This paper considers the task to determine the basic characteristics of bulker vessels.

Available studies that define the main characteristics of ships were analyzed. The complexity of the task to choose a prototype in the process of determining the basic characteristics of bulk carriers at the initial stages of design has been shown.

The accepted assumption implied that each ship in the sample (a ship database) was designed in such a way that some operational qualities and/or economic characteristics of a vessel are optimal in a certain sense.

To solve the task, stochastic factor analysis was used, and the method of principal components was applied, which makes it possible to identify numerical characteristics that affect the parameters of the entire system but are not clearly observed. The factors were compiled that are dimensionless sets of characteristics of sample vessels, which are, in fact, new phase orthogonal coordinates of the problem with unequal variances of the sample along the coordinate directions. Evaluating the significance of correlations of factors and vectors of the sample data makes it possible to discard factors that have a weak influence on the values of the parameters of the sample vessels. Thus, the model of the problem is simplified while preserving essential connections. The analysis of the calculation data revealed that the number of significant factors due to the use of the principal component method decreased from 7 to 3. It was determined that the ships of the same designers are indeed grouped by the value of the factor sets. In the absence of an original algorithm for determining the main characteristics of the vessel at the initial stage of design, the application of the described factor analysis procedure should help with the definition of the main factors that affect the quality of the design. Further, it is possible to apply more subtle mechanisms for project optimization according to selected economic and technical parameters. The data of the numerical calculation of integrated characteristics of the ship are given

Keywords: initial design stage, basic characteristics of bulk carriers, factor analysis, method of main components

UDC 629.5.012:303.722.29
DOI: 10.15587/1729-4061.2023.292690

DETERMINING BASIC CHARACTERISTICS OF BULK CARRIERS AT THE EARLY STAGES OF DESIGN

Oleksandr Demidiuk
Corresponding author
PhD, Assosiate Professor*
E-mail: alexanderdemidiuk@gmail.com
Michail Kosoy
PhD, Senior Lecturer*
Anastasiia Zaiets
PhD, Assosiate Professor*
*Department of Shipbuilding and Ship Repair named after Prof. Y. L. Vorobyov
Odesa National Maritime University
Mechnikova str., 34, Odesa, Ukraine, 65029

Received date 22.09.2023
Accepted date 29.11.2023
Published date 14.12.2023

How to Cite: *Demidiuk, O., Kosoy, M., Zaiets, A. (2023). Determining basic characteristics of bulk carriers at the early stages of design. Eastern-European Journal of Enterprise Technologies, 6 (1 (126)), 68–77. doi: https://doi.org/10.15587/1729-4061.2023.292690*

1. Introduction

Ship design is a complex engineering task of a high level of complexity. In the process of solving this problem, it is necessary to take into account the features of the designed vessel, possible technical, navigational, and operational limitations. The task to determine the basic characteristics of a ship at the initial stage of design is complicated by the lack of information, the amount of which in open sources is limited. One of the ways that makes it possible to obtain missing information about the ship's characteristics is the designing technique based on a prototype. The selection of a prototype or several of the most appropriate prototypes is usually carried out by an expert or a group of experts based on available experience, knowledge, and available information on the object of interest.

A common tool used by designers is the relational or empirical method based on the comparison of data from a similar built vessel. According to this method, the characteristics of the database of ships are accepted for analysis and regression formulas of the dependence of the derivative characteristics of the ship on the main ones are derived. The limitation of the method is the lack of reverse influence of the derived parameters on the basic parameters, which is fundamental for the iterative design method.

As a result, the expert's choice is based on a subjective assessment of the optimality of the selected characteristics. Attempts to find out these optimal qualities and characteristics by a simple comparison of the technical parameters of vessels, as a rule, do not lead to successful results due to the very complex structure of the mutual influence of these parameters. Therefore, the use of mathematical methods of processing statistical data is an urgent task for clarifying the relationship between the studied characteristics. This will allow designers to make informed project decisions.

2. Literature review and problem statement

The task of designing a ship can be reduced to the task of optimizing the ship's characteristics according to one or several selected criteria. Economic or integrated technical parameters can usually be used as criteria. Then, after the construction of a mathematical model of the vessel, which explicitly takes into account the specified criteria, the task of minimizing certain selected criteria and determining the characteristics of the vessel is formed. The solution to such a problem depends on the type of equations and can be performed by known numerical methods. Thus, in works [1, 2], the authors combined the solution to the optimization problem and simulation modeling to determine the characteristics of a dry cargo ship making irregular routes.

In work [3], conceptual design is considered as a synthesis of the multiattributive decision-making process. Conceptual design is considered here as a process of decision-making synthesis. In this process, a huge number of possible designs are randomly generated using adaptive Monte Carlo selection. Vessel properties are assigned using metamodeling methods. The non-dominated solutions are filtered as a set of Pareto optimal solutions that are entered into the «best» design. It is characterized by high-level technical characteristics (geometry, vessel characteristics, capital costs, operating costs, etc.). This information is significantly limited due to what constitutes the designer's intellectual property.

Research [4] implements an innovative methodology of complex design with the help of a certain number of developed software tools aimed at integrating multi-stage methods into a holistic process.

In work [5], an integrated design method is considered, which is built on the basis of the CAESES software and consists of a geometric model core with several integrated modules that cover stability, strength, power, and driving force, safety, economy, as well as a work simulation module, which allows the user to model the response in variations, geometric, structural variables of the vessel under conditions of uncertainty.

Research is being conducted in this promising direction [6, 7]: in [6], multi-criteria optimization is applied to study the extended design space using genetic algorithms to determine the characteristics of medium-tonnage container carriers. The same approach was used in work [7] for the complex design of container ships of the Post-Panamax class. It should be noted that the use of the methods described in [4–7] require considerable time spent on development, testing, and implementation in the work of design institutions.

Of course, simulation modeling is a modern and promising method for choosing the optimal ship parameters. But the simulation results depend to a large extent on the chosen mathematical model of the vessel. In order to determine a number of values of economic indicators for comparing the efficiency of ships, a justified systematic change of the parameters of the ship model is required. This method requires the implementation of a special information and calculation system. The construction of such a system requires significant time and financial costs and is inaccessible to many project organizations. Therefore, improvement of traditional design methods remains a priority.

