

**Prediction Model Using Multiple Regression Analysis
for Relationship Between Forward Speed and Propeller Rotational Speed
of Cruising AUV in Actual Seas**

by

UMEDA Jun^{*}, FUJIWARA Toshifumi^{*}

Abstract

We devised a prediction model for the relationship between forward speed and propeller rotational speed of an autonomous underwater vehicle (AUV) using multiple regression analysis. When designing and operating an AUV, predicting its propulsion performance is important for estimating diving distance and operating time without power supply. To estimate the propulsive performance of the AUV, it is important to predict the propeller rotation speed and propeller power required for the propeller thrust to balance the resistance of the AUV at a given speed using experiments and numerical simulations. Additionally, although data on the propulsive performance of AUVs in actual seas needs to be measured and compared with tank tests and numerical simulations, the comparison requires processing that takes into account factors such as scaling, steering, and AUV motion. Thus, the purpose of this research is to develop a prediction model for the relationship between forward speed and propeller rotational speed considering other factors using the regression analysis based on data observed in actual seas. This predictive model should be comparable to the self-propulsion point estimated by tank tests and computational fluid dynamics (CFD). To demonstrate an example application of the regression model, we compared the results predicted by the regression model with those predicted by the self-propulsion simulation. The results of the regression model agreed with those calculated by CFD. The regression model proposed in this research can be used to validate the CFD results in an AUV instead of comparing with measured data directly.

^{*} Offshore Advanced Technology Department

Received January 26th, 2022.

Accepted March 8th, 2022

Contents

1. Introduction	63
2. Analysis method	64
2.1 Subjected AUV	64
2.2 Regression analysis method	65
2.3 Data set for regression analysis	67
3. Validation of the regression model	68
4. Application example of the multiple regression model	69
5. Conclusion	70
Acknowledgments	71
References	71
Appendix	72

List of symbols

Nomenclature

U	: Forward speed of the AUV [m/s]
w_0	: Intercept of the regression model [-]
\mathbf{w}	: Vector of regressive coefficients of the regression model [-]
\mathbf{x}	: Vector of independent variables of the regression model [-]
y	: Dependent variable [-]
M	: Number of the train data set [-]
$S_e(x)$: Standard error of x [-]
t	: t-statistic value [-]
λ	: Parameter that controls the strongness of the regularization term [-]
R^2	: The coefficient of determination [-]
V	: Variance inflation factor [-]
U	: Forward speed of the AUV [m/s]
n	: Propeller rotational speed [rpm]
ϕ	: Roll [deg]
θ	: Pitch [deg]
ψ	: Yaw [deg]
p	: Roll rate [deg/s]
q	: Pitch rate [deg/s]
r	: Yaw rate [deg/s]
δ_s	: Angle of starboard rudder [deg]
δ_b	: Angle of bottom rudder [deg]
δ_p	: Angle of port rudder [deg]
δ_t	: Angle of top rudder [deg]
(X, Y, Z)	: Position in a cartesian coordinate system [m]
R_n	: Reynolds number [-]
L	: Representative length of the AUV [m]
ν	: Kinematic viscosity [m ² /s]
D_p	: Diameter of the propeller [m]
T_p	: Thrust force of the propeller [N]
Q	: Propeller torque [Nm]

J	: Advanced coefficient [-]
ρ	: Density of the water [kg/m ³]
K_t	: Thrust coefficient [-]
K_q	: Torque coefficient [-]
η_0	: Open water propeller efficiency [-]

Subscriptions and Superscripts

$\bar{\quad}$: Mean value
$\hat{\quad}$: Predicted value
$\ \quad \ _p$: p -norm of vector
$'$: Standardized value
i	: i -th number of data in data set
j	: Element number of vectors
T	: Transpose operator

1. Introduction

Underwater robots are one of the important underwater technologies for inspections of marine structures, defenses, and ocean surveys. An autonomous underwater vehicle (AUV) is one type of underwater robot, and it is unmanned and has no external connections for powering and control ¹⁾. Estimation of the propulsion performance of an AUV is important to assess cruising distances and operating time without power supply, but this estimation is not easy because of complex hydrodynamic interaction between an AUV body and a propeller. Propulsion performance of an AUV can be estimated by tank tests and numerical simulation²⁾⁻⁵⁾.

