MAPPING OF SALINITY LEVEL USING SPLINE INTERPOLATION TECHNIQUES OVER THE WATER OF SUNGAI MERBOK, KEDAH

Sharir Aizat Kamaruddin[1], Khairul Naim Abd Aziz[2], Muhammad Akmal Roslani[3], Jamil Tajam[4], Shaıdatul Najwa Zaınolabdın [5], Nur Fatin Aqilah Mohd Razman[6]

[1,2,3,4]Marine Research and Excellence Centre, Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Perlis, Kampus Arau, 02600 Arau, Perlis

[5,6]Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Perlis, Kampus Arau, 02600 Arau, Perlis

E-mail: shariraizat@perlis.uitm.edu.my

Received:18 July 2018
Accepted:22 October 2018
Published: 31 December 2018

ABSTRACT

Interpolation map of salinity is a helpful scientific instrument for environmental monitoring and for social economic development especially for the community who lived in Sungai Merbok. This research was conducted to develop a spatial model of salinity using spline interpolation technique. 20 sampling stations were randomly set up to measure the level of salinity using YSI 650 Multiparameter Display System (MDS). Quantitative analysis of standard regression and error index were used to evaluate the developed model. The research found that the tension splines type performed better than regularized splines type. The local government and community, who live in Sungai Merbok, can use the developed map of salinity for guidelines and future development of Sungai Merbok, Kedah.

Keywords: *Spline, Salinity, Sungai Merbok, GIS*

INTRODUCTION

Estuary and coastal areas have complex and vibrant aquatic environment. Numerous saltwater and brackish animals and plants species lived in the area, thriving and supporting the complex marine ecosystem. Mudskippers, for an instant, are an important biological indicator for coastal and estuarine areas as they are sensitive to the ambient environment and are significant biomonitor for the health of coastal and mangrove area. Understanding the importance of the relationship between the ecological role of these areas and the density of animals that lived in the area would open the opportunity for local and government bodies to construct strategic plans and frameworks for the purpose of sustainability and conservation in order to maintain the areas.

Marine vegetations that include marine grasses, seaweeds, mangroves, and algae are the important medium in maintaining the oxygen and carbon dioxide level on earth. From root to the canopy, mangrove area serve the utmost important role as a natural habitat for many marine fishes, birds, reptiles, insects, and mammals. Furthermore, with an increasing number of mangrove plants, the coastal erosion can be reduced over the coastal and riverbank zones. Preservation and protection of marine vegetations should be encouraged as coastal and estuary regions are highly be exposed to massive anthropogenic activities. Besides, some mangrove species require much time to grow and to achieve maturity phase. This could hamper the process of conservating.

Estuary and coastal areas are also well known for tourist attractions whether from local or abroad tourists. Among the local society, these marine locations are important for agriculture and eco-tourism activities. Shrimp, oyster and fish farming flourish along the riverbanks in many Malaysian rivers. Eco-tourism activities such as boating, sports fishing, and floating hotels make up large livelihood for local residents. However, these activities can also give negative impacts towards the ecosystem and the resident who live here, if these activities were promoting and encouraging pollution that could be in the form of noise, air, and water pollution.

Among these types of pollutions, water pollution contributes the most considerable effect towards the marine flora and fauna. Deteriorating water quality would modify or affect other water quality parameters. Vegetation species might be reduced and marine animals might have to migrate to other locations. Thus, creating an unbalanced ecosystem. Moreover, mariculture and aquaculture activities may face disappointment growth and would negatively influence the socio-economical progress among local communities.

One of the important water parameters used to investigate the health of coastal and mangrove area is salinity. As an ecological control, salinity plays an important role in influencing the types of marine animals and plants that live in a body of water (Atwood et al, 2012). The animal and plant species lived here need to withstand the tidal activities. During the low tide, the freshwater would dominate along the river. In contrast, during the high tide phenomena, seawater would flow into the river and change the salinity level. Conventionally, measuring of salinity level is done using portable water quality probe and analysis, which is carried out by determining the lowest and highest level of salinity for every sampling stations and later comparing these results with the water quality index. However, this assessment fails to connect the salinity level with the geographical feature of the sampling stations. One way to observe the pattern of salinity in the coastal and estuarine water is to produce a map of salinity level in the area of interest.

Mapping provides an alternative way for effective communication geospatial data whether the emphasis is on the location or the distribution pattern of data. Water quality map can be developed by using many types of geospatial software, for example, ArcGIS software. ArcGIS is a software used for spatial analysis and developing spatial models. In order to develop spatial models and map the salinity level, the spatial interpolation technique is applied.

