Disentangling associations between vegetation greenness and dengue in a Latin American city: Findings and challenges


Observatory for Urban health in Belo Horizonte, School of Medicine, Federal University of Minas Gerais, Brazil
Pontifical Catholic University of Minas Gerais, Brazil
Institute of Urban and Regional Development, University of California, 316 Wurster Hall, University of California, Berkeley, Berkeley, CA 94720, USA
Belo Horizonte Municipal Health Department, Brazil
Department of Environmental Science, Policy, and Management, Department of Landscape Architecture and Environmental Planning, University of California, Berkeley, USA
Laboratory of Ecology of Diseases and Forests, Nucleus of Biology/NUPEB and Institute of Exact and Biological Sciences, Federal University of Ouro Preto, Brazil
Faculty of Veterinary Medicine, Federal University of Pelotas, Brazil
Federal Center for Technological Education of Minas Gerais, Brazil
Department of City and Regional Planning and Institute of Transportation Studies, University of California, Berkeley, USA

HIGHLIGHTS

- Vegetation greenness is negatively but insignificantly associated with dengue incidence.
- This association is stronger in socioeconomically vulnerable areas.
- Improved vegetation quality has the potential to reduce dengue risk.
- Openly accessible, coarse resolution environmental data have limitations in environmental health studies.

ARTICLE INFO

Keywords:
- Dengue
- Greenness
- Socioeconomic vulnerability
- Reemerging infectious disease
- Planning and policy implications

ABSTRACT

Being a Re-Emerging Infectious Disease, dengue causes 390 million cases globally and is prevalent in many urban areas in South America. Understanding the fine-scale relationships between dengue incidence and environmental and socioeconomic factors can guide improved disease prevention strategies. This ecological study examines the association between dengue incidence and satellite-based vegetation greenness in 3826 census tracts nested in 474 neighborhoods in Belo Horizonte, Brazil, during the 2010 dengue epidemic. To reduce potential bias in the estimated dengue-greenness association, we adjusted for socioeconomic vulnerability, population density, building height and density, land cover composition, elevation, weather patterns, and neighborhood random effects. We found that vegetation greenness was negatively associated with dengue incidence in a univariate model, and this association attenuated after controlling for additional covariates. The dengue-greenness association was modified by socioeconomic vulnerability: while a positive association was observed in the least vulnerable census tracts, the association was negative in the most vulnerable areas. Using greenness as a proxy for vegetation quality, our results show the potential of vegetation management in reducing dengue incidence, particularly in socioeconomically vulnerable areas. We also discuss the role of water infrastructure, sanitation services, and tree cover in lowering dengue risk.
1. Introduction

Approximately 390 million dengue cases occur globally every year (Bhatt et al., 2013). In South America, there is a high prevalence of dengue in many urban areas. However, dengue once appeared to be under control, and recurring outbreaks make it a reemerging infectious disease. The elimination of *Aedes aegypti*, the main dengue vector, by mosquito eradication programs effectively controlled the dengue epidemic in the Americas in the 1950s and 1960s (Gubler, 2011). However, dengue returned to the Antilles in 1960 and reached South America at the end of that decade. Since then, dengue has become more intense and has been affecting mainly large urban centers in different regions (Gubler, 2011; San Martín et al., 2010; Wilcox & Colwell, 2005). Unplanned urbanization, along with ineffective mosquito control, changes in lifestyles, globalization, and international travel, have contributed to dengue’s reemergence in both endemic and non-endemic areas (Buonanno et al., 2014; Gubler, 2011). *Aedes aegypti* is a domesticated mosquito that develops in man-made containers in the urban environment, and the mosquito’s abundance raises with urbanization (Higa, 2011; Powell & Tabachnick, 2013). Urban planning and management have the potential to reduce the risk of vector-borne diseases, through improving the built environment, green spaces, housing conditions, and sanitation (Degroote, Zinszer, & Ridde, 2018; Lee, 1994; Lindsay, Wilson, Golding, Scott, & Takken, 2017; Ogden, 2016).

Studying the associations between dengue incidence and the aforementioned factors at the urban neighborhood scale is critical to identify markers of elevated risk and opportunities for intervention (Spencer, Finucane, Fox, Saksena, & Sultana, 2020), refine disease prevention guidance, and increase resident’s participation in managing natural and human habitats.

Controlling emerging and reemerging infectious diseases, such as vector-borne diseases, including dengue, requires interdisciplinary knowledge and an expanded concept of causality in epidemiology (Ellis & Wilcox, 2009). Under these principles, researchers have developed frameworks including biocomplexity (Colwell, 1998; Wilcox & Colwell, 2005) and the Eco-Bio-Social strategy (de Macedo et al., 2021). Central to these frameworks is that environmental and socioeconomic conditions interact to create elevated risks of diseases such as dengue and the recent COVID-19. Dengue incidence rises with temperature, precipitation, and relative humidity (Xu et al., 2017). It is linked to vector biology: rising temperatures expand the spatial distribution of *Aedes aegypti* both globally (Bustamante et al., 2019) and locally (Bona, Twersdochilb, & Navarro-Silva, 2011; Degallier et al., 2003). *Aedes aegypti* is a highly domesticated mosquito that lays eggs in artificial containers such as old tires, water storage, and trash (Gubler, 1998). Therefore, socioeconomic conditions including lower household accessibility to piped water (Pedrosa et al., 2020; de Teixeira, 2008), increased domestic water storage (Cordeiro et al., 2011), irregular or no trash collection (Cordeiro et al., 2011; Honorato et al., 2014; Vargas et al., 2015), and lack of basic sanitation (Cordeiro et al., 2011; de Almeida & de Medronho, 2009), have been linked to a higher incidence of dengue. In addition, lack of effective vector control (Gubler, 2011), viral serotype, local population immunity (Marti et al., 2020; Salje et al., 2012), and human mobility (Cordeiro et al., 2011; Zhu, Liu, Tan, & Shi, 2016) have been associated with dengue incidence.

