UAV-Based PM2.5 Monitoring for Small-Scale Urban Areas

Jumaah, H. J.,[1,2*] Mansor, S.,[2] Pradhan, B.[3] and Adam, S. N.[2]

[1]Department of Environment and Pollution Engineering, Technical College of Kirkuk, Northern Technical University, Kirkuk 36001, Iraq, E-mail: Huda.rs2018@gmail.com

[2]Department of Civil Engineering, Geospatial Information Science Research Centre (GISRC) Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Selangor, Malaysia

[3]The Centre for Advanced Modelling and Geospatial Information Systems(CAMGIS), Faculty of Engineering and Information Technology, University of Technology Sydney, 2007 New South Wales, Australia

*Corresponding author

Abstract

Air quality data such as Particulate Matter PM2.5 collection near the ground is difficult, particularly in small complex regions. This study aims to introduce a PM2.5 prediction algorithm based on measurements from Unmanned Arial Vehicle (UAV)-based sensing system and validate the model at a specified low altitude. Observations were applied around 1.6 km² area in University Putra Malaysia. This study uses an empirical method via applying amassed records of PM2.5 and meteorological parameters to produce a predictive Geographically Weighted Regression (GWR) model. An accuracy value is computed from the probability value given by the regression analysis model. To validate this approach, we have utilized training and testing data. To evaluate and validate the suggested model, we applied the model to the training set. The obtained result indicated that there is a good statistical correlation, and demonstrated that the characteristics obtained by analysis are able to predict the concentration of PM2.5.

1. Introduction

Air quality data in the urban area such as fine particle matter PM2.5 is of high importance to control contamination of air and to preserve human life (Zheng and Hsieh, 2013). Many recent epidemiological types of researches have displayed PM2.5 particles, which in populated areas are emitted primarily from anthropogenic and biogenic sources, are associated with various human health effects (Meng et al., 2005, Guo et al., 2016, Pope III et al., 2002, Hu et al., 2013 and Dockery et al., 1993). Expanded urbanization, industrial development and more vehicular utilization in most Asian countries, combined with trans-limit haze contamination and dust spreading in the air have added to the expansion of concentrations of PM in the air (Tahir et al., 2013 and Mahmud, 2017). Usually, air quality observations so, monitoring and mapping are directed by costly monitoring stations at fixed locations (Tian and Chen, 2010 and Sevusu, 2015). Which frequently fairly sparse and irregularly set apart, so interpreting the statistics from these monitoring stations can rarely display a complete explanation of the regional air quality (Tian and Chen, 2010). In Malaysia, they have specific Guidelines intended for monitoring air quality which depends on the Recommended Malaysia Ambient Air Quality Guidelines (RMAAQG) (Amir, 2007). Satellites are utilized to estimate air contamination concentrations along wide areas. However, they are inconvenient for an application on a small scale ranges (Ende, 2016). The meteorological impact on PM2.5 particle matter concentrations was used to study the estimation of PM2.5 concentration (Shith et al., 2017). Use of the Unmanned-Aerial Vehicles UAVs is an evolving tool to obtain different information. UAVs are practical in all manners of environments of various sizes on a moment monitoring (Hemmelder, 2016).