There are two important points to note about these methods. The first is how to evaluate performance on a full-scale AUV. The performance of a small model in a water tank test may differ from that of a full-scale AUV due to scale effects. For AUVs over 4 m in length, tank tests of a full-scale AUV are difficult due to the limited dimensions of the tank, so numerical calculations such as CFD are often used to estimate the performance of a full-scale AUV. Therefore, it is important to evaluate reliability of the results of CFD calculations on a full-scale AUV. The second is the influence of steering, motion, and other factors. Estimation by tank tests and numerical simulation usually does not consider rudder and AUV motion. Performance estimated by tank tests and numerical simulation is, namely, the performance of a AUV under ideal conditions. On the other hand, propulsion performance of an AUV in tank test and numerical simulation differs from that in actual conditions because that of an AUV in actual seas varies continuously due to multiple factors such as steering, motion, and ocean currents.

It is simplest to measure data on the propulsive performance of AUVs in actual seas and compare it with tank tests and numerical simulations. Direct comparisons are difficult because of the variety of factors involved in the measured data. In other words, it is necessary to process the data in some way that considers factors such as scaling, steering, and AUV motion in order to make the data comparable to the measured data.

In previous research on the performance estimation of an AUV in actual seas, the hydrodynamic coefficients in the mathematical model describing the dynamic behavior of an AUV were estimated by using a system identification ^{5), 6)}. An advantage of system identification is that the coefficients obtained by system identification can be directly compared with those obtained from model tests or numerical simulations. System identification including model of steering and AUV motion makes it possible to consider effects of steering and AUV motion. Problems with system identification are overfitting and overparameterization. Overfitting can occur when observed data are noisy, and it can cause overestimation of coefficients obtained by system identification. Overparameterization can cause convergence ratio to become slow and unstable. Using a simplified model, therefore, is desirable to avoid overparameterization and overfitting.

As another situation, an AUV may have flexible cables for ocean exploration in several cases. Investigation of AUV performance in these cases requires the consideration of effects of towed flexible cables in addition to the dynamic characteristics of an AUV. The dynamic model of the towed flexible cables can be modeled by the lumped-mass method, but the model tends to be complicated ⁷⁾. Accordingly, a simplified method is required to consider the effects of the flexible cables in addition to steering and AUV motion.

Propulsion performance of an AUV includes various factors, however it denotes a self-propulsion point of an AUV in this paper. The self-propulsion point is defined as the condition of an AUV and a propeller when propeller thrust balances AUV resistance. In general, the propeller rotation speed and propeller power that are required for the propeller thrust to balance the resistance of the AUV at a given speed can be obtained from propeller-body interaction characteristics (wake fraction and thrust deduction coefficient) determined from water tank tests and CFD. In this research paper development of a prediction model for the relationship between forward speed and propeller rotational speed of a subjected cruising AUV is addressed because this predictive model is comparable with the self-propulsion point estimated by tank tests and CFD.

This model presented in the paper can consider factors such as flexible cables, steering, and motion of the AUV. It is because the prediction model considering these factors leads to an estimation of propulsion performance of the AUV in various conditions. This research introduces two multiple regression analyses based on variable selection and sparse modeling to predict the relationship between forward speed and propeller rotational speed. The prediction performance of the two regression models is compared with each other.

Although measured data cannot be directly compared with numerical analysis or experimental results, predictive models based on measured data using multiple regression analysis may be compared with numerical analysis or experimental results. As an example of application, the self-propulsion points of the subjected AUV are calculated from the self-propulsion simulation by computational fluid dynamics (CFD). The results of the CFD and the predictive model are compared with the regression model for validation of the CFD.

2. Analysis method

2.1 Subjected AUV

Figure 1 and Table 1 show appearances and principal particulars of NMRI Cruising AUV#3, respectively. The National Maritime Research Institute, Japan (NMRI) has developed four cruising AUVs ⁸⁾. In them, NMRI Cruising AUV#3 especially has a high motion characteristic because the AUV has the capacity that the maximum navigating pitch angles during descending and ascending are -45 and $+70$ degrees, respectively. Actuators of the AUV are cross-type four rudders for motion control and a thruster with a duct as a propulsion system.

The onboard sensors shown in Table 1 are used for navigation and exploration. The sensors for navigation are an attitude heading reference system, a doppler velocity log, and a depth sensor. The survey sensors are a multi-beam echo sounder, a CT sensor, and a self-potential sensor. The multi-beam echo sounder is used to obtain seafloor topography. The self-potential sensor is used to detect naturally occurring electric fields in seafloor mineral deposits ⁹⁾. This self-potential sensor is removable depending on the purpose of the survey.

