Atmospheric Research 193 (2017) 142–162

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Estimating Particulate Matter using satellite based aerosol optical depth and meteorological variables in Malaysia

Nurul Amalin Fatihah Kamarul Zaman[a], Kasturi Devi Kanniah[a,c,*], Dimitris G. Kaskaoutis[b]

[a] Tropical Map Research Group, Faculty of Geoinformation and Real Estate, Universiti Teknologi Malaysia, 81310 Skudai, Johor Darul Takzim, Malaysia
[b] Atmospheric Research Team, Institute for Environmental Research and Sustainable Development, National Observatory of Athens, GR-11810 Athens, Greece
[c] Centre for Environmental Sustainability and Water Security (IPASA), Research Institute for Sustainable Environment (RISE), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor

A R T I C L E   I N F O

Keywords:
PM_{10}
Satellite remote sensing
Artificial Neural Network
Multiple linear regressions
Meteorology
Malaysia

A B S T R A C T

The insufficient number of ground-based stations for measuring Particulate Matter < 10 μm (PM_{10}) in the developing countries hinders PM_{10} monitoring at a regional scale. The present study aims to develop empirical models for PM_{10} estimation from space over Malaysia using aerosol optical depth (AOD_{550}) and meteorological (surface temperature, relative humidity and atmospheric stability) data (retrieved or estimated) from Moderate Resolution Imaging Spectroradiometer (MODIS) during the period 2007–2011. The MODIS retrievals are found to be satisfactorily correlated with ground-based measurements at Malaysia. Multiple linear regressions (MLR) and Artificial Neural Network (ANN) techniques are utilized to develop the empirical models for PM_{10} estimation. The model development and training are performed via comparison with measured PM_{10} at 29 stations over Malaysia and reveal that the ANN provides slightly higher accuracy with $R^2 = 0.71$ and RMSE = 11.61 μg m^{-3} compared to the MLR method ($R^2 = 0.66$ and RMSE = 12.39 μg m^{-3}). Stepwise regression analysis performed on the MLR method reveals that the MODIS AOD_{550} is the most important parameter for PM_{10} estimations ($R^2 = 0.59$ and RMSE = 13.61 μg m^{-3}); however, the inclusion of the meteorological parameters in the MLR increases the accuracy of the retrievals ($R^2 = 0.66$, RMSE = 12.39 μg m^{-3}). The estimated PM_{10} concentrations are finally validated against surface measurements at 16 stations resulting in similar performance from the ANN model ($R^2 = 0.58$, RMSE = 10.16 μg m^{-3}) and MLR technique ($R^2 = 0.56$, RMSE = 10.58 μg m^{-3}). The significant accuracy that has been attained in PM_{10} estimations from space allows us to assess the pollution levels in Malaysia and map the PM_{10} distribution at large spatial and temporal scales.

1. Introduction

Air pollution has become a serious environmental problem in the developing southeast Asian countries, especially in Malaysia that is aiming to become an industrial nation by 2020. Malaysia is ranked as the 117th worst country among 180 nations worldwide in terms of air quality (EPI, 2016) with the major sources of air pollutants to be open biomass burning, motor vehicles and industries (Afroz et al., 2003; Hyer et al., 2013; Alias et al., 2014; Khan et al., 2016a). Besides the local sources, biomass-burning aerosols from wildfires in Indonesia are also transported over Malaysia particularly during the dry season and southwest monsoon (Kanniah et al., 2014a; Khan et al., 2016b). Furthermore, Malaysia lies in the main pathway of the southeast Asian pollution outflow (Lawrence and Lelieveld, 2010; Reid et al., 2013; Wang et al., 2013), which contributes significantly to the local aerosol and pollutant emissions. The air pollution levels in Malaysia are expressed via the Air Pollution Index (API) that is calculated using surface concentrations of carbon monoxide (CO), tropospheric ozone (O_3), nitrogen dioxide (NO_2), sulphur dioxide (SO_2) and Particulate Matter < 10 μm (PM_{10}) (Awang et al., 2000; Dominick et al., 2012).

