ESTIMATING PM_{10} CONCENTRATIONS IN THE MAIN ROUTES OF HANOI, VIETNAM BY USING VNREDSat-1 DATA

Nguyen Nhu Hung[(1)], Le Minh Hang[(1)], Tran Van Anh[(2)], Du Vu Viet Quan[(3)], Van-Manh Pham[(3)]

[1] Le Quy Don Technical University, 236 Hoang Quoc Viet, Bac Tu Liem, Ha Noi, Viet Nam
[2] University of Mining and Geology, 18 Pho Vien, Duc Thang, Bac Tu Liem, Ha Noi, Viet Nam
[3] VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Ha Noi, Viet Nam
Email: nhuhungdhktlqd@gmail.com; leminhhang81@gmail.com; tva_ninh@yahoo.com; duvuvietquan@gmail.com; manh10101984@gmail.com

KEYWORDS: PM_{10}, VNREDSat-1 imagery, Linear regression model

ABSTRACT: Estimation of Particulate matter 10 micrometers (PM_{10}) from remote sensing data is a practical approach for long-term monitoring quality of air pollution, which assist the urban management and environmental sustainability. This study presents the use of VNREDSat-1 image data for retrieving environmental parameters to estimate the spatial variability of PM_{10} in Hanoi city, Vietnam. linear regression models were performed by using input from field survey data acquired on May, 7th, 2013. In total, 14 points were collected and 10 points were used in the regression models; the remaining 4 points were used to validate the results. Quality control was conducted by comparing the coefficient of determination (R squared - R^2) of the best-fit model to estimate the PM_{10} accurately. The correlation coefficient of this proposed model archived 0.943 and RMSE was 52 $\mu g/m^3$. The study area is the main routes of Hanoi, Vietnam where have a high traffic density and high air pollution. This approach is an effective way to use satellite data in estimating PM_{10} in urban regions.

1. INTRODUCTION

Air pollution is currently an interesting research topic in the world. Recently, the World Health Organization (WHO) has released a new report, which warns that serious pollution in many cities around the world is killing millions of people. According to the WHO, urban air pollution causes 1.3 million deaths and about 4.6 million people of reduction of life expectancy each year in the world due to air pollution (Report of WHO, 2014). With high urbanization, Hanoi, the capital of Vietnam is a hotspot for air pollution, especially for construction dust and vehicle emissions. According to statistics from ARIA Technologies, a French company that specializes in providing software for calculating, simulating air pollution and weather forecasting services, the average increase of vehicles is from 12% to 15%, which contributes to toxic emissions such as SO_2 and NO_x. According to data from the Center for Environmental Monitoring, Vietnam Environment Administration, many intersections such as Kim Lien – Giai Phong, Phung Hung - Ha Dong, are densely populated areas with high dust concentrations more than allowed, sometimes up to seven times. Other pollutants such as CO, SO_2 are also increasing.

Research on particle matter less than ten micrometers in the atmosphere (PM_{10}) concentrations is often based on spatial and temporal data as measured by ground stations spread across cities (Retalis *et al*., 1999; Sifakis *et al*., 1998). However, ground stations require high installation and maintenance costs. Besides that, the data collected through these methods is just shown the air pollution in a small area around the station. Therefore, ground measurements cannot provide a detailed spatial distribution of PM_{10} air pollutants in large areas (Wald *et al*., 1999). Recent satellite remote sensing provides quantitative information of dust particle in the air with acceptable accuracy compared to measurements at surveying stations. Studies in the world have shown that the use of multispectral satellite imagery can completely detect air pollution in the areas of interest. Some studies have given out the relationships between satellite data and air pollution (Ung et al., 2001; Weber *et al.*, 2001). There have been many types of research on application of satellite data for environmental and atmospheric study such as (Asmala *et al*., 2002) using NOAA-14 AVHRR for haze detection that was a key for API evaluation at Peninsular Malays; other researches used Landsat 7 ETM^+ imagery (Ung *et al*., 2001; Lim *et al*., 2004; Nadzri *et al*., 2010) for determining the PM_{10}; or used SPOT images (Sifakis *et al*., 1992) or MODIS images (Péré *et al*., 2009) for air quality evaluation.

In Vietnam, there are some scientists have studied the possibility of monitoring air pollution using

Landsat 7 satellite imagery (Nguyen *et al*., 2014; Tran *et al*., 2014) or using SPOT5 images (Luong *et al*., 2010). However, most of the studies used only dust measurement points at fixed ground stations in the city area; therefore, the number of points were quite small. For Hanoi area, there are only five dust measurement stations or six dust measurement stations located in Ho Chi Minh City (Tran *et al*., 2014). With such a small number of measurement points, the calculation of the regression function will be less accurate and in addition, there were no redundant measurement points to validate the regression model. VNREDSat-1 (Vietnam Natural Resources, Environment and Disaster - monitoring Satellite-1) is Vietnam's first optical Earth Observing satellite which was successfully launched on May 7, 2013. There have been many studies on applications of the images in land use and land cover but few ones have used the VNREDSat-1 image in the PM_{10} determination. Thus, in this study, we present the potential of using VNREDSat-1 satellite imagery to estimate concentrations of PM_{10} in Ha Noi region. In addition, 16 measurement points by Dust Track II were used for calculating the regression function and validating the accuracy of the regression model.