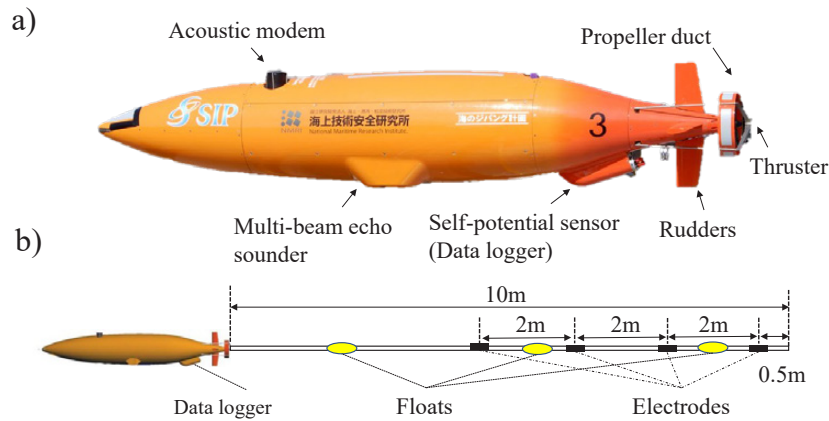


Figure 1 Appearances of NMRI Cruising AUV#3 with a self-potential sensor.

Table 1 Principal particulars of NMRI Cruising AUV#3.

Item	Value and contents
Length overall	3.90 m
Maximum diameter of hull	0.65 m
Weight in air	545 kg
Maximum operation depth	2000 m
Cruising speed	1.80 m/s
Actuators	One thruster, Cross-type four rudders
Sensors	Attitude heading reference system, Doppler velocity log, Depth sensor, Multi-beam echo sounder, CT sensor, Self-potential sensor

2.2 Regression analysis method

Multiple linear regression is a common modeling technique to analyze the correlation between independent variables and dependent variables. Multiple linear regression models can be expressed as follows:

$$\hat{y} = \mathbf{w}\mathbf{x} + w_0 \tag{1}$$

The regression coefficients in the model can be obtained by solving the optimization problem defined as follows ¹⁰⁾:

$$\min. \sum_{i=1}^M (y_i - \hat{y}_i)^2 \tag{2}$$

Minimal independent variables are desirable to avoid overfitting and multicollinearity in the regression model. Therefore, only the significant independent variables were selected. Variable selection was based on backward elimination ¹¹⁾. First, a regression model using all independent variables was established. The independent variable with the smallest contribution factor was selected for removal. If the variable satisfied a removal criterion, the variable was removed from the candidates of the significant variables. The process was repeated until there are no variables in the equation that satisfy the removal criterion.

The removal criterion is applied to a hypothesis test based on the t-Test. The statements for the hypothesis test are expressed as:

$$\begin{aligned} H_0 &: w_j = 0 \\ H_1 &: w_j \neq 0 \end{aligned} \tag{3}$$

where, H_0 and H_1 mean the null and alternative hypotheses, respectively. The t-statistic is an index for the null hypothesis via the t-Test. The t-statistic value is defined by

$$t = \frac{w_j}{S_e(w_j)} \quad (4)$$

The P-value was calculated from the t-statistic and the t-distribution table. H_0 in the null hypothesis is rejected if the P-value was less than a significance level. When the null hypothesis for each coefficient is rejected, the result supports the hypothesis that there is a significant relationship between the independent variable and the dependent variable. In this research, the significance level was set to 0.05.

Multicollinearity causes two serious problems in the multiple regression analysis. Due to multicollinearity, the predictions of the regression model are extremely sensitive to even slight changes. Multicollinearity prevents the regression model from predicting the dependent variable accurately. Moreover, since multicollinearity affects the P-value, the evaluated significance of independent variables is unreliable. Dependent variables with high correlation are eliminated to solve multicollinearity.

The variance inflation factor (VIF) provides an index that assesses multicollinearity. The VIF of each variable was estimated to select the variables. First, the multiple linear regression model for x_j as a function of all the other variables was determined. The linear regression model for x_j that is an element in the vector of \mathbf{x} is expressed by as follows:

$$\hat{x}_j = \mathbf{w}_j \mathbf{x}_{j,i} + w_{j0} \quad (5)$$

The VIF and the coefficient of the determination of x_j are defined as follows:

$$V_j = \frac{1}{1 - R_j^2} \quad (6)$$

$$R_j^2 = 1 - \frac{\sum_{i=1}^M (x_{j,i} - \hat{x}_{j,i})^2}{\sum_{i=1}^M (x_{j,i} - \bar{x}_j)^2} \quad (7)$$

The VIF of 5 or more indicates multicollinearity in the regression model. The solution for multicollinearity is to remove one (or more) of the highly correlated dependent variables.

There is regression analysis that can automatically select significant independent variables. It is the least absolute shrinkage and selection operator (LASSO)¹²⁾. In the multiple linear regression, the regression coefficients are determined to minimize the residual sum of squares of Eq. (2). In the LASSO, the 1-norm was added Eq. (2) to avoid overfitting. The regression coefficients in the LASSO can be determined by solving an optimization problem defined as follows:

$$\min. \sum_{i=1}^M (y_i - \hat{y}_i)^2 + \lambda \|\mathbf{w}\|_1 \quad (8)$$

This optimization problem can be solved numerically with several software packages¹³⁾.