Urban areas are complex mosaics of heterogeneous environmental and socioeconomic conditions, making these areas hotspots of dengue. Focused and unplanned urbanization have composition of sub-standard sanitation services and underdeveloped water infrastructures, leading to favorable conditions for *Aedes aegypti* infestation. The high population density in cities also increases exposure to mosquitoes and the likelihood of dengue transmission. Together, these conditions make dengue primarily an urban disease that reemerges in large tropical urban centers at a faster frequency (Gubler, 2011). With the increasing transmission of vector-borne diseases in cities, public policy has turned to multisectoral approaches, including environmental management, expansion of sanitation services, housing improvement, and educational campaigns (Donaliso et al., 2017) to reduce disease risks. In this context, modifying the built and natural environment through vegetation management and land use planning have emerged in the urban planning sector as promising approaches to reduce the spread of vector-borne diseases such as dengue (Degroote et al., 2018; Lee, 1994; Lindsay et al., 2017; Ogden, 2016).

Vegetation management can be an effective strategy in highly urbanized areas characterized by high population density and dominance of built-up areas, where other interventions including land use control and housing improvement may be more challenging and costly. Observational studies have examined dengue incidence and vegetation within and between cities. The studies found a mixture of positive, negative, and non-linear associations (Araujo et al., 2015; Cao et al., 2017; Huang et al., 2018; Martínez-Bello, López-Quílez, & Torres Prieto, 2017; Penso-Campos, Fraga, Caldas, Sommer, & Périco, 2018; Qi et al., 2015; Sari, Adelwin, & Rinawan, 2020). These inconsistent findings can be attributed to different spatial scales of analysis, measures of vegetation, and analytical approaches. Experimental studies have found that vegetation management, including planting certain species (Jiannino & Walton, 2004) and optimizing planting configuration (Thullen, Sartoris, & Walton, 2002), can reduce mosquito populations. However, these studies (Jiannino & Walton, 2004; Thullen et al., 2002) were conducted in wetlands and did not use dengue incidence as an outcome. Therefore, the effectiveness of the proposed management strategies in reducing dengue risk in cities remains uncertain. Different vegetation types, including farmland, forests, and grassland, co-exist in cities. These vegetation types could increase or decrease the *Aedes aegypti* population due to changes in competing species (Medeiros-Sousa, Fernandes, Ceretti-Junior, Willeke, & Marrelli, 2017) and shading (Barrera, Amador, & Clark, 2006), consequently modifying dengue transmission patterns. Therefore, while local governments can employ vegetation management to reduce the risk of dengue transmission, effective ways to achieve this goal remain uncertain.

In this study, we examine Belo Horizonte, Brazil, which has experienced several dengue outbreaks since 1996. The city represents a complex spatial distribution of dengue cases and environmental and socioeconomic conditions (Fig. 1). Research has linked seasonal peaks in temperature and precipitation (Almeida, Assunção, Proietti, & Caiaffa, 2008; Horizonte, 2016; Campos et al., 2019), socioeconomic, demographic, and urban infrastructure characteristics (Bavia et al., 2020; Campos et al., 2019; de Mattos Almeida, Caiaffa, Assunção, & Proietti, 2007; Pessanha, Caiaffa, Kroon, & Proietti, 2010) with dengue incidence in Belo Horizonte. In 2010, twelve years after the previous epidemic, the city recorded an epidemic with the concurrent circulation of three viral serotypes (DENV-1, DENV-2, and DENV-3) and an annual incidence of 2053 cases per 100,000 inhabitants. Dengue generally exhibits a seasonal and cyclical behavior over the years due to climatic conditions, circulating viruses, and population susceptibility (Ximenes et al., 2016). Consequently, Belo Horizonte has experienced subsequent epidemics in 2013, 2016, and 2019, with progressively increasing incidence, disease severity, and fatality rates (Horizonte, 2020; Campos et al., 2019). These recurring outbreaks highlight the urgency of interventions, including those in vegetation management, to curb dengue risk.

Accordingly, to understand how vegetation management could reduce dengue risk, we investigated the association between vegetation greenness and dengue incidence in 3826 census tracts in Belo Horizonte, Brazil, during the 2016 dengue epidemic. To our knowledge, only one prior study has analyzed the relationship between dengue and vegetation at this scale (Martínez-Bello et al., 2017). This fine spatial scale is critical for communicating research findings with local decision-makers. We measured vegetation greenness using the Normalized Difference Vegetation Index (NDVI), a popular satellite-based indicator for vegetation coverage, health, and photosynthetic activities. NDVI has been widely used in environmental health studies (Rojas-Rueda, Niewenhuysen, Gascon, Perez-Leon, & Mudu, 2019).
We hypothesize that vegetation greenness has a negative association with dengue incidence, as higher greenness after controlling for vegetation coverage indicates healthier vegetation with the presence of predators and competing species of the *Aedes aegypti*, leading to a protective effect against dengue. We also explore whether the relationship between dengue incidence and greenness is modified by socioeconomic vulnerability. We hypothesize that more vulnerable areas would be at greater risk due to lower quality green spaces that may be a better habitat for the vector. To reduce potential omitted variable bias in the estimated relationships between dengue incidence and greenness, we controlled for a set of covariates, including socioeconomic vulnerability, population density, building height, percentage of census tract area covered by building footprints, land cover composition, elevation, and weather patterns. The findings are intended to inform urban planners, engineers, and public health professionals regarding vegetation management as a tool to mitigate the transmission of dengue and other arboviruses.

