Carpathian Journal of Earth and Environmental Sciences August 2022, Vol. 17, No. 2, p. 275 – 286; DOI:10.26471/cjees/2022/017/221

HIGH-RESOLUTION SPATIAL SOLUTION FOR AIR QUALITY ASSESSMENT IN THE TROPICAL CITY OF HO CHI MINH CITY, VIETNAM

Van TRAN THI[1,2*], Bao HA DUONG XUAN[1,2], Bao VO QUOC[1,2], Phuong DINH THI KIM[1,2] & Hieu NGUYEN TRINH DUC[1,2]

[1]Department of Environmental and Resources Information System, Faculty of Environment and Natural Resources, Ho Chi Minh City University of Technology, (HCMUT), 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam; ∗Corresponding author: tranthivankt@hcmut.edu.vn

[2]Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Ho Chi Minh City, Vietnam

Abstract: Air pollution is the leading environmental agent that poses a severe threat to human health and is one of the most severe problems in cities. Of the pollutants, particulate matter (PM), especially particles less than 2.5 microns in diameter, has the most profound health impacts. Urbanization and industrialization in cities have made the air quality up there worse and worse. Our study was based on the Aerosol Optical Depth (AOD) feature, a parameter obtained by remote sensing that relates to the presence of airborne particles potentially associated with PM. In this study, the $PM_{2.5}$ concentration data from the ground monitoring station and the MODIS AOD product of 3 km resolution were correlated to build a suitable regression function to simulate the spatial distribution of $PM_{2.5}$ concentrations. Next, the AOD was retrieved from the Landsat image based on the characteristics of the decrease in atmospheric clarity caused by the pollution particles. Landsat AOD has a 600m higher resolution than MODIS AOD. Research results on air quality (AQ) were simulated on Landsat AOD image through $PM_{2.5}$ concentration distribution and air quality index (AQI), in which AQI was determined based on USEPA standards. The analysis shows that the linear regression function between $PM_{2.5}$ concentration and MODIS AOD correlated best with the correlation coefficient R=0.9. Then $PM_{2.5}$ distribution was established on Landsat AOD image with higher spatial resolution. Case analysis for March of 2018 reflected that the average concentration of $PM_{2.5}$ across Ho Chi Minh City (HCMC) was higher than the allowable threshold specified in QCVN05:2013/BTNMT. $PM_{2.5}$ concentration in central districts tended to be higher than in suburban districts. The study also found that the city average AQI-$PM_{2.5}$ was around 97.38, peaking at 159, which was in the Unhealthy range, especially for sensitive groups. The result of the study provides potential solutions for AQ monitoring at the city level with a detailed spatial distribution.

Keywords: air pollution; AOD; AQI; high spatial resolution; $PM_{2.5}$; regression analysis

1. INTRODUCTION

Air pollution is now a problem in developing countries. It causes respiratory and chronic diseases in people living in polluted air and affects crops and forests. Human activities and processes in the natural environment are sources of air pollution. Many factors cause air pollution dispersion, such as meteorological conditions combined with wind flow from traffic activities, topography, and geomorphology of the area due to construction works. Therefore, air pollution is often very spatially variable, even at short distances. Aerosols' air pollutants can be SO_2, NOx, CO… or Particulate Matter (PM). $PM_{2.5}$ are suspended dust particles with a diameter of 2.5µm or smaller, which can easily penetrate human cells, destroying self-protection and immune mechanisms from within the cells. $PM_{2.5}$ is an environmental pollutant and the causative agent of many dangerous diseases to human health. In addition to causing a series of acute diseases, they are also toxic to essential organs such as the lungs, heart, brain, etc. Compared with PM_{10}, $PM_{2.5}$ fine dust is much more pathogenic. PM_{10} dust particles can only enter the nose and respiratory system and can be coughed up by themselves to reduce the level of damage. However, smaller particles can enter the trachea and bronchi, even the end of the bronchi and blood, significantly affecting health (Linh, 2017). Therefore, AQ has a significant impact on human health, and it is necessary to monitor it to ensure a better quality of life for people, especially urban residents.

Accelerated industrialization and modernization have faced severe environmental challenges in developing and underdeveloped countries, especially large cities. The ongoing operation of outdated technology factories, increased traffic density during peak hours, indiscriminate burning activities, unshielded constructions, and accompanied unreasonable economic development policies from state agencies are the direct and indirect causes leading to the generation of a large amount of $PM_{2.5}$ in the atmosphere. In large cities, to control the concentration of $PM_{2.5}$ in the atmosphere, many rapid monitoring stations are installed in critical areas. Communities can view dust and air condition data displayed on separate websites or through apps downloaded with smartphones. Based on the monitoring data set, people can know the pollution status in the living area, and managers can easily control the fluctuation of $PM_{2.5}$ dust.

Besides ground-based monitoring, air pollutants can be monitored using remote sensing technology. Aerosol optical depth (AOD) is a parameter obtained by remote sensing that relates to the presence of airborne particles potentially associated with PMs. Optical depth is a measure of the radiation transmission of a vertical column of air per unit cross-sectional area. The high optical depth means less radiation transmission through the atmosphere. Optical depth is the result of the combined effect of vertical scattering and absorption, caused mainly by aerosols and air molecules. Optical depth due to aerosol is called aerosol optical depth (AOD) (Satheesh, 2002). Moderate Resolution Imaging Spectroradiometer (MODIS) provides a daily AOD product, a valuable data source to support AQ monitoring spatially distributed in the study area. Previous studies have shown that since the 1990s, there have been studies to determine AOD from satellite images Landsat TM/ETM+, ASTER, SPOT, ALOS, IRS, MODIS (Kaufman et al., 1990; Sifakis & Deschamps, 1992; Retalis, 1998; Retalis et al., 1999; Hadjimitsis & Clayton, 2009; Hadjimitsis, 2009a; Hadjimitsis, 2009b). Based on the AOD values from the images, some studies have examined the correlation between them and $PM_{2.5}$. Wijeratne (2003) found a correlation between AOD calculated based on Landsat 7/ETM+ satellite images and pollutants, including PM_{10}, CO, NO, NO_2, SO_2, NH_3, O_3, and BP. From there, a distribution map of pollutants was established throughout the Netherlands. Kumar (2007) studied using a grid with 1×1.5 km^2 spread throughout the study area. 113 $PM_{2.5}$ dust samples were collected at different locations and times. Then it is considered to correlate with AOD extracted from MODIS images (5-km resolution). Research results have shown a positive correlation between MODIS-AOD and $PM_{2.5}$. When AOD changes 1%, $PM_{2.5}$ measured in 45 and 150 minutes changes 0.52±0.202% and 0.39±0.15%, respectively. Studies have also been estimating $PM_{2.5}$ concerning AOD on high-resolution data, allowing better discernment of daily and long-term impacts of fine dust for urban, suburban, and rural areas, demonstrating applicability for city-scale (Lee et al., 2011; Kloog & Alexandra, 2014).