The LASSO model was compared with the multiple regression model to validate the significant variables. The LASSO performs both determination of regressive coefficients and variable selection. The LASSO tends to produce regressive coefficients of unnecessary dependent variables as zero.

Setting the proper value of λ is essential to the performance of the LASSO. When the parameter, λ , increases, significant variables may be eliminated, and regression coefficients can be shrunk excessively. In this research, the parameter was 0.15 on the basis of several trials that maximum the coefficient of the determination defined by¹⁰⁾

$$R_y^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \quad (9)$$

2.3 Data set for regression analysis

The data set for the regression analysis were observed in several sea trials from the year 2018 to 2020^{8), 14)}. The sampling time of the data is 0.25 s, and in total, five hours of the data were used for regression analysis. The data set was randomly split into 85% for training and 15% for testing to validate the regression model. The number of training data in this research is about 500.

Figure 2 shows observed propeller rotational speeds versus forward speeds of the AUV. The blue and red markers show the results without and with the self-potential sensor. The forward speeds increased roughly in proportion to the propeller rotational speed. However, there were variations in the propeller rotational speeds and forward speeds depending on the steering and the presence or absence of the self-potential sensor. Propeller rotational speeds with the self-potential sensor were higher than ones without the self-potential sensor due to the drag of the self-potential sensor.

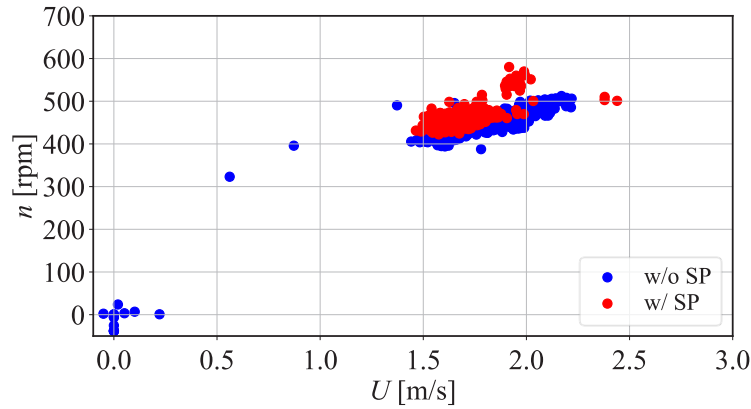


Figure 2 Propeller rotational speeds versus forward speeds of the AUV. Red and blue markers show data observed with and without the self-potential sensor, respectively.

Regression analysis requires pre-processing for the training data set. The observed data were filtered by a simple moving average every minute. Moreover, normalization is required to adjust values measured on different scales to a notionally common scale. Several normalization techniques were proposed in statistics¹³⁾. In this research, the maximum absolute scaling was selected as follows:

$$x'_j = \frac{x_j}{\max(|x_j|)} \quad (10)$$

The maximum absolute scaling translates each variable into a value between -1.0 and 1.0 by dividing every observation by its maximum absolute value.

Since the prediction of the self-propulsion point was focused on in this research, the relationship between propeller rotational speed and forward speed were modeled considering disturbances such as steering. In this paper, U means the forward speed through the water to consider the effects of tidal currents. The independent variables and the dependent variable in the two regression models are defined by the following equations:

$$\mathbf{x} = [U' \quad \phi' \quad \theta' \quad \psi' \quad p' \quad q' \quad r' \quad \delta'_s \quad \delta'_b \quad \delta'_p \quad \delta'_t \quad s]^T \quad (11)$$

$$\hat{y} = n \quad (12)$$

where, s in Eq. (11) means the dummy variable whether the self-potential sensor is installed or not. When the sensor is installed to the AUV, the dummy variable equals one, and when it is uninstalled, the dummy variable equals zero.

3. Validation of the regression model

Table 2 shows the determined regression coefficients and R_y^2 by using the multiple linear regression model and the LASSO model on the basis of the analysis introduced in the previous section. The regression coefficient of zero means that the independent variable in Eq. (11) is determined to be insignificant. Since R_y^2 of all models were more than 0.90, the two models were agreed with the training and test data set. R_y^2 for test data of the multiple linear regression was larger than that of the LASSO.

Table 2 Regression coefficients and R_y^2 of the multiple linear regression model and LASSO model.