2. Data and methods

2.1. Study area

Our study area initially covers 3828 census tracts (setores censitários) nested in 474 neighborhoods (bairros) in Belo Horizonte, Brazil (Fig. 1). We dropped two census tracts due to missing land cover data, and we used the remaining 3826 census tracts throughout the study. As defined by the Brazilian Institute of Geography and Statistics (IBGE), a census tract is a continuous area located in a single urban or rural setting, with varying sizes and numbers of households (Bueno, 2017; IBGE, 2010). In Belo Horizonte, a census tract contains an average population of 5886 (Interquartile Range, IQR: 3800 – 7886) and an average area of 7.94 ha (IQR: 2.61 – 7.87). Neighborhood, officially defined by the city, is a cluster of adjacent census tracts by which the city communicates with its residents and provides services (Bairros de BH, 2009). We obtained neighborhood boundary data from BHGeo (https://bhgeo.pbh.gov.br/home). A neighborhood contains between 1 and 58 census tracts (Table 1).

2.2. Outcome: dengue incidence

The outcome variable is dengue incidence (cases per 100,000 population) in 2010. Dengue cases are reported to the Notifiable Diseases Information System, and we obtained the data for Belo Horizonte through the City Health Department. As part of the BH-VIVA database project (Friche, Dias, Reis, Dias, & Caiaffa et al., 2015), the Urban Health Observatory of Belo Horizonte performed data cleaning and consistency check, and it georeferenced each case to a census tract using the coordinates of the patient’s residence. Population counts for the census tracts in 2010 were obtained from WorldPop (Sorichetta et al., 2015).

2.3. Exposure: vegetation greenness

Our exposure variable is vegetation greenness derived from NDVI. NDVI represents the combined effect of vegetation quantity, including coverage and biomass, and quality, capturing photosynthetic activities and vegetation health (Tucker, 1979). NDVI takes values between –1 and 1, with higher values indicating the larger quantity, better quality of vegetation, or both.
We calculated vegetation greenness as the census-tract average of annual mean of monthly maximum NDVI values. We first computed NDVI every 16 days in 2010 from the red and near-infrared electromagnetic bands (630–690 nm and 760–900 nm, respectively) of the Landsat-5 TM satellite surface reflectance images. Landsat-5 TM images have 30 m spatial resolution and 16-day temporal frequency. We masked out image areas covered by cloud, shadow and water, which would otherwise cause bias in NDVI and greenness estimates. Then for each month, we generated a monthly maximum NDVI image reflecting the greenest condition of each image pixel and averaged the monthly images to produce an NDVI image for 2010. We calculated the census-tract average of this 2010 NDVI image, which we subsequently used as vegetation greenness in our analyses. While, by definition, NDVI ranges between –1 and 1, for better interpretation of the model results, we rescaled greenness to a range between -10 and 10 (Fig. 1c). We accessed Landsat-5 data and performed these calculations in Google Earth Engine, a cloud-based platform for remote sensing image processing (Gorelick, 2013).

Because NDVI measures vegetation quantity and quality simultaneously, a census tract with healthy green vegetation that occupies a small proportion of its area might have a similar greenness value to a fully vegetated census tract with stressed or senescent and therefore less green vegetation. To better estimate the effect of vegetation quality on dengue incidence, we introduced additional covariates controlling for vegetation quantity within a census tract. These covariates include the percentage of census tract covered by forest, non-forest natural land, farmland, and building footprints. These covariates, along with others, are introduced below.

### 2.4. Other covariates

To reduce potential bias in the estimated association between dengue incidence and greenness, we controlled for a comprehensive set of covariates including socioeconomic vulnerability, population density, building height, the percentage of census tract area covered by building footprints, land cover composition, elevation, weather patterns, and neighborhood random effects. We identified these covariates initially using a directed acyclic graph (DAG) (Fig. S2), and we narrowed them down to a final set by examining multicollinearity with variance inflation factor (VIF). We collected these covariates from the year 2010 when possible to match the timeframe of the outcome and exposure variables. We used the Health Vulnerability Index (HVI) to measure socioeconomic vulnerability. This composite index, used by the City Health Department (Secretariat, 2018), is based on the 2010 census and encompasses infrastructure indicators and the socioeconomic status of residents (Fig. 1d). HVI considers infrastructure conditions, including water supply, sewage connection, and garbage collection, and it measures socioeconomic status by literacy rate, poverty, income, and percentage of the population in different racial groups. HVI is on a scale from 0 to 10, with higher values indicating more vulnerable areas. HVI influences both greenness and dengue incidence: vulnerable areas are often less resourceful to build and manage green space (Casey, James, Cushing, Jedsale, & Morello-Frosch, 2017). These areas also have undesirable water and sanitation infrastructures, which are suitable for the reproduction of Aedes aegypti.

We included population density, building height, the percentage of census tract covered by building footprints (i.e. building density), and land cover composition to further characterize the urban environment. These factors could be modified by urban planning and affect vegetation greenness in several ways. Areas with dense populations and buildings may have less open space for vegetation, leading to lower greenness. Land cover composition, including forest, non-forest natural land, farming, urban, water, and others, helps to separate the effect of vegetation quality from quantity on greenness magnitude. These urban environment characteristics also influence dengue incidence. Large building density and the presence of high-rise buildings are indicative of better housing conditions and less on-site water storage in the study area, which reduce breeding sites for Aedes aegypti (Seidahmed, Lu, Chong, Ng, & Elnaih, 2018; Troyo, Fuller, Calderon-Aruguadas, Solano, & Beier, 2009). Land cover composition predetermines the suitability of Aedes aegypti habitats and the dynamics of the human population.

We calculated population density by dividing population counts from the 2010 WorldPop dataset (Soricbeta et al., 2015) by census tract area. We obtained height and footprint of individual buildings from BHGeo (https://bhgeo.pbh.gov.br/home). We retrieved 2010 land cover maps from the MapBiomass Brasil dataset (https://mapbiomas.org/en), a dataset that maps land cover annually with satellite remote sensing images at 30 m spatial resolution.