Different studies have used statistical methods to evaluate air pollutants levels and some have relied on mobile sensors (Alvear et al., 2015). Empirical statistical models outcomes consider more accurately distributions for retrieving of PM2.5 concentrations when compared with the different models such as scaling factor models and physical analysis models (Jiang et al., 2017). The underlying concept of the Geographically Weighted Regression (GWR) is that variables can be predicted anywhere in the location given a dependent parameter and one or more independent parameters that have been recorded at sites whose position is known (Charlton et al., 2009). GWR is defined as analyzing of spatial variation relationships (Fotheringham et al., 2002). A process of exploring spatial non-stationary by calibration of the multivariate regression models that permits different correlations to occur at various points in the space (Wankie, 2013 and LeSage, 2004). An alternate approach reflects GWR as an econometric misspecification detecting tool (McMillen, 2004). Statistical analysis can be conducted several runs of variable selection for model building to clarify the correlation and conclude outcomes (Mao et al., 2012). The analysis of satellite data is a helpful tool for monitoring PM2.5 levels particularly in areas where the ground measurements are not available. Resultant coefficient correlation R^2 between PM2.5 and Moderate Resolution Imaging Spectroradiometer (MODIS) Aerosol Optical Thickness (AOT) increased from 0.52 to 0.62 when, hourly data of particle matter concentrations were utilized instead of daily mean data of PM2.5 concentrations, and the regression equations which were estimated are applicable to calculate PM2.5 levels (Gupta and Christopher, 2008). Satellite remote sensing has a strong potentiality extended to ground measurement in local ambient Air Quality Parameters (AQPs) monitoring. A positive correlation of 0.75 R^2 concluded by MODIS-AOT and data of 24 hourly mean PM2.5 levels in Kuala Lumpur. Through the resultant map, the retrieved AOT distribution from MODIS can have significant potentials to support air quality monitoring by environmental agencies (Youssef et al., 2016). Otherwise a study of predicting PM2.5 concentrations introduced in San Francisco Bay Region using products of Aerosol Optical Depth (AOD), concluded that multiple regression equations can assist monitoring of air quality and decision making, but are not true with R^2 value of 0.11, which mean that 11% of the variance can be interpreted with regression equation in spite of the results were comparable to a previous study with same R^2 (Jennings, 2013). Though a variety of advanced models of PM2.5 prediction have been made, maximum of studies were restricted to certain cities or areas. In addition, maximum of these studies have used lower resolution AOD products in predicting PM2.5 levels. This generally used resolution of AOD products is frequently so coarse and therefore inadequate to define exposure estimations in urban areas (Youssef et al., 2016). A multivariate analysis is applied in this study to calculate PM2.5 concentrations in a small-scale area in University Putra Malaysia.

2. Methodology

2.1 Study Area Description

The study area is about 1.6 km² that lies between 101° 42′ 26″ E to 101°43′ 40″ E longitude and 3° 00′ 43″ N to 2° 59′ 46″ N latitude in University Putra Malaysia. Most of the measurements were done in the Faculty of Engineering in University Putra Malaysia which lies at the main Serdang Campus, some 22 km south of Kuala Lumpur (Figure 1).

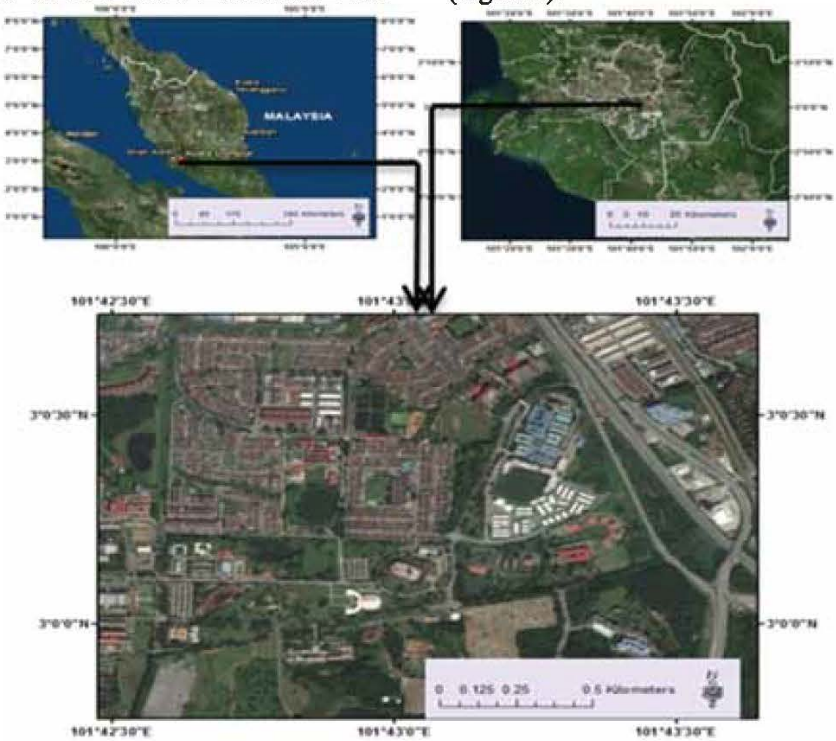

Figure 1: The location map of the study area

This area was selected in this study for some reasons; the availability and accessibility of the location and different weather from other places in addition to ease of using the instruments and data acquisition.