Item	Multiple linear regression	LASSO
w_0	7.77	15.3
ϕ'	-75.9	-59.3
θ'	-41.1	-39.7
ψ'	0.0	-1.94
p'	0.0	0.0
q'	0.0	0.0
r'	0.0	0.0
δ'_s	0.0	0.0
δ'_b	24.8	25.1
δ'_p	-31.8	-29.9
δ'_t	0.0	0.0
s	27.4	32.7
U'	563	546
R_y^2 for training data	0.921	0.921
R_y^2 for test data	0.927	0.925

The regression coefficients of the two regression models showed a similar trend, but there were some differences, such as ψ' in the multiple regression was zero and removed, while it was not zero in LASSO. The angular speeds (p' , q' , r') did not affect predicted values because observed data were averaged every minute. δ'_s and δ'_p are the rudder angles of elevator rudders, and the two rudders often were manipulated simultaneously. Since the correlation between δ'_s and δ'_p became high because of simultaneous operations, δ'_s was removed to solve multicollinearity due to the high correlation. δ'_t also was removed for the same reason.

Since residuals of values predicted by the regression model should agree with a normal distribution, the distribution of the residuals is shown to verify the regression model. Figure 3 shows a quantile-quantile plot for the residuals of the multiple regression and the LASSO model. The quantile-quantile plot visualizes how the shapes of the two distributions are similar. The line and markers in Fig. 3 show the theoretical normal distribution and residuals of the regression model. When the markers agree with the line, the distribution of residuals agreed with the normal distribution. Figure 3 shows the residuals of the regression model agreed with the normal distribution.

Since the LASSO model was equivalent to the multiple regression model with variable selection, the LASSO is compatible to the multiple regression with variable selection. The LASSO is recommended because LASSO can remove insignificant independent variables with few steps when the number of independent variables is large.

Figure 4 shows the predicted results and observed data. Since R_y^2 for test data in the multiple linear regression was larger than that in the LASSO, the multiple regression model was selected as the prediction model. Black cross and red circle markers in Fig. 4 show the predicted results and the data observed with the self-potential sensor. Yellow cross and blue circle markers in Fig. 4 show them without the self-potential sensor. The predicted results agreed with the observed data. That is, the statistical model based on regression analysis shows the relationship between the propeller rotation speed and the forward speed of the AUV considering several factors such as flexible cable dynamic behavior, steering, and AUV motion.

The data used in the regression analysis was obtained at a constant speed of 1.6~2.4 m/s. Outside this range, the regression model has low prediction accuracy. Note that the model is also not applicable to large rudder angles (>10 degrees) such as when turning.

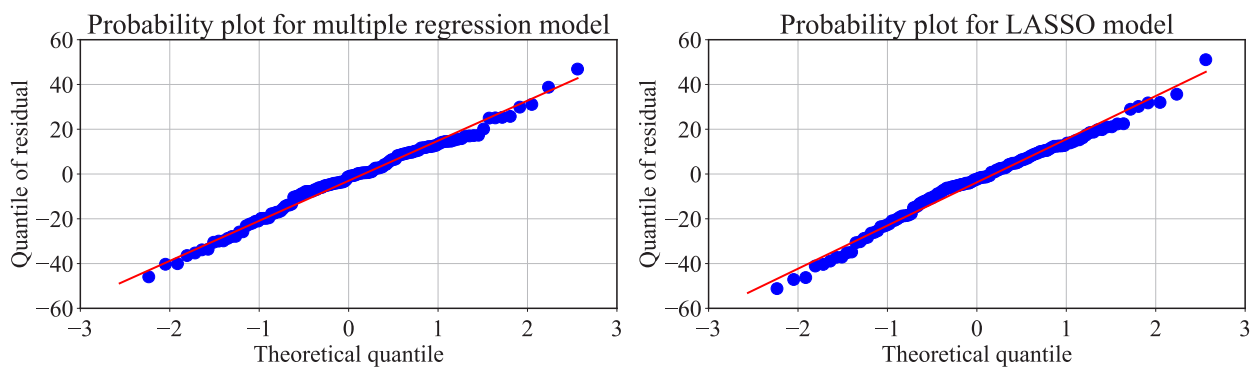


Figure 3 Quantile quantile plots of the residuals (Left: multiple regression model, Right: LASSO model).

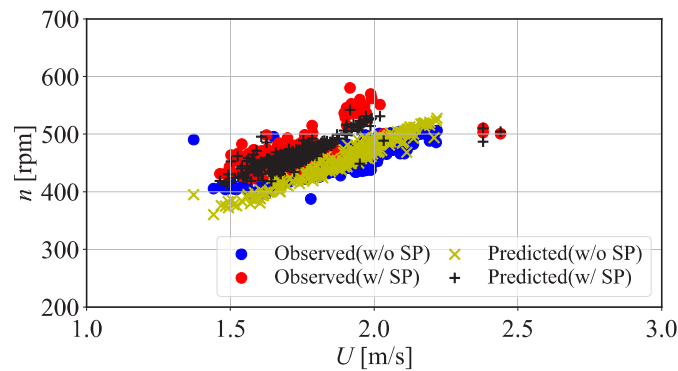


Figure 4 Comparison of observed data and results predicted by the multiple regression model.

