PRELIMINARY BULBOUS BOW DESIGN TOOL APPLYING K NEAREST NEIGHBOURS CLASSIFICATION AND REGRESSION MODEL

(Reference No: IJME698, DOI No: 10.5750/ijme.v163iA3.806)

A Grech La Rosa, **E Anderlini** and **G Thomas,** University College London, UK

KEY DATES: Submitted: 20/01/21; Final acceptance: 06/08/21; Published 16/11/21

SUMMARY

Designing bulbous bows for ships remains a challenging task. Their impact on different design attributes as well as their change in performance when operating off their intended design condition renders this as a multidimensional problem. This paper explores the application of machine learning techniques to a sample of in-service vessel data to develop a preliminary design tool. The ships' data was analysed together with their bulbous bow data to generate machine learning models using a supervised approach. The K Nearest Neighbours Classifier and Regression models were used as the basis of the tool. Together, these models can be used to predict whether to install a bulbous bow and the recommended dimensionless coefficients for new vessels. Generating this preliminary bulbous bow design tool required the introduction of new dimensionless coefficients that discretise the bulbous bow's longitudinal section. The preliminary design tool gives the designer the ability to determine whether a bulbous bow should be fitted and, if so, to obtain an initial estimate of the bulbous bow required for the vessel being designed, based on key input parameters that relate to the ship and its operation. The new design tool is demonstrated to provide preliminary design details for bulbous bows through the case studies.

NOMENCLATURE

A	Administrative
CFD	Computational Fluid Dynamics
IHS	Information Handling Services
KNN	K Nearest Neighbours
MAE	Mean Absolute Error
NYK	Nippon Yusen Kabushiki
PF	Performance Features
QQ	Quantile - Quantile
SP	Ship Particulars
SRP	Ship Route Prediction
TP	Taxonomic Parameters
$\frac{B_{Moulded}}{D_{Moulded}}$	Beam to Depth Ratio (-)
$\frac{B_{Moulded}}{T_{Design}}$	Beam to Draught Ratio (-)
$\frac{L_{BP}}{B_{Moulded}}$	Length to Beam Ratio (-)
$\frac{L_{BP}}{D_{Moulded}}$	Length to Depth Ratio (-)
C_{AR}	Aspect Ratio Coefficient (-)
$C_{HBottom}$	Ratio of the bulb's bottom height to the bulb's total height (-)
$C_{HFairing}$	Ratio of the bulb's fairing height to the bulb's total height (-)
C_{HFlat}	Ratio of the bulb's top flat height to the bulb's total height (-)
C_{HTop}	Ratio of the bulb's top height to the bulb's total height (-)
$C_{LFairing_{BP}}$	Ratio of the bulb's fairing length to the Vessel's L_{BP} (-)
$C_{LFairing_{Max}}$	Ratio of the bulb's fairing length to the bulb's L_{Max} (-)
$C_{LFLAT_{BP}}$	Ratio of the bulb's flat proportion over the vessel's L_{BP} (-)
$C_{LFLAT_{Max}}$	Ratio of the bulb's flat proportion to the bulb's L_{Max} (-)
C_{LPR}	Ratio of the bulb's L_{Max} to the ship's length between perpendiculars (-)
C_{Top}	Ratio of the bulb's top draught to the Vessel's draught (-)
C_T	Ratio of the bulb's draught centre point to the Vessel's draught (-)
H_{Bottom}	Vertical height from the base of the bulbous bow to the bottom of H_{Flat} (m)
$H_{Fairing}$	Height from the keel line to the base of the bulb (m)
H_{Flat}	Vertical height of the flat proportion of the bulbous bow (m)
H_{Top}	Vertical height from the base of the bulbous bow to the top of H_{Flat} (m)
H_{Total}	Total bulb height at the base of the bulb (m)
k	Number of neighbours (-)
L_{BP}	Ship's length between perpendiculars (m)
$L_{Fairing}$	Length of bulb fairing (m)
L_{Flat}	Length of the top flat proportion of the bulbous bow (m)
L_{Max}	Longitudinal distance between the maximum longitudinal offset from the forward perpendicular (m)
T_{Centre}	Draught to the bulb's central point on H_{Flat} (m)
T_{Design}	Ship's design draught (m)
T_{Top}	Head of water present above the base of the bulb (m)

1. INTRODUCTION

A ship's hullform is a complex composition of curves and contours. Its final shape has a major impact on the ship's energy performance, payload carrying capacity, sea keeping and manoeuvring ability. When a vessel's hullform is modified it has an effect on the whole design that is not straight forward to recognise or quantify. With multiple design activities intertwined to ensure the final design is safe and feasible, iterative procedures are commonly used to enable convergence to a solution that satisfies all the necessary requirements for that particular vessel. The desire to have optimised vessels that perform well in a range of different operating conditions, such as multiple speeds, adds another layer of complexity to the design challenge.

Bulbous bows are a widely-used feature in hullform design aimed at reducing a vessel's wave-making resistance. They are expected to generate a wave forward of the bow that interacts with the vessel's wave system, resulting in destructive interference (Kracht, 1978; Larsson, 2010). This interaction reduces the wave-making component of the vessel's total resistance typically by up to 5% but more generous savings of up to 15% have also been recorded (Tsakalakis and et al., 2014).

In addition to the wave making component, bulbous bows can also be designed to impact the hullform's viscous resistance. This can be done by extending the hullform's length to facilitate a better flow transition between the hullform's main body and the bow (Kracht, 1978). Such bulbous bows are therefore typically found on vessels having block coefficients larger than approximately 0.7, typically on tankers and bulk carriers.

Designing a bulbous bow is a multidimensional problem and despite their extensive use, bulbous bows are still challenging to design for a particular vessel's operating profile. Their shape is dependent on several attributes such as resistance, hydrostatics and seakeeping. Should the design of the bulbous bow be comprehensive enough to factor in these attributes, their in-service performance is influenced with the way the vessel is operated. Factors such as the vessel's loading condition, speed and operational environment are all known to alter the bulbous bow's performance (Grech La Rosa *et al.*, 2015). Despite their importance, the tools currently available to design bulbous bows are limited and few sources provide assistance for their early stage design (Kracht, 1978; R Sharma and Sha, 2005).

Whilst vessels are expected to perform well in different scenarios with varying operating conditions (Banks *et al.*, 2013; Christensen *et al.*, 2018), a bulbous bow has a fixed, rigid shape. Therefore, its contribution to resistance reduction will not be constant given the change in conditions it experiences. Companies such as Maersk (Jonathan, no date; Klimt-Møllenbach, no date; Cerup-Simonsen *et al.*, 2009), Wilh Wilhelmsen and Nippon Yusen Kabushiki Kaisha's (NYK) (*New Bulb to Save Fuel*, 2014; Tolstrup, 2017) have replaced some of their vessel's bulbous bows with alternatives to suit a new operating profile, i.e. reduced speed. Such measures show the importance of having a bulbous bow appropriately designed for a particular vessel and operating condition.

One of the principal resources is still the publication by Kracht (1978), which gives guidance on designing bulbous bows focusing on their wave-making performance. This work involved systematically altering the shape of the bulbous bow and measuring the changes in vessel resistance in towing tank experiments. Whilst this work is invaluable, a desire to have a more accessible means of predicting a bulbous bow's expected performance exists amongst industry stakeholders.

One way of achieving this is by Computational Fluid Dynamics (CFD) analysis. This route is typically taken to carry out in-depth assessment of a bulbous bow's flow features to determine whether it is performing as expected (Atreyapurapu, Tallapragada and Voonna, 2014; Park *et al.*, 2015; Hakan Ozdemir *et al.*, 2016). In order to be able to conduct this analysis, the designer must already have a geometric model of both the hullform and the bulbous bow. The baseline bulbous bow model can then be modified in an attempt to improve its performance. This could be done manually or by using other bespoke software or algorithms to facilitate a more mathematical approach to the problem (Renaud and Berry, 2013).

