Application of Remote Sensing Instruments in Air Quality Monitoring in Malaysia

Mohd Muzammil, SALAHUDDIN, a Zulfa Hanan, ASHAARIib*

a b Department of Environmental Science, Faculty of Environmental Studies, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia
*zulfa@upm.edu.my

Abstract – The use of remote sensing in detecting aerosol or air pollution is not widely applied in Malaysia. The large area of coverage provided by remote sensing satellite may well be the solution to the lack of spatial coverage by the local ground air quality monitoring stations. This article discusses the application of remote sensing instruments in air quality monitoring of Malaysia. The remote sensing data is validated using ground truths either from local ground air monitoring stations or the Aerosol Robotic Network (AERONET). The correlation between remote sensing is relatively good with R from 0.5 to 0.9 depending on the satellite used. The correlation is much improved using the mixed effects algorithm applied on MODIS Aerosol Optical Depth (AOD) data. Accuracy of predicted air quality data by remote sensing is generally tested using the Root Mean Squared Error (RMSE) against the ground truths data. Besides the Geographic Information System (GIS) tools are used in manipulating the data from both remote sensing and ground stations so as to produce meaningful results such as spatio-temporal pattern mapping of air pollution. Overall the results showed that the application of remote sensing instruments in air quality monitoring in Malaysia is very useful and can be improved further.

Keywords: Aerosol Optical Depth, Air Quality, GIS, Malaysia, Remote Sensing.

Introduction
Atmospheric aerosol is becoming more important each day because of its various negative effects not only on human beings but also to the environment. By definition, aerosols near the ground can be called atmospheric particulate matter (PM) (Pope et al. 1995). Particulate matters smaller than 10 μm can cause or worsen respiratory and cardiovascular diseases by entering human lungs via respiration (Sardar et al. 2005; Franck et al. 2011). If the PM is more than 10 μm (PM10), it would not be able to enter human body through respiration due to human’s nose and throat, which filters the large particles via cilia and mucus. If the diameter is between 2.5 μm and 10 μm, they can enter the upper respiratory tract, but some of them will be blocked by inside the nose, thus relatively does less harm to human health. Particulate matters with diameter less than 2.5 μm (PM2.5), can be inhaled into the body and is
not easy to be blocked. Hence it can cause devastating effect to human health such as asthma, bronchitis and cardiovascular diseases. These particles may go into the bloodstream and together with metal and gases dissolved in blood and they can be very dangerous and harmful (Pope et al. 1995; Wang and Christopher, 2003; Fang et al. 2006).

A study has shown that acid deposition due to the atmospheric aerosol transportation influences the river and ocean chemistry (He et al. 2011). Also aerosol particle radiative effect impacts photochemical rates in plants (Tang et al. 2003) and causes food contamination (Srinivas et al. 2009). Aerosol particles in the atmosphere are important in the atmospheric radiative budget because of their ability to absorb and scatter the solar and terrestrial radiation (Davison et al. 2004; Rajeev et al. 2008; Feng and Christopher 2013). Aerosol indirectly impacts the ocean photochemistry as a result of the radiation perturbation in the atmosphere and can be seen in the coral record (Risk et al. 2003).

Aerosol pollution in Malaysia mainly originates from the local production of aerosols. They can be emissions from millions of automobiles on the road, from factories throughout Malaysia, burning activities including fossil fuels in power plants and open burning activities of leaves, wood and trash. Apart from the local sources of anthropogenic aerosol, Malaysia is also exposed to the trans-boundary air pollutants resulting from the Southeast Asia biomass burning and forest fires. The transboundary air pollutants contribute to the high load of aerosol and pollution especially during the dry season from June to September (Kanniah et al. 2013). This suggests that the knowledge of spatial and temporal pattern of the aerosol is crucial to prevent, manage and control any hazardous events that occur due to the movement and transportation of aerosol in the atmosphere.

