Yordan Garbatov
Petar Georgiev

http://dx.doi.org/10.21278/brod74404

ISSN 0007-215X
eISSN 1845-5859

Principal component analysis of containership traffic in the Black Sea

UDC 629.544:656.025.4
Original scientific paper

Summary

A novel quantitative analysis employing the Principal Component Analysis (PCA) of containership traffic in the Black Sea from 2018 to 2021 is performed. The study uses a matrix covering five ship size classes from A to E for four years of operation, from 2018 to 2021, accounting for ship traffic, CO_2, fuel consumption (FC), shipping intensity, and eco and traffic efficiency. Only the first two principal factors are analysed because of their total variation weight. Shipping intensity, FC intensity, and CO_2 intensity plays a significant role in the first factor, while Eco efficiency, FC efficiency, and Traffic efficiency are considered for the second factor. Notably, the set of parameters pertains to time and is strongly associated with DWT. Two principal components were identified, F1 and F2, where F1 integrates efficiency and intensity. At the same time, F2 separates the intensity from the efficiency conditional on the ship size and the year of operations. In the principal component F1 the activities of ships A and C differ from B, D and E, separating more efficiently from less efficiently used ships, and in F2, the activities of class sizes of ships C and D and E contrast A and B ships, distinguishing the big-size class ships from small ones. It was concluded that the most intensively used ships are the ship size classes C and D, and the most efficient are ship size classes A and B. The most intensive use of the ships was in 2020, followed by 2021, and the most efficient were in 2018, 2019. Based on the ship activities and using the Within-class variance, ships are grouped into two clusters of similar activities, where the first one, with lower variance and more homogeneous, includes only the ship size class A. The second one with a relatively large variance consists of the rest size of the ships. Linear relationships considering the intensity and efficiency are derived as a function of the main variables, where the factor loading represents the variable's coefficient, given as a relative weight to any factor.

Keywords: *ship traffic; fuel consumption; shipping; CO_2; Eco-efficiency; PCA*

1. Introduction

Over the past ten years, the International Maritime Organization (IMO) has undertaken remarkable initiatives to curb greenhouse gas (GHG) emissions from maritime operations. In April 2015, the European Parliament sanctioned a strategy to progressively integrate this issue into the European Union's ongoing endeavours to mitigate domestic GHG emissions. On July

1 2015, Regulation 2015/757 April 29 [1] was enacted. This strategy primarily focuses on the monitoring, reporting, and verification (MRV) of carbon emissions from ships.

The developed EU-MRV database [2] started collecting data on January 1, 2018. Data for various emission indicators of ships with GT > 5,000 are monitored, reported, and verified before being included in the database.

For the first year, the database included information for more than 11,600 ships, about 38% of the world's merchant fleet. The total amount emitted CO_2 in 2018 was over 138 MT, which is comparable to the CO_2 emissions of Belgium (105.1 MT) and Chechia (108.3 MT) for the same year [3]. Container ships have the largest CO_2 emission share, about 30% of the total amount, followed by bulk carriers and tankers [4].

Data on container ships visiting six Black Sea ports: Varna, Burgas, Constanta, Odesa, Novorossiysk, Poti and Ambarli [5] was collected to identify the demand and potential need for ships [6]. Ship traffic data is taken from the Marine Traffic Platform [7], and vessel characteristics are from the Equasis database [8]. The types of containerships operated by the Black Sea ports are shown in Fig. 1.

Type	No of TEU
Small feeder	Up to 1000
Feeder	1001 - 2000
Feeder max	2001-3000
Panamax	3001-5000
Post -Panamax	5001 – 8000
New Panamax	8001 – 1000
Ultra-large	above 10000

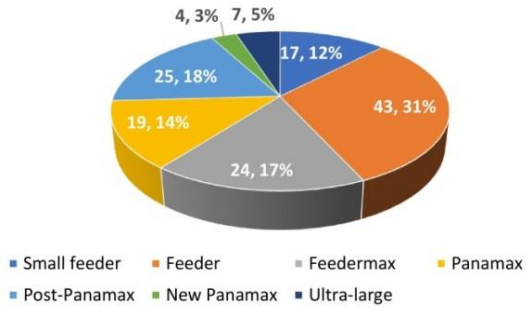

Fig. 1 Types of containership operated in Black Sea ports [5]

The emission data for these ships for 2018 to 2021 were extracted from the EU-MRV database based on the ship's name (or IMO number in the case of a name change).

The amendments to the EU MRV Regulation 2015/757 from 2023 May 10 [9] enlarge the scope of this implementation. Regarding emissions, the term CO_2 is replaced by "greenhouse gas". From January 1, 2025, the application is enforced for offshore ships above 400 GT and general cargo ships between $400 \leq GT \leq 5{,}000$. Attention to ship emission databases such as EU-MRV and IMO DCS (Data Collection System) from October 28 [10] will increase because they can be used to estimate the CII (Carbon Intensity Indicator) for existing ships.

Large datasets are becoming increasingly prevalent, presenting challenges in terms of interpretation. To address this issue, PCA offers a valuable technique for reducing the complexity of such datasets [11]. By doing so, it enhances interpretability while minimising the loss of information. The origins of PCA can be traced back to the work in [11]. However, the computational feasibility of applying PCA to non-trivially small datasets was realised only with the widespread availability of electronic computers in the following decades. A brief review of PCA's state of the art can be found in [12].

Kawashima et al. [13] introduced a novel approach for characterising and generating ship traffic flow through PCA. This method is employed explicitly on AIS data, and the resulting ship traffic data exhibit similar characteristics to the original AIS data. Consequently, a simulation method is developed to assess the likelihood of encounters between ships. By extracting encounter conditions associated with collision risks, it becomes possible to calculate the encounter probability.

To enhance operational efficiency and service quality in container handling, [14] suggested the utilisation of a hybrid prediction model that combines Principal Component Analysis (PCA) and Extreme Learning Machine (ELM), which is optimised using Improved Particle Swarm Optimisation (IPSO). This approach aims to develop scientific and practical berth plans. By employing PCA, the model reduces the dimensionality of the investigation data, which encompasses uncertain factors that impact the operation time of container ships at the berth.

Perera and Mo [15] introduced an approach based on PCA to evaluate ship performance and navigation behaviour in specific marine engine operating regions. A data set containing ship performance and navigation information was analysed to uncover the underlying structure of the selected operating region. The data set was initially categorised based on the engine's operating points, identifying three central operating regions. Subsequently, one of these operating regions was examined to determine the principal components (PCs) associated with it. These PCs represented the relationships between various ship performance and navigation parameters and their connection to the operating region, enabling the evaluation of ship performance and navigation behaviour.

