Use of statistical model for zooplankton density in sea grass bed at Pulau Tinggi, Malaysia

Nasir Shuaib*, Mohd Saifullah Rusiman†, Maryati Mohamed†, Hazel Monica Metias-Peralta†, Saba Rizwan‡, Muhammad Mumtaz Khan§

Abstract

Zooplankton density model was developed using multiple linear regression. The Sea grass bed in Pulau Tinggi, Malaysia was selected as sampling site. Sampling of water was carried out for six months (April, June, October, 2015 and April, June, August, 2016). The statistical model was developed for physico-chemical parameters as well as environmental parameters with the density of zooplankton. Ten stations were fixed for water sampling. The results of statistical analysis showed that large sized density model (> 2100 µm) was the best model due to its lower value of mean square error (MSE) for physico-chemical parameters among three size fractions. The student's t-test concluded that there was no significant difference (p > 0.05) between mean density of zooplankton in dry and wet season. The mean zooplankton density was 2442 ind/m^3 and 2683 ind/m^3 in dry and wet season respectively in the present research period.

Key words: zooplankton density, physico-chemical parameters, sea grass bed, Pearson correlation, multiple linear regression.

Introduction

Statistical analysis is a common method whose usage is in common practice. The application of statistics is common in many fields of science like medicine, biology, geology, economics, agriculture, sociology and others. Statistical models can be used for the estimation and prediction of zooplankton. Regression is one of the standard tools in analyzing data. It is used in different disciplines. The relationship between the dependent and independent variables can be calculated through a mathematical equation which widely available in computer packages and easy to interpret. The multiple linear regression (MLR) modeling is a popular model in analyzing the linear model (Wahab *et al*., 2018). Statistical analysis was performed using SPSS software (IBM SPSS Statistics Version 20). All statistical tests were considered significant at $p < 0.05$.

* Faculty of Applied Sciences and Technology (FAST) Universiti Tun Hussein Onn Malaysia, Kampus Pagoh, Johor, Malaysia. nasirshuaib123@gmail.com

† College of Fisheries, Central Luzon State University, Nueva Ecija, Philippines

‡ Department of Botany, Jinnah University for Women, Nazimabad Karachi-74600, Pakistan

§ Department of Microbiology, University of Haripur, Haripur, Pakistan

Prior to analyses, all variables were tested for normality, homogeneity of variances and multi-collinearity. Zooplankton plays a critical role in the marine environment, models need to catch satisfactory the dynamics of zooplankton. Use observations of abundance and density for the determination of zooplankton models (Everett and Baird, 2017). Multiple linear regression models were developed for density of zooplankton during the study duration.

Material and methods

Sampling site

Pulau Tinggi sea grass bed was selected as sampling site. It is situated in the southeast of Mersing, Johor, Malaysia having 16 km2 area. The sampling area was divided in to ten fixed stations at Pulau Tinggi. Bimonthly samples collections were carried out in April, June, October, 2015 and April, June, August, 2016, along Pulau Tinggi Mersing, Johor. Water samples were collected from ten fixed stations in the sea grass area of Pulau Tinggi (Shuaib *et al*., 2019). The water samples were analyzed in the laboratory of Universiti Tun Hussein Onn Malaysia to identify and quantify zooplankton. Two season namely dry season (April to July) and wet season (August to January) were observed during the study period (Marine Meteorological Department, 2016).

Statistical modeling

Data was analyzed by using multiple linear regression for density of zooplankton. Density model was developed to determine the effect of physico-chemical and environmental parameters. The physico-chemical parameters were temperature, salinity, dissolved oxygen, turbidity, conductivity and pH. According to Huo *et al*. (2012) regression models can give a rapid estimate of production and the outcome results based on the models can still provide an understanding of the zooplankton production. Multiple linear regression model was developed for physical and chemical as well as for environmental parameters of small sized (<500 µm), medium sized (501-2000 µm) and large sized (>2100 µm) zooplankton. The density of zooplankton was dependent variables and the rest of them were independent variables. Multiple linear regression model used by ecologists to determine effects of several ecological factors on population and community ecology (Rice and Stewart, 2016).

