

The Influence of Air Temperature Controls in Estimation of Air Temperature over Homogeneous Terrain

Fariza Yunus¹, Jasmee Jaafar^{2,*}, Zamalia Mahmud², Nursalleh K Chang¹

¹Malaysian Meteorological Department, Jalan Sultan 46667 Petaling Jaya, Selangor, Malaysia

²Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor, Malaysia

*Corresponding author: jasmee@salam.uitm.edu.my

Received September 22, 2014; Revised September 30, 2014; Accepted December 22, 2014

Abstract Variation of air temperature from one place to another is caused by air temperature controls. In general, the most important control of air temperature is elevation. Another significant independent variable in estimating air temperature is the location of meteorological stations. Distance to coastline as well as land use type also contributes to significant variations in the air temperature. On the other hand, in homogeneous terrain direct interpolation of discrete points of air temperature work well to estimate air temperature values in un-sampled areas. In this process the estimation is solely based on discrete points of air temperature. However, this study presents that air temperature controls also play significant roles in estimating air temperature over the homogenous terrain of Peninsular Malaysia. An Inverse Distance Weighting (IDW) interpolation technique was adopted to generate continuous data of air temperature. This study compared two different datasets, observed mean monthly data of T, and estimation error of T-T', where T' is the estimated value from a multiple regression model. The multiple regression model considered eight independent variables. They are elevation, latitude, longitude, coastline, and four land use types consisting of water bodies, forest, agriculture and build up areas to represent the role of air temperature controls. Cross validation analysis was conducted to review the accuracy of the estimated values. Final results show that the estimated values of T-T' produced lower errors for mean monthly mean air temperature over homogeneous terrain in Peninsular Malaysia.

Keywords: air temperature control, interpolation analysis, Peninsular Malaysia, regression model, air temperature

Cite This Article: Fariza Yunus, Jasmee Jaafar, Zamalia Mahmud, and Nursalleh K Chang, "The Influence of Air Temperature Controls in Estimation of Air Temperature over Homogeneous Terrain." *Applied Ecology and Environmental Sciences*, vol. 2, no. 6 (2014): 141-145. doi: 10.12691/aees-2-6-3.

1. Introduction

An air temperature control determines air temperature variation from one region to another region [1]. Elevation is a main control of air temperature in heterogeneous regions. Temperature lapse rate is an inverse correlation between air temperature values and elevation, where air temperature decreases when elevation is increased [1,2]. In most of the previous studies, [3-7], the lapse rate values were approximated from data utilized in the studies. In Peninsular Malaysia for estimating mean air temperature, estimation by Yunus [8] considered only elevation as the independent variable. However in homogeneous terrain, such as climate region II of Peninsular Malaysia [9], the role of elevation is less significant in the variation of air temperature element [10]. Furthermore, it is possible to perform direct interpolation procedure for air temperature in areas with minor topographic variation [11].

Other air temperature controls are location, coastline and build up areas. These factors are significant environmental independent variables to estimate air temperature values [12]. Study by Choi [13] discovered

that errors in interpolation of air temperature can be reduced up to 30% by considering urbanization. Previous studies [14,15,16,17] suggested that there is a vital contribution of land use types in air temperature distribution. Main types of land use are agriculture, forest, build up, and water bodies [14,16,18]. The study conducted by Stahl [5] considered the location of meteorological stations as an independent variable in modeling air temperature over British Columbia, Canada. In general, important controls of air temperature are elevation, station's location, coastline and land use types. However, there were limited studies that combined all the eight significant air temperature controls in estimation of air temperature element. Furthermore, the existing estimation by Yunus [8], for estimating mean air temperature in Peninsular Malaysia only considered elevation as an independent variable.

A multiple regression model is commonly developed to explain the effect of air temperature controls on the air temperature element [5,19,20]. To predict air temperature element, Stahl [5] and Brown & Comrie [22] employed station's coordinate as independent variables in multiple regression. Jarvis & Stuart [19] developed a multiple regression model to examine the responses of elevation

and land use over air temperature. Predictor of station's locations and elevation were also selected to develop a multiple regression model in estimation of air temperature element [20]. The developed multiple regression model was integrated by an interpolation technique that interpolates the residual of $T-T'$, where T represents the observed air temperature and T' is the estimation of air temperature from the multiple regression model [19,20].