In practice, bulbous bows are typically designed using a combination of these two methods, increasing the complexity of the approach as the design matures. Although not explicitly published, it is known that hullform designers extensively use their professional judgement and utilise previously designed vessels as case studies when defining new designs. The methods employed in each company to carry out this analysis are commercially sensitive protocols. Such procedures may aid the designer in identifying the most effective bulbous bow from the set being considered, but not necessarily the most effective bulbous bow for that vessel.

Finding solutions to multidimensional optimisation problems is a challenge that is experienced in many fields. Different data analysis methods could be used to address these tasks, which use recorded data to predict a desired outcome based on predefined input parameters. These approaches vary according to the volume of data available and whether the prediction is related to classification for categorical outputs or regression for numerical outputs. Machine learning solutions can be supervised, unsupervised, semi-supervised or use reinforcement learning (Sarker, 2021). These strategies differ according to how the data is used to generate a prediction. Supervised learning creates a mapping between the input and output variables using the data source's matched input and output values. If data is unlabelled, unsupervised learning techniques map data according to similarities, clusters and patterns without human intervention. Datasets having partially labelled data can benefit from semi-supervised learning which is a combination of two. Reinforcement learning on the other hand uses models that learn directly from the environment that they are interacting with, and are thus most suited to control applications. K Nearest Neighbours (KNN) Classifier, Linear Support Vector Classification (SVC) and Naïve Bayes are examples of classification models, whilst Stochastic Gradient Descent (SGD) Regressor, Lasso and Elastic Net are examples of models used for regression analysis.

Data science is being widely used in various industries such as business (Bose and Mahapatra, 2001), health care (Norgeot, Glicksberg and Butte, 2019), aviation (Zhang and Mahadevan, 2019) and manufacturing (Wuest *et al.*, 2016). It is also being used in the shipping industry, for example Cui *et al* (2012) applied the Atkinson and Shiffrin model, as well as the memory concept by Baddeley to aid the optimisation process of their algorithm through deep learning. They carried out a case study on the structural optimisation of a bulk carrier to highlight the benefits of using such an approach. The model catered for important attributes related to a ship's structure and recognised the various limitations imposed by different regulating bodies. The results show an 8.15% reduction in structural weight. Other applications include the work carried out by Trodden (Trodden *et al.*, 2015) showing how data analysis can be used to monitor and predict a vessel's fuel usage for energy efficient operations, whilst Duca *et al* (2017) used the K Nearest Neighbour classifier (KNN) to generate a Ship Route Prediction (SRP) algorithm.

Machine learning methods have also been applied to the development of preliminary design tools. In the study carried out by Cepowski (2017), the data from a fleet of container ships was used to predict the main engine power of newly proposed container ships by means of the linear regression equations plotted through the data used. Abramowski (2010) applied neural network theory to established empirical equations from experiments. The input parameters to generate the result set were retrieved from available vessel databases. The neural network consisted of five input parameters and made use of the Levenberg-Marquard learning algorithm. Gaspar (2019) proposed the increased use of data driven analysis for ship design applications.

Given the complexity involved in bulbous bow design, this study aims to apply data science approaches to aid the designer in determining whether a bulbous bow should be considered for the vessel being designed and, if so, what longitudinal profile it should have. In this work, data analysis combined with statistical examination and data visualisation is used to suggest initial bulbous bow dimensionless coefficients from a collection of in-service vessels, thus developing a tool that aids early-stage bulbous bow design.

The geometries of bulbous bows of different vessels are scrutinised to update the characterisation of bulbous bows. New variables pertaining to different bulbous bow dimensions are recorded to define the shape of a bulbous bow by means of new dimensionless coefficients. Spatial plots are used to illustrate the results from an initial study to identify trends between the different parameters considered. Statistical methods are then used to provide evidence of any features that may aid in generating repeatable results of the bulbous bow dimensionless coefficients.

2. GEOMETRY OF BULBOUS BOWS

A clear definition of a bulbous bow is first required. A bulbous bow is a bulging formation in the vessel's structure at the forward end of the vessel that typically extends beyond the forward perpendicular.

The standard measurements typically used for bulbous bows were defined by Kracht (1978). These measurements define the primary parameters that influence a bulbous bow's shape when the vessel is sitting at its design waterline. Since then, the number of differing bulbous bows has increased, some having distinctly dissimilar shapes to the ones considered by Kracht. The bulbous bow's aspect ratio as well as the longitudinal profile's curvature vary greatly. These differences impact the bulbous bow's performance making each style of bulbous bow more appropriate to different ship types having different requirements.

This study builds on the work of Kracht (1978) and introduces a series of new variables, measurements and dimensionless coefficients. By decomposing the main lines that form the shape of the bulb, a more extensive representation of modern-day bulbous bows is proposed that caters for a larger variety of bulbous bows. Unconventional bulbous bows that do not exhibit typical longitudinal profiles, such as those found on ice breaking ships or special purpose vessels, are outliers to the new representation.

These new dimensionless coefficients allow the additional geometric features of modern bulbous bows to be captured. This is principally required since the selection of bulbous bow longitudinal profiles has increased since Kracht's study. Additional parameters to discretise the bulbous bow's transverse section along the bulbous bow's length should be considered in future work.

The reference lines used in the analysis are the design waterline and the forward perpendicular. These two lines act as a means of sectioning off the bulbous bow from the rest of the hullform. The proportion of the forward perpendicular that intersects the hull's bow will be referred to as the base of the bulbous bow.

Figure 1 shows a marked-up sketch of the newly proposed measurements. Table 1 records each of the variables labelled on the diagram together with a brief description of what they represent.

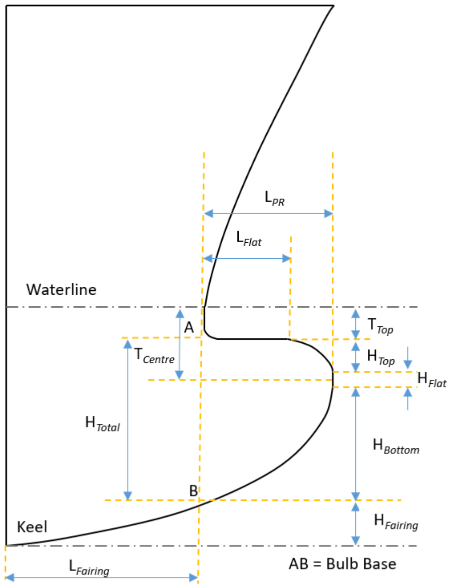


Figure 1 - Diagram of a bulbous bow with the proposed dimensions

Table 1 - Proposed bulbous bow dimensions

Variable	Description
L_{Max}	The longitudinal distance between the maximum longitudinal offset from the forward perpendicular
H_{Total}	The total bulb height at the base of the bulb
L_{Flat}	The length of the top flat proportion of the bulbous bow
$L_{Fairing}$	The length of bulb fairing
T_{Top}	The head of water present above the base of the bulb
T_{Centre}	The draught to the bulb's central point on H_{Flat}
H_{Flat}	The vertical height of the flat proportion of the bulbous bow
H_{Top}	The vertical height from the base of the bulbous bow to the top of H_{Flat}
H_{Bottom}	The vertical height from the base of the bulbous bow to the bottom of H_{Flat}
$H_{Fairing}$	The height from the keel line to the base of the bulb

The fairing of the main hullform into the bulbous bow is also important (R Sharma and Sha, 2005a; Larsson, 2010). This fairing is defined as the proportion of the bow arrangement that streamlines the bulbous bow aft of the forward perpendicular to the vessel's keel.

Whilst a bulbous bow may not necessarily be totally submerged, the complete bulbous bow profile must still be accounted for. As the waterline changes during the vessel's operating profile, some of its dimensionless coefficients will also change resulting in a change in the bulb's performance. The base of the bulb though remains constant. By introducing new parameters to discretise a bulbous bow, new dimensionless coefficients may be defined that allow the similarities or differences between bulbous bows to be identified. These have been recorded in two sets to reflect the fact that some of these dimensionless coefficients would vary as the waterline changes.