2. STUDY AREA AND METHOD

2.1. Description of the study area

Hanoi city is a financial and trading center of Vietnam. For the last 15 years, the city has witnessed the rapid expansions of population, finance, constructions, and urbanization. These developments result in the growth of material transporting vehicles, which are the main reason for causing air pollution in the city.

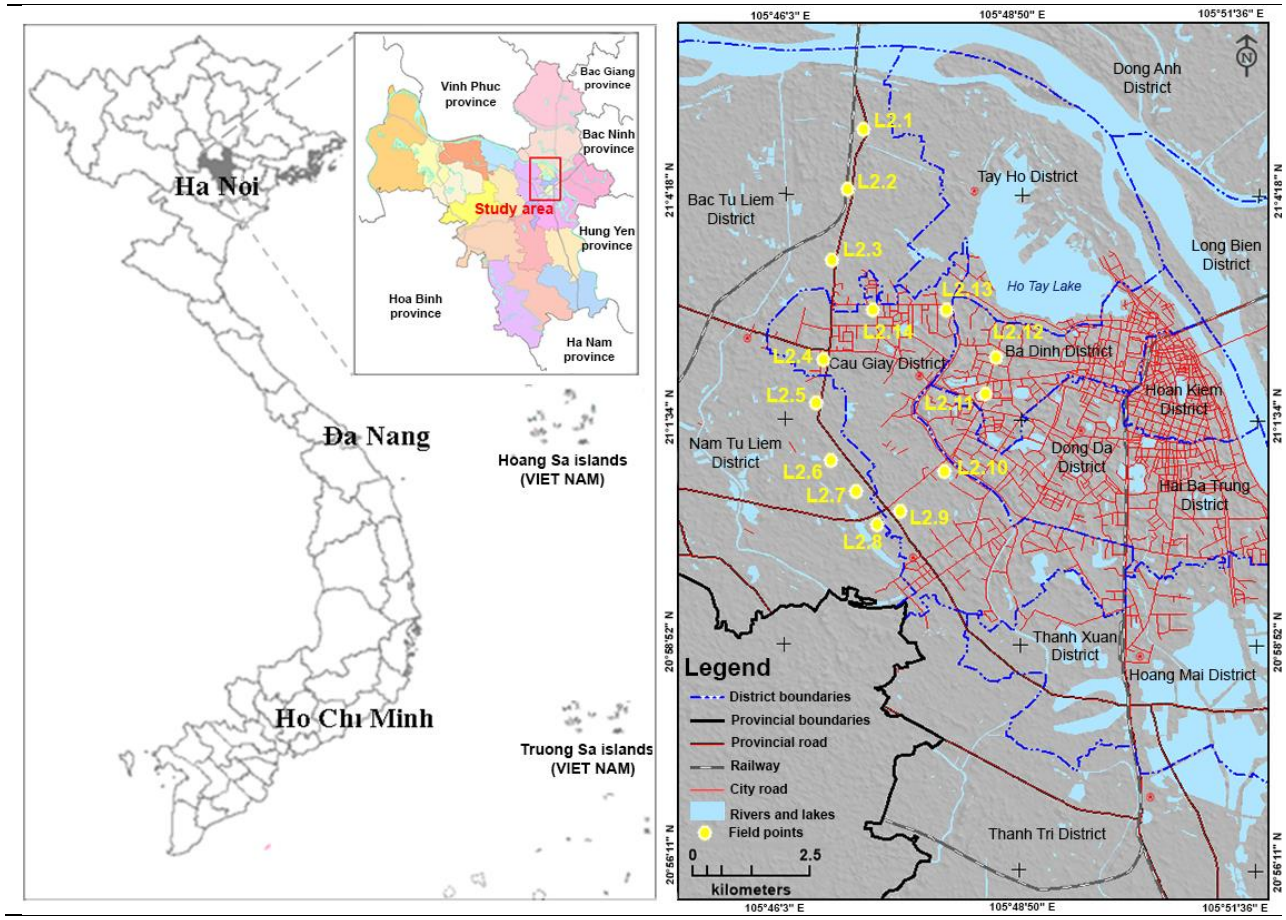


Figure 1. Study area and location of PM_{10} measuring points.

In addition, weather in Hanoi presents for typical Northern weather which is humid, hot and rainy in the summer, cold and dry in the winter. In the winter, northeast's wind carries dust and pollutants from neighbor regions combining with cold dry weather that cause the PM_{10} dust concentration increase. In contrast, in the summer, southwest and southeast winds bring frequent rains washing out the dust in the air which lower the PM_{10} dust concentration compared to it in the winter. Furthermore, PM_{10} dust concentration is also affected by city transportation. There are very few monitoring stations in Hanoi (two fixed stations and eight sensor stations) which make it is hard to evaluate the air quality of the whole area. That is why there a need for PM_{10} dust concentration monitoring and

evaluating techniques for a large area in order to give early warnings.