Multiple regression analysis is often used by ecologists to investigate the impact of various environmental factors on organismal, population, and community ecology (Wahab *et al*., 2018). The Regression equation is given as below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_k X_{ik} + \varepsilon_I \qquad (1)$$

Where, Y_i is dependent variable, $ß_0$, $ß_1$, $ß_{2...}$, $ß_k$ are regression coefficients with respect to observations, X_1, X_2….X_{ik} are predictor variables. ε_i is error which describe the discrepancy between predicted data and observed data, whereas ε_i are independent with $\varepsilon_i \sim N(0, \sigma^2)$ and i= 1…, n. The least square estimators $\hat{\beta}$ that minimize the sum of the squared errors in MLR are represented by equation (2).

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (2)$$

Where, Y: a n×1 vector of observation, X: an $i \times k$ matrix of the levels of the regressor variables and ß: k×1 vector of the regression coefficients. ANOVA was observed if probability for the data is < 0.05. The MSE was also take in consideration to compare the different size fractionated model. The significance of the different variables was studied. Data were obtained from the water samples collected from Pulau Tinggi sea grass bed.

Student's t-test for density in dry and wet season

The *t*-test was applied to density data having in wet and dry season. Where H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$. The decision rule is rejecting H_0 when significant value is < 0.05. H_0 is null hypothesis (Jorge *et al.*, 2016 and Mc Donald and John, 2008).

Results

The zooplankton samples analysis was made on their size during the study duration from the collected samples at Pulau Tinggi, Malaysia. The smallest size fraction is for individuals with measurement of <500µm. The largest is for those with >2100µm. In between are the medium size with the measurement between 501 to 2000µm.

The normality and multi-collinearity test were analyzed for large size zooplankton density. The residual plot has shown in Figure 1. There was high multi-collinearity among the independent variables. So, rainfall was deleted. The MSE for density of large size (> 2100µm) zooplankton model was 4572.778 with $r = 0.516$ and $R^2 = 0.267$. The equation is as below as in equation (3).

$$Y = -566.830 + 8.411 \text{ conductivity} + 29.663 \text{ turbidity} - 15.431 \text{ salinity} - 3.178 \text{ water current} \quad (3)$$

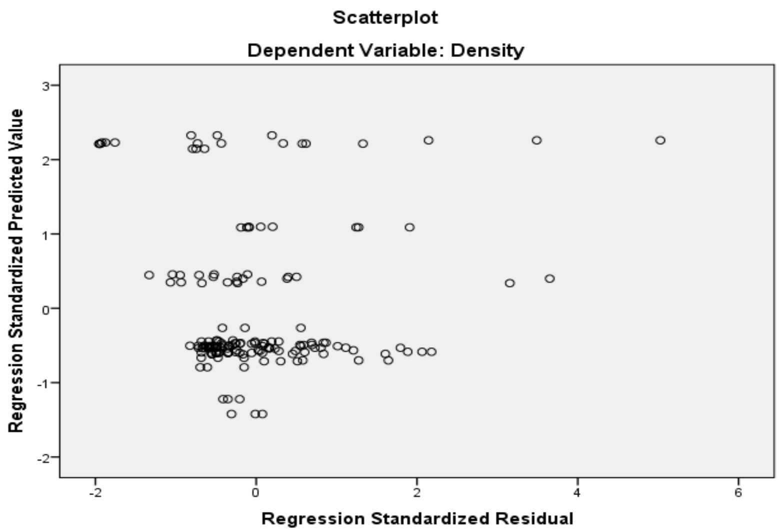

Figure 1. The residual plot showing the distribution of data for large size zooplankton density model

Statistical analysis revealed that large size zooplankton density increases with higher conductivity and turbidity. Moreover, the density of zooplankton is inversely proportional to salinity and water current. The estimated parameters have shown in Table 1.

Table 1: Parameters estimated for density of large size zooplankton

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-566.830	147.132		-3.853	.000
Conductivity	8.411	1.960	.308	4.292	.000
Turbidity	29.663	6.192	.544	4.791	.000
Salinity	-15.431	4.707	-.335	-3.278	.001
Water current	-3.178	1.071	-.263	-2.968	.003

The student's *t*-test indicated that the zooplankton mean density for the wet season is almost similar to the zooplankton density mean in the dry season. The value of mean zooplankton density for the wet season and dry season is 2682.6 individual/m^3 and 2442.3 individual /m^3 respectively.

Discussion

Modeling is a beneficial tool to compare abundance, distributions, fluctuations, and production of living organisms to variations in the ecological environment, food conditions, and predation (Harris *et al.*, 2000). Regression model use gives rapid estimates of production and the results obtained based on the models can still provide an understanding of the zooplankton production (Nakajima *et al.*, 2014). Statistical models have been developed for density of zooplankton in large (> 2100 µm), medium (501-2000 µm) and small (< 500 µm) size fractions. A total of 12 statistical models were developed for density of zooplankton in three size fractions. Among them the density for large size zooplankton have described here in this paper. The large size (>2100 µm) density model is the best model when analyzed statistically. The large size density model having the lowest value of mean square error (MSE).

