Novel Decomposition Scheme for Characterizing Urban Air Quality with MODIS

Prakhar Misra *,†, Aya Fujikawa and Wataru Takeuchi †

Department of Civil Engineering, Institute of Industrial Science, The University of Tokyo, Meguro, Tokyo 153-8505, Japan; watt770@gmail.com (A.F.); wataru@iis.u-tokyo.ac.jp (W.T.)

* Correspondence: mprakhar@iis.u-tokyo.ac.jp; Tel.: +81-809-982-0620
† Current address: Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan

Received: 30 May 2017; Accepted: 2 August 2017; Published: 7 August 2017

Abstract: Urban air pollution is one of the most widespread global sustainability problems. Previous research has studied growth or fall of particulate matter (PM) levels using on-ground monitoring stations in urban regions. However, studying this worldwide is difficult because most cities do not have sufficient infrastructure to monitor air quality. Thus, satellite data is increasingly being employed to solve this limitation. In this paper, we use 16 years (2001–2016) of aerosol optical depth (AOD) and Angstrom exponent (α) datasets, retrieved from MODIS (Moderate Resolution Imaging Spectroradiometer) sensors on the National Aeronautics and Space Administration’s (NASA) Terra satellite to study air quality over 60 locations globally. We propose a novel technique, called AirRGB decomposition, to characterize urban air quality by decomposing AOD and α retrievals into ‘components’ of three distinct scenarios. In the AirRGB decomposition method, using AOD and α dataset three scenarios were investigated: ‘R’—high α and high AOD, ‘G’—high α and low AOD, and ‘B’—low α and low AOD values. These scenarios were mapped and quantified over a triangular red, green and blue color scale. This visualization easily segregates regions having a high concentration of industrial aerosol from only natural aerosols. Our analysis indicates that a sharp divide exists between North American and European cities and Asian cities in terms of baseline pollution and slopes of R and G trends. We found that while pollution in cities in China has started to decrease (e.g., since 2011 for Beijing), it continues to increase in South Asia and Southeast Asia. e.g., R offset of Beijing and New Delhi was 54.98 and 50.43 respectively but R slope was −0.04 and 0.08 respectively. High offset (≥45) and slope (≥0.025) of B for New York, Tokyo, Sydney and Sao Paolo shows that they have clean air, which is still getting better.

Keywords: global city comparison; aerosol; least squares; AOD; Angstrom exponent; anthropogenic emissions; AirRGB

1. Introduction

Goal number 11.6 of the ‘2030 Agenda Sustainable Development Goals’ (SDG) adopted by the United Nations General Assembly [1] targets ‘reducing environmental impact of adverse urban air quality’. In order to achieve this, these regions need to be identified where air quality is reaching unacceptable levels. Air quality species that are commonly considered to characterize overall urban air quality are PM2.5 and PM10 (particulate matter of sized 2.5 µm and 10 µm), NO2, SO2, CO and O3. Both PM particles and NOx constitute aerosols. Particulate matter (PM2.5, PM10) is commonly found to exceed limits set by the World Health Organization, especially in developing countries. Recent reports by World Health Organization [2] indicate that across several urban locations globally polluting aerosol species is rising against the target of the SDG.

Studies have extensively employed in situ observations and satellite datasets to investigate the factors leading to a build-up of pollutants borne out of human activities. Depending on the scale
at which it is being studied, urban air pollution and its growth can be a result of trans-boundary transport \[3\], anthropogenic emissions and meteorology \[4\], and socio-economic development \[5\]. Source apportionment studies from various cities \[6–10\] have routinely reported that vehicular, industrial emissions and resuspended soil-dust are main contributors in raising mean annual pollution levels.

Ground infrastructure is required to monitor urban air quality reliably. This is seldom found in most cities globally. In recent years, several countries have instituted public air quality monitoring programs in the form of Air Quality Index (AQI) which is basically an index to report air quality at fixed time intervals using a set of air pollutants. However, different methodologies for AQI make it difficult to compare air quality across countries, e.g., US’s AQI and China’s AQI (called ambient AQI) consider different pollutants as well as a different formula. Several research works have monitored the trend of particulate matter levels over cities on a global scale \[5,11–13\]. An important component of these studies was availability of long-running temporal data from ground monitoring sensors in their research locations. Availability of historical dataset is also important to monitor long-term growth or a fall in pollutants to assess efficacy of mitigation policies.

Satellite sensors such as MODIS (Moderate-resolution imaging spectroradiometer) and MISR (Multi-angle imaging spectroradiometer) can provide aerosol retrievals that are suitable for studying air pollution trends. Although the satellite sensors cannot measure particulate matter levels directly, they can retrieve variation in aerosol concentration and composition fraction of fine mode and coarse mode particles, which can then be related to particulate matter. Aerosol concentration is represented by aerosol optical depth (AOD) \(\tau_\lambda\), which is the integrated extinction coefficient over columns of unit cross section measured at a particular wavelength \(\lambda\). Change in AOD relative to various wavelengths of light is quantified by the Angstrom exponent, \(\alpha\). By observing \(\tau_{\lambda_o}\) and \(\tau_\lambda\) at wavelength \(\lambda_o\) and \(\lambda\), respectively, \(\alpha\) can be approximated as follows \[14\]:

\[
\alpha = -\frac{\log \tau_\lambda}{\log \lambda},
\]

