ANALYZING LONG-TERM AVAILABILITY OF URBAN GREEN SPACE BY SOCIOECONOMIC STATUS IN MEDELLIN, COLOMBIA, USING OPEN DATA AND TOOLS

Jorge E. Patiño [1]

[1] RiSE-Group, Dept. of Mathematical Sciences, School of Sciences, EAFIT University, Medellin, Colombia - jpatinoq@eafit.edu.co

KEY WORDS: Urban Green Space, Socioeconomic Status, Long-term Availability, Environmental Justice, NDVI, Landsat

ABSTRACT:

The availability of green spaces is an important issue for urban populations worldwide, given the benefits that the green spaces provide for health, well-being, and quality of life. But urban green spaces are not always distributed equally for different population groups within cities. Latin America is the second most urbanized region of the world, but there are few published studies analysing the green space availability for different urban population groups, and less so analysing the long-term trends. This work presents an analysis of long-term availability of urban green spaces by different socioeconomic status population groups in Medellin city, Colombia, using open geospatial data and open software tools. The results indicate that disparities between different groups have been decreasing in the last years, but there are still efforts to do. Showing this kind of analysis based on open data and tools is essential as it opens the possibility for replicating it in other cities with scarce budgets.

1. INTRODUCTION

The characteristics of urban form influence the daily life of citizens and it is recognized as an important factor in both quality of life and environmental impact (Panagopoulos et al., 2016). According to Tzoulas et al. (2007), poverty and social factors are the main determinants of human health in urban areas. Environmental features are recognized too, with many academic studies providing evidence of a positive correlation between well-being, health, and green space (Rigolon, 2018; Tzoulas et al., 2007). The mechanisms that explain the relationship between urban green space availability, well-being, and health are related to some environmental services of urban green areas such as decreasing air pollution, noise, and the urban heat-island effect (Meerow and Newell, 2017) as well as providing pleasant and relaxing views (Nordbø et al., 2018; Tomita et al., 2017; Wood et al., 2017). These mechanisms also include additional opportunities to perform physical activities and improve social engagement, as well as reduced psychological stress and depression (James et al., 2016). As the urban population grows, the availability of urban green spaces gets reduced. Ecosystem quality tends to decrease as urban density increases (Panagopoulos et al., 2016) and city planning has to address trade-offs between city development and the preservation of green or natural spaces within the urban fabric (Bertram and Rehdanz, 2015).

The availability of green spaces throughout the city is becoming an important topic for urban populations worldwide. Urban green spaces are often not distributed equitably across different population groups, which is recognized as an issue of environmental justice that requires attention (Wolch et al., 2014). The US Environmental Protection Agency (EPA) defined environmental justice as the "fair treatment and meaningful involvement of all people regardless of race, colour, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations and policies" (Nelson and Grubesic, 2018). As for urban green space, environmental justice means equal access to green spaces for all population groups in the city. This concept fits directly into the Sustainable Development Goal 11: "*Make cities and human settlements inclusive, safe, resilient and sustainable,*" specifically to the 11.7 target: "*By 2030, provide universal access to safe, inclusive and accessible, green and public spaces, in particular for women and children, older persons and persons with disabilities.*"[1]

Latin America is the second most urbanized region of the planet, with 81% of its population living in urban areas (United Nations, 2018). But Latin American cities are still segregated both at the social and spatial level (UN-Habitat, 2012), which is expressed in the high urban income inequality, the persistence of informal settlements, and uneven access to green and public spaces, among other issues (UN-Habitat, 2012). Since 1985 Latin America has gone through a process of fast urbanization that has compromised the availability of green spaces for urban populations (Wolch et al., 2014; Nor et al., 2017; Dobbs et al., 2018). The study and defence of environmental justice are particularly important for urban populations in Latin America, as most of the literature on the topic have addressed cities from the Global North and there are very few published works from Latin American cities in the academic literature (Rigolon et al., 2018). To the best of the author's knowledge, there are no published works analysing long-term availability of urban green spaces for different population groups by socioeconomic characteristics in Latin American cities. Therefore, the importance of contributing with empirical evidence from Latin American cities to better inform local authorities and urban planners about long-term disparities in the urban green space provision and insights about how to achieve more inclusive and healthier cities in the future.