Spatial interpolation techniques have become the popular method in the GIS field in investigating and analyzing the pattern of any attribute data along with its geospatial information. For instance, there are two classes of spatial interpolation methods namely deterministic and stochastic interpolation methods. The stochastic method has been widely used to interpolate many attribute data. However, the method requires much skill and effort to run the analysis. Less research has been found on using spline Interpolation technique to map water pollutant such as salinity.

The aim of this research is to map the surface water salinity level over Sungai Merbok, Kedah using the spline interpolation method. The specific objectives are (1) to develop spatial models of surface salinity using spline interpolation techniques and (2) to determine the accuracy of the developed map using statistical analysis. The scope of this research is to investigate the applicability of spline interpolation technique in predicting spatial pattern of salinity over Sungai Merbok, Kedah. Two methods of spline interpolation were used namely regularized spline and tension spline. Sungai Merbok, Kedah was chosen as the study area in this research because the area contains a complex natural environment and anthropogenic activities. Conservation in this area is critical for the sustainability of natural resources and future development of this area.

LITERATURE REVIEW

Salinity is the quantity of dissolved salt content found in waters. Even though the sea contains many different kinds of salt, but the most plentiful is sodium chloride since it is the most dominant ions in seawaters. Thus, salinity can be classified as an important water quality parameter that influences and engineers the characteristics of the deep ocean, coastal waters, estuary, and freshwater.

It is important to observe and investigate the effects of salinity in the prospect of time and location as they would provide valuable information regarding the distribution pattern of salt in water. Salinity varies temporary and spatially according to freshwater and saltwater input (Mitsch & Gosselink, 2000). Freshwater contains less salt in comparison to the ocean. On the other hand, for brackish water, the level of salinity is found to be less salty than the ocean since brackish water contains the mixture between the saltwater and the freshwater. Thus, estuarine water which is dominated by brackish water forming the area suitably but complex for mangrove animals and plants. Moreover, the concentrations of salt are found to increase during the high tide and decrease during the low tide over the estuary area. Tidal activities are also affecting the salinity level in the coastal area.

Coastal water and estuary are well known to home for many marine animals and plants. The greatest threat when the salinity changes drastically to marine faunal and vegetation is lost of habitat (Gehrke et al, 2011). Mangrove area served as a valuable nursery location for crustaceans, shrimps, fishes, and mollusk. The juveniles will be concentrated under the roots or shoots of the mangrove trees. Reducing the number of mangrove tree will have a direct impact on the diversity and density of these animals.

Next, in terms of agriculture and economic perspective, fish farmers may be experiencing difficulty in culturing the fish in the river and estuary area. The increasing of salt in water, especially in the river, has a negative impact on fish diversity (Gain et al, 2008). These effects are seriously observed in animals and plants that live in the upstream where the concentration of salt is low. Freshwater fish that are unable to tolerate with the high concentration of salt ions, would have to migrate to another habitat. Consequently, the migration of these freshwater species would be disrupting the food chain and ecosystem nearby. Higher salinities are also found to lower vegetation richness in the estuarine environment (Bantilan-Smith et al, 2009).

Spatial interpolation is an alternative way to map the salinity level spatially and temporally. Spatial interpolation is the procedure of using points with known values to predict values at other points (ESRI, 2017a). By using this technique, the cost of expense and time can be reduced especially during the sampling period. The interpolation is based on method opted by the users based on the objective and mission of the research carried out.

Spatial interpolation methods may be deterministic or geostatistics (Adhikakary & Dash, 2017). The example of a deterministic method is spline interpolation technique that is based on spline mathematical model while for geostatistics is the kriging interpolation. Even though the kriging interpolation technique has been widely used, there is still less information on the development of surface water salinity using spline interpolation technique.

Sungai Merbok (Merbok river) is located at the state of Kedah Darul Aman. The length of the river is about 35 km, the width is approximately 2 - 20 km and the depth of the river could reach from 4 to 15 meter (Kaniz Fatema et al, 2014). Thus, thick mangrove in the area sheltered many species of flora and fauna along the river. Species of Rhizophora mucronata, Rhizophora apiculata, Brugueira parvifolia, Brugueiria spp., Avicennia spp., and Sonneretia spp. (Ong et al. 1992) dominated the mangrove area. Along the river, tourism activities flourish with increasing floating hotels, floating restaurants, fish and oyster farming, boating and ecotourism activities.