Spatial interpolation technique plays a significant role in most environmental studies. The technique is used to produce a continuous surface from discrete data. There are some established techniques of spatial interpolation such as inverse distance weighting (IDW), kriging, spline, and trend surface regression [21]. Although kriging is the best known spatial interpolation technique in earth science [21,22], IDW is frequently used in spatially located climate data [3,4,5,20,23]. Furthermore, earlier studies have shown that the performance of kriging and IDW are roughly equal [5,8,19,24].

Currently, there are limited studies carried out in estimating air temperature at un-sampled sites in Malaysia. Yunus [8] proposed the IDW interpolation technique with constant temperature lapse rate as the appropriate technique in generating mean 10-day mean air temperature. However, the study only used data over a one year period and elevation was the only environmental factor considered in the study. Furthermore, the influence of elevation factor on air temperature variation becomes less significant in homogeneous terrain with minor topographic variation. Hence, the purpose of this study is to understand and evaluate all eight air temperature control of elevation, latitude, longitude, coastline and four types of land use, which can be used to estimate the air temperature over homogeneous terrain of climate region II in Peninsular Malaysia [9]. This process would also include the integration of multiple regression model and interpolation technique to generate continuous surface of air temperature.

2. Study Area and Data Source

The study area covers climate region II in the eastern part of Peninsular Malaysia. Generally climate in the peninsula is hot and humid, with relative humidity of between 70 and 90 percent. Annual air temperature is around 26.5°C [25]. This region experiences tropical climate. It has two annual monsoon seasons namely the southwest (May to September) and northeast (November to February). This area receives up to 2,500 mm of rain annually [25].

This study investigated the meteorological element of mean monthly mean air temperature. These monthly data are provided by the Malaysian Meteorological Department. The data were observed from 17 meteorological stations in the study area of climate region II. Figure 1 shows the distribution of 17 stations in the study area.

Monthly data of mean air temperature were acquired over a 10 year period of January 1999 to December 2008. All stations had sufficient data, where 16 stations or 94.1% of the total stations had at least 80% of completed data [26]. Additionally data of 90-meter Digital Elevation Model and year 2000 land use map were also used in this

study. These data were also obtained from the Malaysian Meteorological Department.

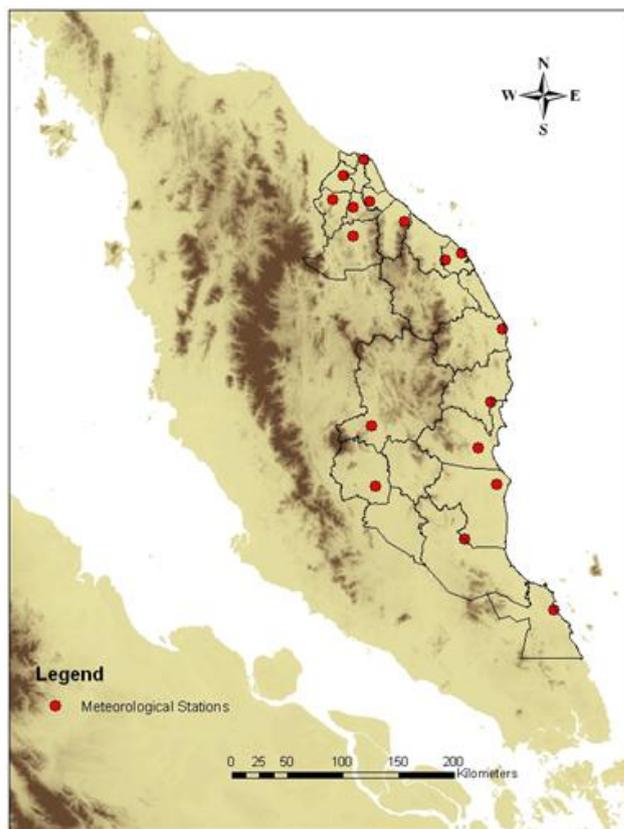


Figure 1. Distributions of meteorological stations in the study area

3. Methodology

In this study, eight independent variables [3,4,5,6,14,16] were evaluated in developing the regression model for the mean monthly mean air temperature element. These independent variables included station's location of latitude and longitude, station's elevation, the nearest distances of station to coastline and the nearest distances of station to each of the four main types of land use, which were water bodies, forest, agriculture and build up area [16]. Independence variables of latitude, longitude and elevation; for each station were obtained together with air temperature element data from MMD. However, the other five independent variables have to be determined from the point's data of station, line data of coastal area and polygon data of land use. Application in Geographical Information System (GIS) of ArcGIS 9.3 to measure the nearest distances was carried out to estimate all these distances.