This study quantifies the availability of urban green space for different socioeconomic status (SES) groups in the period between 1984 and 2018 in Medellin city using open geospatial data and tools. The use of open data and tools to analyse this issue is of particular importance for many Latin American cities with scarce resources, as this approach could be implemented at a relatively low cost.

---

[1] https://sustainabledevelopment.un.org/sdg11

2. STUDY AREA AND DATA SETS

2.1 Medellin city

Medellin city is the second largest city in Colombia (Figure 1). It is located in the Andean Mountain Range at 1500 meters above sea level in the north-western region of the country and it has a population of 2.4 million according to the 2018 National Census (DANE[2]). This city extends over a narrow valley crossed by Medellin River from South to North, with the most western and eastern neighbourhoods of the city built over slopes steeper than 20%. Rapid urban growth in the last decades has led to a high degree of spatial heterogeneity in both the socioeconomic and physical characteristics of its neighbourhoods (Duque et al., 2013): the more affluent population is located in the West and South, while the less affluent are in the North and towards the urban-rural fringe (Patino et al., 2014, Garcia Ferrari et al., 2018).
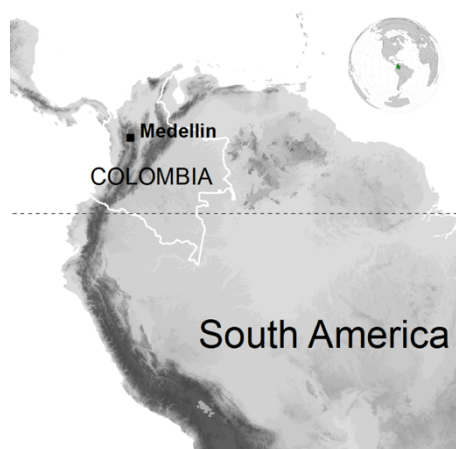

Figure 1. Medellin location map.

2.2 Administrative neighbourhoods' boundaries

Urban administrative neighbourhoods were used to select the areas of the city that were already part of the urban fabric in the first year of the analysis. This dataset was obtained from the Open Data website of the city (https://geomedellin-m-medellin.opendata.arcgis.com/).

2.3 Socioeconomic status

In Colombia, housing areas are categorized by a socioeconomic classification into "strata" for taxation of public services (water, sewer, electricity, and gas). This is done to differentially charge the home public services based on the socioeconomic stratum and to allocate subsidies and charge for contributions. The rationale behind the system is that high-income people groups pay more for the public services in order to contribute to subsidize low-income groups to have the same services at an affordable cost (DANE, 2019). This system classifies residential households into 6 classes, being 6 the wealthier and 1 the poorest, and it is a very good proxy for the socioeconomic status of the urban population (Garcia Ferrari et al., 2018). The socioeconomic stratification map at block level was obtained from the city planning office. For the purpose of this analysis, the 6 socioeconomic strata were reclassified into three classes of

socioeconomic status (SES) as showed in Table 1 (and Figure 2).

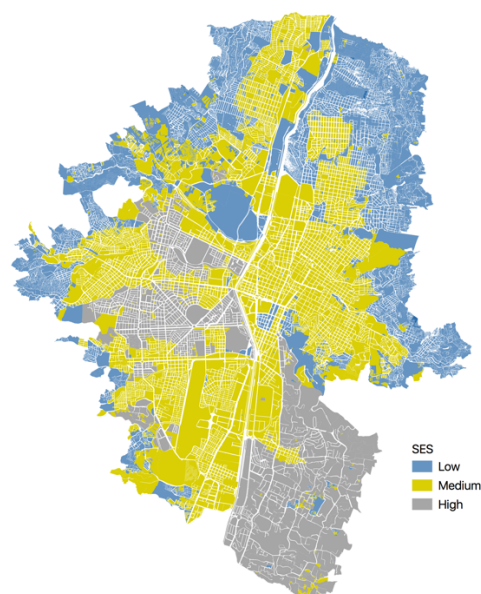| SE strata	SES class
1 - 2	Low
3 - 4	Medium
5 - 6	High

Table 1. Reclassification of strata to SES class.

Figure 2. Urban area of Medellin by SES.

2.4 Residential areas

Even though the stratification system is intended to classify only residential areas, the map of socioeconomic strata shows that every urban cadastral parcel in the city was allocated into some socioeconomic stratum. The Medellin's land use map obtained was also from the Open data portal of the city and was used to identify and extract only residential blocks from the SES map for this analysis (Figure 3).