METHODOLOGY

20 sampling points were established based on human activities and ecological importance around Sungai Merbok on October 17th, 2015 during the high tide. The locations of each station were determined using the Global Positioning Satellite (GPS) device. The in-situ measurement of salinity was recorded using the YSI 650 Multiparameter Display System (MDS). YSI 650 MDS was deployed at every station from the surface until 1 meter depth. 1 minute idle time was waited before reading the data displayed on the screen. A database of salinity with their geo-coordinates was downloaded later into Microsoft Excel. The data were carefully examined to eliminate the error. Then, the data were equally divided for testing and validating sets. The testing set was used to develop the spatial model while the validation set was used to validate the developed spatial model. The method of model development and validation is summarized in Figure 1.

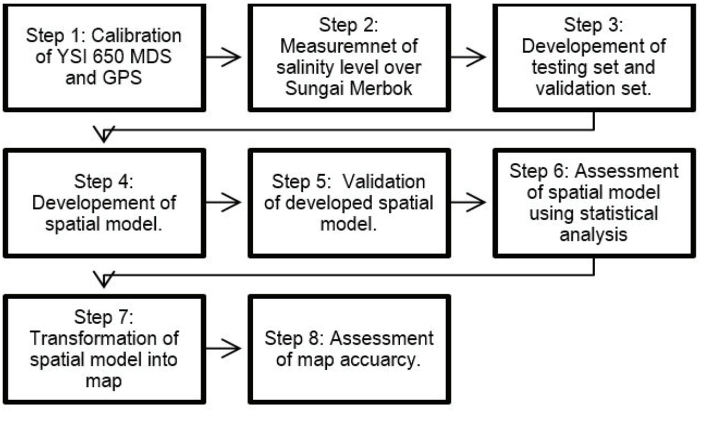

Figure 1 Methodology

Spline interpolation methods

Splines are interpolators that fit a function to sampled points and can be classified as tension splines and regularized splines (Garnero & Godone, 2014). Tension splines type creates a less smooth surface with value more closed constrained by the sample data range compared to regularized spline type. Section A shows the general mathematical formula for spline model. T(x,y) and R(r) are defined differently, depending on the selected option: regularized or tension type. The formula for splines methods was referred to ESRI (2017b).

A. Splines Equation

$$S(x,y) = T(x,y) + \sum_{j=1}^{N} \lambda j \quad R(r_j)$$

Where:-
J =1,2, ……… N
N =Number of points
λj =Coefficients found by the solution of a system of linear equations.
r_j =Distance from the point (x,y) to the jth point.

B. Tension spline method

$$T(x,y) = a1$$

where:-
a1 = coefficients found by the solution of a system of linear equations.
and,

$$R(r) = \frac{1}{2\pi\varphi^2}\left[1n\left(\frac{r\varphi}{2}\right) + c + K_{\circ}(r\varphi)\right]$$

φ^2= Weight parameter
K_°=Modified Bessel function
C=constant equal to 0.0577215

C. Regularized spline method

$T(x,y) = a1 + a2x + a3y$

where:-
a1 = coefficients found by the solution of a system of linear equation.
and,

$$R(r) = \frac{1}{2\pi}\{\frac{r^2}{4}\left[1n\left(\frac{r}{2\pi}\right) + c - 1\right] + \tau^2\left[K_\circ\left(\ \frac{r}{\tau}\ \right) + c + 1n\left(\frac{r}{2\pi}\right)\right]\}$$

where
r = Distance between the point and the sample
τ^2= Weight parameter
K_°=Modified Bessel function

Statistical Analysis

Pearson correlation coefficient (r) is a statistics between -1 and 1 that measures the association between two variables (Zou, Tuncali & Silverman, 2003). In this research, the developed spatial model should yield a positive correlation coefficient.

$$Pearson\ \ correlation\ \ coefficient\ \ (r) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_1 - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

where:-

X¯ and Y¯are the sample means of xi and yi

The coefficient of determination (R2) was used to interpret the proportion of the variance in the dependent variable (predicted salinity value) that was predictable from the independent variable (observed salinity value). In this research, the R2 was written in percentage form. A typical value of 0.5 is considered acceptable (Santhi et al. 2001). Thus, the R2 value in this research is expected to be more than 0.5.