The fixed set of dimensionless coefficients is the larger of the two. It comprises of Kracht's bulbous bow overall length ratio as can be seen in Equation 1 as well as the new dimensionless coefficients being defined here.

$$C_{LPR} = \frac{L_{Max}}{L_{BP}} \quad \text{Eq1}$$

Equation 2 shows the bulbous bow aspect ratio coefficient which is the ratio of the bulbous bow's maximum length to its total height. This coefficient can be used to distinguish between bulbous bow types.

$$C_{AR} = \frac{L_{Max}}{H_{Total}} \quad \text{Eq2}$$

Equation 3 and Equation 4 relate the top flat proportion of the bulbous bow to the ship's length between perpendiculars and the bulbous bow's maximum length. Should the flat proportion not be parallel to the waterline, the angle between the waterline and top of the bulbous bow, μ, was proposed as a parameter. A positive μ represents a bulbous bow having its nose pointing down.

$$C_{LFlat_{BP}} = \frac{L_{Flat}}{L_{BP}} \quad \text{Eq3}$$

$$C_{LFlat_{Max}} = \frac{L_{Flat}}{L_{Max}} \quad \text{Eq4}$$

The changes in bulbous bow longitudinal profile over its height have been recorded by means of Equations 5 to Equation 7 which express the top, flat and bottom proportions of the bulbous bow as a fraction of the total height. Combined, these dimensionless coefficients can be used to understand how the bulbous bow can be subdivided along its height.

$$C_{HTop} = \frac{H_{Top}}{H_{Total}} \quad \text{Eq5}$$

$$C_{HFlat} = \frac{H_{Flat}}{H_{Total}} \quad \text{Eq6}$$

$$C_{HBottom} = \frac{H_{Bottom}}{H_{Total}} \quad \text{Eq7}$$

The length and height of the fairing are discretised by means of Equation 8 to Equation 10. These dimensionless coefficients express the fairing's length and height to the ship's length between perpendiculars as well as the bulbous bow's maximum length and height.

$$C_{LFairing_{BP}} = \frac{L_{Fairing}}{L_{BP}} \quad \text{Eq8}$$

$$C_{Fairing_{Max}} = \frac{L_{Fairing}}{L_{Max}} \quad \text{Eq9}$$

$$C_{HFairing} = \frac{H_{Fairing}}{H_{Total}} \quad \text{Eq10}$$

The second set of dimensionless coefficients relate the bulbous bow with its loading condition. Equation 11 expresses the bulbous bow's draught as a ratio of the vessel's draught. This dimensionless coefficient is a modified version of Kracht's where the draughts being considered are defined from the design waterline instead of the keel. When analysing the bulbous bows performance over a range of operating conditions, the design waterline could be replaced with the operating waterline.

$$C_T = \frac{T_{Centre}}{T_{Design}} \quad \text{Eq11}$$

The head of water on top of the bulbous bow is factored in by means of Equation 12. Should a proportion of the bulbous bow be out of the water, a negative value is assigned to T_{Top}.

$$C_{Top} = \frac{T_{Top}}{T_{Design}} \quad \text{Eq12}$$

3. DATA

3.1 DATA SOURCE

To investigate whether data analysis techniques can be used to obtain initial estimates of bulbous bow characteristics for different ships with varying operating variables, a data sample is needed.

The only available source to collect this data was identified to be Significant Ships (Royal Institution of Naval Architects, 2012, 2013, 2014, 2015, 2016, 2017, 2018). This is an annual publication which presents a variety of vessels that are deemed to be at the forefront of the market for that particular year. As well as the ship's particulars, the publication also provides a longitudinal section of the vessel. Unfortunately, no lines plans or transverse sections that could be used to quantify the bulbous bow's offsets, sectional areas or volume are provided. In the absence of additional resources to conduct this study, this data source was deemed suitable to investigate this proof-of-concept preliminary design tool.

The publications from 2012 to 2018 were used as the basis of this work resulting in a total of 298 ships (Royal Institution of Naval Architects, 2012, 2013, 2014, 2015, 2016, 2017, 2018). The data manually recorded from these publications could be categorised into four categories, being:

1. Administrative – *A*
2. Taxonomic Parameters - *TP*
3. Ship Particulars - *SP*
4. Performance Features - *PF*

Examples of the fields used in each category are recorded in Table 2. Administrative variables helped manage data through the data manipulation stages whilst the taxonomic data was used to filter through the different types of ships. Ship particulars and performance features were essential to create the machine learning model to record trends in the data set.

Table 2 - Examples of fields used in each category

Administrative *(A)*	**Taxonomic Parameters *(TP)***	**Ship Particulars *(SP)***	**Performance Features *(PF)***
Publication Year	Ship Family	Overall Length	Design Speed
IMO Number	Ship Type	Design Draught	Total Power
Ship Name	Designer	Block Coefficient	Fuel Consumption

The bulbous bow measurements required to quantify the dimensionless coefficients were manually extracted from the drawings available. This was done by:

1. importing the ship drawing image into CAD software;
2. manually taking measurements of known reference lengths, typically the overall length;
3. marking the design draught on the drawing by scaling the recorded draught to that of the drawing;
4. measuring and recording each bulbous bow measurement;
5. automatically scaling the bulbous bow measurements taken from the drawing to the full-scale length by means of a python code.

This procedure was repeated for each vessel equipped with a bulbous bow. Those vessels without a bulbous bow only had their other variables recorded.

3.2 DATA DISTRIBUTION

The distribution of vessels with and without bulbs is 82.9% and 17.1% respectively. This clearly demonstrates the bias towards vessels being designed and built fitted with bulbous bows. Such an imbalance in the dataset could lead to inaccurate predictions favouring vessels
with bulbous bows. For the purposes of this initial study, it was decided not to artificially balance the data sets to record what the outcome would be. To see whether a vessel type had the likelihood of being equipped with or without a bulb, a histogram showing the distribution of vessel types with and without bulbs for the whole of the sample was plotted.

A total of 43 ship types were recorded with the maximum population being 55 for container ships. The sample
was too restricted for some ship types, with some having a population of one. To rectify this, a modified Information Handling Services (IHS) Statcode5 classification system as shown in Figure 2 was used to generate clusters according to Level 2, which is referred to as the *Family* (IHS Markit, 2018). This subdivision resulted in larger populations for each family with 28 passenger vessels, 90 dry cargo vessels, 39 bulk carriers, 110 tankers and 31 miscellaneous vessels. No fishing or offshore vessels were considered in this study. As can be seen from Figure 3, a more uniform distribution between the vessels is recorded.

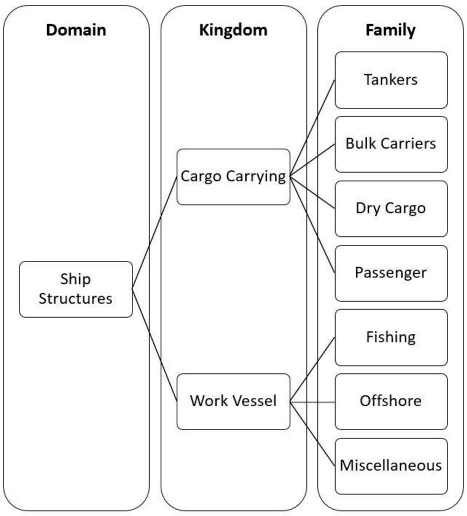


Figure 2 - Modified IHS Classification System

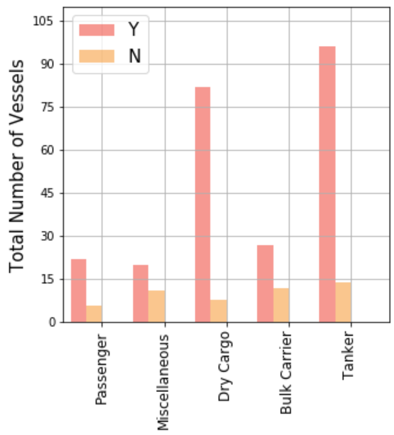


Figure 3 - Distribution of vessel families with and without bulbs

A total of 88 ship design companies were recorded in the sample, but no particular company had designed more than 9.7% of the vessels, meaning that there is not a bias towards the design methodologies employed by a particular ship designer in this study.

