Determination of the Compensated Gross Tonnage factors for Super Yachts

Report, Februari, 2010
Version = 9.0 Final

On behalf of: SYBAss, Mr. Theo Hooning,
By: TU Delft, J.J. Hopman, J.F.J. Pruyn, R.G. Hekkenberg

Acknowledgements

The authors would like to thank SYBAss for their trust in us in executing this project and the following yards for supporting the project and providing the information that was indispensable for performing this research:

- Royal Huisman Shipyard, The Netherlands
- Koninklijke De Vries Scheepsbouw (Feadship), The Netherlands
- Perini Navi, Italy
- Trinity Yachts, USA
- Benetti Spa., Italy
- Royal Van Lent (Feadship), The Netherlands
- Heesen Yachts, The Netherlands
- Oceanco, The Netherlands
- Vitters Shipyard, The Netherlands
- Amels, The Netherlands
- Lürssen, Germany

Further thanks go out to the University of Genoa and especially mr D. Boote for cooperating with us on this research and providing verification of the results.

Management Summary

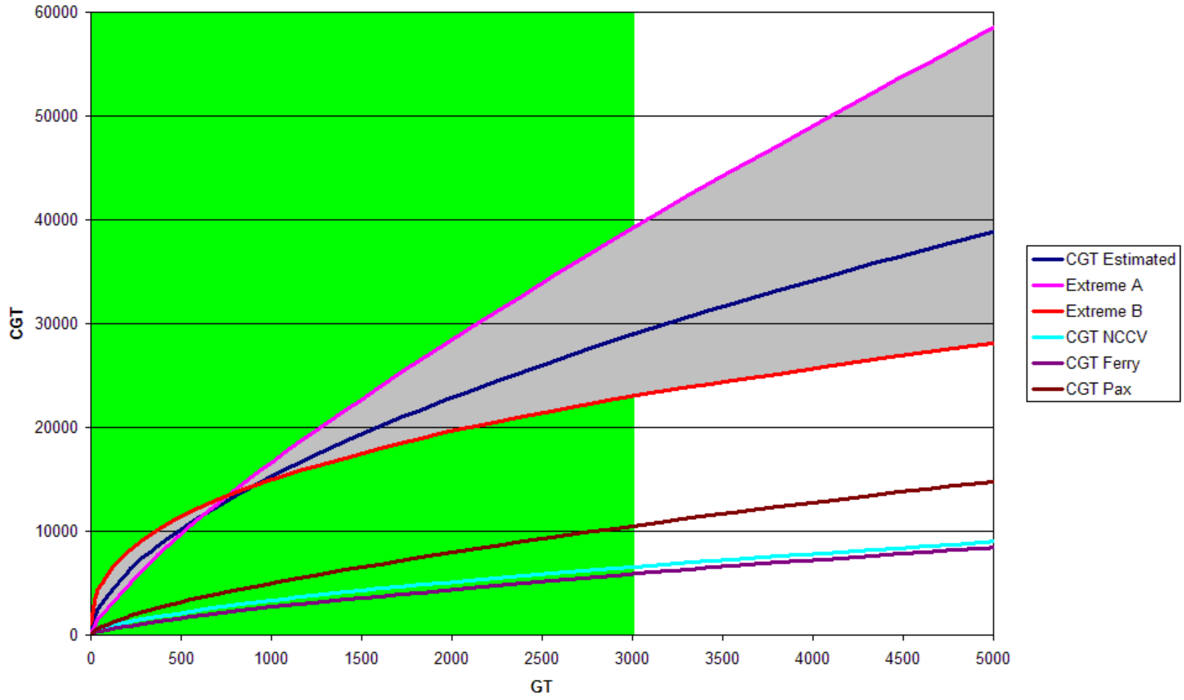
To compare the output of shipyards and shipbuilding industries, the comparison of the size (e.g. in terms of Gross Tonnage, GT) of the delivered vessels alone is an insufficient approach: How does one compare a (very large) crude oil tanker to a (smaller but much more complex) passenger ship?
In order to provide the basis for a fair comparison, an indicator was developed to measure the amount of work that goes in the construction of a vessel, called Compensated Gross Tonnage. As the basis of this factor, parameters have been agreed upon in the 1970's by the OECD to compensate for the differences in work involved in producing a gross ton (GT) of ship for different sizes and types of ships, thereby creating a more fair basis for comparison. From 2007 onwards CGT of a vessel of a certain type as a function of its size (measured in GT) is expressed as $CGT = A*GT^B$.

For super yachts, no internationally accepted A and B values to convert from GT to CGT exist. It is the belief of the super yacht building industry that this results in an under appreciation of the importance of the sector. As a result it is the desire of a large part of the industry, pooling their interests within the SYBAss organisation, to develop these values. To develop the appropriate A and B-values for super yachts, it must be realized that CGT is an industry-wide, macro-economic indicator that provides an estimate of the amount of work involved in building a ship of a certain type and size, ultimately expressed in man-hours per GT. In determining the appropriate CGT coefficients for super yachts, it is necessary to compare the effort involved to the effort required to build other ship types. Exact figures for a specific ship or yard situation can therefore not directly serve as the basis for calculation of CGT factors, since these figures are influenced by yard-specific factors. As a result, yard specific input from a number of yards is necessary to arrive at an industry-wide average.

A method was developed, in cooperation with the University of Genoa, to express the work involved in building a vessel in terms of a man-hour equivalent. This method was tested for a group of 18 control vessels. These are vessels for which a proper CGT value was available. These tests confirmed the applicability of the method to create a man-hour equivalent that could serve as an estimation of the CGT value of a vessel.

Applying this method on the 41 super yachts for which data was provided, it was possible to create an estimate of their "true" CGT values. Using these values and the Wald-test it was convincingly proven that both the CGT factors for ferries, passenger vessels and NCCVs, which are generally believed to be the most applicable ship types to use for CGT determination for super yachts, could not be applied to our data set.

As a result, new CGT parameters for super yachts should be created. Based on the currently available dataset, the factors would be the following; A = 278 and B = 0.58. However, due to a limited dataset and significant scatter in data points, some deviation from these factors may still be considered as scientifically plausible. This is best presented with the coefficient ellipse of Figure 3 in 5.2. At the extremes of these potential other values are A = 76, B = 0.78 and A = 1012, B = 0.39. As can be seen in the graph below, the effects of this for the super yachts (in the green area) are quite limited. The results for larger sizes leave a lot of room for uncertainty.



The last check performed was to see if motor and sailing yachts would require separate CGT factors. Based on the result of this test, this idea was rejected based on the current data set. The likelihood of two datasets having different CGT factors is expressed as a percentage. While a value of more than 95% is needed to be considered valid only 15% was achieved, indicating that based on the current data there is no ground to have two different sets of CGT factors.