Until the end of 2018, throughout HCMC, AQ monitoring is conducted at 15 locations with the semi-automatic operation. The sampling frequency is only 10 days per month. The reflection of the citywide AQ situation is mainly based on monitoring measurements from semi-automatic stations. Data from ground stations can observe continuously and many times a day. However, they have the limitation that they incorrectly reflect the spatial distribution of the whole area. Besides, historical data for monitoring concentrations of pollutants observed by MONRE are not publicly available, making it difficult for many analytical studies (Thu & Blume, 2018). Since 2016, the monitoring station at the United States (US) consulate has been installed to continuously measure hourly results, including $PM_{2.5}$ and AQI-$PM_{2.5}$ dust measurements. Therefore, this study focuses on the problem of $PM_{2.5}$ dust and the air quality index (AQI) in HCMC, using data from the US Consulate monitoring station for analysis. However, there is only one monitoring point in the whole city, and it is impossible to evaluate the spatial distribution of $PM_{2.5}$ concentrations and AQI throughout the area. Remote sensing technology provides satellite images, which allow various physical features to be monitored simultaneously over vast areas, with an observational density proportionate to the sensor spatial resolution (Sifakis & Deschamps, 1992). On the other hand, remote sensing is also a modern, popular technology with a reasonable cost and can be extracted from many different sources. Remote sensing techniques can build a map of the spatial distribution of $PM_{2.5}$ concentration. Thereby assessing AQ to warn of the impact of dust on human health and provide an overview of the pollution status in the whole area, helping environmental managers have a basis for proposing solutions to reduce fine dust concentration, contributing to the improvement of AQ in HCMC.

2. DATA AND METHODS

2.1 Study area

HCMC is the largest city in Vietnam, with a natural area of about 2095 km^2 (Fig. 1). Located in the tropical region affected by monsoon, HCMC has a hot and humid climate with only two seasons: dry and rainy. Rainfall is high, and the average/year is 1,949 mm. About 90% of annual rainfall is concentrated in the rainy season months from May to November, in which June and September usually have the highest rainfall. On the spatial scale of the city, the rainfall is unevenly distributed and tends to increase gradually along the southwest-northeast axis. Most urban districts and northern suburban areas usually receive higher rainfall than districts in the south and southwest of the city.

Regarding wind, HCMC is influenced by two main wind directions: the west-southwest and north-northeast monsoons. West-southwest wind from the Indian Ocean blows in during the rainy season, from June to October. The north-northeast wind blows from the East Sea during the dry season from November to February. In addition, there is a trade wind from the south-southeast direction from March to May. In general, HCMC is in a storm-free area. The average air temperature is 27˚C with abundant radiation. There are over 330 days every year with an average temperature of 25-28˚C. HCMC has a well-developed network of rivers and canals located downstream of the Dong Nai-Saigon River system. Regarding hydrology, most of HCMC's rivers and canals are influenced by the semi-diurnal tidal fluctuations of the East Sea. Every day, the water rises and falls twice, whereby the tide penetrates deeply into the canals in the city, causing a significant impact on agricultural production and limiting water drainage in the inner city.

The soil in HCMC is formed on two sedimentary facies: Pleistocene and Holocene. Pleistocene sediments (ancient alluvial deposits): occupy most of the north, northwest, and northeast of the city, including most of Cu Chi, Hoc Mon, Bac Binh Chanh, Thu Duc, north-northeast of District 9, and most of the old inner city. Holocene sediments (young alluvial deposits): in HCMC, this sediment has many origins, including coastal areas, bays,

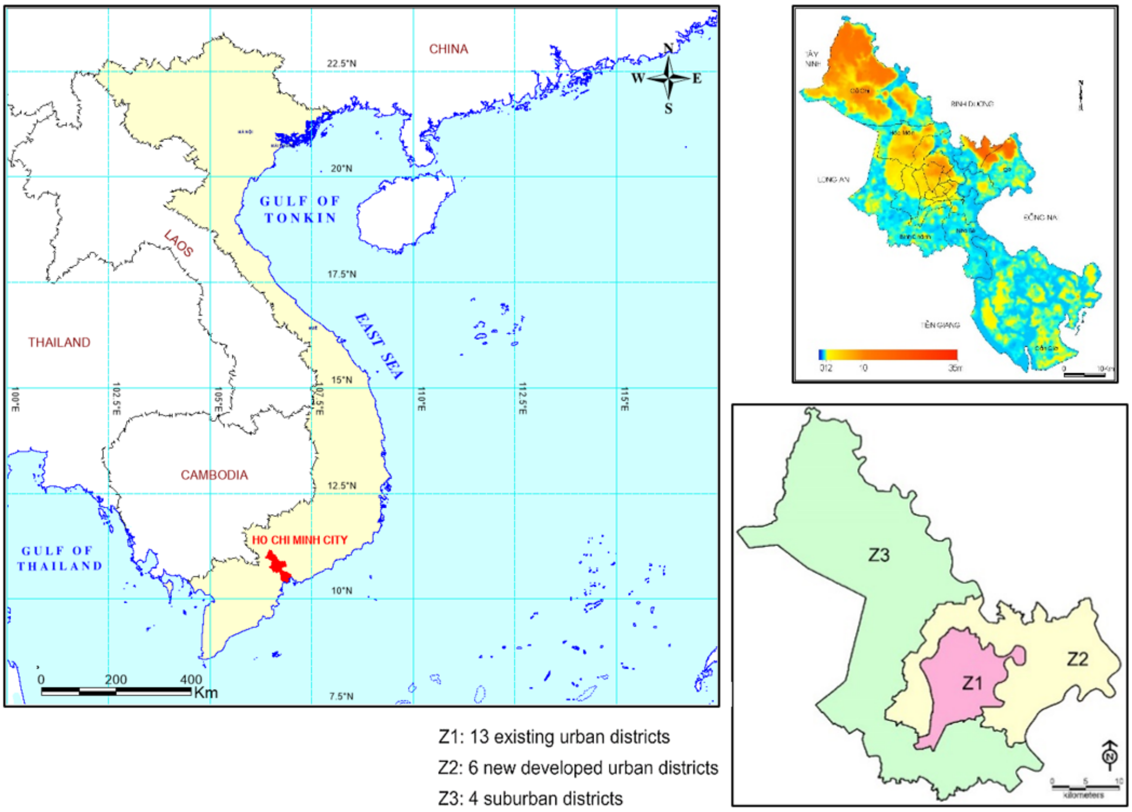


Figure 1. The extent of the study area on the right is shown in the topography map of the whole HCMC, which was the red point in the Vietnam map on the left. This study area is divided into three administrative zones Z1, Z2, and Z3, for analysis.

rivers, riverbed alluvial and alluvial flats, etc., so it has formed many different types of soil such as groups of alluvial soil, groups of alkaline soil and saline soil. HCMC is located in the transition zone between the Southeast region and the Mekong Delta. The general topography has a lower shape from North to South and East to West. It can be divided into three topographical sub-regions. The highland is located in the north-northeast and part of the northwest (in the north of Cu Chi district, northeast of Thu Duc district, and District 9), with wavy topography, the average altitude of 10-25 m, and interspersed with hills with the highest elevation up to 32m, such as Long Binh hill (District 9). Low-lying areas in the South-Southwest and Southeast of the city (in districts 9, 8, 7, and suburban districts of Binh Chanh, Nha Be, and Can Gio). This area has an average height of about 1m; the highest is 2m, and the lowest is 0.5m. The medium zone, distributed in the city center, includes most of the old inner city, a part of District 2, Thu Duc, District 12, and Hoc Mon district. This area has an average elevation of 5-10m. In general, the topography of HCMC is not complicated, but it is also quite diverse, with conditions for multi-faceted development.