The Principal Component Analysis was also employed in elaborating ship hull form. A study in [16] used PCA to compress the geometric representation of a set of existing vessels. The resulting principal scores were then used to generate numerous derived hull forms, which are computationally evaluated for their performance in still-water conditions.

Helmsmar [17] proposed an integrated approach for identifying the efficient hull configuration by combining a data compression method with Computational Fluid Dynamics (CFD). To achieve this, a comprehensive database of normalised ship offsets was compiled through a literature survey, focusing on a specific hull form. The table of offsets was compressed using PCA, resulting in a set of scores. These compressed scores are then utilised as inputs in the CFD solver to evaluate the still-water performance.

The present study aims to employ PCA for characterising the container ship traffic flow. The analysis used AIS data for container ship traffic in the Black Sea. Only the first two principal components are analysed, and the eigenvectors of the variables, factor loadings and correlation between variables are established. Several relations of the essential intensity and efficiency factors in containership traffic in the Blac Sea are identified.

2. Data and analysed variable

The analysis is performed for the container ship traffic in the Black Sea for 2018, 2019, 2020 and 2021 years. The number of voyages, speed and DWT are taken from AIS (Automatic Identification System) and given in Fig. 2 to 7. The registered containerships are classified into five classes, where Class A is from 8 to 29 DWT×1,000, B is from 29 to 50 DWT×1000, C is from 50 to 72 DWT×1,000, D is from 72 to 93 DWT×1,000 and E is from 93 to 114 DWT×1,000.

In the Black Sea region from 2018 to 2021, the containerships slightly reduced the number of voyages and the voyage time (Fig. 2 and Fig. 6), increasing the DWT (Fig. 3). The speed (Fig. 7) was slightly increased and almost unchanged for the fuel consumption (Fig. 4) and CO_2 emissions (Fig. 5).

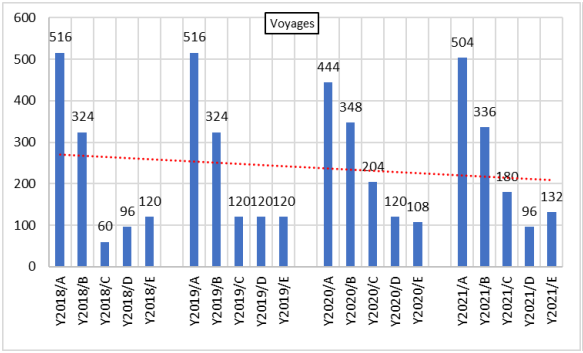

Fig. 2 Voyages as a function of ship size class and year.

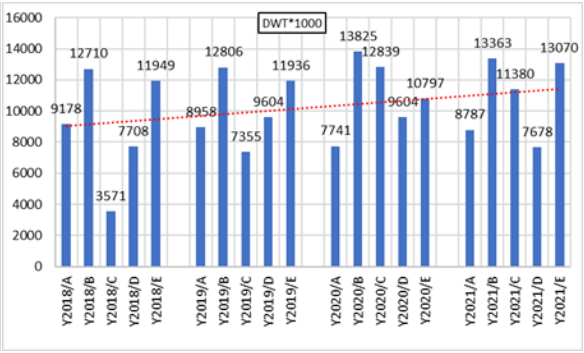

Fig. 3 DWT as a function of ship size class and year

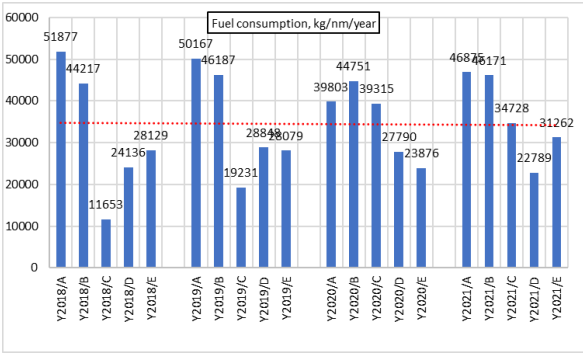

Fig. 4 Fuel consumption as a function of ship size class and year.

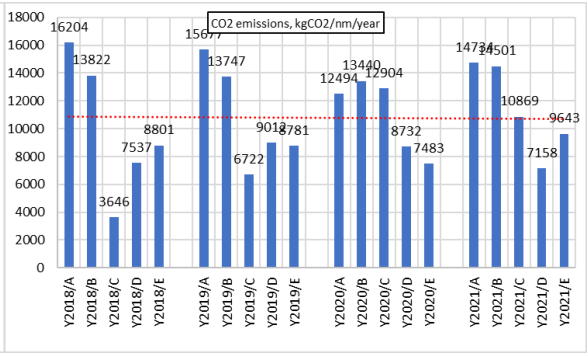

Fig. 5 CO_2 emissions as a function of ship size class and year

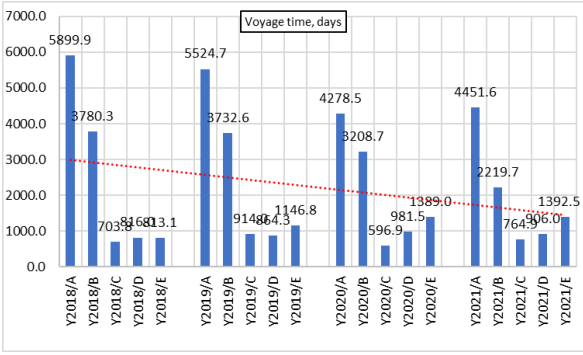

Fig. 6 Voyage time as a function of ship size class and year.

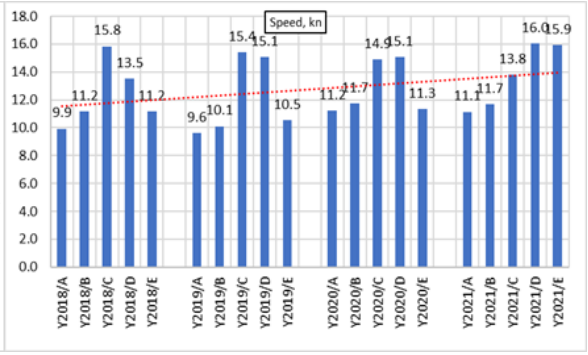

Fig. 7 Speed as a function of ship size class and year

Fuel consumption intensity is defined as the fuel consumption needed for a unit of time in operation (see Fig. 8):

$$FC_{int\,i,j} = \frac{FC_{i,j}}{Time_{i,j}} \tag{1}$$

where $FC_{i,j}$ is the average fuel consumption of the i^{th} class ship in the j^{th} year and $Time_{i,j}$ is the time, days, spent in one year of operation of the i^{th} class ship in the j^{th} year.