A considerable amount of research considers obtaining sets of linear and non-linear regressions to estimate the basic dimensions, mass, and various parameters of the vessel. Thus, in [8, 9] the method of structural-parametric design using the results of one-parametric regression analysis of vessel parameters from the database is applied. The base length of the vessel is used as a regression parameter. In study [6], the carrying capacity of the vessel was used to determine the characteristics of container carriers, in [10], the number of *TEU* containers was used as a basic parameter, and in [11], deadweight, the number of *TEU* containers, and the speed of the vessel were used as basic parameters in nonlinear regression formulas. In the regression search procedure itself, the method of neural networks was used.

The above sources [6, 8–11] use regression formulas with a different number of basic parameters (depending on the study). This approach is widely used in the field of ship design, but it is not without systemic shortcomings, namely:

– instability to random emissions of sample data, which requires «smoothing» of data before processing, as a result of which the already limited information about ships is lost;

– the impossibility of establishing the statistical dependence of the basic parameters and the mutual influence of the sought values, the choice of whose values is a largely contradictory task, which is a necessary requirement of an integrated approach to design [3–5, 12];

– regression coefficients cannot be interpreted from an objective point of view.

Within each of the methods, algorithms for selecting the optimal solution from available ones obtained as a result of database processing have been developed.

To take into account the complex interdependences of the main characteristics of a ship, it is advisable to use preliminary design methods that take into account the mutual influence of the selected operational, weight, and geometric characteristics. Such methods consider the design process as a whole, due to which the iteration cycle is shortened. These methods are developed on the basis of the general method of parallel design [14]. The basic concept is that in the design process there is an exchange of knowledge and interaction of departments at all stages of the design. The accumulated experience is the property of the design company. In problems with many local extrema of the objective functions, which is the design of the ship, a good choice of the initial approximation of the parameters is significant.

Project development experience remains the intellectual property of a company. Despite the successful implementation of the methodology [14] within the development company, successful design solutions of competitors remain inaccessible because they are intellectual property. Therefore, the problem of the initial approximation of the project is still relevant.

To implement a holistic approach to design, first of all, a rational systematization of data on ships is required. That is, the expected result of processing the database should be the distribution of the sample of vessels into groups according to the sets of values of their relative parameters. The usual grouping method, the list method, is not rational. The number of combinations of sets of these parameters obeys the law $\prod_{k=1}^{N} k_n$, where n is the number of parameters, and k_n is the number of value intervals of the n-th parameter. So, for example, with 7 dimensionless parameters accepted for analysis and 4 intervals of changes in the values of each, the number of combinations is 2401. Such a number of combinations does not lend itself well to review, and there is no reason to hope for some kind of systematization.

The problematic situation in the problems of ship design is highlighted by the results of the analysis of the considered studies and implies the following. In the existing implementations of design tasks, the systematization of data on ships by sets of values of their parameters and calculations of the main characteristics of the ship on its basis is not implemented.

3. The aim and objectives of the study

The purpose of this study is to determine integrated characteristics, the values of which will be used to group ships from the sample according to certain criteria and to calculate based on the results of grouping the main characteristics of the ship. This will make it possible at the initial design stage to choose a prototype (prototypes) from the grouping of sample ships that correspond to the selected class of vessels, created by experienced designers, and have close to the desired values of combinations of the technical parameters of the ship.

To achieve the goal, the following tasks were set:

– at the first stage, as a result of the analysis of statistical data of the prototype vessels, find the factors that determine the peculiarities of the design characteristics;

– at the second stage, carry out a substantive analysis (by grouping vessels according to the values of the found factors) and interpretation of the found factor combinations;

– to develop an algorithm for using the results of the factor analysis of the ship database to determine the main characteristics of bulk carriers based on the values of the found factor combinations and the performed grouping of the sample ships.

4. The study materials and methods

The characteristics of 32 existing vessels for the transportation of mass cargo in bulk – bulk carriers – were chosen as objects of research for the analysis. By increasing deadweight, bulk carriers are divided into the following main classes: HANDYSIZE, HANDYMAX, SUPRAMAX, PANAMAX, KAMSARMAX (larger-size Panamax bulk carrier), POST-PANAMAX, and CAPESIZE.

Data from work [13] were partially used to construct a sample group. The main characteristics of bulk carriers are given in Tables 1, 2.

The study used the method of principal components. This method of factor analysis of statistical data [14–17] makes it possible to reduce the number of parameters, according to the values of which it is possible to carry out an overview grouping of vessels with similar technical and operational characteristics.

Table 1

Main characteristics of bulk carriers (part one)

