Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 4 (2015), pp. 2131-2135

http://www.ripublication.com

Multiple Linear Regression for Estimation of Marine Fish Landing in Mueang Chon Buri District

J. Mekparyup[1], K. Saithanu[2*] and P. Sombat[3]

[1,2,3]Department of Mathematics, Faculty of Science, Burapha University
169 Muang, Chonburi, Thailand
[1]jatupat@buu.ac.th, [2]corresponding author: ksaithan@buu.ac.th,*
[3]52030789fon@gmail.com

Abstract

The purpose of this study was to estimate monthly marine fish landing in Mueang Chon Buri District, Chonburi, Thailand. Using multiple linear regression analysis to build the regression equation consists of one dependent variable, marine fish landing (Y), and seven independent variables, average sea level pressure (X_1), rainfall (X_2), average relative humidity (X_3), average air temperature (X_4), wind speed (X_5), average wind direction (inward: X_6) and average wave direction (outward: X_7). The results of this study found that the multiple regression equation for estimation of monthly marine fish landing in Mueang Chon Buri District was $\hat{Y}' = 0.0589 + 0.000001X'_4 - 2.86X'_7$ with standard error of estimation (S) 79.488 and adjust coefficient of determination (R^2_{adj}) 34.1.

AMS Subject Classification: 62J05

Key Words: multicollinearity, variance inflation factor, marine fish landing

INTRODUCITON

Current fishing is extremely important because it is the world's major food source. Moreover, fishing operations and processes also help to create enormous revenue for the country and the country people in each year. In addition, fishing also affects the country's economy which is major exports in the form of fresh and processed seafood. The volume and value of seafood tend to increase because the needs of consumers both inside and outside the country. In contrast, the volume of marine animals reduces sharply because the impact of changes in the marine ecosystem caused by both

changing environment, including the capture of marine animals over demand, and pollution made by human. Mueang Chonburi, a district in Chonburi province, is an area adjacent to the Gulf of Thailand in the East and also locates of the port which is important for the country. The present study analyzes the amount of marine fish landing, saltwater fishing, in Mueang Chonburi District using multiple linear regression (MLR) analysis to determine the relationship between climate factors, sea level pressure, rainfall, relative humidity, air temperature and wind speed, and sea condition factors, wind direction and wave direction, which effect to amount of marine fish landing [1][2][3][4].

MATERIAL AND METHODS

The amount of marine fish landing (Y) was collected from the Department of Fisheries and five climate factors, the average sea level pressure (X_1), rainfall (X_2), the average relative humidity (X_3), average air temperature (X_4) and wind speed (X_5), were collected from the Thai Meteorological Department and two sea condition factors, average wind direction (inward: X_6) and average wave direction (outward: X_7), were collected from the Hydrographic Department.

1. DETERMINING THE SIMPLE CORRELATION COEFFICIENTS

Simple correlation coefficients (R) were firstly calculated to identify relationship among amount of marine fish landing, climate factors and sea condition factors

2. BUILDING THE MLR EQUATION

The MLR equation to estimate amount of marine fish landing was generated by MLR model as of Equation 1 and the best subset method was used to choose the equation [5].

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \varepsilon \qquad (1)$$

3. CHECKING ASSUMPTIONS FOR MLR ANALYSIS

After obtaining the appropriate MLR equation, checking all assumptions of MLR analysis was proceeded. There are four assumptions to be tested; (I) normality of the error distribution using Anderson-Darling statistic (AD) [6]; (II) independence of the errors using Durbin-Watson statistic (DW) [7]; (III) homoscedasticity of the errors using Breusch-Pagan statistic (BP) [8]; (IV) multicollinearity among predictor variables using Variance Inflation Factor (VIF) [9].

4. COMPARISON BETWEEN THE REAL AND ESTIMATED VALUES

After testing all assumptions, the comparison between the real values (RV) and the estimated values (EV) of the amount of marine fish landing from the obtained MLR equation was plotted.

RESULTS

The correlation coefficient values (R) for the eight variables, Y, X_1, X_2, X_3, X_4, X_5, X_6 and X_7, were calculated. The results found that R ranged from –0.470 to 0.524 which was the same previous studies [1][2][3][4] and the highly positive significant correlation was between Y and X_6 (R=0.524, p-value=0.000). Then possible MLR equations were generated by best subset method to select the best fitted equation. It was shown that X_1, X_4 and X_6 were selected to build the MLR equation with the Mallow C-p = 3.1, S = 76.191 and R^2_{adj}=35.20. After obtaining the equation, checking all assumptions of MLR analysis was determined. However, the test was not satisfied. The Box-Cox method was then used to transform the data [10]. After transforming data, the MLR equation was regenerated and the fitted equation was $Y' = 0.0589 + 0.000001X'_4 - 2.86X'_7$ where $Y' = 1/\ln Y$, $X'_4 = X_4^2$ and $X'_7 = 1/X_7$ with S= 79.488 and R^2_{adj}=34.10. Checking all assumptions of the analysis was then retested; (I) The test of normality: AD was calculated and the value was 0.385 (p-value=0.448) so the distribution of error term was normal, (II) The test of independence: the results showed that the DW=1.672 was more than critical values (DL=1.662) so the errors were independent, (III) The test of homoscedasticity: BP was 5.780 which less than the critical values (5.991) so the variance of errors was constant, (IV) Test of multicollinearity: all VIF values were 1 which less than 5 then there was no relationship among independent variables in MLR equation [11]. The amount of marine fish landing was then estimated by using this equation and the comparison between the RV and the EV was displayed by graph as of Figure 1.

