Correlation of PM2.5 based on Landsat 8 satellite imagery with ground measurements in Bandung City

R Hernawati and S Darmawan

Department of Geodesy Engineering, Institut Teknologi Nasional, Jalan PKH. Mustofa no. 23, 40124 Bandung, Indonesia

Email: rikah@itenas.ac.id

Abstract. The purpose of this research is to find the correlation of PM2.5 with ground measurements. The calculated of PM2.5 were using the algorithm derived from Landsat imagery in the year of 2019 and the concentration of PM2.5 ground measurements was collected using tools of sensor pocket PM2.5. The Implementation of methodology were include geometric and radiometric correction, calculation for PM2.5 algorithm include AOD from AERONET, calculated LST, and calculated TVDI. The calculated data on PM2.5 were compared with ground measurements of PM2.5 from different locations in Bandung and the result show that a linear correlation of PM2.5 derived from Landsat between ground measurements having coefficient of determination value is 0.65. It means, the data of PM2.5 derived from Landsat 8 imagery between the data of PM2.5 ground measurements indicates a moderate positive correlation with linear regression model $y = 0.8023x + 3.8578$.

1. Introduction

In recent years, the problem of air pollution has become a major concern in industrial and urban areas, such as Bandung City, because air pollution in Bandung is already at the limit of the Indonesian National Ambient Air Quality Standard value [1]. The cause of this air pollution occurs due to the incomplete combustion process from motor vehicle fumes, one pollutant that causes pollution is particulate matter 2.5 (PM 2.5). PM2.5 refers to the particles which is less than or equal to 2.5 μm, also known as inhalable particles that are very dangerous in air quality.

According to the Environmental Management Agency, one of the sources of PM2.5 that is spread in the atmosphere comes from the incomplete combustion fumes from motor vehicles [1]. Motorized vehicles add PM2.5 concentration levels from exhaust emissions, brake and tire abrasion, road use, and road and soil dust resuspension [2]. High levels of PM2.5 concentrations are usually found in urban areas with heavy traffic and populations. This causes the number of people exposed to high, thus worsening the level of health.

Previous studies showed that the best way to identified of PM was measured by in situ ground measurements, while field observation measurements require high costs for installation and maintenance and the data collected is only effective in small spaces with coverage around the observer station. Thus, measurements at the observer station cannot provide detailed spatial distribution of PM over large areas [3]. Now a day, the studies on PM concentrations are usually based on spatial data and temporal data series measured at the location of air pollution monitoring stations in cities and rural areas [4]. The development of remote sensing technique can assist in detection and analysis of air pollutant data derived from satellite imagery data. The estimation or identification of PM from remote sensing satellite data had been applied in recent years in many area. Several studies have shown the relationship of satellite data with air research for atmospheric studies and applied it with the algorithm to estimation or
identified the air pollutant. Many researcher had used different algorithm model based on satellite imagery to find out the identification of PM2.5.

Based on the research conducted by Chen [5], they used Landsat 8 to estimate the concentration of PM2.5 on the ground surface using 3 parameters include Aerosol Optical Depth (AOD), Land Surface Temperature (LST), Temperature Vegetation Dryness Index (TVDI). In this study the algorithm model was developed from [5] has been improved to identification of PM2.5 for Landsat 8 imagery over urban area in Bandung. But in this research the AOT data was using the data from AERONET (Aerosol Robotic Network). AERONET is a NASA project that a version of a ground-based aerosol monitoring system that offers a standardization for a ground-based regional to global scale aerosol monitoring and characterization network [6]. It provides continuous cloud-screened observations of spectral aerosol optical depth (AOD), precipitable water, and inversion aerosol products in diverse aerosol regimes [7]. Ground measurements data of PM2.5 using sensor pocket required for correlation between reflectance values of Landsat 8.

2. Methodology

The technique used in this research based on a model developed by Chen [5], which had been modified to estimation of PM2.5 using Landsat 8. The methodology include study area of Bandung City, data acquisition, data processing include geometric correction, estimating of Top of Atmospheric (ToA), NDVI, LST, TVDI, calculation of PM2.5, and correlation with PM2.5 ground measurements.

2.1. Study Area

The study located in Bandung city, West Java with 167.7 km² in area with geographic coordinates: 107° 36' E and 6° 55' S. This region was chosen because it has a high population density, rapid economic development, and infrastructure, so that hypothetically it would be affected the quality and quantity of vegetation land cover which is will be affected the PM2.5 concentrations.