2.5 Urban green space availability, Landsat imagery

We used Landsat imagery and the normalized difference vegetation index (NDVI) to estimate the availability of urban green spaces at different locations. According to Rhew et al. (2011), the NDVI is a useful measure of greenness, which show strong correlation with expert ratings. It can be computed easily from open multispectral satellite imagery, and it is widely used in epidemiologic and urban research (Gupta et al., 2012; Jorgensen and Gobster, 2010; Nordbø et al., 2018). The Google Earth Engine platform (Gorelick et al., 2017) was used to search and download Landsat surface reflectance images from different years from 1984 to 2018 (one every 5 or 6 years) from the Landsat 4, 5, 7 and 8 collections (Table 2). The surface reflectance images were preferred because they are already atmospherically corrected. The main criteria for the selection of the images was the cloud cover over the city's urban area and that their anniversary date is as close as possible or at least from the same season of the year. However small clouds and shadows were present in some of the images as it was not possible to find

100% cloud-free images for all the required years. Clouds and shadows were masked using the pixel quality bands. The masked areas and the gaps from the Landsat 7 SLC-Off images were filled using the pixel values at the same location of the previous or following images in the same image series to create cloud-free composites for each selected date. Figure 4 show infrared colour composites of the Landsat images used in the analysis.
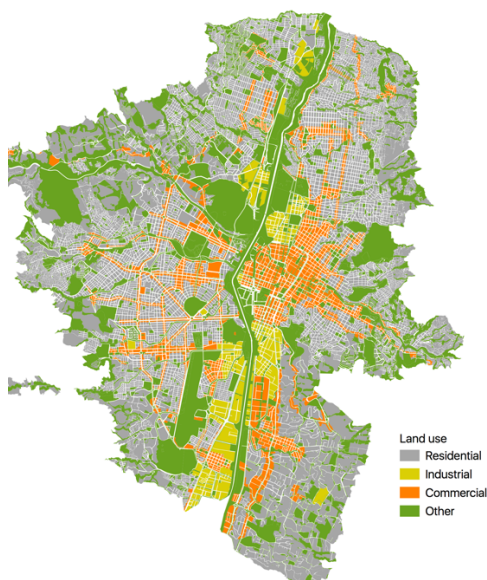

Figure 3. Medellin land use map.

Date	Landsat collection
1984-09-02	LANDSAT/LT05/C01/T1_SR
1989-08-07	LANDSAT/LT04/C01/T1_SR
1996-08-02	LANDSAT/LT05/C01/T1_SR
2000-07-12	LANDSAT/LT05/C01/T1_SR
2005-07-18	LANDSAT/LE07/C01/T1_SR
2011-07-19	LANDSAT/LE07/C01/T1_SR
2015-09-16	LANDSAT/LE07/C01/T1_SR
2018-06-12	LANDSAT/LC08/C01/T1_SR

Table 2. Landsat image composites and used collections
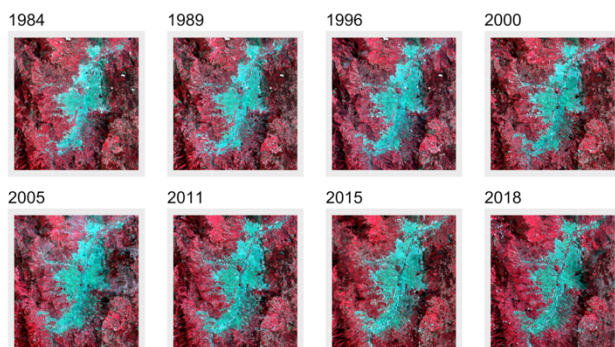(Landsat imagery courtesy of US Geological Survey).

Figure 4. Infrared colour composites of the Landsat images.

3. METHODS

This analysis was implemented using the following open tools: R (R Core Team, 2013), RStudio (RStudio Team, 2015) and QGIS (QGIS Development Team, 2018). The general workflow is composed of three main parts: sampling, image processing and feature extraction, and descriptive data analysis. The availability of urban green space was computed for each year at three different buffer sizes from the centroid of residential blocks: 100, 300, and 500 meters.

3.1 Sampling

The aim of the sampling was to obtain a balanced sample of residential areas that were already built in 1984. The administrative neighbourhood boundaries were overlaid over the 1984 Landsat image to identify those neighbourhoods that were mostly built-up in that year. The urban neighbourhoods in 1984 were manually selected and merged in QGIS to obtain the reference area for the selection of urban blocks to be analysed.

