



A robust calibration approach for PM₁₀ prediction from MODIS aerosol optical depth

X. Q. Yap and M. Hashim

Institute of Geospatial Science & Technology (INSTeG), Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor Bahru, Malaysia

Correspondence to: M. Hashim (mazlanhashim@utm.my)

Received: 31 October 2012 – Published in Atmos. Chem. Phys. Discuss.: 6 December 2012

Revised: 28 February 2013 – Accepted: 12 March 2013 – Published: 27 March 2013

Abstract. Investigating the human health effects of atmospheric particulate matter (PM) using satellite data are gaining more attention due to their wide spatial coverage and temporal advantages. Such epidemiological studies are, however, susceptible to bias errors and resulted in poor predictive output in some locations. Current methods calibrate aerosol optical depth (AOD) retrieved from MODIS to further predict PM. The recent satellite-based AOD calibration uses a mixed effects model to predict location-specific PM on a daily basis. The shortcomings of this daily AOD calibration are for areas of high probability of persistent cloud cover throughout the year such as in the humid tropical region along the equatorial belt. Contaminated pixels due to clouds causes radiometric errors in the MODIS AOD, thus causes poor predictive power on air quality. In contrary, a periodic assessment is more practical and robust especially in minimizing these cloud-related contaminations. In this paper, a simple yet robust calibration approach based on monthly AOD period is presented. We adopted the statistical fitting method with the adjustment technique to improve the predictive power of MODIS AOD. The adjustment was made based on the long-term observation (2001–2006) of PM₁₀-AOD residual error characteristic. Furthermore, we also incorporated the ground PM measurement into the model as a weighting to reduce the bias of the MODIS-derived AOD value. Results indicated that this robust approach with monthly AOD calibration reported an improved average accuracy of PM₁₀ retrieval from MODIS data by 50 % compared to widely used calibration methods based on linear regression models, in addition to enabling further spatial patterns of periodic PM exposure to be undertaken.

1 Introduction

The interest in using earth observation satellites to measure atmospheric aerosols has progressed from climate studies to the more important topic of human health. This is due to a satellite's unique ability of providing a synoptic view over large areas in a uniform, repetitive and quantitative way. Atmospheric aerosols originate from both natural and anthropogenic emission sources. The latter are considered to have major implications on human health as they are highly related to mortality and morbidity as already shown by many researchers around the world (Wan Mahiyuddin et al., 2013; Sahani et al., 2011; Bell et al., 2007; Dominici et al., 2006; Franklin et al., 2007; Gent et al., 2003, 2009; Schwartz et al., 1996; Slama et al., 2007; Hu, 2009). Most of the recent studies highlighted PM_{2.5} as the main contributor towards health effects. However, PM_{2.5} is a portion of PM₁₀ and can be estimated with a known constant (Marcazzan et al., 2001). In many developing countries PM₁₀ is still being measured instead of PM_{2.5} due to limited resources. For example, in Malaysia, PM₁₀ is being measured and used in the Air Pollution Index (API) to assess regional air quality.

Satellite data can be used as a surrogate to monitor regional air quality due to the fact that there are limited ground monitoring stations where many regions are left unmonitored (Schaap et al., 2009; Engel-Cox et al., 2004b). The widely used method of predicting PM concentration from satellite data is by empirical analysis, where in situ PM measurement are linearly regressed with the corresponding satellite AOD. In order to improve the predictive power of the linear regression models, related parameters such as local meteorological and land use information were also used as an input into PM

prediction (Liu et al., 2009). However, these models generally predict < 60 % of the PM variability (Hoff and Christopher, 2009). The latest model developed by Lee et al. (2011) uses a mixed effect model to establish daily specific AOD–PM_{2.5} relationship and predicts the site mean PM_{2.5} concentrations with $R^2 = 0.62$. Furthermore, Lee et al. (2011) also hypothesized that the relationship between PM_{2.5} and AOD varies daily due to time-varying parameters influencing the AOD–PM relationship, such as PM vertical and diurnal concentration profiles, PM optical properties, and others. Therefore, a linear AOD–PM relationship in a long-term daily monitoring is rather limited (Yap et al., 2011), and in fact the time-varying assumption by Lee et al. (2011) that varies minimally spatially on a given day over a specific spatial scale is rarely valid for humid tropical weather over the equatorial regions, where a high probability of cloud-cover exists and is also dependent on the surroundings maritime environment. Thus it is more practical and efficient for the calibration of the satellite data to be based on a monthly basis.

The monthly calibrated satellite data are useful in improving the air pollution indicators of Environmental Performance Index (EPI) reporting in this region. The EPI is a system used to evaluate countries based on 22 performance indicators that focus on environmental issues for which governments can be held accountable (Emerson et al., 2012). Atmospheric PM derived from monthly average satellite data is one of the performance indicators used in EPI evaluation for environmental health. Without the robust calibration, errors in the datasets resulting from systematic error and local climatic effects such as the monsoon and site specific error may occur. These will lead to poor representation of PM concentration and EPI derived from satellite measurements, having consequences of misinterpretation by policy makers around the world.

In this paper, a robust calibration approach is introduced by incorporating a simple adjustment technique into a mixed effects model that is developed to predict PM concentrations using MODIS AOD monthly average datasets. The MODIS AOD is calibrated by minimizing the inherent systematic and random errors (i.e. from sensor and site specific ones) in order to improve the AOD–PM relationship. The adjustment in the mixed effects model was made based on a long-term (2001–2006) analysis of the residual bias of MODIS AOD. In addition, this mixed effects model was adjusted for site errors which accounted for time varying parameters on a monthly basis. From the literature search, there are no specific similar robust calibration approaches for satellite AOD which have been reported to date. The result of this study can provide an improved AOD–PM prediction for an EPI and PM human health exposure study as well as for the investigation of PM spatial patterns.