3.3 DATA VERIFICATION

Since the data sourced was not intended to be used for such an application, verification of data against other sources was carried out to determine the quality of the sample. Ship particulars were also cross referenced against its drawing to determine whether the data was consistent.

Where data was missing, alternative sources were consulted to populate the missing fields. If this was not possible, interpolation or the mean absolute value for vessel variables from the same ship family were used.

3.4 DATA UNCERTAINTY

Obtaining the parameters from a set of published drawings meant that there is some inherent uncertainty in the parameter values due to: line thickness, scaling issues and poor, or inconsistent, definition of the vessel design draught. Table 3 records an estimation of the percentage uncertainty associated with some of the dimensionless coefficients considered.

Table 3 - Sample of errors recorded for dimensionless coefficients

Dimensionless Coefficient	Error
C_{LPR}	5%
C_{AR}	2%
$C_{LFlat_{Max}}$	7%

4. DATA ANALYSIS

The basis of the preliminary design tool is a Python script applying data analytical techniques. To determine whether a vessel should have a bulbous bow, a supervised classification model will be modelled whilst a supervised regression model will be applied to quantify the bulbous bow dimensionless coefficients. The quality and volume of data available has a direct impact on the choice of machine learning algorithms that can be selected to develop this tool. Given the limitations of the dataset, the K Nearest Neighbour algorithm for both classification and regression was identified to be the most suitable to develop this concept. This model has also been applied to other maritime applications such as Abbasian et al's (2018) study on Offshore Support Vessel (OSV) design robustness. This model was selected for its simplicity and applicability to smaller, non-gaussian datasets (Parthasarathy and Chatterji, 1990).

The data analysis process followed the key steps as portrayed in Figure 4, with this process being repeated for each target variable considered in this study.

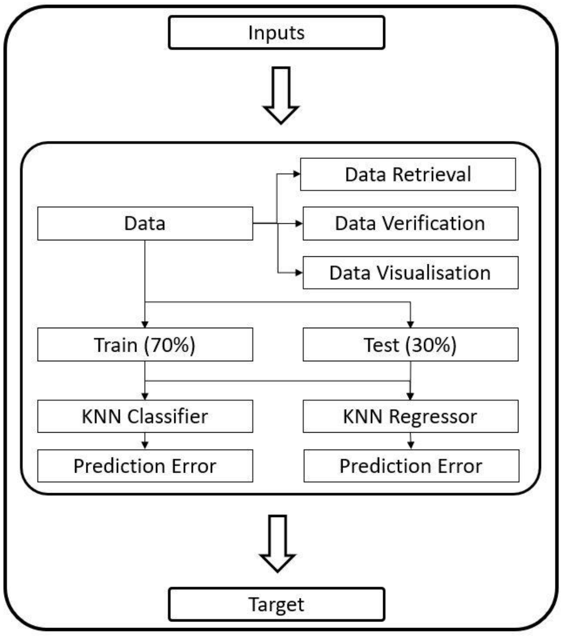


Figure 4 – Flowchart of the analysis process

4.1 DATA VISUALISATION

Two and three-dimensional scatter plots were used to identify trends between variables. The plot markers represented an additional dimension by associating each point with a ship family. Visualising the data in this manner permitted the identification of the following phenomena:

1. Supervised clusters - a cluster is made up of points representing vessels that have similar characteristics to each other but are distinct enough from others to form a subset. The clusters were defined as supervised since they were assigned to a group.
2. Overall trends - when the points plotted formed trends, curve fitting or regression analysis was carried out to predict bulbous bow dimensionless coefficients for other vessels. Trends were identified in multiple dimensions.
3. Trends within supervised clusters - when trends were found within a cluster, separate curve fitting or regression analysis could be carried out by isolating the effects of each ship family. Such trends enabled the regression analysis to be carried out for each ship family.

4.2 STATISTICS

The statistical limitations of the data set were established. The statistical distributions of the variables in the sample set were assessed by means of histograms, skewness and kurtosis, the Shapiro Wilk test and the Anderson test to determine whether any variables were Gaussian.

The relationships between variables were captured by plotting a correlation matrix of all the fields. Recognising the variables that exhibited better conformance with each other aided in selecting the inputs for the machine learning model.

The Chi Squared test was also carried out on each of the ship families to determine whether the difference between the expected and observed quantities of vessels with and without bulbous bows in each category display any relationship in the population.

4.3 OUTLIERS

The main outliers in the dataset were identified to determine whether they could have originated from any of the errors outlined in Section 3.4 and where possible corrected.

Other outliers that were identified to be factually correct were kept in the dataset. The machine learning model was left to tackle outliers instead of manually eliminating vessels from the dataset.

4.4 K NEAREST NEIGHBOURS CLASSIFICATION

A supervised learning approach was selected due to the small volume of data available. A *1* was assigned to vessels having a bulb whilst a *0* was assigned to vessels without.

The KNN Classifier was the machine learning model selected; it is an algorithm that can be used to predict the group that a set of input parameters should be assigned to. This is based on feature similarity and computing the distance between the point being queried and the *k* number of adjacent points (Goldberger *et al.*, 2004; Parsons and Scott, 2004).

The classification of the k points is then consulted to determine which group the point being investigated should be assigned to. Depending on the degree of clustering present in the dataset, the ambiguity of which subset the data point should be assigned to varies. The magnitude of *k* therefore varies according to the type and size of data available.

A study to determine the optimal k was conducted by monitoring the mean binary classification error of the predicted values to the test set for each classification model having different values of k. This was quantified by calculating the mean of a vector of booleans. A *1* is assigned when the test and prediction do not match whilst a 0 is allocated when they do. The model recording the least mean binary classification error for the least k was identified to be the optimal model for the study. Basing a prediction on a fewer number of k vessels reduces the chances of having vessels not in the same range to influence results.

Different distance metrics can be used to compute the distance between the respective points to identify the *k* points needed to carry out the computation. This study applied the Euclidean distance between each point which is used to determine the straight line distance between 2 points.

The KNN Classifier is also able to apply weighting to the points being considered. This provides the option of making the points closest to the data point being considered to have more of an impact than the other points further away. Other functions that apply weighting according to different criteria may also be custom defined. Here a uniform weighting was applied since *k* and the sample size were small.

The input parameters used to build the model were Froude number, L_{BP}/B and B/T. Although additional features would characterise the differences between vessel requirements better, KNN models are known to perform best with the least number of features, especially if the data set is small (Parthasarathy and Chatterji, 1990). The selected features were identified to be the least number of features possible to isolate the performance of each ship family. Other inputs such as block coefficient and displacement were considered but rejected since they had lower record populations.

The data was randomly split into a training set (70%) and test set (30%). The random selection of vessels for each set was inspected to ensure that a similar proportion of vessels with and without bulbs was included. Future iterations of the tool would seek to treat the imbalance in the dataset by gathering more data or by artificially augmenting the population by bootstrapping or by applying Synthetic Minority Oversampling Technique (SMOTE)(Chawla *et al.*, 2002). Cross validation was used on the training set to check whether the model was over fitting; enabling better use of the small sample being used.

A confusion matrix based on the complete sample set was conducted to capture the performance of the classification model. The accuracy, misclassification, sensitivity, false positive rate, specificity and precision were also calculated.

4.5 K NEAREST NEIGHBOURS REGRESSION

The K Nearest Neighbour approach was also used to determine the shape and size of a bulbous bow. Having the same benefits of being able to function with smaller datasets that do not necessarily follow a normal distribution, this model was identified to be one of the most valid for this proof-of-concept preliminary design tool. In a similar way to the KNN classifier, this model considers k sample points adjacent to the point being investigated to determine what its target output should be. A new model was generated for each target considered.

The attributes used to form this model are similar to the classifier. The Euclidean distance was selected to be the most appropriate means of quantifying the distance between points and no weighting was assigned to them. Once the points required to carry out the analysis were outlined, the regression analysis was carried out. The model was made by randomly selecting 70% of the sample as a training set and the remainder as the test set. This process was repeated for each target considered.