The SES map spatial unit is the block (n = 25,942). We first obtained the centroids of the urban blocks, and then performed a spatial join with the land use map to assign the land use category to each block centroid. We kept only the residential block centroids (full sample, n = 14,719). Then we used the 1984 reference polygon to narrow that selection to the urban areas in that year and obtained 10,750 residential centroids. We then counted the number of features by SES class (Table 3) and used the minimum to balance the sample: 784 residential blocks in each SES class (Figure 5). The feature selection and geometric operations were done in R using the "sf" package (Pebesma, 2018).

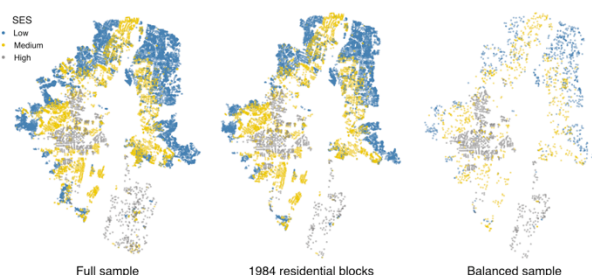| SES class	Number of blocks in 1984
Low	6,354
Medium	3,577
High	784

Figure 4. Block centroids count by SES in 1984.

Figure 5. Sample of urban block centroids.

3.2 Image processing

The image processing workflow was fully implemented in R. Three different process were applied to the Landsat images: relative radiometric normalization, NDVI calculation, and mean NDVI extraction for different buffer sizes. The relative radiometric normalization was done using the "RStoolbox" package (Leutner, et al., 2018), using the automatic pseudo-invariant feature match technique with the "pifMatch" function and the Euclidean distance method to compute similarity between image pixels. Then the normalized difference

vegetation index was calculated for each year using the "raster" package (Hijmans, 2019) with the red and infrared bands of the Landsat images, following the classic formula (Equation 1). Figure 6 shows the NDVI images.

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \qquad (1)$$

The extraction of mean values was done using the "exactextracr" package (Baston, 2019). We first calculated the three different buffers from the centroids with the "st_buffer()" function (from the "sf" package), and then extracted the mean values within the buffers with the "exact_extract()" function. This function has two advantages to compute zonal statistics for vector geometries: it is faster than most of the available implementations in R, and it allows to assign each extracted value directly to the original centroid it belongs, without writing it to a different output dataset.
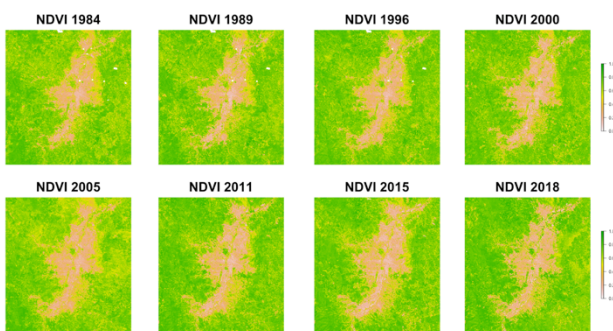

Figure 6. NDVI images.

We used the block centroid attribute table with the SES class and NDVI values by year and by buffer size as input to the descriptive analysis.

3.3 Descriptive data analysis

The first exploration of the data was the preparation of boxplots of mean NDVI values by SES for each buffer size and year. The boxplots help to identify the differences on the values distributions by groups. Then we calculated the average of all extracted values by buffer size by SES class by year and plotted them as a series to analyse the evolution of urban green space availability from 1984 to 2018. The data plots were created in R using the "ggplot2" package (Wickman, 2016).

4. RESULTS AND DISCUSSION

The boxplots and the temporal plots as well show interesting trends in the green space availability by SES class and by different buffer size. Figure 7 shows the boxplots, and the temporal trend is better illustrated in Figure 8. The boxplots show approximately the same trend of differences by SES each year, and the difference between SES groups of green space availability gets reduced as the buffer size increase. Between 1984 and 1996 one can observe almost the same trend: the medium SES class had the lower availability of urban green spaces in the closer neighbourhood (100 m), and the difference with the other two SES classes decreases as the buffer size increases (300 and 500 m). In 2000 that trend starts to show changes, with the high SES class gaining higher availability of green space than the other two classes, and from 2005 to 2018 the availability reorders to show the higher green space