\(\alpha\) reflects particle size mode of aerosol \[14\]. \(\alpha\) is inversely related with the particle size, i.e., a high \(\alpha\) indicates dominance of fine mode particles (radii \(\leq 0.6 \mu m\)) \[15\]. Such fine mode particles are associated with biomass burning or anthropogenic combustion activities \[16\]. The following Nakajima and Higurashi \[17\] high values of \(\alpha\) (\(>1\)) indicate the presence of fine urban-industrial aerosol, while prominence of mineral or marine aerosol is indicated by small \(\alpha\) (\(\sim 0\)). Small \(\alpha\) and AOD values result in high visibility of the sky, which gets reduced under large \(\alpha\) and low AOD conditions. This happens due to Mie scattering \[18\] of incoming light by fine aerosol particles from industrial and vehicular emissions (black carbon, sulphates, etc.) \[19\]. Thus, inferences can be drawn about concentration and the dominant mode of particulate matter (and therefore the nature of the polluting source) by considering AOD and \(\alpha\) together.

Current online portals for global aerosol visualizations using satellite datasets e.g., NASA Worldview \[20\] and NASA Giovanni \[21\] can depict concentration of specific sized aerosol or global concentration of aerosol optical depth (e.g., Giovanni) or AOD and \(\alpha\) in the same map (e.g., Worldview). Their capabilities can be enhanced by representing dominant aerosol mode (\(\alpha\)) and aerosol concentration (AOD) on a single map in such a manner that not only aerosol level but also its type can be distinguished. This can be achieved by quantifying it over a scale that considers both AOD and \(\alpha\) information.

The objective of this study is to characterize urban air quality by decomposing urban AOD and \(\alpha\) levels into ‘components’ of three distinct scenarios and investigate their trend over the last 16 years. The originality of this research lies in utilizing the long-running 16 year satellite AOD and the \(\alpha\) dataset to monitor air quality in urban locations globally.
2. Methodology

2.1. Study Location and Flowchart

Forty-six country capitals are selected for this research, as shown in Figure 1. These cities are selected on the basis of countries with the most car ownership in 2010 or country capitals that have the highest growth rate of car ownership in 2006–2010. In addition, 14 other highly populated cities known to be undergoing variation in air pollution are also selected e.g., Chongqing, Mumbai, Los Angeles, etc. All locations are shown in Figure 1. In the figure, cities are also color coded by developing or a developed country based on the United Nations’ country classification [22].

![Figure 1](image1.png)

Figure 1. Cities considered for analysis in this study. The city markers are sized as per vehicle population. The markers are shown in red if the national gross domestic product per capita (pc GDP, year 2015) is less than $12,000 otherwise, it is shown in green.

Figure 2 lays down the flowchart of data processing and analysis to obtain results.

![Figure 2](image2.png)

Figure 2. Flowchart of study outlining data, processing and analysis. AOD and ANG refer to the aerosol optical depth and Angstrom exponent products respectively.

2.2. Datasets Used

Retrieval of AOD and $\alpha$ was based on Level2 Atmosphere Daily Global Product ‘MOD04L2’ Collection6 obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard NASA’s Earth Observing System (EOS) Terra satellite. Terra has a sun synchronous, near polar circular orbit that passes over at approximately 10:30 a.m. local
Remote Sens. 2017, 9, 812 4 of 19

time. ‘Deep_Blue_Aerosol_Optical_Depth_550_Land’ dataset was used for AOD images and ‘Deep_Blue_Angstrom_Exponent_Land’ dataset was used for α images. MODIS was considered for its ability to provide spatially and temporally contiguous global coverage. To develop global daily products, cloud-free pixels (quality flag QA = 3) were considered and composited spatially. This daily data has a 10 km resolution and observes between 82° North to 82° South latitudes. The acquisition period consists of 16 years of datasets from January 2001 to December 2016. In this paper, we are converting MODIS observations to integer numbers by multiplying the float values by 1000.

Over bright surfaces (like urban areas) aerosol retrievals obtained from Deep Blue (DB) algorithm of MODIS show good agreement with those obtained by AERONET ground stations and are useful for urban areas [23]. DB performed better than the other dominant aerosol retrieval algorithm, Dark Target (DT), in Indo-Gangetic Plains [24]. The uncertainty of DB retrieval is dependent on geographical location and meteorological season. For example, Ginoux et al. [25] found that, while DB based AOD values were within the standard deviation of AERONET data in Africa, the Arabian Peninsula, and India, they were overestimated in California, Australia, and Israel. Payra et al. [26] observed dependence of estimation accuracies on Monsoon in a semi-arid urban region in India. In 2013, the second generation of DB (as part of Collection 6) was proposed [27], which used a hybrid approach to account for inhomogeneously vegetated urban/built-up areas among other improvements. In various regions of China, Tao et al. [28] found that although DB tends to underestimate AOD compared to DT, it provides a much larger sampling frequency in high aerosol loading conditions (AOD > 1) like regional haze pollution. Bilal and Nichol [29] found that DB retrieved AOD performed well compared to AERONET in the national capital region of China on both hazy and clear days, yet tended to overestimate during very polluted days (AOD > 1.5). The level of uncertainty of DB Collection 6 AOD retrievals (QA = 3) is approximately ± (0.03 + 0.23τ) [30], an improvement over Collection 5.1 uncertainty of ± (0.05 + 0.2τ) [31]. In Collection 6, a ‘best of’ AOD product also exists that combines AOD from DB and DT algorithms to estimate aerosol distribution over densely vegetated areas as well as bright surfaces. Sayer et al. [32] compared DB, DT as well as the combined DB-DT product and found that neither ‘consistently outperforms the other’. Although DB does give a higher fraction of observations agreeing with AERONET than DT, in very high AOD areas of South Asia DB tends to underestimate while DT tends to overestimate. The authors suggest that, in such areas, the combined product is likely to perform better than either DB or DT by canceling the opposing errors [32]. Therefore, further verification of the combined DB-DT product is desirable in terms of seasonality and long term trends, especially for Asian cities with evolving trends. Since, in this research, the area of interest is only urban/built-up regions, we have considered DB only, but there is also scope for the merged DB-DT product.