An iterative procedure was carried out to determine which variables should be used as input parameters for the models. This process was conducted in consultation with the correlation matrix. The variables that showed the best outcomes for the targets were L_{BP}/B and B/T, together with the design Froude number. All of the bulbous bow dimensionless coefficients defined earlier were set as targets.

4.6 UNCERTAINTY ANALYSIS

After creating the KNN models, the uncertainly for each model was quantified by calculating the mean absolute error (MAE) between the train and test set. This error was recorded whilst ensuring that the model was not over fitting the data. The number of inputs selected for each model was limited to maintain the error in a tolerable range.

5. RESULTS AND DISCUSSION

Trends, clusters and trends within clusters were recorded, with better conformance noted when bulbous bow dimensionless coefficients were plotted against ship particulars and performance features.

5.1 TWO-DIMENSIONAL SCATTER PLOTS

A sample of the two-dimensional plots is shown in Figure 5 where the bulbous bow aspect ratio dimensionless coefficient is plotted against the vessel's Froude number. The plot shows that as the Froude number increases, the shape of the bulb becomes longer and slenderer. This observation is in good conformance with wave theory. This plot shows supervised clustering between tankers and dry cargo vessels. The bulbous bow aspect ratio dimensionless coefficient for bulk carriers and tankers are similar with some over spill into the dry cargo regime. Passenger vessels do not cluster as well as other ship families but do inhabit a distinct region. Increasing the sample's population of passenger vessels could result in the formation of a clear cluster. The miscellaneous family consists of a variety of ship types that have differing characteristics resulting in no distinct clusters being recorded.

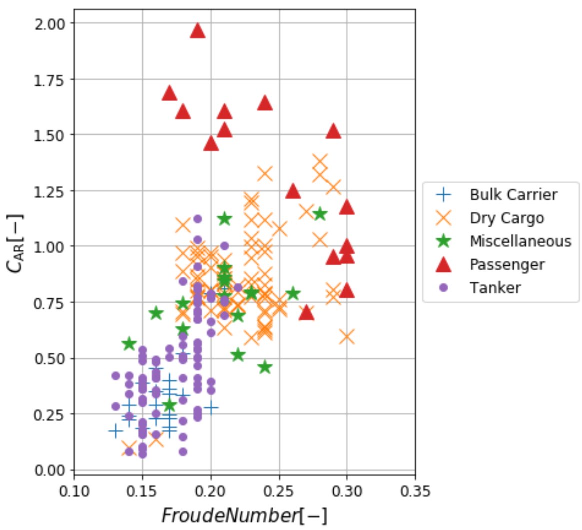


Figure 5 - Scatter plot C_{AR} against Fn

5.2 STATISTICS RESULTS

The ship and bulbous bow variables investigated did not follow a normal distribution. This was concluded after analysing the results of the statistical tests. Table 4 shows a sample of the statistical results recorded, showing p values of 0 for all three variables. This was one of the key reasons why a non-parametric method was selected as the basis of the tool. To simplify the model's complexity, some variables were discounted from the study to stop them influencing the rest of the analysis. Such variables include gross tonnage and deadweight.

Table 4 - Sample of Statistical results recorded for Input variables.

Variable	Skewness	Kurtosis	Shapiro statistic	Shapiro p value
L_{BP}/B	3.62	31.40	0.75	0.00
B/T	2.38	10.57	0.81	0.00
Fn	2.17	17.64	0.86	0.00

Figure 6 shows an example of the histograms recorded. In this case, the distribution recorded is not strictly gaussian but does show the correct trend. Increasing the sample size could improve this distribution.

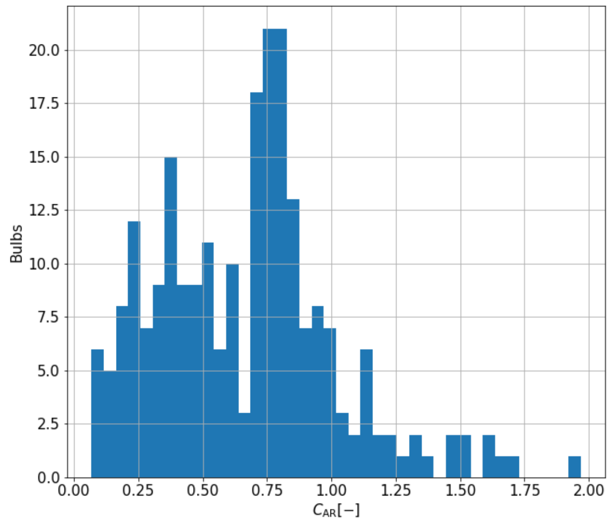


Figure 6 – Histogram of C_{AR}

A correlation matrix was plotted to understand which variables conformed more strongly to each other.

Figure 7 shows the strongest correlations are recorded between the vessel's Froude number, block coefficient and the bulbous bow dimensionless coefficients. These correlations can be confirmed with the two- and three-dimensional plots highlighted earlier.

The Chi Squared test conducted showed dependency between vessels with and without bulbous bows. The null hypothesis was therefore rejected.