The city has a significant average annual population growth rate (Giang, 2015). Along with the population growth, the demand for transportation also increases. By 2017, the city is managing more than 8 million motorbikes, the highest number of motorbikes in the country (Phu, 2017). HCMC has an abundant labor force due to its growing industries. As of 2015, HCMC has 3 Export Processing Zones and 13 Industrial Parks in the inner and outer suburbs. According to the planning and development, by 2020, HCMC will have 23 export processing and industrial zones (HEPZA, 2017). In addition, the number of vehicles is also increasing. All have contributed to releasing a large amount of gas into the atmosphere, causing the $PM_{2.5}$ concentration to increase rapidly year by year. From 2017 to 2018, the $PM_{2.5}$ concentration level rose from 23.6 to 26.9 µg/m³ (Dang, 2020; Tra, 2020). These levels are more than double those recommended by the World Health Organisation (WHO) as an annual guideline (IQAIR, 2020). The causes of air pollution in HCMC are mainly from 3 sources: traffic, industrial production, and construction activities (Thu & Blume, 2018). Emissions from these sources generate many toxic compounds into the atmosphere, including fine dust $PM_{2.5}$.

2.2 Data

The observed data is $PM_{2.5}$ concentration. Ground monitoring data $PM_{2.5}$ was extracted at two stations in the HCMC area. (1) Station at the US consulate at coordinate 10°46'59.9" N 106°42'03.2" E was installed and put into operation from 2016 until now. It is an hourly automated station that records $PM_{2.5}$ and AQI-$PM_{2.5}$ measurements. (2) The Environmental Source Samplers company station at coordinate 10°48'54.5" N 106°43'11.0" E was installed and implemented in 2016. However, at the beginning of 2018, the company moved its location out, so the equipment stopped monitoring HCMC. So, we focused our review on data for 2016-2018. Ground data should be collected on the same day as the images are acquired. Besides, a good relationship between images and ground data can be expected if ground data is available when the satellite is overpassed.

The Terra/Aqua MODIS daily captures the earth's surface. MODIS AOD image extracted from MOD04_3K product of the MODIS C6.1 dataset has a resolution of 3 km with 7 different bands in wavelengths from 0.47 µm to 2.13 µm. The atmospheric scattering and absorption causing smog on satellite images only occur at visible and near-infrared wavelengths (Gustavo et al., 2011). In addition, previous studies (Sifakis & Deschamps, 1992; Hadjimitsis & Clayton, 2009; Van et al., 2012; Van et al., 2014) on the distribution of PM_{10} on Landsat and Spot 5 images have shown that the suspended particles scatter well with light at visible wavelengths. Therefore, this study focused on AOD at three wavelengths 0.47 µm, 0.55 µm, and 0.66 µm. The collected dataset was divided into two parts: training data to establish the model (50 values in 2016-2018) and testing data to check the reliability and accuracy of the model (25 values in 2020). MODIS products are sinusoidal projections, so MOD04_3k must be reprojected as WGS84 UTM 48N for HCMC using MODIS Conversion Toolkit (MCTK).

Landsat images have a higher resolution of 30m for reflectance bands. For calculating the AOD, it is necessary to select Landsat images of two days: one under clear atmospheric conditions and one under polluted conditions. The selected clean day must be on a clear, pollution-free day with no clouds in the entire image. We chose the period when the Vietnam industry, especially heavy industry, was not developed, and the traffic was not dense. The Landsat 5 image acquired on 12 February 1996 was selected as the clean day based on the above conditions. The selected pollution image must coincide with the date and time acquired with the MODIS AOD image to perform validation and is also a case image for analyzing the AQ. Landsat 7 image acquired on 13 March 2018 is appropriate. In 2003, the Scan-Line

Corrector (SLC) on Landsat 7 with an ETM+ sensor failed, with black strip lines restricting around 22% of the pixels per scene from being scanned (Maxwell et al., 2007; Chen et al., 2011; Yin et al., 2016). So, the black strip lines were gap-filled to fix the SLC error before further processing.

2.3 Method

This study needs to create detailed spatial distribution patterns of AQ over the city. Therefore, the approach of the study is to integrate ground observation data with satellite data to achieve the following objectives: (1) Build a regression function to estimate $PM_{2.5}$ based on ground observation and daily available MODIS AOD data with 3km-resolution; (2) Estimate AOD from Landsat image with higher resolution than MODIS AOD image; (3) Simulate the spatial patterns of $PM_{2.5}$ and AQI for AQ assessment over the city.

2.3.1 Statistical analysis for building regression function

Statistical analysis was conducted based on ground $PM_{2.5}$ and MODIS AOD variables. It is necessary to determine their relationship to simulate the spatial distribution of $PM_{2.5}$ concentration throughout the study area. Pearson analysis was conducted to check the correlation between $PM_{2.5}$ concentration and AOD values at the three different wavelengths 0.47, 0.55, and 0.66 µm. The Pearson coefficient (R) ranges from -1 to +1. The closer R is to +1 or -1, the closer AOD and $PM_{2.5}$ are related. Based on the Pearson coefficient of correlation analysis, any wavelengths that show AOD was not correlated with $PM_{2.5}$ were removed. Conversely, the remaining AODs with a high correlation with $PM_{2.5}$ were applied to build the least-squares regression model. Then, we examined the test with different regression models to find the best-fit function. All steps were conducted by statistical software SPSS. In SPSS analysis, selecting the best model that mathematically describes the relationship between two variables is based on three criteria. The Sig Anova and Sig coefficient (or p-value) results must be less than 5% (significant level) (Trong & Ngoc, 2008). It means that the null hypothesis can be discarded and that there exists an association between the two variables. After that, R^2 was the final criterion used to determine the best model.

The error between the estimated and valid (measured) values was evaluated using some evaluation indices, including the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RSME). The formula for each index is as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (6)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_t\right| \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_t)} \quad (8)$$

where y_i, $\hat{y}_t$ are estimated and valid values, respectively, and n is the number of samples. SS_{res} is the error between the regression data and mean value, SS_{tot} is the error between the actual data and the mean value, referring to the mean value of the valid value. Valid values in this study are the ground measured $PM_{2.5}$ and the available MODIS AOD 3km.

2.3.2 Landsat AOD retrieval

Landsat AOD retrieval was based on the theory developed by Sifakis & Deschamps (1992). In the visible spectrum of electromagnetic waves, the scattering of shortwave radiation by air particles (as polluted particles) causes a decrease in contrast and distortion of the spectral response patterns in satellite images. This effect can be estimated using the AOD derived from at-satellite apparent reflectance Eq. (1). Sifakis & Deschamps (1992) assumed that the standard deviation expresses the analogical contrast in images for urban targets. The relationship between the standard deviation of apparent reflectance $\sigma(\rho^*)$ and the standard deviation of real reflectance $\sigma(\rho)$ was derived using Eq. (2).

$$\rho^* = \rho\frac{T(\theta_s)T(\theta_v)}{1-\rho S} + \rho_a \quad (1)$$

$$\sigma(\rho^*) = \sigma(\rho)\frac{T(\theta_s)t_{dir}(\theta_v)}{1-\rho S} \quad (2)$$

where ρ^* is the apparent reflectance; ρ is the intrinsic surface reflectance; ρ_a is the atmospheric or sky reflectance; S is the spherical albedo of the atmosphere, defined as the ratio of the scattered to the total attenuated radiation (i.e., absorbed + scattered); θ_s is the solar zenith angle; θ_v is the zenithal observation angle; $T(\theta_s)$ and $T(\theta_v)$ are the total transmission function on the downwelling path and the upwelling path, respectively; $t_{dir}(\theta_v)$ is the direct transmission functions.