In Malaysia, data of air quality, including that of PM10 is generally recorded by ground stations located across the country. The monitoring networks are managed by the Malaysian Department of Environment (DOE) via a private company known as Alam Sekitar Malaysia Sdn Bhd (ASMA) and the Malaysia Meteorological Department (MMD). They both have 51 (Juneng et al. 2009) and 22 (Kanniah et al. 2013) ground stations respectively nationwide. The continuous air quality monitoring

Figure 1 Location of CAQM stations in Peninsula Malaysia (Malaysian environmental quality report 2013; Doe (2014))
(CAQM) of DOE consists of 5 types of stations which are Industrial, Residential, Traffic, Background and PM10. PM10 parameter is measured by all type of stations. The automatic CAQM data will be collected hourly during the monitoring period. The CAQM is also supplemented by manual air quality monitoring stations (High Volume Sampler) located at 19 different sites. Parameters like Total Suspended Particulates (TSP), PM10 and several heavy metals such as lead mercury, sodium, iron, copper are measured once every six days. These data are collected monthly by ASMA.

Ground stations offer a continuous high frequency of highly quality data but still lack in the spatial coverage that is required to derive a good synoptic spatial pattern. Besides the basis of aerosol detection used by ground stations are surface based, which means any variability of columnar aerosol in the atmosphere cannot be detected. The use of remote sensing in detecting atmospheric aerosol has gained popularity in recent years due to its large spatial coverage, high frequency data, large database and highly accurate data which are being improved all the time with the introduction of better algorithms. Remer et al. (2008) reported the accuracy of MODIS Aerosol Optical Depth (AOD) is ±0.05 AOD under clear skies and ±0.15 AOD under moderately contaminated atmosphere) over land. Apart from MODIS, there are also other sensors such as Landsat satellite, Total Ozone Mapping Satellite (TOMS) and Satellite Pour l’Observation de la Terre (SPOT) which have been used to detect atmospheric aerosol with their own unique algorithms and techniques. These captured satellite images ARE manipulated using various techniques and algorithms in order to extract information according to the applications required. The remote sensing data for aerosol detection is usually validated against the Aerosol Robotic Network (AERONET). AERONET is a program established by NASA and PHOTONS (PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire) which provides network of ground-based remote sensing of aerosol, measuring atmospheric aerosol properties using sun photometers. Using remote sensing together with ground based monitoring stations may provide high frequency and highly accurate data that cover a large spatial area. As a result, a better synoptic spatio-temporal pattern study of aerosol in the atmosphere, especially in the study area can be obtained. Geographic Information System (GIS) software will be used to manipulate the data acquired from the satellites and ground stations so as to create aerosol mapping for better and clearer presentations.

The following sections discuss the data product in the application of remote sensing, the relationship between remote sensing and ground based data, and the contribution of the GIS. The discussion continued with the issues of air quality standards and the comparison of remote sensing application in air quality monitoring in Malaysia

Data product in the Application of Remote Sensing

MODIS AOD & FMF

MODIS Level 2 Aerosol Optical Depth (AOD) data is retrieved globally in a daily basis at the spatial resolution of 10km x 10km. MODIS AOD satellites, which consists of Aqua and Terra overpass Malaysia once A day around 10.30 am and 1.30 pm LST, thus enabling data capturing at two different times per day. Most of aerosol studies via remote sensing in Malaysia utilized MODIS AOD level 2 products. Yap & Hashim (2013), Yap, Hashim et al. (2011), Kanniah et al. (2013, 2010) and Jamil et
al (2011a; 2011b) used MODIS Level 2 AOD product to study the aerosol pattern in their respective study area. Due to the continuous presence of cloud cover in the region, retrieval of aerosol data from satellite sensors in Malaysia is often missing. As explained by Yap & Hashim (2013) and Yap, Hashim et al. (2011), this problem can be overcome by using an averaging technique of 5 x 5 windows. The technique assumes that the neighboring 5 pixels with no AOD retrieval have the same value with the reference pixel with a valid retrieval. This means that if there is no retrieval of AOD in that particular area, then the nearest 5 pixel (50 km) retrieval will be used. In this regard, the number of pixels without AOD information due to cloud cover can be reduced. If there is a continuous valid AOD retrieval, a normal averaging scheme will be applied by ignoring pixels with no AOD retrieval. On averaging multiple pixels, it is expected to reduce the influence of random errors associated with the retrieval of AOD. Furthermore, a 5×5 window averaging has been widely used in MODIS validation work, which is in agreement with the average speed of aerosol air mass transport in the mid-troposphere in the Atlantic (Ichoku et al. 2002; Remer et al. 2005). However, as the average wind speed near the earth surface is much less than mid-troposphere, a 5×5 window is considered appropriate. If a 3×3 window is used, we found that there are many voids left in the imagery that resulted in the poor retrieval of the overall MODIS AOD in Peninsular Malaysia. The projection used by MODIS in its data is the decimal degrees system. Therefore the projections of location of the ground stations are converted into decimal degrees to locate the correct pixels from MODIS image for the corresponding ground stations.

