

Resistance Performance Forecast of Trimaran Based on Multiple Regression Analysis

Xiao ying Xu¹, Chao ying Meng^{2,*}, and KUANG Wang³

^{1,2} Wenhua College, Wuhan, 430074, China

³ China Classification Society Wuhan Branch, Wuhan, 430033, China

Abstract. Based on the multiple linear regression of the relevant data of 29 ships in the basic database, including cargo ship, oil ship, container ship and chemical ship, etc, the empirical estimation formula of resistance coefficient of single ship under different Fr was obtained. The feasibility of applying the above regression formula to the prediction of the resistance performance of the trimaran based on DTMB5415 is verified by comparison with CFD simulation data. This method has a strong engineering applicability and provides a new idea for the study of computer-aided layout optimization of the trimaran.

1 Introduction

In recent years, research on trimaran, especially its layout optimization is taking attention to a growing number of scholars. Many studies have achieved remarkable results, but they generally use CFD technology based on viscosity theory, which takes a long time to calculate. They are not suitable to be extended to a large number of conversion in the layout and speed. In response to the above problems, this paper performed multiple linear regression based on the relevant data of 29 cargo ships, oil tankers, container ships and chemical tankers, and obtained the empirical estimation formula for the resistance coefficient of monohull under different Fr. Using CFD to simulate the trimaran verified the feasibility of applying the regression formula, which predicted the resistance performance of the trimaran. This paper provides a new idea for computer-aided trimaran layout optimization research, which can be widely used preliminary prediction of the trimaran's performance.

2 Regression principle

Regression analysis is a method of statistical calculation. It starts from a set of observation data of independent variables and dependent variables. Finally, find a functional expression that approximates the statistical correlation between the independent and the dependent variables. In various regression analysis, linear regression is relatively common. And some nonlinear problems can also be transformed into linear regression problems through variable substitution. We called them generalized linear regression. It is:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_m x^m + \varepsilon \quad (1)$$

where ε is error. When $x_1 = x, x_2 = x^2, \dots, x_m = x^m$, equation (1) becomes:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon \quad (2)$$

For a set of observations containing n data points ($x_{i1}, x_{i2}, \dots, x_{im}, y_i$), the regression coefficients $\beta_0, \beta_1, \dots, \beta_m$ can usually be determined by the least square method to minimize the residual sum of squares, as shown in the following formula (3):

$$Q = \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im})]^2 \quad (3)$$

After obtaining the regression coefficients, the significance of the regression model must be tested with using the F method. Firstly, define $\hat{y} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im}$ and $\bar{y} = \sum_{i=1}^n y_i / n$. Secondly, put them into the following formula (4):

$$\begin{aligned} SSE &= \sum_{i=1}^n [y_i - \hat{y}]^2 \\ SSR &= \sum_{i=1}^n [\bar{y} - \hat{y}]^2 \\ SST &= \sum_{i=1}^n [y_i - \bar{y}]^2 \end{aligned} \quad (4)$$

Where SSE is error square sum, SSR is regression squares sum, SST is total error square sum.

Finally, suppose $H_0 : \beta_0 = \beta_1 = \dots = \beta_k = 0$. When H_0 holds, the tested statistic F is shown in the following formula (5):

$$F = \frac{SSR/m}{SSE/(n-m-1)} \sim F(m, n-m-1) \quad (5)$$

The probability of test significance is $p = P(F(m, n-m-1) > F)$, when $p < \alpha$ (α generally taken as 0.05) holds, the regression equation is

* Chao ying Meng: 253508105@qq.com

significant. So, the smaller SSE, the higher the accuracy of the regression equation.

In addition to testing the overall significance, it is also necessary to test the significance of a single variable in the regression model. We can eliminate variables that have no significant impact and obtain the optimal equation. It can be carried out according to the following basic idea. Under a given limit, some variables are gradually introduced and deleted in the equation containing only constant. And a significance test is performed for each step. This process is repeated until there are no insignificant variables removed and introduced into the regression equation.

3 Resistance performance of monohull

3.1. Sample database

The database is composed of 29 inland vessels, including 17 cargo ships, 7 container ships and 5 petro-chemical ships. It can be supplemented in follow-up work. Because of the wide distribution range of the main dimension of each ship, basic data is so much. Due to limited space, Table 1 only gives the relevant parameters under $Fr=0.22$.