2 Methodology

Our study is focused on Peninsular Malaysia. The air station across Peninsular Malaysia uses Met One BAM 1020 instrument to collect in situ PM₁₀ concentration in an hourly basis and averaged into daily average. In order to calibrate the MODIS AOD data for this region, PM₁₀ was sampled at 34 air stations as shown in Fig. 1 for a period of six years (i.e. 2001 to 2006). The monthly corresponding PM₁₀ values for each of the 34 air stations were averaged out from the daily PM₁₀ concentration measurements. On the other hand, the calibration on MODIS AOD data was done by using the in situ PM₁₀ measurements. Here, the calibration was performed independently for each monitoring site using multiple regression method to identify the random error to be included into the mixed effects model. Thus, this accounts for the spatial variability of the random errors on a monthly basis. After that, a single monthly AOD–PM₁₀ relationship was established using all the parameters from the 31 monitoring stations. The predicted PM₁₀ concentration from this method was validated independently in three sites, namely, in the northern, central and southern part of Peninsular Malaysia.

2.1 MODIS derived AOD

MODIS is a space sensor aboard NASA's (National Aeronautics and Space Administration) Terra Earth Observing System (EOS) satellite launched in December 1999. Operating at an altitude of approximately 700 km, this polar-orbiting satellite is able to provide aerosol data on a daily basis. MODIS Terra satellite crosses the equator at about 10.30 a.m. (descending orbit) UTC, with a scanning swath of 2330 km (cross-track) by 10 km (along-track at nadir). MODIS has a total of 36 different wavelength channels suited for a wide range of applications. AOD was retrieved by using the second generation operational algorithm (Collection 5) developed by Levy et al. (2009). In general, seven out of 36 wavelength channels (between 0.47 and 2.12 μm) are used during the AOD retrieval.

According to the MODIS AOD retrieval algorithm (Collection 5) by Levy et al. (2009), three different channels of 0.47, 0.66, and 2.12 μm are primarily employed for land aerosol retrievals, while others are used to screen out cloud, snow cover, and ice cover. The reported AOD by MODIS at the wavelength of 0.55 μm is the result of simultaneous inversion from these 3 channels. The accuracy of MODIS AOD data is expected to be $\Delta\text{AOD} = \pm 0.05 \pm 0.15$ AOD over land. More details about the retrieval of MODIS satellite aerosol data are reported in Remer et al. (2005) and Levy et al. (2007, 2009, 2010). The MODIS AOD value ranged from 5.0 to -0.05 . In this study, the full range of MODIS AOD value were taken into consideration to avoid any biases that may occur during the calibration.

To conduct this study, level 2 MODIS Terra AOD product (MOD04) data were collected for a period of six years

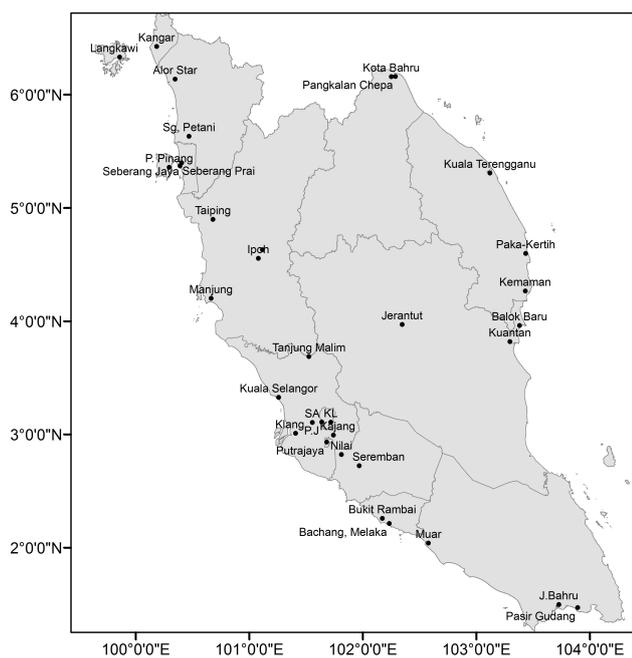


Fig. 1. Spatial distributions of 34 monitoring sites to be used in this study. (SA: Shah Alam; KL: Kuala Lumpur; J. Bahru: Johor Bahru; PJ: Petaling Jaya).

(2001 to 2006). However, aerosol data are often missing due to clouds, high surface reflectance (e.g. snow- and ice-cover), and retrieval errors. For Malaysian climatic conditions, cloud cover is a serious issue that causes failure in AOD retrieval by MODIS in most of the region.

To overcome this problem, an averaging algorithm of a 5×5 window was used (Yap et al., 2011). The algorithm assumes that the neighbouring 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. On the other hand, 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.

2.2 Mixed effects model

Recent work by Lee et al. (2011) states that AOD–PM relationship is influenced by time-varying parameters such as relative humidity, PM vertical and diurnal concentration profiles, and PM optical properties. Thus Lee et al. (2011) developed a mixed effect model which allows for day to day variability with a hypothesis of little spatial variability over the study region. In this study, a monthly observation is performed.

The mixed effects model proposed in this study, therefore, uses a monthly input parameter. Here, we hypothesize that the time-varying parameter exhibits a certain pattern in Peninsular Malaysia as a result of the peninsula's climate. Therefore, the monthly site specific spatial variability error, which is affected by the time-varying parameter, are statistically estimated from the AOD–PM₁₀ relationship and plotted to characterize its overall pattern across Peninsular Malaysia. This site specific spatial variability error is then further included in the mixed effects model to predict PM₁₀ concentrations of the study region.

