Urban Environmental Performance Index: The Quito Pilot Case

Diego C. Manya
Yale University, diego.manya@yale.edu

Ryan Thomas
Yale University, ryan.thomas@yale.edu

Matthew Moroney
Yale University, matthew.moroney@yale.edu

Diego Manya, MEM 2018
Mathew Moroney, MEM 2018
Ryan Thomas, Research Associate

Localizing the Environmental Performance Index: Lessons and Approaches from Piloting an Urban EPI

Existing indices that examine urban sustainability do not analyze which communities gain from improved environmental benefits and which are disproportionately impacted by environmental hazards. We analyzed **49 Indices and 412 urban sustainability indicators**. Standardized metrics and definitions are lacking, making it difficult to compare urban sustainability initiatives from one city to the next. In this data project, we explored methods of using openly available data to determine the distribution of air quality and solid waste collection in relation to different socioeconomic groups in Quito. This represents a first attempt at developing an urban environmental performance framework that accounts for overall performance as well as the distribution of environmental hazards and benefits within cities. Where possible, the data were collected using scripts and scraping rather than relying on human readable data sources. This will help us scale the analysis up to other cities moving forward.

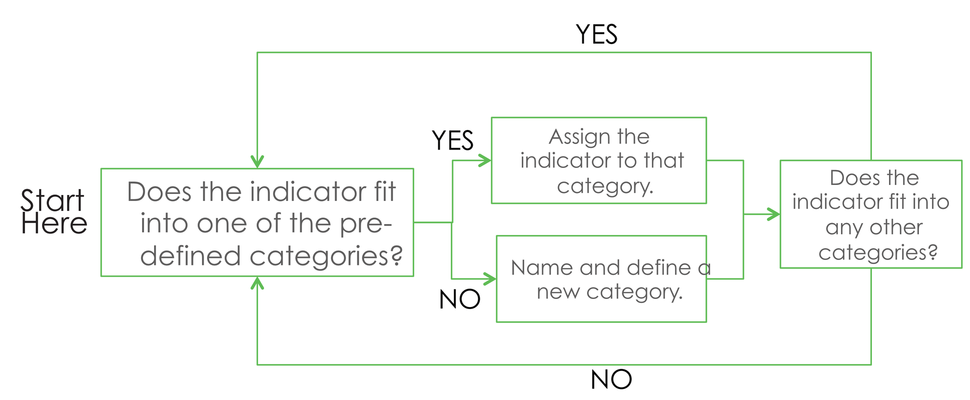
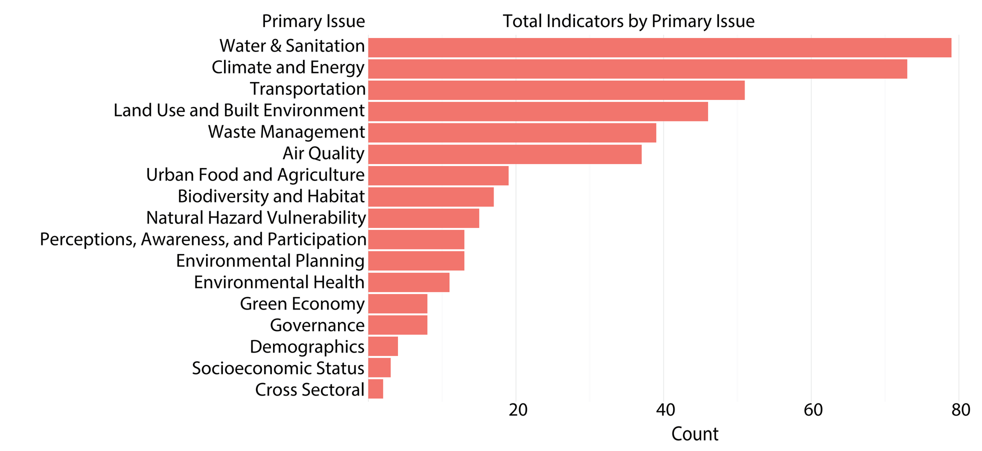


Fig. 1. We used a qualitative, iterative coding process to identify trends in the indicators we reviewed.