PM_{10} is an important component of air-pollution monitoring networks since it affects human health by causing respiratory problems (Balakrishnan et al., 2002; Pope et al., 2011a, b; Ha Trang and Tripathi, 2014), while inhaling PM_{10}, and most particularly Particulate Matter < 2.5 μm ($PM_{2.5}$), can cause cardiovascular diseases (Dominici et al., 2006), birth defects and premature death (Ballester et al., 2010). PM_{10}, as aerosol particles, can also affect the climate system by scattering and absorbing the incoming solar radiation resulting in heating and/or cooling effects (Ramanathan and Carmichael, 2008) depending on the type of particles (i.e. dust, soot or anthropogenic pollution), optical and

* Corresponding author.
E-mail address: kasturi@utm.my (K.D. Kanniah).

http://dx.doi.org/10.1016/j.atmosres.2017.04.019
Received 22 October 2016; Received in revised form 1 April 2017; Accepted 5 April 2017
Available online 06 April 2017

chemical properties.

Satellite remote sensing has been increasingly available for PM_{10} monitoring studies over a large spatial domain. Remote-sensing retrievals of aerosol optical depth (AOD) from multiple satellite sensors, such as Multiangle Imaging Spectroradiometer (MISR) (Liu et al., 2007; Van Donkelaar. et al., 2010; Dey et al., 2012; Sotoudeheian and Arhami, 2014), Spinning Enhanced Visible and Infrared Imager (SEVIRI) (Emili et al., 2010), Moderate Resolution Imaging Spectroradiometer (MODIS) (Gupta et al., 2006; Van Donkelaar. et al., 2010; Jamil et al., 2011; Nordio et al., 2013; Yap and Hashim, 2013; Chitranshi et al., 2014), Medium Resolution Imaging Spectrometer (MERIS) (Kaskaoutis et al., 2010; Kanniah et al., 2014b; Beloconi et al., 2016) and Landsat (Nguyen and Tran, 2014; Nadzri et al., 2010) are commonly used to estimate PM_{10} and/or $PM_{2.5}$ from space.

Among the available remote-sensing data, MODIS AOD is found to have high retrieval accuracy over land (i.e. $\pm$ 0.05 $*$ AOD under clear skies and $\pm$ 0.15 $*$ AOD under moderately cloud-contaminated atmospheres), as well as nearly daily global coverage (Remer et al., 2008). Validation of MODIS AOD retrieved from both Terra and Aqua satellites showed high correlation (R^2 = 0.9) with AERONET AODs worldwide (Levy et al., 2010) and, over Malaysia (R^2 = 0.55) in particular (Kanniah et al., 2014a). Nevertheless, a main drawback for usage of MODIS AOD is its coarse spatial resolution (10 km) that may tend to generalize AOD information spatially. Therefore, especially over urban environments that are characterized by heterogeneous surfaces, satellite AOD of higher spatial resolution is preferred for PM_{10} estimations. In this respect, MERIS AOD and synergized MERIS/AATSR data (North et al., 2009) have been recently used although MERIS's relatively poor performance compared to MODIS (Kaskaoutis et al., 2010; Benas et al., 2013; Kanniah et al., 2014b). Furthermore, MODIS-AOD retrievals at 3-km spatial resolution are also considered as not so suitable for air-quality studies due to improper characterization of the underlying urban surfaces that resulted in significant bias for AODs > 0.1 (Levy et al., 2013; Munchak et al., 2014; Xie et al., 2015).