We also included land elevation and weather patterns to characterize the natural environment. These factors act as the background for vegetation development and influence Aedes aegypti population and dengue incidence. We approximated census-tract average elevation using the ALOS World 3D global digital surface model at 30 m resolution collected between 2006 and 2011 (Tadono et al., 2014). Although a digital surface model measures land elevation combined with the heights of aboveground objects (e.g., trees and buildings), this dataset was still valuable for this study as a proxy for land surface elevation, given the lack of spatial information on individual aboveground objects such as trees. For

### Table 1

Summary statistics of outcome variable, exposure variable, and covariates.

<table>
<thead>
<tr>
<th>Max.</th>
<th>Min.</th>
<th>Mean</th>
<th>Median</th>
<th>25th percentile</th>
<th>75th percentile</th>
<th>Std.</th>
<th>Std., between neighborhood</th>
<th>Std., within the neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence rate (cases/100,000 population)</td>
<td>6756.48</td>
<td>0.00</td>
<td>229.18</td>
<td>138.11</td>
<td>46.03</td>
<td>306.00</td>
<td>307.52</td>
<td>447.49</td>
</tr>
<tr>
<td>Greenness (0–10)</td>
<td>6.69</td>
<td>0.93</td>
<td>2.52</td>
<td>2.39</td>
<td>2.07</td>
<td>2.79</td>
<td>0.71</td>
<td>0.72</td>
</tr>
<tr>
<td>Vulnerability (HVI, 0–10)</td>
<td>6.96</td>
<td>0.47</td>
<td>2.41</td>
<td>2.35</td>
<td>1.70</td>
<td>2.93</td>
<td>0.92</td>
<td>1.06</td>
</tr>
<tr>
<td>Population density (people/ hectar)</td>
<td>8476.80</td>
<td>0.78</td>
<td>1306.15</td>
<td>1110.57</td>
<td>618.82</td>
<td>1578.70</td>
<td>808.10</td>
<td>701.56</td>
</tr>
<tr>
<td>Average building height (m)</td>
<td>96.25</td>
<td>2.43</td>
<td>7.11</td>
<td>5.24</td>
<td>4.74</td>
<td>6.65</td>
<td>5.49</td>
<td>3.15</td>
</tr>
<tr>
<td>% building footprint</td>
<td>80.92</td>
<td>0.19</td>
<td>40.59</td>
<td>34.82</td>
<td>34.53</td>
<td>48.13</td>
<td>12.11</td>
<td>12.03</td>
</tr>
<tr>
<td>% forest</td>
<td>68.60</td>
<td>0.00</td>
<td>0.60</td>
<td>0.00</td>
<td>0.00</td>
<td>4.29</td>
<td>3.93</td>
<td>3.61</td>
</tr>
<tr>
<td>% non-forest natural land</td>
<td>31.65</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.75</td>
<td>0.53</td>
</tr>
<tr>
<td>% farmland</td>
<td>71.09</td>
<td>0.00</td>
<td>1.06</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>5.75</td>
<td>6.76</td>
</tr>
<tr>
<td>Average elevation (m)</td>
<td>1235.49</td>
<td>697.23</td>
<td>867.52</td>
<td>867.51</td>
<td>817.75</td>
<td>924.97</td>
<td>74.95</td>
<td>82.63</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>73.59</td>
<td>70.14</td>
<td>71.35</td>
<td>71.54</td>
<td>71.11</td>
<td>71.90</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>Total precipitation (mm/year)</td>
<td>100.36</td>
<td>83.47</td>
<td>90.93</td>
<td>94.33</td>
<td>83.47</td>
<td>94.33</td>
<td>5.02</td>
<td>5.12</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>21.22</td>
<td>19.44</td>
<td>20.43</td>
<td>20.55</td>
<td>20.41</td>
<td>20.55</td>
<td>0.49</td>
<td>0.52</td>
</tr>
<tr>
<td>Census tract area (hectare)</td>
<td>1288.10</td>
<td>0.08</td>
<td>7.95</td>
<td>4.98</td>
<td>2.61</td>
<td>7.87</td>
<td>27.82</td>
<td>35.15</td>
</tr>
<tr>
<td>Total population (people)</td>
<td>24,298</td>
<td>32</td>
<td>5889</td>
<td>5869</td>
<td>3801</td>
<td>7887</td>
<td>2999.81</td>
<td>2457.93</td>
</tr>
</tbody>
</table>

Number of census tracts = 3826, Number neighborhoods = 474, each containing between 1 and 58 census tracts.
climate variables, we included census-tract averages of daily mean temperature, daily mean relative humidity, and annual total precipitation in 2010. We obtained annual total precipitation from the ERA5 dataset, which has 30 km spatial resolution and hourly temporal resolution (Copernicus Climate Change Service (C3S), 2017). We first generated annual total precipitation for each ERA5 grid cell by summing all hourly values in 2010. We then calculated census-tract precipitation as the area-weighted mean of the values from the ERA5 grid cells intersecting a census tract. We used the same procedure to calculate census-tract averages of daily mean temperature and daily mean relative humidity from the ERA5Land dataset (Copernicus Climate Change Service (C3S), 2019), which had a 9 km spatial resolution. It is worth noting that both ERA5 and ERA5Land are at coarse scales when compared with the census tracts.

After examining multicollinearity using VIF, we excluded several covariates of land cover composition, including % urban, % water, and % others from the model. Our final covariates and their summary statistics are in Table 1. The final covariates have a Spearman’s correlation between −0.64 and 0.37 with the exposure variable, greenness (Fig. S1).

Furthermore, the covariates show low levels of multicollinearity, as their VIFs were all below 5 (Table 2), which is a commonly used empirical threshold for detecting multicollinearity (O’brien, 2007). In addition, the final covariates are confounders to the dengue-greenness relationship, which reduce bias in our estimation even when high multicollinearity between the exposure and these covariates exists. Multicollinearity does lead to wider confidence intervals, but confounding represents a greater concern in studies like ours (Schisterman, Perkins, Mumford, Ahrens, & Mitchell, 2017).