AOD and α cannot be retrieved under cloudy sky conditions, as MODIS is an optical sensor. This results in significant reductions of cloud-free composites over equatorial regions of West Africa, South America, and Asia, as well as boreal regions of Europe and Canada [33]. To overcome this deficiency, usually a mean monthly composite is generated using daily images. Possible techniques include generating simple and quality weighted mean over a period of 16 days or 30 days [34], n-day composites on a rolling basis [33] and pixel coverage threshold monthly mean [35]. However, final values between these images can vary as much as 30% and, in light of different biases involved in each of these methods, Levy et al. [35], have noted that neither technique can be adjudged superior to another. In this research, a simple mean of daily images has been considered to obtain monthly datasets.

For validation of the proposed technique, measurement of PM2.5 from Beijing was used. Beijing was considered, since it is one of the few highly polluted cities with a publicly available and long-running PM2.5 measurements from a ground monitoring sensor. The sensor is installed at the United States’ embassy in the Chaoyang district (latitude −39.95°, longitude −116.47°), Beijing [36] and the observation dataset is available online [37]. The dataset consists of daily hourly measurements since May 2008. Although this data is not fully verified, previous studies have demonstrated its close
correspondence with citywide PM2.5 concentrations [38,39]. For our study, we only considered ‘valid’ (quality flag) measurements made at 10:00 a.m. between May 2008 and December 2016.

2.3. AirRGB Decomposition and Visualization

Global aerosols have been visualized using color schemes in earlier research works to obtain more information. For example, Putman et al. [40] studied the effect of local emissions and large-scale weather systems on column CO$_2$ mixing ratios by portraying global clouds (white) and aerosols (sulfate (blue), dust (brown) and carbon (green) from GEOS-5 model simulation. Penning de Vries et al. [41] determined the aerosol type by making use of trace gas abundance, aerosol absorption, mean aerosol size and MODIS derived AOD and $\alpha$ to represent desert dust (red), biomass burning smoke (gray), secondary biogenic aerosols (green) and sea salt (blue).

Mukai et al. [42] and Sano et al. [43] used about three years of daily AOD and $\alpha$ observation in urban Osaka, Japan to cluster aerosols into dust episodes, anthropogenic aerosol events and background clear atmosphere. The scatter points tended to lie within a triangular space when plotted over an $\alpha$ and AOD orthogonal axes. Vertices of the said triangle represent three diverse scenarios corresponding to very high $\alpha$ and low AOD value regime (signifying high composition of fine mode particles by anthropogenic contribution in emissions), high AOD value regime (signifying high aerosol levels in general), and low $\alpha$ and low AOD value regime (signifying clean air quality).

In this paper, a three component decomposition technique called ‘AirRGB’ for visualizing aerosols to study trend of aerosols in urban regions has been proposed. The key idea was representing proximity of AOD and $\alpha$ values to the three scenarios by decomposing it into three components. The three components are visualized by representing corresponding channels over a three-channel red-green-blue (RGB) scale varying from 0 to 100. In the AirRGB technique, we consider AOD and $\alpha$ value as a $(x, y)$ Cartesian coordinate—(AOD, $\alpha$) and refer to it as an ‘aerosol coordinate’. A bounding triangle is chosen such that, for each location, monthly aerosol coordinates shall lie within the triangle as much as possible. After setting the bounding triangle, aerosol coordinate is then transformed into a three-element vector that contains distance metrics from each vertex of the triangle. Since urban locations seldom show very high AOD and low $\alpha$ unless regularly affected by dust storms, this scenario is not represented. Aerosol coordinates of each location is transformed onto a RGB scale to represent its ‘closeness’ to each scenario. The transformation procedure is discussed in detail in the following steps:

1. For each urban location, the overall mean of monthly (from year 2001 to 2016) $\alpha$ and AOD values of its central pixel were plotted on $y$ and $x$-axis, respectively. The central pixel refers to the pixel corresponding to the city center. It was chosen to deal with cities (specially megacities) that are spread across several MODIS pixels (e.g., New York). This scheme is unlikely to affect the AOD and $\alpha$ estimates since neighboring pixels have similar values over an urban area [44].

