CLEAN Soil Air Water
www.clean-journal.com

Linking PM_{10} and $PM_{2.5}$ Pollution Concentration through Tree Coverage in Urban Areas

David Sierra-Porta, Yady Tatiana Solano-Correa, Miguel Tarazona-Alvarado, and Luis Alberto Nuñez de Villavicencio*

Particulate matter, PM_{10} and $PM_{2.5}$, represents common air pollutants in cities and constitute a considerable threat to public health impacting daily activity of people living in city. In large cities, the main sources of PM_{10} and $PM_{2.5}$ are diesel engine exhaust, brake dust, and particulate matter from vehicle tires. These particles can be deposited, filtered, and considerably reduced if there is a vegetative surface in the neighborhoods, thus eliminating a part of these particles and reducing their harmful footprint. This study evaluates the effect of tree coverage in urban areas on PM_{10} and $PM_{2.5}$ removal considering air quality monitoring stations. Estimation of tree coverage is made by using high spatial and temporal resolution satellite images from Planet constellations. An empirical relationship between these two variables, with an acceptable correlation ($R^2 = 0.478$ and $R^2 = 0.589$ for PM_{10} and $PM_{2.5}$, respectively), is obtained. A higher abundance of green space is associated with significantly lower PM_{10} and $PM_{2.5}$ values. Preliminary results suggest that the amount of tree coverage do cause some degree of air quality improvement and can be used to inform national clean air strategies aimed at reducing pollutant emissions.

1. Introduction

The existence and proliferation of fine particles, PM_{10} and $PM_{2.5}$, is regularly considered as a potential and actual health hazard. These particles can originate in open spaces due to many factors but can be diffused and spread to larger areas through wind geodesics in the atmosphere and can also invade indoor spaces. Relative particle sizes are measured in microns, and to measure the number of particles, it is usually expressed in units of $\mu g\ m^{-3}$.

These particles are especially hazardous to health because they are so fine that they can easily penetrate the respiratory tract. Inhalation of these particles mainly causes chronic bronchitis and asthma. The health impact of fine particles is difficult to measure; however, according to many studies,[1] they reduce life expectancy and are responsible for many deaths each year in many countries.

Each country is responsible for setting maximum allowable limits and health risk limits for prolonged exposure to these particles. However, on December 14, 2012, the United States Environmental Protection Agency (EPA) finalized an update to the National Ambient Air Quality Standard for $PM_{2.5}$ and PM_{10}.[2] Currently, the EPA has primary and secondary standards for $PM_{2.5}$ (annual average standards with levels of 12.0 and 15.0 $\mu g\ m^{-3}$, respectively; 24-h standards with 98th percentile shapes and levels of 35 $\mu g\ m^{-3}$) and the 24-h PM_{10} standard unchanged at 150 $\mu g\ m^{-3}$.

There are three main sources of pollution for the production of this type of particulate matter:[3,4,5] i) road traffic (diesel particles) including brake and tire wear: In many countries in Europe and the United States, steps have been taken in the renewal of automobile traffic with electric and hybrid engines, reducing these pollution factors a little, but Latin America still persists with a very high dependence on this fuel where practically all transportation means use it; ii) the burning of trees and wood combustion, biomass burning in general, fires, etc.; and iii) Industry, responsible for a third of the particles in the air.

In recent decades, there has been much concern about the environmental and health problem caused by air pollution from particulate matter PM_{10} and $PM_{2.5}$, and it is a challenge for country authorities to keep exposure situations to a minimum as increasing global emissions threaten ecosystems[6,7,8] in our daily lives. Even with few statistics so far, partial but careful estimates have shown that today there is greater risk and exposure. Many of the causes of disease and death worldwide are the result of poor air quality, and there is evidence of great vulnerability to human health,[9,10,11] and it has been demonstrated that during the COVID-19 pandemic this has been even more dramatic.[12,13,14] According to results from the World Health Organization (WHO), there is a close quantitative relationship between exposure to high concentrations of small particles (PM_{10} and $PM_{2.5}$) and increased daily and long-term mortality or morbidity.[15,16] Conversely, when concentrations of small and

D. Sierra-Porta, Y. T. Solano-Correa
Faculty of Basic Sciences
Universidad Tecnológica de Bolívar
Parque Industrial y Tecnológico Carlos Vélez Pombo
Km 1 Vía Turbaco, Cartagena de Indias, Cartagena 130010, Colombia
E-mail: dporta@utb.edu.co
M. Tarazona-Alvarado, L. A. N. de Villavicencio
School of Physics
Universidad Industrial de Santander
Car 27 \#9, Bucaramanga, Santander 680001, Colombia

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/clen.202200222

DOI: 10.1002/clen.202200222

fine particles are reduced, the associated mortality also falls, assuming other factors remain unchanged. This allows policymakers to make projections regarding the improvement in population health that could be expected if particulate air pollution is reduced. Particulate pollution has health effects even at very low concentrations. In fact, no threshold below which health harms have not been observed has been identified. It is estimated that ambient air pollution, both in cities and rural areas, is responsible for 4.2 million premature deaths worldwide per year; this mortality is due to exposure to small particles of 2.5 microns or less in diameter ($PM_{2.5}$), which cause cardiovascular and respiratory diseases and cancer.[15]

PM_{10} and $PM_{2.5}$ particles can be transported by the air and deposited on the surface of vegetation, resulting in a partial removal of these particles and eliminating this fraction of the atmosphere.[17,18] From the above, the surrounding vegetation in urban areas can act as a natural filter for this type of particles, and of course, a higher density of trees and vegetation can contribute to increase the amount of effective surface area of pollutants deposition. The problem is not trivial, however, since vegetation can also redirect and alter the airflow such that some areas may have higher concentrations of particulate pollutants.

Reductions in particulate matter concentration have been studied from different perspectives, yielding results of spatial-temporal dependence considering land areas covered by trees[19] in the case of inferring indirect relationships for the effect of air pollution on the occurrence of asthma and other respiratory diseases. This relationship has also been inferred in terms of the number of trees in populated cities or, for example, as reported by Xu et al.[20] for a direct measurement of trees' capacity (in cities) to remove particulate matter concentrations for PM_{10}, NO_2, and SO_2. The amount of surfaces destined to the construction of buildings and houses has also been considered. Some studies make this type of relationship possible also by using satellite imagery and also web-based applications to determine the capacity of trees (shrubs and herbaceous vegetation) to absorb some of the pollution generated by humans in daily life and industrial process facilities.[21,22,23,24,25] Similarly, for example, a recent study[26] has demonstrated a reduction of up to 3 kg $PM_{2.5}$ evaluating particulate matter concentrations in areas covered with approximately 80 trees in the Drumcondra tree alley in Dublin at different times of the day for 5 days representing rush and non-rush hours.