The CO_2 intensity is defined as the average $CO_{2,i,j}$emission generated for the time spent on the i^{th} class ship in the j^{th} year for the $Time_{i,j}$ spent in one year of operation (see Fig. 9):

$$CO_{2,int,i,j} = \frac{CO_{2,i,j}}{Time_{i,j}} \tag{2}$$

The traffic intensity is defined as the number of voyages carried for the time spent by the i^{th} class ship in the j^{h} year of a one-year operation (see Fig. 10):

$$Traffic_{int,i,j} = \frac{Voyages_{i,j}}{Time_{i,j}} \tag{3}$$

where $Voyages_{i,j}$ is the voyages.

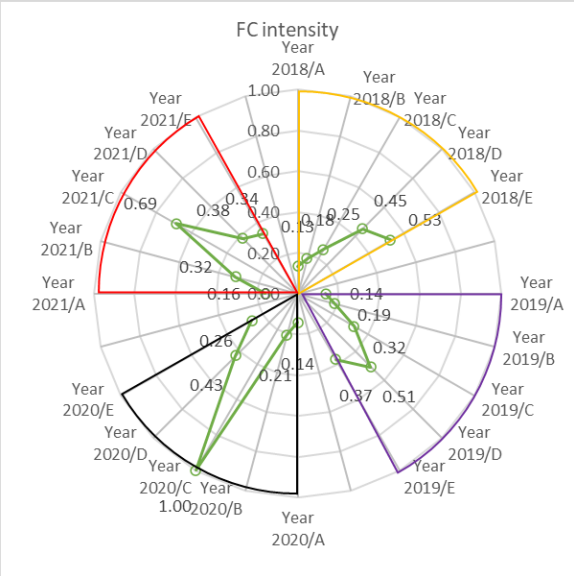

Fig. 8 Normalised fuel consumption intensity as a function of ship size class and year.

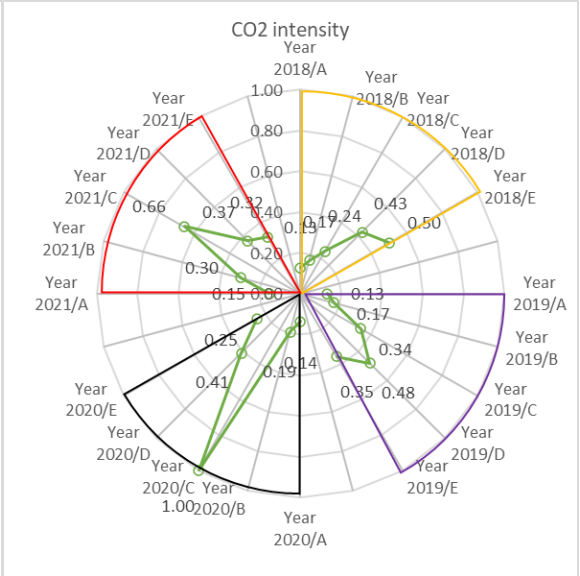

Fig. 9 Normalised $CO_{2,i,j}$as a function of ship size, class, and year.

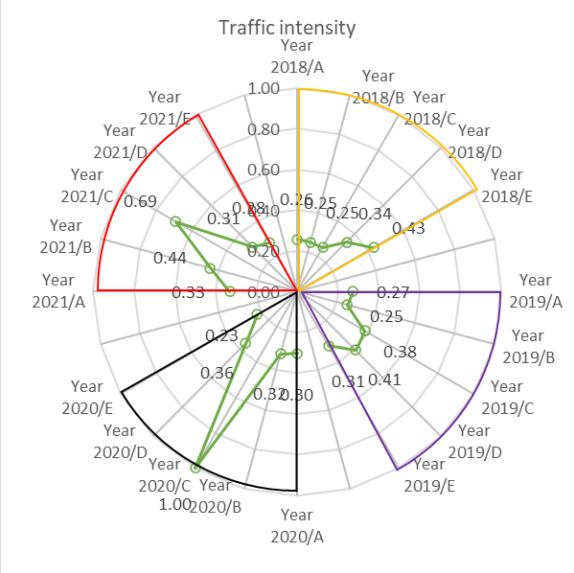

Fig. 10 Normalised traffic intensity as a function of ship size class and year.

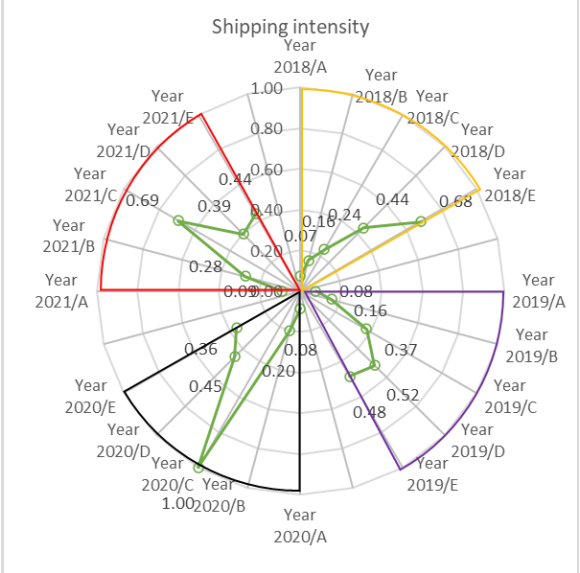

Fig. 11 Normalised shipping intensity as a function of ship size class and year

The shipping intensity is defined by the total amount of $DWT_{i,j}$ transported in the time (see Fig. 11):

$$Shipping_{int\ i,j} = \frac{DWT_{i,j}}{Time_{\ i,j}} \tag{4}$$

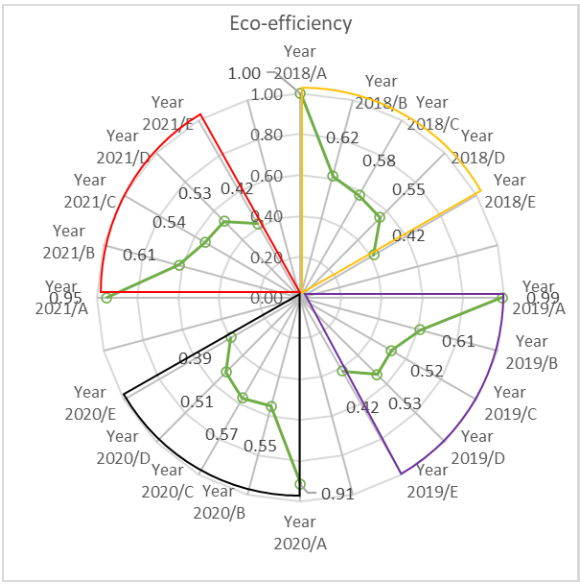

Fig. 12 Normalised eco-efficiency as a function of ship size class and year.