No.	Ship name	L_{oa}	L_{bp}	B	H	T	D_b	D_n	dw
		m	m	m	m	m	t	t	t
1	2	3	4	5	6	7	8	9	10
1	NORD HONG KONG	179.90	171.50	28.40	14.10	10.00	41,748	9,459	32,290
2	DRAGONERA	180.00	176.63	30.00	14.70	10.10	45,009	10,396	34,613
3	HALKI	187.00	178.02	27.80	15.60	10.90	47,246	10,396	36,850
4	THALASSINI AXIA	196.00	189.00	32.26	18.60	12.50	69,919	11,311	58,923
5	ARKADIA	197.08	189.00	32.26	18.50	12.60	68,418	12,070	55,161
6	AMBER CHAPION	199.85	194.50	32.26	18.50	12.70	75,471	11,671	63,800
7	STX ARBORELLA	199.82	192.67	32.26	19.30	12.72	70,605	13,065	57,539
8	SAGE AMAZON	199.99	193.74	32.26	18.50	13.30	74,922	11,621	63,301
9	AP LIBERTAS	225.00	217.00	32.26	19.60	14.20	87,072	11,859	75,213
10	GIEWONT	229.82	222.82	32.26	20.25	14.62	93,377	13,727	79,649
11	NIKI	253.92	241.00	40.00	21.00	14.62	117,602	17,602	100,000
12	OCEAN GARNET	229.20	222.00	38.00	20.70	14.90	108,709	15,690	93,018
13	ANANGEL DAWN	249.85	240.00	43.00	21.30	15.00	134,376	20,285	114,091
14	STX FREESIA	292.00	283.00	45.00	24.80	18.22	207,383	26,647	180,736
15	K. HOPE	330.07	321.00	57.00	25.10	18.00	285,100	35,100	250,038
16	MIEDWIE	190.00	182.60	23.00	14.60	10.40	39,174	9,190	29,984
17	IVS VISCOUNT	178.80	172.00	28.00	14.80	10.65	43,757	9,058	34,699
18	SPAR LYRA	190.00	183.44	32.26	17.40	12.50	65,000	11,600	53,565
19	TAI PROGRESS	225.00	217.00	32.26	19.50	14.10	88,366	10,532	77,834
20	KOHYOHSAN	289.00	280.20	45.00	24.10	17.78	193,802	21,238	172,564
21	WORLDERA 2	189.96	182.00	32.24	16.50	11.62	59,857	10,950	48,907
22	MV MARILIA	190.50	183.68	32.01	16.40	11.85	59,280	10,640	48,640
23	BURGIA	229.00	222.00	32.26	20.25	14.62	93,288	13,885	79,403
24	AGRI GRANDE	229.00	225.32	32.26	20.00	14.45	95,374	13,408	81,966
25	BOAVISTA	229.50	222.90	38.00	20.70	14.90	108,707	15,630	93,077
26	ABY DIVA	224.90	217.00	32.26	19.50	14.14	86,824	10,228	76,596
27	CORAL TOPAZ	225.00	218.00	32.20	19.80	14.23	87,409	10,811	76,598
28	OSMARINE	224.94	217.00	32.26	19.50	14.14	86,890	10,228	76,662
29	TOPEKA TBR	225.00	217.00	32.26	19.60	14.27	87,856	12,627	75,229
30	MG SAKURA	224.95	217.00	32.20	19.15	13.84	85,857	10,460	75,397
31	HOUHENG 5	326.99	319.77	57.00	25.50	18.83	296,699	34,938	261,761
32	PACIFIC QUEEN	291.80	282.00	45.00	24.75	18.10	202,617	26,699	175,918

Table 2

Main characteristics of bulk carriers (part two)

No.	Ship name	V_S	*MCR*	T^*	K^{**}	Type	C_b^{***}
		Knots	kW	Unit	Unit	–	–
1	2	11	12	13	14	15	16
1	NORD HONG KONG	13.80	6,000	5	4	Handysize	0.804
2	DRAGONERA	14.00	7,500	5	4	Handysize	0.789
3	HALKI	14.80	7,800	5	4	Handysize	0.814
4	THALASSINI AXIA	14.60	9,960	5	4	Supramax	0.856
5	ARKADIA	14.50	11,620	5	4	Handymax	0.835
6	AMBER CHAPION	14.48	8,050	5	4	Handymax	0.877
7	STX ARBORELLA	13.90	9,960	5	4	open–hatch	0.834
8	SAGE AMAZON	14.50	8,300	5	4	Supramax	0.846
9	AP LIBERTAS	14.00	8,990	7	–	Panamax	0.822
10	GIEWONT	14.28	11,060	7	–	Kamsarmax	0.834
11	NIKI	15.00	11,180	7	-	Capesize	0.783
12	OCEAN GARNET	14.35	12,240	7	–	Large bulk	0.811
13	ANANGEL DAWN	14.70	15,820	7	–	Handycape	0.814
14	STX FREESIA	15.00	17,900	9	–	Dunkirkmax	0.838
15	K. HOPE	14.95	23,000	9	–	VLOC	0.812
16	MIEDWIE	14.00	7,800	6	3	Great Lakes	0.845
17	IVS VISCOUNT	14.00	6,500	5	4	Handysize	0.805
18	SPAR LYRA	14.00	9,480	5	4	Handymax	0.824
19	TAI PROGRESS	14.00	10,002	7	–	Panamax	0.840
20	KOHYOHSAN	14.70	14,711	9	–	Capesize /dunkirkmax	0.811
21	WORLDERA 2	14.00	7,707	5	4	Handymax	0.816
22	MV MARILIA	12.50	10,400	5	4	Handymax	0.806
23	BURGIA	14.00	11,620	7	–	Kamsarmax	0.836
24	AGRI GRANDE	14.00	9,801	7	–	Kamsarmax	0.852
25	BOAVISTA	14.30	13,569	7	–	Post Panamax	0.808
26	ABY DIVA	14.00	10,327	7	–	Panamax	0.823
27	CORAL TOPAZ	14.00	9,230	7	–	Panamax	0.821
28	OSMARINE	14.00	10,327	7	–	Panamax/self loaded	0.816
29	TOPEKA TBR	14.00	9,699	9	–	Panamax	0.817
30	MG SAKURA	14.00	9,266	7	–	Panamax	0.825
31	HOUHENG 5	15.20	18,240	9	–	Capesize	0.803
32	PACIFIC QUEEN	14.00	16,871	9	–	Capesize	0.820

*Note: * – number of holds; ** – number of cargo cranes; *** – block coefficient of a vessel*

On the way to determining the integrated characteristics, an assumption was put forward, which implies the following. When designing according to a prototype, it is advisable to use data from successful projects of third-party designers. It is assumed that each of the options is designed in such a way that some operational qualities and/or economic characteristics (currently unknown, due to the lack of information) of the vessel are optimal. Attempts to find out these optimal qualities and characteristics by a simple comparison of the technical parameters of the vessels do not lead to successful results due to the very complex structure of the mutual influence of these parameters. Thus, the research concept is based on the logical assumption that the optimality of the ship's characteristics is present in a hidden way and is contained in a combination of the ship's technical parameters (dimensionless sets) – determining factors. Analysis of the technical characteristics of the selected sample of ships makes it possible to identify these factors that significantly affect the operational qualities of the ship but are clearly not observed. Therefore, at the first stage of design, the task was set to identify these determining factors as a result of the analysis of statistical data on the characteristics of vessels of this type.