R^2= r^2

Where:-
r = Pearson correlation coefficient

The Mean Absolute Error (MAE) was used to observe the absolute differences between the independent and dependent variable. In this research, MAE was used to measure the absolute difference between predicted salinity value and observed salinity value. Singh et al. (2005) stated that MAE value was less than half of the standard deviation of the observed data contribute to better model performance. Therefore, the research is expects the value of MAE to be less than half of the standard deviation of the observed salinity levels. Moreover, the spatial model which contributes to the lowest MAE value should be chosen.

$$Mean\ \ Absoulte\ \ Error\ \ (MAE) = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j|$$

where:-
yj is the observed value and y^_j is the predicted value

The Root Mean Square Error (RMSE) was used to measure the differences between the values predicted by a model. In this research, RMSE was used to compare the regularized and tension spline model. The best model should yield a low RMSE value. The RMSE value which is less than half of the standard deviation of the observed data gave an improved model performance (Singh et al. 2005).

$$Root\ \ Mean\ \ Square\ \ Error\ \ (RMSE) = \left[\sum_{i=1}^{N}(z_{fi} - z_{0i})^2 /N\right]^{\frac{1}{2}}$$

Where:-
z_fi is the observed value and z_0i is the predicted value.

Map accuracy assessment

In order to check the precision of the developed map, statistical analysis of overall accuracy (OA) was used. In this research, the range of salinity was developed first. Then, the number of correct occurrences was counted if the predicted value of salinity falls into the range of salinity in the observed value.

$$Overall\ Accuracy\ (OA) = \frac{Number\ of\ Correct\ Occurrences}{Total\ Occurrences} \times 100\%$$

RESULTS AND DISCUSSIONS

Spatial models were successfully developed using regularized and tension spline. Figure 2 shows the spatial models developed using regularized and tension spline methods. The spatial pattern can be observed on both spatial models where the downstream of Sungai Merbok was subject to high salinity contributed from the ocean, while the middle-stream and upstream were dominated by brackish water. Brackish water is the suitable environment for mangrove vegetation to grow and subsequently it becomes the nursery for marine juvenile and habitat for fish and prawn (Sasekumar et al, 1992). Wafar (2009) found that there is a linear relationship exists between shrimp production and the size of the mangrove forest.

Then the predicted values from the validation set were compared with the observed values. For the regularized spline model, the research found that the correlation coefficient was determined at 0.955. On the other hand, the correlation coefficient for tension spline was calculated at 0.969. Both models exhibit a positive and strong correlation. Moreover, the research also found that only 91.14% of the variance was explained by the regularized model compared to 93.36% of the variance explained by tension spline model. Thus, the spatial model developed using tension spline has stronger correlation compared to the regularized spline.

The linear equation developed for regularized spline model was $y = 0.8555x+3.469$, while for tension spline model was $y = 0.8646x + 3.315$. Figure 3 shows the graph developed when the observed value was regressed with the predicted value. The standard deviation calculated for the regularized spline model was 7.248, while for tension spline was determined at 7.215. The regularized and tension interpolation type recorded lower value of standard deviation compared to the observed value.

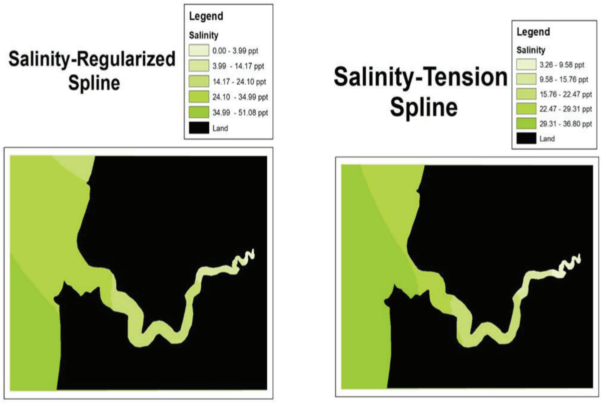

Figure 2 Spatial models of salinity developed using regularized (left) and tension (right) model

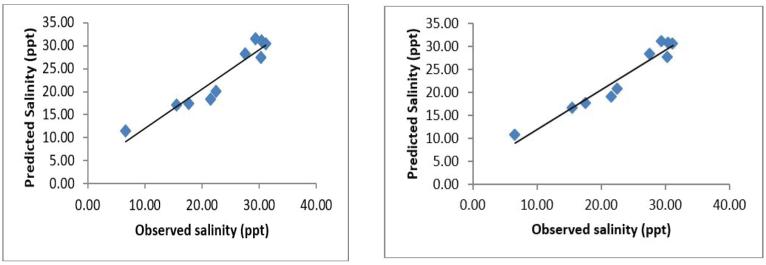

Figure 3 The graph on the top was generated based on the regularized spline type while the graph on the bottom was produced using the tension spline type.