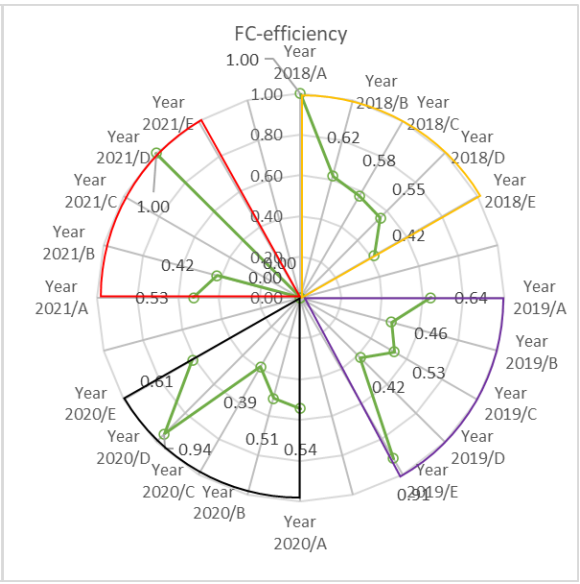

Fig. 13 Normalised fuel consumption efficiency as a function of ship size class and year

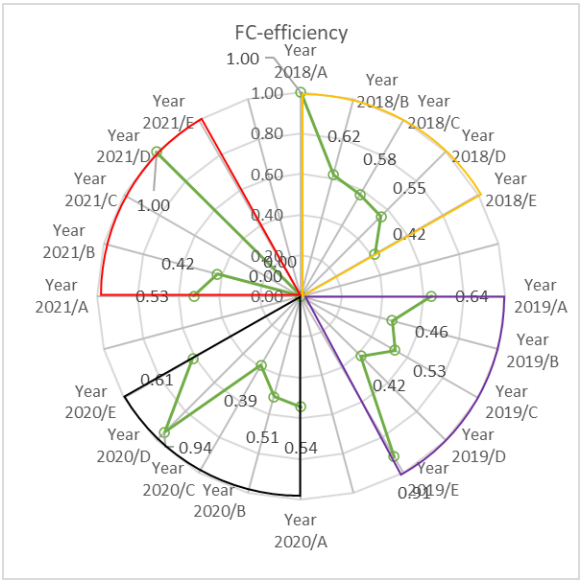

Fig. 14 Normalised traffic efficiency as a function of ship size class and year

Eco-efficiency is defined as the $CO_{2,i,j}$ emission generated by the i^{th} class of ships in the j^{th} year for transporting a total amount of $DWT_{i,j}$ (see Fig. 12):

$$Eco_{eff\ i,j} = \frac{CO_{2,i,j}}{DWT_{i,j}} \tag{5}$$

where $CO_{2,i,j}$ is the CO_2 emissions, kgCO_2/nm/year.

Fuel consumption efficiency is defined as the fuel consumption needed for transporting a unit of DWT in operation (see Fig. 13):

$$FC_{eff\ i,j} = \frac{FC_{i,j}}{DWT_{i,j}} \tag{6}$$

Traffic efficiency is defined as the number of voyages for transporting a unit DWT (see Fig. 14):

$$Traffic_{eff,i,j} = \frac{Voyages_{i,j}}{DWT_{i,j}} \quad (7)$$

It can be seen from Fig 8 and 9 that the normalised FC intensity and normalised CO_2, intensity as a function of ship size class and year are very similar but vary with the years and ship size class. The most intensive FC and CO_2 intensity are seen in 2020 for class size ships C, accompanied by the most intensive traffic and shipping (Fig10, 11). On the contrary, the smallest FC and CO_2 intensity are observed in 2018 for class-size ships A. Whit respect to the Eco-efficiency (Fig. 12), as defined by Eqn (5), the most efficient class-size ships are A-ships for all analysed years. The most efficient FC (Fig. 13) for 2018 is shown for class size ships A, for 2019 is E, and for 2020 and 2021 is D, which looks like the traffic efficiency presented in Fig. 14.

3. Principal Component Analysis

The Principal Component Analysis is used to analyse the original information of ship traffic containerships in the Black Sea taken from the AIS as given in Fig. 2 to 7 into smaller factors [18] and identify the critical variables influencing the traffic intensity and efficiency.

PCA is performed for seven variables "Shipping intensity", "FC intensity", "CO_2 intensity", "Traffic intensity", "Eco efficiency", "FC efficiency", and "Traffic efficiency" for the objects "Y2018/A", "Y2018/B", "Y2018/C", "Y2018/D", "Y2018/E", "Y2019/A", "Y2019/B", "Y2019/C", "Y2019/D", "Y2019/E", "Y2020/A", "Y2020/B", "Y2020/C", "Y2020/D", "Y2020/E", "Y2021/A", "Y2021/B", "Y2021/C", "Y2021/D" and "Y2021/E". The summary statistics, including observations (Obs.) minimum (Min) and maximum (Max), mean value (Mean) and standard deviation (Std. Dev) and correlation matrix of variables are given in Table 1 and 2.

Table 1 Summary statistics

Variable	Obs.	Min.	Max.	Mean	Std. Dev.
Shipping intensity (SI)	20	1.556	21.510	7.734	5.239
FC intensity (FCI)	20	8.793	65.868	23.027	14.153
CO_2 intensity (CO2I)	20	2.747	21.619	7.275	4.602
Traffic intensity (TRI)	20	0.078	0.342	0.127	0.062
Eco-efficiency (EE)	20	0.693	1.766	1.079	0.341
FC efficiency (FCE)	20	2.211	5.652	3.440	1.098
Traffic efficiency (TRE)	20	0.010	0.058	0.024	0.018

Table 2 Correlation matrix

Variables	SI.	FCI	CO2I	REI	EE	FCE	TRE
SI	1	0.969	0.966	0.800	-0.630	-0.655	-0.689
FCI	0.969	1	0.999	0.899	-0.476	-0.501	-0.565
CO2I	0.966	0.999	1	0.909	-0.463	-0.492	-0.552
TRI	0.800	0.899	0.909	1	-0.155	-0.189	-0.225
EE	-0.630	-0.476	-0.463	-0.155	1	0.996	0.975
FCE	-0.655	-0.501	-0.492	-0.189	0.996	1	0.976
TRE	-0.689	-0.565	-0.552	-0.225	0.975	0.976	1

It can be seen that "Shipping intensity (SI)", "FC intensity (FCI)", "CO_2 intensity (CO2I)", and "Traffic intensity (TRI)" are highly positively correlated and negatively correlated to "Eco efficiency (EE)", "FC efficiency (FCE)", and "Traffic efficiency (TRE)". "Traffic intensity" is low and negatively correlated to "Eco efficiency", "FC efficiency", and "Traffic efficiency".