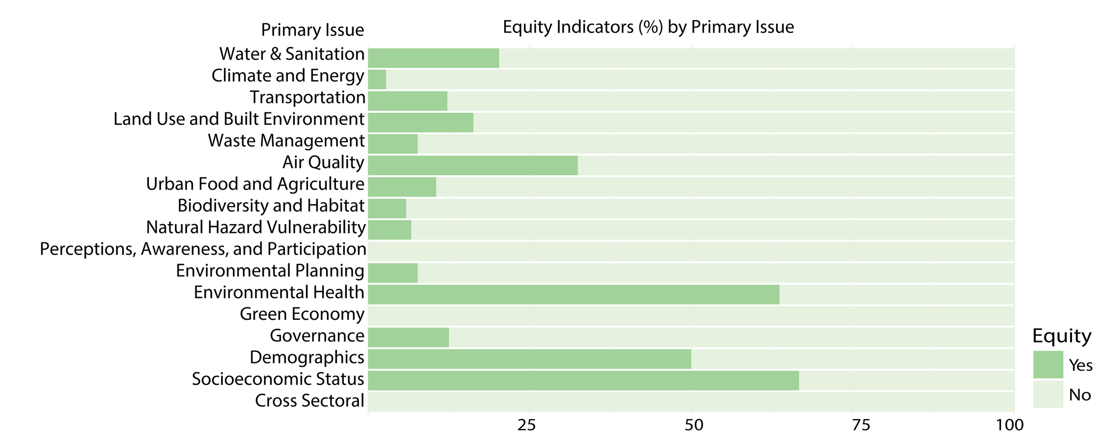


Fig. 2. After coding the indicators, we analyzed them to see whether they emphasize equity. The few indicators that explicitly incorporate equity are from the same indices: Boston Indicators Project (The Boston Foundation, 2016), and Cumulative Environmental Vulnerability Assessment (Huang and London, 2012).

Air Quality

Air quality is an integral component to examining a city's environmental performance, being an useful measure to compare cities. Urban areas with more industry and more automobiles typically have worse air quality. Air quality data is normally collected for several air pollutants such as Nitrogen Dioxides (NO), Sulfur Dioxides (SO2), Particulate Matter (PM), Ozone (O3), and Carbon monoxide (CO). These pollutants reduce visibility in cities and cause health effects to people who inhale them. Air quality data is normally very time-dense -- hourly observations for years with many stations. This data resolution allows detailed examinations of temporal trends: daily, seasonal, and long term trends.

1. Collection

The raw air quality data was collected from the Quito government's website, **quitoambiente.ec.gob**. There are a total of eight air monitoring stations distributed throughout the city of Quito. The city government monitors for four common air quality contaminants: Sulfur Dioxide (SO2), Carbon Monoxide (CO), Particulate Matter of 2.5 microns or less (PM 2.5), and Ozone (O3). There are over 7 years of data.

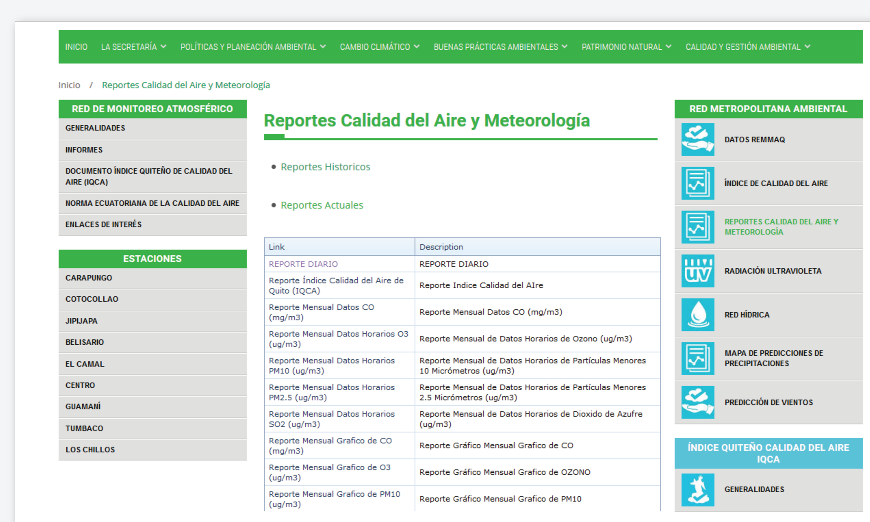


Figure 3. The data was stored in separate spreadsheets for each chemical and each monitoring station. A bash script was used to download and compile approximately 500 data files into a single csv for analysis.