The only major recommendation for further research is the collection of a larger data set to perform the calculations on. The authors do realise that the collection of the current data set was already a long and difficult path to walk, but more data would arguably allow a better estimation of the factors A and B for super yachts. While more data will not necessary lead to a smaller spread in the data points, it will become easier to identify outliers and by eliminating them reduce the spread as well.

Reading Guide

Background to this research in particular and of CGT in general is provided in the chapters 1 and 2. The fundamentals applied to the problem are presented in chapter 3 and these are validated in chapter 4. Chapter 5 determines and validates our proposed factors to be applied to super yachts. Lastly chapter 6 summarises the most important conclusions of this research

TABLE OF CONTENTS

Acknowledgements ... 2
Management Summary ... 3
Reading Guide ... 5
1. Background: CGT factors ... 7
2. Problem definition ... 8
3. Approach ... 9
   3.1. Fundamentals of the approach ... 9
   3.2. The process of data collection ... 10
4. Validation ... 13
   4.1. Validation of the method ... 13
   4.2. Validation of CGT methods for super yachts ... 15
5. Determination of factors for a Superyachts Group ... 18
   5.1. Determination of A and B for Superyachts ... 18
   5.2. Validation of the factors ... 19
   5.3. Differentiate between motor and sailing yachts? ... 22
6. Conclusions and Recommendations ... 24
   6.1. Conclusions ... 24
   6.2. Recommendations ... 24
Literature List ... 25
Appendix 1: CVs of research staff ... 26
Appendix 2: UNAS Code and assignment to systems ... 27

1. Background: CGT factors

To compare the output of shipyards and shipbuilding industries, the comparison of the size (e.g. in terms of Gross Tonnage, GT) of the delivered vessels alone is an insufficient approach: How does one compare a (very large) crude oil tanker to a (smaller but much more complex) passenger ship?
In order to provide the basis for a fair comparison, an indicator was developed to measure the amount of work that goes in the construction of a vessel, called Compensated Gross Tonnage. As the basis of this factor, parameters have been agreed upon by the OECD [1] to compensate for the differences in work involved in producing a gross ton (GT) of ship for different sizes and types of ships, thereby creating a fairer basis for comparison.

Up to 2007, the C-factor (linking GT to directly CGT) was set as a constant value over a certain range of GT for a certain ship type, resulting in stepwise changes of C factors. From 2007 onwards, this system was changed to a continuous function, expressed as $CGT = A*GT^B$. Within the formula, factor A primarily reflects the differences between the various ship types, while B reflects the influence of size within a certain ship type.

For super yachts, no A and B values exist to convert from GT to CGT. Since this is believed to lead to an under appreciation of the importance of the super yacht building sector, it is the desire of a large part of the industry, pooling their interests within the SYBAss organisation, to develop these values.

2. Problem definition

For super yachts no A and B values exist to convert from GT to CGT. It is the desire of a large part of the industry, pooling their interests within the SYBAss organisation, to develop these values. To develop the appropriate A and B-values for super yachts, it must be realized that CGT is an industry-wide, macro-economic indicator that provides an estimate of the amount of work involved in building a ship of a certain type and size, ultimately expressed in man-hours per GT. In determining the appropriate CGT coefficients, it is necessary to compare the effort involved to that for other ship types.

A major challenge here is that exact figures for a specific ship or yard situation can not directly serve as the basis for calculation of CGT factors, since these figures are influenced by yard specific factors.

Quoting OECD's document "Compensated Gross Tonnage (CGT) system 2007": "*…However, when it comes to details, practically no shipyard builds a ship the same way as its competitors. One of the major differences is the production depth, i.e. the amount of parts and blocks produced in the shipyard relative to the amount which is subcontracted to outside suppliers. …… similarly the degree of rationalization and the range of shipbuilding equipment such as cranes and machine tools will vary considerably and these will also influence the man-hours necessary for the building of a specific ship*"

A further complicating factor is that super yacht builders, unlike many other shipyards, do not build other ship types, certainly not within the same business unit, employing the same organisation and supply chain. Thus direct comparison of man-hours spent to build a super yacht and, say, a ferry is not possible within the same company and a more subtle approach is required, which is elaborated in the next chapter. Using the results obtained by this approach, representative CGT curves need to be derived, producing the desired A and B values

3. Approach

To overcome the issues mentioned above, a fair comparison needs to be made of the amount of work involved in building ships of various types and sizes. Since it was concluded above that each shipyards works in its own way and hardly any shipyards build both yachts as well as ships for which a CGT-factors do exist, this is challenging. The solution is found in a fortunate coexistence of this project together with a second project investigating the CGT values of smaller vessels at yards where also larger vessels (falling within a CGT group) are built. Using the same methodology for both the super yachts and the regular vessels, the authors were able to arrive at sufficiently accurate estimates of the amount of work in a vessel and with this an average amount of work per GT.

The remainder of this chapter is used to provide an overview of all the steps taken. This is followed by a validation of the results in chapter 4. In chapter 5 a first estimation of the CGT factors is presented. Due to the sensitive nature of the information used to arrive at the results of this research, no values that can be related back to the individual vessels will be presented.

3.1. Fundamentals of the approach

As discussed before, the OECD issues a warning concerning of the difficulties in assessing the production effort of various yards. However, it does not elaborate the way in which these are compensated for in OECD's methodology. This is however believed to be of minor importance, as long as a fair comparison can be made. As a result, we are left with the task of formulating a way of comparing the production effort in such a way that we can compare not only different yards, but also completely different vessels. At the basis of the approach followed there are two fundamental assumptions:

1. The CGT should be based on the total effort of building the vessel, including the work of all subcontractors and suppliers.
2. The money spent on the sub-contractors and major suppliers is a fair basis for assessing the effort they put into the vessel.

The production depth of every yard is different; while yards that do everything in-house still exist, the other extreme, where the yard only facilitates the building of the vessel, also exists. To be able to compare various vessels it was decided to create an estimate of the total effort involved in the vessel (assumption 1). To come to an estimation of the total vessel is was necessary to be able to assess the effort involved in the work of all the sub-contractors.

However, as yards are generally reluctant to provide information on the man-hours involved in their product, we can safely assume that the sub-contractors would be at least as reluctant to provide this data to the yards that employ them. As a result, another measure for their effort was needed. As the only information the yards will be able to provide about their sub-contractors is their price, this was to be used as an estimate for their effort (assumption 2). It is understood that the price paid by the yards is not the costs made by the sub-contractor (i.e. there will be a profit margin), two more assumptions make it possible to create an estimation of this effort:

3. The profit margin of the yard, can be used as a proxy of the profit margin of the sub-contractors
4. For specific disciplines, in which part of the work is done by the yard itself and part is done by subcontractors, division of costs between labour and materials of the yard is assumed the same for the sub-contractor

Yards usually use sub-contractors from the region they are located in. Reasons for this can be many, but language and distance are definitely important ones. If this would not be the case, they are operating both in the same global market. As a result of this, they share the same economic heights and lows, which makes it safe to assume that profits of both companies follow a similar trend; high when the economy is high and low when the economy is low (assumption 3).