MODIS Fine Mode Fraction (FMF) is a product derived from the MODIS AOD. FMF is the ratio of the aerosol optical depth (AOD) of small aerosols to the AOD of all aerosols. FMF can be used together with AOD to distinguish aerosol types and sources regions based on the sensitivity of the AOD to aerosol column density and the dependence of FMF on the aerosol particles’ size (Kanniah et al., 2013).

**SPOT 5**

The very high spatial resolution images of SPOT-5 were used by Lim, MatJafri et al. (2008) to study the relationship of SPOT-5 images to PM2.5 using the dark surface targets algorithm in the blue band in Penang Island, Malaysia. The surface reflectance from the raw image is determined by using ATCOR2 image processing technique. The digital number (DN) were retrieved for all the bands and converted to radiance and then to reflectance values (satellite reflectance). The atmospheric reflectance was then estimated by subtracting the surface reflectance from the satellite observation reflectance values. The atmospheric reflectance values are then used to map the PM2.5 concentration using the calibrated algorithm. For validation purposes, 25 ground truth data for PM2.5 are collected using handheld PM2.5 meter (DustTrak) and the locations of each ground points are noted with the help of handheld GPS device. The digital signal from the satellite was correlated against the recorded ground truths data using the regression algorithm analysis.

**LANDSAT 7-ETM**

Ibrahim & Kamarulzaman (2012) used the high resolution multispectral Landsat 7-ETM satellite image to study air pollution in the southern part of Peninsula Malaysia, Johor Bahru. The Aerosol
Optical Thickness (AOT) is derived from the raw satellite images which includes manipulation of the visible, near infrared and thermal infrared bands of Landsat 7-ETM satellite. The technique applied here is to do radiometric comparison between satellite image on a polluted day and the one on a clean or less polluted day. The satellite images were processed and corrected for geometric and radiometric distortion and were adapted to Malaysia’s geo-reference system. In this study, the simplified model known as the “Apparent Reflectance” model for converting the DN image into at-satellite reflectance image is used due to limited parameters availability. Several assumptions were made to the original algorithm because of the lack of data on parameters. By applying the ratio of the “clean day” image pixel value to the “polluted” image pixel value, the AOT can be calculated using a technique called the “blurring effect”.

NOAA AVHRR

The National Oceanic and Atmospheric Administration (NOAA)’s Advance Very High Resolution Radiometer (AVHRR) data, retrieved from SEAFDEC (Southeast Asia Fishery Development Centre) were used by Asmala Ahmad et al. (2006) to estimate the air quality index (AQI) in Peninsular Malaysia. The data offer high spectral and temporal resolution with minimum cost making them very attractive in environmental study. The satellite is also used for other various applications such as sea-surface temperature determination, snow and ice detection, daytime cloud and surface mapping and haze mapping. Similar with other remote sensing satellites, the satellite detects air pollution by measuring radiation affected by scattering and absorption of molecules in the atmosphere. This prevents any false interpretations due to the visual similarity between cloud and haze. Before the data being processed, a calibration of the visible Band 1 of the NOAA-14 AVHRR was done to compensate a known data degradation issue from the sensor. Distinguishing cloud from haze is done using the thresholding technique. The model used in this study exploit the skylight to indicate the existence of haze. Occurring especially during hazy seasons, skylight is an indirect radiation scattered by elements in the haze layer. The calibrated data were collected within a 2.5km radius from each of the participating air pollution stations on the ground. The relationship between the satellite and ground-based air quality data are analysed using linear regression.