Dimensionless parameters of sampling vessels are accepted for statistical processing:

$$X_1 = \frac{L}{B}; X_2 = \frac{B}{H}; X_3 = \frac{D_b}{\rho LBT}; X_4 = \frac{D_b}{\rho LBH};$$

$$X_5 = \frac{D_b - D_n}{D_b}; X_6 = \sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}; X_7 = \frac{P}{gD_bV}.$$

Here *L, B, H, T* are the length, width, height of the side and the average draft of the vessel, respectively, m; D_b – weight water capacity of the vessel at full load, (metric tons) mt;

D_n is the weight of the empty vessel, mt; P is the total power of the main propulsion unit, kW; V – maximum speed, m/s; g=9.81 m/s^2 – gravitational constant; ρ=1024 kg/m^3 is the density of sea water. The method of factor analysis is used to determine the factors F_j that affect the value of parameters X_i in the case of their linear relationship of the following form:

$$X_i = \sum_j b_{ij} F_j,\ i = 1,n;$$

$$F_j = \sum_i a_{ji} X_i,\ j = 1,n;\ \|a_{ji}\| = \|b_{ij}\|^{-1}. \quad (1)$$

Further, the designations X_i for the technical characteristics of vessels from the sample, and F_j for the factors affecting their choice in the project of each vessel, are introduced. Based on the meaning of F_j factors, they should be sets of technical characteristics and correspond to the basic principle of dimensionality theory, i.e. $F_j = \prod_i X_i^{a_{ji}},\ j = 1,n.$ To move from the multiplicative form of the dependence of F_j on X_i to form (1), the parameters Y_i=$Ln(X_i)$ are accepted for processing instead of X_i:

$$Y_i = \sum_j b_{ij} F_j,\ i = 1,n;$$

$$F_j = \sum_i a_{ji} Y_i,\ j = 1,n;\ \|a_{ji}\| = \|b_{ij}\|^{-1}. \quad (2)$$

For the convenience of presentation, the matrix of values of parameters Y_i for a sample of length N is marked through $\|Y_{ij}\|_{n\times N}$, the matrix of factor weights through $\|b_{ij}\|_{n\times n}$, the matrix of N values of n factors through $\|F_{ij}\|_{n\times N}$. The correlation matrix of parameters Y_i=(Y_{im}), m=1,N, where N is the length of the sample, is calculated by the formula $K_Y = \frac{\|Y_{jl}\| \cdot \|Y_{jl}\|^T}{N-1} = \|b\| \cdot \frac{\|F_{jl}\| \cdot \|F_{jl}\|^T}{N-1} \|b\|^T = \|b\| \cdot K_F \|b\|^T$. Using the method of principal components assumes statistical independence of factors F_j=(F_{jm}), i.e., imposing independence restrictions on factors, it will take a diagonal form. Then the matrix K_Y will take the form $K_Y = \|b\| \cdot \begin{Vmatrix} \lambda_1 0..0 \\ 0..0\lambda_n \end{Vmatrix} \cdot \|b\|^{-1}$. In terms of statistics, the diagonal elements of the matrix $K_F = \begin{Vmatrix} \lambda_1 0..0 \\ 0..0\lambda_n \end{Vmatrix}$ are the variances of factors F_j. The ranking of the eigenvalues of the K_Y matrix by magnitude from larger $K_F = \left\| \frac{(F_j, F_k)}{N-1} \right\|$ to smaller determines the significance of the influence of each factor on the ship's parameters. Thus, the greater the eigenvalue of λ_j, i.e., the variance of the factor F_j, the more intensively it affects the value of the parameter Y_i in the statistical dependence.

The methods of statistical assessment of the significance of this influence make it possible not to take into account the factors that are weakly manifested, and thus to simplify and make the subject analysis of a sample of vessels available for inspection. However, this simplification is associated with a loss of information. If all eigenvalues take comparable values, then their entire set has to be taken into account. In order not to lose information and not to reject factors with small variance values (they correspond to small values of the eigenvalues of the K_Y matrix), it should be assumed that the system of factors F_j is normalized and has a diagonal correlation matrix with a single diagonal. Then we obtain the canonical form $K_Y = \|b_{ij}\| \cdot \|\sqrt{\lambda_j}\| \cdot \begin{Vmatrix} 10..0 \\ 0..01 \end{Vmatrix} \cdot \|\sqrt{\lambda_j}\| \cdot \|b_{ij}\|^{-1}$. Thus, $\|a_{ji}\| = \|\sqrt{\lambda}_j\| \cdot \|b_{ij}\|^{-1}$ and (2) can be used to calculate the factors.

In order to still be able to simplify the system by reducing the factors, it is necessary to calculate the correlation matrix $K_{FY} = \|k_{ij}\| = \frac{\|a_{ik}\| \cdot \|Y_{kl}\| \cdot \|Y_{kl}\|^T}{N-1}$. The assessment of the significance of the values of the correlation elements is carried out according to the Student's t-test with a probability of, for example, 5 %: $\frac{|k_{ij}\sqrt{N-2}|}{\sqrt{1-k_{ij}^2}} > 2.8$. For greater clarity, the significance results are entered in the incidence matrix $IK_{FY} = \|\gamma_{ij}\|$: if the inequality is satisfied, then the correlation k_{ij} of the parameter Y_i and the factor F_j is considered not zero and γ_{ij}=1, if it is not fulfilled, then it is considered that the parameters Y_i and the factor F_j are statistically independent and γ_{ij}=0. Therefore, a new matrix $\|\alpha_{ij}^*\| = \|\alpha_{ij}\| \cdot \|\gamma_{ij}\|$ should be adopted in expansion (2).

The found values of the factors for ships make it possible to group them. The method of this grouping consists of the following steps:

– set the intervals for changing the values of the selected factors;

– group the vessels of the sample according to the aggregate values of the factors;

– select the most numerous groups and note the intervals of the factor values for them (the number in the group indicates the unmanifest quality common to the vessels of the group);

– if m is the number of accepted factors, and n is the number of parameters, and $n>m$, then out of n unknown quantities, only (n–m) will be independent, the others are connected by relations (2) for factors with accepted values.

Independent parameters are selected based on the results of a case study, and dependent parameters, if necessary, will be determined in such a way as to ensure some optimal quality of the vessel.