The eigenvalue, variability and cumulative contribution of the principal components are given in Table 3, and the eigenvectors of the variables, factor loadings and correlation between variables and factors are in Table 4.

The 70% rule introduced in [19] decides the number of significant components. It was recommended that the retained variables capture 70% of the total variances, and if their eigenvalue is greater than one, the factors are considered important. Only the first two principal factors will be considered since they demonstrate eigenvalues bigger than one and represent 98.39% of the total weight.

The eigenvalues show the projection quality from the seven-dimensional variables analysis into two dimensions: intensity and efficiency. The first eigenvalue is 5.07 and covers 72 % of the total variability. The total variability can still be observed if the data are presented on only one axis. Each eigenvalue is related to a one-dimensional factor, a linear combination of the initial variables, and all-uncorrelated factors.

Table 3 Eigenvalue, variability, and cumulative contribution of principal components

	F_1	F_2
Eigenvalue	5.070	1.818
Variability (%)	72.425	25.967
Cumulative %	72.425	98.392

Table 4 Factor descriptors

Variable	$F_1 -$ Eigenvector	F_2- Eigenvector	F_1 - Loadings	F_2- Loadings	F_1 - Correlation	F_2- Correlation
Shipping intensity	0.967	0.203	0.967	0.203	0.967	0.203
FC intensity	0.917	0.389	0.917	0.389	0.917	0.389
CO_2 intensity	0.912	0.403	0.912	0.403	0.912	0.403
Traffic intensity	0.710	0.675	0.710	0.675	0.710	0.675
Eco efficiency	-0.784	0.614	-0.784	0.614	-0.784	0.614
FC efficiency	-0.804	0.587	-0.804	0.587	-0.804	0.587
Traffic efficiency	-0.835	0.535	-0.835	0.535	-0.835	0.535

The first two eigenvalues correspond to 98.39 % of the variance, guaranteeing that the maps based on the first two factors are a good quality projection of the initial multi-dimensional problem. The factor scores for the first two components are shown in Fig. 15.

The correlation circle belonging to axes F_1 and F_2 from Fig. 15 shows that the Second Component separates the intensity from efficiency, and the First Component integrates all.

It is also shown that when two variables are far from the centre and close to each other, they are significantly positively correlated. For example, "FC intensity - CO_2 intensity" and "Eco efficiency - FC efficiency - Traffic efficiency". However, if they are orthogonal, they are not correlated.

The analysis of the values of the contributions of the variables and squared cosines of the variables refines this conclusion. The squared cosines of variables show that the first component highly contributes to Shipping intensity (0.934), FC intensity (0.841), CO_2 intensity (0.832) and Traffic intensity (0.504), and the Second Component contributes mainly to Eco-efficiency (0.615), FC-efficiency (0.646) and Traffic efficiency (0.697).

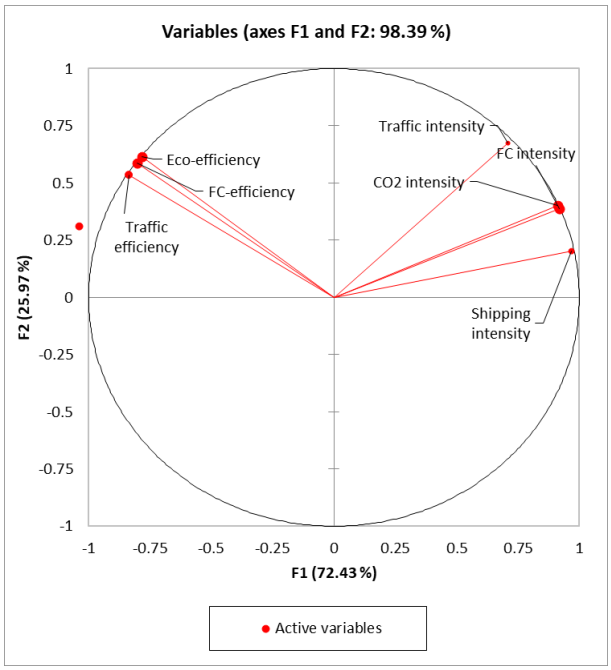

Fig. 15 Factors score of active variables.

The principal component analysis combines variables in a linear relationship considering a reduced number of components while maximising the variances. The component F_1 associated with the intensity and F_2 with efficiency are described by a linear model field [20]:

$$F_1(Intensity) = 0.967\overline{(Shipping - intensity)} + 0.917(\overline{FC - intensity}) + 0.912\left(\overline{CO_2 - intensity}\right) + 0.710\left(\overline{Traffic - intensity}\right) - 0.784\left(\overline{Eco - efficiency}\right) - 0.804\left(\overline{FC - efficiency}\right) - 0.835\left(\overline{Traffic - efficiency}\right) \quad (8)$$

$$F_2(Efficiency) = 0.203\left(\overline{Shipping - intensity}\right) + 0.389(\overline{FC - intensity}) + 0.403\left(\overline{CO_2 - intensity}\right) + 0.675\left(\overline{Traffic - intensity}\right) + 0.614\left(\overline{Eco - efficiency}\right) + 0.587\left(\overline{FC - efficiency}\right) + 0.535\left(\overline{Traffic - efficiency}\right) \quad (9)$$

where $\overline{(x_i)}$ is the normalised variable, and it can be transformed into the normal one by using the following:

$$\overline{(x_i)} = \frac{x_i - E(x_i)}{\sigma(x_i)} \tag{10}$$

where $E(x_i)$ is the mean value and $\sigma(x_i)$ is the standard deviation. The factor loading represents the variable's coefficient, given as a relative weight to any factor. The factor loadings are shown in Table 4.

To identify the impact of the observation from different years and class size of ships that contribute to these differences, the factor score of the components is presented in Fig. 16. The Second Component contrasts most of the class sizes of ships (C and D and E) to (A and B), meaning distinguishing the big-size class ships from small ones. The first component differs from the ships (A and C) from (B, D and E), separating more efficiently from less efficiently used ships.

It may also be concluded that the first and fourth quadrant of the map (see Fig. 16) identifies ships with high annual intensity (Traffic, CO_2, FC, and Shipping intensity. The first and second Quadrants of the map identify ships with high annual efficiency (Eco, FC and Traffic). Quadrant one covers the class size ships' performance (Y2020/C and Y2021/C), and Quadrant three includes the worst performance of class size ships (Y2018/B, Y2018/C, Y2019/B and Y2020B).