4. Application example of the multiple regression model

In this section, an application example of the multiple regression model is presented. Regression analysis yielded the relationship between the propeller rotational speed and the forward speed of the AUV, including other factors. When the independent variables except for the forward speed are equal to zero, the regression model can predict the self-propulsion points under the same ideal conditions as CFD and other models. Therefore, the regression model predicts the propeller rotational speed using only the forward speed with and without the self-potential sensor, so that it can be compared with CFD.

Self-propulsion simulations by the CFD were conducted to estimate the speed-rpm curve of the subjected AUV. The detailed information of the CFD calculation is shown in Appendix. Note that CFD does not take into account the flexible cable of the self-potential sensor because of the difficulty of modeling it.

Figure 5 shows the regression model and the results calculated by the CFD with and without self-potential sensor. The regression model without the self-potential sensor derived from the data set observed in the actual seas agreed with the results calculated by the CFD without the self-potential sensor. Due to the increased resistance of the AUV caused by the self-potential sensor, the propeller rotation speed is increased in the regression model. On the other hand, the propeller rotation speed is almost unchanged in the CFD calculation.

The self-potential sensor consists of a flexible cable and a data logger. However, fluid phenomena in the cable section were not considered in CFD calculations because modeling the flexible cable was difficult. Therefore, the difference between the regression model and CFD calculations is reasonable. In CFD calculation, the cable must be modeled using methods such as those proposed in previous research.

Measured data were not directly compared to the numerical analysis, but the predictive model was created from the measured data using multiple regression analysis and compared to the CFD calculations. The results show that the regression model can be used to validate against the CFD results.

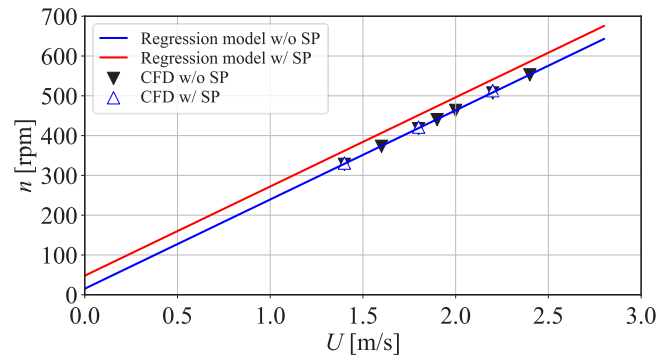


Figure 5 Comparison of the regression model and the results calculated by the CFD w/o and w/ self-potential sensor

5. Conclusions

This research addressed the development of a prediction model for the relationship between forward speed and propeller rotational speed of a subjected cruising AUV to make this predictive model comparable with the self-propulsion point estimated by tank tests and CFD. The model developed in the paper can consider factors such as flexible cables, steering, and AUV motion.

The multiple regression with variable selection and the LASSO model were applied for regression analysis. The two regression models were agreed with the data observed in the various conditions. The prediction performance of the two analysis was equivalent. Although the LASSO and the multiple regression with variable selection can be utilized, the LASSO that can automatically remove insignificant independent variables is recommended to suppress the computational cost for variable selection increases.

As an application example of the regression model, the regression model was compared with the self-propulsion points calculated from self-propulsion simulations by CFD. The accuracy level of CFD method on self-propulsion simulation was able to be confirmed using the regression model presented in this paper. The regression model proposed in this research are available for validation of CFD simulation results in an AUV instead of comparing with measured data directly.

Acknowledgments

This research was supported by Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), “Innovative Technology for Exploration of Deep Sea Resources” (Lead agency: Japan Agency for Marine-Earth Science and Technology, JAMSTEC).

We would like to acknowledge Dr. Kim, K., Dr. Sasano, M., Dr. Sakamoto, N., Mr. Sato, T., Mr. Inaba, S., et.al. in the National Maritime Research Institute for the collection of the in-service data and the CFD calculation.

