Estimating Ship Repair Time for a Particular Ship Type and Shipyard Using Multivariable Linear Regression

Deepak Sebastian, Dr. K Sivaprasad

Research Scholar, Department of Ship Technology, CUSAT, Ernakulam-682022, India

joedeepak@gmail.com

Professor, Department of Ship Technology, CUSAT, Ernakulam-682022, India

sivaprasad@cusat.ac.in

Abstract

Introduction: Evaluating the total duration required for a ship repair project is a key element towards estimating the cost of repair by a shipyard and also for allocation of resources. The duration of ship repair is the period of downtime for the ship and that information is crucial for ship owners for operations planning

Objectives: The objectives include identification and analysis of the various factors affecting ship repair time, and deducing a reliable empirical formula to estimate the total repair time for a given set of repair works in a particular shipyard.

Methods: Multivariable Linear Regression and 3D Surface Plots are used to analyse the factors affecting ship repair time, and arrive at the formula.

Results: The independent variables exhibit a positive correlation with Ship Repair Time and the empirical formula is seen to be having an accuracy of about 70%.

Conclusion: The study has further scope, to expand to study for various types of vessels and for different shipyards, and accuracy of the empirical formula can be improved with larger data sets.

Key Words: Ship, repair time, shipyard, repair quantum, age, ship type, displacement, deadweight, tankers, multiple variable linear regression

INTRODUCTION

Ships all over the world are required to be annually or periodically taken for repair/maintenance work as mandated by Classification Societies. This work is undertaken at a shipyard as decided by the ship owner/ management. The ship will not be available for operations during this period of work, and implies loss of revenue for the owner. The duration of ship repair engages the facilities in the yard limiting use for other requirements. Therefore, both the shipowner and shipyard are interested in reducing the total repair time. The ship owner seeks quotations for the repair from different shipyards. The duration of the repair period has a direct bearing on the cost of repair. Accurate estimation of ship repair time is a valuable tool to predict repair costs and commit in the shipyard quotation.

BACKGROUND

The author has spent considerable time in ship repair activities and estimation of repair time was a challenge faced repeatedly. Similarly, in different shipyards, this is a grey area, where there are no empirical formulae to estimate the ship repair time. It is so, because the governing parameters forms a large set. The types of repair work are so many and their quantum varies from ship to ship. The fixed parameters of the ship such as age, type, displacement, dimensions etc. also add to the complexity. The facilities, equipment and technology available vary between shipyards and the total repair time taken depends on that as well. Only limited studies have been undertaken in this largely unexplored domain, motivating analysis on the same.

Previous Studies

Surjandari, I., et al (2013) [1] have prepared predictive models to estimate the duration of drydocking of ships for maintenance, in Jakarta, Indonesia. Dev, A.K., et al (2015) [2] have evaluated the factors affecting repairing time of crude oil tankers, at a single shipyard and developed a mathematical model for the relationship using multiple linear regression analysis. Dev, A.K., et al (2017) [3] have derived a relationship between the repairing time and labour for a ship and its influencing factors, again using the data from a single shipyard. Dev, A.K., et al (2017) [4] have analysed a particular component of ship repairing work, i.e hull coating renewal and its impact on the total ship repairing time. Karim, M.S., (2019) [5] has identified the factors that affect ship repairing labour (in man days) and defined their inter-relationship based on data of 190 ships of different types, that were repaired in Chittagong Dry Dock Limited (CDDL) Bangladesh. Manea, M.G., et al, (2020) [6] in their paper, have derived an analytical function for defining the dependence of time period for ship repair, on the different types of maintenance works carried out in a shipyard. Awal, Z.I., et al (2020) [7] have proposed a mathematical model for estimating the repair time for cargo ships by investigating its dependency on various parameters associated with ship repair. Data for the study has been obtained from Chittagong Dry Dock Limited(CDDL), Bangladesh. It can be seen that the previous studies are few and specific to a particular shipyard and geographical location. There is immense scope for multiple types of analysis on a host of parameters, based on different shipyards, types of ships and types/quantum of work. In this study, a shipyard has been identified for analysis of the total ship repairing time, details of which cannot be provided due to confidentiality requirements.

