

Література

[1]. Blintsov V.S., Klochkov O.P. Generalized method of designing unmanned remotely operated complexes based on the system approach. Scientific journal «EUREKA: Physics and Engineering». 2019. Vol. 2 (21). P. 43-51.

[2]. Kris Osborn. The Navy's Unmanned Surface Vessels Will Be Hungry for Energy. <https://nationalinterest.org/blog/buzz/integrated-power-and-energy-systems-will-power-21st-century-weapons-199462>

[3]. Joel Coito. Maritime Autonomous Surface Ships: New Possibilities—and Challenges—in Ocean Law and Policy. Published by the Stockton Center for International Law. Volume 97, 2021. 49 Pages. <https://digital-commons.usnwc.edu/cgi/viewcontent.cgi?article=2955&context=ils>

[4]. The Future is Now: Unmanned and Autonomous Surface Vessels and Their Impact on the Maritime Industry. Benedict's Maritime Bulletin December 2017. <https://www.blankrome.com/publications/future-now-unmanned-and-autonomous-surface-vessels-and-their-impact-maritime-0>

INFORMATION MODEL OF AN UNMANNED SURFACE VESSEL FOR OCEAN EXPLORATION

Volodymyr B., Doctor of Technical Sciences, Professor, Kherson Educational-Scientific Institute of Admiral Makarov National University of Shipbuilding, Kherson, Ukraine.

Nadtochii V., PhD, Kherson Educational-Scientific Institute of Admiral Makarov National University of Shipbuilding, Kherson, Ukraine.

Sabutskyi I., PhD, Admiral Makarov National Shipbuilding University, Mykolaiv, Ukraine.

Unmanned surface vehicles (USVs) are an effective technical means of collecting and transmitting spatial information about the marine environment. In the process of creating such a vessel, the information model forms a theoretical basis for building effective algorithms and programs for its operation. The synthesis of the information models is an important stage in the creation of vessels for the study of the aquatic environment using unmanned technologies.

Keywords: information model, unmanned surface vessel, information flow, aquatic environment.

UDC 004.942:519.25

PRELIMINARY PROCESSING DATA FOR BUILDING A MATHEMATICAL MODEL OF BUSINESS APPLICATIONS METRICS DEVELOPED ON C#

Makarova L.M.

candidate of technical sciences

Associate Professor of the Department of Software of Automated Systems

lidiia.makarova@nuos.edu.ua

Liutyi D.A.

PhD student of the Department of Software of Automated Systems

ljutyj.dima@gmail.com

Admiral Makarov National University of Shipbuilding

Mykolaiv, Ukraine

Abstract. The preliminary processing data for building a mathematical model for estimating the software size of open-source of business applications C# projects was observed in the paper. For more accurate estimation multivariate nonlinear regression must be used. The preliminary data processing to build the model is done by means of Variable Inflation Factors and Mahalanobis squared distance.

Keywords: software size estimation, C# projects, nonlinear regression analysis, multivariate regression, VIFs, Mahalanobis squared distance.

Introduction. Software size is very important metric in software development. The estimation of software size plays a key role during the prediction of the project effort and cost. COCOMO II model can be used for this [1].

Regression analysis is a classical statistical technique for building estimation models. It is concerned with describing and evaluating the relationship between a dependent variable and one or more independent variables. The model is built and evaluated through collecting sample data for these variables.

Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable. Linear regression fits a straight line or surface that minimizes the discrepancies between predicted and actual output values. Hence low number of outliers or more accurate dependency can be achieved using nonlinear regression.

The purpose of the work is the preliminary processing data for building a mathematical model for estimating the software size of open-source of business applications C# projects.

Main part. Multivariate nonlinear regression is a statistical method used to model the relationships between multiple independent variables and a dependent variable when the relationship is nonlinear. To build a nonlinear regression model, a preliminary data processing such as checking multicollinearity of independent variables and detecting outliers in source data must be accomplished.

Variable Inflation Factors (VIFs) is used for multicollinearity detection, and it shows if there is a dependency between variables. The variables must not have a multicollinearity to build a good regression model [2].

For detecting outliers in the data, the Mahalanobis squared distance (MSD) with the decimal logarithm for normalizing transformations can be used [3].

In this paper the data samples for estimation are taken from 33 open-source projects of business application developed using the C# programming languages. The data set to process: actual software size in the thousand lines of code (KLOC) Y , the total number of classes X_1 , the total number of relationships X_2 and the average number of attributes per class X_3 .

The data check on VIFs showed that there is no multicollinearity in the observed data set. To detect outliers MSD technique was applied. One outlier has been found using decimal logarithm as a normalizing transformation and F -distribution for $\alpha=0.005$, and it has been removed.

Conclusions. Thus, the preliminary data processing has been done, and we can proceed to the next step is to building multivariate nonlinear regression model. In the following research other normalizing transformation will be applied such as Johnson's and Box-Cox transformations.

REFERENCES

[1] Hee Beng Kuan Tan, Yuan Zhao, Hongyu Zhang. Estimating LOC for information systems from their conceptual data models, in Proceedings of the 28th International Conference on Software Engineering (ICSE '06), May 20-28, 2006, Shanghai, China, pp. 321-330.

[2] Chatterjee S., Price B. Regression Analysis by Example. New York: John Wiley & Son, 1977. 228 p.

[3] S. Prykhodko, N. Prykhodko, L. Makarova, A. Pukhalevych. Application of the squared Mahalanobis distance for detecting outliers in multivariate non-Gaussian data, in Proceedings of 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv, 2018, pp. 962-965.