2.2 Methodology Adopted for PM2.5 Estimation.
The main scheme of this study is expressed in Figure 2.

2.3 In Situ Data Collection and UAV
In situ data measurements were collected from 80 points chosen randomly on the proposed horizontal path along 3.12 km track around 1.6 km^2 area in University Putra Malaysia during July. We used 80 points involving Particle Matter (PM2.5), Temperature of the air (T), Humidity (H), and the Wind Speed (WS) records. In addition to 20 points measured in August. A ground measurement of (PM2.5, H, T) of points selected randomly in engineering faculty UPM to perform the model. Unmanned Aerial Vehicles outfitted with various sensors have been presented to monitor the quality of air, for instance, they can suggest new approaches and research openings in air contamination and monitoring (Barnard, 2006). The UAV properties used in this research is Tarot 680PRO six axis of folding vehicle TL68P00 Hexacopter, is provided by UPM Geospatial Information Research Centre GISRC Laboratory. Figure 3 represents the UAV-based PM2.5 monitoring in the study area.

2.4 Predicted Model
ArcGIS offers significant tools for inclusive, raster-based spatial analysis (Ajaj et al., 2017). For model prediction and data analysis, we used ArcGIS (version 10.3) and its spatial statistics for cross-validation and modeling spatial relationship and spatial prediction of PM2.5 concentration using (GWR). A multivariate linear regression technique is the most frequently applied statistical method for relating one or a set of variable (more than one) specified by the equation (Shareef et al., 2014, Seber and Lee 2012 and Weisberg, 2005):

$$y = X\beta + \varepsilon$$

Equation 1

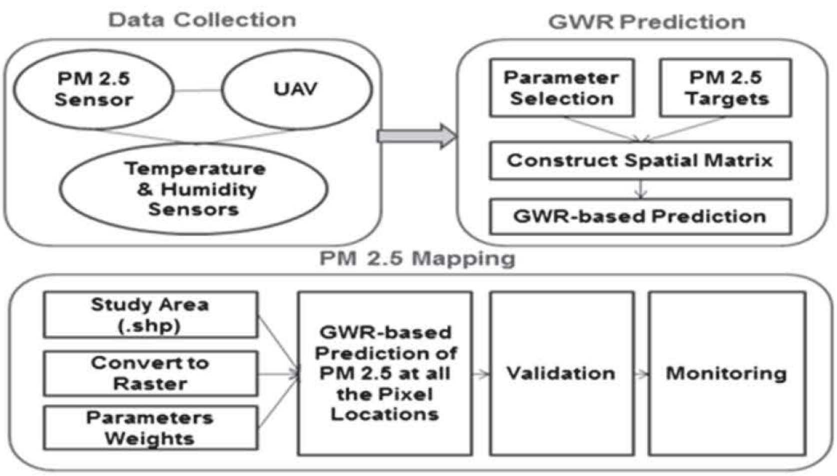

Figure 2: Methodology adopted for PM2.5 estimation

Figure 3: UAV-based PM2.5 monitoring

Where; y is a vector of response, X is the design matrix of regression variables, β is the vector of the parameters, and ε is the vector of the random error. GWR is an extension of the global regression model, and the comprehensive theory is presented in the equation (Andersson, 2017):

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Equation 2

Where y; dependent variable, x; independent variable, β_0; intercept, β_1; slope coefficients, and ε: error term. Based on the collected data by the UAV based PM2.5 monitoring system in July, we performed GWR model to estimate the concentrations of PM2.5 at all points. Depending on the basic formula of GWR, equation (2). The statistical analysis for model constructing was conducted through three runs of variable selection. The first case, the GWR was employed to analyze the relationship between each independent variable (except for temperature variable) and dependent variable PM2.5. The second case, the GWR was used to study the relationship between the independent variable humidity and the dependent variable PM2.5. The third case of variable selection was the independent variable wind speed with the dependent variable PM2.5 to construct the model. Cause no correlation we excluded temperature variable from the equation and model constructing.

After GWR analysis has been done, three predictive models to calculated PM2.5 have been constructed. The best models that we have obtained based on July observations, represented in equations (3, 4, and 5) respectively:

$$\mathrm{PM2.5_{calculated}} = \beta_0 - \beta_1 * H + \beta_2 * Ws$$

Equation 3

$$\mathrm{PM2.5_{calculated}} = \beta_0 - \beta_1 * H$$

Equation 4

$$\mathrm{PM2.5_{calculated}} = \beta_0 + \beta_2 * Ws$$

Equation 5

Based on August observations, ground measurements of PM2.5, T, and H parameters were used to predict PM2.5 concentrations, the model that we have obtained represented in equation (6):

$$\mathrm{PM2.5_{calculated}} = \beta_0 - \beta_1 * H$$

Equation 6

Also, cause no correlation we excluded temperature variable from the equation and model constructing. We used two kinds of data as training and testing. Training data were collected by PM2.5 monitoring system on the ground in deferent positions. While tested data collected by using Air Quality Multimeter.