2. Cleaning

The compiled csv data file was imported to R statistical coputing software. Before analysis, the data had to be cleaned into proper variable types. Two main packages were used to clean and tidy the data prior to analysis: lubridate to standardize dates and times and dplyr to compile the data into a tall format. All non-application and non-detect (NA and ND) values were dropped from the analysis. Data covered the period between April 2008 and September 2015, however several stations reported no data during the period.

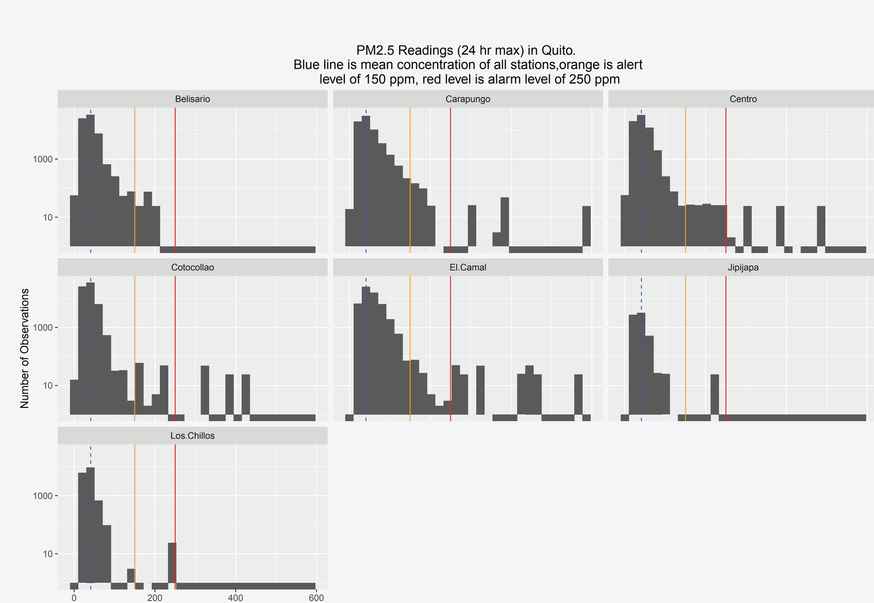


Figure 4. Each histogram above represents the entirety of PM data for each station in Quito. Some stations never see concentrations that exceed the alarm level (such as Belisario and Jipijapa), whereas some see it exceed somewhat more often (such as Carapungo and El Camal). This is one indicator of unequitable pollution distribution. The orange line represents the alert level and the red level represents the alarm level. Many of the individual high values (between approximately 300 and 600 ppm) occurred during New Year's Eve. This is most likely because of cultural traditions of lighting fireworks and burning effigies. These higher concentrations represent increase risk to human health if inhaled.

3. Analyzing

The data was compared with the official air quality standards. Safe concentration levels are based on rolling means from hourly data. PM2.5 is monitored on a 24 hour rolling mean. The data showed that only particulate matter had a significant number of exceedances. Some stations (such as Belisario) have a tighter distribution of PM values where some stations (such as Carapungo and El Camal) has a much wider distribution. These could be indicators of cultural factors such as increased vehicle traffic, increased cooking with low quality fuels, or increased manufacturing. Our research intends to explain some of the observed air quality variability with socioeconomic variables that substitute as proxies for poverty.

4. Visualization

Each air quality contaminant was plotted to tease out visual trends. Plots were constructed using the ggplot2 package for R. Three main plots were used: time-series plots, histograms, and boxplots. These plots were chosen to examine the temporal trends and distributional trends of the dataset. Plots for PM are shown below as a typical example of how air quality data looks visually.