2. Consider a bounding triangle, $\triangle$RGB, such that its vertices corresponding to (a) high $\alpha$ and high AOD, (b) high $\alpha$ and low AOD and, (c) low $\alpha$ and low AOD are labeled as $R(x_R, y_R)$, $G(x_G, y_G)$, and $B(x_B, y_B)$, respectively. $R$, $G$ and $B$ correspond to the colors red, green and blue (the color assigned to the vertex is irrelevant, but, in this paper, it is considered in the same fashion to ensure consistency).

3. For any coordinate $C(x_C, y_C)$ lying inside $\triangle$RGB, decompose it over 3-band RGB-scale by first computing its distance metric from each vertex ($R$, $G$ and $B$) as $l_R, l_G$ and $l_B$. Thereafter, a distance vector $l$ is derived by normalizing each distance metric by the side opposite to the corresponding vertex in $\triangle$RGB, i.e., $L_R, L_G$ and $L_B$, respectively. $l$ is given for a general point coordinate $C$ as:

$$l = \begin{bmatrix} l_R \\ l_G \\ l_B \end{bmatrix}$$

where
\[ l_R = |(y_B - y_G) \times (x_C - x_B) + (y_B - y_C) \times (x_B - x_G)|, \]
\[ l_G = |(y_R - y_B) \times (x_C - x_R) + (y_R - y_C) \times (x_R - x_B)|, \]
\[ l_B = |(y_G - y_R) \times (x_C - x_G) + (y_G - y_C) \times (x_G - x_R)|. \]  \( (3) \)

For intuitive understanding of Equation (3), suppose that proximity of \( C \) is increasing to a vertex of \( \triangle RGB \). Then, components of \( l \) corresponding to other vertices will start decreasing to zero. The distance components for any vertex is in fact the sum of the area of two rectangles formed between \( C \) and each of the other vertices. For example, in Figure 3, component \( l_G \) is the sum of yellow and green rectangles. Note that as \( C \) approaches towards either \( R \) or \( B \) vertices, \( l_G \) linearly reduces to zero.

![Diagram](image.png)

**Figure 3.** Finding \( l_G \) (distance metric corresponding to vertex \( G \)) for a point \( C \) lying within bounding \( \triangle RGB \). The vertices \( R, G \) and \( B \) refer to scenarios of (a) high \( \alpha \) and AOD, (b) high \( \alpha \) and low AOD and, (c) low \( \alpha \) and low AOD. In this figure, \( l_G \) is the absolute sum of areas of dotted rectangle \( ((y_R - y_B) \times (x_C - x_R)) \) and hatched rectangles \( ((y_R - y_C) \times (x_R - x_B)) \) (Equation (3)). Note that, since \( x_C \leq x_R, l_G \) will effectively be the difference of the areas of rectangles, \( l_G \) will decrease towards zero as \( C \) approaches any vertex other than \( G \).

4. For RGB color representation, values for channel \([R, G, B]\) are normalized by maximum possible distance vector and multiplied by 100. It is given as:

\[ [R, G, B] = l \cdot \frac{100}{l_R + l_G + l_B}, \]  \( (4) \)

where \( C \subseteq \triangle RGB \)

5. For locations whose aerosol coordinates lie outside \( \triangle RGB \), \([R, G, B]\) is set as per the following rule:

\[ [R, G, B] = \begin{cases} 
(100, 0, 0) & x_C \geq x_R \text{ & } y_C \geq y_R, \\
(0, 0, 100) & x_C \leq x_B \text{ & } y_C \leq y_B \end{cases} \]  \( (5) \)

AirRGB decomposition can be interpreted as a transformation technique using any aerosol coordinate (i.e., AOD and \( \alpha \) combination) that can be mapped onto a RGB scale. On this scale, high values (\( \geq 50 \)) of a component imply significant closeness towards the corresponding scenario.
that, if any component has a value $\geq 50$, other components will be $<50$. Additionally, if values of all three components are $<50$, then a careful interpretation is needed since aerosol coordinates lie in the central region of the triangle and do not show proximity to any scenarios. For example, if $B$ is trending to high values ($B \geq 50$), it indicates that the aerosol concentration is decreasing and the dominant aerosol mode is gradually shifting towards mineral aerosol (e.g., dust, etc.). This holds irrespective of $R$ or $G$ value. On the other hand, decreasing value or a low value of $B$ only indicates that it is approaching a poor air quality regime. Only considering this information does not tell us whether this is due to anthropogenic emissions or mineral dusts unless $R$ or $G$ values are also probed simultaneously. It should be noted that, since it is a triangular space, the values of RGB components are not independent. Therefore, in Figure 3, it is impossible for $C$ to trend in such a way that only one of the components change.

2.4. Uncertainty Estimation and Validation

The uncertainty in the AOD ($\delta x_C$) from DB is $\pm (0.03 + 0.23\tau)$ [30] as discussed before. We used the $\alpha$ uncertainty calculated from DT, which is $0.4$ [45], since we could not find literature supporting uncertainty of $\alpha$ ($\delta y_C$) derived from DB. Equation (6) was derived to estimate uncertainty ($\delta l_R, \delta l_G, \delta l_B$) in $R, G$ and $B$ based on error propagation in Equation (3):

$$\delta l_R = 100\sqrt{[(y_B - y_G) \times \delta x_C]^2 + [(x_B - x_G) \times \delta y_C]^2},$$

$$\delta l_G = 100\sqrt{[(y_R - y_B) \times \delta x_C]^2 + [(x_R - x_B) \times \delta y_C]^2},$$

$$\delta l_B = 100\sqrt{[(y_G - y_R) \times \delta x_C]^2 + [(x_G - x_R) \times \delta y_C]^2}.$$ (6)

The $R, G$ and $B$ values obtained in the previous Section are validated with PM2.5 values using Beijing as a test case. Daily observations of $R, G$ and $B$ were plotted against corresponding PM2.5 measurements. Since PM2.5 sensor provides roadside measurements rather than background measurements, we expected consideration variation in the distribution of the plot. To overcome these mean values at range intervals of 25, $\mu g/m^3$ of PM2.5 were also plotted.