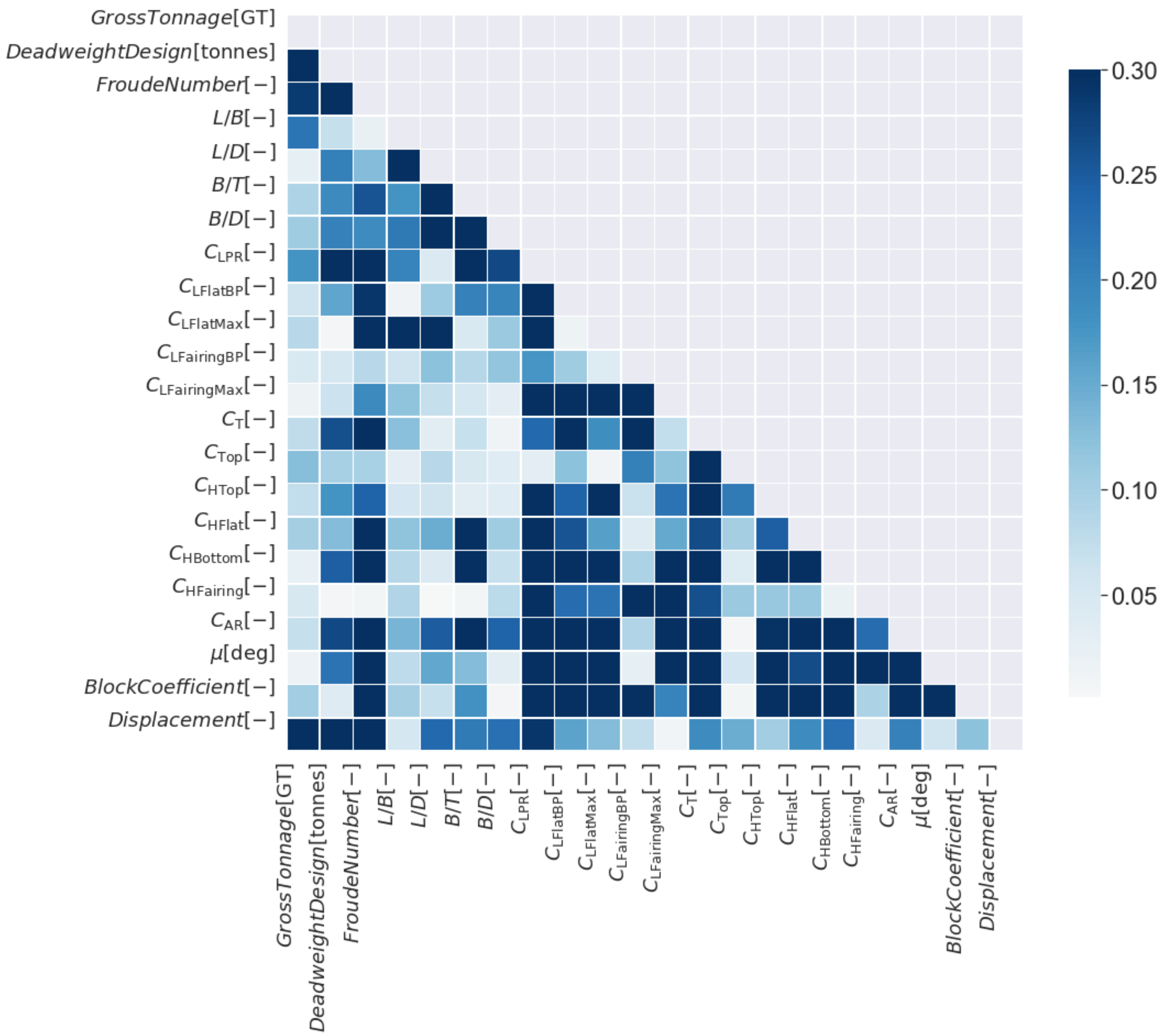


Figure 7 - Correlation Matrix for complete sample set

5.3 KNN CLASSIFICATION RESULTS

The classification model generated can predict whether a vessel should or should not have a bulbous bow fitted to it. A study to determine the optimal k value for the KNN Classification model by monitoring the mean binary classification error was carried out. Figure 8 shows the plot of the mean binary classification error against k that shows that a k value of 5 is the smallest value that would generate the least error.

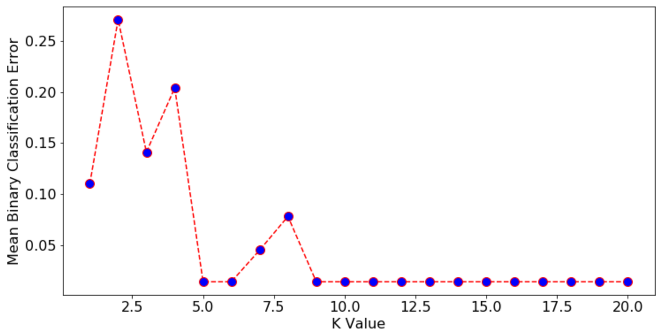


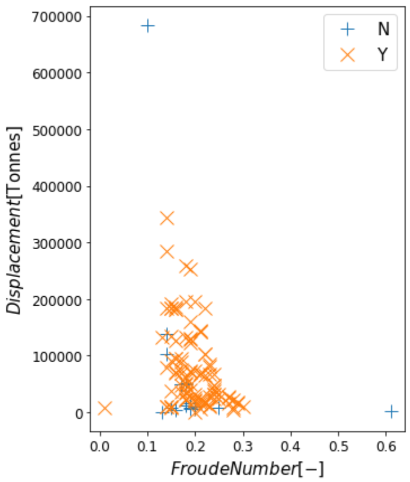
Figure 8 - Mean binary classification error against k value for different classification models

Table 5 records the results of the confusion matrix together with the evaluation metrics used to understand the machine learning model's performance. The matrix demonstrates that the model was able to predict whether a vessel should or should not have a bulbous bow fitted based on the three features selected with an accuracy of 87.2%. The model was able to predict 243 true positives with a precision of 87.4%. These results demonstrate that a supervised learning approach as outlined in Section 1 produced good results for this application.

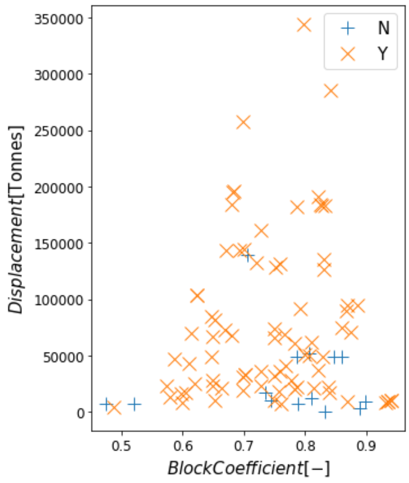
Table 5 - Confusion Matrix for complete sample

		Predicted	
	n=297	**0**	**1**
Actual	**0**	16	35
	1	3	243
Accuracy		87.2%	
Misclassification		12.8%	
Sensitivity		98.8%	
False Positive Rate		68.6%	
Specificity		31.4%	
Precision		87.4%	

Figure 9 shows how the population of vessels within the sample predominately have bulbous bows. The vessels do not form distinct subsets of vessels with or without bulbous bows. Improving the size and quality of the data together with implementing additional features would improve the tool's performance. As highlighted previously, the data could also be artificially modified to reduce the bias on vessels with bulbous bows.



(A)



(B)

Figure 9 – Scatter plots showing the distribution of vessels with and without bulbous bows.

5.4 KNN REGRESSION RESULTS

A machine learning model for each of the bulbous bow dimensionless coefficients was successfully generated. Based on the inputs submitted, the model suggests an initial bulbous bow parameter to create the bulbous bow section. The case studies presented later in this paper demonstrates how these results can be utilised.

5.5 UNCERTAINTY RESULTS

A sensitivity study was carried out to determine whether the sample size was sufficient to reliably predict bulbous bow parameters. By varying the number of vessels used to generate the machine learning model, the error between the predicted bulb parameter and the actual bulb parameter was calculated.

Figure 10 shows the standardised mean absolute error results for the sensitivity study. The errors plotted show that whilst certain errors relating to some bulb parameters plateaued, others had not and would benefit from more vessels. Based on the study carried out, it was decided to use all vessels in the dataset for the tool.

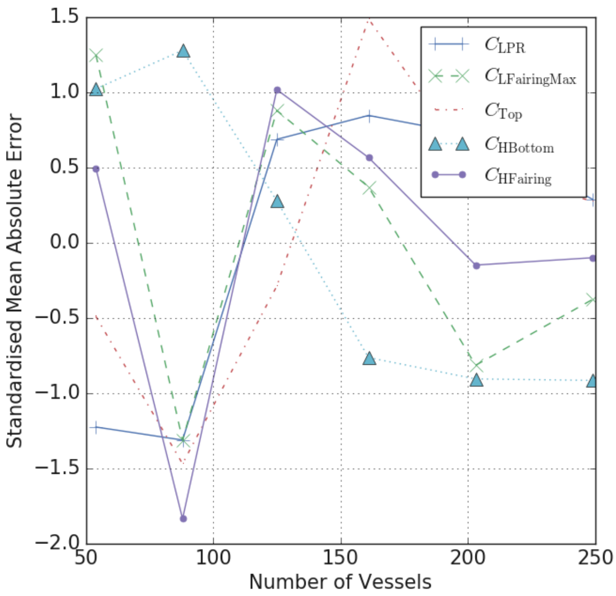


Figure 10 – Standardised mean absolute error for sensitivity study

Table 6 records a selection of some of the errors quantified for each dimensionless coefficient. For the purposes of this study, this error was deemed acceptable.

Table 6 - Error quantification for a selection of the machine learning algorithms created

Variables	Mean Absolute Error	Mean Square Error
C_{LPR}	5%	5%
C_{AR}	2%	2%
$C_{LFlat_{Max}}$	7%	7%

The percentage error for each predicted variable was not the same. This is due to the difference in spatial spread that each variable has. If the similarity between one bulbous bow and the other for a set of inputs was higher, the calculated error was less.

Should additional inputs be used to predict the same target variables listed, with the same restricted dataset, the error in the prediction will increase. This means that for the tool to cater for other input variables, such as block coefficient and length to displacement ratio, the sample size must increase.

By addressing the data extraction errors outlined in Section 3.4, the quality of the data would improve and the quality of the results would also improve. This means that increasing the sample size is not the only approach for reducing the uncertainty.

6. CASE STUDIES

To demonstrate the performance of the bulbous bow initial design tool, the ship and bulbous bow data was extracted from Significant Ships 2019 (Royal Institution of Naval Architects, 2019) following the same procedures outlined earlier. The details of these vessels were used as the input parameters in the model, whilst testing was enabled by knowing whether a bulbous bow was fitted to the vessel and, if so, what its profile was. A total of 43 vessels were recorded from this publication but the case studies will focus on four vessels. Table 7 records the details of these vessels where there first two vessels do not have a bulbous bow fitted and the remaining two have bulbous bows.