The Varimax orthogonal transformation has been employed to adjust the principal component axes and to facilitate the interpretation of the results [18, 21], which gives a clockwise rotation, providing a new set of rotation loadings, as shown in Fig. 17. The improvement in the simplicity of the interpretation is straightforward. The first dimension remains linked to Traffic, CO_2, FC, and Shipping intensity, and the second dimension now appears more clearly to be related to the Eco, FC, and Traffic efficiency.

Fig. 17 shows that the most intensively used ships are ship size categories C and D, and the most efficient are ship size categories A and B. The most intensive use of the ships was in 2020, followed by 2021, and the most efficient were in 2018, 2019.

Agglomerative Hierarchical Clustering (AHC) [22] is used to identify the dissimilarities between the objects to be grouped in clusters employing the dendrogram showing the progressive grouping.

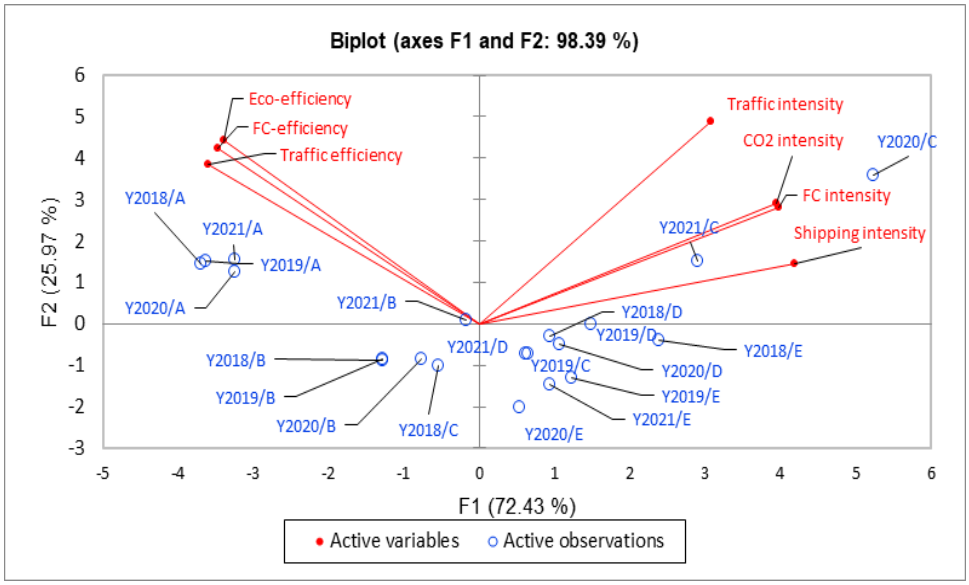

Fig. 16 Impact of active observation and variables before the Varimax rotation.

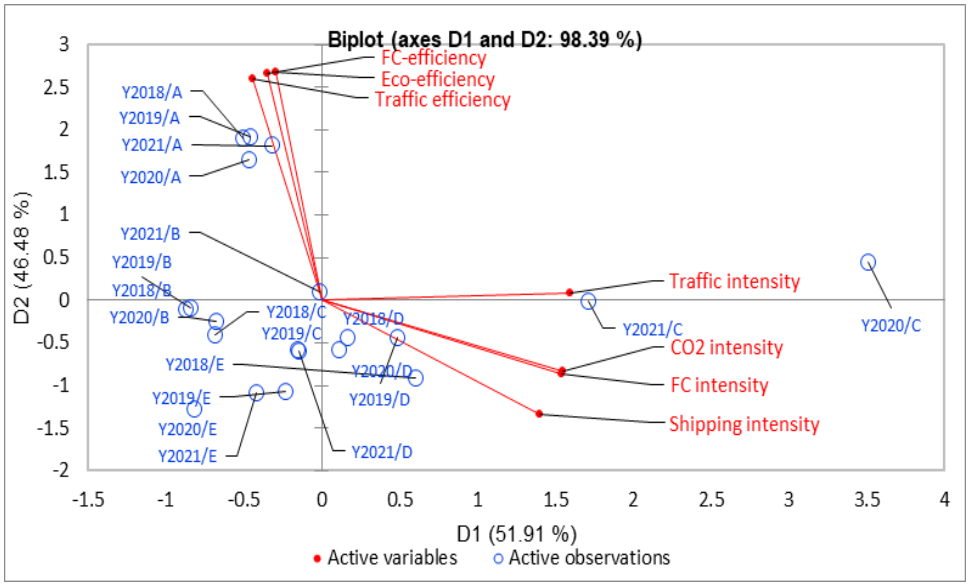

Fig. 17 Impact of active observation and variables after the Varimax rotation.

Fig. 18 presents the dendrogram showing how the subgroups of objects are grouped in two clusters. The first cluster (left) is relatively homogeneous compared to the second one (right). The first cluster covers the class ship size of A, and the second one cover the rest. Fig. 18 also shows the objects classified into each cluster. Based on the Within-class variance, the first cluster, variance=0.133, is more homogeneous than the second, variance =4.606.

The proximity between objects is identified by estimating how much they are similar (similarity) or dissimilar (dissimilarity) using the distance to the centroid, as seen in Fig. 19. Similarity is observed between the same class size of ships during the analysed years, but not seen similarity between the years.

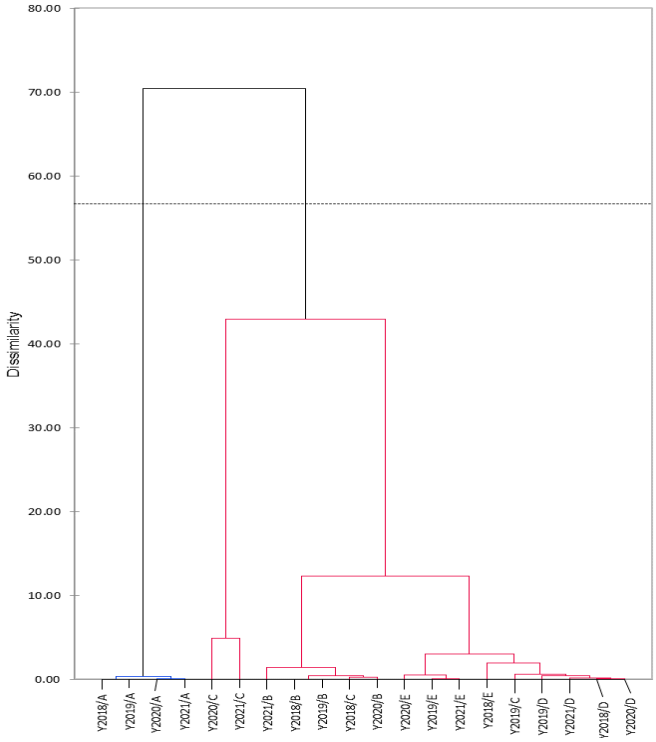

Fig. 18 Dendrogram.