2.2. Methodology

2.2.1. Satellite image pre-processing and Field data collection

Since the ultimate objective of this study is estimating PM_{10} concentrations in the main routes of Hanoi, high resolution, and multispectral data should be used. Indeed, the selection of satellite images is much dependent on the availability of data and date of capture. For that reason, the VNREDSat-1 image, captured on May 7th, 2013, was considered as high-resolution data and suitable for this research (Table 1).

Table 1. Description of satellite data.

Type of image	Bands	Spatial resolution (m)	Central wavelength (µm)	Spectral range (µm)
Multispectral	Blue	10	490	0.45 – 0.52
	Green	10	550	0.53 – 0.59
	Red	10	660	0.62 – 0.69
	Near-Infrared	10	830	0.76 – 0.89
Panchromatic	-	2.5	600	0.45 – 0.75

Converting DN value to surface reflectance value according to COST method: The COST (Cosine of the Solar Zenith Angle) method combines the DOS method assumptions with the fact that very few objects on the surface of the earth are dark objects (Zhang et al., 2014).

Therefore, it is usually equivalent to 1% of the reflected image. The radiation value of an object that is completely dark and unaffected by the shadow is calculated as follows:

$$L_{1\%\lambda} = \frac{0.01 E_\lambda \cos\theta_i}{d^2 \pi} \quad (1)$$

$L_{1\%\lambda}$ is the value of 1% of the dark object's radiation according to the assumption. The radiation value is then converted into a surface reflection of the Earth using the following formula:

$$\rho_\lambda = \frac{(L_{sen,\lambda} - (L_{haze,\lambda} - L_{1\%\lambda}))d^2\pi}{E_\lambda cos\theta_i} \quad (2)$$

In which: ρ_λ-reflection value on the satellite for wavelength λ; d-the distance between the Earth and the Sun; E_λ-illuminance of upper atmosphere from the average Sun; θ_i-angle of the sun.

Field data collection: Ground measurement data collection was conducted at the same time as satellite imaging by the Dust Trak II instrument. Each measured point was determined location by GPS device and the measurement time for collecting dust was also recorded. The locations for ground measurements were illustrated in Fig. 1. The PM_{10} values taken from field data collected in sample points ranged from 108 µg/m^3 to 270 µg/m^3, with a mean of 170 µg/m^3 and a standard deviation of 52 µg/m^3. The whole set of field observation points was divided into two parts, one for training models (select the interval 70% of points), and the others for validation and evaluating the prediction accuracy (about 30% of points).

2.2.2. Optical Thickness (AOT) and PM_{10} correlation

Solar energy enters the aerosol of the troposphere, under the impact of pollutant gas molecules and dust particles, partially reflected in the aerosol layer and then toward the satellite's receiver, the sun's rays reach the objects on the surface and then reflect to the satellite sensor. Based on the energy loss to the satellite receiver due to the absorption, scattering of pollutant gases and dust particles from which the dust content in the air is calculated (Sam Appadurai *et al.*, 2016).

After conducting the radiometric correction the reflectance measured from the satellite (reflectance at the top of the atmosphere, TOA) was subtracted by the amount given by the surface reflectance to obtain the atmospheric reflectance. The atmospheric reflectance was then related to PM_{10} using the regression Algorithm analysis. PM_{10} maps were generated using the proposed algorithm based on the highest R and lowest RMSE values. So the algorithm of AOT for single band or wavelength is simplified as (Sam Appadurai *et al.*, 2016):

$$AOT(\lambda) = a_0 R(\lambda) \quad (3)$$

Equation (3) is rewritten into three band equation as Equation (4).

$$AOT(\lambda) = a_1 R_{\lambda 1} + a_2 R_{\lambda 2} + a_3 R_{\lambda 3} \quad (4)$$

Where $R_{\lambda i}$ is the atmospheric reflectance (i = 1, 2 and 3 corresponding to wavelength for satellite), and a_i is the algorithm coefficient (i = 1, 2 and 3) determined experimentally.

The relationship between PM and AOT originates from a homogeneous atmosphere that contains globular spheres. Concentration at the surface is obtained after the drying of the air sample given by (Koelemeijer *et al.*, 2006). Therefore, PM content can correlate better with direct AOT. By replacing AOT with PM_{10} in Equation (4) we obtain Equation (5) (Lim HS *et al*., 2004; Nadzri *et al*., 2010).

$$PM_{10} = a_1 R_{\lambda 1} + a_2 R_{\lambda 2} + a_3 R_{\lambda 3} \quad (5)$$

Where $R_{\lambda i}$ is the atmospheric reflectance (i = 1, 2, and 3) corresponding to wavelength for satellite), and a_i is the algorithm coefficient (i = 1, 2, and 3) obtained from the experiment.