3.2. Sample regression analysis

After selecting the regression variable, multiple linear regression and testing are repeated many times. The significance test result of residual resistance is obtained.

The root mean square error (RMSE) is 0.0878, the statistical value F is 10.9071 and the corresponding probability p is 0.0001.

Since p is less than α , it shows that the regression equation is highly significant. And $RMSE = 0.0878$, indicates that the regression effect is good. The final regression equation is as follows (6):

$$\begin{aligned} C_r = & 4.4284 + 7.3579\left(\frac{L}{B}\right) + 5.1019\left(\frac{B}{T}\right) - 95.2222(C_b) \\ & - 9.6765\left(\frac{L}{B}\right)(C_b) - 4.1209\left(\frac{B}{T}\right)(C_b) - 0.1861\left(\frac{B}{T}\right)^2 \\ & + 110.4263(C_b)^2 \end{aligned} \quad (6)$$

For the convenience of observation, the following Figure 1-3 respectively shows the comparison between the final regression result and the original C_r about L/B , B/T and C_b . It can be seen that the regression value and the original data are both in good agreement with the distribution trend and numerical value.

Table.1 the Relevant Parameters under $Fr=0.22$

num	L/B	B/T	C_b	C_r	C_t
1	6.4779	4.25	0.796	2.062	4.133
2	5.82	3.3333	0.83	1.962	3.946
3	5.5740	3.0167	0.837	2.143	4.117
4	5.6961	4.6428	0.750	2.105	4.234

5	6.1617	4.8571	0.765	1.257	3.344
6	5.9520	5.2142	0.801	1.75	3.825
7	6.3023	5.375	0.831	2.415	4.421
8	5.7129	3.0851	0.833	2.283	4.250
9	5.1687	6.5	0.771	1.826	3.989
10	6.1764	4.8571	0.766	1.201	3.288
11	4.9843	6	0.769	1.965	4.208
12	5.4044	5.6666	0.772	1.835	3.967
13	5.838	4.7692	0.796	1.927	4.064
14	5.9635	4.6666	0.843	3.110	5.199
15	5.9714	5	0.833	2.459	4.547
16	5.7661	4.8571	0.790	2.147	4.257
17	5.2518	4.8571	0.784	2.262	4.403
18	5.9147	5.0714	0.827	2.475	4.562
19	6.0421	4.6666	0.782	2.104	4.188
20	5.9057	4.3783	0.842	2.932	4.977
21	5.4122	5.1034	0.785	2.51	4.621
22	5.5789	4.3428	0.807	2.643	4.727
23	5.1093	6.4	0.809	2.246	4.419
24	5.4737	4.875	0.782	2.536	4.617
range	4.98-6.48	3.01-6.5	0.75-0.86	1.20-3.12	3.34-5.20

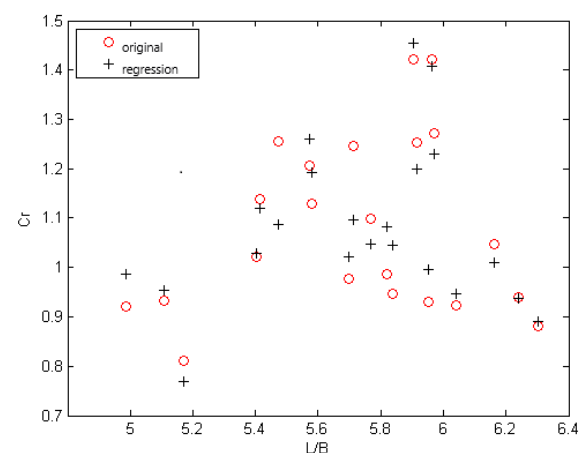


Fig. 1. Comparison between the Regression Result and the Original C_r about L/B ,