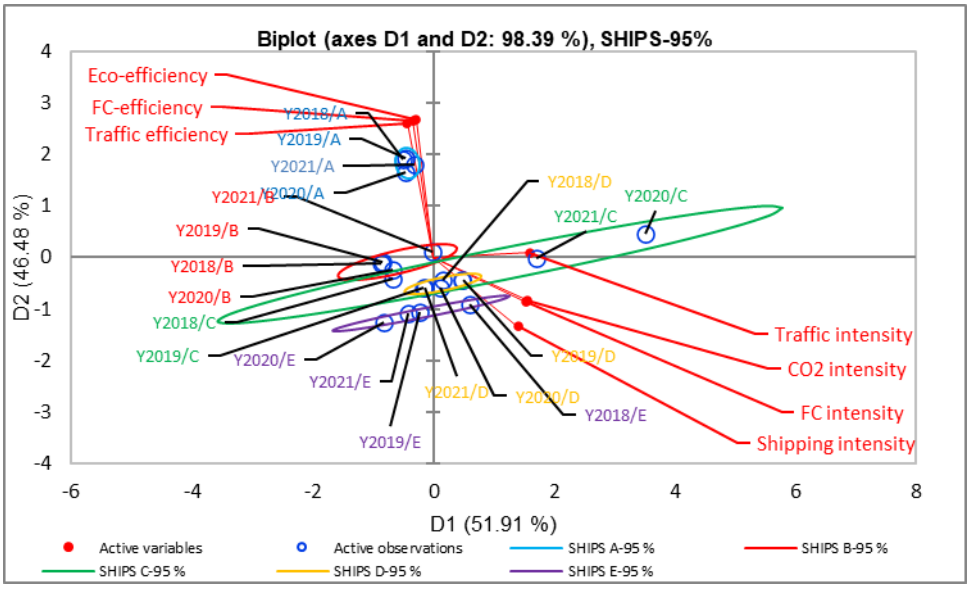

Fig. 19 95% confidence ellipse for ship class sizes

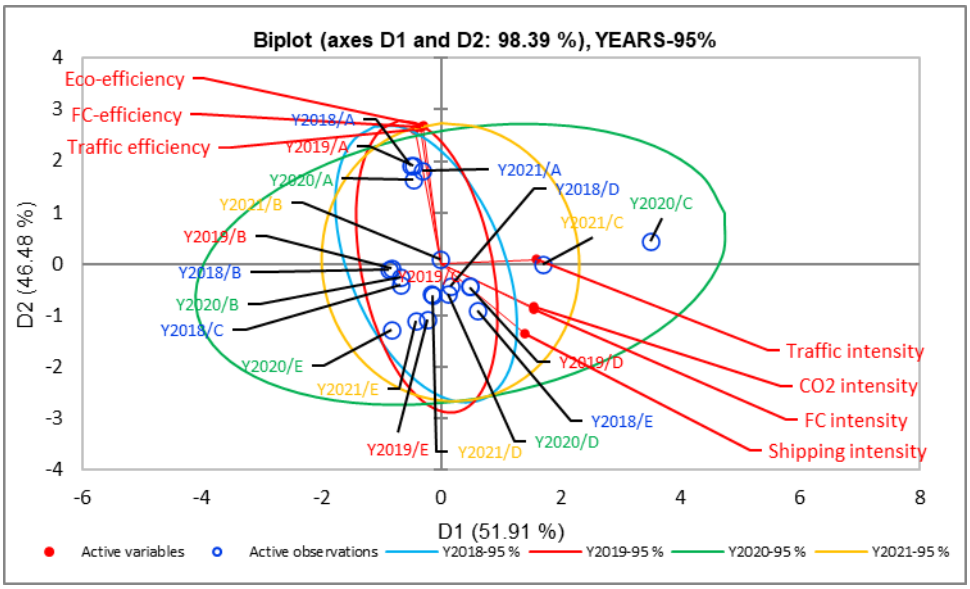

Fig. 20 95% confidence ellipse for ship traffic from 2018 to 2021.

Additional clustering was performed using a 95% confidence ellipse for ship class sizes and years presented as a function of the two factors in Fig. 19 and 20 [23, 24]. The mean point of the ellipse is located at coordinates $E(F_1)$ and $E(F_2)$, representing the mean of all F_1 coordinates and all F_2 coordinates and the confidence ellipse is defined by its centre at $E(F_1)$ and $E(F_2)$, covering a percentage of the F_1, F_2 points. The semi-major and minor axes of the ellipse are defined $k\sigma(F_1)$ and $k\sigma(F_2)$ representing the standard deviations of the independent random variables F_1, F_2 and k is the elliptical scale factor. The elliptical scale factor k is defined as 2.45 for 95 % confidence, and the elliptic equation is defined as:

$$\frac{F_1^2}{\sigma^2(F_1)} + \frac{F_2^2}{\sigma^2(F_2)} = k^2 \tag{10}$$

The results in those figures confirm the initial conclusions that the ships category A and B were the most efficient and D and E less efficient concerning eco, FC and traffic. It can also be seen that the ship class categories C and E were most intensively employed.

Regarding the year performance, it seems that the average ship efficiency was very similar during the analysed period—the intensity was significant in 2020 and 2021 and less in 2018 and 2019.

4. Conclusions

Initially, the study showed that the containerships slightly reduced the number of voyages, voyage time and increased the DWT in the Black Sea region from 2018 to 2021. The speed was increased somewhat and almost unchanged for the fuel consumption and CO_2 emissions. The normalised FC intensity and normalised $CO_{2,i,j}$ are very similar but vary with the years and ship size class. The most intensive FC and $CO_{2,i,j}$ are seen in 2020 for class size ships C, accompanied by the most intensive traffic and shipping. On the contrary, the smallest FC and $CO_{2,i,j}$ are observed in 2018 for class size ships A. Regarding Eco-efficiency, the most efficient class-size ships are A-ships for all analysed years. The most efficient FC for 2018 is shown for class size ships A, for 2019 is E, and for 2020 and 2021 is D, which also looks like the traffic efficiency. Employing the Principal Component Analysis established the principal component relationships used in a novel analysis of the relationship between containership ship size classes operated in the Black Sea for four years starting from 2018 up to 2021, accounting for traffic, CO_2, fuel consumption and shipping intensity, as well as Eco, Fuel consumption and Traffic efficiency. It has been identified that ship size categories A and B were the most efficient and D and E less efficient concerning Eco, Fuel consumption and Traffic. Additionally, it was seen that the ship size categories C and E were intensively used. What concerns the years of analysis, the ship categories of ship efficiency were very similar during the analysed period. However, more intensity was seen in 2020 and 2021 and less in 2018 and 2019. It was also observed that the ship size category A behaves similarly in different years, which is not the case for other ship size categories. Ships were clustered based on the ship activities and using the Within-class variance. The first cluster with a variance of 0.133 is more homogeneous and includes only the ship size class A, and the second one with a variance of 4.606 consists of the rest size of the ships. Linear relationships considering the intensity and efficiency are defined as a function of the main variables, where the factor loading represents the variable's coefficient, given as a relative weight to any factor. Additionally, the confidence ellipse approach has been applied to identify more information about the intensity, efficiency and similarity of the container ships operating in the Black Sea.