Another example[27] involves modeling PM_{10} in the West Midlands and Glasgow (UK), concluding that particulate matter concentrations could be reduced by up to 10%, thus providing evidence that green spaces drive changes and modifications in PM_{10} concentrations, including particle trapping on leaf surfaces.[28]

Some other studies make use of much more complicated models for spatio-temporal estimation of particle deposition on vegetation-covered surfaces using tree coverage data, hourly weather and pollution concentration to quantify the hourly amount of pollution removal, and the corresponding percentage of improvement in air quality.[29,30,31,32] In this case, they used a model known as i-Tree Eco.[33,34,35] The model is complicated and uses many variables to determine the atmospheric pollutant deposition rate which varies with wind speed. A fairly detailed description of how the model works for a particular pollutant can be found in Hirabayashi et al.[36]. As seen recently in a report of Lin et al.,[32] the model could have high uncertainties that could be at best 12.3% for leaf area, 13.4% for carbon storage, 11.1% for carbon sequestration, 40.7% for isoprene emissions, and 25.0% for monoterpene emissions. The model also seems to work under U.S. climate conditions; however, its application in other places, such as Europe, is really uncertain and even more so in Latin America dominated by particular dynamics different from the rest of the world. This makes the model inapplicable to the Americas, thus not being possible to use it for decision making. An additional limitation has to do with the availability of labeled data for tree species in cities and also a stringent need to have many air quality monitoring stations in a small study area. In addition, the consolidation of meteorological data is necessary for the calibration and final implementation of the model. Such sort of data is not usually available for Latin American countries, making it difficult or impossible to apply them in current studies.

In America, particularly in Colombia, there is a limitation of data from monitoring stations that have completeness of measurements of particulate matter concentrations, not to mention meteorological data at the same site of the air quality monitoring stations and/or in neighboring areas. Each region has ensured the consolidation and completeness of monitoring stations in the main cities; however, the goal remains complicated. Additionally, the mechanisms for estimating the amount of trees and vegetation are also limitations mainly due to the lack of resources necessary for accurate and efficient remote sensing in the territory. Different from developing countries, no forest inventories exist at national or regional levels with enough spatial resolution to perform detailed analysis. Access to financial resources to invest in technology and the acquisition and use of this technology is still a pending problem in most Latin American countries. Freely available data from NASA and ESA constellation exist but do not comply with the spatial resolution required for mapping tree coverage at city scales. All the above, together with the lack of enough computational resources, makes it necessary to present simple models that allow estimating the amount of particulate matter reduction with respect to tree coverage in cities. Such information could help in decision making and establishing mechanisms and goals for the reorganization and adaptation of green spaces, at least in the main and most populated cities.

Accordingly, this research aims at taking as reference several monitoring stations of PM_{10} concentrations (located in Bogotá, Colombia) in order to obtain an empirical relationship for the dependence of particulate matter concentrations and tree coverage. The model is meant to be simple, yet precise and exploits freely available tools combined with high spatial resolution data to guarantee tree coverage mapping. Based on observations of a 4-month period of PM_{10} and $PM_{2.5}$ concentration data and using satellite imagery to measure the tree coverage on the surface around 20 monitoring stations in a major city in Colombia, we attempt to provide a relationship for PM_{10} concentration that allows extrapolating a reduction ratio of particulate matter depending on tree coverage.

2. Experimental Section

Bogota (4° 42' 40'' N, 74° 4' 20'' W), capital of the Republic of Colombia and the Department of Cundinamarca, as well as being the political, administrative, economic, industrial,

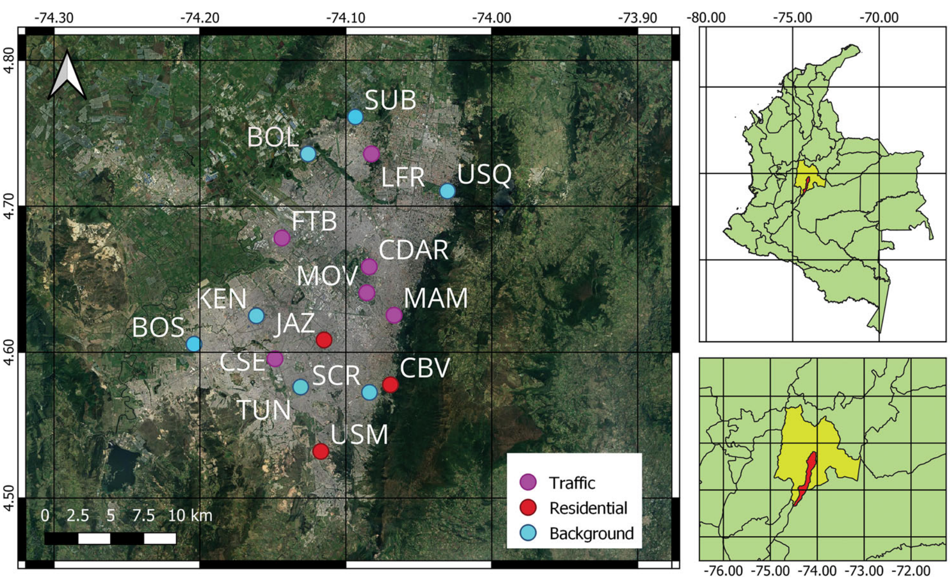

Figure 1. Geographical distribution of monitoring stations. On the map panel, monitoring stations in Bogotá, the distance between the two farthest stations is 25.88 km. The total area of the stations included in this study covers about 400 km^2. Visualization is performed using Folium 0.12.1 in Python 3.7. Google Earth. Copyright 2022 Maxar Technologies, CNES/Airbus.

artistic, cultural, sports and, of course, tourist epicenter of the country, is also the fourth most populous city in South America[37] (see also www.worldatlas.com/articles/biggest-cities-in-south-america.html). It is a diverse and multicultural city in which modern constructions are combined with those of the colonial past. In addition, it is important to note that thanks to its hills and parks, it is a green city dominated by large extensions of vegetation in some parts of its geography, mainly those that are farther away from the city center. Toward the center of the city, constructions and buildings dominate the landscape. It covers a total area of 1587 km^2 and an average altitude of 2640 m above sea level. Bogotá has an estimated population of 7 412 566 inhabitants in the capital area while it could reach up to 10 700 000 in the metropolitan area, including the neighboring municipalities on the suburbs of the city. The average temperature is 14.5 °C (58 °F), varying from 6 to 19 °C (43 to 66 °F) on sunny days to 10 to 18 °C (50 to 64 °F) on rainy days. Dry and rainy seasons alternate throughout the year. The traffic infrastructure cannot handle the masses, and the old buses sometimes let entire streets go down in thick smoke.

2.1. Material Particulate Concentrations Data

PM_{10} and $PM_{2.5}$ concentration data have been acquired from 18 stations located in Bogotá. The monitoring stations of the Bogota Air Quality Monitoring Network (Red de Monitoreo de Calidad del Aire de Bogotá, RMCAB) have been operating since 1997 and consist of monitors, analyzers, and automatic sensors that collect hourly data on the state of air quality in Bogota. This information is stored and sent via Internet to the central server of the Secretariat of Environment, where it is pre-filled and published in real time on the entity's web page. The network's monitors and analyzers operate under specific measurement methods established in Title 40 of the Code of Federal Regulations, which are approved by the U.S. EPA. For each pollutant, a specific reference method is defined, according to the equivalent technique for the operation of each monitor.[38]

Daily air quality data (PM_{10} and $PM_{2.5}$ concentrations) were acquired throughout 2020 and 2021. Meteorological data, including precipitation, wind direction, and wind speed, were acquired from the Institute of Hydrology, Meteorology and Environmental Studies (Instituto de Hidrología, Meteorología y Estudios Ambientales, IDEAM) meteorological partner stations. These data were used to calculate the monthly mean, maximum, minimum, median, and standard deviations of PM_{10} and $PM_{2.5}$ concentrations. **Figure 1** and **Table 1** show the main geographical information and distribution of the monitoring stations available in this study.