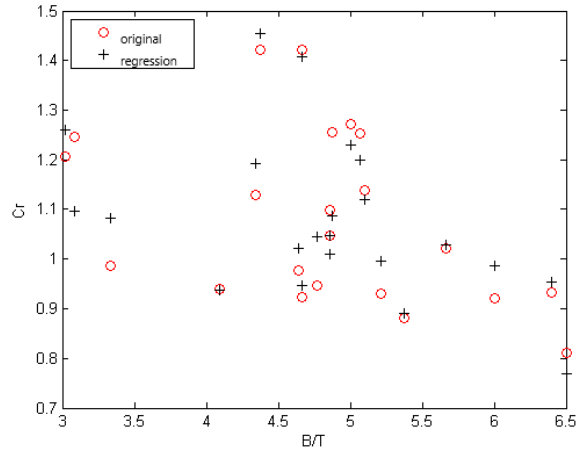


Fig. 2. Comparison between the Regression Result and the Original Cr about B/T

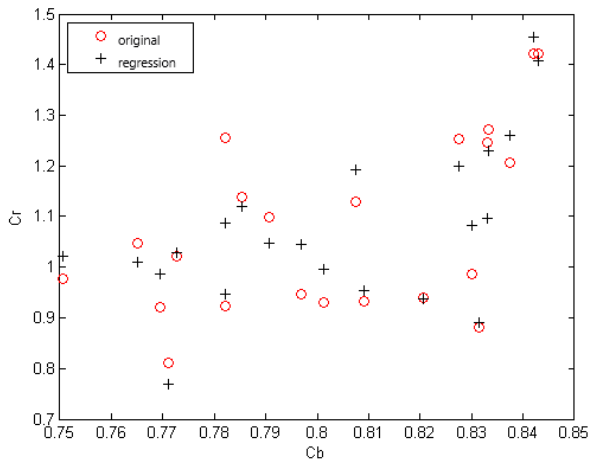


Fig. 3. Comparison between the Regression Result and the Original Cr about Cb

The friction drag coefficient of the real ship can be calculated by the ITTC57 formula. Using the regression value of C_r obtained above, the total drag coefficient of the ship can be obtained. Table 2 is the error comparison between calculated C_t calculated by regression C_r and the original test data. As the rejection points of the sample data under different Fr may be different, the number column in the table is different according to different Fr .

Comparing it with the original experimental data, the error is within $\pm 5\%$. It can be seen that the two data fits well, and the method has a certain degree of credibility.

3.3. Regression results under each Fr

For each Fr , the author calculates its coefficient according to the above process with the following formula (7):

$$C_r = b_1 + b_2\left(\frac{L}{B}\right) + b_3\left(\frac{B}{T}\right) + b_4(C_b) + b_5\left(\frac{L}{B}\right)(C_b) + b_6\left(\frac{B}{T}\right)(C_b) + b_7\left(\frac{B}{T}\right)^2 + b_8(C_b)^2 \quad (7)$$

Table.2 Regression value and original test data under $Fr=0.15$

num	C_b	Original data	Regression data	Error
3	0.83	3.049	3.147	3.22%
4	0.8375	3.2602	3.3128	1.61%
6	0.7505	3.198	3.2426	1.40%
7	0.765	3.2218	3.1851	-1.14%
8	0.8012	3.0935	3.1592	2.12%
9	0.8315	2.9688	2.9768	0.27%
10	0.833	3.2903	3.1412	-4.53%
11	0.771	3.067	3.0258	-1.34%
12	0.7694	3.264	3.33	2.02%
13	0.7727	3.2443	3.2506	0.19%
14	0.7968	3.1749	3.2735	3.10%
15	0.843	3.5982	3.5832	-0.42%
16	0.8332	3.4475	3.4068	-1.18%
17	0.7906	3.2968	3.2456	-1.55%
19	0.8275	3.4281	3.3731	-1.60%
20	0.7822	3.0962	3.1178	0.70%
21	0.8421	3.5497	3.5838	0.96%
22	0.8206	3.0303	3.0273	-0.10%
23	0.7853	3.331	3.3103	-0.62%
24	0.8075	3.3	3.363	1.91%
25	0.809	3.199	3.2205	0.67%
26	0.7822	3.423	3.2558	-4.88%

After repeating linear regression and eliminating the abnormal point, the following table 3 is finally obtained. This result makes RMSE, the statistical value F and the corresponding probability p meet certain requirements.

Table.3 Regression Equation Coefficients under Each Fr

coefficients	0.15	0.16	0.17	0.18
b1	4.428	22.312	42.181	68.047
b2	7.357	4.917	0.865	-3.602
b3	5.101	2.964	1.390	-1.821
b4	-95.222	-108.53	-119.55	-131.96
b5	-9.676	-6.810	-1.840	3.688
b6	-4.120	-1.690	0.135	4.015
b7	-0.186	-0.172	-0.164	-0.161
b8	110.426	101.578	85.696	62.175
coefficients	0.19	0.2	0.21	0.22
b1	101.256	82.135	134.723	127.868
b2	13.241	-2.609	-6.901	5.497
b3	-2.946	-0.790	-4.849	-0.570
b4	-140.22	-181.46	-257.17	-352.75
b5	16.021	2.161	7.405	-9.256
b6	4.728	3.398	8.382	6.022
b7	-0.099	-0.208	-0.215	-0.466
b8	20.698	102.156	116.668	246.722

It can be seen from above the constant term in the formula increases gradually, which determines that C_r

will show an increasing trend with the increase of Fr . While the coefficients containing only C_b terms (namely b_4 and b_8) have larger absolute values, which shows a greater impact on the rest resistance.