2.2.3. The accuracy assessment of regression models

In this step, the r-square and the RMSE (Root Mean Squared Error) were applied for assessing the regression models' accuracies. The r-square was used to evaluate the correlation between the observed PM_{10} and the predicted values. Moreover, the Root Mean Squared Error was used to examine model performance by comparing the deviation between observed and predicted data. The following formula was calculated by Equations (6) and (7):

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(Y_{ij} - \hat{Y}_{ij})^2}{\sum_{i=1}^{n}(Y_{ij} - \overline{Y}_{ij})^2} \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Y_{ij} - \hat{Y}_{ij})^2}{n - p}} \quad (7)$$

where $\hat{Y}$ is predicted the value and Y is the observed value of the dependent variable (PM_{10}). The r^2 of the independent validation was calculated as the square of correlation coefficient (R) for a linear transformation of the model. The model with the highest r^2 and lowest RMSE was selected for use in predictive mapping of PM_{10}.

3. EXPERIMENTAL RESULTS

3.1. Correlation and regression analysis

The first step in the modeling process is determining the model structure to predict PM_{10} concentration as the dependent variable and ρ_λ Atmospheric reflectance values were considered as independent variables. Based on Equation (5), the correlation and regression analysis of PM_{10} model with VNREDSat-1 was conducted. Previous studies focused on the single-variable linear regression between PM_{10} concentration and AOT data measured by different satellite sensors (Lim HS *et al*., 2004; Nadzri *et al*., 2010), mainly used the blue or green or red bands with a wavelength from 0.45 µm ÷ 0.69 µm. In this study, we used three bands (Blue, Green, and Red). From the original image data, a series of processing steps were taken, then pre-processed imagery data was combined with ground measurement data to compute the coefficient of the linear regression equation with the models shown in Table 2.

Table 2. Linear regression models used in this study

	PM_{10}							
Band 1	$a_1.R_1$	-	-	-	$a_1.R_1$	$a_1.R_1$	-	$a_1.R_1$
Band 2	-	+ $a_2.R_2$	-	-	+ $a_2.R_2$	-	+ $a_2.R_2$	+ $a_2.R_2$
Band 3	-	-	+ $a_3.R_3$	-	-	+ $a_3.R_3$	+ $a_3.R_3$	+ $a_3.R_3$

Among the 14 points of dust measurement sites in the field using the Dust Track II, 10 points were used for linear regression analysis available in MS Excel; the other 4 points were used for validating the regression model. The results of the regression calculation were shown in Table 3. The results of

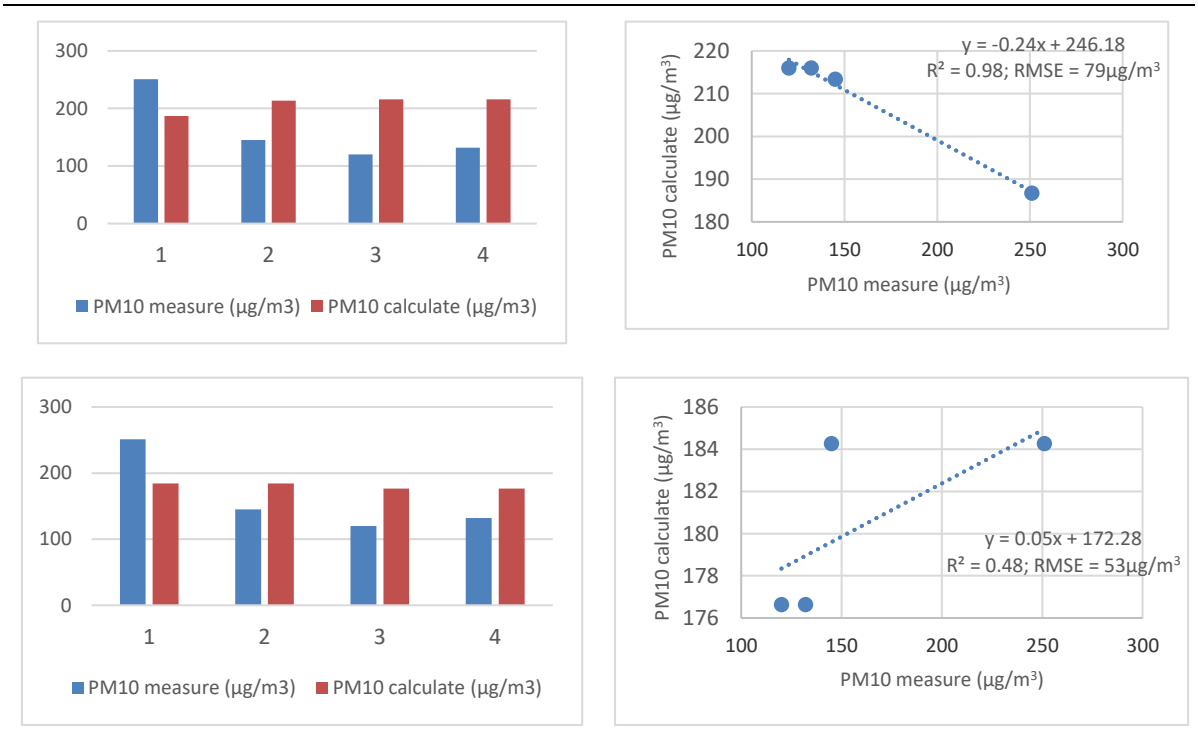
correlation and regression showed that when using independent spectral bands, the Coefficient of determination (r^2) was very low and the Root mean squared error (RMSE) was high, combining the spectral bands together gave better results. Coefficient of determination showed that when combining atmospheric reflectance was calculated from 3 bands. The results were the best seven models considered suitable for the PM_{10} prediction model. Thus, Model 7 can be used to calculate PM_{10} through three independent variables Band 1 (Blue), Band 2 (Green), and Band 3 (Red), with the coefficient of determination, was highest (r^2 = 0.943) and Root mean squared error (RMSE= 52 $\mu g/m^3$) was lowest.