The mixed effects model used to predict PM₁₀ concentration is summarized by the following equation:

$$E(Y)_{mn} = \alpha_{\text{fix}} + \beta_{\text{fix}} (\text{AOD}_{mn} - [\varepsilon_{mn} + \varepsilon_{\text{fix}}]), \quad (1)$$

where

$$\varepsilon_{\text{fix}} = \alpha - \beta(\text{AOD}_{mn}), \quad (2)$$

where $E(Y)_{mn}$ is the estimated PM₁₀ concentration in month m , at site n ; AOD_{mn} is the MODIS AOD value in the grid cell corresponding to month m , at site n ; α_{fix} and β_{fix} is the intercept and slope; ε_{mn} is the random error for month m , and site n ; ε_{fix} is the fix error or adjustment derived from long-term observations of MODIS AOD. Here, the fix error represents the average effect of AOD on PM₁₀ concentrations as shown in Eq. (2). The α and β in the Eq. (2) is the intercept, and slope for the relationship of the long-term observations of MODIS AOD and residual error measurement for PM₁₀ concentrations. This equation is obtained from a long-term observation of the MODIS AOD residual effects on the in situ PM₁₀ concentrations measurement. The relationship between the fixed error, ε_{fix} and AOD_{mn} is statistically significant with $R^2 = 0.653$ where intercept: $\alpha = 0.214$ [(SE = 0.00379), $p < 0.0001$] and slope: $\beta = 0.653$ [(SE = 0.0105), $p < 0.0001$]. These observations showed that there is a linear pattern of error in the AOD data where the error is directly proportional to the AOD data. Therefore, we added a constant, ε_{fix} , that is derived from this observation to minimize this error.

On the other hand, the random error represents the monthly site bias in the AOD–PM₁₀ relationship. The site bias may arise since an AOD value in a 10×10 km grid cell is an average optical depth in the given grid cell, while the PM₁₀ concentrations measured at a given site may not be representative

of the whole grid cell (Lee et al., 2011). In short, it represents the bias due to their spatial locations and meteorological condition in relation to the surrounding attribute. Therefore, the site bias is different for every location. To control for this site bias, we added a site term as a random error into the statistical model (Lee et al., 2011). The bias value was computed from ground measurements and interpolated to represent the approximate ground conditions. From the monthly observations, the spatial pattern of the site bias exhibit three general patterns due to the meteorological conditions, i.e. the monsoon effect. From here, we average the spatial distribution of the site bias (random error) according to the monsoon period. Once this parameter has been entered into the mixed effects model, the PM₁₀ concentration was estimated for the whole study area using the MODIS AOD.

2.3 Model validation

This model is analyzed throughout Peninsular Malaysia by using a cross-validation (CV) method to examine whether the mixed effects model is applicable to our study region. There are a total of 31 sampling sites which were used in establishing the model and three independent sampling sites were used to validate the model. From the 31 sampling sites, a mixed effects model was developed to predict PM₁₀ in Peninsular Malaysia. To assess the relationship between the predicted and measured PM₁₀ concentrations for each site, the Pearson correlation coefficients were used. A high correlation indicates that the MODIS AOD data can be used to assess human health exposure investigations and can be applied in establishing the EPI for Malaysia. In addition, root mean square error (RMSE) will be used to quantitatively assess the accuracy of final output. This validation is important to investigate the reliability and accuracy of the predicted PM₁₀ concentration to assess the spatial accuracy of the predicted PM₁₀.

2.4 PM₁₀ in Malaysia

In Malaysia, severe cases of air pollution generally affect our neighboring countries as a result of forest fire and monsoon wind (Hashim et al., 2004). This event usually occurs during southwest monsoon season between May till September, which brings haze from the Sumatra region to the western side of Peninsular Malaysia. Other local sources of air pollution include vehicle emissions, power generation, industrial emissions, open burning and forest fires (Afroz et al., 2003; Azmi et al., 2010; Dominick et al., 2012). Furthermore, west Peninsular Malaysia, where major cities are located, usually has a higher PM₁₀ concentration in comparison to other regions due to anthropogenic activities (Azmi et al., 2010).

3 Results and discussion

3.1 Descriptive statistics

The mean concentrations of PM₁₀ in our study site are summarized in Table 1. From Table 1, there are several sites that exhibit high mean (SE) PM₁₀ concentrations across Peninsular Malaysia. For example, Perai, Melaka, Kuala Selangor, Klang, KL, Shah Alam and Manjung. These sites are mainly industrialized regions that are affected by heavy traffic and seasonal haze. Surprisingly, Bukit Rambai, Melaka, has the highest mean (SE) PM₁₀ concentration at 74.9 (1.81) $\mu\text{g m}^{-3}$ followed by Klang at 72.3 (3.04) $\mu\text{g m}^{-3}$. The exceptionally high PM₁₀ concentration in Bukit Rambai is mainly a result of local anthropogenic activities as it is situated in an industrial district with a secondary impact from seasonal haze (Mahmud and Iza, 2010). The range and average number of monthly sample points, (*n*) across peninsular Malaysia was between 35 to 72 and 61, respectively. A total of 433 (17.69%) monthly samples points were discarded due to the unavailability of a corresponding point with the MODIS AOD samples.

3.2 PM₁₀ prediction

In the mixed effects model, the random error for all 72 months were generated (from year 2001 to 2006) and are summarized in Table 2. The random error was attributed to site and time varying errors. It has a seasonal pattern across Peninsular Malaysia as shown in Fig. 2. From Table 2 and Fig. 2, regions of densely developed sites have a high negative random error. This shows that these regions tend to have an overestimated AOD value. Therefore, it is necessary to include the random error into the mixed effects model to perform the adjustment. The fixed error or adjustment effect, ε_{fix} , represents the monthly effect of MODIS AOD on PM₁₀ for all study days. This constant is derived from Eq. (2).