With the profits taken care of, it is still important to realise that the costs of a product are based on the capital costs, the (direct) labour costs and the material costs. For each of the systems individually, it is assumed that the total costs divided by the direct man-hours, are comparable within a system (assumption 4). This implies that cost reductions are the results of balancing demand over multiple projects, rather than the performance on a single project. In other words, costs per man-hour are comparable as both parties strive for optimality.

Using these 4 assumptions it is possible to express the work involved in the building of a vessel in a man-hour equivalent. This equivalent is an indication of the work involved; its value does not represent the actual number of man-hours (direct or indirect) in any way! For this reason its absolute value is not of importance, only its relation to other values.

3.2. The process of data collection

The process of data collection and transformation was done in three steps. The first step was a table with confidential data on the vessel that was filled in by the shipyards. Especially the cost and price data is highly sensitive; therefore this part of the data would remain at the shipyards. As can be seen in Table 1, which is filled with completely fictional data, the work is divided over 11 different systems in order to provide more insight into the major cost drivers as well any major deviations from the norm for the cost of a specific ship. To make sure data was collected in the same way, the details of each system group are provided in Appendix 2, the definitions were based on the Dutch UNAS coding system.

Data Input						
	Total Costs Estimated		Total Costs from Subcontractors		Own Manhours Estimated	
General object cost	450	1000$	0	1000$	800	hours
Hull, ships body	2750	1000$	700	1000$	1300	hours
Small steel work	500	1000$	50	1000$	550	hours
Pipe elements	700	1000$	550	1000$	200	hours
Painting	450	1000$	400	1000$	50	hours
Propulsion	2000	1000$	1200	1000$	90	hours
HVAC	180	1000$	120	1000$	100	hours
Electrical	380	1000$	380	1000$	0	hours
Hatch cover	500	1000$	500	1000$	0	hours
Joinery	350	1000$	300	1000$	75	hours
Other equipement	600	1000$	550	1000$	120	hours
Total cost price	8860	1000$	4750	1000$	3285	hours
Salesprice	9500	1000$				

Dimensions Vessel		
Loa	50	m
B	10	m
D	10	m
T	5	m
GT	2500	Ton
CGT	3000	Ton
Steelweight	1000	Ton
Propulsion Power	7500	kW
Auxilary Power	2500	kW

Dimensions Superstructure		
L	10	m
H	10	m
B	10	m
Crew + Passengers	20	persons
Miscelaneous		
Vessel Type	Imaginairy vessel	
Cargo Capacity	1200	Units

Table 1: Example of the input sheet (values are fictitious)

To be able to exchange the data of the shipyard with the research team, without exposing the confidential details an intermediate sheet was set up, using the data filled in by the shipyards. As can be seen in Table 2, the technical data concerning the vessel itself was transferred, as well as the man-hours used in each of the systems by the shipyard. The financial data however has been translated into fractions of the cost price for the systems and corrected fractions for the sub-contractors, based on assumption 3 (i.e. similar profit margins for yard and subcontractors).

	Fraction Of Price	Fraction subc	Subc * Price	Manhours
General object cost	0.05	0.00	0.00	800
Hull, ships body	0.32	0.23	0.07	1300
Small steel work	0.06	0.09	0.01	550
Pipe elements	0.08	0.76	0.06	200
Painting	0.05	0.87	0.04	50
Propulsion	0.22	0.56	0.13	90
HVAC	0.02	0.63	0.01	100
Electrical	0.04	1.00	0.04	0
Hatch cover	0.05	1.00	0.05	0
Joinery	0.04	0.84	0.03	75
Other equipement	0.06	0.90	0.06	120
Total	1.00		0.500	3285

Dimensions Vessel		
Loa	50	m
B	10	m
D	10	m
T	5	m
GT	2500	Ton
CGT	3000	Ton
Steelweight	1000	Ton
Propulsion Power	7500	kW
Auxilary Power	2500	kW
Dimensions Superstructure		
L	10	m
H	10	m
B	10	m
Crew + Passengers	20	persons
Miscelaneous		
Vessel Type	Imaginairy vessel	
Cargo Capacity	1200	Units

Table 2: Delivered data by the yard (values are fictitious)

Making use of assumption 4, we use the fraction of subcontracted work to calculate the total man-hours for a certain system. It is important to notice that the actual cost of a man-hour is irrelevant for our case. The total value of this is calculated at the bottom of the first green column. If a yard subcontracts a system completely, no own man-hours are available to make an estimation. To compensate for this, the total hours are increased for this fraction of the costs that is missing. This result is the value in yellow in table 3. As a last step these identical man-hours are divided over the various systems based on the fraction of costs they represent. To take pipe elements as an example; based on the extrapolation of the man hours this task would be represented by 835 man-hours. The cost fraction however is 547 man-hour equivalents. These are not true man-hours, but rather a representation of the costs. We get the cost fraction of pipe element by dividing them by the total man-hour equivalent (547 / 7190 =) 7.6 %. The last column divides the standard man-hours of a system by the ones calculated from the input. In our example of the pipe elements, this fraction is 0.66, while not of use on its own; it is handy to track inconsistencies in the data between various vessels and yards.

Ship 1	System Manhours	Total Manhours	Manhour Value
General object cost	800	394	0.49
Hull, ships body	1684	2323	1.38
Small steel work	603	431	0.72
Pipe elements	835	547	0.66
Painting	396	346	0.87
Propulsion	207	1608	7.78
HVAC	273	143	0.52
Electrical	0	288	NA
Hatch cover	0	378	NA
Joinery	464	271	0.58
Other equipement	1262	460	0.36
Total	6524	7190	1.10
Value of Particulars			

Table 3: Calculation of the man-hour equivalent of a vessel (values are fictitious)

4. Validation

Validation of results consists of several steps. First, the results of the method are validated by matching the data from the 18 control vessels for which CGT values are known to their CGT curves. Next, a visual validation will be used to demonstrate that the existing curves do not apply to super yachts. Finally the available super yacht-related data will be regressed using the Ordinary Least Square method to determine realistic A and B values, which are then used to demonstrate statistically that the values of the available CGT-curves can be rejected.

4.1. Validation of the method

A first step in the validation of the method was a discussion with the University of Genoa on our proposed method. During a meeting it was concluded that given the boundary conditions posed by limited data availability, the method was deemed the most viable way of reaching the desired results.