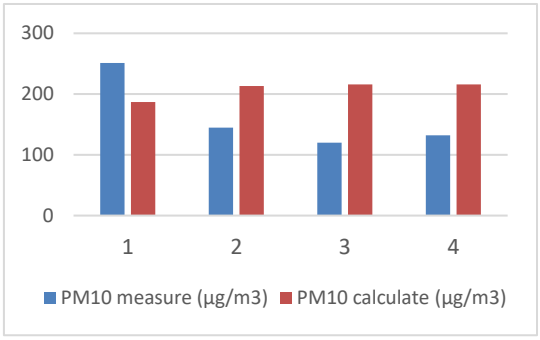
Table 3. Results of correlation and regression analysis

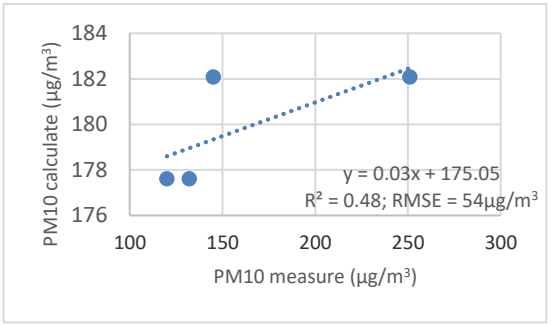
	a_1	a_2	a_3	r^2	RMSE ($\mu g/m^3$)
Model 1	1320	-	-	0.935	49
Model 2	-	1139	-	0.935	49
Model 3	-	-	936	0.929	51
Model 4	630	595	-	0.935	52
Model 5	3128	-	-1286	0.937	51
Model 6	-	5061	-3232	0.942	49
Model 7	738	4441	-3246	**0.943**	52

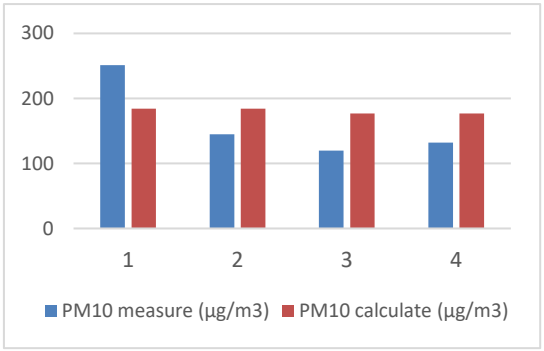
3.2. Validation of regression model

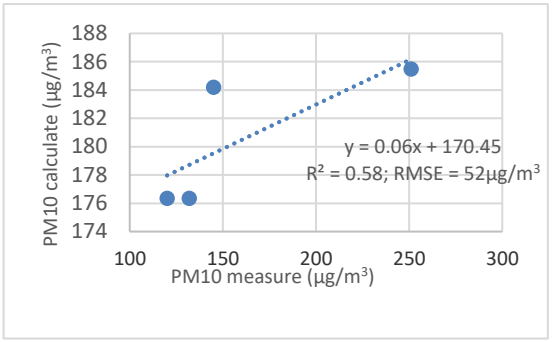
Based on the results calculated from Table 3, by using use the Raster Calculator tool in the ArcGIS software to establish a map of PM_{10}, 4/14 sample points (30% of points) were randomly selected to assess the performance of these models, of which the PM_{10} ranged from 120 to 251 $\mu g/m^3$. Figure 2 and 3 illustrates scatter plots of predicted versus observed PM_{10} to show the accuracy of the estimated PM_{10} by regression model from the best algorithm (lowest RMSE and highest r^2) were selected for the study area. In addition, a high correlation was shown between predicted and observed PM10 with the value of r^2 (0.773), and Root Mean Square (RMSE) of 49 $\mu g/m^3$ based on (Band 1 + Band 2 + Band 3) regression equation (linear model) for VNREDSat-1 imagery data.