<table>
<thead>
<tr>
<th>Study/Paper</th>
<th>Data Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanniah, Yaso et al. (2013)</td>
<td>MODIS FMF</td>
</tr>
<tr>
<td>Lim, MatJafri et al. (2008)</td>
<td>SPOT 5</td>
</tr>
<tr>
<td>Ibrahim and Kamarulzaman (2012)</td>
<td>Landsat 7 ETM</td>
</tr>
<tr>
<td>Ahmad, Hashim et al. (2006)</td>
<td>NOAA AVHRR</td>
</tr>
</tbody>
</table>
Ground Truths

Ground truths air pollution and ancillary data (such as wind direction, wind speed etc.) are provided by the local authorities in Malaysia which are the DOE and MMD. This private company ASMA maintains the Malaysian Continuous Air Quality Monitoring (CAQM) stations in Malaysia which consists of 52 stations for the DOE, established since 1996. Using the β-ray attenuation mass monitor (BAM-1020) by Met-One Instrument Inc which provides air pollution data collection accuracy of ±8% of indication for 1 h mode (Kanniah et al., 2013). MMD operates 22 stations primarily for meteorological purposes, but are also capable of recording air quality data. The highly dependable, high quality data from the ground stations are vital to study the relationship between data from ground stations and data from remote sensing and are used very widely in air quality studies in Malaysia. They are also used in validating any model and algorithm produced using the remote sensing data. Ancillary data is incorporated in air pollution model to increase its efficiency and the accuracy of the prediction.

The other ground-based data used in environmental studies in Malaysia are retrieved from the Aerosol Robotic Network (AERONET). The sun-photometer measures directly the sun radiation from each of the network stations located all around the world and gives AOD reading. While the accuracy of data is very high, there are not that many AERONET stations in this region to be statistically manipulated to produce conclusive results of spatial pattern of AOD. But its high accuracy of data makes them very suitable for validating the remotely sensed AOD data from the satellites. Kanniah et al. (2013) utilized the hourly AOD data from AERONET to study the diurnal pattern of AOD and also as validation for the MODIS AOD in the study area.

The Relationship between Remote Sensing and Ground-Based Data

They are couple of statistical analyses used in remote sensing environmental studies in Malaysia, usually to determine the relationship between remotely sensed data and ground-based data. Statistical tools can also be used to test the accuracy of the air prediction model, which are applied in some of the studies.

Linear Regression Analysis

Linear Regression analysis is an important tool in studying the relationship between the remote sensing data and the ground truths. It is a multivariate analysis, which means the analysis can be done between multiple variables. In most of the studies, for instance the studies conducted by Kanniah & Yaso (2010), Yap, Hashim et al. (2011), Ibrahim, Ismail et al. (2012) and Yap & Hashim (2013), the linear regression is applied to find the relationship between two parameters, normally PM10 from ground stations against remote sensing data such as AOD or atmospheric reflectance values. The resulting coefficient from the analysis is an indicator of the correlation between the sets of data and can be used to predict the PM10 value in areas without the grounds stations using the remote sensing data.
Non-Linear Regression Analysis
Another type of regression analysis is the non-linear regression analysis. The main purpose is still the same as of any regression analysis that is to determine the relationship between two parameters. The difference is non-linear regression uses polynomial equation instead of linear equation, as the tool to evaluate correlation between two sets of data. Non-linear is usually applied when the result of linear regression is weak. Amanollahi, Abdullah et al. (2011a) and Amanollahi, Abdullah et al. (2011b) applied non-linear regression analysis to correlate PM10 data from study area to MODIS AOD data.

Root Mean Squared Error (RMSE)
The root-mean-squared error is usually applied to measure the error or differences between real/sample/observed values and predicted values. It is essentially the standard deviation of the errors between these two values. It is useful in these studies where remote sensing data are used to predict the value of air pollution in areas without ground monitoring stations. This quantification of errors can be used to compare the prediction results of different studies using different methods and data and determine which studies produce better results and apply better techniques. Ahmad, Hashim et al. (2006) used RMSE to evaluate the accuracy of the air quality index (AQI) using the regression coefficient from the analysis between NOAA AVHRR data and ground-based data. Yap and Hashim (2013) applied RMSE to assess the accuracy of the mixed effect model, used to predict PM10 using MODIS AOD data.