Using the mixed effects model approach, the seasonal pattern of random error clearly shows the effects of the monsoon wind on our study region. The negative (red) region denotes the overestimated value from MODIS AOD that needed to be trimmed down. Similarly the positive (blue) region indicates an underestimation of the MODIS AOD, so that enhancement is needed. The overestimation of MODIS AOD may be due to the effect of unscreened clouds resulting from the MODIS cloud screening algorithm (Lee et al., 2011). This was also demonstrated in the work of Holben et al. (1998) where level 2 (AERONET) data were compared with MODIS AOD where unscreened clouds cause a positive bias in the predicted particulate matter concentration. Furthermore, bright surface condition may also increase the error as a result of poorer visible to infrared (2.12 μm) band relationship (Levy et al., 2009). In addition, the overestimation of the AOD may also be related to the natural multiple scattering effect of the atmospheric particulate matter (pollutant).

Table 1. Long-term (2001–2006) descriptive statistics of PM₁₀ concentration ($\mu\text{g m}^{-3}$) and corresponding MODIS AOD observed at 34 monitoring stations.

Site Location	Latitude	Longitude	<i>n</i>	PM ₁₀ ($\mu\text{g m}^{-3}$)		MODIS AOD	
				Mean	SE	Mean	SE
Pasir Gudang	1.4704	103.8940	63	48.19	1.09	0.33	0.01
Teluk Kalung, Kemaman	4.2658	103.4323	60	42.75	1.42	0.28	0.01
Taman Inderawasih, Perai	5.3711	100.3891	51	67.96	2.94	0.37	0.02
Bukit Rambai, Melaka	2.2586	102.1727	57	74.91	1.81	0.39	0.02
Jerantut	3.9706	102.3477	66	40.98	1.28	0.29	0.01
Jalan Tasik, Ipoh	4.6294	101.1166	69	51.55	1.27	0.28	0.01
Perai	5.3982	100.4039	61	62.89	1.77	0.33	0.02
Nilai	2.8216	101.8114	73	58.99	1.51	0.36	0.03
Klang	3.0100	101.4085	61	72.33	3.04	0.42	0.02
Indera Mahkota, Kuantan	3.8193	103.2965	46	35.69	1.13	0.30	0.02
Balok Baru, Kuantan	3.9607	103.3822	64	58.63	1.18	0.34	0.02
Petaling Jaya (PJ)	3.1092	101.6387	66	56.74	1.85	0.37	0.02
Sg. Petani	5.6315	100.4697	67	52.18	1.37	0.3	0.02
J.Bahru	1.4974	103.7268	57	41.98	1.45	0.31	0.01
Taiping	4.8987	100.6792	65	46.65	1.27	0.31	0.01
Pangkalan Chepa, Kota Bahru	6.1591	102.2880	61	44.61	1.34	0.31	0.01
Kota Bahru	6.1587	102.2510	63	41.74	1.35	0.30	0.01
Kajang	2.9939	101.7417	64	49.26	1.72	0.35	0.03
Paka-Kertih	4.5980	103.4349	58	34.85	0.89	0.26	0.01
Shah Alam (SA)	3.1047	101.5563	65	62.40	2.35	0.38	0.02
Langkawi	6.3316	99.8583	57	41.18	1.25	0.25	0.01
Kangar	6.4240	100.1841	63	50.20	1.39	0.32	0.02
Kuala Terengganu	5.3076	103.1202	61	55.60	1.32	0.29	0.01
P. Pinang	5.3575	100.2944	60	41.50	1.41	0.31	0.02
Alor Star	6.1372	100.3466	45	37.03	1.21	0.28	0.02
Manjung	4.2003	100.6633	69	64.97	2.16	0.36	0.02
Bachang	2.2131	102.2343	62	45.02	1.73	0.33	0.02
Muar, Johor	2.0397	102.5769	69	54.57	1.79	0.37	0.02
Tanjung Malim	3.6878	101.5244	70	44.21	1.29	0.31	0.01
Pegoh 4, Ipoh	4.5533	101.0802	69	50.52	1.37	0.30	0.01
Seremban	2.7236	101.9684	71	46.28	1.33	0.32	0.02
Kuala Selangor	3.3265	101.2589	66	65.94	2.08	0.37	0.02
W.P Putrajaya	2.9319	101.6818	44	52.12	2.78	0.38	0.04
W.P K. Lumpur (KL)	3.1062	101.7178	35	60.77	3.11	0.42	0.04

SE: Standard error.

n: Number of monthly samples points.

Thus most of the well-developed regions (having bright surface) tend to be overestimated. In Fig. 2, the effect of the monsoon wind was clear as it drifted the overestimated region further inland towards east Peninsular Malaysia. In contrast, the occurrence of the underestimated MODIS AOD values was common only during the intermonsoon season and in some rural areas. The intermonsoon is the interval when a change of monsoon wind direction occurs. During this period, most of the atmospheric particulate matter concentrations recorded originated from local anthropogenic activities due to the stagnant wind condition. Therefore, the underestimated MODIS AOD value at this period could be due to lower pollution levels in that particular area, compared to its

surroundings that resulted in a plunge of offset in the observed MODIS AOD value below the apparent value.

The predicted PM₁₀ concentrations from the mixed effects model are also prone to errors attributed from the difference in AOD retrieval and in situ measurements of the PM₁₀ concentrations. This is due to the fact that the in situ measurements were point measurements, whilst the AOD was based on 10×10 km grid cells. However, this error was not taken into account due to the fact that the in situ measurements were a 24 h average. Here, if the surrounding (within 10×10 km) PM₁₀ concentration of a particular station was to represent the 10×10 km grid cell, it would most probably have been measured by the monitoring station within 24 h. Thus the 24 h in situ PM₁₀ concentrations averaged

Table 2. Estimation of long-term (2001–2006) monthly mean random error (AOD) for 34 monitoring stations.