2.5. Regression model

We constructed four negative binomial mixed-effects models to estimate the association between dengue incidence and greenness. The mixed-effects model contains fixed effects that are the same across all census tracts and random effects that vary by neighborhood. The mixed-effects model makes use of our nested data structure (i.e., census tracts in neighborhoods, Fig. 1a), reduces omitted variable bias, and allows regression coefficients to vary by neighborhoods.

The first model included vegetation greenness with no covariates other than a neighborhood random intercept, β0j (equation (1)). Random intercept allows neighborhoods to have different baseline-level dengue incidence caused by unmeasured neighborhood characteristics such as vegetation species, management practices, and biophysical and built environments.

In the remaining models, we included additional covariates for socioeconomic vulnerability, population density, building height, percentage of census tract area covered by building footprints, land cover composition (% forest, % non-forest natural land, % farmland), elevation, and weather patterns (equation (2)). We further added a neighborhood random slope, β1j, for vegetation greenness to allow its coefficient vary by neighborhood (equation (3)). In this model, we allowed the random intercept, β0j, and random slope, β1j, to be correlated. Finally, we included an interaction term between vegetation greenness and socioeconomic vulnerability to test whether the association between dengue incidence and greenness was modified by socioeconomic vulnerability (equation (4)).

\[
\log(\text{IR}_0) = \beta_0 + \beta_1 \text{greenness}_j + \beta_0 + \epsilon_j
\]

\[
\log(\text{IR}_0) = \beta_0 + \beta_1 \text{greenness}_j + \beta_2 \text{vulnerability}_j + \beta_3 \text{PD}_j + \beta_4 \text{buildingheight}_j + \beta_5 \text{buildingfootprint}_j + \text{landcover}_j \delta + \beta_6 \text{elevation}_j + \text{weather}_j + \beta_0 + \epsilon_j
\]

\[
\log(\text{IR}_0) = \beta_0 + \beta_1 \text{greenness}_j + \beta_2 \text{vulnerability}_j + \beta_3 \text{PD}_j + \beta_4 \text{buildingheight}_j + \beta_5 \text{buildingfootprint}_j + \text{landcover}_j \delta + \beta_6 \text{elevation}_j + \text{weather}_j + \beta_0 + \epsilon_j
\]

\[
\log(\text{IR}_0) = \beta_0 + \beta_1 \text{greenness}_j + \beta_2 \text{vulnerability}_j + \beta_3 \text{agreeness}_j \times \text{vulnerability}_j + \beta_4 \text{PD}_j + \beta_5 \text{buildingheight}_j + \beta_6 \text{buildingfootprint}_j + \text{landcover}_j \delta + \beta_7 \text{elevation}_j + \text{weather}_j + \beta_0 + \epsilon_j
\]

where: IR0 is the incidence rate of dengue of census tract i in neighborhood j; greennessj is census-tract average of annual mean of monthly maximum NDVI; vulnerabilityj is Health Vulnerability Index measuring socioeconomic vulnerability; PDj is population density; buildingheightj is average building height; buildingfootprintj is percentage of census tract area covered by buildings; landcoverj is a vector for land cover composition including the percentage of census area that is either forest, non-forest natural land, or farming; elevationj is average land elevation in a census tract; weatherj is a vector of census-tract average of daily mean temperature, daily mean relative humidity, and annual total precipitation; β0j is the neighborhood random intercept; β1j is the neighborhood random slope for greennessj; εj is a random error.

We used incidence rate ratios (IRRs) to interpret the model results. IRR is a factor by which the outcome, dengue cases per 100,000 population, changes with a one-unit change in a covariate, when holding all other covariates constant. Therefore, IRRs greater than 1 indicate a positive association, whereas IRRs less than 1 indicate a negative association.

In addition, to account for some of the spatial autocorrelations between the census tracts and heteroscedasticity of model residuals, we estimated neighborhood-cluster-robust standard errors for the model coefficients. The neighborhood-cluster-robust standard error assumes that census tracts within the same neighborhood (cluster) are correlated, but those between different neighborhoods are independent. This assumption is likely to result in larger standard errors, wider confidence intervals, and more statistically insignificant (conservative) inference when compared with assuming the census tracts are independent of each other. In this way, we partially controlled for spatial autocorrelation in our dataset.

2.6. Sensitivity tests

In addition to the main models (Eqs. (2)-(4)) using a mixed-effects negative binomial model with cluster-robust standard errors, we tested several alternatives to justify our choice. We started with a pooled model without neighborhood random intercepts, testing whether

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greenness (-10–10)</td>
<td>2.88</td>
</tr>
<tr>
<td>Vulnerability (HVI, 0–10)</td>
<td>1.64</td>
</tr>
<tr>
<td>Population density (people/hectare)</td>
<td>1.65</td>
</tr>
<tr>
<td>Average building height (m)</td>
<td>1.61</td>
</tr>
<tr>
<td>% building footprint</td>
<td>2.41</td>
</tr>
<tr>
<td>% forest</td>
<td>1.51</td>
</tr>
<tr>
<td>% non-forest natural land</td>
<td>1.07</td>
</tr>
<tr>
<td>% farmland</td>
<td>1.71</td>
</tr>
<tr>
<td>Average elevation (m)</td>
<td>3.33</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>2.60</td>
</tr>
<tr>
<td>Total precipitation (mm/year)</td>
<td>2.48</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>3.74</td>
</tr>
</tbody>
</table>
omitting such effects could bias the dengue-greenness relationship. We then estimated the models assuming the census tracts were independent of each other and model residuals were independently and identically distributed to illustrate how cluster-robust standard errors correct the statistical inference.