The research found that regularized spline type overestimated the level of salinity at ST1, ST5, ST13, ST15, and ST17 and underestimated the level of salinity at ST3, ST5, ST7, ST9, and ST19. Conversely, for tension spline type, the model overestimated the level of salinity at ST1, ST3, ST5, ST13, ST15, and ST17, while the model underestimated the level of salinity at ST7, ST9, ST11, and ST19. Table 1 shows the list of observed and predicted value generated using regularized and tension spline type. It was worth to point out that the largest differences between the observed and predicted value can be noticed at station 1 (ST1) that was located in the upstream of Sungai Merbok. The observed value of ST1 was considered an outlier. An outlier can have a significant effect on spline interpolator giver that the spline minimizes overall surface curvature and generates gently varying surface (ESRI, 2017b).

Table 1 Observed and Predicted Value Generated by dDeveloped by Both Spatial mModels

Station (ST)	Observed value (ppt)	Predicted value (ppt)	
		Regularized spline	Tension spline
1	6.61	11.38	10.90
3	15.49	17.04	16.67
5	17.61	17.54	17.82
7	21.51	18.40	19.11
9	22.43	19.95	20.78
11	30.24	27.62	27.68
13	27.50	28.35	28.33
15	30.36	30.96	30.85
17	29.34	31.58	31.17
19	31.10	30.48	30.30

The research also found that the MAE and RMSE value for the spatial model developed using regularized interpolation type was determined at 1.891 and 2.331. In contrast, the MAE and RMSE value for the spatial model developed using tension interpolation type was calculated at 1.594 and 1.984 respectively. Since MAE value was lower than the RMSE,there were some variations in the magnitude of errors for both spatial models. The RMSE and MAE were also found to be less than half of the standard deviation of the observed value which indicate better model performance. Consequently, tension spline type was selected to be the best in predicting salinity over Sungai Merbok, Kedah since the spatial model demonstrates the lower value of MAE and RMSE compared to regularized spline type.

Next, the spatial model using tension spline type was transformed into a map. Legend, map title, scale and north arrow were inserted into the map for better communication and visualization. Only three data classifications were chosen for the map (<15 ppt, 15-30 ppt and >30 ppt). Figure 4 shows the completed map of salinity over Sungai Merbok, Kedah. The map showed that there was a significant pattern of an increasing level of salinity from upstream to downstream. The upstream of Sungai Merbok is dominated by brackish waters while the downstream was dominated by saltwater.

The accuracy of the developed map was calculated at 80% of accuracy. Two stations which were Station 11 (ST11) and Station 17 (ST17) did not achieve correct occurrence. The reason for incorrect occurrence for these stations was due to overestimating and underestimating of predicted salinity by tension spline type. Table 2 shows the overall accuracy (OA) calculated from the developed map.

Table 2 The Overall Accuracy of the Developed Map of Salinity over Sungai Merbok, Kedah

Observed value	Station	Predicted value
<15 ppt	1	<15 ppt
15 ppt - 30 ppt	3	15 ppt - 30 ppt
15 ppt - 30 ppt	5	15 ppt - 30 ppt
15 ppt - 30 ppt	7	15 ppt - 30 ppt
15 ppt - 30 ppt	9	15 ppt - 30 ppt
>30 ppt	11	15 ppt - 30 ppt
15 ppt - 30 ppt	13	15 ppt - 30 ppt
>30 ppt	15	>30 ppt
15 ppt - 30 ppt	17	>30 ppt
>30 ppt	19	>30 ppt

CONCLUSION

The aim of this research has been achieved. A spatial model of surface water salinity has been successfully developed using spline interpolation technique. The research found that tension spline type performed better than the regularized spline type. The developed tension spline model achieved a positive correlation, a high coefficient of determination, low MAE and RMSE value. The developed spatial model has been successfully transformed into a map. Figure 4 shows the map of salinity over Sungai Merbok, Kedah. Overall,the developed map demonstrated high accuracy. The local government, private sector, and community, who live in Sungai Merbok, can use the developed map of salinity for guidelines and future development of Sungai Merbok, Kedah.