Site location	Latitude	Longitude	Mean random error	<i>p</i> -value
Pasir Gudang	1.4704	103.8940	0.0266	
Teluk Kalung, Kemaman	4.2658	103.4323	0.0017	
Taman Inderawasih, Perai	5.3711	100.3891	−0.0362	
Bukit Rambai, Melaka	2.2586	102.1727	−0.0692	
Jerantut	3.9706	102.3477	0.0168	
Jalan Tasik, Ipoh	4.6294	101.1166	−0.0348	
Perai	5.3982	100.4039	−0.0539	
Nilai	2.8216	101.8114	−0.0166	
Klang	3.0100	101.4085	−0.0091	
Indera Mahkota, Kuantan	3.8193	103.2965	0.0521	
Balok Baru, Kuantan	3.9607	103.3822	−0.0290	
Petaling Jaya (PJ)	3.1092	101.6387	0.0176	
Sg. Petani	5.6315	100.4697	−0.0302	
J.Bahru	1.4974	103.7268	0.0234	
Taiping	4.8987	100.6792	0.0046	
Pangkalan Chepa, Kota Bahru	6.1591	102.2880	0.0217	<i>p</i> <0.0001
Kota Bahru	6.1587	102.2510	0.0273	
Kajang	2.9939	101.7417	0.0324	
Paka-Kertih	4.5980	103.4349	0.0279	
Shah Alam (SA)	3.1047	101.5563	−0.0051	
Langkawi	6.3316	99.8583	−0.0264	
Kangar	6.4240	100.1841	0.0054	
Kuala Terengganu	5.3076	103.1202	−0.0652	
P. Pinang	5.3575	100.2944	0.0443	
Alor Star	6.1372	100.3466	0.0284	
Manjung	4.2003	100.6633	−0.0423	
Bachang	2.2131	102.2343	0.0440	
Muar, Johor	2.0397	102.5769	0.0329	
Tanjung Malim	3.6878	101.5244	0.0265	
Pegoh 4, Ipoh	4.5533	101.0802	−0.0254	
Seremban	2.7236	101.9684	0.0213	
Kuala Selangor	3.3265	101.2589	−0.0369	
W.P Putrajaya	2.9319	101.6818	0.0379	
W.P K. Lumpur (KL)	3.1062	101.7178	0.0269	

to represent the 10 × 10 km grid cell, would most probably resemble the predicted PM₁₀ concentration from MODIS AOD. Furthermore, the comparison of a 10 × 10 km grid cell with a point measurement was a common practice among researchers such as Chu et al. (2003), and Koelemeijer et al. (2006). However, for a monitoring station that is close to the pollution source such as Bukit Rambai, Melaka, the random error would appear higher due to the point measurement as it does not represent the 10 × 10 km² grid cell. Therefore, it is important to avoid sampling in close proximity to a pollution source, when the aim is to compare it to a large grid cell.

3.3 Accuracy assessment

In order to examine the accuracy of the predicted monthly PM₁₀ concentrations, the monthly in situ PM₁₀ concentrations and the monthly predicted PM₁₀ concentrations were regressed as shown in Fig. 3. The mixed effects model ex-

plained 77 % of the variability in the monthly measured PM₁₀ concentration for a period of six years (i.e. 2001 till 2006). From Fig. 3, the relationship of the predicted PM₁₀ concentration using the mixed effects model approximate the ground condition [slope = 1; intercept = 2 × 10^{−05}; *n* = 1895, *p* < 0.0001]. Furthermore, validation from three independent ground stations (i.e. Johor Bahru, Shah Alam and Pengkalan Chepa which are situated in southern, central and northern parts of Peninsular Malaysia) that were chosen to assess the predicted PM₁₀ concentration also show a promising result when regressed with the measured PM₁₀ concentrations [slope = 1.085; intercept = −5.515; *n* = 181; *p* < 0.0001] (Fig. 4). This slope presented in Fig. 4 shows that the predicted PM₁₀ concentration had a high agreement with the in situ PM₁₀ concentration measurement and the intercept represent the noise in the predicted PM₁₀ concentration dataset which is considerably lower.