availability for the high SES class, and the lower green space availability for the low SES class, for all buffer sizes.
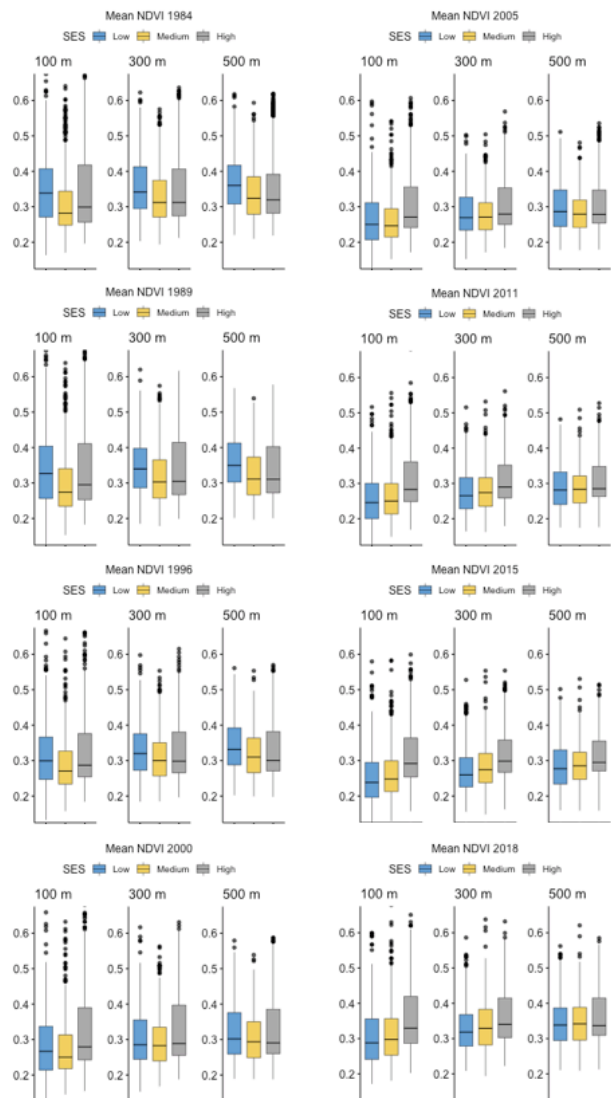

Figure 7. Boxplots showing the distribution of mean NDVI values by SES, buffer size, and year.
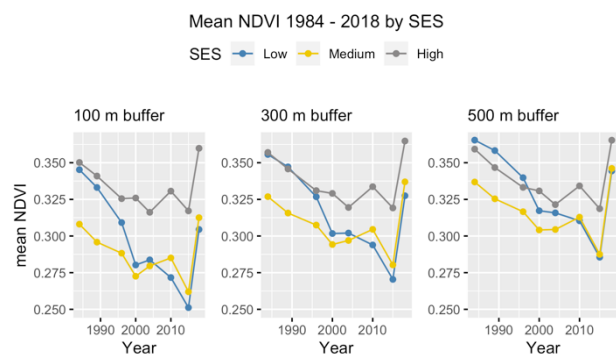

Figure 8. Temporal plots showing the evolution between 1984 and 2018 of the average value of mean NDVI by SES class at different buffer sizes.

The temporal plots confirm those trends. They show that in average, the high SES class has the higher green space availability at close distances (100 m) in the whole analysed period. It is also very interesting to note that for larger buffer sizes (300 and 500 m) the average green space availability was very similar for the low and high SES classes between 1984 and 1996, and the difference increase in the 2000 to 2018 period.

Another interesting fact to note is that the average values for low and medium SES classes are very close in the last five measures (2000 - 2018) for all buffer sizes, which indicates that the disparity between those two classes has been decreasing. But the difference between low and medium SES classes with regard to the high SES class remains in that period and it is higher at smaller distances. The difference in green space availability between the high SES class and the other two SES classes increased between 1996 and 2015, and it is decreasing again to 2018.

The latter is an important finding as it points out the environmental justice in Medellin with regard to the availability of urban green spaces remains, although it is decreasing in recent years. And that the closer urban environment for low and medium SES classes has way less green space availability than the high SES class, which is important because regarding the health benefits of urban green spaces, the closer the better (Ward Thompson et al., 2012; Wood et al., 2017; Coppel and Wüstemann, 2017).