Using Lambert-Bouguer's transmission law as the Eq. (3) for the Eq. (2), we obtain the Eq. (4), which can be applied to reference (clear) day and a polluted day, respectively, with the AOD denoted as τ_1 and τ_2.

$$t_{dir}(\theta_v)=\exp(-k_\lambda m) \quad (3)$$

$$\sigma(\rho^*)=\sigma(\rho)\frac{T(\theta_s)\exp(-\tau/\cos(\theta_v))}{1-\rho S} \quad (4)$$

where $-k_\lambda m$ is AOD (dimensionless) denoted as τ (Iqbal 1983).

Landsat is nadir looking satellite, and the zenith angle can be assumed 0^o. So, Eq. (4) can be applied to the reference clear and real polluted images as the Eq. (5) as follows:

$$\Delta\tau=\tau_2-\tau_1=\ln\left[\frac{\sigma_1(\rho)}{\sigma_2(\rho)}\right] \quad (5)$$

where $\Delta\tau$ is the AOD; σ_1 (ρ) and σ_2 (ρ) are the standard deviation of apparent reflectance on a clear and polluted day, respectively. With a clean day AOD approximately equal to 0 due to no or minimal pollution components, then $\tau_1 = 0$, only the component τ_2 remained in Eq. (5), which is also the AOD on the polluted day image.

This equation uses the standard deviation of apparent reflectance to calculate AOD. Therefore, defining the grid to calculate on the Landsat image pixel is necessary. The basis of the selection of grid cells is as follows (Sifakis & Paronis, 1998): (1) to include some visible ground structure, it needs to be large enough; (2) Sufficiently small to allow consideration of a homogeneous atmosphere inside the grid cells. We tested with (20 x 20) pixel grid cells (at 30 m resolution) and found it suitable. At this point, one grid cell represents (600 x 600) m^2 on the ground. First, the Landsat images need to go through the image preprocessing steps to convert the digital number (DN) to reflectance values as instructed by the Data Users Handbook for Landsat 5 and 7 (Chander & Markham, 2003; Ihlen, 2019). Next, AOD was calculated using Eq. (5).

2.3.3 Air quality analysis

$PM_{2.5}$ is a toxic ingredient that affects public health. It is necessary to show $PM_{2.5}$ in AQI (so-called AQI-$PM_{2.5}$) to inform and warn the community about AQ and care for health. In this study, we applied the EPA Air Quality Guide for Particle Pollution (USEPA, 2018) to determine the AQI-$PM_{2.5}$ (for brevity, we sometimes use AQI). The AQI measure is specified in the range of 0 - 500. The higher the AQI index, the greater the pollution level and potential health effects. The AQI is classified into six categories, from Good to Hazardous (Table 1). The corresponding values of each parameter are looked up in Table 2. The formula for calculating AQI-$PM_{2.5}$ is as follows:

$$\mathrm{AQI}=\frac{(\mathrm{AQI_{Hi}})-(\mathrm{AQI_{Lo}})}{(\mathrm{C_{Hi}})-(\mathrm{C_{Lo}})}\times(\mathrm{C_i}-\mathrm{C_{Lo}})+(\mathrm{AQI_{Lo}}) \quad (9)$$

where C_i is the truncated concentration of pollutant $PM_{2.5}$; C_{Lo} is the concentration breakpoint that is less than or equal to $Conc_i$; C_{Hi} is the concentration breakpoint that is greater than or equal to $Conc_i$; AQI_{Lo} is the AQI value corresponding to C_{Lo}; AQI_{Hi} = the AQI value corresponding to C_{Hi}

3. RESULTS AND DISCUSSION

Terra satellite flies over HCMC at about 10:00 - 11:00 am, so $PM_{2.5}$ observation data from ground measuring stations and AOD value extracted from MODIS images were also concurrently extracted. The statistical analysis included this data pair to build a suitable regression function for simulating the $PM_{2.5}$

Table 1. AQI Basics for Particle Pollution (USEPA, 2018)

Colour	Levels of Concern	Values of Index	Description of Air Quality
Green	Good	0 to 50	Air quality is satisfactory, and air pollution poses little or no risk
Yellow	Moderate	51 to 100	Air quality is acceptable. However, there may be a risk for some people, particularly those who are unusually sensitive to air pollution
Orange	Unhealthy for Sensitive Groups	101 to 150	Members of sensitive groups may experience health effects. The general public is not likely to be affected
Red	Unhealthy	151 to 200	Some members of the general public may experience health effects; members of sensitive groups may experience more serious health effects
Purple	Very Unhealthy	201 to 300	Health alert: The risk of health effects is increased for everyone
Maroon	Hazardous	301 to 500	Health warning of emergency conditions: everyone is more likely to be affected

Table 2. Breakpoints for the AQI-$PM_{2.5}$ (USEPA, 2018)

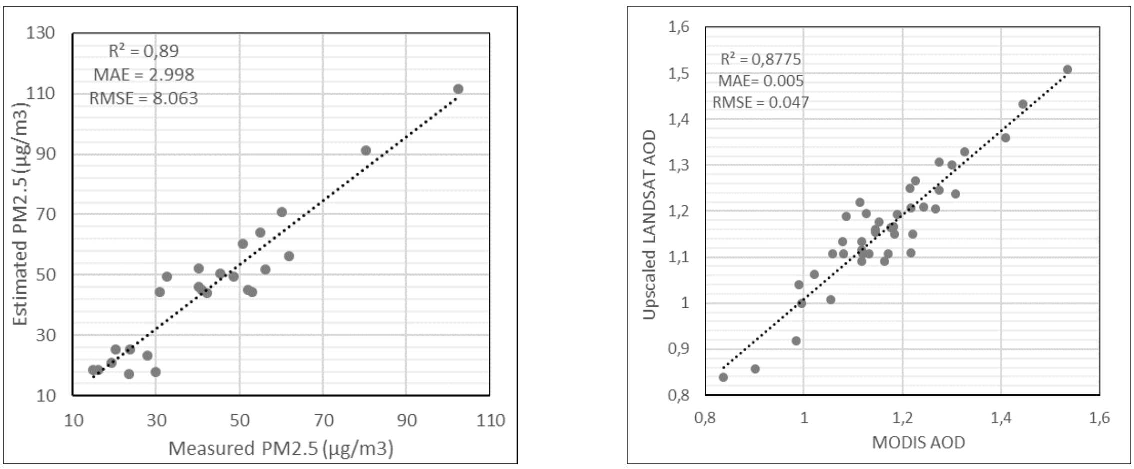
AQI-$PM_{2.5}$ Categories	C_{Lo} (µg/m^3)	C_{Hi} (µg/m^3)	AQI_{Lo}	AQI_{Hi}
Good	0.0	12.0	0	50
Moderate	12.1	35.4	51	100
Unhealthy for Sensitive Groups	35.5	55.4	101	150
Unhealthy	55.5	150.4	151	200
Very Unhealthy	150.5	250.4	201	300
Hazardous	250.5	500.4	301	500

distribution. This study analyzed the correlations between MODIS AOD and $PM_{2.5}$ concentration on three wavelengths 0.47 µm, 0.55 µm, and 0.66 µm. The Pearson correlation analysis generally demonstrated positive relationships between $PM_{2.5}$ and AODs at three wavelengths with p=0.000. However, the $PM_{2.5}$ concentration and AOD at a wavelength of 0.55 µm indicated the most significant correlation at the 0.01 level (2-tailed) with R=0.900, compared to at wavelength 0.47µm and 0.66 µm showed lower correlation levels of R=0.870 and 0.886, respectively. Therefore, we chose the AOD at 0.55 µm to build the regression model for $PM_{2.5}$. Different regression models were conducted. It told that all models were suitable to use at a confidence of 95% because of all Sig. ANOVA results were less than 5%. However, the Sig. Coefficients of the quadratic and cubic models were more significant than 5%, so these models were rejected. The inverse had the lowest R^2 (0.608) among the remaining models, while linear had the highest value (0.810). Therefore, the linear model was selected to build the regression model given by Eq. (10).