Geographic Information System (GIS)
GIS is designed to capture, store, manipulate, manage and present all types of spatial or geographical data. Applied in almost everything, including traffic control, weather mapping, population mapping, flood risk mapping, smart farming, GIS has found its way into the environmental management. Also, GIS provides a convenient tool to manipulate and present remotely sensed data, for example Aerosol Optical Depth. Integration with other types of data such as air quality data from ground stations or also meteorological data make it more attractive in environmental researches. In the studies of air pollution and aerosol detection via remote sensing, the function of GIS is normally to put the value of related parameters to air pollution gathered by the CAQM stations and also those from the satellites on similar proper coordinate or map. Not only the value of parameters will be presented more efficiently and clearly on a map, it also allows better comparison with remotely sensed air pollution parameters. Once in the GIS environment, the data can be manipulated in various ways to produce more meaningful results which are not possible before. Functions like ‘fill’ in the GIS environment can be used to fill in gaps in which no data is available, which is very common in satellite detection due to cloud cover. Spatial distribution on a GIS environment can be easily recognized because of the many available tools in the GIS software. Amanollahi, Abdullah et al. (2011a, 2011b) imported the PM10 data from the ground stations and remote sensing satellite into GIS maps to produce a clear distribution map of PM10. Ahmad, Hashim et al. (2006) used GIS interpolation to map haze based on iso-lines and to determine spatial distribution of PM10 in Peninsular Malaysia.

Figure 2 shows the typical flowchart of data processing involved in using remote sensing data from satellites in air quality studies in Malaysia. Initially the retrieval of remotely sensed data is usually
done by downloading the data from the corresponding websites which handles the data from the satellites. MODIS AOD Level 2 product for example can be downloaded from the website https://ladsweb.nascom.nasa.gov. Next, depending on the type of data downloaded, the satellite image might need to be processed and corrected before it can be used. This includes geometric correction, radiometric correction and atmospheric correction. Models or algorithms are then applied to the satellite data retrieved to acquire the desired data product such as Aerosol Optical Depth data or surface reflectance data. This data are used as the primary data. Ground stations provide air quality data such as PM10 that are vital in validating the remotely sensed air quality data. The satellite data products are usually regressed against the ground-based data to evaluate their relationship. Sometimes air quality data predictions are made using the remote sensing data using known models and equations. To test the accuracy of the predicted air quality data, a RMSE is the common tools applied. All of these processes can be done in a GIS environment, by which the satellite data product and the ground-based data are imported together. GIS environment provides powerful tools in manipulating the data such as interpolation, which give better results and graphical presentation in maps.

![Figure 2: Typical flowchart of data processing in remote sensing analysis of air quality monitoring](image-url)
Air Quality standards in Malaysia

Before interpreting the results of the air pollution studies in Malaysia, it is important to know about the standard of air quality applied. Generally there are two standards being used to gauge air quality and pollution in Malaysia, namely; 1) Malaysian Ambient Air Quality Standard and 2) Air Pollution Index. The MAAQS is a more specific standard of known parameters that are hazardous to human health and is measured from one to 24 hours for different air pollutants. The MAAQS is shown in Table 2. The Air Pollutant Index (API) on the other hand is a simple index to indicate air quality by taking considerations of all the parameters in the MAAQS except for TSP and lead. This simple gauge is very useful in quickly determining the air quality for the general public. In Table 3 the standard API is shown with different values corresponding to different air qualities.

Table 2 Malaysian Ambient Air Quality Standard (MAAQS), (DOE, 2015)

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>Averaging Time</th>
<th>Malaysian Ambient Air Quality Standard (MAAQS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ppm ppb</td>
</tr>
<tr>
<td>Ozone</td>
<td>1 Hour</td>
<td>0.10 200</td>
</tr>
<tr>
<td></td>
<td>8 Hour</td>
<td>0.06 120</td>
</tr>
<tr>
<td>Carbon Monoxide</td>
<td>1 Hour</td>
<td>30.0 35**</td>
</tr>
<tr>
<td></td>
<td>8 Hour</td>
<td>9.0 10**</td>
</tr>
<tr>
<td>Nitrogen Dioxide</td>
<td>1 Hour</td>
<td>0.07 320</td>
</tr>
<tr>
<td></td>
<td>24 Hour</td>
<td>0.04</td>
</tr>
<tr>
<td>Sulphur Dioxide</td>
<td>1 Hour</td>
<td>0.13 350</td>
</tr>
<tr>
<td></td>
<td>24 Hour</td>
<td>0.04 105</td>
</tr>
<tr>
<td>Particulate Matter (PM10)</td>
<td>24 Hour</td>
<td>0.13 350</td>
</tr>
<tr>
<td></td>
<td>12 Month</td>
<td>0.04 105</td>
</tr>
<tr>
<td>Total Suspended Particulate (TSP)</td>
<td>24 Hour</td>
<td>0.13 350</td>
</tr>
<tr>
<td></td>
<td>12 Month</td>
<td>0.04 105</td>
</tr>
</tbody>
</table>