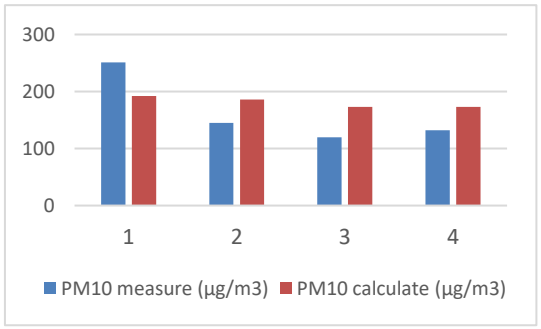


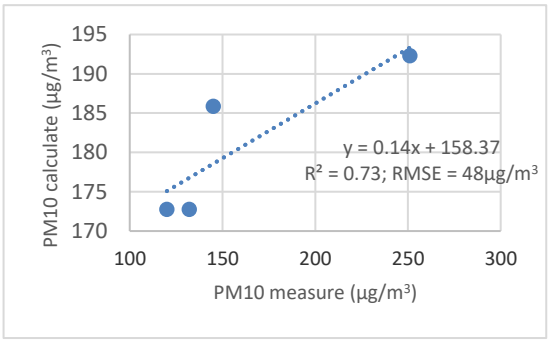


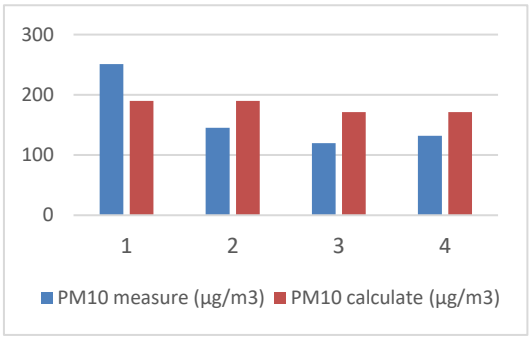


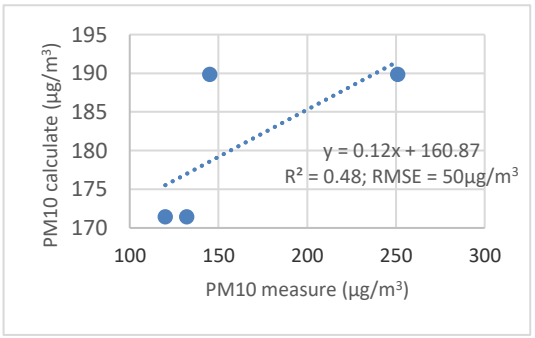


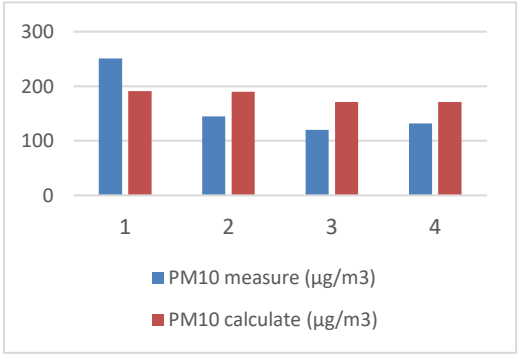












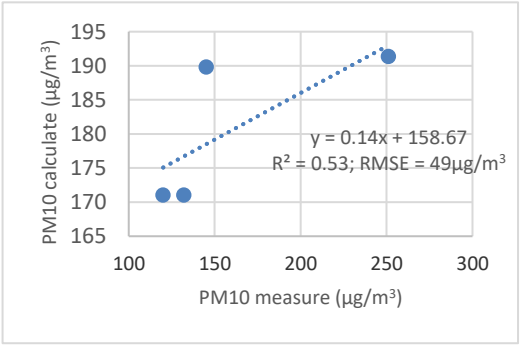


Figure 2. Comparisons of PM_{10} dust at check sites and values computed from images

Figure 3. Correlation between the PM_{10} values measured by the model (Table 3) and the field values

Among different models, the Linear regression model (with three variables: Band 1, Band 2, and Band 3) performed the optimal prediction for PM_{10} estimation and mapping of the study area. Based on the defined linear regression algorithm, a PM_{10} estimation map was established. The PM_{10} map was displayed by seven classes (Figure 4) shows: (i) less than 50 ($\mu g/m^3$), (ii) from 50 to 100 ($\mu g/m^3$), from 100 to 150 ($\mu g/m^3$), from 150 to 200 ($\mu g/m^3$), from 200 to 300 ($\mu g/m^3$), from 300 to 500 ($\mu g/m^3$), and over 500 ($\mu g/m^3$).