Finally, instead of a neighborhood random intercept model (equation (2)), we estimated a neighborhood fixed effects model. Both random intercept and fixed effects models are ways to control unobserved, group-level (neighborhood in this study) effects in hierarchical datasets by estimating a set of intercepts for the groups. The two models require different assumptions. A random intercept model assumes that these intercepts are uncorrelated with the covariates, whereas the fixed effects model relaxes such an assumption (Greene, 2018). While more versatile, we did not prioritize the fixed effects model as we would have a large number of fixed effects (n = 474, one per neighborhood), which leads to large confidence intervals for the coefficient estimates.

### 2.7. Limitations of the model

The first limitation of the study is that we used secondary data where cases were reported by health professionals, health services, and the public. Therefore, this case reporting was subject to misclassification bias. Dengue cases were reported to the Notifiable Diseases Information System, where cases could be misdiagnosed (de Mattos Almeida et al., 2007; Harris et al., 2000). In Belo Horizonte between 1996 and 2017, the surveillance system could confirm 46% of reported cases through laboratory diagnosis, whereas the remaining cases were diagnosed based on epidemiological criteria (Campos et al., 2019). Recent studies have shown improved accuracy in dengue case reporting in the Notifiable Diseases Information System (Barbosa, Barrado, Zara, & Siqueira, 2015; Goto et al., 2016)

Secondly, although we controlled for a comprehensive set of covariates, additional factors may further reduce omitted variable bias. These factors include vegetation species, land use, and irrigation practices, which affect both dengue incidence and greenness. However, quality datasets for these covariates were lacking when we conducted this study.

Thirdly, the environmental datasets in this study are at coarse spatial resolutions, which likely contain measurement errors and fail to capture fine-scale variations. For example, the 30 m resolution MapBiomass Brasil land cover product can only identify the main land cover, and classification accuracy is reduced in heterogeneous landscapes like wetlands due to the 'mixed pixel' problem (Raur et al., 2019). This imprecise estimation of vegetation quantity prevents us from fully isolating the effect of vegetation quantity on greenness. The coarse spatial resolution also means a lack of within-neighborhood variation of the environmental variables to produce precise regression estimates. We consider that this issue exists in NDVI (30 m resolution) and land cover (30 m resolution), and it magnifies in weather pattern variables (9–31 km resolutions). Therefore, collecting fine-resolution environmental

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Results of models examining the association between dengue incidence and greenness, controlling for socioeconomic and environmental covariates.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fix effects</td>
<td>(1) IRR</td>
</tr>
<tr>
<td>Greenness (-10–10)</td>
<td>0.834***</td>
</tr>
<tr>
<td>[0.789,0.882]</td>
<td>[0.847,1.025]</td>
</tr>
<tr>
<td>Vulnerability (HVI, 0–10)</td>
<td>1.254***</td>
</tr>
<tr>
<td>[1.162,1.354]</td>
<td>[1.176,1.368]</td>
</tr>
<tr>
<td>Greenness × Vulnerability</td>
<td>0.879***</td>
</tr>
<tr>
<td>[0.825,0.938]</td>
<td>[0.849,0.998]</td>
</tr>
<tr>
<td>Population density (people/hectare)</td>
<td>0.999***</td>
</tr>
<tr>
<td>[0.9997,0.9998]</td>
<td>[0.9997,0.9998]</td>
</tr>
<tr>
<td>Average building height (m)</td>
<td>0.982</td>
</tr>
<tr>
<td>[0.960,1.005]</td>
<td>[0.956,1.005]</td>
</tr>
<tr>
<td>% building footprint</td>
<td>1.017***</td>
</tr>
<tr>
<td>[1.013,1.022]</td>
<td>[1.013,1.022]</td>
</tr>
<tr>
<td>% forest</td>
<td>1.007*</td>
</tr>
<tr>
<td>[1.000,1.015]</td>
<td>[0.999,1.016]</td>
</tr>
<tr>
<td>% non-forest natural land</td>
<td>1.017</td>
</tr>
<tr>
<td>[0.980,1.055]</td>
<td>[0.983,1.058]</td>
</tr>
<tr>
<td>% farmland</td>
<td>0.982***</td>
</tr>
<tr>
<td>[0.980,0.995]</td>
<td>[0.981,0.995]</td>
</tr>
<tr>
<td>Average elevation (m)</td>
<td>0.993***</td>
</tr>
<tr>
<td>[0.992,0.995]</td>
<td>[0.992,0.995]</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>0.707***</td>
</tr>
<tr>
<td>[0.605,0.827]</td>
<td>[0.598,0.817]</td>
</tr>
<tr>
<td>Total precipitation (mm/year)</td>
<td>1.012</td>
</tr>
<tr>
<td>[0.957,1.027]</td>
<td>[0.995,1.025]</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>0.775*</td>
</tr>
<tr>
<td>[0.593,1.013]</td>
<td>[0.601,1.023]</td>
</tr>
</tbody>
</table>

Note: A coefficient, or incidence rate ratio (IRR), is the factor by which the dengue incidence (dengue cases per 100,000 residents) changes for a one-unit increase in the corresponding covariate when holding other covariates constant. Vulnerability is measured by the Health Vulnerability Index (HVI) that encompasses infrastructure indicators related to basic sanitation and socioeconomic status of residents, recorded by the 2010 census. *, **, and *** indicate significant at p-value < 0.10, p-value < 0.05, and p-value < 0.01. 95% confidence intervals are in square brackets.

+Random effects are the variances of log-transformed coefficients (IRRs).
datasets is highly beneficial for environmental health studies to capture detailed variations in heterogeneous urban landscapes, reducing measurement error, and producing statistically sound results.

3. Results

3.1. Insignificant protective effect of greenness on dengue

We found a negative and statistically significant association (IRR < 1) between greenness and dengue incidence, when controlling only for neighborhood random effects (model 1, Table 3). In this model, a one-unit increase in greenness (40% of the mean of 2.52) was associated with a 16.6% (10.834 – 1) × 100% = −16.6%, similar calculations hereafter) reduction in dengue incidence rate.