3. Result and Discussion

3.1. AirRGB Decomposition Parameters

Figure 4 shows a scatter plot of mean AOD and mean $\alpha$ and respective standard deviations (over the mean monthly values from January 2001 to December 2016) for all locations. The aerosol coordinates are scattered within an approximate triangular space, vertices of which correspond to the three scenarios discussed earlier in Section 2.3. It can also be seen that many cities from developed countries, e.g., Busan, Sydney, Los Angeles, Madrid, etc., have low AOD yet high $\alpha$ values, while cities of rapidly developing countries (India and China) generally constitute the rightmost vertex of the triangle (R) that signifies high AOD and high $\alpha$. Most cities have large standard deviation along $\alpha$ than AOD. This is expected to result in variations in time-series of $G$ and $B$ values. However, if a large standard deviation in AOD is also present, it will result in fluctuations of $R$ and $G$ values. Generally, a high standard deviation would imply a large change in aerosol levels either compared to its historic levels or strong annual seasonal variations. In Figure 4, if a standard deviation bar of a city exceeds the AirRGB triangle, then extreme values of either of $R, G$ or $B$ shall be obtained.
Figure 4. Scatter plot of mean of monthly aerosol optical depth (AOD) and Angstrom exponent, $\alpha$ (ANG) value of each city under study. The corresponding standard deviations are depicted as horizontal and vertical lines. Cities for which availability of valid data is less than 100 months are shaded light (e.g., Lagos, Singapore). This figure also shows the bounding triangle, with each vertex corresponding to extreme scenarios of aerosol optical depth and Angstrom exponent value, and are labeled as $R$, $G$ and $B$.

Based on the distribution of scatter points as discussed in Section 2.3 and as observed in Figure 4, aerosol coordinates of vertices of AirRGB bounding triangle were identified as $R(600, 800)$, $G(20, 1000)$ and $B(50, 50)$ for red, green and blue channels, respectively. Two cities lay outside the triangular space: Lagos and Dhaka. Lagos’ valid observations were less than 48 months (four years) out of the total 16 years of observations and were not considered as reliable.

The transformation scale based on the vertex parameters of the AirRGB triangle is visualized in Figure 5. The $[R, G, B]$ values of the vertices $R$, $G$ and $B$ (based on Equation (4)) are respectively: $[100, 0, 0]$, $[0, 100, 0]$ and $[0, 0, 100]$. Points lying far outside the triangle are not covered by the three scenarios, except for those discussed in Equation (5).
Figure 5. Proposed AirRGB scale for transformation based on bounding triangle shown as a three-band RGB image. On this scale, proximity to a vertex results in a higher value of the color corresponding to that vertex.

3.2. Global AirRGB Visualization

The global AirRGB visualization map is shown at five-year intervals beginning in 2001 in Figure 6. It is easy to locate the highly polluted regions across East China and the Indian subcontinent, which show up in a bright red color (high \( R \) value). Furthermore, between the year 2010 and 2015, falling anthropogenic pollution in the Chinese urban region can be seen by a decreased intensity of red over Eastern China. In the same period, man-made pollution increased to an intense red in North India, West Asia and Central-West Africa and Central South America. Most regions across Western Europe, North America and South Japan are combinations of blue and green, inclined towards the former. This shows that, although industrial emissions are active, their levels are relatively lower than the regions discussed before.

While this paper does not focus on non-urban areas like dense vegetation and deserts, we would like to point out that application of the AirRGB technique over such regions may sometimes cause difficulty in interpreting values. This is likely to happen since AOD and \( \alpha \) characteristics of such areas lie far beyond the bounding AirRGB triangle. For instance, the dusty Australian desert has an AOD in the range of 100 to 250 [25], but its \( \alpha \) lies far above 1000, resulting in bright green (high value of \( G \)). This is clearly misleading. Values of AOD and \( \alpha \) over other prominent desert regions like the Sahara desert, the Arabian desert, the Thar desert and parts of the Gobi desert also lie outside the AirRGB triangle but still show up in red due to proximity to the vertex \( R \). Furthermore, DB retrieves only under bright surfaces in the visible wavelength. Thus, data over regions with extensive biomass or dense vegetation like Central Africa and Amazon cannot be retrieved completely since their surface reflectances in the visible are less than 0.15 [25]. Here, the DB sampling rate is much less and often of poor quality as compared to DT. This also leads to underestimation in the mean monthly images prepared by us. Furthermore, a bright green strip can be seen in the Western Asia in the Red Sea as well as the Persian Gulf. Although there are no urban centers of interest here, we suspect it has high \( G \) value on account of fine mode aerosols of sulphur dioxide emissions from oil refineries in this area [46,47]. Thus, to incorporate effective representation of non-urban regions, the AirRGB triangle with different vertex parameters can be specified using aerosol coordinates from such pixels.
Figure 6. Global visualization of AirRGB using mean annual datasets of 2001, 2005, 2010 and 2015. Over this period, industrial regions across North India tend towards more intense R while Eastern China shows a decrease in intensity of R after year 2010.