2.5 Validation of the Multivariate Predictive Algorithm

Three frequently models were considered using the GWR method based on the measured AQPs. Next, we evaluated the validation (also called fitting) of each station using the tested data. Measured AQPs denotes the observations were taken from the positions of points in the study area and, calculated AQPs refers to the parameters predicted via applying the model once substituting of tested data. In order to get a strong validation, the validation applied for 24 different stations have same coordinates of measured values to find the differences between calculated and measured values. The same fitting and validity process was applied to the additional model where considered using the GWR technique based on ground measured AQPs (PM2.5 and H). In order to acquire a strong validation, the validation applied to 7 different stations has same coordinates of measured values to find the differences between calculated values and the measured.

2.6 Validation by Fitting and Confidence Bound Models

Data fitting technique includes mathematical equations and nonparametric processes. A polynomial model is a function that can be specified by:

$$P(x) = c_0 + c_1 x + ... + c_n x^n$$

Equation 7

For some coefficients $c_0, \dots, c_n$. If $c_n = 0$ then, the polynomial is suggested to be of order n. The confidence bounds for fitting coefficient specified by:

$$C = b \pm t\sqrt{S}$$

Equation 8

Where: b is the coefficient created by the fit, t dependence on the level of confidence, and is calculated using the inverse of t accumulative distribution function, and S is the diagonal elements vector from the expected covariance matrix. Simultaneously prediction bounds for the predictors value and the function are specified by:

$$P_{s,p} = y \pm f\sqrt{xSx^T}$$

Equation 9

Where, f is related to confidence level and is computed using the inverse of the F cumulative distribution function (Shareef et al., 2014).

3. Result and Discussion

3.1 Generation of the Multivariate Predictive Algorithm.

The results of the best models that we have achieved denoted in the equations (10, 11, 12 and, 13) respectively and as follows:

$$\mathrm{PM2.5_{calculated}} = 63.67 - 0.52 * \mathrm{H} + 2.48 * \mathrm{Ws}$$

Equation 10

$$\mathrm{PM2.5_{calculated}} = 57.645 - 0.167 * \mathrm{H}$$

Equation 11

$$\mathrm{PM2.5_{calculated}} = 44.250 + 3.457 * \mathrm{Ws}$$

Equation 12

$$\mathrm{PM2.5_{calculated}} = 78.978 - 0.596 * \mathrm{H}$$

Equation 13

Where PM2.5 ($\mu g/m^3$) represents the particulate matter concentration at each point, H (%) represents the relative humidity and WS(m/sec) represents the wind speed. Table 1 shows the descriptive statistics created by the GWR tool. From the report created by the GWR tool, we can obtain the local R^2 and the local R^2 Adjusted (the adjusted R^2).

The local R^2 values for July are 0.41, 0.51 and 0.65, respectively. The local R^2 values for August is 0.73. The values for the two months are denoted that the overall performance of the model relatively good in each model. From this view confirms previous findings relationship within parameters in the study area, that show including meteorological parameters, such as WS, H would Contribute the PM2.5 estimation but, meteorological data alone is more useful for estimating PM2.5. Using only PM2.5 and WS or only PM2.5 and H, preamble height accounted for the models' ability to estimate PM2.5 and therefore, WS height should be included when utilizing to estimate PM2.5 concentrations. H was the next most important independent variable, then followed by the other variable(T). In case 1, equation (10) and, due to the existence of points outside the boundary condition made the model strongly discriminate among used points. Though, these points effect if have different values compared with other points values (Shareef et al., 2014). For this reason, we noticed that if we omit undesirable points it will increase the R^2 from 0.41 to 0.50, this result obtained after neglecting five points. The values of the predicted PM2.5 concentration in July and August can be mapped; these are shown in Figure 4.

3.2 Validating of the Predictive Algorithm

Two processes of validation have been done to evaluate the strength of equations. Validation of training and testing points. Figure 5 shows the scatterplot of GWR model validation with training and with testing regions. Table 2 shows all GWR models with different variable combinations and validation results in this study.