References

- [1] Sahoo, A., Dwivedy, S. K., and Robi, P. S., Advancements in the Field of Autonomous Underwater Vehicle, *Ocean Engineering*, Vol. 181 (2019), pp. 145–160.
- [2] Allen, B., Vorus, W. S., and Prestero, T., Propulsion System Performance Enhancements on REMUS AUVs, *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings (Cat. No.00CH37158)*, IEEE, Providence, RI, USA (2000), pp. 1869–1873.
- [3] Sezen, S., Dogrul, A., Delen, C., and Bal, S., Investigation of Self-Propulsion of DARPA Suboff by RANS Method, *Ocean Engineering*, Vol. 150 (2018), pp. 258–271.
- [4] Phillips, A. B., Turnock, S. R., and Furlong, M., Comparisons of CFD Simulations and In-Service Data for the Self Propelled Performance of an Autonomous Underwater Vehicle, *27th Symposium of Naval Hydrodynamics (2007)*.
- [5] Hegrehaes, O., Hallingstad, O., and Jalving, B., Comparison of Mathematical Models for the HUGIN 4500 AUV Based on Experimental Data, *Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, IEEE, Tokyo (2007), pp. 558–567.
- [6] Cardenas, P., and de Barros, E. A., Estimation of AUV Hydrodynamic Coefficients Using Analytical and System Identification Approaches, *IEEE J. Oceanic Eng.*, Vol.45, No.4 (2020), pp. 1157–1176.
- [7] Du, X., Cui, H., and Zhang, Z., A Numerical Method for Analyzing the Influence of Underwater Vehicle Flow Field on Dynamic Behavior of Towed Sonar Cable Array, *Ocean Engineering*, Vol. 175 (2019), pp. 163–175.
- [8] Sato, T., Kim, K., Inaba, S., Matsuda, T., Takashima, S., Oono, A., Takahashi, D., Oota, K., and Takatsuki, N., Exploring Hydrothermal Deposits with Multiple Autonomous Underwater Vehicles, *IEEE Underwater Technology (2019)*, pp. 1–5.
- [9] Kawada, Y., and Kasaya, T., Self-Potential Mapping Using an Autonomous Underwater Vehicle for the Sunrise Deposit, Izu-Ogasawara Arc, Southern Japan, *Earth Planets Space*, Vol.70 No.1 (2018), pp.1 – 15.
- [10] Eisenhauer, J. G., *Regression through the Origin*, *Teaching Statistics*, Vol. 25, No.3 (2003), pp. 76–80.
- [11] Heinze, G., Wallisch, C., Dunkler, D., *Variable selection – A review and recommendations for the practicing statistician*, *Biometrical Journal*, Vol. 60, No.3 (2018), pp. 431-449.
- [12] Tibshirani, R., *Regression Shrinkage and Selection via the Lasso*, *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol.58, No.1 (1996), pp. 267–288.
- [13] Scikit-learn developers, 6.3. Preprocessing data, scikit-learn, Dec. 08, 2021, <https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing>.
- [14] Umeda, J., Kim, K., Sato, T., Inaba, S., and Fujiwara, T., Numerical Simulations and Sea Trial for Leader Follower Formation Control of Cruising-Type AUVs, *Proceedings of the ASME International Conference on Ocean, Offshore and Arctic Engineering 2020*, Vol.6 (2020).

Appendix

Self-propulsion simulation using CFD were conducted for investigation the flow around an AUV body and interaction between the AUV body and propeller in detail. The viscous CFD code, NAGISA, was used. This code is an in-house code developed at National Maritime Research Institute (NMRI) in Japan ^{A1)}. The code solves mass-conservation and incompressible Reynolds-averaged Navier Stokes (RaNS) equations by overset structured-grids based finite-volume method. Domain connectivity information for the overset grid was computed by the in-house overset grid assembler developed at NMRI, UP_GRID ^{A2)}.

Figure A1 shows the computational domains and the boundary conditions in CFD. The computational grid consists of the rectangular grid of the entire computational domain, the body of the AUV, the acoustic modem, the self-potential meter, the four rudders (top, bottom, left, and right), and a propeller duct. The effect of the propeller is considered by applying a body force to the RaNS equation. The size of the external zone in X, Y, and Z directions are $-3.0 \leq X/L \leq 2.0$, $-1.5 \leq Y/L \leq 1.5$, and $-2.5 \leq Z/L \leq 0$, respectively. The AUV was located at the center of the computational zone.

Table A1 shows the number of cells in the block. Minimum spacing next to the body surface satisfied that the non-dimensional wall distance was smaller than 1.0 according to the criteria suggested by the ITTC ^{A3)}. It means that the low-Reynolds treatments were utilized for full-scale simulations. The total number of cells was approximately 4.9 million.

The no-slip condition was applied to the boundary condition on the surface of the AUV. The symmetry condition was imposed on the top and bottom surfaces of the external zone. The outflow was imposed on the starboard and afterward surfaces of the external zone. The inflow was imposed on the port and forward surfaces of the external zone.