3.3. AirRGB Trend of Cities

We further discuss variation in urban air quality due to anthropogenic causes using AirRGB decomposition. We first investigated representative cities lying near the vertices of AirRGB bounding triangle in Figure 4: New Delhi (highest R value), Madrid (highest G value) and Manila (highest B value). A four-month moving average of their R, G, B values was plotted and shown in Figure 6.

Air quality in New Delhi, which is located in the Indo-Gangetic Plains, undergoes strong seasonal variations. From Figure 7a, rapidly rising R values (slope is 0.08) can be noticed distinctly. High R (≥50) values are a seasonal phenomena that have been rising even further since 2009. This establishes that contribution and concentration of fine mode particles is rising gradually in these places. The annual maximum value of R occurs around the December–January period, while the lowest values are seen in August–September. B values have been decreasing (slope is −0.04), with high B values (~50) rarely being seen except peak Monsoon season month (July). This signifies a constant presence of aerosols due to anthropogenic emissions, which undergo wet scavenging in rain. High peaks of G are observed frequently in the September months, which is caused by biomass burning after the harvesting season in neighboring farmlands.

Figure 7. Cont.
Figure 7. Comparison of AirRGB trend of offset and slope for cities with extreme RGB values (New Delhi, Manila and Madrid) and other Asian cities (Beijing, Tokyo and Dhaka) over a period from January 2001 to December 2016. Values of R, G, B components are presented as a four-month moving average to ease visual interpretation.

In contrast to New Delhi, RGB components of Madrid (Figure 7b) did not show any seasonality. Low R values (<10) signify low aerosol loading, but G values are much higher (~70) than other components, signifying a large contribution of anthropogenic emissions (mostly vehicular) to urban
aerosol levels. This has also been confirmed by Borge et al. [48] as a violation of NO₂ standards. At certain events, air quality also tends to become very clean, as evidenced by transition from high G value to high B value in January–February of 2003, 2004, 2010, 2013, 2014 and 2016.

Manila (Figure 7c) has very clean air with low aerosol emissions in the early part of the observation period. This was seen from high offset of B (∼76) and very low offset of R (∼4) and G (∼12) values. However, there is also a sharp fall in B level (slope = −0.06).

As cases, we focus on the Asian cities of Beijing, Tokyo and Dhaka. In Figure 7d, it was observed that traditionally high R values are decreasing in Beijing. This has especially become apparent after the year 2011 when higher values of B can also be seen. It was noteworthy that, during the month of August in 2008, R value suddenly drops to much lower than what is expected during this time of year. This was likely due to the 2008 Beijing Olympics Games, during which several traffic and emission restrictions were strictly enforced from the period of April 2007 to January 2009, as has also been confirmed by Wang et al. [49].

Dhaka also presented an interesting case of varying air pollution. The initial decrease in R between 2001 and 2005 has been attributed to government policy intervention factors e.g., a ban on two stroke vehicles, and a promotion of compressed natural gas (CNG) as fuel and the control of industrial emission [50]. A rise in R can be seen post 2006, which has been attributed to a rise in emissions from brick factories, while the fall in R after 2011 is a result of efforts to convert petrol/diesel vehicles to CNG emissions [7]. It appears then that a rise in G despite a fall in R could largely be a cause of brick factories.

Among the four cities, Tokyo has the lowest R offset (∼30). Its B offset (47.83) is similar to that of Dhaka (50.28), but recent B values are much higher than the latter. Furthermore, air quality in Tokyo has been consistently improving, especially in the last 10 years, as represented by a positive slope of B (0.03) and negative slopes of R (−0.01) and G (−0.02).

3.4. Global AirRGB Component Trend

Global trends of AirRGB components of all locations under study were computed for each component using the least squares method over monthly composites from the year 2001 to 2016. Offset and slope of the best fit line for each city so obtained is shown in Figure 8. The offset corresponds to baseline levels of each RGB component while its slope depicts trends of the components.