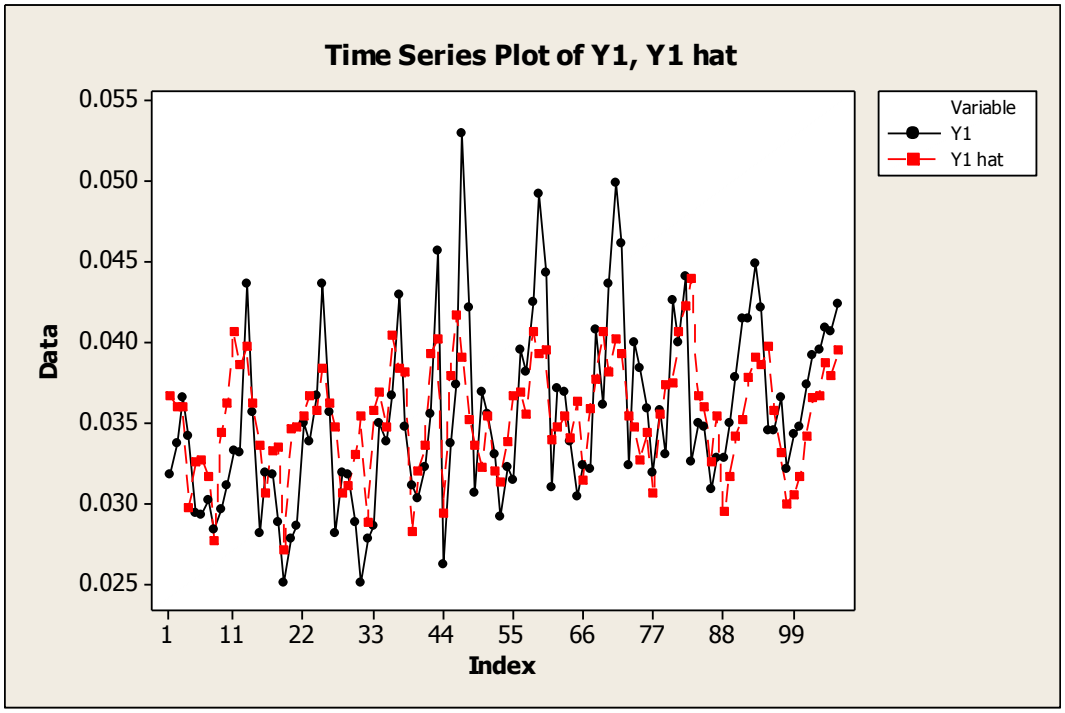

Figure 1: The comparison between the RV and EV of the amount of marine fish landing.

DISCUSSION

The independent variables used to estimate the amount of marine fish landing in Mueang Chonburi District, Chonburi province, Thailand, were average air temperature (X_4) and average wave direction (outward: X_7) with the standard error of the estimation (S) 79.488 and adjusted coefficient of determination (R^2_{adj}) 34.10. The accuracy of estimation was shown by comparing the graph between the real and estimated values from the MLR equation.

ACKNOWLEDGEMENTS

We are grateful to the Department of Fisheries, the Thai Meteorological Department and the Hydrographic Department, Thailand for kindly providing all data.

REFERENCES

[1] Meynecke, J. O., Lee, S. Y., Norman, C. D., & Warnken, J., 2006, "Effect of rainfall as a component of climate change on estuarine fish production in Queensland, Australia," Estuarine, Coastal and Shelf Science, 69, 491-504.

[2] Teixeira, C. M., & Cabral, H. N., 2009, "Time series analysis of flatfish landings in the Portuguese coast," Fisheries Research, 96, 252–258.
[3] Ligas, A., Ranieri, D. S., Micheli, D., Reale, B., Sartor, P., Sbrana, M., & Belcari, P., 2010, "Analysis of the landings and trawl survey time series from the Tyrrhenian Sea (NW Mediterranean)," Fisheries Research, 105, 46–56.
[4] Martina, I. M., & Ahmad, S., 2010, "Stepwise Multiple Regression Method to Forecast Fish Landing," Procedia Social and Behavioral Sciences, 8, 549–554.
[5] Hocking, R. R., & Leslie, R. N., 1967, "Selection of the best subset in regression analysis," Technometrics, 9(4), 531-540.
[6] Lewis, P. A. W., 1961, "Distribution of the Anderson-Darling Statistic," The Annals of Mathematical Statistics, 32(4), 1118-1124.
[7] Durbin, J., & Watson, G. S., 1951, "Testing for Serial Correlation in Least Squares Regression. II," Biometrika, 38(2), 159-177.
[8] Breusch, T. S., & Pagan, A. R., 1979, "A Simple Test for heteroscedasticity and Random Coefficient Variation," Econometrica, 47(5), 1287-1294.
[9] O'brien, R. M., 2007, "A caution regarding rules of thumb for variance inflation factors," Quality & Quantity, 41(5), 673-690.
[10] Sakia, R. M., 1992, "The Box-Cox transformation technique: a review," The statistician, 169-178.
[11] M. H. Kutner, J. N. Christopher, & J. Neter, 1996, "Applied liner regression models, 4th ed," McGraw-Hill/Irwin, USA.