Data at 17 stations with all eight independent variables then were used to generate a multiple regression model for mean monthly mean air temperature element. In creating this multiple regression model, the hierarchical regression approach was adopted since this method is the most flexible method among regression analysis techniques [27], [28]. The software used in this analysis was SPSS version 16.0. All five regression assumptions of linearity, homoscedasticity, independence term of error, normality of error distributions and multicollinearity [27,28] has to be complied in developing this multiple regression model.

In this study, the residual method of $T - T'$ [19,20] is applied to develop the multiple regression model. By using estimated values of T' , new data fields of $T - T'$ were generated for mean monthly data. This study utilized the Inverse Distance Weighting (IDW) interpolation technique since it is most frequently used in interpolating air temperature element data and is amongst the most accurate interpolation method [3,4,5,20,23]. In this IDW technique, the number of station selected was eight and number of power was two [4,20]. The developed multiple regression model was integrated with IDW technique by interpolating of $T - T'$ data. As a comparison, direct IDW interpolation of mean monthly data of T was also performed. In this analysis 70% of both types of data in every month were involved in this spatial interpolation analysis. To apply this analysis, the ArcGIS 9.3 software was used. This analysis was repeated for all 120 months for both types of main data.

Table 1. Results of Hierarchical Multiple Regression for Mean Monthly Mean Air Temperature

R ²	Significant Independent Variables (alpha = 0.05)	
0.14	Constant	26.989
	Lat	0.086
	Long	-
	Elevation	-0.014
	C_line	-
	W_body	0.015
	Forest	0.062
	Build Up	-0.016
	Agri	-

In every case for both types of data, cross validation analysis was carried out by using the rest of 30% of the data, where five errors values of mean absolute error (MAE), root mean square error (RMSE), systematic root mean square error (RMSEs), unsystematic root mean square error (RMSEu) and d values [29,30] were calculated. Maps of mean monthly mean air temperature were also compared and evaluated. IDW interpolation of T main data was directly used to produce the map of mean monthly air temperature, whereas to generate the map of $T - T'$ main data, $T - T'$ data was added to the surface of T' , where this T' was produced by using the developed multiple regression model, DEM and raster data of nearest distances to coastline and the four types of land use.

Performance of the developed estimation was also compared with the previous estimation developed by Yunus [8]. Based on cross validation analysis, generated map of mean monthly mean air temperature and comparison with the previous estimation [8], the appropriate estimation was recommended to predict mean monthly mean air temperature.

4. Results and Discussion

Distances of each station to the nearest coastal area and distances of each station to each type of land use were estimated. Complete information for each station included all the eight independent variables, together with data of mean monthly mean air temperature were used in developing multiple regression models. In the development of multiple regression model which employed hierarchical multiple regression method, all the five assumptions were complied.

Table 1 illustrates the parameters used in the multiple regression model developed to predict the mean monthly mean air temperature element using the hierarchical multiple regression approach. Results in Table 1 shows that among the eight environmental factors, five variables were significant in developing the multiple regression model of mean monthly mean air temperature. Coefficient of determination, r^2 of this model is relatively low with value of 0.14. This low value of r^2 is due to modeling air temperature over homogeneous areas in which the role of environmental factors are reduced [11]. The r^2 of 0.14 means that by adding five independent variables, 14% of variability in mean monthly mean air temperature can be explained. Referring to Table 1, a complete multiple regression model for mean monthly mean air temperature is as follows:

$$\begin{aligned} & \text{Mean monthly mean air temperature} \\ & = 26.989 - 0.086\text{Latitude} \\ & - 0.014\text{Elevation} + 0.015\text{Water Bodies} \\ & + 0.062\text{Forest} - 0.016\text{Build Up} \end{aligned} \tag{1}$$

Estimated values of T' for mean monthly mean air temperature data were calculated by using the developed multiple regression model. New data fields of $T - T'$ were also developed by utilizing calculated values of T' and observed data for mean monthly mean air temperature of T . Complete data of T and $T - T'$ were utilized in further analysis using the IDW spatial interpolation technique.