Table 1: Descriptive statistics created by the GWR tool based on July and August measurements

	July				August
ID	**VARNAME**	**VARIABLE**	**VARIABLE**	**VARIABLE**	**VARIABLE**
1	Bandwidth	0.011	0.006	0.001	0.002
2	Residual Squares	484.312	402.091	287.760	234.349
3	Effective Number	4.115	4.331	24.072	5.117
4	Sigma	2.526	2.305	2.268	3.968
5	AICc	381.013	365.720	385.789	120.710
6	R^2	0.415	0.514	0.652	0.730
7	R^2Adjusted	0.391	0.493	0.509	0.655
8	Dependent Field	PM2.5	PM2.5	PM2.5	PM2.5
9	Explanatory Field1	H	H	WS	H
10	Explanatory Field2	WS	-	-	-

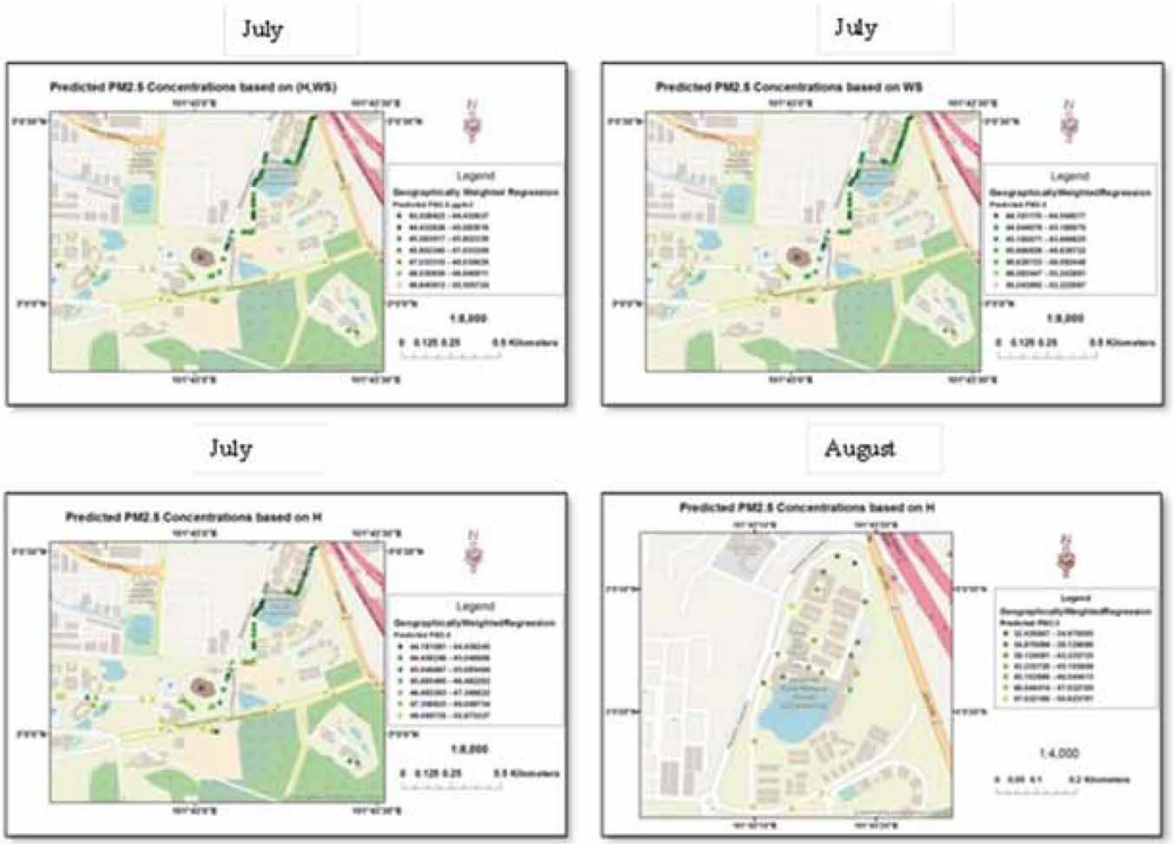

Figure 4: Predicted PM2.5 concentrations of the GWR model based on H&WS July and August 2017

Table 2: The list of GWR models with the variable combinations and validation result

Predicted model	variables	R^2 (July)		R^2 (August)	
		Validation with training	**Validation with testing**	**Validation with training**	**Validation with testing**
PM2.5	H, WS	0.415	0.820	-	-
PM2.5	H	0.514	0.919	0.730	0.946
PM2.5	WS	0.652	0.934	-	-

Validation with training points has been done to evaluate the accuracy of all predictive algorithms. This validation uses the polynomial linear fitting to calculate PM 2.5 concentrations using predictive models and comparing them with measured (trained data) of AQPs. Validating of predictive algorithms applied again with a testing region. 30% of training points measured again as tested AQPs. These points also selected randomly. This validation also uses the polynomial linear fitting to calculate the AQPs, but using predictive models and comparing them with measured (tested data) of AQPs. All estimated parameters using this data appeared a good agreement with ground parameters.