Using the method described above, a man-hour equivalent for the 41 super yachts and the 18 control vessels was determined. As mentioned already before the man-hours equivalents are used as an estimation of the CGT of the vessel. Since the control vessels have a known CGT, it is possible to calculate the number of man-hour equivalent units per CGT for each of them. CGT is regressed using the Ordinary Least Squares (OLS) method on these man-hours to find an average[1]. This will be presented in a box plot. A box plot, also known as a box and whisker diagram, summarizes the distribution of a set of data by displaying the centering and spread of the data using a few primary elements [2].

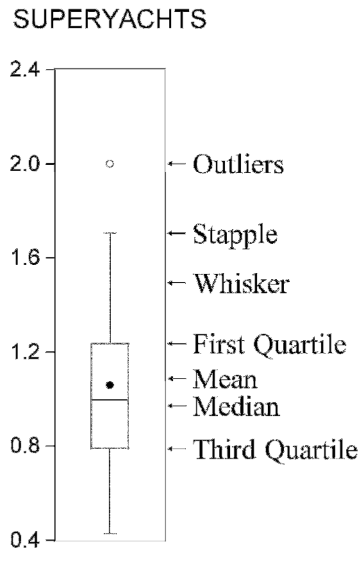


Figure 1: Box plot explained

[1] The difference between a "normal" average and a regressed average, is that the higher values have a greater influence. Regression was chosen for all values for the CGT displayed.

The box portion of a box plot represents the first and third quartiles (middle 50 percent of the data). These two quartiles are collectively termed the *hinges*, and the difference between them represents the *inter-quartile range*, or IQR. The median is depicted using a line through the center of the box, while the mean is drawn using a symbol.

The *inner fences* are defined as the first quartile minus 1.5*IQR and the third quartile plus 1.5*IQR. The inner fences are typically not drawn in box plots, but graphic elements known as *whiskers* and *staples* show the values that are outside the first and third quartiles, but within the inner fences. The staple is a line drawn at the last data point within (or equal to) each of the inner fences. Whiskers are lines drawn from each hinge to the corresponding staple.

Data points outside the inner fence are known as *outliers*. To further characterize outliers, we define the *outer fences* as the first quartile minus 3.0*IQR and the third quartile plus 3.0*IQR. As with inner fences, outer fences are not typically drawn in box plots. Data between the inner and outer fences are termed *near outliers*, and those outside the outer fence are referred to as *far outliers*. A data point lying on an outer fence is considered a near outlier. Assume 0.5 till 2 to be a reasonable bandwidth.

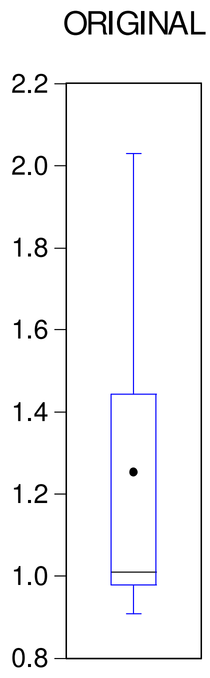


Figure 2: The spread in man-hours/CGT for the control vessels

The overview of spread of manhours per CGT for the control vessels is presented in figure 2. Assume the factor 2 range of the value (between divided by 2 and times 2 of the centre value (1)), to be reasonable. The whiskers (almost) fall within this range and there are no outliers, which is a very good sign. Another benefit of regression is that we can determine a value called R^2, the closer this value is to1, the better the chosen line (Man-hours = c x CGT) fits the data. In this case R^2 = 0.9584, quite a good fit. Another measure for the validity of the fit is the standard deviation of the error. Since the provided data is confidential, no

absolute values will be provided, but the standard deviation is presented as a percentage of the abovementioned factor c. This results in a deviation of 2.72 %. This is indeed a very small standard deviation. The confidence interval is a range around c between which a certain amount of variations can be found. The usual is the 95% confidence interval, which in this case is no wider than 11 % of the factor value, a very narrow band.
Although based on the available data, the correctness of the 4 assumptions made can not be proven individually, the results discussed above do support the conclusion that together they provide sufficiently accurate results. The small spread and error deviation make us confident that the chosen method yields acceptable results for taking the man-hour equivalent as a representation of the CGT value of a vessel.

4.2. Validation of CGT methods for super yachts

Super yachts are currently usually assigned to the CGT-curve of passenger vessels. Other options to assign them to would be the Non Cargo Carrying Vessels (NCCV) or ferries. An inspection of the graphs of these ship types showed that for the range of GT in which super yachts are built, the CGT value is higher for passenger vessels than that of the other two. All CGT-groups are investigated in this paragraph.

Using the factors of each of the 3 possible categories, a CGT value is calculated with the GT of the super yachts made available. In figure 3, the results, using the average number of man-hours per CGT found in the previous paragraph as a value of 1, are presented visually. It is immediately clear that none of the boxes are within the range of 0.5 to 2, defined as the maximum limits in the previous paragraph. This implies that if for instance the CGT curve for passenger vessels were applied to super yachts, on average three times as many man-hours are required to build a superyacht than indicated by the CGT. Since we previously stated that the effort per CGT should be equal across all ship types, this would imply that the factor used for passenger ships is roughly three times too low to be valid for super yachts. The number of builders involved in this data set is roughly one third of the yards involved in super yacht building worldwide and the data also covers roughly the same fraction of yachts built, though not clear from the box plot, the data also covers the full range of sizes in super yachts as they are build today.

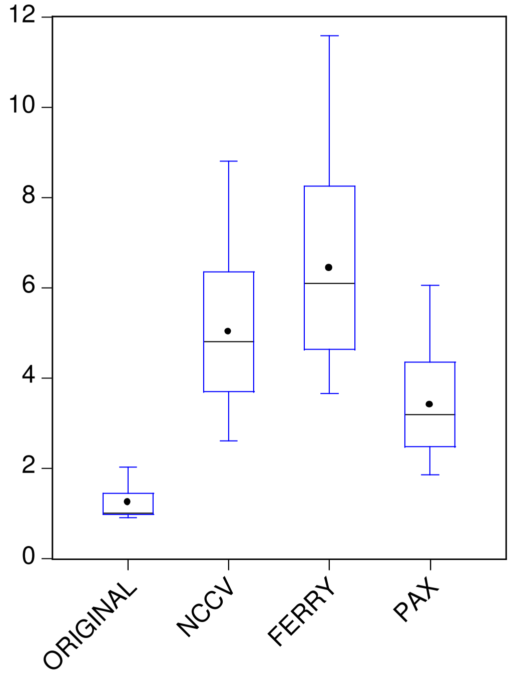


Figure 3: The spread in man-hours/CGT for the superyachts

The almost complete miss of the acceptable interval is a clear sign that the CGT factors currently in use for super yachts are nowhere near their correct value. Continuing our investigation in the same line as in 4.1, the results in the table below are obtained when trying to link our man-hour equivalents to CGT by a single factor through OLS-regression (man-hours = c x CGT). The control group column presents the results for the control group of 4.1. The ferry group presents the results with the CGT of the super yachts calculated using the ferry factors and our calculated man-hours equivalents. The NCCV column, does the same with the CGT calculated with the NCCV factors and lastly the Pax Column does this for the passenger vessels CGT. It is clear that the R^2 values are very low, representing the poor applicability of the CGT curves for these vessel types to super yachts.