$$PM_{2.5} = 36.684 \times AOD + 6.839 \qquad (10)$$

We performed two validation processes for the results, as shown in Figures 2(a) and 2(b). Firstly, we validated the method to calculate $PM_{2.5}$ from the regression function built between ground observed $PM_{2.5}$ and MODIS AOD. The number of observed samples was 25 extracted in the dry season of 2020. Validation showed that the estimated $PM_{2.5}$ from MODIS AOD and ground measurements had good consistency, with a determination coefficient R^2 of 0.89 and MAE and RMSE of 2.998 µg/m^3, 8.063 µg/m^3, respectively. The estimated $PM_{2.5}$ values were slightly lower than the ground measurements and had high fitting accuracy. The comparison results show the ability of the regression model to estimate the $PM_{2.5}$ concentration in this study. Secondly, we validated the retrieved Landsat AOD with the available MODIS AOD by the green band at a wavelength of 0.55 µm. The 600m-resolution Landsat AOD was upscaled to 3 km by the MODIS AOD resolution. A total of 42 pairs of match-up data were obtained. The comparison validation indicated that the AODs retrieved from the Landsat and MOD04_3K AOD products displayed a relatively good correlation, with R^2, MAE, and RMSE of 0.877, 0.005, and 0.047, respectively.

Our study is an evaluation based on Landsat image data with higher resolution than MODIS images. However, the flight period of the Landsat satellite (16 days) is longer than that of the Aqua/MODIS and Terra/MODIS satellites (per day). Therefore, constructing the regression function to simulate the distribution must go through the MODIS image to ensure that the number of samples



(a) (b)

Figure 2. Validation results: (a) Comparison of the estimated and valid value of $PM_{2.5}$ from MODIS AOD, (b) upscaled Landsat AOD with the MODIS AOD at a wavelength of 0.55 µm

selected for the regression function is large enough. We assume that any satellite imaging of the surface captures the same surface state over the same area, on the same day, and simultaneously. Though the resolutions are different, they all reflect the same surface state. Furthermore, the above validation for the Landsat AOD retrieval case shows a good agreement compared with MODIS AOD data. So, this regression function can be applied to the Landsat AOD image.

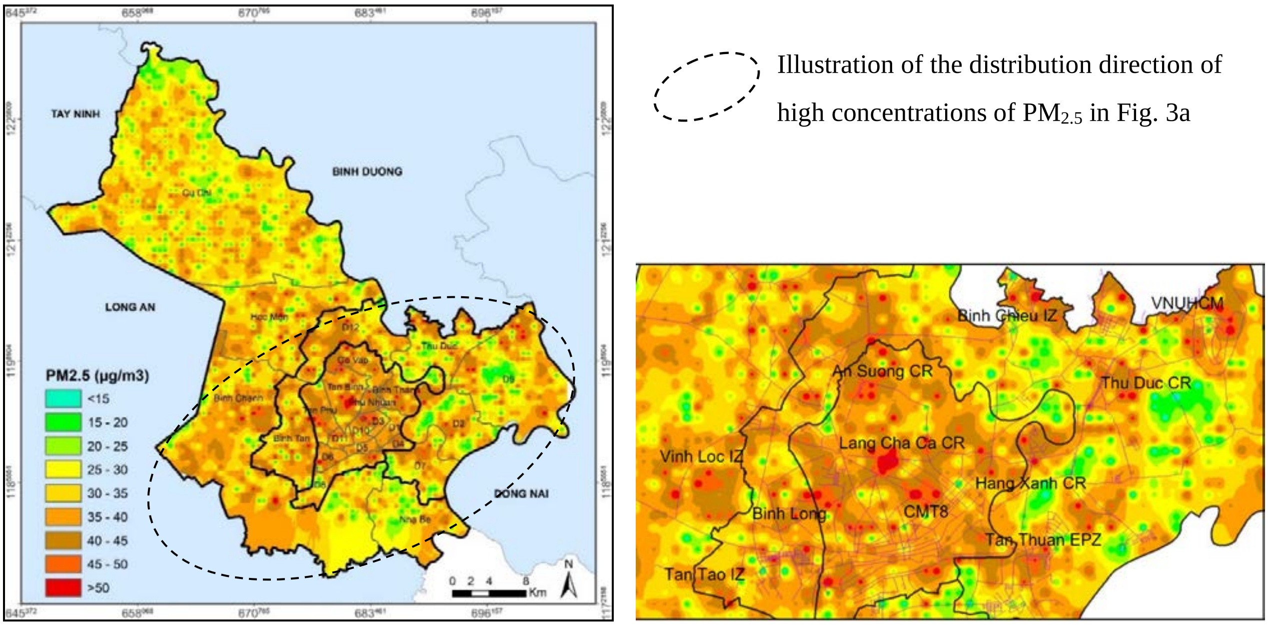
The regression function at Eq. (10) was used to simulate the distribution of $PM_{2.5}$ on the Landsat AOD image at the wavelength of green light acquired on 13 March 2018 as a case study, with the independent variable being the Landsat AOD pixels, the dependent variable being the concentration of $PM_{2.5}$. The $PM_{2.5}$ spatial distribution map for the HCMC area (Fig. 3) was established at around 10:00 am as Landsat image acquisition time. It is the time when factories have started operating, traffic density is high, and trucks are allowed to circulate in the inner city. In general, from the distribution map, it can be seen that $PM_{2.5}$ is represented by regions in the form of locality, not far away, and high concentrations focused in the city center. High-concentration $PM_{2.5}$ focused on the urban center (Z1) and stretched in the northeast-southwest direction (Z2), and gradually decreased when going north in the suburban (Z3). The average $PM_{2.5}$ for the entire northern area of HCMC has exceeded the permissible limit of QCVN 05:2013/BTNMT (33.47 $\mu g/m^3$ versus 25 $\mu g/m^3$) (Table 3). It is the National Technical Regulation on Ambient Air Quality. This regulation applies to monitoring and evaluating ambient air quality in Vietnam.

Zone Z1 in the city center was found with the highest mean concentration (38.17 $\mu g/m^3$), with the highest population density and heaviest traffic. Considering the particular locations in Figure 3(b), we detected hot spots with high concentrations of $PM_{2.5}$. The highest point is 71.22 $\mu g\ m^{-3}$ in the Cau Cho Dem area in the southwest (Binh Chanh district). In most crossroads (CR) and industrial zones (IZ), the $PM_{2.5}$ concentration was very high, exceeding the allowable threshold of QCVN 05:2013/BTNMT. In contrast, the $PM_{2.5}$ concentration still reached the allowable level in some places, such as Bau Chua hamlet in the north suburban Cu Chi district (9.95 $\mu g/m^3$), areas far from CRs and IZs such as parks, tourist areas (District 9), areas near rivers in the northwest of District 2, some new residential areas in Nha Be district. These were locations where industrial activities were limited, the population was sparse, and the air was clear. The $PM_{2.5}$ concentration was within the allowable range. The spatial pattern of $PM_{2.5}$ over HCMC is generally stretched in the northeast-southwest direction, resulting from the prevailing monsoon direction over the study area.