4. Conclusion

There is an urgent necessity to know about the changing levels of air contamination in cities because of their significant impact on the health and to take the precautionary measures. For achieving this, attempts were made to develop a model that is useful for obtaining air quality information. This study aims to introduce a predictive model of PM2.5 concentration and validate the model at low altitudes. A GWR equation is introduced to predict PM2.5 from metrological parameters (T, H, and WS). Based on the study results, the performance of GWR model specified that the model was relatively accurate in predicting PM2.5 levels.

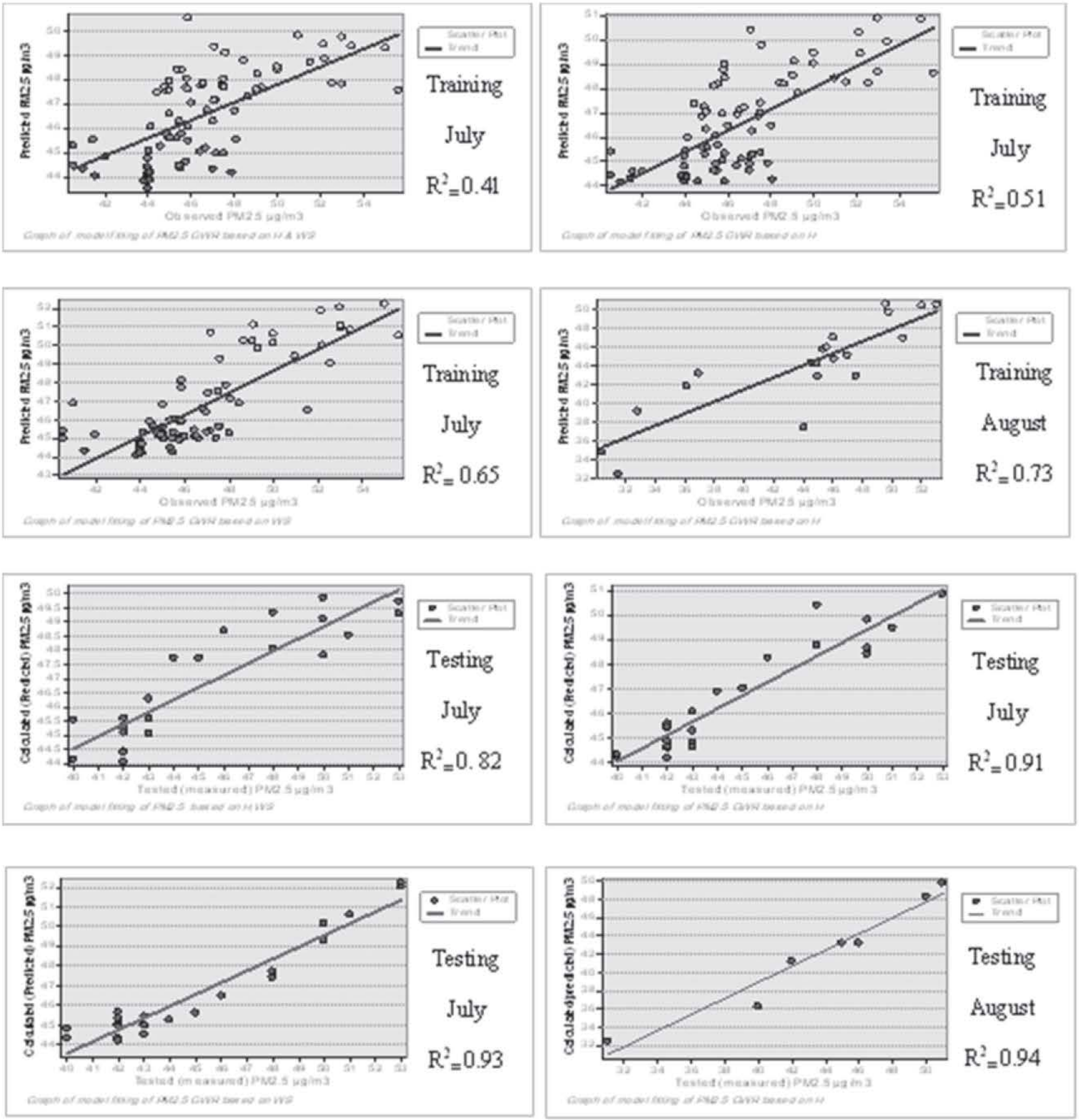

Figure 5: Scatterplots of GWR model validation with training and testing regions July and August 2017