5. Results of determining the integrated characteristics of a vessel

5. 1. Combinations of observed parameters that allow revealing the characteristics and qualities of a vessel

Based on the results of the factor analysis of the parameters of the sample of vessels, a matrix of coefficients $\|\alpha_{ij}^*\|$ and, accordingly, expressions for the F_j factors were obtained:

$$\|\alpha_{ij}^*\| = \begin{Vmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -0.83 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.91 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.24 \\ 0.41 & 0 & -0.11 & 0 & -0.52 & -0.3 & 0.22 \\ 0 & 0 & 0 & 0.2 & 0.22 & -0.25 & 0 \\ 0 & -0.2 & 0 & -0.07 & 0.014 & -0.25 & 0 \\ -0.04 & -0.1 & 0.09 & 0.21 & -0.07 & 0.007 & -0.008 \end{Vmatrix}; \quad (3)$$

$$F_1 = Ln\left(\left(\frac{P}{gD_bV}\right)^{-0.83}\right);\ F_2 = Ln\left(\left(\frac{P}{gD_bV}\right)^{-0.91}\right);$$

$$F_3 = Ln\left(\left(\frac{P}{gD_bV}\right)^{-0.24}\right);$$

$$F_4 = Ln\left(\begin{matrix}\left(\frac{L}{B}\right)^{0.41}\left(\frac{D_b}{\rho LBT}\right)^{-0.11}\left(\frac{D_b-D_n}{D_b}\right)^{-0.52}\times \\ \times\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{-0.3}\left(\frac{P}{gD_bV}\right)^{0.22}\end{matrix}\right);$$

$$F_5 = Ln\left(\left(\frac{D_b}{\rho LBH}\right)^{0.2}\left(\frac{D_b-D_n}{D_b}\right)^{0.22}\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{-0.25}\right);$$

$$F_6 = Ln\left(\begin{matrix}\left(\frac{B}{H}\right)^{-0.2}\left(\frac{D_b}{\rho LBH}\right)^{-0.07}\times \\ \times\left(\frac{D_b-D_n}{D_b}\right)^{0.04}\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{-0.25}\end{matrix}\right);$$

$$F_7 = Ln\left(\begin{matrix}\left(\frac{L}{B}\right)^{-0.07}\left(\frac{B}{H}\right)^{-0.1}\left(\frac{D_b}{\rho LBT}\right)^{0.09}\left(\frac{D_b}{\rho LBH}\right)^{0.21}\times \\ \times\left(\frac{D_b-D_n}{D_b}\right)^{-0.07}\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{0.007}\left(\frac{P}{gD_bV}\right)^{-0.008}\end{matrix}\right). \quad (4)$$

It can be seen from (3), (4) that with the selected minimum number of factors, the most informative will be the selection of factors F_4, F_6 and F_7. For example, geometric characteristics $\frac{L}{B}, \frac{B}{H}, \frac{D_b}{\rho LBH}$. can be determined by the parameters $\frac{P}{gD_bV}$, $\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}$ and $\frac{D_b}{\rho LBT}$ and $\frac{D_b-D_n}{D_b}$, which are defined from the technical task.

According to the values of these factors, the vessels from the sample were grouped. Subjective analysis of vessel characteristics in groups should confirm the assumption that vessels with close factor values have similar operational characteristics, regardless of different sizes and geometric coefficients.

The insufficient amount of information, especially about the operational qualities of ships (which is a consequence of commercial considerations), leads to the fact that from open sources, in addition to the type and parameters of the ship, only the year, place of construction, country, and shipyard where the ship was built can be determined. In a significant number of cases, the designer can also be identified. This information can be a good guide in the initial stages of design, as every design bureau «with a history» has developed design methods, authentic signature of the implementation of successful projects.

5. 2. Grouping of sample vessels

For further calculation, a meaningful analysis of groups of vessels was carried out in order to determine the desired parameters (characteristics). The grouping of vessels according to the most informative factors determined from expression (4) is given in Tables 3, 4.

Next, groups of vessels are considered, from the point of view of the number of vessels in them. The data on ships lacks information on the specific cargo capacity, which affects the required volume of cargo spaces. Therefore, an independent parameter $C=dw/(LBH)$ was introduced, which characterizes the ratio of deadweight dw to the cubic module LBH and is a certain approximation to the specific cargo capacity. This parameter was not used in the process of statistical processing of input data.

Table 3

Grouping of bulkers by the values of factors F_4, F_6 by their numbers in the sample

Range of factor values F_4 and F_6	Number of vessels in the group	Numbers of the group's vessels in the sample
$-1.07<F_4<-0.84$; $0.97<F_6<1.08$	1	14
$-0.60<F_4<-0.37$; $0.64<F_6<0.75$	7	4, 6, 8, 12, 13, 21, 25
$-0.60<F_4<-0.37$; $0.75<F_6<0.86$	5	11, 15, 20, 31, 32
$-0.37<F_4<-0.13$; $0.64<F_6<0.75$	16	1, 2, 3, 5, 7, 10, 16, 17, 18, 19, 22, 23, 26, 28, 29, 30
$-0.37<F_4<-0.13$; $0.75<F_6<0.86$	3	9, 24, 27

Table 4

Grouping of bulkers by the values of factors F_4, F_6, F_7 by their numbers in the sample

Factor F_4 value intervals	Factor F_6 value intervals	Factor F_7 value intervals			
		(−0.27; −0.26)	(−0.26; −0.24)	(−0.24; −0.23)	(−0.23; −0.22)
(−1.07; −0.84)	(0.97; 1.08)	–	–	–	14
	(0.64; 0.75)	–	–	–	–
	(0.75; 0.86)	–	–	–	–
(−0.60; −0.37)	(0.97; 1.08)	–	–	–	–
	(0.64; 0.75)	–	12, 13, 21, 25	4,8	6
	(0.75; 0.86)	11, 15, 31	20	32	–
(−0.37; −0.13)	(0.97; 1.08)	–	–	–	–
	(0.64; 0.75)	1, 2, 22	3, 5, 7, 17, 26, 28	10, 16, 18, 19, 23, 29, 30	–
	(0.75; 0.86)	–	27	9,24	–

From Table 4, we established value intervals of the factors that distinguish these groups. One group consists of vessels with numbers 10, 16, 18, 19, 23, 29, 30. In this group, vessels with numbers 19, 29, 30 are of the Panamax type and were built, respectively, in Taiwan (2004), in China (2000), and Japan (2006). Vessels numbered 10, 23 belong to the Kamsarmax type, which are larger in size than the Panamax type vessels. Both vessels were built in China in 2010 and have similar *C*-factor values of 1.828 and 1.826, respectively. Vessel number 18 is a Handymax type bulker built in 2005 in China. Vessel 16 should be considered separately, which was built in 2010 in China for the specific Great Lakes region and has an unusual number of holds for bulkers – 6. Also, this vessel has the highest value of the *C* coefficient – 2.045 among the vessels of this group.