The selected data correspond to daily averages in the time interval between October 2020 and March 2021. **Figure 2** shows the distribution of daily PM_{10} measurements for each of the stations. Based on the observation in Figure 2, we see that the concentrations are not normally distributed; they rather closely follow a Maxwell–Boltzmann type distribution. Due to this observation, for the purposes of calculating the averages in this period, we determined the cumulative distribution of concentrations and noticed that it can be well fitted with a distribution such that:

$$P\ (X\ =\ x) = \sqrt{\frac{2}{\pi}}\ \frac{x^2 \exp\left(-x^2/2a^2\right)}{a^3} \tag{1}$$

and cumulative probability density function:

$$P\ (X < x) =\ \mathrm{erf}\left(\frac{x}{\sqrt{2}a}\right) - \sqrt{\frac{2}{\pi}}\frac{x^2 \exp\left(-x^2/2a^2\right)}{a} \tag{2}$$

where erf is the error function (also called the Gauss error function), and in this case, a is the only free parameter, and we call it

Table 1. Information about distribution of monitoring stations in Bogotá; the distance between the two farthest stations is 25.88 km.

Code	Latitude (N)	Longitude (W)	Altitude (m.a.s.l)	Kind
GYR	4.783756	−74.041683	2580	Background
USQ	4.710350	−74.030417	2570	Background
SUB	4.761247	−74.093461	2571	Background
BOL	4.735867	−74.125883	2574	Background
LFR	4.735867	−74.082483	2552	Traffic
FTB	4.678183	−74.143819	2621	Traffic
PTE	4.631767	−74.117483	2590	Traffic
KEN	4.625050	−74.161333	2580	Background
CSE	4.595617	−74.148583	2563	Traffic
TUN	4.576225	−74.130956	2589	Background
JAZ	4.608500	−74.114944	2559	Residential
USM	4.532056	−74.117139	2593	Residential
BOS	4.605611	−74.204056	2546	Background
CBV	4.577806	−74.069444	2661	Residential
CDAR	4.658467	−74.083967	2552	Traffic
MAM	4.625486	−74.066981	2621	Traffic
SCR	4.572553	−74.083814	2688	Background
MOV	4.642431	−74.083967	2583	Traffic

Altitude is expressed in meters above sea level (m.a.s.l). The total area of the stations included in this study covers about 400 km^2. The last column specifies the type of measurement station (background, traffic, and residential).

scale parameter and it is related to the mean and standard deviation so that:

$$\mu = 2a\sqrt{\frac{2}{\pi}}, \sigma = \sqrt{\frac{a^2(3\pi - 8)}{\pi}} \tag{3}$$

From Equation (3), averages and their corresponding uncertainties are calculated for particulate matter concentrations over the period under study. In addition to determining the mean particulate matter concentrations, we considered estimating the number of days that the daily measurement exceeded the mean at each station. According to the results, we observed that for PM_{10}, 48% of the days the measured concentration exceeded the mean with a standard deviation of 2%. Similarly, for $PM_{2.5}$, the results are 47% ± 3% in which the concentration exceeded the mean. The minimum and maximum concentrations for the monitoring stations are in the range of $(PM_{10}^{min}, PM_{10}^{max}) = (9 \pm 4, 69 \pm 19)$ μg m^{-3} and $(PM_{2.5}^{min}, PM_{2.5}^{max}) = (4 \pm 2, 34 \pm 6)$ μg m^{-3}, respectively

2.2. Satellite Imagery: Spatial Distribution and Tree Coverage Mapping

For this study, tree coverage in the vicinity (500 m box) of the environmental monitoring stations was measured through high spatial resolution satellite images acquired by Planet Scope constellations. Planet Scope (or Dove) is a constellation of 200 nanosatellites owned by the imaging company Planet Labs (www.planet.com), launched in 2017 that provides multispectral data on a daily basis and with a spatial resolution of 3 m. The initial satellites acquired data with four multispectral bands: red, green, blue, and near-infrared, but newest (2022) constellations acquire data with up to eight bands. Each multispectral image covers an area of 24 × 8 km requiring to applying homogenization steps between images acquired by different satellites on the constellation. Even though Planet data is of commercial purposes, many of its products are free for research purposes (https://www.planet.com), thus being possible to use them for the type of analysis presented in this research (in developing countries).

Due to Bogotá's location, cloud coverage is usually persisting all across the year. This makes finding cloud-free images over a given area for any time of the year a critical task. Thanks to the daily temporality of Planet Scope data, it was possible to find free cloud images covering each of the stations´ location in the period January–February 2021. This period goes in line with the acquisition period for the analyzed stations. As seen in Figure 1, the distribution of the stations makes it necessary to use more than one image to apply an algorithm for tree's detection. This is translated in the requirement to homogenize different images,

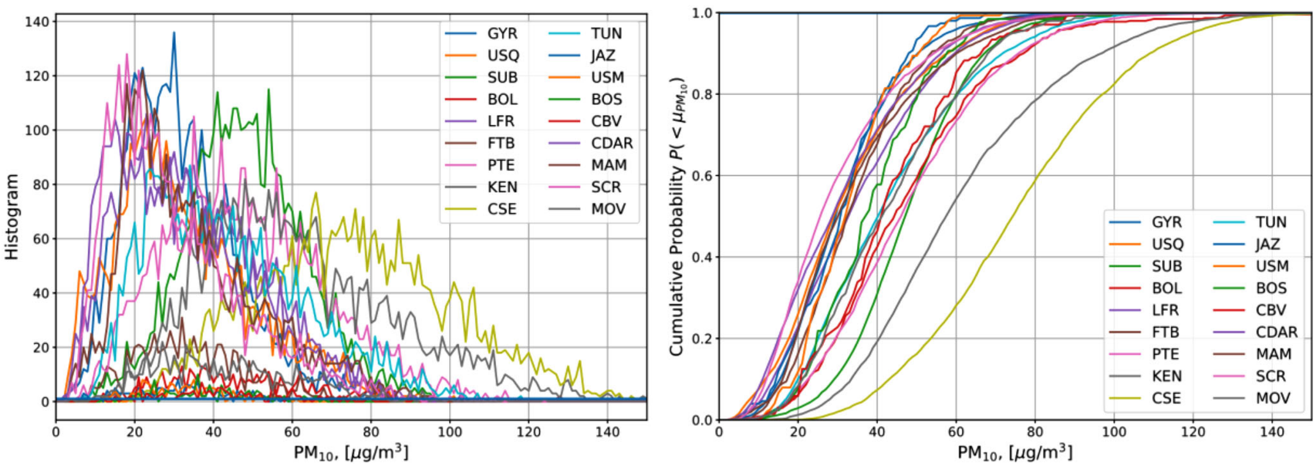

Figure 2. Absolute and cumulative distribution for measured PM_{10} concentrations from analyzed data from Bogotá stations. The distribution of measurements is well fitted to a Maxwell–Boltzmann distribution with free parameter a.

and this is done by applying a relative radiometric normalization based on histogram matching across different images.[39]

Once images have been homogenized, the first step is detecting vegetation and separating it from any other type of land cover. Vegetation detection here is the process of finding pixels and regions of green color, or rather a range of green colors in the image. In order to carry out this step, the normalized difference vegetation index (NDVI) is used. Vegetation indices are combinations of spectral bands recorded by remote sensing satellites, whose function is to enhance vegetation according to its spectral response and to attenuate the details of other elements such as soil, illumination, and water, among others. They are images calculated from algebraic operations between different spectral bands. The result of these operations allows obtaining a new image where certain pixels related to parameters of vegetation cover are graphically highlighted.