The results for large size zooplankton density model with the physico-chemical parameters concluded that the conductivity and turbidity were positively correlated to density of zooplankton. Besides, zooplankton density was inversely proportional to salinity and water flow. The highest recorded density of zooplankton during the present study duration is lower than the density reported in Tropical mangroves estuary in the Strait of Malacca (Balqis, 2016). While the density in the present study is higher than the density in the Straits of Malacca during four oceanographic cruises (Rezai *et al*., 2004). The density of zooplankton in the present study was also higher than the density measured in the East coast of Peninsular Malaysia in Terengganu, Pahang and Johor areas (Zaleha, 2006).

Conclusion

The large size (> 2100 µm) density model is the best model when treated with the physico-chemical parameters during the present research duration. Statistical analysis revealed that large size zooplankton density increases with higher conductivity and turbidity. Moreover, the density of zooplankton is inversely proportional to salinity and water flow. The student's *t*-test indicated that the zooplankton mean density for the wet season is similar to the zooplankton density mean in the dry season. Research in sea grass area in the future can only be focus on density model.

Acknowledgement

The authors would wish to offer their appreciation to Research Management Centre (RMC), University of Tun Hussein Onn Malaysia providing fund for research.

References

Wahab, N.S., Rusiman, M.S., Mohamad, M., Azmi, N.A., Che Him, N., Kamardan, M.G., Ali, M. (2018). "A Technique of Fuzzy C-Mean in

Multiple Linear Regression Model toward Paddy Yield''. *Journal of Physics: Conference Series*, 995 (1), Art. No. 012010.

Everett, J.D and Baird, M.E. (2017). "Modeling What We Sample and Sampling What We Model: Challenges for Zooplankton Model Assessment''. *Frontiers in Marine Science.* pp. 1-19.

Shuaib, N., Mohamad, M., Monica, M-P. H., Rusiman, M. S., Babaji, S. S. (2019). "Copepods status in seagrass area of Pulau Tinggi Marine Park, Johor, Malaysia''. *Proc. International conference on Biodiversity. IOP Conference Series: Earth and Environmental Science*, 269 (1), Art. No. 012044

Marine Meteorological Department. (2016). Jabatan Meteorlogi, Malaysia.

Hue , Y., Sun , S., Zhang, F., Wang, M., Li, C., & Yang, B. (2012). "Biomass and estimated production properties of size-fractionated zooplankton in the Yellow Sea, China''. *Journal of Marine Systems*, pp 1-8.

Rice, E., & Stewart, G. (2016)."Decadal changes in zooplankton abundance and phenology of Long Island Sound reflect interacting changes in temperature and community composition''. *Marine Environmental Research*, 154-165.

Jorge L. , P., Gilmar , P.-N., & Marcos G. , N. (2016). "Zooplankton community and tributary effects in free-flowing section downstream a large tropical reservoir''. *International Review of Hydrobiology*, pp. 48-56.

Mc Donald and John , H. (2008). *Handbook of Biological statistics.* Sparky House Publishing Baltimore, Maryland, 287.

Harris, R., Wiebe, P., Lenz, J., Skjoldal, H. R., & Huntley, M. (2000). *zooplankton Methodology Manual.* 684 pp.

Nakajima, R., Yoshida, T., Ross Othman, B., & Toda, T.(2014) "Biomass and estimated production rates of metazoan zooplankton community in a tropical coral reef of Malaysia''. *Marine Ecology*, pp. 112-131.

Balqis, A. (2016). "Seasonal variations of zooplankton biomass and size-fractionated abundance in relation to environmental changes in a tropical mangroves estuary in the Straits of Malacca''. *Journal of Environmental Biology*, pp. 685-695.

Rezai, H., Yusoff, F. M., Arshad, A., Kawamura, A., Nishida, S. and Othman, B. H. R. (2004). "Spatial and temporal distribution of copepods in the Straits of Malacca''. *Zool. Stud.* 43: pp. 486–497.

Zaleha, K., Roswati, M. N. and Iwasaki, N. (2006)"Distribution of some species of Harpacticoid Copepods in East Coast of Peninsular Malaysia''. *Coastal Mar. Sci.* 30: pp. 141–145.