Table 2 contains the general results of cross validation analysis using the IDW spatial interpolation technique for both types of data. In Table 2, results of cross validation analysis shows that interpolation of $T - T'$ main data had better performance compared to direct interpolation of T main data, with lower error values and a higher d value of 0.91. The RMSE value for $T - T'$ main data was smaller with smaller RMSEs and RMSEu value approaching the RMSE value. These results indicate that the model of $T - T'$ main data is a 'good' model [29] since both the MAE and RMSE values were small, with RMSEs near to zero and RMSEu is close to the RMSE value which showed that most of the RMSE values were from unsystematic errors. Furthermore, the coefficient of determination, r^2 , of the interpolated $T - T'$ data was 0.67 which is higher compared to the r^2 for the direct interpolation of T .

Table 2. Results of Cross Validation Analysis

Stn. No.	Season	Type of Error	Main Data	
			$T - T'$ ($R^2 = 0.67$)	T ($R^2 = 0.47$)
All	All	MAE	0.4478	0.61537
		RMSE	0.5673	0.73331
		RMSEs	0.2302	0.44052
		RMSEu	0.5184	0.58625
		d	0.9083	0.81337

Figure 2 and Figure 3 show sample maps of mean monthly mean air temperature in May 2000 generated from both analyses.

As in Figure 2, the map generated by interpolating $T - T'$ data produced more reliable estimated values of mean monthly mean air temperature compared to the map generated by direct interpolation of T main data (Figure 3). The map created by interpolating $T - T'$ was producing a smooth visualization and nicely presented topographical patterns of Peninsular Malaysia.

The performance of the model developed in this study was compared with the previous estimation by Yunus [8]. It was found that interpolation of T – T’ main data is a better model with lower error values as compared to the previous model with errors of 1.06°C, 1.29°C, 1.14°C and 0.56°C, for MAE, RMSE, RMSEu and RMSEs respectively, and d value of 0.9. The model developed in this study produced lower errors with significant reductions of 58% for MAE, 56% for RMSE and 80% for RMSEs and showed an increase of 1.1% for d value. Meanwhile the RMSEu showed a slight decrease of about 7%, indicating that the developed model is a better model since most of the RMSE error was related to the unsystematic error of RMSEu, as compared to the previous model, where most of the RMSE was related to the systematic component.

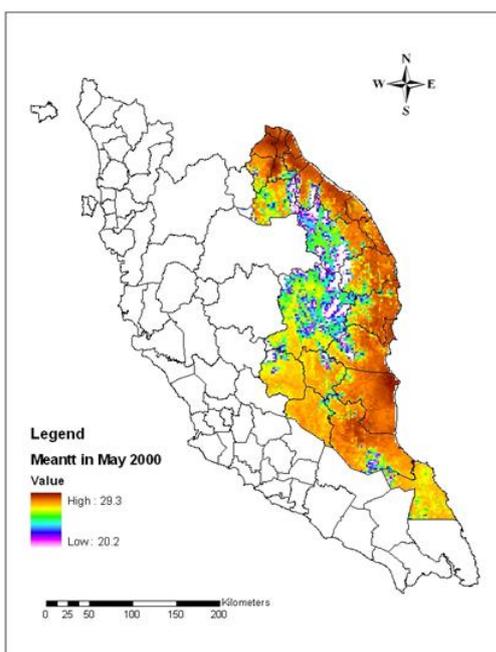


Figure 2. Sample of Generated Map for Interpolation Analysis of T-T’ Main Data

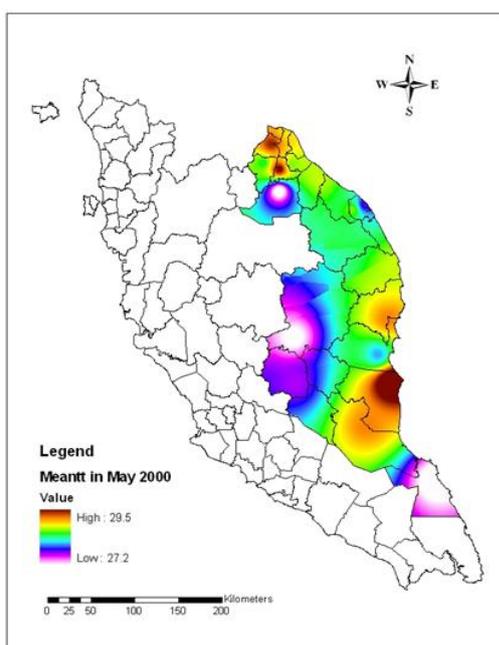


Figure 3. Sample of Generated Map for Interpolation Analysis of T Main Data

The performance of the model developed in this study was compared with the previous estimation by Yunus [8]. It was found that interpolation of T – T’ main data is a better model with lower error values as compared to the previous model with errors of 1.06°C, 1.29°C, 1.14°C and 0.56°C, for MAE, RMSE, RMSEu and RMSEs respectively, and d value of 0.9. The model developed in this study produced lower errors with significant reductions of 58% for MAE, 56% for RMSE and 80% for RMSEs and showed an increase of 1.1% for d value. Meanwhile the RMSEu showed a slight decrease of about 7%, indicating that the developed model is a better model since most of the RMSE error was related to the unsystematic error of RMSEu, as compared to the previous model, where most of the RMSE was related to the systematic component.