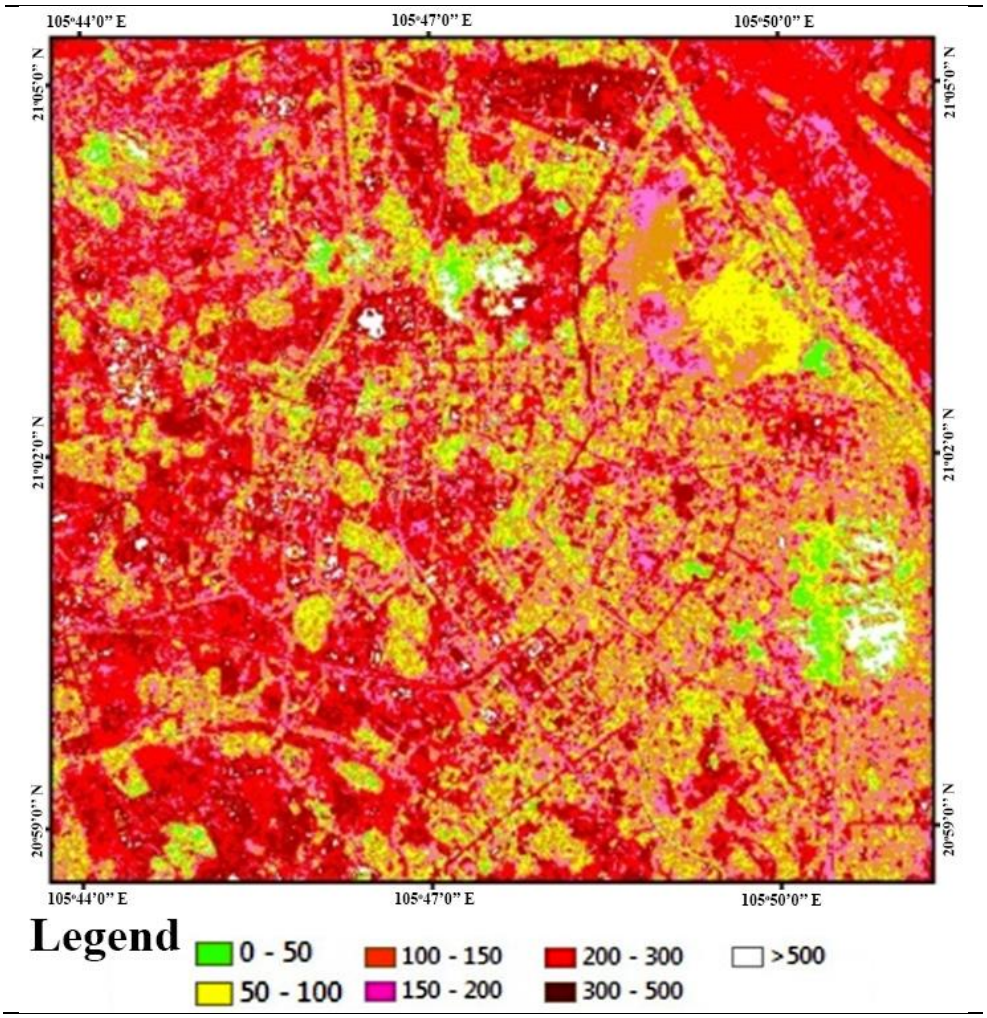


Figure 4. Map of PM_{10} distribution of traffic routes in inner-city areas of Hanoi, Vietnam, calculated from three bands of VNREDSat-1 imagery (color table is classified by Standard AQI-Air Quality Index, unit: $\mu g/m^3$.

The predicted spatial pattern of PM_{10} value is consistent with actual observations. The white color areas represent a larger PM_{10} (over 500 μg/m3) is an area mixed with clouds and roof areas made of corrugated iron, which causes sun glare. Red-brown areas (from 300 to 500 μg/m3) and red color areas (from 200 to 300 μg/m3). This is related to the fact that this area distributed large impervious surface density and crowded transportation infrastructure, mainly concentrated in areas with ongoing construction activities or high traffic density such as Pham Van Dong street, Hang Dau street, Minh Khai ward (Bac Tu Liem district),... These are also areas with high traffic density, ongoing construction, and upgrading of urban infrastructure, generating a large amount of dust into the environment. In contrast, the areas are represented by a pink color (from 150 to 200 $\mu g/m^3$), orange color (from 100 to 150 $\mu g/m^3$), and yellow (from 50 to 100 $\mu g/m^3$) have medium and low PM_{10} values, respectively. These are areas with low construction density and lots of greenery around.

This study tested linear regression models exploit the information from VNREDSat-1 to integrate their advantages for estimation of Particulate matter 10 micrometers. The regression model is used widely for PM_{10} estimation. However, it has some limitation, e.g. the accuracy is not always improved when supplementary variables are added because not all the variables are linearly related to PM_{10} (Table 3). VNREDSat-1 imagery data can give admissible results in the modeling of PM_{10} estimated in urban areas. The correlation between forest PM_{10} and spectral value from three of Blue, Green, and Red on

VNREDSat-1 image was analyzed and calculated for either the linear regression model. The results showed that the coefficient of determination (r^2=0.773) of the linear regression model is the optimal one for PM_{10} prediction.

4. CONCLUSION

The results of this study indicated that PM_{10} air pollution could be determined by using VNREDSat-1 satellite imagery data combined with PM_{10} measurement from ground stations. The remote sensing method has the advantage of being able to identify dust pollution on a large scale. The calculation of PM_{10} depends on the number of ground stations and the location of these stations. With fewer ground stations in Hanoi area, the mobile dust measurement devices can be used instead. However, at a particular time when a satellite is passing, simultaneous conducting the ground measurement at different points at the same time is a problem. Furthermore, based on the results of the PM_{10} distribution map, we found that the effects of clouds, atmosphere and objects on the surface for determining PM_{10} dust from the VNREDSat-1 satellite image were much. So our research direction is to study the methods of enhancing the quality of images to improve the accuracy of PM_{10} computation.