Table 3 Air Pollution Index (API), (DOE 2015)

<table>
<thead>
<tr>
<th>API</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-50</td>
<td>Good</td>
</tr>
<tr>
<td>51-100</td>
<td>Moderate</td>
</tr>
<tr>
<td>101-200</td>
<td>Unhealthy</td>
</tr>
<tr>
<td>201-300</td>
<td>Very</td>
</tr>
<tr>
<td>&gt;300</td>
<td>Unhealthy</td>
</tr>
<tr>
<td>&gt;500</td>
<td>Emergency</td>
</tr>
</tbody>
</table>
The results of different studies utilizing the remote sensing data and GIS environment in detecting air quality in Malaysia are compared. Different techniques and algorithms are used in the respective studies yielding different results.

Comparison of Air Remote Sensing Application in Air Quality Monitoring of Malaysia
The most recent application of remote sensing detection of aerosol is done by Kanniah et al. (2013) whereby the properties of aerosol in Malaysia is investigated using remote sensing as primary data source from 2000 to 2009. A significant correlation of $R=0.55$, $p<0.05$ was found between the obtained MODIS AOD and the AERONET AOD. This gives some level of confidence in using the MODIS level 2 data in determining the aerosol loading in Peninsula Malaysia.

Yap & Hashim (2013) used long term observation of MODIS AOD and ground stations’ PM10 data to develop a robust algorithm to predict PM10 using MODIS AOD. The more robust mixed effects model successfully increased the correlation of linear regression model between PM10 from ground stations and MODIS AOD from 0.60 with RMSE ±12.9 to 0.88 with RMSE of ±7.32. The mixed effects model used in this study assumed that AOD-PM relationship is influenced by time-varying parameters such as relative humidity, pm vertical and diurnal concentration profiles and PM optical properties. It is hypothesized in this study that time-varying parameters exhibit a certain pattern in Peninsular Malaysia due to the climate factor. The mixed effects model used a monthly input parameter to statistically estimate the monthly site specific random error, which are affected by time-varying parameter. The random error was a parameter generated to be applied in the mixed effects model to predict PM10 concentration in the study region and its spatial and temporal distribution across Malaysia are shown in figure 3. The negative (red) region represents the overestimated value from aod that needed to be trimmed down. The positive (blue) region indicates an underestimation of the MODIS AOD, so that an enhancement is required. Overestimation of MODIS AOD may be due to the effect of unscreened clouds in the MODIS cloud screening algorithm, bright surface condition (mostly found in developed region), effects of the monsoonal winds and natural multiple scattering effect of the atmospheric particulate matter. Underestimation of AOD occurred primarily during the intermonsoon season and in rural areas. During the intermonsoon period, most of the pollutants recorded originated locally due to stagnant wind condition. Lower pollution level in rural areas with stagnant wind condition caused the AOD to be underestimated well below the apparent value.

Yap, Hashim et al. (2011) studied the relationship of MODIS estimated PM10 data against the PM10 data from the air quality monitoring stations in Peninsula Malaysia from 2000 to 2006. Unlike in the mixed effect model used by Yap & Hashim (2013), this model doesn’t consider time-varying parameters in the calculation, but only spatial variation of the correlation coefficient. Linear regression was used to retrieve the PM10 concentrations in Malaysia from MODIS AOD. Monthly average shows higher correlation of $R= 0.6$ with 12.90 μg/m3 RMSE for 2000 to 2006. In Figure 4, Yap, Hashim et al. (2011) mapped the variation of correlation of MODIS AOD in PM10 monitoring across Malaysia. Using the correlation coefficient obtained from the MODIS AOD and PM10 relationship analysis, the estimated pm10 concentration is computed. The spatial distribution of the estimated PM10 concentration is shown in figure 5.
Amanollahi, Abdullah et al. (2011a) used MODIS AOT data to study haze and PM10 concentration over Klang Valley between 2004 and 2006. The non-linear correlation coefficient (NLCC) were applied to study the relationship between MODIS AOT and hourly PM10 data from eight monitoring stations around Klang Valley. Highest $R^2$ value was recorded in Kajang (0.66) while the lowest in Petaling Jaya (0.37). Figure 6 shows AOD retrievals in Klang Valley before, during and after the haze occurrence. The red pixels during the haze occurrence indicates a very high reading of MODIS AOD, while AOD’s readings outside of haze period are generally low.