Table 7 – Case studies input parameters

Vessel	Fn	L_{BP}/B	B/T	Bulb?
Vessel 1	0.17	9.06	1.80	N
Vessel 2	0.13	5.63	2.66	N
Vessel 3	0.19	6.86	3.52	Y
Vessel 4	0.15	5.54	3.22	Y

The KNN Classifier was first applied to determine whether these vessels should have a bulbous bow installed. Table 8 shows the results of the confusion matrix using all the data recorded from Significant Ships 2019. The matrix and data metrics recorded show a similar performance to what was achieved with the complete sample set.

Table 8 - Confusion Matrix for sample of vessels from 2019

		Predicted	
	n=43	**0**	**1**
Actual	**0**	5	6
	1	1	31
Accuracy	83.7%		
Misclassification	16.3%		
Sensitivity	96.9%		
False Positive Rate	54.5%		
Specificity	45.5%		
Precision	83.8%		

The tool recommended that no bulbous bows should be fitted to Vessel 1 and Vessel 2 which conforms to the ship's status. The tool also recommended that Vessel 3 and Vessel 4 should have bulbous bows fitted, which also conforms with the ship’s geometry.

The KNN Regressor was then applied to predict the bulbous bow dimensionless coefficients for Vessel 3 and Vessel 4. These, together with the input parameters, were subsequently used to quantify the absolute bulbous bow dimensions as well the bulbous bow’s draught.

The bulbous bow dimensions were first used to manually generate a grid. This grid acted as the basis to build the longitudinal section of the bulbous bow. The uncertainty in the model was captured by drawing an additional two grids representing the upper and lower limits of the model, as is illustrated by means of the dashed lines in Figure 11.

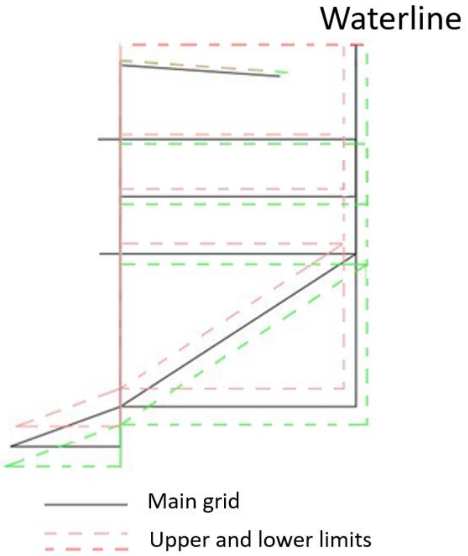


Figure 11 - Montage of Construction grids

A combination of lines and splines were then used to manually create the bulbous bow profile. The intersection points of the main grid acted as the control points whilst the upper and lower limits grid acted as the boundaries for the curves to stay within.

Figure 12 shows the predicted bulbous bow profiles (solid green) together with the actual bulbous bow profile (dashed black) selected for that vessel. The design waterline for each of the vessels was also drawn and contrasted to the actual design waterline.

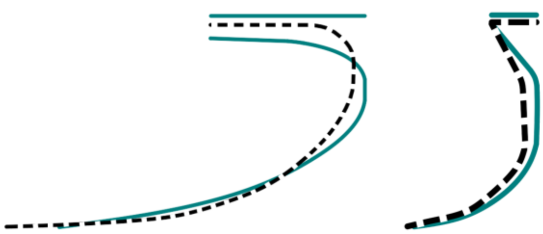


Figure 12 – Predicted bulbous bows for vessels

The profile's shape, scale and features show good conformity with the original bulbous bow profiles. The recommended design waterline was also well predicted where the design and actual waterline of Vessel 3 are superimposed on top of each other.

7. CONCLUSIONS

The development of bulbous bow design, during early-stage ship design, has been revisited in this paper by applying machine learning techniques to determine an initial estimate for a vessel's bulbous bow's design. The key steps in creating this proof-of-concept tool were:

1. developing a sample of ship and corresponding bulbous bow data;
2. developing a parametric geometric model to discretise the bulbous bow's longitudinal profile
3. carrying out a statistical analysis of the data;
4. applying the KNN classifier and regressor to the sample;
5. calculating the associated uncertainties of the models;
6. Conducting a case study to evaluate the tool's performance.

This study proposed a more comprehensive definition for bulbous bows to capture the wide variety of bulbous bows in the world's fleet. New dimensions and dimensionless coefficients were proposed to better characterise the shape of bulbous bows.

The design tool proposed is composed of two machine learning models. The first determines whether a bulbous bow should be fitted to a vessel whilst the second quantifies the bulbous bow dimensionless coefficients if the first model recommended that a bulbous bow should be fitted. These models were assessed by means of case studies that consisted of the vessels from Significant Ships 2019 with focus on four of those vessels. This concept design tool complements other procedures that could be implemented to design bulbous bows.

The results of the KNN Classification model showed good results where an accuracy of 87.2% was recorded. The bias in the dataset towards vessels without bulbous bows resulted in three false negative results from a complete sample of 297 vessels. The data from 2019 recorded a similar overall performance where one false negative was recorded from a sample of 43 vessels. Two of the four vessels, Vessel 1 and Vessel 2, were correctly recommended not to have a bulbous bow fitted whilst the remaining two were correctly assigned a bulbous bow.

Future iterations of the tool would aim to address the bias in the dataset to improve its accuracy. This could be done by having better data, artificially augmenting the population of vessels without a bulbous bow by bootstrapping or SMOTE. Additional challenges relating to the non-gaussian distribution of the data at hand is a challenge that would also to be addressed.

The KNN Regressor model was then applied to quantify the dimensionless coefficients for those vessels that were assigned a bulbous bow with focus on Vessel 3 and Vessel 4. These were used to generate an initial sketch of the longitudinal section for each vessel's bulbous bow. The sections recorded were compared to those of the actual vessel and showed good conformity.

The sample set recorded from Significant Ships was used as the basis for creating the preliminary bulbous bow design tool. This sample was deemed appropriate as a proof of concept, but a more robust dataset would be recommended for further analysis. This would enable additional input parameters to be included that would cater for the vessel's design requirements better. The model is currently limited to the bulbous bow's longitudinal section. Including data relating to the bulbous bow's transverse section would widen the scope of the tool making it possible to investigate volumetric changes in the bulbous bows shape.

Should other elements be factored into the models, predictions of the expected performance of a bulbous bow could be derived. These initial estimates could act as a catalyst in design studies to identify the relative change between different options.

Future iterations of the tool would aim to gather more ship specific data to create specialised sub tools for specific ship families or ship types. This would ensure that predicted bulbous bows having certain design requirements are not influenced by other vessels having different requirements.

Another feature that should be considered for future iterations is to automatically generate the bulbous bow profiles from the dimensionless coefficients. If the model made use of three-dimensional data, the bulbous bow geometry could be exported into other software for further analysis.

The preliminary bulbous bow design tool demonstrates that data analysis techniques can be applied to determine whether a vessel should have a bulbous bow installed and to obtain estimates on what shape and size it should be. Although this work has focused on the design of the bulbous bow, its scope could be widened to consider other vessel attributes or even the complete hullform.