4 Calculation example analysis

The DTMB5415 selected in this paper is a preliminary design of a warship in Taylor tank around 1980. It has a bow sonar housing and a square stern. The main dimension parameters are shown in Table 4 below.

Tab.4 Main Dimension Parameters

Parameters	values
L_{pp}/m	5.72
B/m	0.76
T/m	0.248
S/m^2	4.861
C_b	0.506
Re	1.26×10^7
λ	24.8

In my previous study on the optimization of trimaran layout^[3], the drag coefficient of the model under each Fr was simulated and it is shown in Figure 4 below.

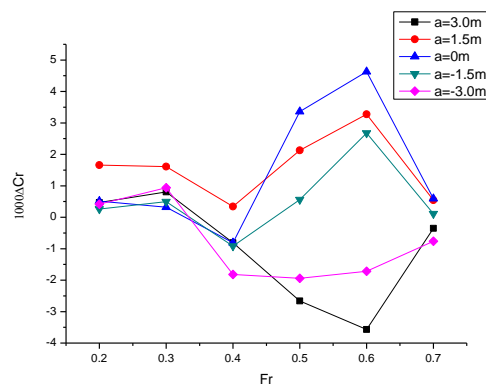


Fig. 4. C_r of Trimaran in Different Layouts

In my previous study on the optimization of trimaran layout, the coefficient of this model under each Fr was simulated. Firstly, take the optimal layout (the side hull 1.5m behind the main hull) in my simulation results. And then, the regression calculation value of the monohull is converted into C_t of the trimaran. Finally, after comparing the CFD simulation results with the regression prediction value, it is found that the error is within 10% in the following Table.5.

The main reasons for the error are as follows: (1) Because the theoretical analysis method fails to take into account the influence of the main and side hull interference on the rapid performance of the trimaran. (2) Due to the certain differences in the number of sample and ship types, the regression formula has a certain limitations.

Tab.5 Contrast Error

$C_t \times 10^3$	$Fr=0.2$	$Fr=0.25$	$Fr=0.3$
CFD simulation results	4.57	4.66	4.91
regression prediction value	4.28	4.51	4.79
error	6.77%	3.33%	2.50%

5 Conclusion

For the resistance prediction of the trimaran, the experimental method is undoubtedly accurate, but the cost is relatively high. The simulation method is also very common, and the interference between the main and side hull is also involved. But the modeling workload is relatively large and the conclusion is affected by the quality of the grid. The regression method involved in my article can be used in the preliminary design stage to estimate the resistance value quickly and roughly. The above case proved that the numerical error is not too large. This method has certain reference value, which can predict the resistance performance of the ship.

Compared with the typical CFD calculation, the method of multiple regression saves a lot of time on the basis of close accuracy, and the ability of both the sample simulation and the large-range prediction is much better than the CFD simulation. To some extent, it can be used as an effective means to predict ship resistance performance in the early stage.

Acknowledgement

PROGRAM: Hubei Provincial Department of Education Scientific Research Guidance Program (num as B2018341)

References

1. Lars Larsson, Frederick Stern and Volker Bertram. Benchmarking of Computational Fluid Dynamics for Ship Flows: The Gothenburg 2000 Workshop. Iowa City. USA. 2003.
2. Leonardo Di Gangi, M. Lapucci, F. Schoen, A. Sortino. An efficient optimization approach for best subset selection in linear regression, with application to model selection and fitting in autoregressive time-series[J]. Computational Optimization and Applications, 2019, 74(3)
3. Xiao Ying Xu, YAN Hong Sheng, LIU Feng Di. Tension hybrid optimization algorithm for automatic mooring positioning system[J]. Ship Engineering, 2013(04) (in Chinese)
4. Zhen Bang Sheng, Principles of Ships. Shanghai Jiaotong University Press. 2020.7 (in Chinese)