The NDVI is a vegetation index used to estimate the quantity, quality, and development of vegetation based on the measurement of the intensity of radiation from certain bands of the electromagnetic spectrum that vegetation emits or reflects.[40] For the calculation of vegetation indices, the information found in the red and near-infrared bands of the electromagnetic spectrum is necessary. NDVI is calculated using the following equation:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{RED}}{\mathrm{NIR} + \mathrm{RED}} \tag{4}$$

From Equation (4), it is possible to understand that NDVI varies in the range [–1,1], with vegetation being all NDVI values >0.

Once vegetation has been separated from other land covers, images are checked in order to evaluate the results and correct possible issues remaining due to large presence of small lakes/ponds in the area with high vegetation content on the surface. Such lakes/ponds get confused with vegetation. Thus, another spectral index, the normalized difference water index (NDWI), is used in order to separate water from vegetation.[41] Similar to the NDVI, NDWI is calculated using the following equation:

$$\mathrm{NDWI} = \frac{\mathrm{GREEN} - \mathrm{NIR}}{\mathrm{GREEN} + \mathrm{NIR}} \tag{5}$$

From Equation (5), it is possible to understand that NDWI varies also in the range [–1,1], with water content being all NDWI values >0.

Once vegetation and water bodies with surface vegetation have been separated, the last step is to separate trees from any other type of vegetation. To this aim, the green band is used by exploiting the higher radiometric response from trees (with reference to other types of vegetation), which is higher due to vegetation density. The threshold is automatically selected by using the Otsu's thresholding method,[42] and it is different from one station (image) to another due to difference in tree's coverage in the areas. Finally, tree coverage is quantified according to the number of pixels detected as such and according to the spatial resolution of Planet Scope data (3 m × 3 m). In order to measure the precision of results in tree detection, we used the Intersection over Union method,[43] where precision is provided as a percentage.

Table 2. Estimated values for number of pixels in image (1 × 1 km^2) and tree coverage in square meters for each of the stations using multispectral band and processing analysis.

Name	Code	Number of pixels	Coverage [m^2]	Kind
Guaymaral	GYR	7811	70 299	Background
Usaquén	USQ	22 435	201 915	Background
Suba	SUB	10 429	93 861	Background
Bolivia	BOL	2535	22 815	Background
Las Ferias	LFR	12 751	114 759	Traffic
Fontibón	FTB	804	7236	Traffic
Puente Aranda	PTE	1936	17 424	Traffic
Kennedy	KEN	1310	11 790	Background
Carvajal	CSE	843	7587	Traffic
Tunal	TUN	2100	18 900	Background
El Jazmín	JAZ	2033	18 297	Residential
Usme	USM	6576	59 184	Residential
Bosa	BOS	3935	35 415	Background
Centro de alto rendimiento	CDAR	11 700	105 300	Traffic
MinAmbiente	MAM	11 936	36 828	Traffic
San Cristobal	SCR	2577	23 193	Background
Movil 7ma	MOV	4092	36 828	Traffic

3. Results

In order to apply the proposed models for quantifying particulate matter concentrations and their correlation to the presence of trees, tree coverage detection was first carried out by means of the proposed method in Section 2.2. The detailed results can be found in **Table 2**, where the wide variability of tree coverage across the different stations location can be noticed. In order to validate the results, tree coverage information was created by photo interpretation for two stations (CDAR and LFR). These stations were selected because of showing extreme cases of analysis, the first one with high tree coverage and presence of a lake/pond with vegetation on the surface, whereas the second station presents low tree coverage and mainly buildings in the surrounding areas.

The obtained accuracies were of 95% and 97%, respectively, and detailed results can be found in **Figure 3**, where the panels (a) and (c) in Figure 3 show the RGB images for CDAR and LFR stations, whereas panels (b) and (d) in Figure 3 show the same image with all the detected trees highlighted in yellow color for comparison. The remaining error, which is low, is attributed to the spatial resolution of the images used, since some trees are smaller and can get confused with other sorts of information. Yet, and as expected, the results are reliable from both a quantitative and qualitative perspective.

From the data collected (and presented above), we calculated the exceedance of pollution days defined as the number of days in which the 24-h average of particulate matter is above a critical range of total suspended particulate matter. The results clearly show that at stations where there is an abundance of trees around the monitoring stations, the number of pollution days is significantly lower than at sites with insufficient vegetation cover. This can be seen in **Figure 4**.

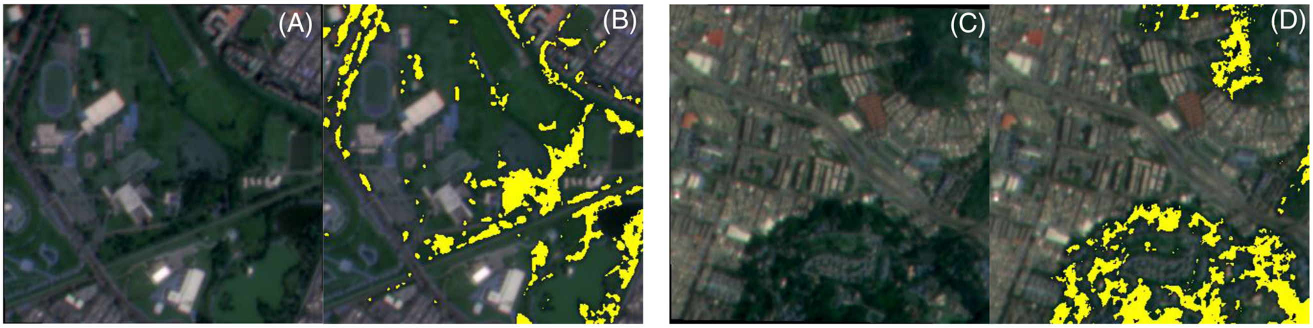

Figure 3. Example of original capture (a,c) and detection result (b,d) with yellow representing detected trees by proposed method for CDAR monitoring station panels (a) and (b) with 95% accuracy. In panels (c) and (d), the same as above but for LFR monitoring station with 97% accuracy. In all cases, the box has 500 m size around the monitoring station.