8. REFERENCES

1. ABBASIAN, N. S. *et al.* (2018) *'Improving early OSV design robustness by applying "Multivariate Big Data Analytics" on a ship's life cycle'*, Journal of Industrial Information Integration. Elsevier, 10(May), pp. 29–38. doi: 10.1016/j.jii.2018.02.002.
2. ABRAMOWSKI, T. (2010) *'Application of Artificial Intelligence Methods for Improving Ship Transport Efficiency'*, Zeszyty Naukowe Akademia Morska w Szczecinie, 21(93), pp. 5–11.
3. ATREYAPURAPU, K., TALLAPRAGADA, B. and VOONNA, K. (2014) *'Simulation of a Free Surface Flow over a Container Vessel Using CFD'*, International Journal of Engineering Trends and Technology (IJETT), 18(7), pp. 334–339.
4. BANKS, C. *et al.* (2013) *'Understanding Ship Operating Profiles with an Aim to Improve Energy Efficient Ship Operations'*, in Low Carbon Shipping Conference, pp. 1–11.
5. BOSE, I. and MAHAPATRA, R. K. (2001) *'Business data mining - A machine learning perspective'*, Information and Management, 39(3), pp. 211–225. doi: 10.1016/S0378-7206(01)00091-X.
6. CEPOWSKI, T. (2017) *'Prediciton of the Main Engine Power of a New Container Ship at the Preliminary Design Stage'*, Management Systems in Production Engineering, 25(2). doi: 10.1515/mspe.
7. CERUP-SIMONSEN, B. *et al.* (2009) *'An integrated approach towards cost-effective operation of ships with reduced GHG emissions'*, Transactions - Society of Naval Architects and Marine Engineers, 2, pp. 1–18.
8. CHAWLA, N. V. *et al.* (2002) *'SMOTE: Synthetic Minority Over-sampling Technique'*, Journal of Artificial Intelligence Research, 16, pp. 321–357. doi: 10.1002/eap.2043.
9. CHRISTENSEN, L. B. R. *et al.* (2018) *'The Effect of Performance Operational Factors on Container Ship Fuel'*, in Full Scale Ship Performance. London: Royal Institution of Naval Architects.
10. CUI, H., TURAN, O. and SAYER, P. (2012) *'Learning-based ship design optimization approach'*, Computer Aided Design. Elsevier Ltd, 44(3), pp. 186–195. doi: 10.1016/j.cad.2011.06.011.
11. LO DUCA, A., BACCIU, C. AND MARCHETTI, A. (2017) *'A K-Nearest Neighbor Classifier for Ship Route Prediction'*, in OCEANS 2017 - Aberdeen. Aberdeen: IEEE, pp. 1–6. doi: 10.1109/OCEANSE.2017.8084635.
12. GASPAR, H. M. (2019) *'A Perspective on the Past, Present and Future of Computer-Aided Ship Design'*, in 18th Conference on Computer and IT Applications in the Maritime Industries. Technische Universität Hamburg-Harburg Tullamore.
13. GOLDBERGER, J. et al. (2004) *'Neighbourhood Components Analysis'*, in Advances in Neural Information Processing Systems 17.
14. GRECH LA ROSA, A. (UCL) *et al.* (2015) *'Bulbous Bows for Energy Efficient Ships : Towards A Novel Design Approach'*, in Energy Efficient Ships 2015. Rotterdam: Royal Institution of Naval Architects.
15. HAKAN OZDEMIR, Y. *et al.* (2016) *'A Numerical Application To Predict the Resistance and Wave Pattern of KRISO*

Container Ship', Brodogradnja, 67(2), pp. 47–65.

16. IHS MARKIT (2018) *'StatCode 5 Shiptype Coding System.'* Available at: https://cdn.ihs.com/www/pdf/Statcode-Shiptype-Coding-System.pdf.
17. JONATHAN, W. (no date) *'Maersk Line Social :: The nose job: Why 10 of our ships are getting a new bulbous bow.'* Available at: http://maersklinesocial.com/nose-job/ (Accessed: 17 April 2015).
18. KLIMT-MØLLENBACH, C. (no date) *'Retrofit - A New Bulbous Bow.'* Available at: http://messe.no/ExhibitorDocuments/191291/6889/4353-1-en Retrofit - a new bulbous bow_web.pdf?ExhibitionId=422 (Accessed: 12 December 2018).
19. KRACHT, A. M. (1978) *'Design of Bulbous Bows'*, SNAME Transactions. SNAME Transactions, 86, pp. 197–217.
20. LARSSON, L. (2010) *'Ship resistance and flow'*. Jersey City, N.J.: Society of Naval Architects and Marine Engineers (Principles of naval architecture).
21. CLIPPER GROUP (no date) *'New Bulb to Save Fuel'* (2014). Available at: http://www.clipper-group.com/about-us/news/all-news/news-archive/2014/clipper-bulb-modification-saves-fuel ACCEPT (Accessed: 17 April 2018).
22. NORGEOT, B., GLICKSBERG, B. S. AND BUTTE, A. J. (2019) *'A call for deep-learning healthcare'*, Nature Medicine, 25(1), pp. 14–15. doi: 10.1038/s41591-018-0320-3.
23. PARK, H.-S. *et al.* (2015) *'A Study on Resistance Performance for Various Trim Conditions and Bulb Shapes on a Container Ship Under Slow Steaming'*, in ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering, p. V007T06A012. Available at: https://www.researchgate.net/publication/301418838_A_Study_on_Resistance_Performance_for_Various_Trim_Conditions_and_Bulb_Shapes_on_a_Container_Ship_Under_Slow_Steaming.
24. PARSONS, M. G. AND SCOTT, R. L. (2004) *'Formulation of Multicriterion Design Optimization Problems for Solution With Scalar Numerical Optimization Methods'*, Journal of Ship Research, 48(1), pp. 61–76.
25. PARTHASARATHY, G. AND CHATTERJI, B. N. (1990) *'A Class of New KNN Methods for Low Sample Problems'*, IEEE Transactions on Systems, Man and Cybernetics, 20(3), pp. 715–718. doi: 10.1109/21.57285.
26. RENAUD, M. AND BERRY, L. (2013) *'Ship design and optimisation With Computational Fluid Dynamics.'*
27. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2012) *'Significant Ships of 2012'*, Significant Ships.
28. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2013) *'Significant Ships of 2013'*, Significant Ships.
29. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2014) *'Significant Ships of 2014'*, Significant Ships.
30. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2015) '*Significant Ships of 2015'*, Significant Ships.
31. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2016) '*Significant Ships of 2016'*, Significant Ships.
32. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2017) *'Significant Ships of 2017'*, Significant Ships.
33. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2018) *'Significant Ships of 2018'*, Significant Ships.
34. ROYAL INSTITUTION OF NAVAL ARCHITECTS (2019) *'Significant Ships of 2019'*, Significant Ships.
35. SARKER, I. H. (2021) '*Machine Learning: Algorithms, Real-World Applications and Research Directions'*, SN Computer Science. Springer Singapore, 2(3), pp. 1–21. doi: 10.1007/s42979-021-00592-x.
36. SHARMA, R AND SHA, O. P. (2005a) *'Hydrodynamic Design of Integrated Bulbous Bow/Sonar Dome for Naval Ships'*, Defence Science Journal, 55(I), pp. 21–63.
37. SHARMA, R AND SHA, O. P. (2005b) *'Practical Hydrodynamic'*, Naval Engineers Journal, pp. 57–76.
38. SHARMA, R. AND SHA, O. P. (2005) *'Practical Hydrodynamic Design of Bulbous Bows for Ships'*, Naval Engineers Journal, 117(1), pp. 57–76. doi: 10.1111/j.1559-3584.2005.tb00321.x.
39. TOLSTRUP, M. M. (2017) *'Retrofitting a New Bulbous Bow.'* Available at: https://forcetechnology.com/en/maritime-industry/cases/retrofitting-a-new-bulbous-bow.
40. TRODDEN, D. G. *et al.* (2015) *'Fuel usage data analysis for efficient shipping operations'*, Ocean Engineering. Elsevier, 110, pp. 75–84. doi: 10.1016/j.oceaneng.2015.09.028.
41. TSAKALAKIS, N. *et al*. (2014) *'Guidelines For Energy Efficient Ships.'* Edited by P. Tsichlis.

42. WUEST, T. *et al.* (2016) *'Machine learning in manufacturing: Advantages, challenges, and applications'*, Production and Manufacturing Research. Taylor & Francis, 4(1), pp. 23–45. doi: 10.1080/21693277.2016.1192517.

43. ZHANG, X. AND MAHADEVAN, S. (2019) *'Ensemble machine learning models for aviation incident risk prediction'*, Decision Support Systems. Elsevier, 116(September 2018), pp. 48–63. doi: 10.1016/j.dss.2018.10.009.