But relating to the R^2 values which were set as (0.41,0.51 and, 0.65) respectively for July and 0.73 for August, indicated that the meteorological parameters; WS and H would contribute PM2.5 predicting but, meteorological data alone is more convenient for estimating PM2.5 in the study area. Therefore, WS highly should be included when utilizing to predict PM2.5 levels. H was the next most important independent variable, then followed by the variable (T). Additionally, we concluded a good statistical correlation among the measured and testing data by validation with testing points. Resultant R^2 values are (0.82, 0.91 and, 0.93) of July and, (0.94) R^2 value of August 2017.

Acknowledgments

The author is grateful to Prof. Dr. Shattri Bin Mansor for the beneficial suggestion. The author also would like to thank University Putra Malaysia GISRC Lab for providing the UAV for data collection. Also the author thanks Mr. Azman A. Ghani for his help to fly the UAV for data collection that used in some of the analyses were applied in this work. Extended thanks go to Dr. Biswajeet Pradhan and Dr. Muntadher A. Shareef for their valuable comments which helped in improving the quality of this paper. This study is a part of MSc research of Huda Jamal.

References

Ajaj, Q. M., Pradhan, B., Noori, A. M. and Jebur, M. N., 2017, Spatial Monitoring of Desertification Extent in Western Iraq Using Landsat Images and GIS. *Land Degradation & Development*, Vol. 28, 2418–2431.

Alvear, O., Calafate, C. T., Hernández, E., Cano, J. C. and Manzoni, P., 2015, Mobile Pollution Data Sensing Using UAVS. *In Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia*, ACM, 393–397.

Amir, A., 2007, Air Pollution Trends in Petaling Jaya, Selangor, Malaysia. MSc thesis, Universiti Putra Malaysia, Malaysia.

Andersson, J., 2017, *Using Geographically Weighted Regression (GWR) to Explore Spatial Variations in the Relationship between Public Transport Accessibility and Car Use: A Case Study in Lund and Malmö, Sweden*. Student Thesis Series INES, Lund University, Sweden.

Barnard, J. A., 2006, Unmanned Air Vehicle Features, Applications and Technologies. *Barnard Microsystems Limited*, 1-146.

Charlton, M., Fotheringham, S. and Brunsdon, C., 2009, Geographically Weighted Regression. White Paper. National Centre for Geocomputation. National University of Ireland Maynooth, 1-14.

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G. and Speizer, F. E., 1993, An Association between Air Pollution and Mortality in Six US Cities. *New England Journal of Medicine*, Vol. 329, 1753–1759.

Ende, P., 2016, Modelling Air Pollution And Personal Exposure in Bangkok and Mexico City Using a Land Use Regression Model. MSc thesis, University of Twente, Netherlands.

Fotheringham, A. S., Brunsdon, C. C. and Charlton, M. M., 2002, *Geographically Weighted Regression: The Analysis Of Spatially Varying Relationships*. 1st edition, (England: John Wiley & Sons, Ltd).

Guo, H., Cheng, T., Gu, X., Chen, H., Wang, Y., Zheng, F. and Xiang, K., 2016, Comparison of Four Ground-Level PM2. 5 Estimation Models Using Parasol Aerosol Optical Depth Data from China. *International Journal of Environmental Research and Public Health*, Vol. 13, 180.

Gupta, P. and Christopher, S. A., 2008, Seven Year Particulate Matter Air Quality Assessment from Surface and Satellite Measurements. *Atmospheric Chemistry and Physics*, Vol. 8, 3311–3324.

Hemmelder, S. D., 2016, The Application and Suitability of Unmanned Airborne Vehicle Based Imagery to Determine River Dynamics: A Case Study of the Buëch in France. MSc thesis, Utrecht Universit, France.

Hu, X., Waller, L. A., Al-Hamdan, M. Z., Crosson, W. L., Estes Jr, M. G., Estes, S. M., Quattrochi, D. A., Sarnat, J. A. and Liu, Y., 2013, Estimating Ground-Level PM2. 5 Concentrations in the South Eastern US Using Geographically Weighted Regression. *Environmental Research*, Vol. 121, 1–10.

Jennings, C. A., 2013, Estimating PM2. 5 Concentrations Using MODIS and Meteorological Measurements for the San Francisco Bay Area. MSc thesis, San Francisco State University, San Francisco, California.