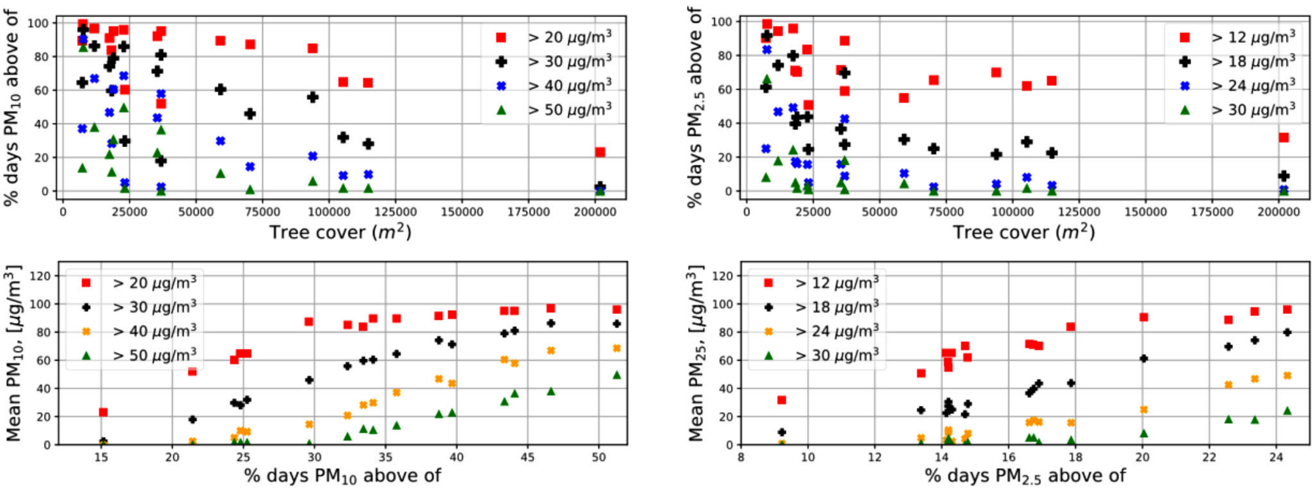

Figure 4. Average number of exceedance days of particulate matter from stations related with tree cover (upper panels) and average particulate matter (bottom panels.).

In the case of PM_{10}, almost all stations exceeded concentrations of >30 μgm^{-3} on more than 40% of the days, while in the case of $PM_{2.5}$, half of the stations exceeded concentrations of >18 μg m^{-3} on more than 30% of the days.

From the different analysis, it was found that the tree coverage amount in the vicinity of monitoring stations has resulted in a filtering effect on the concentrations of particulate matter; that is, when the trees' coverage increases, the average concentration of particulate matter decreases. Indeed, it becomes evident that this effect is also present when considering the number of days in which concentrations exceeded certain critical measurement thresholds. For example, in the case of PM_{10}, it is found that when concentrations exceeded thresholds of 10, 20, 30, 40, and 50 mg, a 90% decrease in tree cover (reduction between stations with lower and higher tree population) can increase the number of days exceeding these thresholds by 27%, 76%, 97%, 100%, and 100%, with averages of 95%, 79%, 55%, 31%, and 15%, respectively. Recall that the EPA standard for 24-h PM_{10} is 150 mg. Likewise, for $PM_{2.5}$, the same reduction suggests that for 6, 12, 18, 24, and 30 mg, the reductions are in the order of 39%, 67%, 89%, 98%, and 100%, with averages of 92%, 70%, 40%, 17%, and 6%, respectively. Recall that the EPA standard for 24-h $PM_{2.5}$ is 35 mg. In general, we can say that the city's air quality is good, with a very small percentage of stations exceeding the international permitted thresholds (see bottom panel in Figure 4).

No correlation is observed between the number of days that stations reported measurements above mean with tree coverage in the vicinity of the monitoring stations. In this case, we obtained correlation coefficients of $r = -0.10$ and $r = -0.33$ for PM_{10} and $PM_{2.5}$, respectively. Furthermore, there is evidence of a statistically significant effect of tree coverage dominance on maximum and minimum PM concentrations from analyzed data for these stations. We found that the correlation coefficients for these cases are $r = -0.44$ (PM_{10}^{min}), $r = -0.64$ (PM_{10}^{max}) and $r = -0.57$ ($PM_{2.5}^{min}$), $r = -0.50$ ($PM_{2.5}^{max}$).

As an important result, there was a clear signal that the abundance of green spaces affected mean PM_{10} concentrations ($r = -0.63$, p-value = 0.0061), as well as a significant effect on mean $PM_{2.5}$ concentration ($r = -0.62$, p-value = 0.0078), considering Pearson's correlation coefficient. In other words, a higher abundance of green space is associated with significantly lower PM_{10}

CLEAN Soil Air Water
www.clean-journal.com

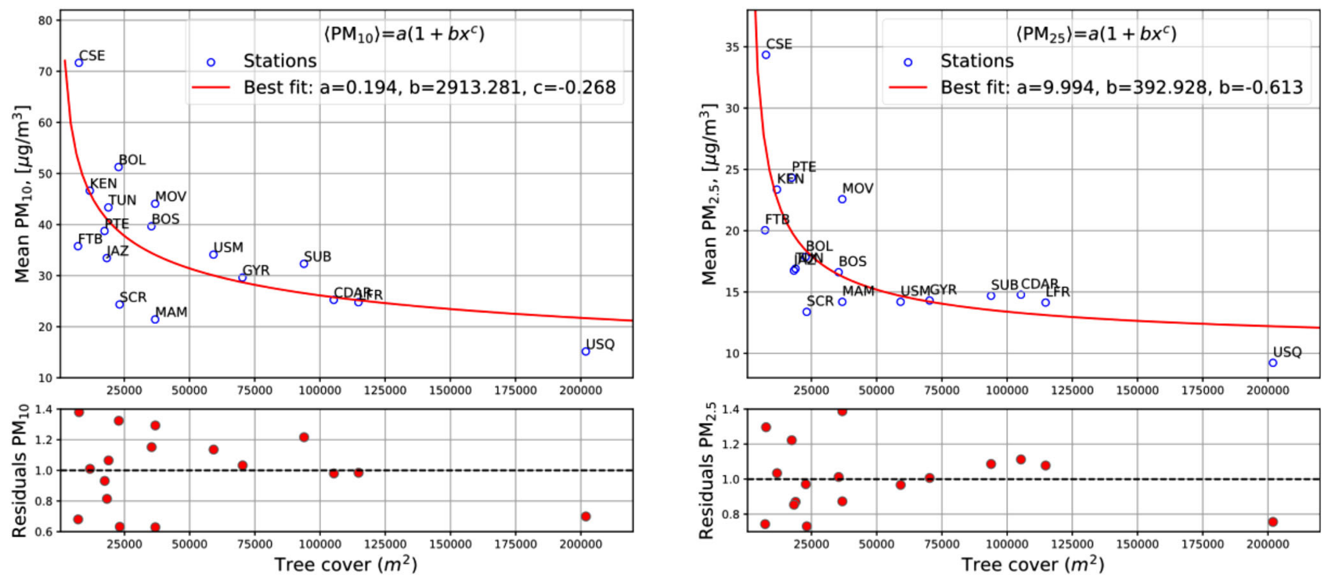

Figure 5. Dependence of mean particulate matter concentrations as a function of tree cover in the vicinity of each monitoring station. Residual is defined in the usual way as difference between data and fit values ($y_{\mathrm{data}} - y_{\mathrm{fit}}$).