To better understand the intensity and the efficiency of the ships operated in the Black Sea, the interaction with the local economy needs to be considered, including the societal acceptance of the pollution level due to the ship activities, voyage, anchoring and port operations. Additionally, since the ships pass through different exclusive economic zones, different national governmental rules and constraints must be accounted for. The parameters used as bases of the present analysis relied on the data provided by the global Automatic Identification System, which may present inaccurate data in some cases. For this reason, data from port monitoring and control for ship traffic needs to be used to calibrate the global date and enhance the input date of the analysis in the future.

ACKNOWLEDGEMENTS

This first author has been supported by the Strategic Research Plan of the Centre for Marine Technology and Ocean Engineering, financed by the Portuguese Foundation for Science and Technology (Fundação para a Ciência e Tecnologia-FCT). The first author has

been funded by the Portuguese Foundation for Science and Technology (Fundação para a Ciência e Tecnologia - FCT) under contract UIDB/UIDP/00134/2020.

This work was performed within the Technical University of Varna Research Plan, funded by the State Budget under the contract NP13 for 2023.

REFERENCES

[1] EU, 2015. REGULATION (EU) 2015/757 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of April 29 2015, on the monitoring, reporting and verification of carbon dioxide emissions from maritime transport and amending Directive 2009/16/EC. Official Journal of the European Union, L 123:156-76.

[2] EMSA\THETIS-MRV. CO_2 emission report. https://mrv.emsa.europa.eu/#public/emission-report

[3] EUROSTAT. Greenhouse gas emissions by source sector. https://ec.europa.eu/eurostat/databrowser/view/ENV_AIR_GGE/default/table?lang=en

[4] Spiraj, S., 2020. EU MRV Analysis of 2018 Data on CO_2 Emissions from Maritime Transport. LinkedIn.

[5] Georgiev, P, Naydenov, L, Garbatov, Y., 2022. Carbon emissions from container shipping in the Black Sea. *Sustainable Development and Innovations in Marine Technologies*, 85-92. https://doi.org/10.1201/9781003358961-12

[6] Georgiev, P, Garbatov, Y., 2021 Multipurpose vessel fleet for short black sea shipping through multimodal transport corridors. *Brodogradnja*, 72(4), 79-101. https://doi.org/10.21278/brod72405

[7] MarineTraffic. https://www.marinetraffic.com/en/ais/home/centerx:-12.0/centery:25.0/zoom:4

[8] Equasis. https://www.equasis.org/EquasisWeb/public/HomePage

[9] EU, 2023. Regulation (EU) 2023/957 of the European Parliament and of the Council of May 10 2023, amending Regulation (EU) 2015/757 in order to provide for the inclusion of maritime transport activities in the EU Emissions Trading System and for the monitoring, reporting and verification of emissions of additional greenhouse gases and emissions from additional ship types (Text with EEA relevance). Official Journal of the European Union, L 130,105–14.

[10] RESOLUTION MEPC.278(70) (Adopted on October 28 2016) Amendments to MARPOL Annex VI (Data collection system for fuel oil consumption of ships), 2016.

[11] Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A*, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

[12] Jolliffe, I.T., 2002. Principal Component Analysis. 2nd edition, Springer.

[13] Kawashima, S., Kawamura, Y., Itoh, H., Fukuto, J. 2016. Characterization of ship traffic flow by Principal Component Analysis of AIS data and its application to ship traffic simulation for evaluation of encounter probability. *Journal of the Japan Society of Naval Architects and Ocean Engineers*, 23(0), 55-63. https://doi.org/10.2534/jjasnaoe.23.55

[14] Li, Z., Wang, L., Piao, W., Jia, H., Dong, S., Zhang, J. 2022. Prediction of Operation Time of Container Ship at Berth under Uncertain Factors Based on a Hybrid Model Combining PCA and ELM Optimised by IPSO. *Journal of Marine Science and Engineering*, 10(12). https://doi.org/10.3390/jmse10121919

[15] Perera, L.P., Mo, B. 2016. Marine Engine Operating Regions under Principal Component Analysis to evaluate Ship Performance and Navigation Behavior. IFAC-PapersOnLine 49-23, 512–517. https://doi.org/10.1016/j.ifacol.2016.10.487

[16] Yu, D., Wang, L. 2018. Hull Form Optimization with Principal Component Analysis and Deep Neural Network. arXiv:181011701v1.

[17] Helmsmar. Prinicipal Component Analysis & Ship hull optimisation 2021. https://www.helmsmar.com/2021/08/01/pca-and-ship-hull-optim/

[18] Hair, J., Anderson, R., Tathan, R., Black, W. 1998. Multivariate data analysis. New York, Pearson.

[19] Solanas, A., Manolov, R., Leiva, D. 2011 Retaining principal components for discrete variables. Anuario de Psicología, Facultat de Psicologia, Universitat de Barcelona. 41(1-3), 33-50.

[20] Field, A. 2015. Discovering statistics using IBM SPSS statistics. London.

[21] Merenda, P.A. 1997. Guide to the Proper Use of Factor Analysis in the Conduct and Reporting of Research: Pitfalls to Avoid. *Measurement and Evaluation in Counseling and Development*, 30, 156-64. https://doi.org/10.1080/07481756.1997.12068936

[22] Kaufman, L., Rousseeuw, P.J. 1990. Finding Groups in Data: An Introduction to Cluster Analysis. New York, John Wiley.

[23] Coe, D. 2009. Fisher Matrices and Confidence Ellipses: A Quick-Start Guide and Software. *ArXiv: Instrumentation and Methods for Astrophysics*, 10/09/06.

[24] Rocchi, M., Sisti, D., Ditroilo, M.C., Panebianco, R. 2005. The misuse of the confidence ellipse in evaluating statokinesigram. *Italian Journal of Sport Sciences*, 169-172.

Submitted: 22.06.2023.
Accepted: 08.08.2023..

Yordan Garbatov, yordan.garbatov@tecnico.ulisboa.pt
Centre for Marine Technology and Ocean Engineering (CENTEC), Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Petar Georgiev, petar.ge@tu-varna.bg
Technical University of Varna, Bulgaria