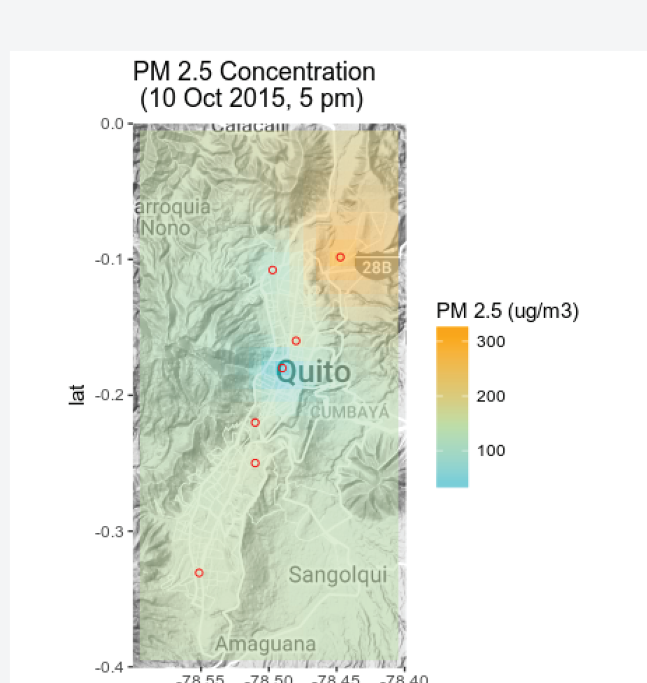


Figure 5. Based on the few points available, we calculated a spatial interpolation. Due to limited readings, the results carry a large error. However, when data collection is scaled up to include cities with higher spatial density of data points, this techinque will be replicable.

