Poverty Population and Its Educational Accessibility: An Evaluation using Geospatial Database in Ecuador

Gustavo Navas, Anthony Loyaga, Jefferson Barahona and Julio Proaño Orellana

Universidad Politécnica Salesiana, Morán Valverde S/N y Rumichaca, Quito, Ecuador

Keywords: Geospatial, GIS, Voronoi, Educational Centers, Spacial Analysis, Poverty.

Abstract: Education is the most critical aspect against inequality. So, the accessibility to educational centers is one of the biggest problems that governments and academia are trying to solve in Latin America. Finding an optimal location is challenging because different variables can be considered, such as serving the poor, accessing to roads, and optimal distribution. This work presents a methodology that uses geospatial techniques to evaluate the level of access of poor people to educational centers considering location and accessibility. This methodology is multi factorial and can be generalized to multiple scenarios in resource planning. In this work, a case study of Ecuadorian educational centers is analyzed. So, an Ecuadorian geospatial database was created that includes educational centers’ location, the distribution of the poor population, Voronoi diagrams, and road networks. All the analysis were carried out using the PostGIS tool and Google Cloud Platform for the implementation. The results show that 90% of educational centers serve at least 50% of the most impoverished population.

1 INTRODUCTION

Education is the most critical aspect to mitigate inequality problems. However, according to UNICEF (UNICEF, 2019), more than 60 million children will not access primary education in 2030. On the other hand, urban and regional areas have experienced considerable growth (Navarrete and Luján-Mora, 2018). It has led to the authorities’ different planning and distribution of educational facilities to find the most suitable location (Brock and Schaefer, 2015). Educational institutions in Ecuador are characterized by low educational quality and the absence of minimum conditions for girls, boys, and young people. On the other hand, education in Ecuador has lacked an academic level of quality in the poorest populations. In 2005, the educational project of the Educational Units of the Millennium began, which aims to reduce this gap and reach children with limited economic resources. The Millennium Educational Units (EMU) is a project for the endowment of public educational institutes of primary and secondary levels. It was created to improve its education and reach the poor sectors. For this reason, its buildings are located in sectors historically relegated to the national level, which present high indices of unsatisfied basic needs and social problems. Each EMU is built to guarantee access for the school population in rural areas permanently excluded from educational services. EMUs are based on various location criteria for their construction. However, it is challenging to choose the optimal location of educational centers based on spatial analyses. It is essential to fulfilling all social demands considering poverty and accessibility. Various methodologies and approaches help planners and governments to address system inefficiency problems through Geographic Information Systems (GIS) (Al-Sabbagh, 2020), (Bulti et al., 2019), (Boix and Olivella, 2007). The Geographic Information System provides reliable data with spatial references that decision-makers could use to guarantee the optimal distribution of schools based on the impact of the children served. The study of location optimization is not new. Thus, since 1909 Weber has studied the optimization of the location of a company in a region (Okabe and Suzuki, 1997). However, in recent works, there are applications of these studies related to the location of schools. In (Jabbar and Laffta, 2020), the authors propose a methodology to analyze location sites of private secondary schools...
in the Al-Jihad neighborhood in Baghdad-Iraq. This methodology determines the spatial distribution patterns of the private secondary schools in the study area and the distance between the secondary school sites in the neighborhoods and, finally, in the appropriate locations by analyzing the spatial distribution of the schools and determining suitable locations. In another study conducted in Jeddah, Saudi Arabia (Murad et al., 2020), the authors used a geodatabase that incorporates data on education and population. Based on these data, spatial and network analyzes were used to understand the distribution of schools, student density, and accessibility. However, the authors do not mention the optimal location and whether their methodology can be applied to develop a new methodology that incorporates a variety of parameters. Therefore, the current supply should be sufficient to cover the demand. However, the author does not consider other parameters such as road accessibility.

Other approaches include additional parameters. For example, in (Al-Sabbaghi, 2020), the authors propose a study that uses GIS tools to locate elementary schools based on different strategies such as current elementary school distribution, average nearest neighbor, hotspot analysis, and clustering analysis. Furthermore, in (Sezer et al., 2018) the study applies the network analysis to locate the closest schools in the city of Usak. The study also considers the number of students, teachers, classrooms, roads, and buildings. These data were obtained from the Usak municipality zoning plan and the OpenStreetMap vector. In all of these previous works, the use of bigdata is missing.

Therefore, the main contribution of this work is to extend previous methodology (Navas et al., 2019) to evaluate the impact of the school’s location in the population with attendance poverty in Ecuador, but taking into account the road accessibility and the distribution of the Voronoi Diagram as optimal distribution. It is worth noting that the combination of economic data (poverty) and GIS tools such as Voronoi and road accessibility is a challenge because our study belongs to all of the country (bigdata).

So, in this work, an extension of a novel methodology is proposed for evaluating the location of schools based on multiple factors such as poverty level, road accessibility, access distance, and a Voronoi Diagram. It is challenging because the use of mixed bigdata, including demographic and spacial data (Yan et al., 2010).

2 METHODS

Before starting the description of the methodology, terms and basic operations of spatial data analysis are briefly explained.