Area of Study

Due to the nature of the study, wherein the type of ship and the shipyard is highly relevant, and limited access to data exists, a non-probability convenience type data sampling has been adopted. Accordingly, for this study, repairs of oil/chemical tankers undertaken in a single shipyard has been taken up for analysis of the total ship repairing time. From the ship repair invoices, the age, displacement and individual repair quantum have been collated for analysis. Individual repair refers to each repair activity, such as underwater hull washing, tank cleaning, painting, cutting, welding, piping renewal, insulation renewal. The individual repair quantum refers to the amount of work, that is quantified by different units such as weight, area, volume etc..

Analysis Methodology

The objective is to carry out a multiple regression analysis, so as to arrive at a regression equation which can be used to estimate the Ship Repair Time.(SRT). For the study, the Ship Repair Time(SRT) has been considered as the dependent variable (DV) and the other parameters are treated as the independent variables (IDV) as mentioned in Table 1:

Independent Variables	**Units**
Age (A)	year
Displacement (D)	tonne
Hull Washing (HW)	m^2
Tank Cleaning (TC)	m^3
Cutting (C)	m
Welding (W)	m
Painting(P)	m^2
Piping(PP)	m
Insulation(I)	m^2

Table 1. List of Independent Variables and Units

The definition of the parameters and independent variables in order to quantify them is as follows:

Ship Repair Time(SRT) : calculated from the date of commencement of refit to the date of completion of refit.

Age(A) : Years since delivery of the ship

Displacement(D) : Measure of the weight of the ship.

Hull Washing(HW) : Hydrojet cleaning of marine fouling of underwater hull.

Tank Cleaning(TC) : Clean all the cargo holds and tanks

Cutting(C) : Undertake structure cutting (includes onboard and shopfloor cutting)

Welding(W) : Undertake welding (includes onboard and shopfloor welding)

Painting(P) : Apply all coats of paint at all location (Internal and External)

Piping(PP) : Remove repair and refit piping

Insulation(I) : Remove repair and refit insulation.

1 day : 12 hours (Max) in a day [8 hours (working hours) + 4 hours (overtime)].

Sample data pertaining to 10 ships, repaired in a single shipyard, was collated. The total Ship Repair Time for each of the 10 ships is graphically represented in Fig 1.

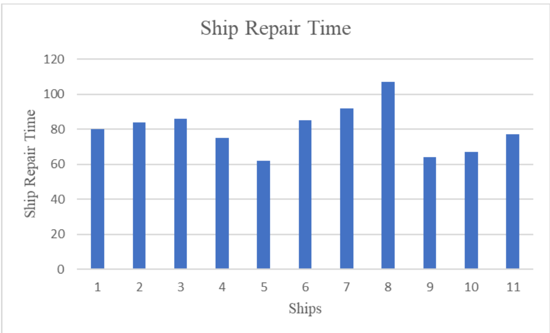


Fig 1. Total Ship Repair Time for Sample data

Multiple Regression Analysis

The steps followed in the multiple regression analysis are as follows:

(a) Generate a scatter plot of each independent variable with the dependent variable.

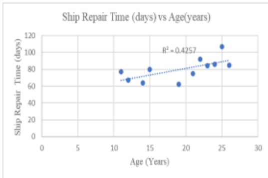
(b) Determine the correlation between each independent variable with the dependent variable.

(c) Generate scatter pots and correlation among the independent variables to identify collinearity and remove redundant variable.

(d) Use relevant independent variables to frame a multiple linear regression equation.

(e) Deduce co-efficients of the independent variables for the regression equation.

The scatter plots of each independent variable with the dependent variable are shown below:



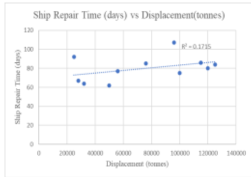
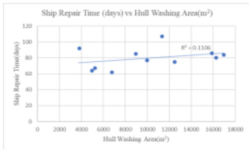


Fig 2. Age vs SRT Fig 3. Disp vs SRT



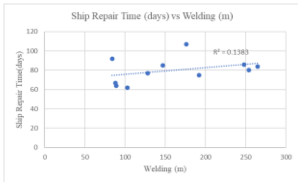
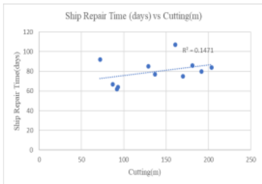