In their other study, Amanollahi, Abdullah et al. (2011B) compared daily MODIS AOD data from 2004 with PM10 measured at three ground stations (Victoria Kl, Cheras KL and Gombak) using NLCC with polynomial equation. The NLCC between MODIS AOD and PM10 concentration is higher in Victoria Kl ($R^2$=0.594) than in Gombak ($R^2$=0.359) and Cheras KL ($R^2$=0.332). The PM10 data are also imported in the GIS environment to produce PM10 map of Kuala Lumpur. Figure 7 shows the PM10 concentration map in Klang Valley on a weekday (Monday) and a weekend (Saturday). The high traffic counts during weekdays (Monday) may be the factor of the higher reading on Monday, as compared to on Saturday. PM Concentration in Kuala Lumpur is mostly influenced by motor vehicles and industry during the non-haze period.

Ibrahim & Kamarulzaman (2012) used the Landsat 7-ETM images to retrieve the aerosol optical depth in Johor Bahru by comparing the image on a ‘polluted’ and a ‘clear’ day. PM10 has a relatively high correlation of 0.63.

In 2008, Lim et al. (2008) retrieved the aerosol measurement from Penang Island, Malaysia using SPOT 5 satellite images. A very good correlation was found between the predicted PM2.5 data using raw SPOT5 image of Penang and the ground truth PM2.5 data recorded using the DustTrak Meter. GPS instrument was applied to provide the exact location. Figure 8 shows the map of PM2.5 in Pulau Pinang as the result of the study.

Kanniah et al. (2010) used MODIS AOD from 2000 to 2009 to study the spatio-temporal pattern of AOD in Peninsula Malaysia. Figure 9 shows the distribution of AOD value across Peninsula Malaysia in 20 June 2009. Note the black area without data which is due to cloud coverage which is a TYPICAL problem for remote sensing data retrieval.