	Control group	Ferry	NCCV	Pax
R^2	0.9584	0.4484	0.4903	0.4698
Standard Deviation of error	2.72 %	5.12 %	4.90 %	5.01 %
Size of 95% confidence interval	11 %	20.49 %	19.62 %	20.04 %

Table 3: Regression results applying the CGT values of Pax, NCCV and Ferries.

The Wald-test calculates the probability that a certain proposed coefficient is correct. We will be using this test as a final check to see if there is any chance that the CGT factors of passenger vessels, ferries or NCCVs could fit the data for super yachts. The results are presented in Tables 4a, 4b and 4c. For those readers unfamiliar with this statistical method, the rightmost column is the most important column. Here the chance (as a fraction of 1) is displayed that the coefficient found valid for the control group can be applied to the super yachts CGT calculated by either of the other values. In all cases for both test statistics, this chance is (virtually) zero. We assume that this is not due to faulty assumptions in the conversion of data to man-hours equivalents, as we validated the method using the 18 control vessels. Yet we contribute this bad fitting due to the fact that the CGT calculations currently in use for super yachts do not suffice and highly underestimate the workload involved in these special vessels.

Ferry	Value	df	Probability
F-Statistic	264.1280	(1,40)	0.0000
Chi-square	264.1280	1	0.0000

Table 4a: Wald-test for the Ferry value.

NCCV	Value	df	Probability
F-Statistic	262.9235	(1,40)	0.0000
Chi-square	262.9235	1	0.0000

Table 4b: Wald-test for the NCCV value.

Pax	Value	df	Probability
F-Statistic	189.7113	(1,40)	0.0000
Chi-square	189.7113	1	0.0000

Table 4c: Wald-test for the Passenger vessel value.

5. Determination of factors for a Superyachts Group

Due to the fact that the CGT factors currently used by OECD so greatly underestimates the actual workload for superyachts, it is deemed worthwhile to propose more acceptable values for the A and B factors for superyachts. Currently the dataset contains only 5 sailing yachts and 36 motor yachts, which is not enough to calculate separate values for both types, but still a check is performed to determine if the values found for just motor yachts is significantly different from the one found for the entire group. If this is the case, that would be a strong indication for the possibility of two different sets of values.

5.1. Determination of A and B for Superyachts

The known equation for CGT ($CGT = A * GT^B$) can be rewritten in a linear form, by taking the natural logarithm (LN) of both sides and rewriting the logarithm on the right side of the equation:

$$LN(CGT) = LN(A) + B * LN(GT) \quad (1)$$

Using the entire sample first the following results are obtained for the regression of the equation presented above. Be aware that the factor c(1) in Table 5 represents LN(A), it will need to be converted to get A. However the confidence range is therefore also a factor and not an added term (see equation 2). The figures will present the value of LN(A), but in the text this will be translated to the value of interest, A.

$$e^{A+SD} = e^A * e^{SD} \quad (2)$$

Below, table 5 presents the results of the regression. The top row shows the equation that was regressed. As mentioned already before, all observations are used, the program confirms this in the second row of the table. The R^2 is a measure for how well the chosen line fits the data, 1 is a perfect fit and a value below 0.5 is not a very good fit and below 0.25 we shouldn't consider it as a fit. The C-factors are the regression factors LN(A) and B, their value presented in the row below. The probability checks if there is reason to believe the values of the coefficients could be zero (and thus left out). The standard deviation is used to determine the spread in the data and is used to calculate the 95% confidence interval. This interval tells us the maximum and minimum value the coefficient might be, without violating the current data set.

LN(CGT) = C(1) + C(2)*LN(GT)		
Included observations: 41		
R^2	0.5920	
	C(1)	C(2)
Coefficient	5.6268	0.5845
Probability of value being 0	0.0000	0.0000
Standard Deviation of error	0.5106	0.0777
Standard Deviation of error %	9.07%	13.29%
Size of 95% confidence interval	36.30%	53.17%

Table 5: Results of determination of A and B value

To start with the results; A = 278 and B = 0.58. However the value found for R^2 here is close to 0.5, so we should be careful in interpreting the found values as true for all super yachts. The standard deviation is also quite high, giving us a large range for the "true" values of A and B to be in. At least the probability of either of the factors being zero can be dismissed. Although there is no proof of this, there is no reason to suspect the fit for the vessels originally used to determine the original OECD CGT factors is very good either. Since for example geared and regular bulk carriers are in the same CGT group, a large spread should also be expected here, as it is definitely more work to build a self unloading ship than one without this equipment.

It would be easy to say that this uncertainty in the values found can be solved using more data points, in essence reviewing more super yachts, but since there are many disturbing factors at play here, there is no guarantee of that. Though after a first review round, we actually added 5 new vessels, but the outcome did not change significantly, the confidence intervals improved by 4% and 17% (Standard deviation and 95% confidence interval respectively).

5.2. Validation of the factors

While the percentages at the 95% confidence interval might seem like a good indication of unreliable results, it is important to realize that the two coefficients are related and plotting the confidence interval for the combination of the two would help us understand the true range of possibilities for the A and B values. This is done in Figure 4 below.

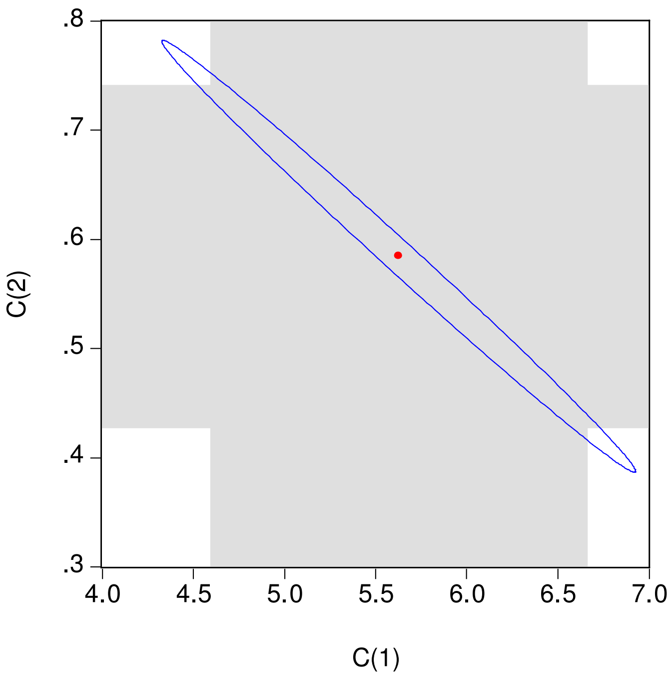


Figure 4: Confidence Ellipse for C(1) and C(2)