Several studies have used only the satellite AOD via simple linear regressions for PM estimations from space resulting in correlation coefficients around 0.52 to 0.77 between measured and estimated PM_{10} concentrations (Gupta and Christopher, 2008; Schaap et al., 2009; Kanniah et al., 2016). According to Gupta and Christopher (2009a, b), AOD alone is not a good predictor of PM, since AOD corresponds to the total columnar aerosol, whereas PM refers to particle concentration at the surface; therefore, the correlation between PM and AOD strongly depends on the vertical distribution of aerosols and several meteorological factors, i.e., profiles of potential temperature and wind regime (Chitranshi et al., 2014; Sinha et al., 2015). On the other hand, the satellite snap-shots usually cover large spatial areas and are subject to cloud contamination (Zhang et al., 2005; Gupta et al., 2006). Therefore, incorporation of other meteorological parameters, such as ambient relative humidity (RH), fractional cloud cover and mixing layer height, improved the prediction of PM, since weather conditions can greatly affect the aerosol dilution and dispersion processes (Gupta and Christopher, 2009b; Benas et al., 2013; Chitranshi et al., 2014). More specifically, RH influences the hygroscopic growth of the particles (Benas et al., 2013), while strong wind speeds affect the PM dispersion (Xiao, 2011) or contribute to additional emissions over arid terrains (Rashki et al., 2012). Furthermore, inclusion of surface temperature for PM estimates has been found to significantly improve the results (Chitranshi et al., 2014; Benas et al., 2013), since temperature may modulate the photochemical reactions and play a major role in the boundary-layer dynamics that control the PM concentrations at the ground via dilution and/or trapping of pollutants (Dumka et al., 2015). In this respect, the atmospheric stability index and boundary layer height (BLH) have been used as surrogates for atmospheric stability resulting in more accurate estimates of PM since the vertical variability of aerosol can be determined (Rohen et al., 2010; Emili et al., 2010). A literature overview (e.g. Hoff and Christopher, 2009) shows that various techniques, such as simple linear regressions (Wang and Christopher, 2003; Engel-Cox et al., 2004), multiple linear regressions (Kanniah et al., 2014b; Chitranshi et al., 2014; Gupta and Christopher, 2009a), nonlinear regressions (Sotoudeheian and Arhami, 2014; Benas et al., 2013), mixed effects models (Kloog et al., 2012; Nordio et al., 2013; Yap and Hashim, 2013; Hu et al., 2014; Xie et al., 2015; Beloconi et al., 2016), statistical and chemical transport models (Van Donkelaar. et al., 2010), as well as complex nonlinear regressions such as Artificial Neural Network (ANN) (Kanniah et al., 2014b; Gupta and Christopher, 2009b) have been applied for PM estimations from space at local, regional and global scales. These advanced statistical techniques significantly improve the prediction capability of the simple AOD-PM relationships that were initially used for PM estimation from space. The continuous renewal and improvement of the databases, methods and techniques for PM monitoring from space justifies the high importance that such studies have for air-pollution mitigation efforts and human-health perspectives, especially in the developing countries (Snider et al., 2016).

In Malaysia, the most widely measured air-quality parameter is the PM_{10}. However, the PM_{10} monitoring network is rather sparse accounting for 74 stations over the whole Malaysian territory of 330,290 km^2 and, therefore, PM_{10} estimates from space are considered as absolutely necessary for warning the local population about the air-quality standards. Nevertheless, no studies covering the entire Malaysian territory have been conducted to map PM_{10} from space. A recent study by Yap and Hashim (2013) attempted to estimate PM_{10} over Peninsular Malaysia (only) via a mixed-effects model, which took into consideration the monthly variability of the coefficients (i.e. slope and intercept) in association with MODIS-AOD and PM_{10}. The use of daily measurements of AOD allows for better assessment of space-time interactions than models that only have spatially resolved time invariant land use terms (Kloog et al., 2012) because the relationship between AOD and $PM_{2.5/10}$ varies on daily basis due to differences in mixing height, RH, particle composition, vertical profiles etc. (Lee et al., 2011; Sinha et al., 2015). Moreover, the PM_{10} concentrations exhibit a remarkable intra-seasonal variation at bi-weekly frequencies of 10–20 days (Juneng et al., 2009) over Malaysia and, therefore, the existing model may discount the PM_{10} daily variation and limits the possibility of studying the acute effects of air pollution. Furthermore, the mixed-effects model may introduce uncertainties with the assumption of linearity and using AOD as the single predictor (Xie et al., 2015). Finally, the model assumes that there is little spatial variability in the relationship, which is not necessarily true especially when the modelling domain gets larger (Hu et al., 2014) as in our study.

Thus, for improving the accuracy of PM_{10} estimates, the present study integrates atmospheric/meteorological parameters along with advanced statistical methods, i.e., multiple linear regressions (MLR) and ANN, for PM_{10} monitoring from space for the first time over the whole Malaysian territory. Although a better knowledge has been gained during the last years on the spatial and temporal evolution of aerosols and pollutants over Peninsular Malaysia (Khan et al., 2016a, b), monitoring the aerosol loading and composition over the East Malaysia (Borneo Island) remains a real challenge and not so well documented. There are only 15 ground monitoring stations in this region compared to its size of 198,080 km^2. Thus, PM_{10} estimations via satellite remote sensing is very essential in this region, which is strongly affected by pollution especially during the southwest monsoon. The seasonality of the spatial distribution of PM_{10} is provided at 10 × 10 km spatial resolution during 2007–2011 and is discussed against the potential sources of aerosol and pollutants. The PM_{10} estimates are validated against ground-based measurements from several stations, while the results are useful for air quality, meteorological and health studies as well as for the calculation of the Environmental Performance Index (EPI) over Malaysia.