Table 3. Goodness of fit in linear estimation of mean values in PM concentrations versus tree cover in the neighborhood of stations.

Parameter/ Contaminant	PM_{10}	$PM_{2.5}$
$a \pm \sigma_a$	$0.194 \pm 6.816 \times 10^{-5}$	$9.994 \pm 7.488 \times 10^{-6}$
$b \pm \sigma_b$	$2913.281 \pm 9.978 \times 10^{-1}$	$392.928 \pm 3.129 \times 10^{-5}$
$c \pm \sigma_c$	$-0.268 \pm 7.986 \times 10^{-6}$	$-0.613 \pm 4.801 \times 10^{-6}$
Determination coefficient of fit (R^2)	0.478	0.589
R^2-adjusted	0.445	0.555
Data/fit, MIN	0.6	0.7
Data/fit, MAX	1.4	1.4
Root mean squared error	9.349	3.655

and $PM_{2.5}$ concentrations. **Figure 5** shows the dependence of the mean concentrations for all the monitoring stations studied as a function of tree coverage obtained by satellite images centered on each station.

We attempted to reproduce a simple empirical model for the dependence of tree cover, measured in square meters, and the averages of PM_{10} and $PM_{2.5}$ concentrations. In this case, we have used a minimization model of a cost function defined as $\chi^2 \sim \Sigma(y_{\mathrm{data}} - y_{\mathrm{model}})^2$. For this, we have proposed the model such that $y_{\mathrm{model}}^{(\mathrm{PM10,PM2.5})} = a\,(1 + bx^c)$ as the best fit performed. The results are shown in **Table 3**.

All nonlinear regression models are statistically significant ($p < 0.01$ for a 95% confidence interval) with moderate high fitting precision. The R^2-value of each regression model varies from 0.478 to 0.589, and the adjusted R^2-value of cross validation results is a little lower, varying from 0.445 to 0.555 for both contaminants.

4. Discussions

From the analysis done in this study and the results presented in Figure 5, it is evident that the mean PM_{10} and $PM_{2.5}$ concentrations vary significantly according to the proportion of tree coverage in the area near and surrounding the monitoring stations. However, despite this observation and when considering the coefficient of variability of the stations ($= \sigma_{\mathrm{PM}}/\langle \mathrm{PM} \rangle$), the average coefficient is 37.19 ± 4.76 μg m^{-3} for PM_{10} and 40.97 ± 8.28 μg m^{-3} for $PM_{2.5}$. This suggests that the dispersion of concentrations is really large. This is because half of the stations considered are traffic stations (or very close to massive traffic sources[44,45,46,47]). So, those areas are affected by local sources of air pollution. However, this is not discriminatory in establishing a relationship with the number of trees in the area.

The discrimination of these separate concentrations in time intervals corresponding to peaks of automobile traffic activity has not been considered for this study. Nevertheless, it is evident that PM_{10} concentrations at least double when tree coverage is ten times lower and triple in the case of $PM_{2.5}$. This is in full correspondence and agreement with previous work.[26,48]

In terms of the established models, we can observe that a 50% decrease in tree coverage in square meters of trees implies (using the regression model) that there is an increase in pollutant concentration of about 20% and 15%, respectively, for PM_{10} and $PM_{2.5}$, while an increase in tree coverage implies a decrease in pollutant concentration of about 10% and 6%, respectively, for PM_{10} and $PM_{2.5}$. It is clear that the effect of trees is beneficial. However, the effect of their decrease in square meters is more drastic than the effect of further increasing the cover to be accounted for by the model.

In terms of quantitative modeling for the exploration of the relationship between land use (considering the number of trees in urban areas of cities) and air quality, we found that the biggest challenge is precisely the scarcity of good data with control monitoring stations of air quality in the cities. Therefore, to improve the robustness of the regression models limited by the scarcity of locations, we have collected a sequence of data that are available in one of the main cities in Colombia for the time interval in the years 2020 and 2021 with about 32 records from air quality monitoring stations. This allows for a greater association between locations and the time trend of air quality. Another difficulty is the limitation of accessibility to the data of other independent variables in each site. Additionally, the greatest limitation is the effective and precise identification of tree coverage in urban areas in Latin American countries. The geographic identification records of trees by the competent authorities are very scarce and are not properly updated; there are no databases of trees planted in urban and rural areas, at least in terms of usability, despite being a responsibility of local authorities to maintain control and inventory of these.

To overcome this last difficulty, satellite records can be used that could help predict said variable. However, this is another limitation. Despite having satellites and geographically labeled data, said data are generally not free, with restricted access, and in most cases, also due to the above, methodologies with very poor resolution data are used. In order to investigate this effect, we have obtained satellite data for the detection of tree coverage in urban areas in Colombia using high spatial resolution data (3 m) and automatic strategies with radiometric indices segmentation. Results were tested with photointerpretation information, showing detection accuracies above 95%.

Although other variables involved in the process of reducing air pollutants are necessary, here we have obtained a model based on data that describes at a moderately high level the ability of trees to reduce air pollution, at least for PM_{10} and $PM_{2.5}$. In order to obtain a more robust model, in addition to predicting the amount of tree coverage, some other variables such as the number of motor vehicles registered in each of the locations could be considered. The appropriate variable for emission or traffic volume is vehicle kilometers traveled (VKT) in each cell, which has been widely used in developed countries.[49] It may also be important to incorporate the number of companies at the sites to establish patterns of pollution sources when apportioning total releases to each location. Different size gradients for areas surrounding the stations could be also considered. All of the above can affect the results of the quantitative modeling of the impact of land use on air quality. This is a future work that will be carried out with an adequate experimental design.

Regarding the modeling methodology, in this study, the air quality and the geographical environment are ordered with approximately the same period. This is convenient and can be done in most of the tropics where the conditions of humidity, temperature, air temperature, pressure, precipitation, and other atmospheric variables do not vary significantly throughout the year, and only two well-differentiated seasons exist. However, changing the geographic environments could take some time to result in air quality changes. Although mean quarterly air quality concentrations are used, which are the aggregated impacts of geographic settings, the lagged effect of geographic settings on air quality may affect the association between them, which can be studied in further research. The nonlinear regression model that we have used to account for the quantitative relationship between land use impacts on air quality in our study is slightly different from others used in studies where linear regressions are used to model.[50,20] The independent variable, tree coverage, is accepted in the statistically significant model ($p < 0.01$ for a confidence interval of 95%).

However, some variables have not been taken into account, such as the determination of the area covered by constructions and buildings, nor the uncovered area. Although these factors affecting air quality were not considered in our study, nor were traffic, traffic volume, or the number of businesses in the areas, which may increase the uncertainty of the final regression tests of the models, these factors may be considered in a more in-depth study to be conducted soon.