Figure 1. Study area of Bandung City, picture was taken from Google Earth.
2.2. Data Acquisition

The data used in this research were using Landsat 8, AERONET data, and ground measurements data. Landsat has been selected for the research in the year of 2019 of May 22nd in Figure 2, the LST and NDVI calculated from the TIR band and the visible band were used to calculated the concentration of PM2.5. The detailed of Landsat 8 spesification is given in Table 1.

![Figure 2. Path/row 122/065_22 Mei 2019 of Landsat 8 imagery](image)

Table 1. Band Spesification of Landsat 8 [8]

| Spectral Band                  | Wavelength (µm) | Spectral Resolution (m) |
|-------------------------------|-----------------|-------------------------|
| Coastal Aerosol               | 0.43 – 0.345    | 30                      |
| Blue                          | 0.45 – 0.51     | 30                      |
| Green                         | 0.53 – 0.59     | 30                      |
| Red                           | 0.64 – 0.67     | 30                      |
| Near Infrared (Nir)           | 0.85 – 0.88     | 30                      |
| Shot Wavelength Infrared (SWIR) 1 | 1.57 – 1.67 | 30                      |
| Shot Wavelength Infrared (SWIR) 2 | 2.11 – 2.29 | 30                      |
| Panchromatic (PAN)            | 0.50 – 0.68     | 15                      |
| Cirrus                        | 1.36 – 1.38     | 30                      |
| Thermal Infrared (TIR) 1      | 10.60 – 11.19   | 100                     |
| Thermal Infrared (TIR) 2      | 11.50 – 12.51   | 100                     |
The AERONET data were obtained from AERONET website [7], accessed on the year of 2019. The data are have the categorized into different level of quality [9]. The version of AERONET data for Bandung City are available in 2019, this research were using AERONET data level 2 for cloud–screened and also quality assured [9].

2.3. Data Processing

The first process Geometric Correction involves modeling relationships between imagery and ground coordinate systems. All two images have been geometrically corrected at level one (L1T) to ensure better accuracy. Two images were rectified to a Universal Transverse Mercator (UTM) 48 South Zone coordinate system and datum of WGS 84. Image transformation results have a Root Mean Square (RMS) error of 0.48 in each image, it showed that the image was accurated to one pixel. After geometric correction the imagery was cropping on the study area, it means to focus on the research area and aims to make image interpretation and processing easier.

Radiometric correction applied by converting the Digital Number (DN) to Top of Atmosferic (TOA) Radiance and conversion to TOA reflectance through the algorithm from [10].

2.3.1. Conversion to TOA Radiance

The first thing to do is to convert the digital number on Landsat 8 becomes spectral radiance by using the following equation [10].

\[ L_\lambda = ML.Qcal + AL \]

Where :
- \( L_\lambda \) = TOA spectral radiance (Watts/( m² * srad * μm))
- \( ML \) = Band-specific multiplicative rescaling factor from the metadata RADIANCE_MULT_BAND_x, where x is the band number
- \( AL \) = Band-specific additive rescaling factor from the metadata (RADIANCE_ADD_BAND_x, where x is the band number)
- \( Qcal \) = Quantized and calibrated standard product pixel values (DN)

2.3.2. Conversion to TOA Reflectance

The following equation is used to convert DN values to TOA reflectance as follows [10]:

\[ \rho'_\lambda = M\rho.Qcal + A\rho \]

Where :
- \( \rho'_\lambda \) = TOA planetary reflectance.
- \( M\rho \) = Band-specific multiplicative rescaling factor from the metadata (REFLECTANCE_MULT_BAND_x, where x is the band number)
- \( A\rho \) = Band-specific additive rescaling factor from the metadata (REFLECTANCE_ADD_BAND_x, where x is the band number)
- \( Qcal \) = Quantized and calibrated standard product pixel values (DN)

2.3.3. Particulate Matter (PM\(_{2.5}\))

To get the estimation of PM2.5 this research refer to [5] which is the algorithm using NDVI, LST and TVDI parameter, formula PM2.5 :

\[ PM2.5 = e^{b_{box} (AOD)} b^{\Lambda\alpha} (TVDI)^{\beta_{TVDI}} x (LST)^{\beta_{LST}} \]

Data AOT are retrieved from situs NASA AERONET, LST is land surface temperature (LST) retrieved from the thermal infrared band data (band10) of the Landsat 8 data using Image-based Method. \( \beta_0 \), \( \beta_{AOT} \), \( \beta_{TVDI} \) and \( \beta_{LST} \) are regression coefficients for AOD, LST, TVDI and other factors, respectively. Before calculating PM2.5, first compute NVDI, LST and TVDI.
2.3.4 Calculation of Normalized Difference Vegetation Index

Normalized Difference Vegetation Index (NDVI) Algorithm At this stage data processing is carried out using the Normalized Difference Vegetation Index (NDVI) method to produce a vegetation index [11].

\[
\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})}
\]

Where, NIR and Red are the DN value of band NIR and band Red, respectively. To NDVI, the larger the value, the higher the vegetation coverage is, while the smaller the value, the closer the bare soil. In order to improve the accuracy of NDVI.