In this study, direct interpolation of T was found to have lower errors compared with the previous estimation by Yunus [8], with error reduction of 44%, 56% and 70% for MAE, RMSE and RMSEs. RMSEu value for direct interpolation of T produced a slightly higher error of 3% compared to the previous model. This result shows direct interpolation of T data is a better model because the majority of RMSE error was related to unsystematic error of RMSEu. In descending order, performances of these three models are: interpolation of T – T’ data, direct interpolation of T and previous estimation by Yunus [8].

5. Conclusion

Air temperature controls of elevation, latitude and nearest distances to three main types of land use which are water bodies, forest and build up are significantly vital in estimating mean monthly mean air temperature over homogeneous terrain in Peninsular Malaysia. Cross validation analysis shows, integration model of the developed multiple regression interpolation and IDW interpolation technique, in interpolation of T – T’ main data, is a ‘good’ model to estimate mean monthly mean air temperature. This model produced acceptable error values [11] of 0.45°C, 0.58°C, 0.23°C, 0.52°C for MAE, RMSE, RMSEs and RMSEu, respectively, with high d value of 0.91. Although the generated multiple regression model had low r² of 0.14, the role of the five independent environmental factors are obviously significant with reduction of errors between 7% and 80% compared with direct interpolation. Furthermore, map generated from interpolation of T – T’ main data estimated reliable values of mean monthly mean air temperature values with smooth s that followed topographic pattern of Peninsular Malaysia. The model developed in this study is also a better estimation with lower error values, compared to the previous estimation developed by Yunus [8] which considered only elevation as an independent variable. This study discovers that in estimating air temperature element in homogeneous areas, all air temperature controls have to be evaluated. By considering only elevation as a predictor, the final value of air temperature is not an appropriate estimation. Furthermore, the direct interpolation of T performed better with lower errors compared to the estimation using only the elevation predictor. This study revealed that there are significant contribution of five independent variables of latitude, elevation and the three

nearest distances of station to water bodies, forest and build up areas, in the estimation of air temperature element over homogeneous terrain areas of Peninsular Malaysia.