The horizontal and vertical grey boxes represent the confidence interval for the single coefficients. The blue line is the border of their combined confidence interval, with the red dot presenting their current values. This area is much smaller than the central square formed by the two grey boxes. However the blue line does extend to outside this box into the white areas. The direction of the ellipse (top-left to bottom-right) can be explained by the fact that if one of the values increases, the other one will have to decrease. The narrow bandwidth of the ellipse further indicates that this is a very strong relation. From the graph it is possible to derive the extreme points of the combination of A and B. In the top left edge of the ellipse A = 76 and B = 0.78. In the bottom right edge of the ellipse A = 1012 and B = 0.39.

Would we compare the possible values to those of passenger vessels and ferries, we could assume the same power (B-factor) for the super yachts as the group it is compared with. The A-factor is then determined by the confidence interval. In the case of the passenger vessels it is 3 times as large (+/- 155 vs. 49). In the case of the ferries it is even 6 times as large (+/- 120 vs. 20). These values are not the best fit to the available data, but with the current data available, these values can also not be rejected. This example serves mainly to indicate the big difference between super yachts and the groups to which they have been assigned so far.

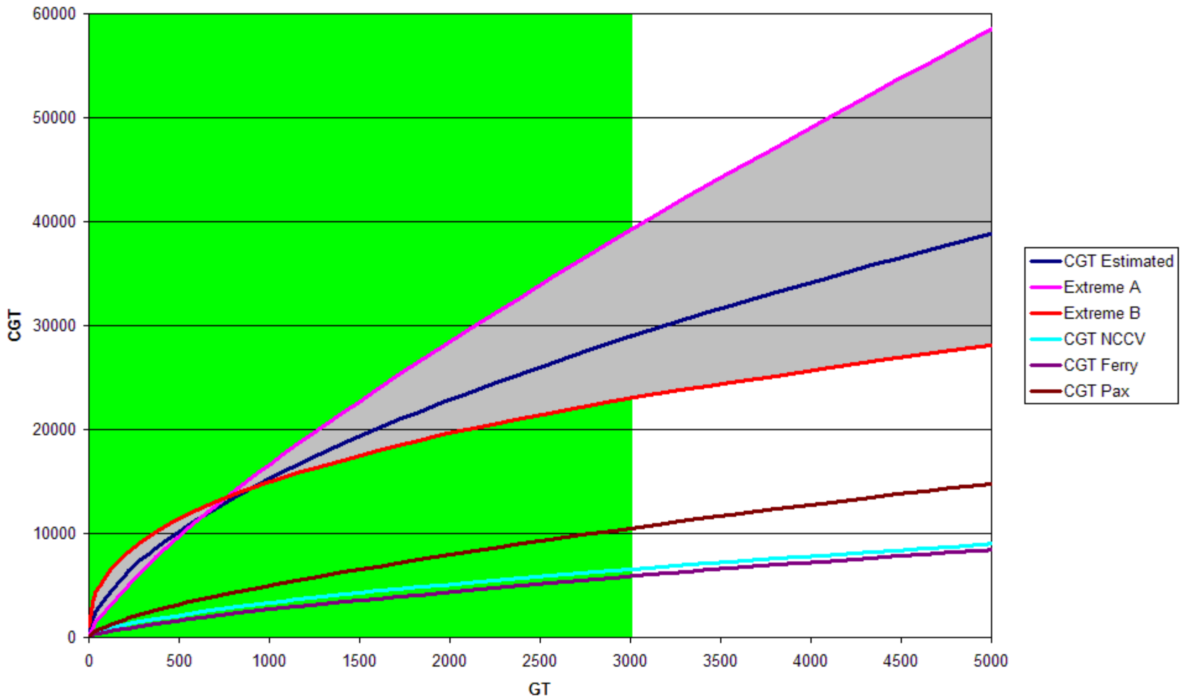


Figure 5: 95% confidence area for Super yachts CGT.

While this seems to result in quite extreme values, one should realise that the data is based on ships within the green area (and also meant for ships within this bar) in Figure 5. The spread, represented by the grey area in this part of the graph is actually quite limited. The effects with higher GT values on the other hand are increasing rapidly. It is expected that more data will converge the two extremes on a smaller bandwidth.

Next to the estimated value and the two extremes, the figure also shows the three CGT lines of the current categories available to super yachts. These all lay below even the minimum as was derived from our sample

A further visual check is done going back to figure 3, where the man-hour equivalent per CGT was presented with the CGT values calculated according to the factors of pax vesels, ferries and NCCV. We repeat that picture here in figure 6, but now adding the results for the CGT calculated by the values found in 5.1. With only 1 outliers outside the range 0.5-2, it is very clear that a CGT calculated according to these factors represents the work involved in the yachts a lot better.