The NDVI value between vegetation and bare soil indicates that the pixel consists of a certain percentage of vegetation and a certain percentage of bare soil. So, we determined the vegetation coverage of every pixel in the image with the following formula [9]:

\[
P_v = \left[ \frac{(\text{NDVI} - \text{NDVI}_{\text{min}})}{(\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}})} \right]^2
\]

Where:
- \( P_v \) = Proportion of Vegetation
- \( \text{NDVI} \) = DN values from NDVI Image
- \( \text{NDVI}_{\text{min}} \) = Minimum DN values from NDVI Image
- \( \text{NDVI}_{\text{max}} \) = Maximum DN values from NDVI Image

In this study, \( \epsilon \) of water surface pixels is 0.995, \( \epsilon \) of town and natural surface can be calculated with the following formula

\[
\epsilon = 0.985P_v + 0.960(1-P_v) + 0.06P_v (1-P_v)
\]

Where:
- \( \epsilon \) = Land Surface Emissivity
- \( P_v \) = Proportion of Vegetation

2.3.5 Land Surface Temperature

Where, \( \epsilon \) surface is the land surface emissivity of natural surface, and \( \epsilon \) town is the land surface emissivity of town surface. In this study, band TIR on Landsat 8 to retrieve the brightness temperature of the land surface, and then calculated the land surface temperature using \( \epsilon \). The land surface brightness temperature can be calculated with the following formula:

\[
T_b = \frac{K_2}{\ln\left(\frac{K_1}{T_b} + 1\right)}
\]

\( T_b \) : Brigthness Temperature satelit (°C)
\( K_1 \) : Band calibration constant 10 (Landsat 8))
\( K_2 \) : Band calibration constant 10 (Landsat 8))
\( L_\lambda \) : Radiance spectral TOA (watts/m2 Srad μm)

Finally, land surface temperature (Ts) was calculated as:

\[
T_s = \frac{T_b}{1 + (\frac{L_\lambda}{\epsilon})\ln\epsilon}
\]
Ts : Land Surface Temperature (ºC)
Tb : Temperature Brightness (ºC)
λ : Mid-wavelength value band 10
δ = hc/σ : 1.438 x 10^-2 mK
h : Konstanta Planck (6.26 x 10^-34 J sec)
c : Speed of light (2.998 x 10^8 m s^-1)
σ : Konstanta Stefan-Boltzman (1.38 x 10^-23 J K^-1)
ε : Emisivitas objek

2.3.6 TVDI (Temperature/Vegetation Condition Index)
A simplified water stress index called the TVDI, based on an empirical interpretation of the NDVI–Ts space, was suggested by [12]. The TVDI is related to the soil moisture status so that high values indicate dry conditions and low values indicate moist conditions. The TVDI is a no-moisture index, reaching values of 1 at the dry edge (water limited condition) and 0 at the wet edge (adequate water availability). Based on remotely sensed observations, the TVDI can be formulated as:

\[ TVDI = \frac{Ts - Ts_{min}}{(a + b \times NDVI - Ts_{min})} \]

Ts : Land Surface Temperature
Tsmin: Land Surface Temperature minimum
Tsmax : Land Surface Temperature maximum (a + b x NDVI)

3. Result and Discussion

3.1. PM2.5 Concentration
PM2.5 concentration estimates are obtained from the extraction results between LST and TVDI. This method refers to the research conducted by [5] in Bandung city derived from Landsat 8 imagery.

Figure 3. Identification of PM2.5 based on Landsat 8 imagery
In determining the concentration of PM2.5, several parameters are needed, include soil surface temperature (LST) and vegetation drought index (TVDI) which aims to show the status of soil moisture so that high values indicate dry conditions and low values indicate moist conditions. In this processing required the value of Aerosol of Depth (AOD), the value of vegetation density, soil surface temperature, and drought index where the AOD value is obtained from AERONET. The AOD from AERONET values in the site are displayed based on observations of seconds per day. The value of aerosol depth, vegetation density, temperature, and drought were affect of PM2.5 concentration. The concentration of pollutants can be affected by wind, humidity, and air temperature.