The group of vessels with serial numbers 3, 5, 7, 17, 26, 28 is a group of bulkers of the Handysize (3, 17) and Panamax (26, 28) types, with the exception of number 13 – Handymax type, and numbers 7 and 17. Vessels 3 and 5 were built at Korean shipyards in 2011 and 2012, respectively. Vessels 7 and 17 are of Chinese production, built in 2012 and 2003. Vessels numbered 26 and 28 were built in 2007 and 2006 at the Japanese shipyard IMABARI SHIPBUILDING Co. LTD in Marugame. These vessels belong to the Panamax class, the feature of number 28 is the presence of self-unloading equipment. The value of the coefficient *C* for both vessels is almost the same – 1.782.

The group of vessels with serial numbers 12, 13, 21, 25 represents a group of large bulkers of various types: 12 – Large bulk, 13 – Handycape, 21 – Handymax, 25 – Post Panamax. The value of the coefficient *C* is close to 1.9. All vessels were built in Chinese shipyards in 2010–2011, with the exception of number 25, which was built in Japan in 2001, and which has a slightly higher *C* factor value of 1.980.

The group of ships with serial numbers 1, 2, 22 is a combination of a group of bulk carriers of the Handysize type (1 and 2) and a bulk carrier of the Handymax type. Vessels 1 and 2 were built in China in 2011 and have similar *C*-factor values.

The group of ships with the serial numbers 14, 15, 31 is a grouping of fairly large ships of the Dunkirkmax (14) Capsize (31) type. Number 15 is a VLOC class ore carrier similar in size. Vessels 14 and 31 were built in China in 2009 and 2017, respectively, and have similar *C*-ratio values (14:1.747 and 31:1.776). The ACP of these vessels is also similar, they each have 9 cargo holds. Vessel 15 was built in Republic of Korea in 2012.

Small (only two vessels) but interesting groupings were also considered.

The first such group contains vessels with numbers 4 and 9, which belong to the Supramax class, which is sometimes called Large Handymax. The vessels were built in Korea (2010) and China (2012), respectively. Values of the coefficient *C* are typical.

The second group contains vessels numbered 9 and 24, and belong to the Panamax and Kamsarmax types, respectively, and also have 7 holds each and a similar architectural and structural type. Built in China in 2008 and 2017, respectively.

Regarding manufacturing companies, the analysis revealed the following:

– three vessels, 3, 5, 15, were built at the enterprises of the HYUNDAI company (Republic of Korea);

– four vessels were built in Japan: 26, 28 were built by IMABARI (Japan), vessel 30 was built by UNIVERSAL SHIPBUILDING CORPORATION (Japan), vessel 21 was built by TSUNEISHI SHIPBUILDING (Japan);

– other vessels, with few exceptions, were built by Chinese shipbuilding companies.

Thus, grouping by the values of the obtained factors made it possible to somewhat systematize disparate information, at least in relation to the structural types of ships, some of their characteristics, and construction companies. Focusing on the success of the selected vessel, the characteristics of the vessel determined by the values of the selected factors should be selected during design.

5. 3. Determining the main characteristics of a bulk carrier based on the results of a factor analysis of a sample of vessels

The calculation of the main characteristics of the bulker based on the results of the grouping is carried out according to the algorithm proposed below and consists of the following steps. According to one of the tables of groups of vessels by factor values, it is necessary to select a group of vessels. A preliminary study of the characteristics of vessels in groups should contribute to this choice. We are talking about the analysis of deadweight, specific cargo capacity, utilization ratio by deadweight, and other characteristics of the vessel. The factors (or groups of factors) in question, the values of which change within narrow limits, will be shown on the groups of vessels belonging to them, as a marker.

For example, according to the table of grouping by the values of factors F_4, F_6 (Table 3), a group of vessels of the same type is established, for which the values of the factors lie within ($-0.6<F_4<-0.37$; $0.75<F_6<0.86$). These vessels are 11, 15, 20, 31, 32. Next, according to the table of grouping of vessels by factor F_7 (Table 4), it is necessary to select a group that contains vessels that fall into the previously selected group. This group contains vessels 1, 2, 11, 15, 22, 31 and is determined by the inequality $-0.27<F_7<-0.26$. For this group of vessels obtained by crossing two previously selected groups, namely for vessels 11, 15, 31, a system of inequalities is formulated with respect to dimensionless parameters (5).

Four of the seven parameters of the system are independent. For example, the geometric characteristics L/B, B/H, $D_b/\rho LBH$, $D_b/\rho LBT$ can be selected as independent. Then the admissible values of the relative power $\frac{P}{gD_bV}$, speed $\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}$, and the characteristics of the load capacity $\frac{D_b-D_n}{D_b}$ are determined from the system of inequalities:

$$\begin{cases} e^{-0.6}\le\left(\frac{L}{B}\right)^{0.41}\left(\frac{D_b}{\rho LBT}\right)^{-0.11}\left(\frac{D_b-D_n}{D_b}\right)^{-0.52}\times \\ \times\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{-0.3}\left(\frac{P}{gD_bV}\right)^{0.22}\le e^{-0.37}, \\ e^{0.75}\le\left(\frac{B}{H}\right)^{-0.2}\left(\frac{D_b}{\rho LBH}\right)^{-0.07}\left(\frac{D_b-D_n}{D_b}\right)^{0.04}\times \\ \times\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{-0.25}\le e^{-0.37}, \\ e^{-0.27}\le\left(\frac{L}{B}\right)^{-0.07}\left(\frac{B}{H}\right)^{-0.1}\left(\frac{D_b}{\rho LBT}\right)^{0.09}\left(\frac{D_b}{\rho LBH}\right)^{0.21}\times \\ \times\left(\frac{D_b-D_n}{D_b}\right)^{-0.07}\left(\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}\right)^{0.007}\left(\frac{P}{gD_bV}\right)^{-0.008}\le e^{-0.26}. \end{cases} \quad (5)$$

The results of calculations using inequalities (5) are given in Table 5.