In Bogotá, 1.2 million public trees are planted on roads, parks, riverbanks, and sidewalks. The tallest trees are pines and eucalyptus (Pinacease, Cupressacease). Consequently, of the giants planted, these are the most common. Regardless of their size or species, they are not enough to mitigate the city's environmental footprint. The eight families of trees that have been planted the most in the capital, and which are the most common, taking into account their growth capacity, ability to clean the air, and resistance to pollution, among others, are: Fabacease (29.3%), Adoxaceae (19.3%), Pittosporacease (12.1%), Oleacease (9.2%), Cupressacease (8.1%), Myrtacease (7.7%), Rosacease (7.4%), and Bignoniacease (6.9%). The city government maintains a policy of maintenance and monitoring of the tree fauna to keep it always in progressive growth. Regardless of the size that plants can reach, the characteristics of the environment cause their growth to be higher or lower than average.

Trees are unquestionably the fundamental elements of Urban Ecosystems, which contain a diversity of interacting elements, each with specific requirements, for which cities must provide their spaces, where animals, plants, and infrastructure can coexist forming pleasant places for their inhabitants, even more so when we are facing a changing city in terms of mobility and infrastructure. The World Health Organization[51] states that there should be 9 m^2 of green areas per inhabitant in cities; we are currently facing a great challenge to maintain and increase the green areas of cities in the face of the changing landscape of modernization that is required to be at the forefront of the needs of the population that resides in them.[52,53]

Trees in cities, whether they are isolated or in areas such as wetlands, parks, riverbanks, or streams, are the first approach to nature for many citizens; they regulate the temperature with the shade they provide and by releasing water vapor through their leaves, they can lower the air temperature by roughly 2 to 8 °C.[54,55] Likewise, their shade reduces the evaporation of water from the lawn, and as they transpire, atmospheric humidity increases, requiring less irrigation, and finally, they become a natural filter for the absorption of some of particles on the leaf surface, significantly reducing the concentrations of these pollutants.[56,57,58]

Additionally, we have taken into consideration for the evaluation of tree coverage square boxes of satellite images of 1 km side. The trees studied in the tropics and in the cities have an average canopy diameter of 12 m^2. Moreover, some studies[59,60] estimate that in Colombia are about 44951 trees km^{-2}. This implies to have an average pollutant rate of PM_{10} of 40 μg m^{-3} and of 15 μg m^{-3} for $PM_{2.5}$; the required density, according to our results, would be around 13% and 29%, respectively. With these estimates, this implies that maintaining this pollutant rate would require about 5715 and 13 084 trees km-^{2} for PM_{10} and $PM_{2.5}$, respectively.

5. Conclusions and Final Comments

In this study, we have evaluated the potential and unquestionable effect that trees have on the cleanliness of urban air in cities and in particular evaluating their effect on the decrease of PM_{10} and $PM_{2.5}$ concentrations. The objective was focused on finding an empirical dependence of the mean pollution concentrations and the tree coverage in the vicinity of monitoring stations. We obtained a simple nonlinear model, as a first approximation with acceptable Pearson's coefficients for PM_{10} ($R^2 = 0.478$) and $PM_{2.5}$ ($R^2 = 0.589$). Thus, the potential of urban trees in removing traffic and background air pollution through deposition and improving air quality is investigated using 18 monitoring stations of the Bogotá-Colombia Air Quality Monitoring Network system. The results indicated that there are no significant differences, regardless of whether the stations are background or traffic stations, in the reduction of PM_{10} and $PM_{2.5}$ averages. Concentrations are significantly lower in places with a higher amount (density) of trees, and vice versa, higher concentrations occur in areas with a lower amount of trees (which implies a higher amount of infrastructure and buildings). Although this study provides an initial exploration of trees and air quality in Bogotá, it is necessary to address this same study now by stratifying the measurements by the hours of heaviest automobile traffic in the city and evaluate differences in mean concentrations at various time intervals. The results of this study confirm that a greater abundance of trees in the city could help mitigate the effects of high elevations of particulate matter concentrations, but more importantly, inform governments and decision makers about the importance of a sustained policy of tree planting and improvement of green spaces in the city as an important management to help achieve emission reduction goals in the future. Finally, the original data for the results in this manuscript are available at web direction,[61] so they are free to replicate and reproduce our results.

Acknowledgements

The authors are thankful to the Secretaría Distrital de Ambiente for the data provided and information about characteristics of the Bogota Air Quality Monitoring Network stations. The authors would like to thank Planet Labs Inc. for providing the images used in this study under the "Education and Research Program."

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available at https://doi.org/10.17632/z6wtdbkc5b.1.

Keywords

air quality, PM_{10}, $PM_{2.5}$, remote sensing, tree coverage

Received: July 10, 2022
Revised: February 3, 2023
Published online: April 5, 2023

[1] World Health Organization, *Air Quality Guidelines for Europe*, WHO Regional Office for Europe, Copenhagen, Denmark **2000**. https://apps.who.int/iris/handle/10665/107335.
[2] Environmental Protection Agency (EPA), *Fed Regist* **2020**, *85*. https://www.govinfo.gov/content/pkg/FR-2020-12-18/pdf/2020-27125.pdf.
[3] O. Ogunkunle, N. A. Ahmed, *Sustainability* **2021**, *13*, 5465.
[4] C.-Y. Lee, L.-Y. Lin, H.-C. Chuang, K.-F. Ho, K.-J. Chuang, *Atmosphere* **2022**, *13*, 631.
[5] M. Calvia, *Cleaner Responsible Consumption* **2022**, *7*, 100090.
[6] M. S. Hammer, A. van Donkelaar, C. Li, A. Lyapustin, A. M. Sayer, N. C. Hsu, R. C. Levy, M. J. Garay, O. V. Kalashnikova, R. A. Kahn, *Environ. Sci. Technol.* **2020**, *54*, 7879.
[7] P. A. Owusu, S. A. Sarkodie, *Sci. Total Environ.* **2020**, *742*, 140636.
[8] I. Eguiluz-Gracia, A. G. Mathioudakis, S. Bartel, S. J. H. Vijverberg, E. Fuertes, P. Comberiati, Y. S. Cai, P. V. Tomazic, Z. Diamant, J. Vestbo, *Allergy* **2020**, *75*, 2170.
[9] Y. Raziani, S. Raziani, *J. Chem. Rev.* **2021**, *3*, 83.
[10] A. Pandey, M. Brauer, M. L. Cropper, K. Balakrishnan, P. Mathur, S. Dey, B. Turkgulu, G. A. Kumar, M. Khare, G. Beig, *Lancet Planet Health* **2021**, *5*, e25.
[11] H. L. Brumberg, C. J. Karr, A. Bole, S. Ahdoot, S. J. Balk, A. S. Bernstein, L. G. Byron, P. J. Landrigan, S. M. Marcus, A. L. Nerlinger, *Pediatrics* **2021**, *147*, https://doi.org/10.1542/peds.2021-051484.
[12] I. E. Isphording, N. Pestel, *J Environ Econ Manage* **2021**, *108*, 102448.
[13] I. Hernandez Carballo, M. Bakola, D. Stuckler, *Environ Res.* **2022**, 114155. https://doi.org/10.1016/j.envres.2022.114155.
[14] K. Ravindra, T. Singh, S. Vardhan, A. Shrivastava, S. Singh, P. Kumar, S. Mor, *J. Infect. Public Health* **2022**, *15*, 187.
[15] M. E. Goldstone, *Air Qual. Clim. Change* **2015**, *49*, 35.
[16] Y. O. Khaniabadi, G. Goudarzi, S. M. Daryanoosh, A. Borgini, A. Tittarelli, A. De Marco, *Environ. Sci. Pollut. Res.* **2017**, *24*, 2781.
[17] K. Corada, H. Woodward, H. Alaraj, C. M. Collins, A. de Nazelle, *Environ Pollut* **2021**, *269*, 116104.
[18] B. Konćzak, M. Cempa, M. Deska, *Environ Pollut* **2021**, *268*, 115465.
[19] T. S. Eisenman, G. Churkina, S. P. Jariwala, P. Kumar, G. S. Lovasi, D. E. Pataki, K. R. Weinberger, T. H. Whitlow, *Landscape Urban Plann.* **2019**, *187*, 47.
[20] G. Xu, L. Jiao, S. Zhao, M. Yuan, X. Li, Y. Han, B. Zhang, T. Dong, *Atmosphere* **2016**, *7*, 62.
[21] C. Zafra, Y. Angel, E. Torres, *Atmos Pollut Res* **2017**, *8*, 660.
[22] P. J. Irga, M. D. Burchett, F. R. Torpy, *Atmos. Environ.* **2015**, *120*, 173.
[23] D. Velásquez Ciro, J. E. Cañón Barriga, I. C. Hoyos Rincón, *Environ Monit Assess* **2021**, *7*, 396.
[24] A. Russo, W. T. Chan, G. T. Cirella, *Land* **2021**, *10*, 788.
[25] W. Lin, Y. Sun, D. Wang, Y. Li, X. Yu, *Sustainable Cities Soc.* **2021**, *70*, 102915.
[26] E. Riondato, F. Pilla, A. Sarkar Basu, B. Basu, *Sustainable Cities Soc.* **2020**, *61*, 102356.
[27] A. G. McDonald, W. J. Bealey, D. Fowler, U. Dragosits, U. Skiba, R. I. Smith, R. G. Donovan, H. E. Brett, C. N. Hewitt, E. Nemitz, *Atmos. Environ.* **2007**, *41*, 8455.
[28] W. Zhang, B. Wang, X. Niu, *Forests* **2017**, *8*, 92.
[29] W. Selmi, C. Weber, E. Rivière, N. Blond, L. Mehdi, D. Nowak, *Urban For. Urban Greening* **2016**, *17*, 192.
[30] J. Wu, Y. Wang, S. Qiu, J. Peng, *Sci. Total Environ.* **2019**, *688*, 673.
[31] M. T. I. Cabaraban, C. N. Kroll, S. Hirabayashi, D. J. Nowak, *Environ Pollut* **2013**, *176*, 123.