The result shows that the estimation of PM2.5 in Bandung (show in Figure3) are in the range 0µm until 15.4 µm which is shown in green color and 15.4 µm until 35.4 µm shown in yellow one. It means with a classification that refers to regulation of Government Regulation No. 41 of 1999 and Head of Environmental Impact Management Agency No. 107 of 1997 based on Air Pollution Standard Index value the air in the city of Bandung is included to category of healthy/good to be inhaled.

3.2. Ground Measurements

PM2.5 ground measurements data collection had done by tracking along the road in the city of Bandung. Ground measurements are carried out in the morning and evening, which are adjusted to the recording of Landsat satellites, while daily activities begin during the morning and evening to measure PM2.5 levels. Measurements had taken at 08.00 am until 10 am in the morning and 16.00 pm until 18.00 pm in the evening. The measurement results obtained as Figure 4 and Figure 5 and the fix data show in Table 2.
Tabel 2. Ground measurements data

| DATE      | TIME     | Latitude   | Longitude  | PM2.5(ug/m³) |
|-----------|----------|------------|------------|--------------|
| 4/25/2019 | 13:27:36 | -6.899882793 | 107.6432114 | 30.5         |
| 4/25/2019 | 13:27:37 | -6.899882793 | 107.6432114 | 30.2         |
| 4/25/2019 | 13:27:38 | -6.899882793 | 107.6432114 | 30           |
| 4/25/2019 | 13:27:39 | -6.899882793 | 107.6432114 | 29.9         |
| 4/25/2019 | 13:27:40 | -6.899882793 | 107.6432114 | 28.7         |
| 4/25/2019 | 13:27:41 | -6.899882793 | 107.6432114 | 28           |
| 4/25/2019 | 13:27:43 | -6.899882793 | 107.6432114 | 27.9         |
| 4/25/2019 | 13:27:44 | -6.899882793 | 107.6432114 | 27.6         |
| 4/25/2019 | 13:27:45 | -6.899882793 | 107.6432114 | 27.3         |
| 4/25/2019 | 13:27:46 | -6.899882793 | 107.6432114 | 27.1         |
| 4/25/2019 | 13:27:47 | -6.899882793 | 107.6432114 | 26.8         |
| 4/25/2019 | 13:27:48 | -6.899882793 | 107.6432114 | 26.5         |
| 4/25/2019 | 13:27:49 | -6.899882793 | 107.6432114 | 26.3         |
| 4/25/2019 | 13:27:50 | -6.899882793 | 107.6432114 | 26.2         |
| 4/25/2019 | 13:27:51 | -6.899882793 | 107.6432114 | 25.9         |
| 4/25/2019 | 13:27:52 | -6.899882793 | 107.6432114 | 25.5         |
| 4/25/2019 | 13:27:53 | -6.899882793 | 107.6432114 | 25.1         |
| 4/25/2019 | 13:27:54 | -6.899882793 | 107.6432114 | 24.8         |

The measurement results for PM2.5 in Figure 6 in the morning show the level of pollutants around 300 µg / m³, where the level of these pollutants shows the classification at the level of 250.4 - 500.4 µg / m³ (dangerous) for the type of pollutants PM2.5. As for the measurements in the afternoon in Figure 7, the level of PM2.5 pollutants is at the level of 350 µg / m³ (dangerous).
3.3 Correlation PM2.5 concentration Between Landsat Imagery and Ground Measurements

After getting the results of PM2.5 processing, an analysis was carried out to determine the correlation between PM2.5 concentration values of Landsat images and PM2.5 values from ground measurements. Both PM2.5 concentration values were analyzed for their correlation using linear regression. The relationship of PM2.5 processing using Landsat Satellite Imagery in 2019 and PM2.5 field measurements of PM2.5 were carried out in 2019. From the results of processing and measurement, there is a very significant difference. PM2.5 processing results using Landsat images obtained relatively small results on average 22.8 µg / m³, while the field measurement data showed a large enough value of 150 µg / m³. As an example, the results from ground measurements in Figure 9 and data processing in figure 8 can be compared from samples taken through the same coordinates. From the sample taken PM2.5 results obtained measurements of 31.9 µg / m³ while the processing obtained by 22.51 µg / m³.
Based on the results of PM processing based on Landsat imagery with the measurement results in the field there are significant differences. The results of PM concentrations obtained based on processing on images are relatively small compared to the results of PM concentrations from field measurements. For PM2.5 processing results obtained a maximum value of 25.47 µg / m³ with an average of 15.21 µg / m³ while based on measurement results in the field there is a maximum value of up to 999 µg / m³. For some factors that cause this significant difference are the influence of the closest object sensor and wind speed. Measuring instrument used (measurement ground) is a measurement tool that is more suitable used to measure the concentration of PM issued by each object. For example, measurements were made on motor vehicle recognition in order to find out how much PM concentration value issued by the object. Following Figure 10 is the results of linear regression between PM concentration based on processing on Landsat imagery and field measurement results.