The AQI-$PM_{2.5}$ index was set up to notify the status of $PM_{2.5}$ fine dust and warn of the impact of dust on human health depending on different pollution levels. USEPA's 6-level AQI scale ranges from Good to Hazardous (Table 1). The distribution map of AQI-$PM_{2.5}$ in Figure 4 showed that on 13 March 2018, the whole northern HCMC area was in the range of 41-159 AQI, corresponding to the quality level from Good to Unhealthy (Table 4). The quality level of Moderate and Unhealthy for Sensitive Groups occupied the majority of over 90% of the city space. The spatial distribution of the AQI increased when entering the city center and stretched in a northeast-southwest direction. AQ level 3 - Unhealthy for Sensitive Groups occupied most space in the central area. At this level, the majority of the community has not been seriously affected. However, some sensitive groups, such as people with lung disease, the elderly, and children, are at high risk of illness when exposed to dust. Especially in area Z1 (mainly in Tan Binh district, districts 1 and 3), level 4 Unhealthy occupied the space area with the highest rate of 4.18% compared to the other two zones. The AQ level makes some members of the general public experience health effects, and members of sensitive groups may experience more severe health effects. Therefore, people who are easily irritated with pollutants, those with respiratory, lung, or cardiovascular diseases, should limit vigorous exercise, overexertion, or outside activities.

In contrast, HCMC suburban districts had better AQ, with the majority in the AQI range from Good to Moderate, notably in Cu Chi district in the last northern part of the city, District 9 in the east, and Nha Be district in the south. These areas have low population density and low traffic activity but a high density of trees compared to the city center. This result was relatively consistent with ground observations and announcements by authority bodies when assessing the causes of air pollution in urban areas (as opposed to suburban areas) (MONRE, 2018; DONRE, 2018).

Finally, we also considered the possible relationship between green space and $PM_{2.5}$ concentration in the study area. Trees are a factor that prevents the diffusion of $PM_{2.5}$ dust in the air from reducing the risk of harmful effects that $PM_{2.5}$ dust causes to human health. To better understand the current land cover status of HCMC, in which vegetation covers represent trees, thereby orienting



(a) (b)

Figure 3. (a) Spatial patterns of $PM_{2.5}$ concentrations in northern HCMC on 13 March 2018; (b) Hotspot areas with high $PM_{2.5}$ concentration

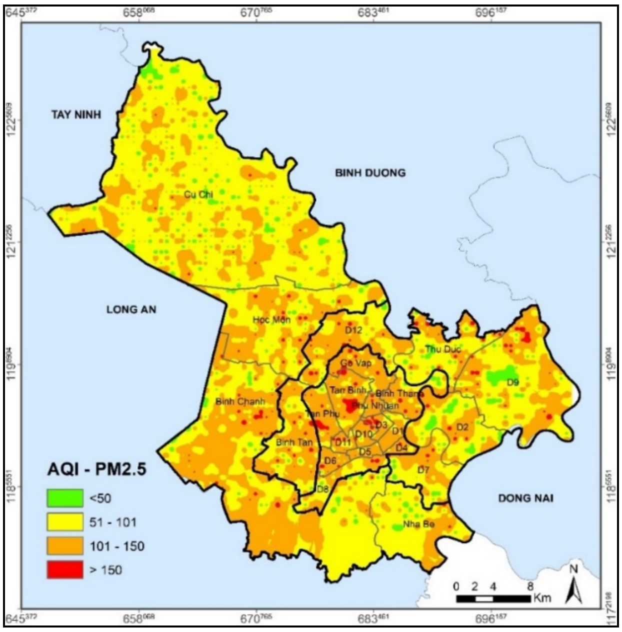


Figure 4. Spatial patterns of AQI-$PM_{2.5}$ in the northern HCMC area

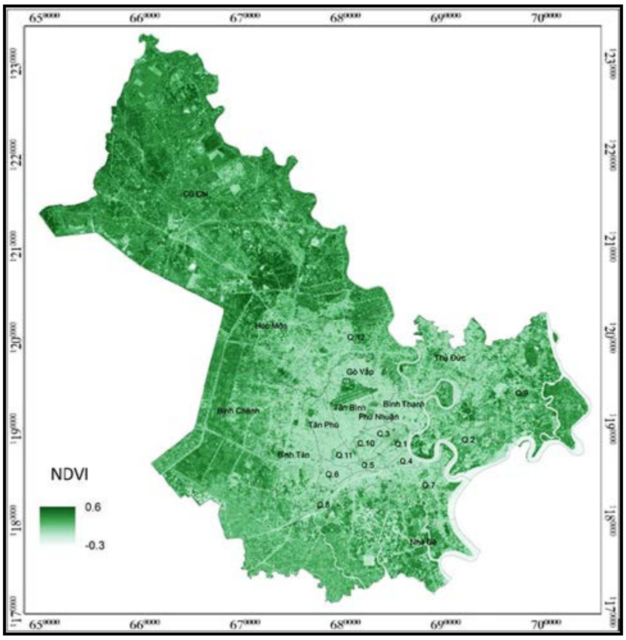


Figure 5. NDVI on 13 March 2018

Table 3. Statistics of $PM_{2.5}$ concentration (µg/m^3) and percentage (%) of AQI-$PM_{2.5}$ by administrative zone on 13 March 2018

Administrative zone	$PM_{2.5}$ (µg/m^3)			AQI-$PM_{2.5}$ (%)		
	Mean	Max	Min	Mean	Max	Min
Z1: 13 existing urban districts	38.17	65.19	10.26	108.29	156	43
Z2: 6 newly developed urban districts	33.83	62.17	10.02	98.45	154	42
Z3: 3 suburban districts	32.32	71.22	9.95	94.58	159	41
Z1 + Z2	35.19	65.19	10.02	101.53	156	42
Whole northern HCMC	33.47	71.22	9.95	97.38	159	41

Table 4. Percentage (%) of the spatial distribution of AQI levels by administrative zone on 13 March 2018

AQI level	USEPA rank	Z1	Z2	Z3
0-50	Good	0.70	2.86	2.79
51-100	Moderate	36.59	52.22	50.00
101-150	Unhealthy for Sensitive group	58.54	42.54	46.64
>150	Unhealthy	4.18	2.38	0.57

appropriate measures to reduce $PM_{2.5}$ concentration, we used the Normalized Difference Vegetation Index (NDVI) to survey the relationship with $PM_{2.5}$ concentration. NDVI is defined by the red and near-infrared wavelengths in the electromagnetic spectrum (NDVI = (NIR-RED)/(NIR+RED)) and represents the distribution density of vegetation and other ground objects (soil, water, urban). The Landsat image dated 13/03/2018 calculated the NDVI (Fig. 5). The NDVI values were divided into intervals of 0.2 in the range of -0.2 to 0.58 as statistics on the NDVI image. Correspondingly in each NDVI interval, we determined the average value of $PM_{2.5}$ concentration. We established a statistical relationship with NDVI as the independent variable and $PM_{2.5}$ concentration as the dependent variable. A survey of the suitable regression functions shows a linear relationship with the compactness coefficient R^2 = 0.6983. This correlation is relatively close.