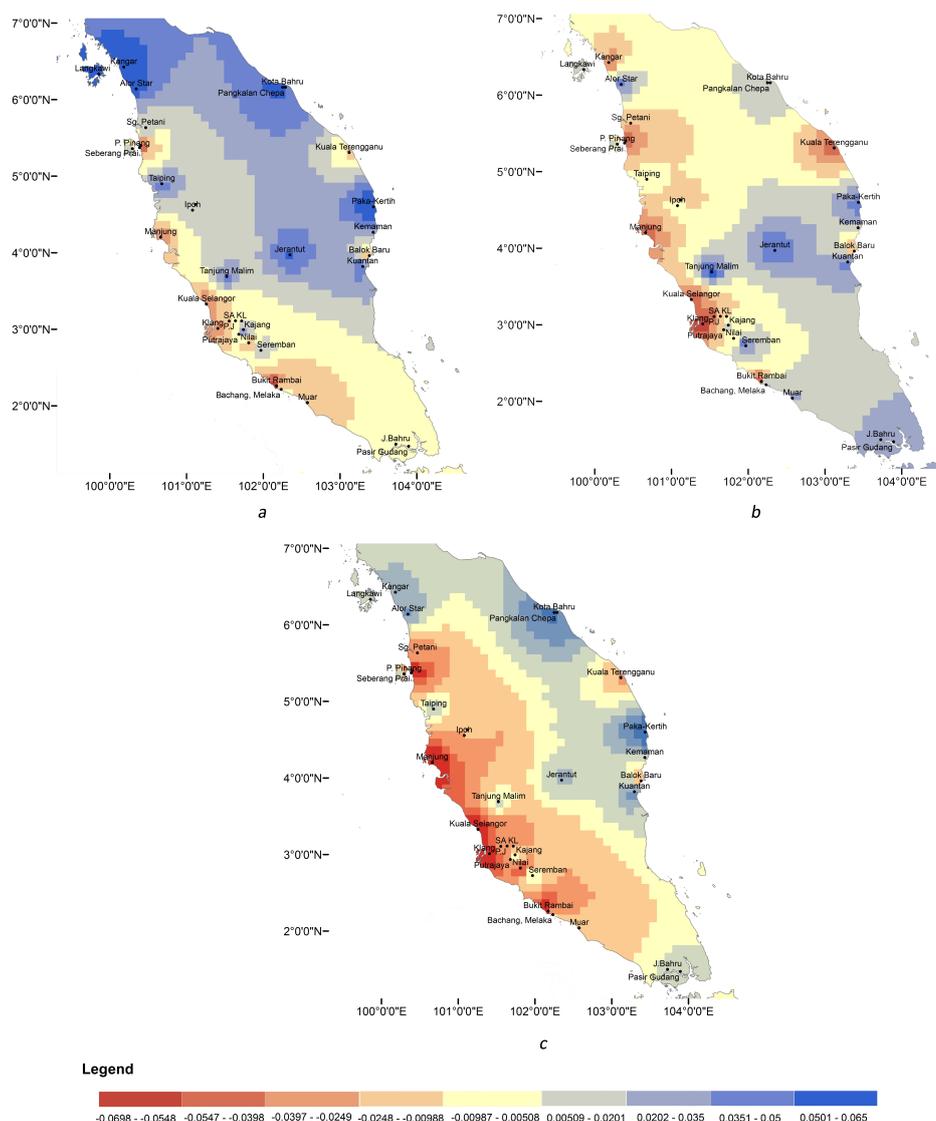


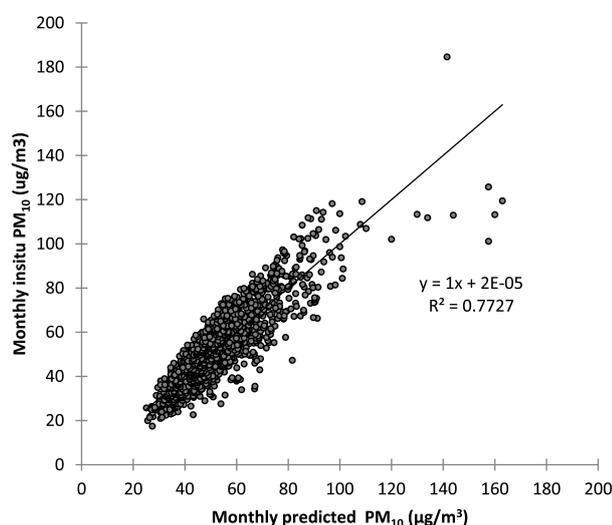
Fig. 2. Seasonal distribution of random error (a) intermonsoon, (b) northeast monsoon, (c) southwest monsoon season.

Furthermore, the ability of the mixed effects model to predict the PM₁₀ concentration was compared to a linear regression model by using Pearson correlation, R , and root mean square error, RMSE (Table 3). The linear regression model has been widely used by many researcher (Chu et al., 2003; Engel-Cox et al., 2004a,b, 2005, 2006) to establish the AOD–PM₁₀ or 2.5 relationship, and therefore is regarded as a common and valid methodology to predict particulate matters of different sizes (10 μm and 2.5 μm in diameter). Since the R does not quantitatively reflect the difference between the measured and predicted PM₁₀ concentrations, RMSE is necessary to better assess both models. In Table 3, the performance of the mixed effects model has significantly improved the accuracy of the predicted PM₁₀, compared to the linear regression model. Overall, the long-term Pearson cor-

relation, R , of predicted PM₁₀ concentration has improved from 0.60 to 0.88 using the mixed effects model. Similarly, the RMSE of the predicted PM₁₀ concentration of the mixed effects model improvised the linear regression model by an average of $\pm 6.18 \mu\text{g m}^{-3}$ annually. In other words, the accuracy of the mixed effects model was superior and has improved approximately 50% compared to the conventional linear regression model. This was further confirmed by the ANOVA test (p value ≈ 1) which suggests that the predicted PM₁₀ concentration using our method is in high agreement with in situ measurements. From this performance test, the mixed effects model appeared to be a better solution in producing a reliable concentration map for both environmental and health effect studies.

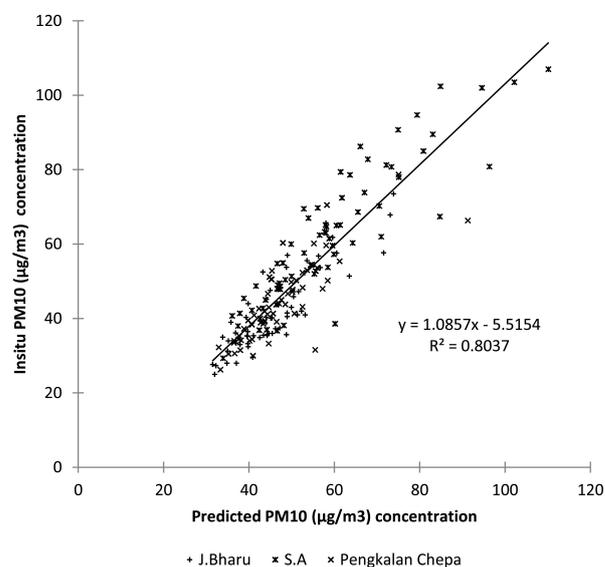
Table 3. Long-term comparison on linear regression model and mixed effects model on PM₁₀-AOD and RMSE ($\mu\text{g m}^{-3}$) of annual MODIS estimated PM₁₀ concentration.