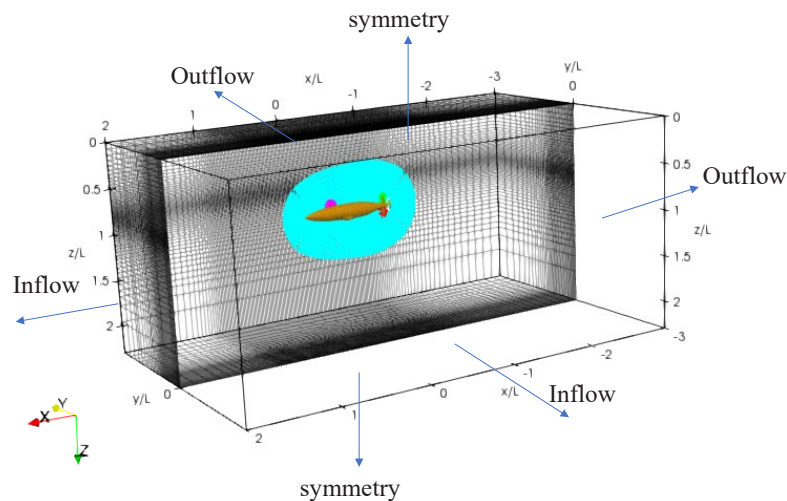


Figure A1 Computational domain and boundary conditions in the CFD.

Since the calculations were conducted in full scale, the model-scale correction was not considered. In the calculation, the free surface was not considered, and $k-\omega$ explicit algebraic stress model was applied to the turbulence model. The Reynolds number (R_n) was $4.54 \times 10^6 - 7.78 \times 10^6$, which corresponds to the forward speed of 1.4 m/s to 2.4 m/s. The definition of the Reynolds number is as follows:

$$R_n = UL/\nu \quad (\text{A1})$$

Table A1 Number of cells in the CFD.

Block name	Duct	Bottom rudder	Top rudder	Starboard rudder	Port rudder	Self-potential sensor	Acoustic modem	AUV body	External zone
Cells	599 K	329 K	329 K	329 K	329 K	638 K	358 K	860 K	1,131 K

The evaluation of the propulsion performance requires the resistance and the open-water characteristics of the propeller. The resistance of the AUV were calculated directly by the CFD. The open-water characteristics of the propeller were calculated by the body force theorem. Self-propulsion points were calculated by the thrust identity method.

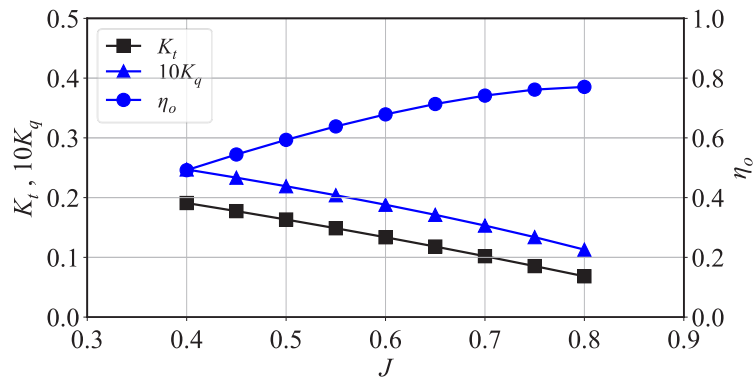
Figure A2 shows the propeller open-water characteristics calculated by the body force theorem. The horizontal axis in Fig. A2 is the advance coefficient, and K_t , K_q , and η_o in the vertical axes are thrust coefficient, torque coefficient, and propulsor efficiency ratio. The definitions of these parameters are as follows:

$$J = \frac{U}{nD_p} \tag{A2}$$

$$K_t = \frac{T_p}{\rho n^2 D_p^4} \tag{A3}$$

$$K_q = \frac{Q}{\rho n^2 D_p^5} \tag{A4}$$

$$\eta_o = \frac{J K_T}{2\pi K_Q} \tag{A5}$$



FigureA2 Propeller open-water characteristics used in the self-propulsion simulations.

References

- [A1] Ohashi, K., Hino, T., Kobayashi, H., Onodera, N., and Sakamoto, N., Development of a structured overset Navier–Stokes solver with a moving grid and full multigrid method, *Journal of Marine Science and Technology*, Vol. 24 (2019), pp. 884-901.
- [A2] Kobayashi, H., and Kodama, Y., Developing spline based overset grid assembling approach and application to unsteady flow around a moving body, *Journal of Mathematical and System Science*, Vol. 6 (2016), pp. 339-347.
- [A3] International towing tank conference: Practical Guidelines for Ship CFD Applications, ITTC Recommended Procedures and Guidelines 7.5-03-02-03 (2014).