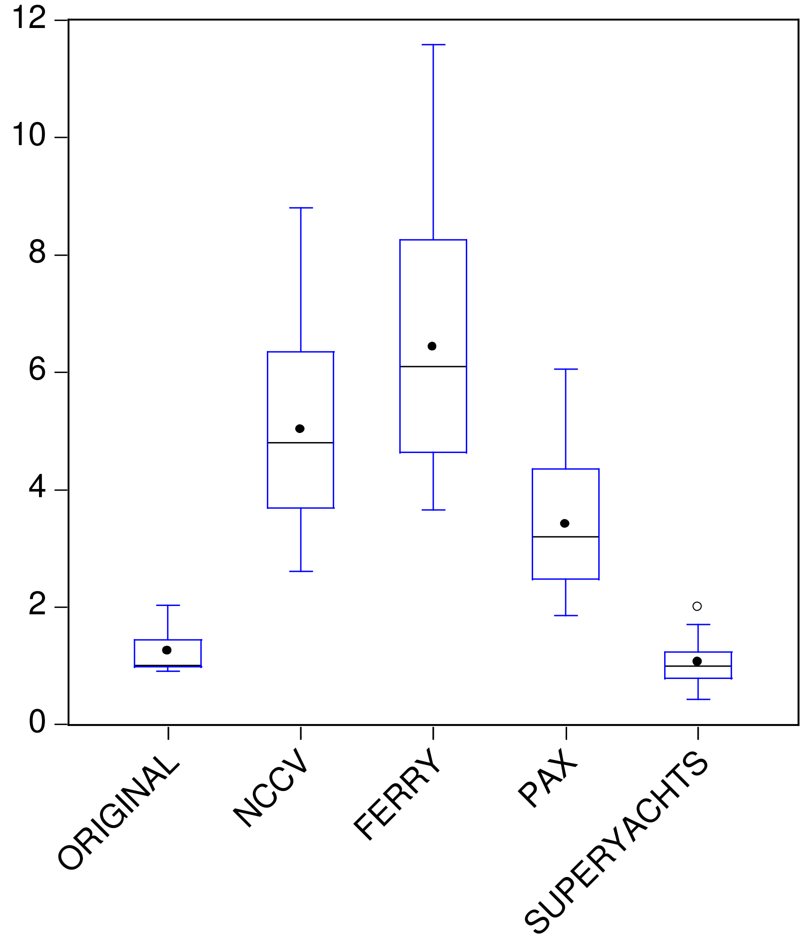


Figure 6: The spread in man-hours/CGT for the super yachts

The factors for ferries are A = 20 (LN(A) = 2.9957) and B = 0.71, for NCCV A = 46 (LN(A) = 3.8286) and B = 0.62 and for passenger vessels A = 49 (LN(A) = 3.8918) and B = 0.67. A quick check on the graph of Figure 4 shows that all combinations are not within the blue ellipse, they are not even on the graph range. Based on this, we can be sure that none of the lines is suited to represent the super yachts. To further confirm this, a Wald-test is performed but this time on the GT to CGT conversion, with the results presented in table 6a, 6b and 6c.

Ferry	Value	df	Probability
F-Statistic	663.7847	(2, 39)	0.0000
Chi-square	1327.569	2	0.0000

Table 6a: Wald-test for the Ferry value.

NCCV	Value	df	Probability
F-Statistic	495.8432	(2, 39)	0.0000
Chi-square	991.6865	2	0.0000

Table 6b: Wald-test for the NCCV value.

NCCV	Value	df	Probability
F-Statistic	280.0776	(2, 39)	0.0000
Chi-square	560.1552	2	0.0000

Table 6c: Wald-test for the Passenger vessel value.

The results of the test further confirm that for this data set the factors for ferries, NCCV and passenger vessels are not applicable, with zero probability of ever being applicable. On the other hand there is still a range of considerable size in which the real A and B values for super yachts can be found.

5.3. Differentiate between motor and sailing yachts?

As only 5 sailing vessels are available at this point in time, it is not possible to determine a valid estimate of the A and B factor for them individually. However with 36 motor yachts available it will be possible to regress an A and B factor for this group. In 5.1 it already became clear that the available data did not allow us to narrowly define an A and B factor for the complete group of super yachts, it is therefore not expected that the results for just motor yachts will be better off. The values presented here should not be considered as a suggestion for a separate CGT group for super motor yachts, but rather as a basis to check if such a group is necessary or based on the currently available data, can be discarded as an option.

The procedure is the same as used for the entire group explained in 5.1 and will not be further described here. The results of the regression are presented in table 7. This translates in the following factors for super motor yachts; A = 199 and B = 0.63.

Included observations: 36		
LN(CGT_MOTOR) = C(1) + C(2)*LN(GT_MOTOR)		
R2	0.5382	
	C(1)	C(2)
Coefficient	5.2908	0.6337
Probability of value being 0	0.0000	0.0000
Standard Deviation of error	0.6744	0.1007
Standard Deviation of error %	12.74%	15.89%
Size of 95% confidence interval	50.98%	63.55%

Table 7: Results of determination of A and B value

The table suggests an even worse fit, and with the wide confidence interval it is not possible to justify two different lines for motor and sailing yachts at this time. It would be possible to check visually if there is a chance that the values found for the entire group fit in the confidence ellipse of our current estimation, but better would be to execute the Wald-test and get quantified result. In this case we skip the confidence ellipse and go straight to Wald-test. The results of this are presented in Table 8.

Superyachts	Value	df	Probability
F-Statistic	0.128223	(2, 34)	0.8801
Chi-square	0.256446	2	0.8797

Table 8: Wald-test for the Superyachts value

Both tests that are part of the Wald-Test report a probability of at least 87 % that the values found for motor yachts could be replaced by the values found for the complete group of super yachts. In statistics, to say that two solutions are different this probability should have been below 5 %. This is not the case and we can conclude that based on the data currently available we cannot justify a different value for motor yachts (yet). Naturally these conclusions might or might not change if more sailing yachts can be included in the data set.

6. Conclusions and Recommendations

This chapter will summarise the conclusions of this report and propose some recommendations for further research.

6.1. Conclusions

The goal of this research was to check if the available CGT factors can be applied to super yachts and, if this is not the case, to define a new set of factors A and B that can be applied to super yachts. The last question was to investigate if there was a ground to differentiate between factors for motor and sailing yachts.

To be able to do all this, it was necessary to develop a method that allows expression of the workload involved in the construction of a vessel in a man-hour equivalent. This equivalent is solely meant to uniquely represent the work involved, it does not represent real (direct or indirect) man-hours. For this reason its absolute value is not of importance, only its relation to other values.

The method was developed in cooperation with the University of Genoa and tested for a group of 18 control vessels for which a proper CGT value was available. These tests confirmed the applicability of the method to create a man-hour equivalent that can serve as an estimation of the CGT value of that vessel.

Applying this method on the 41 super yachts for which data was provided it was possible to create an estimate of their "true" CGT values. Using these values and the Wald-test it was easily proven that the CGT factors for Passenger vessels, Ferries and NCCVs could not be applied to the data set.

As a result it can be concluded that new CGT factors for super yachts should be created. Based on the current dataset the factors would be the following; A = 278 and B = 0.58. However, due to a limited dataset and significant scatter in data points, some deviation from these factors may still be considered as scientifically plausible. This is best presented with the coefficient ellipse of Figure 3 in 5.2. At the extremes of these potential other values are A = 76, B = 0.78 and A = 1012, B = 0.39. The last check to see if motor and sailing yachts would require their own CGT factors was rejected on the basis of the current data set. The probability of two different sets of factors was 13 % a very low likelihood, it needs to be at least 95 % to speak of statistical significance.

6.2. Recommendations

The only major recommendation for further research is the collection of a larger data set to perform the calculations on. The authors do realise that the collection of the current data set was already a long and difficult path to walk, but more data would arguably allow a better estimation of the factors A and B for super yachts. While more data will not necessary lead to a smaller spread in the data points, it will become easier to identify outliers and by eliminating them reduce the spread as well.

With 11 different super yacht builders involved in this data set, variation could still be considered an issue. The greater the diversity of super yacht builders involved, the better the data set represents reality and the easier the found values can be defended.