Ahmad et al. (2006) manipulated the remote sensing image from NOAA-14 AVHRR to predict the PM10 in Peninsula Malaysia. A relatively good correlation was achieved with $R= 0.5563$ and average RMSE of 33. The resulting air quality map of Peninsula Malaysia can be seen in Figure 10. Ahmad et al. (2006) then used GIS interpolation to produce the iso-lines PM10 map in Malaysia using data from ground stations as shown in Figure 11.
<table>
<thead>
<tr>
<th>Study/Research</th>
<th>Method/Technique</th>
<th>Study Period</th>
<th>Relationship/Correlation</th>
<th>Findings/Results</th>
</tr>
</thead>
</table>
| Kanniah, Yaso et al. (2013) | 5x5 window averaging AOT | 1998 - 2012 | Linear correlation (LC) of $R=0.55$, $p<0.05$ between MODIS AOD and AERONET AOD around Peninsula Malaysia | Maximum AOD during dry season (June to Sept) and intermonsoon (Oct)  
• High FMF in June, Sept - Oct |
| Yap and Hashim (2013) | 5x5 window averaging AOT with mixed effect model | 2001 - 2006 | LC of $R=0.88$ with RMSE of $\pm 7.32 \mu g/m^3$ using the Mixed Effects Model applied in Peninsula Malaysia | Improvement of correlation using the mixed effect model from $R(\text{linear})=0.6$ to $R(\text{mixed})=0.88$ |
| Yap, Hashim et al. (2011) | 5x5 window averaging AOT | 2001 – 2006 | LC of $R=0.6$ with 12.90 μg/m3 RMSE for 2000 to 2006 between PM10 from ground stations and MODIS AOD in Peninsula Malaysia | Predicted PM10 data using MODIS AOD is very good and consistent |
| Amanollahi, Abdullah et al. (2011a) | 5x5 window averaging AOT & GIS Mapping | Before, during & after haze August 2005 | Non-Linear Correlation Coefficient (NLCC) of $R^2=0.66$ in Kajang (highest), $R^2=0.37$ in Petaling Jaya (Lowest). Strong correlation during dry season compared to wet. | Haze has caused AOT to increase 3 times than normal reading. |
| Amanollahi, Abdullah et al. (2011b) | 5x5 window averaging AOT & GIS Mapping | 2004 to 2005 | NLCC of $R^2=0.594$ in Victoria KL, $R^2=0.332$ in Cheras KL and $R^2=0.359$ in Gombak KL | Relationship between AOT and PM10 in three stations. |
| Ibrahim and Kamarulzaman (2012) | Apparent Reflectance & Blurring effect | 2 April (clear) & 28 Jan (polluted) 2002 & 2005 | LC of $R(\text{PM10})=0.63$ between predicted PM10 using Landsat 7-ETM and ground data in Johor Bahru  
LC of $R(\text{CO})=0.817$, $R$ of $R(\text{BP})=0.65$ | Carbon Monoxide, PM10 and Black particles concentration can be determined using Landsat 7 ETM image with good results |
| Lim, MatJafri et al. (2008) | ATCOR2 and handheld Dustrak meter | 31 Jan 2006 | LC of $R=0.9288$ between predicted PM2.5 data using SPOT5 image and ground truth PM2.5 data in Penang | Very good high spatial PM10 prediction map using SPOT5 satellite image |
| Ahmad, Hashim et al. (2006) | GIS interpolation and mapping | August to September 1997 | LC of $R=0.5563$ between predicted PM10 data using NOAA AVHRR satellite and ground PM10 data. Average RMSE is 33. | Haze and air quality distribution mapping using satellite data |
| Kanniah and Yaso (2010) | Single pixel extraction | 2000 - 2009 | Highest average MODIS AOD of 0.56±0.16 in October (dry season). Lowest is 0.15±0.02 during the rainy season. | • Higher AOD reading in west coast Peninsula  
• High AOD reading during dry seasons |
Figure 3 Maps of seasonal distribution of random error (a) intermonsoon (b) northeast monsoon, (c) southwest monsoon using MODIS data (Yap and Hashim, 2013)
Figure 4 Correlation of MODIS AOD with ground stations PM10 monitoring across Peninsula Malaysia. (Yap, Hashim et al., 2011)

Figure 5 Map of MODIS estimated PM10 concentration. (Hashim et al. 2011)
Figure 6 Haze in Klang Valley and surrounding area on MODIS image. (a) AOT value on 22nd July 2005 before haze. (b) AOT value on 8th August 2005 during the haze. (c) AOT value on 21st August 2005 after the haze (Amanollahi, Abdullah et al., 2011a)

Figure 7 Mapping of PM10 concentration in Kuala Lumpur on Monday (left) and on Saturday (right). (Amanollahi, Abdullah et al., 2011b)
Figure 8 Maps of PM2.5 in Penang on 31 Jan 2006 using SPOT 5 image. (Lim et al. 2008)

Figure 9 MODIS level 2 AOD product dated 20 June 2009. (Kanniah et al. 2010)
Figure 10 Air Quality Index in Peninsula Malaysia. (Ahmad et al. 2006)

Figure 11 ISO-Lines of PM10 for 10/11 August 2005 using GIS interpolation in Peninsula Malaysia. (Ahmad et al. 2006)
Conclusion
The relationship between remotely sensed data for air quality in Malaysia is in relatively good correlation with data collected from ground stations with R from 0.5 to 0.9. Highest correlation between remote sensing data and ground truth was achieved using the SPOT5 satellite image. The average correlation achieved using MODIS AOD is between 0.5 and 0.6. Landsat satellite image managed a good correlation of 0.6 with the ground truths data. While NOAA AVHRR satellite has a correlation of 0.55 against the ground truth data. The article outlines two main issues in acquiring a high accuracy of air quality prediction in Malaysia. First, most of the issues stems from the lack of ground truths used during the regression analysis which leads to prediction coefficient which are average. The second issue is cloud cover. This affects the amount of days available for the modelling which also leads to an average prediction of air quality. The mixed effect models proposed by Yap & Hashim (2013) tremendously increases the correlation might be a step in the right direction FOR air quality prediction using remote sensing in Malaysia. GIS tools are also important in manipulating the already available data in order to improve the quality of the results.

References