References

- [1] C. D. Ahrens, *Meteorological Today*, 8th Edition. United States of America: Thompson Higher Education, 2007.
- [2] W. P. Cunningham and M. A. Cunningham, *Principles of environmental science inquiry and applications*. United States of America: McGraw Hill International Edition, 2006.
- [3] A. T. DeGaetano and B. N. Belcher, "Spatial Interpolation of Daily Maximum and Minimum Air Temperature Based on Meteorological Model Analyses and Independent Observations," *Journal of Applied Meteorology and Climatology*, vol. 46, p. 13, 2007.
- [4] R. Dodson and D. Marks, "Daily air temperature interpolated at high spatial resolution over a large mountainous region," *Climate Research*, vol. 8, p. 20, 1997.
- [5] K. Stahl, R. D. Moore, J. A. Floyer, M. G. Asplin, and I. G. McKendry, "Comparison of Approaches for Spatial Interpolation of Daily Air Temperature in a Large Region with Complex Topography and Highly Variable Station Density," *Agricultural and Forest Meteorology*, vol. 139, p. 13, 2006.
- [6] Z. Ustrnul and D. Czekierda, "Application of GIS for the Development of Climatological Air Temperature Maps: an Example from Poland," *Meteorological Applications*, vol. 12, p. 8, 2005.
- [7] P. V. Bolstad, L. Swift, F. Collins, and J. Regniere, "Measured and predicted air temperature at basin to regional scales in the southern Appalachian mountains," *Agricultural and Forest Meteorology*, vol. 91, p. 16, 1998.
- [8] F. Yunus, "Assessment of Spatial Interpolation Techniques of Temperature Elements in Peninsular Malaysia," Master of Science Dissertation, University Putra Malaysia, 2005.
- [9] Fariza Yunus, Aziz Shafie, Jasmee Jaafar, and Zamalia Mahmud, "Homogeneous Climate Divisions for Peninsular Malaysia," *Geodinamica Acta*, vol. 24, p. 7, 2012.
- [10] C. Daly, "Guidelines for Assessing the Suitability of Spatial Climate Data Sets," *International Journal of Climatology*, vol. 26, p. 15, 2006.
- [11] S. P. Serbin and C. J. Kucharik, "Spatiotemporal Mapping of Temperature and Precipitation for the Development of a Multidecadal Climatic Dataset for Wisconsin," *International Journal of Climatology*, vol. 48, p. 15, 2009.
- [12] C. H. Jarvis and N. Stuart, "A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part I: The selection of "guiding" topographic and land cover variables," *Journal of Applied Meteorology*, vol. 40, p. 15, 2001.
- [13] J. Choi, "Urban-effect to improve accuracy of spatially interpolated temperature estimates in Korea," *Journal of Applied Meteorology*, vol. 42, p. 9, 2003.
- [14] K. L. Civerolo, G. Sistla, S. T. Rao, and D. J. Nowak, "The Effects of Land Use in Meteorological Modeling: Implications for Assessment of Future Air Quality Scenarios," *Atmospheric Environment*, vol. 34, p. 7, 16th August, 1999.
- [15] K. P. Gallo, T. W. Owen, and D. R. Easterling, "Temperature trends of the U.S. historical climatology network based on satellite-designated land use/land cover," *Journal of Climate*, vol. 12, p. 5, 1999.
- [16] H. Shudo, J. Sugiyama, N. Yokoo, and T. Oka, "A study on temperature distribution influenced by various land uses," *Energy and Buildings*, vol. 26, p. 7, 1997.
- [17] H. Taha, "Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat," *Energy and Buildings*, vol. 25, p. 5, 1997.
- [18] Othman Jaafar, Sharifah Mastura, and Alias Mohd Sood, "Land use deforestation modelling of river catchments in Klang Valley, Malaysia," *Sains Malaysiana*, vol. 38, p. 10, 2009.
- [19] C. H. Jarvis and N. Stuart, "A comparison among strategies for interpolating maximum and minimum daily air temperature. Part II: The Interaction between Number of Guiding Variables and the type of interpolation method," *Journal of Applied Meteorology*, vol. 40, p. 10, 2001.
- [20] D. Kurtzman and R. Kadmon, "Mapping of temperature variables in Israel: a comparison of different interpolation methods," *Climate Research*, vol. 13, p. 11, 1999.
- [21] D. E. Myers, "Spatial interpolation: an overview," *Geoderma*, vol. 62, p. 12, 1994.
- [22] D. P. Brown and A. C. Comrie, "Spatial modeling of winter temperature and precipitation in Arizona and New Mexico, USA," *Climate Research*, vol. 22, p. 13, 2002.
- [23] D. T. Price, D. W. M. I. A. N. M. F. Hutchinson, and J. L. Kestesen, "A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data," *Agricultural and Forest Meteorology*, vol. 101, p. 14, 2002.
- [24] R. D. Valley, M. T. Drake, and C. S. Anderson, "Evaluation of alternative interpolation techniques for the mapping of remotely-sensed submersed vegetation abundance," *Aquatic Botani*, vol. 81, p. 13, 2005.
- [25] L. J. Tick and Azizan Abu Samah, *Weather and Climate of Malaysia*, 1st Edition ed. Kuala Lumpur: University of Malaya Press, 2004.
- [26] B. R. A. Malmgren and A. Winter, "Climate Zonation in Puerto Rico Based on Principal Component Analysis and an Artificial Neural Network," *Journal of Climate*, vol. 12, p. 9, 1999.
- [27] R. Ho, "Multiple regression," in *Handbook of univariate and multivariate data analysis and interpretation with SPSS* ed United States of America: Chapman & Hall, 2006, p. 15.
- [28] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis*, Seventh Edition ed. New Jersey: Pearson Education, 2010.
- [29] C. J. Willmott, "On the validation of model in physical geography," in *Spatial statistic and models*, ed: D. Reidel, 1981, p. 18.
- [30] C. J. Willmott, S. G. Ackleson, R. E. Davis, J. J. Feddema, K. M. Klink, D. R. Legates, J.O'Donnell, and C. M. Rowe, "Statistics for the evaluation and comparison of models," *Journal of Geophysical Research*, vol. 90, p. 11, 20th September 1985.