The temporal plot for the 500 m buffer shows that the average green space availability at that distance was the highest for the low SES class in the period from 1984 to 1996, and then the trend got reversed. That could be the effect of the urban densification without good planning practices, as most of the low SES areas of the city have grown led by self-build processes in informal settlements, not leaving space for public squares and parks, or even for planting trees and shrubs in the streets.

5. CONCLUSSIONS

The results of this work are informative for urban planners, decision makers and local authorities interested in the interplay between public health, wellbeing, and the built environment in the city. They show that the disparities with regard to the availability of urban green spaces by different SES class increased in Medellin city from 1984 to 2010, and although it is slightly decreasing in the later years, more efforts are needed to reduce the disparities in green space availability (and its benefits) between different SES classes.

The analysis also shows that for the low SES class, the availability of green spaces has decreased more between 1984 and 2010 than for the other two SES classes, maybe as a result of urban densification in those areas. The data shows a reversing trend in the last years which means that the several planning tools, like the city master plan and several integrated urban projects, all put in place after 2000, are in effect helping to improve that situation.

This work has demonstrated the usefulness of open geospatial data, in particular the Landsat Archive, and open software tools for analysing long-term availability of urban green space in a Latin American city (Medellin, Colombia) and by different SES groups. Showing this kind of analysis based on open tools is essential as it opens the possibility for replicating it in other cities with scarce budgets. Given that this approach is not very

sophisticated, at least statistically speaking, it could even empower citizens and some interest groups to better understand what is happening in their cities to demand proper action from their local authorities.

ACKNOWLEDGEMENTS

REFERENCES

Baston, D. 2019. exactextractr: Fast Extraction from Raster Datasets using Polygons. R package version 0.1.1. https://CRAN.R-project.org/package=exactextractr

Bertram, C., Rehdanz, K. 2015. The role of urban green space for human well-being. *Ecological Economics*, *120*, 139–152. https://doi.org/10.1016/j.ecolecon.2015.10.013

Coppel, G., Wüstemann, H. 2017. The impact of urban green space on health in Berlin, Germany: Empirical findings and implications for urban planning. *Landscape and Urban Planning*, *167*(March 2016), 410–418. https://doi.org/10.1016/j.landurbplan.2017.06.015

Dobbs, C., Hernández-Moreno, Á., Reyes-Paecke, S., Miranda, M. D. 2018. Exploring temporal dynamics of urban ecosystem services in Latin America: The case of Bogota (Colombia) and Santiago (Chile). *Ecological Indicators*, *85*(November), 1068–1080. https://doi.org/10.1016/j.ecolind.2017.11.062

DANE, 2019. Socio-economic stratification system. https://www.dane.gov.co/index.php/69-espanol/geoestadistica/estratificacion/468-estratificacion-socioeconomica (15 October 2019).

Duque, J. C., Royuela, V., Noreña, M. 2013. A stepwise procedure to determine a suitable scale for the spatial delineation of urban slums. In E. Fernandez & F. Rubiera Morollón (Eds.), *Defining the spatial scale in modern regional analysis. Advances in Spatial Science* (pp. 237–254). Berlin Heidelberg: Springer-Verlag.

Garcia Ferrari, S., Smith, H., Coupe, F., Rivera, H. 2018. City profile: Medellin. *Cities*, *74*, 354–364. https://doi.org/10.1016/j.cities.2017.12.011

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gupta, K., Kumar, P., Pathan, S. K., Sharma, K. P. 2012. Urban Neighborhood Green Index – A measure of green spaces in urban areas. *Landscape and Urban Planning*, *105*(3), 325–335. https://doi.org/10.1016/j.landurbplan.2012.01.003

James, P., Hart, J. E., Banay, R. F., Laden, F. 2016. Exposure to greenness and mortality in a nationwide prospective cohort study of women. *Environmental Health Perspectives*, *124*(9), 1344–1352. https://doi.org/10.1289/ehp.1510363

Jorgensen, A., Gobster, P. H. 2010. Shades of Green: Measuring the Ecology of Urban Green Space in the Context of Human Health and Well-Being. *Nature and Culture*, *5*(3), 338–363. https://doi.org/10.3167/nc.2010.050307

Hijmans, R. 2019. raster: Geographic Data Analysis and Modeling. R package version 3.0-7.
https://CRAN.R-project.org/package=raster