Table 5

Numerical verification of algorithm performance

Vessel number in the sample	$-0.6<F_4<-0.37$; $0.75<F_6<0.86$; $-0.24<F_7<-0.23$									
	Dimensionless parameters of vessels from replica							Calculation		
	$\frac{L}{B}$	$\frac{B}{H}$	C_b	η_{dw}	$\frac{D_b}{\rho LBH}$	$\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}$	$\frac{P}{gD_bV}$	$\frac{D_b}{\rho LBH}$	$\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}$	$\frac{P}{gD_bV}$
11	6.025	1.904	0.844	0.855	0.608	0.340	0.0008	0.670	0.320	0.00077
15	5.630	2.270	0.845	0.765	0.624	0.395	0.0011	0.682	0.303	0.0006
20	6.227	1.867	0.857	0.817	0.603	0.369	0.0014	0.653	0.330	0.0007
31	5.610	2.235	0.844	0.882	0.623	0.307	0.0004	0.731	0.351	0.00041
32	6.267	1.818	0.860	0.868	0.630	0.301	0.0006	0.680	0.355	0.0006

If the resulting intervals of values contradict common sense, then the values of independent quantities L/B, B/H, $D_b/\rho LBH$, $D_b/\rho LBT$ should be adjusted. If they are chosen wisely, the designed ship should fall into the same selected group, that is, it will be a ship of the same type. Also, independent ones $\frac{P}{gD_bV}$, $\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}$, $\frac{D_b-D_n}{D_b}$, $\frac{D_b}{\rho LBT}$ can be chosen. Then, with care in choosing their values, the values of L/B, B/H, $D_b/\rho LBH$ are determined from the system of inequalities so that the designed ship belongs to the same type as ships from the selected group.

6. Discussion of results of using the method of factor analysis in the process of determining the basic characteristics of bulk carriers

There are no relevant quantitative results for comparison in available sources specifically devoted to the design of bulk carriers, so the data obtained in this study can only be compared with sample data. However, it is possible to conduct a comparative analysis of the methods used in this work and alternative methods, to highlight their advantages and disadvantages.

The results of studies [10, 11] for container ships are one-component linear and nonlinear regressions with the basic parameter *TEU* (the number of standard 20-foot containers) of form $\sum k_i TEU^{\alpha_i}$ and three-component regressions with the basic parameters D, *TEU* and V of form $\sum_i \left(a_i D^{\alpha_i} + b_i TEU^{\beta_i} + c_i V^{\gamma_i}\right)$, where D is displacement and V is speed. However, regression formulas have several disadvantages:

– instability to outliers of the sample data, therefore the data should be «smoothed» before processing, i.e., the already limited information about the vessels is lost;

– the statistical dependence of the basic parameters and the mutual influence of the sought values are not taken into account;

– regression coefficients have no physical meaning.

The method of principal components used in our work to determine the characteristics of the designed vessel largely meets the specified requirements for the following reasons. First, the calculated factors, dimensionless sets of characteristics of sample vessels are, in fact, new phase orthogonal coordinates of the problem with unequal dispersions of the sample along the coordinate directions. Evaluating the significance of correlations of factors and vectors of the sample data makes it possible to discard factors that have a weak influence on the values of the parameters of the sample vessels. Thus, the model of the problem is simplified while preserving essential connections. The factors accepted for the study serve as a basis for grouping vessels according to the values of these factors. In the task, out of 7 factors, only 3 were found to be significant (F_4, F_6 and F_7 see expression (3)). The distribution of vessels according to the intervals of the values of these factors made it possible to divide 32 vessels into 12 groups, and only 3 of them are numerous: one has 4 vessels, another has 6 vessels, and the third has 7 vessels. Moreover, the last two groups are neighbors by only one factor, and by the other two they are in the same group. The subject analysis of ships from the same group (11, 15, 20, 31, 32) showed their similarity in terms of design decisions, namely, a similar type of bulk carrier (four out of five ships are of the Capesize class), a similar structural structure (four out of five ships have 9 cargo holds and no cargo devices), similar hull shape: the coefficient of total completeness C_b for four of the five ships is about 0.81. Bulker hull contours are quite conservative within a separate class, so C_b values are stable.

The results of using the method of principal components make it possible to reduce the number of significant factors, as can be seen from expressions (3), (4), which leads to a simplification of the analysis of the sample by ships. This shortens the process of calculating ship characteristics based on the values of the factors that define a group of ships of the same type (using Tables 3, 4).

The analysis of the data obtained by the devised methodology reveals that the vessels of the same designers are indeed grouped by the value of the factor sets. This does not contradict the assumption proposed above about the existence of successful algorithms used by ship project designers. In the event that the design office does not have its authentic algorithm for determining the main characteristics of the designed vessel, the use of the described analysis method should help determine the main factors affecting the quality of the design. Further, it is possible to apply more subtle mechanisms for project optimization according to selected economic and technical parameters.

The reliability of our results is predetermined by the universality of the use of the factor analysis method for processing sample statistical data, and its correct application.

Directions for further research include the following:

– an increase in the number of ships in the sample, which is restrained by the limitation and reliability of data on ships from open sources;

– clarifying and supplementing available technical information on ships (for example, data on the features of the architectural and structural type – the presence of longitudinal bulkheads, etc.);

– addition of information about ships with characteristics of economic efficiency of operation;

– justification of the breakdown of factors into ranges in the process of grouping: both large and small groups of vessels are uninformative.

7. Conclusions

1. Expressions for the dependence of determining factors on the dimensionless characteristics of vessels in the sample were constructed. Statistically independent combinations of geometric, weight, kinematic, and energy parameters of the vessel, namely L/B, B/H, C_b, η_{dw}, were determined.