![Figure 10](image)

Figure 10. The correlation of linear regression PM2.5 ground measurement with PM2.5 Landsat 8 satellite imagery in 2019.

The pattern of relationship has equation model $y = 0.8023x + 3.8578$, which is obtained from regression analysis where $x$ is variable of the PM2.5 ground measurement and $y$ is variable of the PM2.5 based on Landsat-8. The coefficient of determination ($R^2$) is 0.65. It means that the data of PM2.5 derived from Landsat 8 imagery between the data of PM2.5 ground measurements indicates a moderate positive correlation.

4. Conclusion

PM2.5 estimation results derived from Landsat imagery are correlated with ground measurement data with a correlation determination value ($R^2$) is 0.65, it means the correlation of PM2.5 based on Landsat-8 with PM2.5 ground measurement indicates a moderate positive relationship. Based on the results of research that has been done by calculating the concentration of PM2.5 in the city of Bandung, the air quality obtained is included in the category of Good (Healthy) to be breathed by living things. However, the concentration of PM in the city of Bandung is gradually increasing and will lead to several possibilities in the future if there is no anticipation of the living creature itself that is the source of pollutants.
5. Acknowledgments
Thanks to LPPM/ Lembaga Penelitian dan Pengabdian Masyarakat (Research Institutions and Community Service) Institute of Technology Nasional Bandung for providing research funding and this research is still ongoing and develop, also thanks to Prof. Wataru Takeuchi from IIS-Tokyo University for research collaboration, and thanks to Aab, Anisa, Novita, Anis, Reza, Felita, Ichwan, and Derry.

6. References

[1] BPLHD. Badan Pengolahan Lingkungan Hidup Daerah, 2015 *Kondisi Udara di Kota Bandung* https://jabarprov.go.id/index.php/pages/id/1321

[2] Malina, Christiane, and Fischer, Frauke, (2012). *The impact of low emission zones on PM10 levels in urban areas in Germany*. No. 58. CAWM Discussion Paper, Centrum fürAngewandte Wirtschaftsforschung Münster.

[3] Wald, L., and Baleynaud, J. (1999). *Observed Air Quality Over City of Nantes by Means of Landsat Thermal Infrared Data*. International Journal of Remote Sensing, 20, 947-959.

[4] Retalis, A., Cartalis, C., & Anthassioss, 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.

[5] Chen, Y., Han, W., Chen S., Tong, Ling., (2014). *Estimating ground-level PM2.5 Concentration using Landsat 8 in Chengdu, China*. Remote Sensing Of Atmosphere Vol 9259 925917-1.

[6] Holben B.N., T.F. Eck, I. Slutsker, D. Tanré, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov. (1998). AERONET - A federated instrument network and data archive for aerosol characterization, Rem. Sens. Environ., 66, 1-16.

[7] “Aerosol Robotic Network (AERONET) Homepage.” NASA, NASA, aeronet.gsfc.nasa.gov/.

[8] “Landsat Missions.” *Landsat 8*, www.usgs.gov/land-resources/nli/landsat/landsat-8?qt-science_support_page_related_con=0#qt-science_support_page_related_con.

[9] Li, Z., Roy, D. P., Zhang, H. K., Vermote, E. F., & Huang, H. (2019). *Evaluation of Landsat-8 and Sentinel-2A aerosol optical depth retrievals across Chinese cities and implications for medium spatial resolution urban aerosol monitoring*. Remote Sensing, 11(2), 1-14. https://doi.org/10.3390/rs11020122.

[10] USGS. *United States Geological Survey Landsat 8 Product*. http://landsat.usgs.gov/Landsat8_Useing_Product.php.

[11] Weier, J, and Herring, D. 2000. “Measuring Vegetation (NDVI & EVI).” NASA, NASA, 30 Aug. 2000, earthobservatory.nasa.gov/features/MeasuringVegetation.

[12] Sandholt, I., K. Rasmussen, and J. Andersen, A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. Remote sensing of environment, 2002. 79(2): 213-224.