Jiang, M., Sun, W., Yang, G. and Zhang, D., 2017, Modelling Seasonal GWR of Daily PM2.5 with Proper Auxiliary Variables for the Yangtze River Delta. *Remote Sensing*, Vol. 9, 346.

LeSage, J. P., 2004, A Family of Geographically Weighted Regression Models. In Advances in Spatial Econometrics, Springer, Berlin, *Heidelberg*, 241–264.

Mahmud, M., 2017, Active Fire and Hotspot Emissions in Peninsular Malaysia During the 2002 Burning Season. *Geografia-Malaysian Journal of Society and Space*, Vol. 1, 32-45.

Mao, L., Qiu, Y., Kusano, C. and Xu, X., 2012, Predicting Regional Space–Time Variation of PM 2.5 with Land-Use Regression Model and MODIS Data. *Environmental Science and Pollution Research*, Vol. 19, 128–138.

McMillen, D. P., 2004, Geographically Weighted Regression: the Analysis of Spatially Varying Relationships. *Journal of Agricultural Economics*, Vol. 86, 554–556.

Meng, Q. Y., Turpin, B. J., Korn, L., Weisel, C. P., Morandi, M., Colome, S., Zhang, J., Stock, T., Spektor, D., Winer, A. and Zhang, L., 2005, Influence of Ambient (Outdoor) Sources on Residential Indoor and Personal PM 2.5 Concentrations: Analyses of RIOPA Data. *Journal of Exposure Science and Environmental Epidemiology*, Vol. 15, 17-28.

Pope III, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K. and Thurston, G. D., 2002, Lung Cancer, Cardiopulmonary Mortality, and

Long-Term Exposure to Fine Particulate Air Pollution. *Jama*, Vol. 287, 1132–1141.

Seber, G. A. F. and Lee, A. J., 2012. Linear Regression Analysis. 2nd edition, (New Jersy: John Wiley & Sons).

Sevusu, P., 2015, Real-Time Air Quality Measurements Using Mobile Platforms. MSc thesis, Rutgers The State University of New Jersey-New Brunswick.

Shareef, M. A., Toumi, A. and Khenchaf, A., 2014, Prediction of Water Quality Parameters from SAR Images by Using Multivariate and Texture Analysis Models. In SAR Image Analysis, Modeling, and Techniques XIV, *International Society for Optics and Photonics*, Vol. 9243, 924319.

Shith, S., Yusof, N. F. F. M., Ramli, N. A. and Elbayoumi, M., 2017, Characterization of Chemical Composition in Fine Particles (PM2. 5) from Industrial Site in Malaysia. *Sustainability in Environment*, Vol. 2, 104.

Tahir, N. M., Suratman, S., Fong, F. T., Hamzah, M. S., and Latif, M. T., 2013, Temporal distribution and Chemical Characterization of Atmospheric Particulate Matter in the Eastern Coast of Peninsular Malaysia. Aerosol and Air Quality Research, Vol. 13, 584–595.

Tian, J. and Chen, D., 2010, A Semi-Empirical Model for Predicting Hourly Ground-Level Fine Particulate Matter (PM2. 5) Concentration in Southern Ontario from Satellite Remote Sensing and Ground-Based Meteorological Measurements. *Remote Sensing of Environment*, Vol. 114, 221–229.

Wankie, C., 2013, Local Spatial Modeling Using Geographically Weighted Regression (GWR). MSc thesis, California State University, Long Beach, California, United States.

Weisberg, S., 2005, *Applied Linear Regression*. 3rd edition, (Hoboken, New Jersey : John Wiley & Sons).

You, W., Zang, Z., Zhang, L., Li, Y., Pan, X. and Wang, W., 2016, National-Scale Estimates of Ground-Level PM2.5 Concentration in China Using Geographically Weighted Regression Based on 3 km Resolution MODIS AOD. *Remote Sensing*, Vol. 8, 184.

Youssef, K. B., Abdullah, A. M., Shafri, H., Ashaari, Z. H., Gumel, D. U. and Yusuf, H. Y., 2016, Estimation of Aerosols dispersion & Urban Air Quality evaluation over Malaysia using MODIS Satellite. International Journal of Advanced Scientific and Technical Research, Vol. 3(6), 229-238.

Zheng, Y., Liu, F. and Hsieh, H. P., 2013, U-air: When Urban Air Quality Inference Meets Big Data. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining ACM.* 1436–1444.