Fig 4. HW vs SRT Fig 5. Welding vs SRT



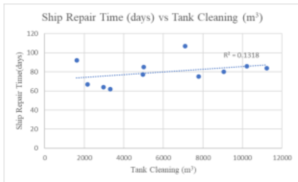
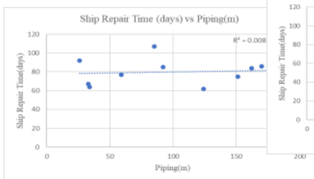


Fig 6. Cutting vs SRT Fig 7. TC vs SRT



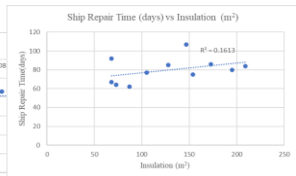


Fig 8. Piping vs SRT Fig 9. Insulation vs SRT

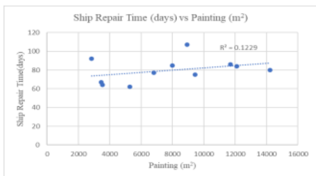


Fig 10. Painting vs SRT

The correlation coefficient(r) and the coefficient of determination (r^2) is tabulated below:

IDV vs SRT	Correlation Co-efficient (r)	Coefficient of Determination (r2)
Age	0.6525	0.4257
Disp	0.4141	0.1715
Hull Washing	0.3325	0.1106
Tank Cleanir	0.363	0.1318
Cutting	0.3835	0.1471
Welding	0.372	0.1383
Painting	0.3506	0.1229
Piping	0.0895	0.008
Insulation	0.4016	0.1613

Table 2 . r and r^2 values

In this process, simple regression is carried out between a pair of variables. The trendline gives the simple regression equation applicable for each pair. Correlation co-efficient determines the strength of a linear relationship. As a thumb rule, if $|r| \geq (2/\sqrt{n})$, where n is the number of samples, then it can be said that a relationship exists. In this case n=10 and hence $(2/\sqrt{n}) = 0.6325$. It can be seen that piping has insignificant correlation and hence is eliminated. The co-efficient of determination indicates how well the trendline fits the data. If the co-efficient of determination, $r^2 > 0.6$, then it can be considered a good fit. It is observed that only Age displays a reasonable correlation co-efficient and none of the IDV's display a requisite r^2 value. Nevertheless, we proceed to the next step in the process, i.e. to determine collinearity if any among the dependent variables.

Using the same process as before, scatter plots of independent variables with each other are plotted to identify correlation. In case, the correlation coefficient is above 0.6325, then only one of the two variables (the one with the higher correlation coefficient with the dependent variable) is to be taken for the regression equation. Cutting and Welding exhibit high collinearity, which is logical since the amount welding is highly correlated with the quantum of cutting.

The typical multiple linear regression equation is

$$y = b_0 + b_1x_1 + b_2x_2 .... + b_kx_k$$

where

y : the predicted value of the dependent variable

$x_1, x_2, .. x_k$: values of the independent variable

b_0, $b_1 .... b_k$: co-efficients of the independent variables

k : number of independent variables

For this study, the repair time is assumed as the summation of the product of the individual repair quantum with a co-efficient. Using multi variable linear regression, a regression equation has been arrived at, to provide an estimate of the ship repair time in days. The multiple linear regression equation would be

$$SRT = b_0 + b_1A + b_2D + b_3HW + b_4TC + b_5C + b_6W + b_7P + b_8PP + b_9I$$

Eliminating the insignificant and redundant independent variables, the equation comes down to

$$SRT = b_0 + b_1A + b_2D + b_3HW + b_4TC + b_6W + b_7P + b_9I$$

Using multiple linear regression in Excel, the co-efficients were evaluated and obtained as follows:

b_0 =34.29, b_1 =1.94, b_2 = 0.009, b_3 =1.31, b_4 =0.0083, b_6 =0.0045, b_7 =0.0039, b_9 = 0.071