2. The most informative factors F_4, F_6 and F_7 were selected for the grouping of vessels and systematic analysis of their design features. Three significant groups of vessels were distinguished. An analysis of the vessels by groups was carried out and the interpretation of the found factors was performed.

The intervals of factor values ($-0.37<F_4<0.13$), ($0.64<F_6<0.75$), ($-0.24<F_7<-0.23$) determined the group of bulkers of the «Panamax» and «Kamsarmax» types.

The intervals of factor values ($-0.6<F_4<-0.37$), ($0.75<F_6<0.86$), ($-0.24<F_7<-0.23$) determined the group of bulkers of the «Capesize» class.

The intervals of factor values ($-0.6<F_4<-0.37$), ($0.64<F_6<0.75$), ($-0.26<F_7<-0.24$) determined a group of large bulkers of different classes that have similar features: «Large bulk», «Handycape», «Handymax», «Post Panamax».

3. An algorithm was developed, according to which based on the values of F_4, F_6 and F_7 factors, which determine the group of ships, and the values of sets L/B, B/H, C_b and η_{dw} the values of $\frac{D_b}{\rho LBH}$, $\sqrt[6]{\frac{\rho}{D_b}}\frac{V}{\sqrt{g}}$ and $\frac{P}{gD_bV}$ were calculated. The aggregate of seven sets determines the geometric, weight, kinematic, and energy parameters of the first approximation for the projected vessel. Dimensionless characteristics of the «Capesize» class bulker were calculated according to the algorithm.

Conflicts of interest

The authors declare that they have no conflicts of interest in relation to the current study, including financial, personal, authorship, or any other, that could affect the study and the results reported in this paper.

Funding

The study was conducted without financial support.

Data availability

All data are available in the main text of the manuscript.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the presented work.

References

1. Nekrasov, V. A., Kabanova, N. N. (2008). Zadacha funktsionirovaniya sudna, ekspluatiruyushchegosya na nezakreplennyh liniyah perevozok, v probleme ego proektirovaniya. NUK, Sbornik nauchnyh trudov NUK, 6 (423), 35–40.
2. Kabanova, N. N. (2012). Reshenie stohasticheskoy zadachi optimizatsii harakteristik trampovogo sudna. Herald of the Odessa National Maritime University, 34, 172–183. Available at: http://visnyk.onmu.odessa.ua/index.php/1/issue/view/36/29
3. Trincas, G., Mauro, F., Braidotti, L., Bucci, V. (2018). Handling the path from concept to preliminary ship design. CRC Press. Available at: https://www.taylorfrancis.com/chapters/edit/10.1201/9780429440533-15/handling-path-concept-preliminary-ship-design-trincas-mauro-braidotti-bucci
4. la Monaca, U., Bertagna, S., Marinò, A., Bucci, V. (2019). Integrated ship design: an innovative methodological approach enabled by new generation computer tools. International Journal on Interactive Design and Manufacturing (IJIDeM), 14 (1), 59–76. doi: https://doi.org/10.1007/s12008-019-00612-4
5. Nikolopoulos, L., Boulougouris, E. (2020). A novel method for the holistic, simulation driven ship design optimization under uncertainty in the big data era. Ocean Engineering, 218, 107634. doi: https://doi.org/10.1016/j.oceaneng.2020.107634
6. Koutroukis, G., Papanikolaou, A., Nikolopoulos, L., Sames, P., Köpke, M. (2013). Multi-objective optimization of container ship design. Developments in Maritime Transportation and Exploitation of Sea Resources, 477–489. doi: https://doi.org/10.1201/b15813-58
7. Priftis, A., Boulougouris, E., Turan, O. (2018). Parametric design and holistic optimisation of post-panamax containerships. Conference: 7th European Transport Research Arena (TRA 2018). Available at: https://www.researchgate.net/publication/324603164_Parametric_design_and_holistic_optimisation_of_post-panamax_containerships
8. Frantsev, M. E. (2014). Ispol'zovanie chennyh metodov pri realizatsii zadachi parametricheskogo proektirovaniya kompozitnogo promyslovogo sudna dlya pribrezhnogo lova. Sudostroenie, 4, 30–34.
9. Frantsev, M. E. (2010). Proektnoe obosnovanie optimal'nyh sochetaniy harakteristik massy, moshchnosti i skorosti dlya skorostnyh sudov iz kompozitov metodami analiza baz dannyh. Transportnoe stroitel'stvo, 3, 53–59.
10. Statistical Analysis and Determination of Regression Formulas for Main Dimensions of Container Ships based on IHS Fairplay Data. Available at: https://docplayer.net/50727912-Statistical-analysis-and-determination-of-regression-formulas-for-main-dimensions-of-container-ships-based-on-ihs-fairplay-data.html
11. Cepowski, T., Chorab, P. (2021). Determination of design formulas for container ships at the preliminary design stage using artificial neural network and multiple nonlinear regression. Ocean Engineering, 238, 109727. doi: https://doi.org/10.1016/j.oceaneng.2021.109727

12. Elvekrok, D. R. (1997). Concurrent Engineering in Ship Design. Journal of Ship Production, 13 (04), 258–269. doi: https://doi.org/10.5957/jsp.1997.13.4.258
13. Larkin, Yu. M., Onishchenko, A. F. (2015). Osobennosti proektirovaniya balkerov. Herald of the Odessa National Maritime University, 3, 219–228. Available at: http://nbuv.gov.ua/UJRN/Vonmu_2015_3_22
14. Pinskiy, A. N. (1977). Ispol'zovanie statisticheskih metodov v algoritmah avtomatizirovannoy sistemy proektirovaniya. Voprosy sudostroeniya. Seriya «Proektirovanie», 13, 131–140.
15. Zhukovskaya, V. M., Muchnik, I. B. (1976). Faktorniy analiz v sotsial'no-ekonomicheskih issledovaniyah. Moscow: «Statistika», 153.
16. Kim, Dzh.-O., Myuller, Ch. U., Klekka, U. R., Oldenderfer, M. S., Bleshfild, R. K. (1989). Faktorniy, diskriminantniy i klasterniy analiz. Moscow: «Finansy i statistika», 216.
17. Louli, D., Maksvell, A. (1967). Faktorniy analiz kak statisticheskiy metod. Moscow: Mir, 144.