[32] J. Lin, C. N. Kroll, D. J. Nowak, *Urban For. Urban Greening* **2021**, *60*, 127062.
[33] D. J. Nowak, D. E. Crane, J. C. Stevens, *Urban For. Urban Greening* **2006**, *4*, 115.
[34] D. J. Nowak, S. Hirabayashi, A. Bodine, R. Hoehn, *Environ Pollut* **2013**, *178*, 395.
[35] D. J. Nowak, S. Hirabayashi, A. Bodine, E. Greenfield, *Environ Pollut* **2014**, *193*, 119.
[36] S. Hirabayashi, C. N. Kroll, D. J. Nowak, *Citeseer* **2012**, *36*. https://www.itreetools.org/documents/60/i-Tree_Eco_Dry_Deposition_Model_Descriptions_V1.5.pdf
[37] R. Weller, D. Gouverneur, Z. Drozdz, B. Ye, *J. Landscape Archit.* **2021**, *16*, 76.
[38] US Environmental Protection Agency, *Fed Regist. EPA Publication 530-SW-88-011*, **1988**, *53*, 33314.
[39] Y. T. Solano-Correa, F. Bovolo, L. Bruzzone, *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7579.
[40] J. W. Rouse, R. H. Haas, J. A. Schell, D. W. Deering, in *Third Earth Resources Technology Satellite-1 Symposium*, Vol. *1*: Technical Presentations Section A, NASA, Washington, DC **1974**, p. 309.
[41] B. C. Gao, *Remote Sens. Environ.* **1996**, *58*, 257.
[42] M. Sezgin, B. Sankur, *J Electron Imaging* **2004**, *13*, 146.
[43] N. C. Chung, B. Z. Miasojedow, M. Startek, A. Gambin, *BMC Bioinf.* **2019**, *15*, 1.
[44] Y. Li, G. Jiang, Y. Wang, Z. Cai, Q. Zhang, *Atmos. Environ.* **2008**, *42*, 2037.
[45] Y. Li, P. Wang, L. Ding, X. Li, T. Wang, Q. Zhang, H. Yang, G. Jiang, F. Wei, *Chemosphere* **2010**, *79*, 253.
[46] E. Menichini, N. Iacovella, F. Monfredini, L. Turrio-Baldassarri, *Chemosphere* **2007**, *69*, 422.
[47] E. Abad, K. Martínez, L. Gustems, R. Gomez, X. Guinart, I. Hernández, J. Rivera, *Chemosphere* **2007**, *67*, 1709.
[48] L. S. Vailshery, M. Jaganmohan, H. Nagendra, *Urban For. Urban Greening* **2013**, *12*, 408.
[49] P. Hidas, G. R. Shiran, J. A. Black, presented at the 30th International Symposium on Automotive Technology and Automation, Florence, Italy **1997**.
[50] L. P. Clark, D. B. Millet, J. D. Marshall, *Environ. Sci. Technol.* **2011**, *45*, 7028.
[51] World Health Organization, *Urban Planning, Environment and Health: From Evidence to Policy Action*, WHO, Geneva, Switzerland **2010**. https://www.cabdirect.org/cabdirect/abstract/20103308516
[52] A. Addas, G. Alserayhi, *SAGE Open* **2020**, *10*, 215824402092060.
[53] W. Kuang, Y. Dou, *Remote Sens.* **2020**, *12*, 1929.
[54] World Health Organization, *Urban Green Spaces: A Brief for Action*, WHO, Geneva, Switzerland **2017**. https://apps.who.int/iris/handle/10665/344116
[55] T. T. Tura, T. Soromessa, S. Leta, M. Argew, Z. Eshetu, *Taiwania* **2016**, *61*, 305. https://doi.org/10.6165/tai.2016.61.305.
[56] T. Gao, F. Liu, Y. Wang, S. Mu, L. Qiu, *Forests* **2020**, *11*, 950.
[57] M. Chen, F. Dai, B. Yang, S. Zhu, *Building Environ.* **2019**, *158*, 1.
[58] S. Arghavani, H. Malakooti, A.-A. Ali Akbari Bidokhti, *J. Cleaner Prod.* **2020**, *261*, 121183.
[59] Y. M. Bar-On, R. Phillips, R. Milo, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 6506.
[60] T. W. Crowther, H. B. Glick, K. R. Covey, C. Bettigole, D. S. Maynard, S. M. Thomas, J. R. Smith, G. Hintler, M. C. Duguid, G. Amatulli, *Nature* **2015**, *525*, 201.
[61] D. Sierra Porta, *Dataset: Air quality Bogota. Mendeley Data*, V1, **2022**. https://doi.org/10.17632/z6wtdbkc5b.1.