Leutner, N. H, Schwalb-Willmann, J. 2018. RStoolbox: Tools for Remote Sensing Data Analysis. R package version
0.2.3. https://CRAN.R-project.org/package=RStoolbox

Meerow, S., Newell, J. P. 2017. Spatial planning for multifunctional green infrastructure: Growing resilience in Detroit. *Landscape and Urban Planning*, *159*, 62–75. https://doi.org/10.1016/j.landurbplan.2016.10.005

Nelson, J., Grubesic, T. 2018. Environmental Justice: A Panoptic Overview Using Scientometrics. *Sustainability*, *10*(4), 1022. https://doi.org/10.3390/su10041022

Nor, A. N. M., Corstanje, R., Harris, J. A., Brewer, T. 2017. Impact of rapid urban expansion on green space structure. *Ecological Indicators*, *81*(September 2016), 274–284. https://doi.org/10.1016/j.ecolind.2017.05.031

Nordbø, E. C. A., Nordh, H., Raanaas, R. K., Aamodt, G. 2018. GIS-derived measures of the built environment determinants of mental health and activity participation in childhood and adolescence: A systematic review. *Landscape and Urban Planning*, *177*(1432), 19–37. https://doi.org/10.1016/j.landurbplan.2018.04.009

Panagopoulos, T., González Duque, J. A., Bostenaru Dan, M. 2016. Urban planning with respect to environmental quality and human well-being. *Environmental Pollution*, *208*, 137–144. https://doi.org/10.1016/j.envpol.2015.07.038

Patino, J. E., Duque, J. C., Pardo-Pascual, J. E., Ruiz, L. A. 2014. Using remote sensing to assess the relationship between crime and the urban layout. *Applied Geography*, *55*, 48–60. https://doi.org/10.1016/j.apgeog.2014.08.016

Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal,
https://journal.r-project.org/archive/2018/RJ-2018-009/

QGIS Development Team, 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org .

R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

RStudio Team, 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. http://www.rstudio.com/.

Rigolon, A., Browning, M. H. E. M., Lee, K., Shin, S. 2018. Access to Urban Green Space in Cities of the Global South: A Systematic Literature Review. *Urban Science*, *2*(67), 18–21. https://doi.org/10.3390/urbansci2030067

Rhew, I. C., Vander Stoep, A., Kearney, A., Smith, N. L., Dunbar, M. D. 2011. Validation of the normalized difference vegetation index as a measure of neighborhood greenness.

Annals of Epidemiology, *21*(12), 946–952. https://doi.org/10.1016/j.annepidem.2011.09.001

Tomita, A., Vandormael, A. M., Cuadros, D., Di Minin, E., Heikinheimo, V., Tanser, F., … Burns, J. K. 2017. Green environment and incident depression in South Africa: a geospatial analysis and mental health implications in a resource-limited setting. *The Lancet Planetary Health*, *1*, 152–162. https://doi.org/10.1016/S2542-5196(17)30063-3

Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., James, P. 2007. Promoting ecosystem and human health in urban areas using Green Infrastructure: A literature review. *Landscape and Urban Planning*, *81*(3), 167–178. https://doi.org/10.1016/j.landurbplan.2007.02.001

UN-Habitat. 2012. *State of Latin American and Caribbean Cities 2012, Towards a new urban transition*. Naples: UN-Habitat.

United Nations. 2018. *World Urbanization Prospects: The 2018 Revision [key facts]*. Retrieved from https://esa.un.org/unpd/wup/Publications/Files/WUP2018-KeyFacts.pdf (6 July 2018).

Ward Thompson, C., Roe, J., Aspinall, P., Mitchell, R., Clow, A., Miller, D. 2012. More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns. *Landscape and Urban Planning*, *105*(3), 221–229. https://doi.org/10.1016/j.landurbplan.2011.12.015

Wickman, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wolch, J. R., Byrne, J., Newell, J. P. 2014. Urban green space, public health, and environmental justice: The challenge of making cities "just green enough." *Landscape and Urban Planning*, *125*, 234–244. https://doi.org/10.1016/j.landurbplan.2014.01.017

Wood, L., Hooper, P., Foster, S., & Bull, F. (2017). Public green spaces and positive mental health – investigating the relationship between access, quantity and types of parks and mental wellbeing. *Health & Place*, *48*(July), 63–71. https://doi.org/10.1016/j.healthplace.2017.09.002