Figure 6 shows that NDVI and $PM_{2.5}$ were negatively correlated. The higher the NDVI, the lower the $PM_{2.5}$ and vice versa, which means that in areas with dense vegetation cover, the $PM_{2.5}$ dust concentration is usually low. In contrast, in areas with very low NDVI (close to 0), such as bare land, roads, and urban areas, $PM_{2.5}$ concentrations are often high. Observe Figure 5, and it can be seen that, in the existing 13 urban districts (Z1), the NDVI index was very low with a mean and standard deviation of 0.10 and 0.07, respectively, because the impervious surface of the construction works was dominant in this area. It can partly explain the reason for the highest $PM_{2.5}$ concentration in the above area. In contrast, in suburban districts, especially Cu Chi and Nha Be districts in zone Z3, the vegetative cover was more abundant and played a dust filter role in reducing the $PM_{2.5}$ dispersed in the air. Trees are a factor that prevents the diffusion of $PM_{2.5}$ dust in the air, which reduces the risk of harmful effects that $PM_{2.5}$ dust causes to human health. These results proved the importance of the vegetation cover in urban.

In addition, climatic conditions also have an important influence on the dispersion and deposition of $PM_{2.5}$. Specifically, there are three possible scenarios: (1) Favorable meteorological conditions (for $PM_{2.5}$ deposition) including weak-medium winds, heavy rainfall, and high humidity; (2) Normal and favorable meteorological conditions (for $PM_{2.5}$ deposition) include weak-medium wind, cloudy rain, moderate-high humidity; (3) Adverse meteorological conditions (which enhance $PM_{2.5}$ emissions) including strong winds (Tuan et al., 2021). Besides, high temperature is also a favorable condition for enhancing $PM_{2.5}$ dispersion (Van et al., 2012; Van et al., 2014). As mentioned above, HCMC is influenced by two main wind directions: the west-southwest and north-northeast monsoons. The north-northeast wind blows in the dry season from November to February. In addition, there is a trade wind blowing in the south-southeast direction from March to May.

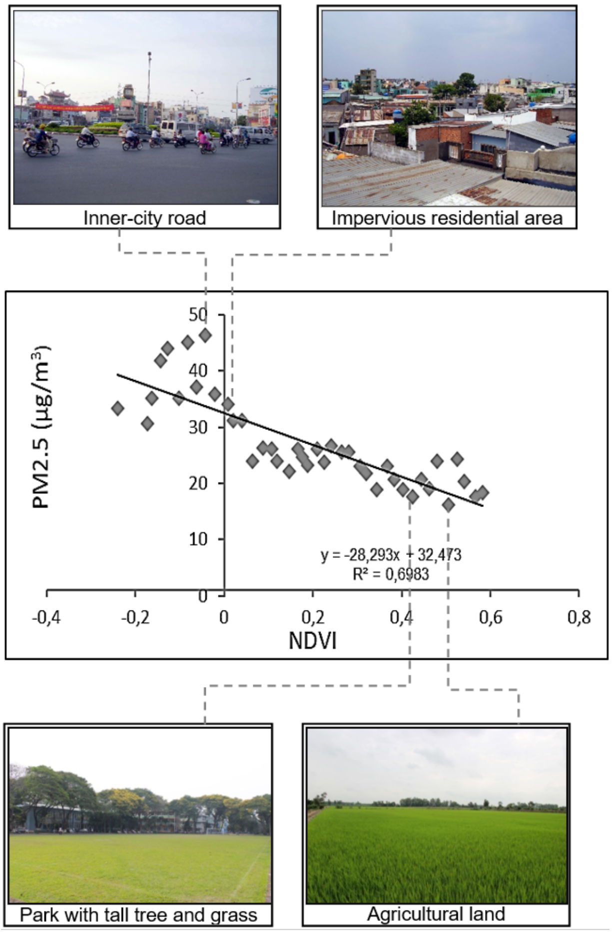


Figure 6. Statistical relationship between $PM_{2.5}$ and NDVI (some photographs representing the NDVI land cover with typical $PM_{2.5}$)

Besides that, the inner-city area is where the temperature is higher than the surrounding suburban areas due to the characteristics of the impervious surface of the building materials (Van, 2010). The time of this study was in March. In HCMC, the image of the high concentration distribution of $PM_{2.5}$ is concentrated in the inner city (where the temperature is high) and has a direction extending to the northeast-southwest and a little bit to the south (in line with the prevailing wind direction) (Fig. 3a).

From there, HCMC needs to pay attention to improving the AQ over the city, such as:

- Enhancing green space, especially in the inner-city area (can refer to the solution of greening the roof), and developing evenly distributed areas of the city to reduce the amount of $PM_{2.5}$ dispersed in the atmosphere.
- Developing industries using clean energy, replacing coal and kerosene to reduce emissions containing $PM_{2.5}$ that pollute the environment. Limit the use of charcoal in daily living and cooking. It is the custom of Vietnamese people that needs to be changed.
- Redesigning the traffic routes at peak hours, avoiding long traffic jams, limiting the amount of $PM_{2.5}$ dust generated, and persisting in a specific area. Moreover, public transport must also comply with the latest European standards to limit $PM_{2.5}$ emissions into the environment.
- In addition to mitigation solutions, it needs a dust monitoring network to monitor and control the amount of $PM_{2.5}$ emitted into the air, making appropriate adjustments. In the limited number of monitoring stations, remote sensing techniques to monitor AQ are considered a quick and timely solution to cope with today's rapidly increasing urbanization rate.

4. CONCLUSIONS

The AQ monitoring in HCMC still has certain shortcomings. Meanwhile, HCMC is a typical city of the tropics, with year-round hot and humid temperatures and favorable conditions for the formation and dispersion of pollutants in the air. This study provides the solution for monitoring $PM_{2.5}$ and AQI over citywide space with the support of high-resolution satellite data. The $PM_{2.5}$ predictive regression function based on AOD validated with R^2 and RMSE of 0.89 and 8.063 $\mu g/m^3$, respectively, is promising for application. This study revealed hot spots of $PM_{2.5}$ dust pollution and AQ warning levels in specific areas, which is beneficial for government managers to orient the focus investigations in a proper place to find out the causes of pollution, thereby having suitable solutions to improve AQ. This study demonstrates the potential for reasonable cost monitoring of the AQ for the city level with the detailed spatial distribution when choosing a solution using high-resolution satellite data combined with ground observation data. It is essential for HCMC, but it is impossible to immediately invest in many ground monitoring stations. Finally, HCMC needs to strengthen the ground monitoring station combined with satellite observations to monitor AQ to promptly detect and handle pollution causes to ensure a good quality of life for urban residents.

Acknowledgments

This research was funded by Ho Chi Minh City University of Technology, Vietnam National University - Ho Chi Minh city, under grant number T-MTTN-2020-39. We would like to thank Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for the support of time and facilities for this study. The authors are grateful to the United States Geological Survey (USGS), the National Aeronautics and Space Administration (NASA), and The US Consultation and Environmental Source Samplers company for accessing the satellite imageries and $PM_{2.5}$ data used for this study

REFERENCES

Chander, G. & Markham, B., 2003. *Revised Landsat-5 TM Radiometric Calibration Procedures and Postcalibration Dynamic Ranges*. IEEE T. Geosci. Remote 41(11), 2674-2677.