Year	Linear regression model				Mixed effects model			
	<i>R</i>	<i>n</i>	RMSE ($\mu\text{g m}^{-3}$)	<i>p</i>	<i>R</i>	<i>n</i>	RMSE ($\mu\text{g m}^{-3}$)	<i>p</i>
2001	0.60	315	± 13.01	<i>p</i> < 0.0001	0.89	315	± 6.25	<i>p</i> < 0.0001
2002	0.60	325	± 15.36		0.89	325	± 8.57	
2003	0.63	322	± 14.30		0.90	322	± 7.19	
2004	0.62	367	± 13.38		0.90	367	± 7.61	
2005	0.60	366	± 12.27		0.88	366	± 7.42	
2006	0.66	381	± 12.52		0.91	381	± 6.75	
2001–2006	0.60	2076	± 12.90		0.88	2076	± 7.32	

**Fig. 3.** Assessment of the monthly in situ PM₁₀ measurement and monthly predicted PM₁₀ by a mixed effects model.

4 Conclusions

To date, there has been an increase in the adoption of satellite AOD data into air pollution, health effects and environmental studies. The awareness of the potential of remote sensing technologies to enhance ground-level particulate matters monitoring networks has further encouraged the many government and private agencies to look into its practicality. In Malaysia, the used of satellite derived parameters as performance indicators in EPI is one of the highlights to bring forward these technologies. However, the application of satellite data has always been received with skepticism in this region due to cloud cover and low predictive power. The proposed mixed effects model suggested in this paper has shown that this calibration method can be reliable in producing a better PM₁₀ concentration map for this region. Taking into account the site specific random error and the fixed errors, the accuracy of the satellite data improve significantly.

**Fig. 4.** Assessment of the monthly in situ PM₁₀ measurement and monthly predicted PM₁₀ from three independent monitoring sites.

Next, we anticipate that the outcome of this method will be increasingly used for health effects, pollution and environmental related studies. Future satellite technologies are expected to improve spatial and temporal resolutions in the near future, resulting in an even more accurate retrieval method. As the satellite data are readily available, monitoring and predicting atmospheric pollution such as PM₁₀ can be made in a cost-effective way. Another focus of our future research will be to study atmospheric particulate matter and other atmospheric trace gases that are harmful to human health.

Acknowledgements. This study was conducted as part of the Research University Grant of Universiti Teknologi Malaysia (UTM). We would like to thank Zamalah Scholarship from UTM in financing the graduate student involved in this study. We are also indebted to Mohd Nadzri Reba for his valuable comments on the data analysis. Comments from referees for improving the manuscript is highly acknowledged.

Edited by: D. Tanre

References

- Afroz, R., Hassan, M. N., and Ibrahim, N. A.: Review of air pollution and health impacts in Malaysia, *Environ. Res.*, 92, 71–77, doi:10.1016/S0013-9351(02)00059-2, 2003.
- Azmi, S. Z., Latif, M. T., Ismail, A. S., Juneng, L., and Jemain, A. A.: Trend and status of air quality at three different monitoring stations in the Klang Valley, Malaysia, *Air Qual. Atmos. Health*, 3, 53–64, 2010.
- Bell, M. L., Ebisu, K., and Belanger, K.: Ambient air pollution and low birth weight in Connecticut and Massachusetts, *Environ. Health Persp.*, 115, 1118–1124, 2007.
- Chu, D. A., Kaufman, Y. J., Zibordi, G., Chern, J. D., Mao, J., Li, C. C., and Holben, B. N.: Global monitoring of air pollution over land from the Earth observing system-terra moderate resolution imaging spectroradiometer (MODIS), *J. Geophys. Res.-Atmos.*, 108, 4661, doi:10.1029/2002jd003179, 2003.
- Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., and Samet, J. M.: Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases, *J. Am. Med. Assoc.*, 295, 1127–1134, 2006.
- Dominick, D., Juahir, H., Latif, M. T., Zain, S. M., and Aris, A. Z.: Spatial Assessment of air quality patterns in Malaysia using multivariate analysis, *Atmos. Environ.*, 60, 172–181, doi:10.1016/j.atmosenv.2012.06.021, 2012.
- Emerson, J. W., Hsu, A., Levy, M. A., de Sherbinin, A., Mara, V., Esty, D. C., and Jaiteh, M.: Environmental Performance Index and Pilot Trend Environmental Performance Index, Yale Center for Environmental Law and Policy, New Haven, 2012.
- Engel-Cox, J. A., Hoff, R. M., and Haymet, A. D. J.: Recommendations on the use of satellite remote-sensing data for urban air quality, *J. Air Waste Manage.*, 54, 1360–1371, 2004a.
- Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., and Hoff, R. M.: Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality, *Atmos. Environ.*, 38, 2495–2509, 2004b.
- Engel-Cox, J. A., Young, G. S., Hoff, R. M.: Application of satellite remote-sensing data for source analysis of fine particulate matter transport events, *J. Air Waste Manage.*, 55, 1389–1397, 2005.
- Engel-Cox, J. A., Hoff, R. M., Rogers, R., Dimmick, F., Rush, A. C., Szykman, J. J., Al-Saadi, J., Chu, D. A., and Zell, E. R.: Integrating lidar and satellite optical depth with ambient monitoring for 3-dimensional particulate characterization, *Atmos. Environ.*, 40, 8056–8067, 2006.
- Franklin, M., Zeka, A., and Schwartz, J.: Association between PM_{2.5} and all-cause and specific-cause mortality in 27 US communities, *J. Expo. Sci. Environ. Epidemiol.*, 17, 279–287, 2007.
- Gent, J. F., Triche, E. W., Holford, T. R., Belanger, K., Bracken, M. B., Beckett, W. S., and Leaderer, B. P.: Association of low-level ozone and fine particles with respiratory symptoms in children with asthma, *J. Am. Med. Assoc.*, 290, 1859–1867, 2003.
- Gent, J. F., Koutrakis, P., Belanger, K., Triche, E., Holford, T. R., Bracken, M. B., and Leaderer, B. P.: Symptoms and medication use in children with asthma and traffic-related sources of fine particle pollution, *Environ. Health Persp.*, 117, 1168–1174, 2009.
- Hashim, M., Kanniah, K. D., Ahmad, A., Rasib, A. W., and Ibrahim, A. L.: The use of AVHRR data to determine the concentration of visible and invisible tropospheric pollutants originating from a 1997 forest fire in Southeast Asia, *Int. J. Rem. Sens.*, 25, 4781–4794, 2004.
- Hoff, R. M. and Christopher, S. A.: Remote sensing of particulate pollution from space: Have we reached the promised land?, *J. Air Waste Manage.*, 59, 645–675, 2009.
- Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – a federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, 66, 1–16, 1998.
- Hu, Z.: Spatial analysis of MODIS aerosol optical depth, PM_{2.5}, and chronic coronary heart disease, *Int. J. Health Geogr.*, 8, 27, doi:10.1186/1476-072X-8-27, 2009.
- Ichoku, C., Chu, D. A., Mattoo, S., Kaufman, Y. J., Remer, L. A., Tanré, D., Slutsker, I., and Holben, B. N.: A spatio-temporal approach for global validation and analysis of MODIS aerosol products, *Geophys. Res. Lett.*, 29, 8006, doi:10.1029/2001GL013206, 2002.
- Koelemeijer, R. B. A., Homan, C. D., and Matthijsen, J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, *Atmos. Environ.*, 40, 5304–5315, 2006.
- Lee, H. J., Liu, Y., Coull, B. A., Schwartz, J., and Koutrakis, P.: A novel calibration approach of MODIS AOD data to predict PM_{2.5} concentrations, *Atmos. Chem. Phys.*, 11, 7991–8002, doi:10.5194/acp-11-7991-2011, 2011.
- Levy, R. C., Remer, L. A., Mattoo, S., Vermote, E. F., and Kaufman, Y. J.: Second-generation operational algorithm: retrieval of aerosol properties over land from inversion of moderate resolution imaging spectroradiometer spectral reflectance, *J. Geophys. Res.*, 112, D13211, doi:10.1029/2006JD007811, 2007.
- Levy, R., Remer, L., Tanré, D., Mattoo, S., and Kaufman, Y.: Algorithm for remote sensing of tropospheric aerosol over dark targets from MODIS: Collections 005 and 051: Revision 2, MODIS Algorithm Theoretical Basis Document, available at: http://modis-atmos.gsfc.nasa.gov/_docs/ATBD_MOD04_C005_rev2.pdf, 2009.
- Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F.: Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, *Atmos. Chem. Phys.*, 10, 10399–10420, doi:10.5194/acp-10-10399-2010, 2010.
- Liu, Y., Paciorek, C. J., and Koutrakis, P.: Estimating regional spatial and temporal variability of PM_{2.5} concentrations using satellite data, meteorology, and land use information, *Environ. Health Persp.*, 117, 886–892, 2009.
- Mahmud, M. and Iza, H. A. L.: Pencemaran udara di Bukit Rambai, Melaka sewaktu peristiwa jerebu tahun 2005 (Air pollution at Bukit Rambai, Melaka during the haze episode of 2005), *Geografia, Malaysian Journal of Society and Space (GMJSS)*, 6,