Literature List

1. OECD paper C/WP6(2006)7, A NEW COMPENSATED GROSS TON (CGT) SYSTEM, 25-Oct-2006, OECD
2. Robert McGill, John W. Tukey, Wayne A. Larsen (February 1978). "Variations of Box Plots". The American Statistician 32 (1): 12-16. http://lis.epfl.ch/~markus/References/McGill78.pdf

Appendix 1: CVs of research staff

Prof. Hans Hopman MSc. holds a degree in Maritime Technology from Delft University of Technology. After a career in the Royal Netherlands Navy, in 2006 he became the head of Delft University of Technology's department of Ship design, production and operation. He has over 25 years of experience in the design of ships and has vast experience in leading multi-disciplinary design teams working on complex ships.

Robert Hekkenberg MSc holds a degree in Maritime Technology from Delft University of Technology. He started working for Delft University of Technology directly after graduation in 2004 as a researcher. In October 2007 he took up a position as assistant professor at the same university. Since 2004 he has been involved in both national and international research projects in the fields of inland shipping, ship production and ship design. Mr. Hekkenberg also teaches various courses at Delft University, dealing with the same topics and is working part-time on a PhD thesis in the field of inland navigation.

Jeroen F.J. Pruyn MSc holds a degree in Maritime Technology from Delft University of Technology. Started working for Delft University of Technology straight after graduation in 2004 as a researcher. During his work he has been and is involved in various projects both on a National and European level focussed on digital information exchange, yard lay-out, yard design and yard process optimisation & simulation. In 2005 he took a break of a year to travel around the world. Recently he has started working part-time on his PhD involving an economic simulation of the shipping industry in order to test policies on a national and company level.

Appendix 2: UNAS Code and assignment to systems

#	Name	KPI	UNAS Groups Present
0	General Object Costs	GT	[000] - [090]
1	Hull, Ship's Body	Lws	[100] - [120]
2	Small Steelwork	Lws	[130] - [150], [170]
3	Piping	Lws	[300] - [360], [380], [390]
4	Painting	LBD^(2/3)	[160]
5	Propulsion	kW	[200] - [220]
6	Ventilation	m3 cargo hold	[371]
7	HAC	crew/passengers	[372] - [374]
8	Electrical	kVA	[400] - [490]
9	Joinery	crew/passengers	[700] - [790]
10	Other Equipment	LxD	[500] - [520], [600] - [690], [800] - [890]
11	Special Equipment	GT	[530] - [590], [900]

[000] General Items
[010] ACQUISITION (COSTS)
[020] SPARE PARTS
[030] CLASSIFICATION (COSTS)
[040] TRIALS AND SUPERVISION
[050] SHIPMENT (COSTS)
[060] MISCELLANEOUS
[070] DESIGN ASPECTS
[080] FINANCING AND BANKING EXPENSES
[090] RE-CHARGEABLE EXPENSES AND ADDITIONAL CHARGES
[100] Shipbuilding
[110] HULL
[120] SUPERSTRUCTURE
[130] HATCHES, DOORS, WINDOWS ETC.
[140] STAIRS, LADDERS, HANDRAILS, PLATFORMS ETC.
[150] ADDITIONS TO SHIP'S CONSTRUCTION
[160] CORROSION PROTECTION AND DECK COVERING (OUTSIDE)
[170] REMAINING ITEMS
[200] Machinery/Propulsion
[210] PROPULSION SYSTEM
[220] STEERING SYSTEM
[300] Primary Shipsystems
[310] BILGE-, BALLAST-, DECKWASH-, AND INTERNAL FIFI SYSTEM
[320] FUEL OIL SYSTEM
[330] COOLING WATER SYSTEM
[340] FRESH/ AND SEA WATER SYSTEM
[350] FILLING/, SOUNDING/ AND DE-AERATION SYSTEM
[360] LUBRICATION OIL SYSTEM
[370] VENTILATION, AIRCONDITIONING AND HEATING SYSTEM
371 NATURAL AND MECHANICAL VENTILATION SYSTEM (supply and exhaust)
372 AIRCONDITIONING SYSTEM
373 HEATING SYSTEM
374 REMAINING ITEMS
[390] REMAINING PRIMARY SHIP SYSTEMS
[400] ELECTRICAL SYSTEM
[410] POWER GENERATING SYSTEM
[420] CABLES AND WIRING
[430] SWITCH BOARDS
[440] ALARM SYSTEM
[450] LIGHTING
[460] INTEGRATED INFORMATION SYSTEMS
[490] REMAINING ITEMS
[500] DECK EQUIPMENT
[510] ANCHOR EQUIPMENT
[520] MOORING SYSTEM
[530] FISHING GEAR
[540] HOISTING EQUIPMENT
[550] ANCHOR HANDLING, TOWING AND PUSHING EQUIPMENT
[560] DIVING SYSTEM
[570] LIFE SAVING/ FIRE PROTECTION EQUIPMENT
[590] REMAINING ITEMS
[600] SECONDARY SHIP SYSTEMS
[610] HYDRAULIC SYSTEM
[620] COMPRESSED AIR SYSTEM
[630] CARGO HANDLING SYSTEM
[640] (OIL) POLLUTION CONTROL
[650] EXTERNAL FIRE FIGHTING AND SALVAGE SYSTEM
[660] PRE-WETTING SYSTEM
[670] FIXED INSTALLATIONS FOR INTERNAL FIRE EXTINGUISHING
[680] COOLING AND FREEZING SYSTEM
[690] SEA KEEPING IMPROVEMENT DEVICES
[700] JOINERY AND ARRANGEMENT OF ACCOMMODATION
[710] JOINERY (GENERAL)
[720] INSULATION
[730] ARRANGEMENT ENGINE ROOM
[740] ARRANGEMENT WHEELHOUSE
[750] ARRANGEMENT LIVING QUARTERS
[760] ARRANGEMENT STORE AND BOATSWAIN EQUIPMENT
[770] ARRANGEMENT ADDITIONAL SPACES AND SANITARY SPACES
[780] TOOLS
[790] ACCOMMODATION PASSENGERS
[800] NAUTICAL, NAVIGATION AND COMMUNICATION EQUIPMENT
[810] NAVIGATION LIGHTING
[820] OPTICAL SIGNALLING EQUIPMENT/ SEARCH LIGHT SYSTEM
[830] ACOUSTICAL SIGNALLING SYSTEM
[840] RADAR SYSTEM
[850] DIRECTION FINDING AND COURSEKEEPING SYSTEM
[860] POSITION FINDING AND POSITIONING SYSTEM
[870] SONAR, SPEEDLOG AND DEPTH SOUNDING SYSTEM
[880] INTERNAL AND EXTERNAL COMMUNICATION (INCL. TV/AUDIO/VIDEO)
[890] METEOROLOGICAL AND OTHER MEASURING AND MONITORING SYSTEMS
[900] SPECIAL EQUIPMENT