The negative association between greenness and dengue incidence attenuated and was not statistically significant after controlling for a set of covariates including socioeconomic vulnerability, population density, building height, percentage of census tract area covered by building footprints, elevation, land cover composition, and weather patterns (model 2, Table 3). In the latter model, a one-unit increase in greenness was associated with a 5.1% reduction in dengue incidence.

3.2. Socioeconomic vulnerability modifies the dengue-greenness association

We observed a statistically significant random slope for greenness, indicating the association between greenness on dengue incidence varied by neighborhood (model 3, Table 3). In addition, this model estimated a stronger and statistically significant protective effect of greenness, where a one-unit increase in greenness reduced dengue incidence by 8.5%.

We observed a statistically significant negative interaction between greenness and socioeconomic vulnerability (model 4, Table 3), suggesting that the association between dengue incidence and greenness was modified by socioeconomic vulnerability. The association between dengue incidence and greenness shifted from positive (IRR > 1) to negative (IRR < 1) as the census tracts in our sample became more vulnerable (Fig. 2). This association was stronger and statistically significant in the census tracts with low (roughly below the 10th percentile) or high (roughly above the 70th percentile) HVI.

3.3. The associations between dengue incidence and the covariates

We found that socioeconomic vulnerability, % building footprint and % forest had positive and statistically significant associations with dengue incidence across the models (models 2–4, Table 3). In addition, % non-forest natural land and total precipitation had positive but statistically insignificant associations with dengue incidence.

We found negative and statistically significant associations between dengue incidence and population density, % farmland, average elevation, relative humidity, and temperature (models 2–4, Table 3). We found negative but statistically insignificant association between dengue incidence and average building height.

3.4. Sensitivity tests

The alternative model specifications (Table S1) generally confirmed the dengue-greenness relationship in our main models (Table 3). However, the sizes of the coefficients and their confidence intervals varied due to different model assumptions. Compared with the main models, the pooled model ignoring neighborhood effects (model 2–1, Table S1) produced a stronger negative and statistically significant association (IRR = 0.812) between dengue incidence and greenness. However, this association was likely biased due to omitted variables. Specifying cluster-robust standard errors led to wider confidence intervals and more conservative statistical inference, when compared with using independent standard errors (models 2–2, 3–1, 4–1, Table S1). Finally, the neighborhood fixed effects model yielded a slightly weaker association between dengue and greenness (model 2–3, Table S1).

4. Discussion

4.1. Protective effect of vegetation greenness on dengue incidence

Our results suggested a negative but statistically not significant association between dengue incidence and vegetation greenness averaged across the census tracts in Belo Horizonte, after controlling for socioeconomic vulnerability, population density, building height, percentage of census tract area covered by building footprints, elevation, land cover composition, weather patterns, and neighborhood random intercept (model 2, Table 3). Furthermore, this association varied by neighborhood (model 3, Table 3). For example, it changed based on the socioeconomic vulnerability of the census tracts: the association was negative and statistically significant for the most vulnerable census tracts, but it was positive and statistically significant for the least vulnerable ones (Fig. 2).

Greenness, as measured by NDVI, alone represents the combined effect of vegetation quantity and quality. We attempted to separate the effect of vegetation quantity from quality by introducing covariates for vegetation coverage, including land cover composition and the percentages of census tract area covered by buildings. Despite some data limitations discussed in Section 2.7, higher greenness likely indicated better vegetation quality in our analyses.

Our results align in part with previous findings indicating a negative

![Fig. 2. Associations (IRR) between dengue incidence and vegetation greenness, conditioned on socioeconomic vulnerability. IRR is the factor by which dengue incidence (annual dengue cases per 100,000 residents) changes for a one-unit increase in a covariate, when holding other covariates constant. Socioeconomic vulnerability is measured by the Health Vulnerability Index. 95% confidence intervals of the IRRs are shaded.](image-url)
association between dengue incidence and vegetation regarding quantity (Araujo et al., 2015; Cao et al., 2017) and the combined effect of quantity and quality (Meza-Ballesta & Gónima, 2014). However, other studies have also reported positive (Martínez-Bello et al., 2017) or non-linear associations (Qi et al., 2015). We suspect that the inconsistencies reflect differences in spatial scales of analysis, measurements of vegetation, and analytical approaches. For instance, some studies only analyzed simple correlations between dengue incidence and vegetation without controlling for a comprehensive set of covariates as in our study (Cao et al., 2017; Martínez-Bello et al., 2017). It is also worth noting that the overall negative association between dengue incidence and greenness identified in our study was small and not statistically significant. One explanation could be that the main factor leading to high Aedes aegypti population (the main regional dengue vector) and consequently high dengue incidence include household water storage, sewage connection, and trash collection, rather than vegetation itself. It may also be the case that the log-linear relationship assumed by negative binomial models is an oversimplification, as threshold effect (“tipping points”) may exist in the actual dengue-greenness relationship. To account for the threshold effect, researchers have employed higher-order terms of independent variables (Spencer, 2013), semi-parametric models (Ding, Cao, Yu, & Ju, 2021), and machine learning models (Zhao, Yan, Yu, & Van Hentenryck, 2020).

To our knowledge, no previous studies have examined whether the association between dengue incidence and greenness is modified by socioeconomic vulnerability. Our results (model 4, Table 3) show that while census tracts with lower socioeconomic vulnerability had a positive association between dengue incidence and greenness, the association was negative in census tracts with higher socioeconomic vulnerability (Fig. 2). This pattern may indicate that other vegetation characteristics, such as vegetation species, type (natural versus cultivated, tree versus grass), structure, and management practices differ between census tracts with different levels of socioeconomic vulnerability (Mitchell et al., 2016). For example, in Belo Horizonte, many vulnerable census tracts are close to the city border (Fig. 1d), which is abundant in natural vegetation not suitable for Aedes aegypti (Dorvillé, 2010; Medeiros-Sousa et al., 2017). The same one-unit increase in the greenness of different vegetation could lead to variable changes in dengue incidence. This also may explain why prior studies have found mixed relationships between dengue incidence and vegetation, as vegetation can be systematically different between their study areas.