REFERENCES

Asmala Ahmad and MazlanHashim, (2002). Determination of haze using NOAA-14 AVHRR satellite data, [Online] available: http://www.gisdevelopment.net/aars/acrs/2002/czm/050.pdf

Lim HS, MatJafri MZ, Abdulla K, Mohd NS, Sultan AS, (2004) Remote Sensing of PM10 From Landsat TM Imagery.25th ACRS 2004 Chiang Mai, Thailand.

Luong C.K, Ho T.V.T, Tran N.T, NguyenL.D, (2010). Detecting Air Pollution In Vietnam By Optical Satellite Images, E-proceedings of The 31th Asian Conference on Remote Sensing (ACRS2010), 1-5 November 2010, Hanoi, Vietnam.

Ministry of Natural Resources and Environment, National environmental Status Report for 2011-2015.

Nadzri, O., Mohd, Z.M.J., Lim, H.S., 2010. Estimating Particulate MatterConcentration over Arid Region Using Satellite Remote Sensing: A CaseStudy in Makkah, Saudi Arabia. Modern applied Science 4: 131-142.

Nguyen, N.H., Tran, V.A., 2014, Estimation of PM10 from AOT of satellite LandSat 8 image over Ha Noi city, International Symposium on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences

Report of WHO, http://www.who.int/phe/health_topics/outdoorair/en//Landsat 8 Data User Handbook,2016.
https://landsat.usgs.gov/sites/default/files/documents/Landsat8DataUsersHandbook.pdf

Retalis A, Cartalis C, Athanassious E, 1999, Assessment of the distribution of aerosols in the area of Athens with the use of Landsat Thematic Mapper data. Int J Remote Sensing 20: 939-945.

Sam Appadurai.A and J.Colins JohnnyM.E,(2016), Satellite based estimation of pm10 from AOT of landsat 7ETM+ over Chennai city, International Journal of Advances in Engineering Research, Vol. No. 11

Sifakis N, Gkoufa A, Soulakellis N.,1998. Integrated Computational Assessment Via Remote Observation System.

Sifakis, N. &Deschamps, P.Y. (1992). Mapping of air pollution using SPOT satellite data, Photogrammetric Engineering & Remote Sensing, 58(10), 1433 – 1437

Tran T.V, Nguyen P.K, Hà D.X.B, 2014. (Remotely sensed Aerosol Optical Thichness determination to simulate PM10 distribution over urban area of Ho Chi Minh city), Journal of Science and Technology Development, Vol.30, No.2 ,p52-62

Tran T.V, Trịnh TB, Ha D.X.B, 2012. (Study of dust pollution detecting ability in urban areas by remote sensing technology to support air environment observation), Journal of Science and Technology Development, Vol. 16, No.2,

Tran X.T, VươngT.K, NguyenV.M 2013, Develop a program to monitor mine air pollution from satellite imagery data, Journal of Mining Industry, Vol.26

Ung, A., Wald, L., Ranchin, T., Weber, C., Hirsch, J., Perron, G. and Kleinpeter, J., 2001b. , Satellite data for Air Pollution Mapping Over A City-ꝸVirtual Stations, Proceeding of the 21th EARSeL

Symposium, Observing Our Environment From Space: New Solutions For A New Millenium, Paris, France, 14 - 16 May 2001, Gerard Begni Editor, A., A., Balkema, Lisse, Abingdon, Exton (PA), Tokyo, pp. 147 – 151

Ung, A., Weber, C., Perron, G., Hirsch, J., Kleinpeter, J., Wald, L. and Ranchin, T., 2001a. Air Pollution Mapping Over A City - Virtual Stations And Morphological Indicators, Proceedings of 10th International Symposium "Transport and Air Pollution" September 17 - 19, 2001 - Boulder, Colorado USA.

Wald L, Baleynaud JM, 1999, Observed air quality over city of Nantes by means of Landsat Thermal Infrared data. International journal of Remote Sensing 20: 947-959.

Weber, C., Hirsch, J., Perron, G., Kleinpeter, J., Ranchin, T., Ung, A. and Wald, L., 2001, Urban Morphology, Remote Sensing and Pollutants Distribution: An Application To The City of Strasbourg, France, International Union of Air Pollution Prevention and Environmental Protection Associations (lUAPPA) Symposium and Korean Society for Atmospheric Environment (KOSAE) Symposium, 12th World Clean Air & Environment Congress, Greening the New Millennium, 26 - 31 August 2001, Seoul, Korea.

Zhang, L., S. Hu, and H. Yang, The Effects of Solar Irradiance Spectra on Calculation of Narrow Band Top-of-Atmosphere Reflectance. APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, 2014. 7(1): p. 49-58