- 30–39, 2010.
- Marcazzan, G. M., Vaccaro, S., Valli, G., and Vecchi, R.: Characterisation of PM₁₀ and PM_{2.5} particulate matter in the ambient air of Milan (Italy), *Atmos. Environ.*, 35, 4639–4650, 2001.
- Remer, L. A., Kaufman, Y. J., Tanré, D., Mattoo, S., Chu, D. A., Martins, J. V., Li, R. R., Ichoku, C., Levy, R. C., Kleidman, R. G., Eck, T. F., Vermote, E., and Holben, B. N.: The MODIS aerosol algorithm, products, and validation, *J. Atmos. Sci.*, 62, 947–973, 2005.
- Sahani, M., Wan Mahiyuddin, W. R., Aripin, R., Latif, M. T., Thach, T. Q., and Wong, C. M.: Health Risks of Air Pollution on Mortality in Klang Valley, Malaysia, ISEE 22nd Annual Conference, Seoul, Korea, 28 August–1 September 2010: Air Pollution – Short-term Health Effects, *Epidemiology*, 22, S159, doi:10.1097/01.ede.0000392160.27116.be, 2011.
- Schaap, M., Apituley, A., Timmermans, R. M. A., Koelemeijer, R. B. A., and de Leeuw, G.: Exploring the relation between aerosol optical depth and PM_{2.5} at Cabauw, the Netherlands, *Atmos. Chem. Phys.*, 9, 909–925, doi:10.5194/acp-9-909-2009, 2009.
- Schwartz, J., Dockery, D. W., and Neas, L. M.: Is daily mortality associated specifically with fine particles?, *J. Air Waste Manage.*, 46, 927–939, 1996.
- Slama, R., Morgenstern, V., Cyrus, J., Zutavern, A., Herbarth, O., Wichmann, H. E., and Heinrich, J.: Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: a study relying on a land-use regression exposure model, *Environ. Health Persp.*, 115, 1283–1292, 2007.
- Wan Mahiyuddin W. R., Sahani, M., Aripin, R., Latif, M. T., Thach, T. Q., Wong, C. M.: Short-term effects of daily air pollution on mortality, *Atmos. Environ.*, 65, 69–79, 2013.
- Yap, X. Q., Hashim, M., and Marghany, M. M.: Retrieval of PM₁₀ concentration from Moderate Resolution Imaging Spectroradiometer (MODIS) derived AOD in Peninsular Malaysia, *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE International, Vancouver, 24–29 July, ISSN:2153-6996, 4022–4025, doi:10.1109/IGARSS.2011.6050114, 2011.