4.2. Benefits of improved built environment and housing

Our models also estimated the associations between dengue incidence and additional covariates of socioeconomic vulnerability, urban and natural environment. The disease process is complex, and particular microenvironments can become “pathogenic landscape” (Martí et al., 2020). While the socioeconomic vulnerability, urban, and natural environment covariates studied here are markers of a “pathogenic landscape” and may lead to interventions, any causality should be carefully interpreted from our results due to omitted variable bias and study design issues (Spencer et al., 2020). We focus our discussion here on socioeconomic vulnerability and urban environment covariates, as they could be improved by social policy and urban planning.

The positive association between dengue incidence and socioeconomic vulnerability (HVI) found in this study (models 2–3, Table 3) is consistent with previous ones in this urban setting (Campos et al., 2019; de Mattos Almeida et al., 2007). We consider this positive association highly likely, as supported by the construct of HVI and vector biology. High HVI indicates inadequate water supply infrastructure, sewage connection, and garbage collection, creating favorable conditions for Aedes aegypti infestation (Gubler, 2011).

Census tract covered by high fractions of building footprint (i.e., building density) and forest showed positive associations with dengue incidence. High building density may be associated with extensive drainage networks (Seidahmed et al., 2018), and forest provides shading (Barrera et al., 2006) – both are favorable conditions for Aedes aegypti. Population density and % farmland showed a negative association with dengue incidence. While increased population density means higher chances of transmission, it may also indicate a better supply of water infrastructures and possibilities of herd immunity. The association between % farmland and dengue incidence is likely dependent on agriculture practices and requires further investigation in our study area.

4.3. Urban planning and policy implications

Local governments should consider vegetation management strategies to improve vegetation quality, particularly in vulnerable areas with greater potential benefits. Preserving green space has been included in Belo Horizonte’s development agendas, such as the Organic Law of the Municipality, the municipal conferences on urban policies in 2014, and the city’s master plan in 2019 (Costa, Álvares, Maciel, Teixeira, Coimbra, & de Simiao, 2009; Cruz, 2020). When preserving green space, we suggest local governments restore typical, native forest understory and arboreal landscapes. When well-managed, these landscapes could largely decrease the presence of vectors for dengue and other infectious diseases (Pedrosa et al., 2020). Most problems with green space in Brazilian cities, as happens in Belo Horizonte, are related to poor management, which leads to habitat degradation and trash accumulation (Cardoso, Vasconcellos Sobrinho, & de Vasconcellos, 2015; Lobo, 2020). These poorly managed areas are therefore more likely to become potential breeding sites for disease vectors. Managing vegetation not only reduces dengue risk but also brings co-benefits such as reducing crime, mitigating urban heat, simulating physical activities, and improving mental health (Cardoso et al., 2015; Casey et al., 2017; James, Banay, Hart, & Laden, 2015). Managing tree canopy to reduce shading, mainly when water bodies are present, may also be an option, although other benefits of trees should be considered.

Green space management should come together with improvements in water supply and drainage infrastructures, and sanitation services to reduce the likelihood of having a “pathogenic landscape”. Furthermore, the city may consider adopting an optimized drainage network to avoid an extensive length of pipes, which further reduces Aedes aegypti habitat. One priority area for these interventions is informal settlements in Belo Horizonte. The city began to improve the living conditions of informal settlements in 1970, and its recent efforts have been led by the Vila Viva Project since 2005 (Friche, Dias, Reis, Dias, & Caiâfa, 2015). This project creates the opportunity for future studies to quantify how bundled interventions, including improvements in sanitation and housing conditions, and provision of green space, together change dengue risk.

5. Conclusions

We found an overall negative but statistically not significant association between dengue and greenness across 3826 census tracts in Belo Horizonte, Brazil, after controlling for a comprehensive set of covariates. We also found that the dengue-greenness association was modified by socioeconomic vulnerability: the association was positive in the least vulnerable census tracts but negative in the most vulnerable areas. Since we controlled for vegetation quantity using land cover composition and percentage of census tract area covered by building footprints, high greenness here likely indicates higher vegetation quality in terms of better vegetation health, more vigorous photosynthetic activities, and less environmental degradation.

Collectively, our results show the potential role of improved vegetation quality, as manifested by higher greenness, in reducing dengue risk, particularly for vulnerable areas. When well-managed, vegetated areas are less likely to become breeding sites for dengue vectors and reduce the risk of local dengue transmission. While we showed the potential role of improving vegetation quality in reducing dengue risk, this
strategy should be incorporated with other multiprofessional approaches, such as improvements of water supply and drainage infrastructures, to reduce the presence of “pathogenic landscape” and dengue risk.

CRediT authorship contribution statement

Maria da Consolação Magalhaes Cunha: Conceptualization, Investigation, Data curation, Writing - original draft, Writing - review & editing. Yang Ju: Conceptualization, Investigation, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. Maria Helena Franco Morais: Conceptualization, Investigation, Data curation, Writing - review & editing. Iryna Dronova: Writing - review & editing. Servio Pontes Ribeiro: Conceptualization, Writing - review & editing. Fabio Raphaël Pascoleti Bruhn: Writing - review & editing. Larissa Lopes Lima: Data curation, Writing - review & editing. Denise Marques Sales: Data curation, Writing - review & editing. Olivia Lang Schultes: Writing - review & editing. Daniel A. Rodriguez: Supervision, Conceptualization. Waleusa Teixeira Caiafia: Supervision, Writing - review & editing.

Acknowledgments


Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2021.104255.

References