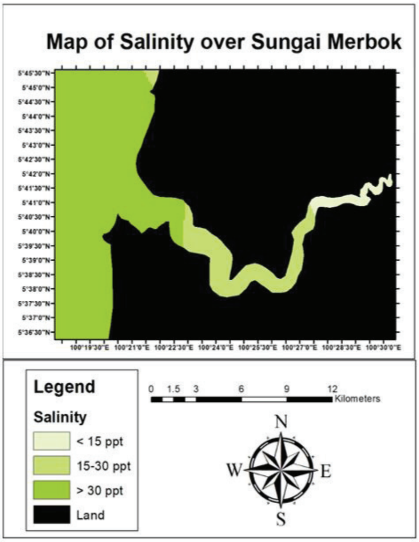

Figure 4 Map of Salinity over Sungai Merbok, Kedah

ACKNOWLEDGMENT

The authors gratefully acknowledge the generous assistance and support from the academic and non-academic staff for their contributions in this research especially to Marine Research Center (MAREC), Universiti Teknologi MARA Cawangan Perlis.

REFERENCES

Adhikary, P. P., & Dash, C. J. (2014). Comparison of Deterministic and Stochastic Methods to Predict Spatial Variation of Groundwater Depth. *Applied Water Science, 7*(1), 339-348. doi:10.1007/s13201-014-0249-8

Atwood, T.B., Wiegner, T.N., MacKenzie, R.A. (2012). Effects of Hydrological Forcing & on the Sructure of a Tropical Estuarine Food Web. *Oikos.121*(2), 277–289. DOI: 10.1111/j.1600-0706.2011.19132.x

Bantilan-Smith, M., Bruland, G.L., MacKenzie, R.A., Henry, A.R. & Ryder, C.R. (2009). A Comparison of the Vegetation and Soils of Natural, Restored, and Created Coastal Lowland Wetlands in Hawai'i. *Wetlands*. 29, 1023–1035. DOI: 10.1672/08-127.1

ESRI (2017a). *An overview of the Interpolation toolset.* Retrieved from http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/an-overview-of-the-interpolation-tools.htm.

ESRI (2017b). *How Spline Works*. Retrieved from http://pro.arcgis.com/en/pro-app/tool-reference/3d-analyst/how-spline-works.htm.

Gain, A.K., Uddin, M.N., & Sana, P (2008). Impact of River Salinity on Fish Diversity in the South-West Coastal Region of Bangladesh. *International Journal of Ecology And Environmental Sciences, 34*(1), 49-54, 2008.

Garnero, G., & Godone, D. (2014). Comparisons Between different interpolation techniques. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W3, 139-144. doi:10.5194/isprsarchives-xl-5-w3-139-2013

Gehrke, P.C., Sheaves, M. J., Boseto, D., Figa, B.S. & Wani, J. (2011). *Vulnerability of Freshwater and Estuarine Fish Habitats in theTropical Pacific to Climate Change* (pp.369-461). Secretariat of the Pacific Community, Noumea, New Caledonia.

Kaniz Fatema, Wan Maznah, W.O & Mansor Mat Isa (2014). Spatial and Temporal Variation of Physico Chemical Parameters in the Merbok estuary, Kedah, Malaysia. *Tropical Life Sciences Research, 25*(2), 1-19, 2014.

Mitsch, W.J.& Gosselink, J.G.(2000). *Wetlands*, 3rd edition. John Wiley & Sons, Inc, New York, NY.

Ong, J. E., Gong, W. K., Wong, Y. P., & Wong, H. K.. (1992). Identification of Mangrove Vegetation Zones using MicroBRIAN and LANDSAT imagery. *Proceedings of the Third ASEAN Science and Technology*

Conference Vol 6.

Santhi, C, J. G., Arnold, J. R., Williams, W. A., Dugas, R. Srinivasan, & L. M. Hauck. (2001). Validation of the SWAT Model on a Large River Basin with Point and Nonpoint Sources. *J. American Water Resources Assoc. 37*(5): 1169-1188.

Sasekumar, A., Chong, V. C., Leh, M. U., & D'Cruz, R. (1992). *Mangroves as a Habitat for Fish and Prawns.* The ecology of mangrove and related ecosystems, 195-207. doi:10.1007/978-94-017-3288-8_21.

Singh, J., Knapp, H.V., Arnold, J. & Demissie, M. (2005). Hydrologic Modeling of the Iroquois River Watershed using HSPF and SWAT. *Journal of the American Water Resources Association. 41*(2), 343-360. DOI: 10.1111/j.1752-1688.2005.tb03740.x

Wafar, S. (2009). Role of Mangroves in Brackish Water Fish Culture. *Biology Education*, 51-52.

Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and Simple Linear Regression. *Radiology, 227*(3), 617-628. doi:10.1148/radiol.2273011499.