3.4.1. Global Trend of R

In Figure 8a, a positive slope of R was observed for most Asian cities. Furthermore, all such cities belong to the developing countries located in South Asia and South East Asia (e.g., India, Nepal, Vietnam). Cities of North America (e.g., New York, Washington DC) and Europe (e.g., Brussels, Paris) were concentrated in a narrow region with a negative slope (between −0.03 and 0.00) and small offset (less than 16) except Athens and Los Angeles. Both Athens and Los Angeles count amongst the most polluted cities in Europe and North America on account of traffic and industrial emissions (Chaloulakou et al. [51], American Lung Association [52]). However, their offset was much less than that of Asian cities, which is mostly greater than 25. Notable outliers with large offsets include Beijing, Urumqi, New Delhi and Kolkata. Of these, only Beijing has a negative slope, as discussed previously. Several other cities in China (Shanghai, Chongqing) also have high offsets but negative slopes. This is likely an indication that although baseline aerosol levels in Chinese cities have historically been high, its air quality is improving. It is of further interest to identify factors leading to increase or decrease of slope in these cities.
Slope and offset values, two clusters of cities are visible (a) 70 to 40 and (b) 35 to 5. In the first cluster, it was hard to distinguish between cities from developed and developing countries. This implied that, even under low aerosol loading conditions. Unlike the plot for R, here European and North American cities were not concentrated in a well-defined region. This implies varying contributions of traffic and industrial emissions to urban pollution in cities in developed countries. However, based on offset values, two clusters of cities are visible (a) 70 to 40 and (b) 35 to 5. In the first cluster, it was hard to distinguish between cities from developed and developing countries. This implied that, even with cleaner air quality, fine mode aerosol from traffic and industrial emission can dominate. South American cities Rio de Janeiro, Buenos Aires and Sao Paulo were present in the former cluster and showed negative slopes. This is in agreement with issues associated with SO$_2$ (additionally CO for Sao Paulo) and PM2.5 levels caused by vehicular and industrial emissions as identified previously.
(Toledo et al. [53], Mage et al. [5], Bogo et al. [54]). The second cluster shows that cities like Tokyo and London have similar offsets (∼25) as Kolkata and Beijing. However, these cities have midrange offsets that lie along and close to the side BR in △RGB (Figure 5), and so analyzing the second cluster based on G values alone can lead to misleading conclusions.

3.4.3. Global Trend of B

In the offset and slope plot of global cities for B (Figure 8c), notable outliers existed in terms of slope. These were Ho Chi Minh, Bangalore, Dhaka and Islamabad. They have relatively high offset of B values (≥40) from historically anthropogenic polluting activities, but a large negative slope (≤0.07) indicates pollutants surging rapidly. In conjunction with their R and G plots, we find that they are rapidly moving towards high pollution and urban emissions in the following descending order: Bangalore, Islamabad, Ho Chi Minh and Dhaka. On the other hand, New York, Tokyo, Sydney and Sao Paolo have high B offsets (>45) but a positive B slope, which indicates that these cities not only had low aerosol levels but are still achieving cleaner air. All South American cities are also on a rapid course to achieving high B values.

In some cities like Kabul and Baghdad, not much literature is available regarding air quality trend of anthropogenic emissions. For instance, we inferred domination of primary fine mode aerosol at Kabul based on its high offset (∼65) and positive slope (∼0.03) of G value. This inference is similar to [55], who collected samples at urban locations in Kabul but unlike that realized by [56], who used two weeks of data monitored at military camps). For Baghdad, high R and G values point to a presence of mineral dust and fine aerosols. This is in conjunction with [57], who identified the dominant sources to be dust and carbonaceous aerosols (gasoline and diesel engines) using source apportionment study.

The AirRGB technique provides a strong hint of emission issues by visualizing which of the three representative scenarios are prime in the city. Trends of R, G and B help identify the effect of a policy intervention, conclusions based on which can lead to establishing more robust ground monitoring sensor systems and systematic control of air quality by city authorities.

3.5. Uncertainty and Validation

3.5.1. Uncertainty

Using Equation (6) and AirRGB parameters, the maximum uncertainties in R, G and B were computed as: ±15.98, ±22.15 and ±23.2. We would like to point that these uncertainties are dependent on values of AOD, and we used the values of the vertex to compute corresponding uncertainties. Although R’s uncertainty is the lowest amongst the three, it can considerably impact interpretation drawn from Figure 8a, especially those locations showing a high standard deviation in their AOD. The high uncertainty in G and B values is on account of the uncertainty in α.

3.5.2. Validation

Nine hundred forty-five data points from Beijing with valid concurrent PM2.5 and AirRGB observations were found. They are plotted in Figure 9. A clear separation between AirRGB components can be seen with respect to PM2.5 values. Many B values occurred between the two bands 20 and 40 for a wide range of PM2.5. This occurs because Deep Blue algorithm caps extremely high α at 1.8 and fixes it to 1.5 during very low AOD conditions [31]. Several R values equal to 100 are easily noticeable because Beijing is located almost near the R vertex in the AirRGB triangle. Thus, aerosol coordinates lying beyond the vertex result in higher prominence of R equal to 100. Contrary to expectations, high R values were found during some low PM2.5 (<50 μg/m³) measurements while B reached moderate levels during high PM2.5 (>100 μg/m³) measurements. This is probably due to the difference between a ground monitoring sensor (which is affected local events) and a coarse resolution satellite sensor (reflecting observation over a 10 km² region).
To further analyze AirRGB, we tried overcoming the difference in spatial sampling by considering mean values, as shown in Table 1. Data points get sparse with higher PM2.5 values. When PM2.5 is between 0 and 50 µg/m³, G has the highest value, implying the absence of coarse particles and aerosol loading, mostly on account of fine mode particles. Despite low values of PM2.5, B rarely gets beyond 25. Thus, even when PM2.5 is low, Beijing’s air quality may still not be as breathable as in other cities with higher B values (e.g., London, New York). As PM2.5 rises beyond 50 µg/m³, high mean R values are seen, which approach 100 as PM2.5 goes further beyond 150 µg/m³.

Table 1. Mean values of PM2.5, R, G and B.

