23(5), 909-921.