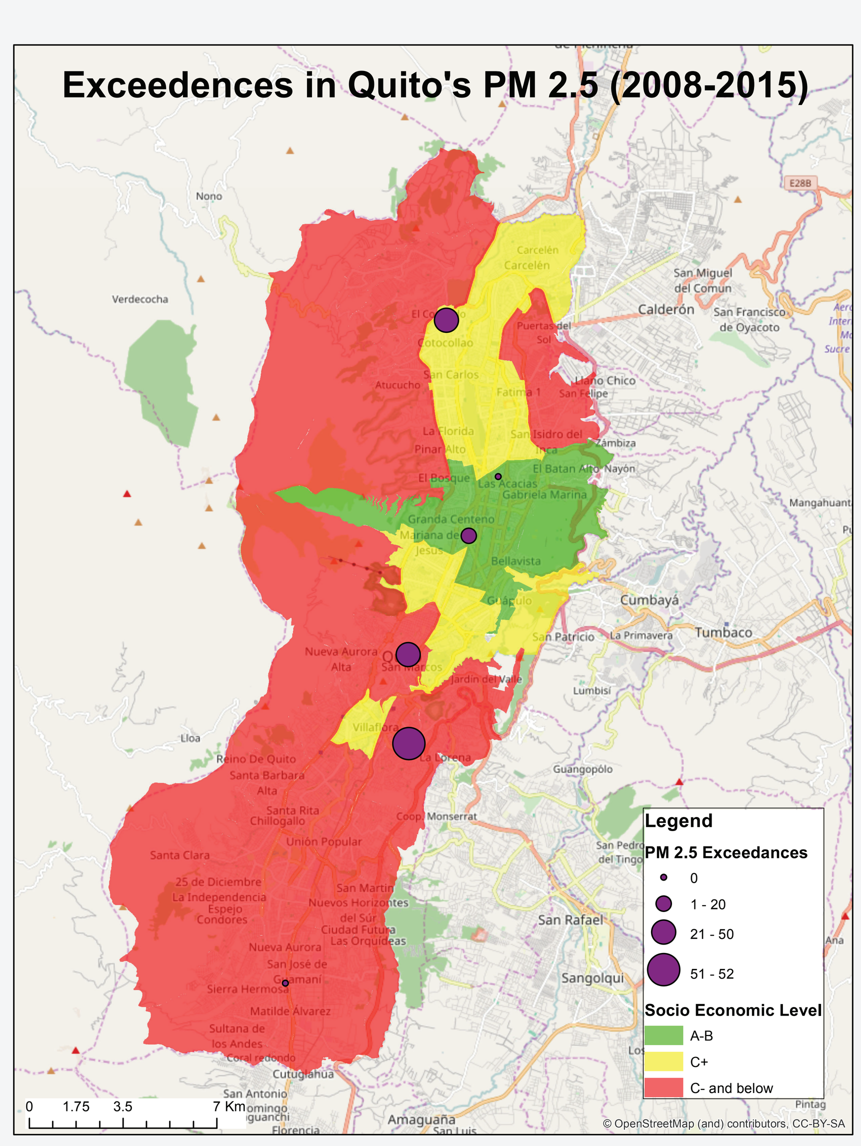


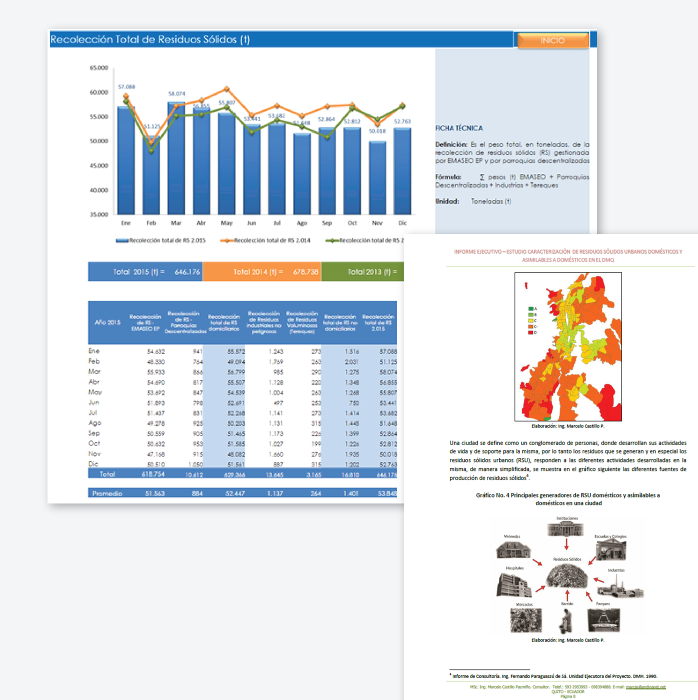
Figure 6. This map shows the socioeconomic level and the number of exceedances recorded from each station. It does appear that sensors in poorer areas report a higher number of exceedances.

Waste Management

1. Data Collection

Waste collection data are typically collected from the local government, landfill tipping fees, and private waste collectors. The demographic and geographic data was collected mostly from Quito's Metropolitan Municipality, and the waste management information for domestic non-hazardous waste from the Public Company of Cleaning Services.

However, the information available through the local government platforms was, in most cases, in formats that didn't allow for readily analysis, such as: (a) Data was stored in non-machine readable formats like PDF; (b) Data was presented as final outputs of previous analysis, without option to access the source data. To overcome this issues, several techniques had to be used, these included: (a) Use of tools for extraction of tables from PDF files to csv or Excel formats (i.e. Tabula); (b) Reconstruction of geographic data using transformation between formats (i.e PDF to DWG to SHP);



0.85 kg generated
person/day

2. Disaggregation

The collected data provided an overview of the **volume** of waste managed by the city of Quito and some outlying admin istrative areas; however, to examine the environmental performance of the city **while addressing equity**, we needed to disaggregate the data and understand how this number varies over socio-economic groups.

To disaggregate these data, we used statistical and mathematical operations to estimate the population, socioeconomic level, waste collection, and uncollected wast in each of Quito's parroquias (parishes).

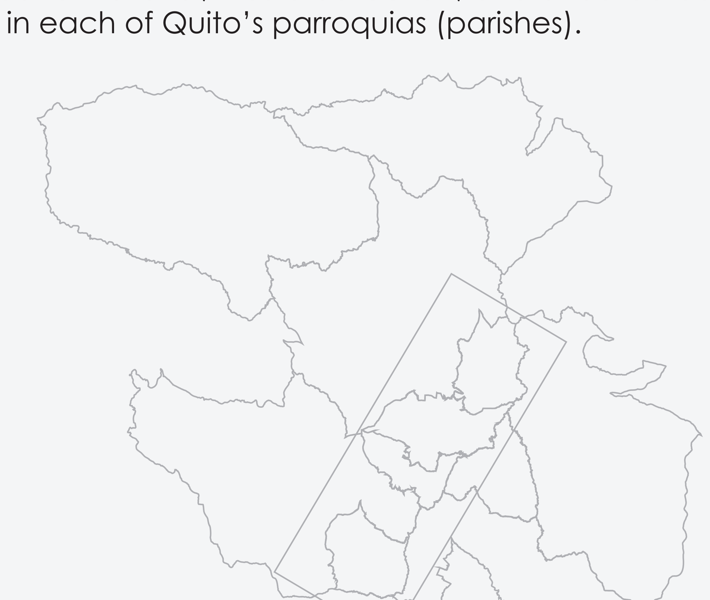


Figure 5. Municipal boundary of Quito with approximate study area (urbanized area) outlined.

Administrative Zone	Parish	2015 Population	Socio-Economic Level	Generation Per Capita (kg/PP/Day)
Eloy Alfaro	**Chilibulo**	50167	4.00	0.861
	Chimbacalle	40997	3.56	0.920
	La Argelia	58982	4.03	0.862
	La Ferroviaria	67562	3.87	0.878
	La Magdalena	31730	3.09	0.983
	La Mena	45663	4.00	0.861
	San Bartolo	65928	3.51	0.926
	Solanda	80588	3.77	0.892

The City of Quito publishes population estimates for each parroquia in 2010 and projects the pronvice's population for each year up to 2020. We used these estimates to estimate the 2015 population of each parroquia.