Chen, J., Zhu, X., Vogelmann, J.E., Gao, F. & Jin, S., 2011. *A simple and effective method for filling gaps in Landsat ETM+ SLC-off images*. Remote Sens. Environ. 115, 1053–1064.

Dang, N.P., 2020. *Assessment of urban air pollution trends and proposed solutions*. Environment 2, 28-31.

DONRE HCMC, the Department of Natural Resources and Environment of Ho Chi Minh City, 2018. *Environmental monitoring report of HCMC in 2018*. Ha Noi, Vietnam. (in Vietnamese)

Giang, T., 2015. *Ho Chi Minh City: Population may be over 10 million in 2020*. http://www.thesaigontimes.vn (in Vietnamese)

Gustavo, C.V., Devis Tuia, D., Luis Gómez-Chova, Sandra Jiménez & Jesús Malo, 2011. *Synthesis Lectures on Image, Video, and Multimedia Processing*, in: Remote Sensing Image processing. Vol. 5, Morgan & Claypool Publishers, pp. 1-192.

Hadjimitsis, D.G. & Clayton, C.R.I, 2009. *Determination of aerosol optical thickness through the derivation of an atmospheric correction for short-wavelength Landsat TM and ASTER image data: an application to areas located in the vicinity of airports at UK and Cyprus*. Appl. Geomatics J., 1, 31-40.

Hadjimitsis, D.G., 2009a. *Aerosol Optical Thickness (AOT) retrieval over land using satellite image-based algorithm*. Air Qual. Atmos. Hlth. 2, 89-97.

Hadjimitsis, D.G., 2009b. *Description of a new method for retrieving the aerosol optical thickness from satellite remotely sensed imagery using the maximum contrast value principle and the darkest pixel approach*. T. GIS, 12(5), 633—644.

HEPZA, Ho Chi Minh City Export Processing and Industrial Zones Authority, 2017. *Planning and projecting the development of export processing and industrial zones in Ho Chi Minh City until 2020*. http://www.hepza.hochiminhcity.gov.vn (in Vietnamese)

Ihlen, V., 2019. *Landsat 7 Data Users Handbook*. EROS. Department of Interior US Geological Survey.

Iqbal, M., 1983. *An Introduction to Solar Radiation*. Academic Press, New York.

IQAIR, 2020. *Air quality in Ho Chi Minh City*. https://www.iqair.com/vietnam/ho-chi-minh-city

Kaufman, Y.J., Fraser, R.S. & Ferrare, R.A., 1990. *Satellite measurements of large-scale air pollution: Methods*. J. Geophys. Res.95, 9895-9909.

Kloog, I. & Alexandra, A., 2014. *A new Hybrid Spatio-Temporal Model for Estimating Daily Multi-Year PM2.5 Concentrations Across Northeastern USA Using High-Resolution Aerosol Optical Depth Data*. Atmos. Environ.95, 581-590.

Kumar, N., 2007. *An empirical relationship between PM2.5 and aerosol optical depth in Delhi Metropolitan*. Atmos Environ, 41(21), 4492–4503.

Lee, H.J. & Liu, Y., 2011. *A novel Calibration Approach of MODIS AOD data to predict PM2.5 concentrations*. Atmos. Chem. Phys., 11, 7991-8002.

Linh, G., 2017. *$PM_{2.5}$ fine dust: Invisible killer*, http://vem.tapchimoitruong.vn/ (in Vietnamese)

Maxwell, S.K., Schmidt, G.L. & Storey, J.C., 2007. *A multi-scale segmentation approach to filling gaps in Landsat ETM+ SLC-off images*. Int. J. Remote Sens., 28, 5339-5356.

MONRE, the Ministry of Natural Resources and Environment, 2018. *National State of the Environment Report 2018*. Ha Noi, Vietnam. (in Vietnamese)

Phu, D., 2017. *Ho Chi Minh City is stuck in traffic not only during rush hour*. https://thanhnien.vn/tp-hcm-ket-xe-khong-chi-gio-cao-diem-post685217.html (in Vietnamese)

QCVN 05:2013/BTNMT, 2013. *National Technical Regulation on Ambient Air Quality*. Ha Noi, Vietnam. (in Vietnamese)

Retalis, A., 1998. *Study of atmospheric pollution in Large Cities with the use of satellite observations: development of an Atmospheric correction Algorithm Applied to Polluted Urban Areas*. PhD Thesis, Department of Applied Physics, University of Athens.

Retalis, A., Cartalis, C. & Athanasiou, E., 1999. *Assessment of the distribution of aerosols in the area of Athens with the use of Landsat TM*. Int. J. Remote Sens., 20, 939-945

Satheesh, S.K. (2002). *Aerosol and Climate. Resonance*. Journal of Science Education, 7(4), 48-59.

Sifakis, N. & Deschamps P. Y., 1992. *Mapping of Air Pollution Using Spot Satellite Data*. Photogramm. Eng. Rem. S., 58(10), 1433-1437.

Sifakis, N. & Paronis, D., 1998. *Quantitative mapping of air pollution density using Earth observations: A new processing method and application on an urban area*. Int. J. Remote Sens., 19, 3289 – 3300

Thu, N.T.A. & Blume, L., 2018. *Air Quality Report 2017*. https://cleanairday.vn/wp-content/uploads/2019/04/AQR_VIE_FINAL.pdf (in Vietnamese)

Tra, M.T.M., 2020. *Air environment quality in some urban areas in 1-2/2020*. Environment 2, 24-25.

Trong, H. & Ngoc, C. N. M. (2008). *Analyze research data with SPSS*. Hong Duc Publisher, 313. (in Vietnamese)

Tuan, T.V., Hoan, B.D., Sang, N.X., 2021. *Study on the influence of meteorological conditions on the concentrations of PM_{10} and $PM_{2.5}$ in the air around Phap Van - Cau Gie highway*. Journal of Marine sciences and technology, 66, 99-102 (in Vietnamese).

USEPA, 2018. *Technical Assistance Document for the Reporting of Daily Air Quality – the Air Quality Index (AQI)*. Publication No. EPA-454/B-18-007

Van, T.T., 2010. *Urban heat island under impact of urbanization in Hochiminh city from remote sensing data*. Sci. Tech. Dev.J. - Sci. Earth Environ., 13(M2), 103-117 (in Vietnamese).

Van, T.T., Binh, T.T. & Bao, H.D.X., 2012. *Study of dust pollution detecting ability in urban areas by remote sensing technology to support air environment observation*, Sci. Tech. Dev.J. - Sci. Earth Environ., 15(M2), 32-47. (in Vietnamese)

Van, T.T., Khanh, NP. & Bao, H.D.X., 2014. *Based on remote sensing optical thickness simulating dust distribution PM_{10} in the inner city of Ho Chi Minh City*. Sci. Tech. Dev.J. - Sci. Earth Environ. 30(2), 62-72 (in Vietnamese)

Wijeratne, I., 2003. *Mapping of Dispersion of Urban Air Pollution Using Remote Sensing*. International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences. Vol. XXXVI - Part 2.

Yin, G., Mariethoz, G. & McCabe, M., 2016. *Gap-Filling of Landsat 7 Imagery Using the Direct Sampling Method*. Remote Sens, 9, 12.

Received at: 04. 04. 2022
Revised at: 04. 06. 2022
Accepted for publication at: 22. 06. 2022
Published online at: 07. 07. 2022