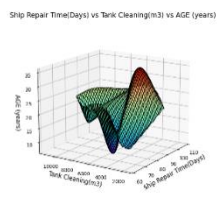
To test the fit of the regression equation, the sample data of all the ten ships was substituted in the equation, to arrive at the predicted value of Ship Repair Time. The percentage deviation or error from the actual Ship Repair Time was evaluated and tabulated below. It can be seen that for three out of the 10 ships, the error percentage is less than ten percent:

Ship	1	2	3	4	5	6	7	8	9	10
Percentage Error	2.68	1.80	-10.20	-7.60	-27.58	-3.51	9.27	16.32	-5.65	1.81

Table 3 . Percentage Error

3-D PLOTS

To assess the correlation between more than 2 variables, it was attempted to create 3D surface plots (using Python code). The dependent variable was kept on one axis, while combination of any two IDV's were placed on the remaining two axes. A few of the plots are as given below:



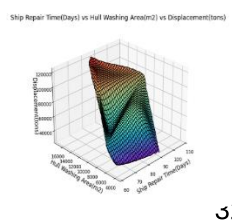
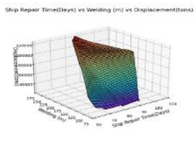


Fig 11. SRT vs TC vs Age Fig 12. SRT vs HW vs D



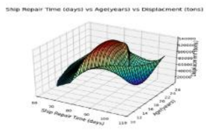


Fig 13. SRT vs W vs D Fig 14. SRT vs A vs D

Results and Analysis

The scatterplots have indicated that all the independent variables have a positive correlation with the dependent variable, i.e. the Ship Repair Time increases with increase in each independent variable. However, only Age is the only variable that is showing a relatively strong correlation, i.e. above 0.65. The remaining independent variables are showing low correlation. Likewise, the fit of the trendline (the simple regression equation) with the data, reflected in the r^2 value, is quite low. Even for Age, the value is less than the min required 60 percent.

However, for the sample data of ten ships, the derived multiple linear regression equation has provided a predicted Ship Repair Time with substantial amount of accuracy of about 60-70%. The 3D plots were carried out on a trial basis and most of the results seem a bit erratic.

The low correlation and error percentages in the results is attributed to extremely low number of the sample data, i.e. the number of ships, for which the data has been collected. Larger data set will help to facilitate a better analysis of trend, correlation and a more accurate prediction of the dependent variable. Also, since the data has been laboriously extracted from invoices, error during collation of data cannot be entirely ruled out. The analysis is applicable only to a single shipyard and for a particular type of ship from where the data has been obtained.

Conclusion

The study has large scope for further research. A generalised equation applicable for any shipyard for any type of ship in any geographical location would be case for future study. An algorithm/program can be prepared to take inputs on repair data from any shipyard and generate a customized prediction for the Ship Repair Time for any type of ship.

References

[1] Surjandari, I., and Novita, R., (2013) "Estimation Model of DryDocking Duration Using Data Mining". *International Journal of Industrial and Manufacturing Engineering Vol:7, No:7. World Academy of Science Engineering and Technology*.

[2] Dev, A.K., and Saha, M., (2015) "Modeling and Analysis of Ship Repairing Time". *Journal of Ship Production and Design, Vol. 31, No. 2, May 2015*, pp. 129–136.

[3] Dev, A.K., and Saha, M., (2017) "Ship Repairing Time and Labour". *https://www.researchgate.net /publication /320126165*.

[4] Dev, A.K., and Saha, M., (2017) "Analysis of Hull Coating Renewal in Ship Repairing Time". *Journal of Ship Production and Design*, Vol. 33, No. 3, August 2017, pp. 197–211.

[5] Karim, M.S., (2019) "Modeling and Analysis Of Labour Cost Estimation For Ship Repairing: A Case Study In Chittagong Dry Dock Limited". *Masters Thesis, Khulna University of Engineering and Technology*, Bangladesh.

[6] Manea, M.G., Zagan, R., and Manea, E., (2020) "Modelling of maritime ship repairs processes in shipyards, using the linear regression and correlation theory". I*OP Conf. Series: Materials Science and Engineering 916 (2020)*

[7] Awal, Z.I., and Abdullah, A., (2020) "Multiple Linear Regression Model for Ship Repair Time Estimation". *International Conference on Maritime Technology*, *MARTEC 20*.