The waste generation was disaggregated by geographically extrapolating the socioeconomic level of each parish, based in the socioeconomic level of the waste collection routes in that area. Then, using the official waste generation ratio for each socioeconomic level, we created a range from which we calculated the specific waste generation (kg per person per day) for each of the parishes.

3. Analysis

With all the information, we could now do a proper analysis to evaluate their environmental performance and its distribution in the city as a way to address equity. The component selected for analysis was the Coverage of waste management collection system in each parish. This component was measured by the following indicator: Kilograms per person per day of domestic waste that is not collected by the waste management system (Kg/person/day).

The results of this analysis, showed that parishes with lower socioeconomic level have a higher level of waste not collected that those with higher socioeconomic level. A statistical analysis of the data shows that there is a 45% correlation between the two variables, and that the Socioeconomic waste is a significant predictor of the amount of waste not collected.

4. Visualization

The final step of the process is to present the information in way that can rapidly convey the differences in the performance of waste management within the city.

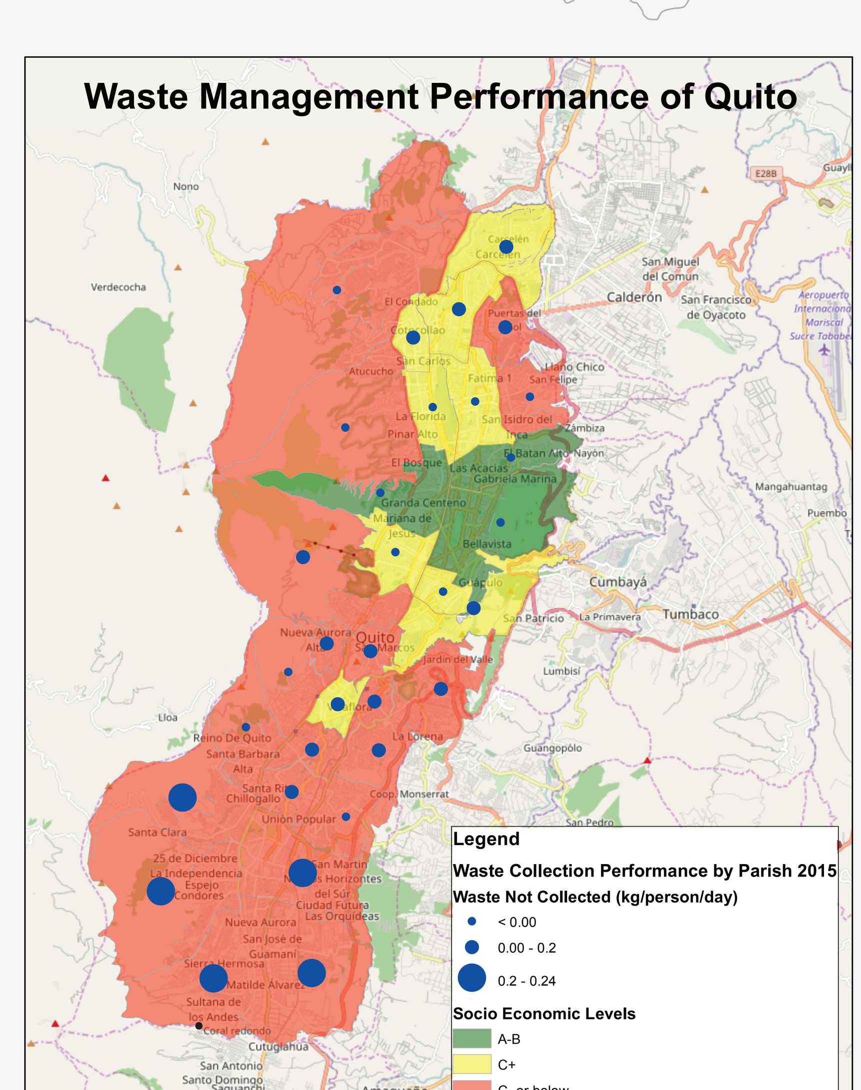


Figure 7. This map shows the socioeconomic level and the quantity of uncollected waste in each parroquia, reflecting the same data but highlighting that uncollected waste is in areas at a lower socioeconomic level.