<table>
<thead>
<tr>
<th>PM2.5 Range (µg/m³)</th>
<th>Count</th>
<th>Mean PM2.5 (µg/m³)</th>
<th>Mean R</th>
<th>Mean G</th>
<th>Mean B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–25</td>
<td>301</td>
<td>13.72</td>
<td>13.12</td>
<td>60.16</td>
<td>26.7</td>
</tr>
<tr>
<td>26–50</td>
<td>196</td>
<td>36.7</td>
<td>29.79</td>
<td>44.7</td>
<td>25.49</td>
</tr>
<tr>
<td>51–75</td>
<td>129</td>
<td>61.79</td>
<td>51.93</td>
<td>30.26</td>
<td>17.79</td>
</tr>
<tr>
<td>76–100</td>
<td>97</td>
<td>86.72</td>
<td>64.62</td>
<td>20.93</td>
<td>14.44</td>
</tr>
<tr>
<td>101–125</td>
<td>61</td>
<td>111.72</td>
<td>74.68</td>
<td>16.2</td>
<td>9.11</td>
</tr>
<tr>
<td>126–150</td>
<td>66</td>
<td>137.59</td>
<td>80.76</td>
<td>10.57</td>
<td>8.64</td>
</tr>
<tr>
<td>151–175</td>
<td>37</td>
<td>161.91</td>
<td>88.04</td>
<td>6.88</td>
<td>5.06</td>
</tr>
<tr>
<td>176–200</td>
<td>25</td>
<td>189.48</td>
<td>95.47</td>
<td>2.13</td>
<td>2.38</td>
</tr>
</tbody>
</table>

In Beijing’s case, R values showed close agreement with PM2.5 \( R = 32.40 \ln(\text{PM2.5}) - 78.57 \) and a high coefficient of agreement (0.97). This relation might be unique for Beijing since other cities may have different emission characteristics and consequently different relationships between PM2.5 and R, G and B. One of the ways to assess this could be to simulate PM2.5 and PM10 generated using chemistry transport models (e.g., WRF-Chem, CMAQ) at different spatial resolutions [58]. A multi-sensor fusion approach can be also be taken to incorporate other crucial information such as planetary boundary layer height, relative humidity, season, and other geographical attributes of monitoring sites. CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations), an active remote sensing satellite, can be useful in this regard. To resolve coverage issues, aerosol data products from VIIRS (Visible Infrared Imaging Radiometer Suite) are gaining popularity.

Figure 9. Scatter plot of AirRGB values versus PM2.5 observations. Whenever R reaches 100, G and B reach zero.

One of the most prospective applications of this research can be to generate high resolution AirRGB maps to depict variation in air quality across regions with different characteristics in a city.
This can bring out the variation in aerosols over residential, commercial and industrial areas, as they are likely to have different emission sources.

4. Conclusions

In this paper, a novel visualization and monitoring scheme called AirRGB has been proposed that decomposes AOD and Angstrom exponent into three components (labelled as $R$, $G$ and $B$) corresponding to three scenarios: high pollution regime ($R$), clean air dominated by fine mode aerosols ($G$) and clean air ($B$). AirRGB contains more interpretable information than a conventional visualization. We used 16 years of AOD and Angstrom exponent ($\alpha$) retrievals from MODIS sensor and analyzed trend of their $R$, $G$ and $B$ components. We identified that, in contrast to European and North American cities, high overall background pollution exists in Asian cities (except Tokyo). Analysis of Asian cities reveals decreasing overall aerosol loading in Tokyo and the opposite scenario for New Delhi and Dhaka. Beijing has started showing an improvement in aerosol concentrations in terms of $R$ values, but rising $G$ values indicate that anthropogenic pollution is still the dominant mode.

In this research, daily levels of AOD and $\alpha$ retrievals were composited at the mean monthly scale by considering only cloud free pixel values. This is likely to bias the composite image to some extent, since aerosol concentrations during cloudy days are not considered. Furthermore, uncertainties in AOD and retrievals themselves also lead to uncertainties in $R$, $G$ and $B$, which were computed to be $\pm 15.98$, $\pm 22.15$ and $\pm 23.2$, respectively. Therefore, AirRGB parameters and inferences must be further investigated by considering datasets derived from other techniques of compositing means, such that this bias can be removed.

AirRGB trends inform us about the nature of urban air quality but not about the underlying cause. As a future step, AirRGB components can be linked with energy consumption and economic growth to understand how long-term air quality evolves in urban environments. Nightlight observed from space, e.g., VIIRS DNB satellite images, can be used in this regard, since it is correlated with socio-economic indicators like vehicle count, energy use, GDP, etc. [59]. Analyzing long-term air quality trends with nightlight brightness can inform about how the cause of pollution is changing with socio-economic development [60]. Similarly, AirRGB can also be analyzed with nightlight images observed from space. It can help in studying which of the three scenarios represented by AirRGB decomposition is affected by socio-economic development.

Acknowledgments: This research was in part supported by the Monbukagakusho (MEXT) scholarship.

Author Contributions: A.F. and W.T. conceived and designed the experiments; P.M. and A.F. performed the experiments; P.M. and W.T. analyzed the data; W.T. prepared the base datasets; P.M. wrote the paper and all authors read the paper and provided revision suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References


© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).