- The Voronoi Diagram. It is generated from \( N \) spatial objects (geographical points or lines), which must generate \( N \) Voronoi polygons that delimit each object, where the number of sides of this polygon is according to the number of closest neighbors, consequently generating a polygon that contains the area of influence of the object (Erwig, 2000).

- The Buffer Operator. Returns a buffer zone at a specified distance from a geographic point or line. The type of geographic object obtained is a polygon, and what is inside the polygon is called a buffer zone.

- The Intersection Operator. Returns a zone of influence corresponding to the intersection or common area between two classes of different geographic entities. The Intersection, is that operator where there are two areas that give place to the intersection or common area between them.

- Population Density. The \( \rho \), corresponds to total population density in number of inhabitants per square kilometer (hab/km\(^2\)).

Next, the mathematical formalization is established in a subsequent subsection to show how it is applied and finally land the case study’s solution.

2.1 Methodology Development

The objective of this methodology is to establish possible locations that are the most appropriate according to the variable considered. In our case study the location corresponds to new sites to locate new pre-university study centers and the parameter to consider is poverty and the accessibility. From step 1 to step 8, the existing locations are analyzed based on the parameter studied to order them from highest to lowest depending on the parameter. Starting from step 9, the procedure is given to establish the new locations. The process is explained bellow (see Figure 1).

Step 1: First, set the generators “1”, that are the fundamental input for steps 2 and 3 that generate the diagrams of Voronoi and Buffer tool respectively. The
generator turns out to be the geographical location of all the units that exist, are located within area \( S \) and which is being studied. In our case study they are educational units and it is always considered a point type.

**Step 2:** The generators from step 1 will be used for the Voronoi diagrams, the first classification of these limits comes given by the result of applying the Voronoi diagram in a established area. Voronoi diagrams cover the entire space of study. **Step 3:** Corresponds to the Buffer tool around the generators set in the first step. One feature is that there must be as many areas generated by the Voronoi diagram as areas generated by the Buffer, and these must be equal in quantity to the number of geometries generated in step 2.

\[
B_{\text{io}}(\nu, \text{generators, distance}) = \beta_{\nu}(\nu, s_{1i}, d)
\]  

where \( B_{\text{io}} \) occupies a subspace of \( S \)

**Step 4:** Intersection operator between the generated areas in steps 2 and 3. For this case there can be two types of polygons resulting.

- That the buffer is completely within the Voronoi diagram correspondent.
- Let the two areas intersect in a common area.

\[
I_i = B_{\text{io}} \cap \nu_i
\]

The \( i \) index traverses each of the generators for both the polygons of Voronoi as well as for the Buffer areas generated.

**Step 5:** It is a new generator ”\( s_{2i} \)” that corresponds to ways land, river, etc. found in area \( S \).

**Step 6:** It is a buffer around the generators ”\( s_{2i} \)” related to the access roads, which were established in step 5. This buffer corresponds to an area of influence of the road, depending on the closeness to her.

\[
B_{\text{il}}(S, \text{generators, distance}) = \beta_{\nu}(\nu, s_{2i}, d)r
\]

**Step 7:** It is an intersection operation between the results from step 4 and step 6. Giving rise to a new result that is:

\[
I_{\text{il}} = B_{\text{il}} \cap I_i
\]

Applying \( I_i \), we have:

\[
I_{\text{il}} = \left( B_{\text{il}} \cap ( B_{\text{io}} \cap \nu_i ) \right)
\]

This relates the areas linked to the units with accessibility.

**Step 8:** It is an intersection between the geopolitical division with the analyzed parameter ”\( \Gamma_j \)” and the step 7 result ”\( I_{\text{il}} \)” . It gave rise to a new result that is formalized as follows manner.

\[
I_{\text{ji}} = \Gamma_j \cap I_{\text{il}}
\]

Applying \( I_{\text{ji}} \), we have:

\[
I_{\text{jil}} = \Gamma_j \cap \left( B_{\text{il}} \cap ( B_{\text{io}} \cap \nu_i ) \right)
\]

\( \Gamma_j \) corresponds to the number of political divisions, which contains the parameter to be analyzed. Therefore, the index \( j \) is different from index \( i \), since \( j \) corresponds to the number of political divisions within ”\( I_{\text{ji}} \)” . The parameter in \( \Gamma_j \), can be diverse, in some cases, it could be the poverty index and in another moment could be the Population Density H. The intersection gives rise to a series of sub-areas for each unit provided in generator ”\( s_{1i} \)” . Every sub-area includes its value of the parameter.

It is analyzed for sub-areas for each generator 1, the procedure described by (Navas et al., 2019) for \( I_{\text{jil}} \) through the formula 8 of weighting a group of data \( X_1, X_2, \ldots, X_n \), considering a series of coefficients or parameters for each political division, called

![Figure 1: Methodology development diagram.](image-url)
weights $W_1, W_2, \ldots, W_n$. It can determine the level of
the parameter in each weighted political division of
each irregular polygon.

$$\bar{X}_n = \frac{(w_1X_1 + w_2X_2 + \ldots + w_nX_n)}{(w_1 + w_2 + \ldots + w_n)}$$  \hspace{1cm} (8)

**Step 9:** Incorporate the population density parameter
$H$ at the level of political divisions it corresponds to
$\Gamma_H$, and select those political divisions that incorpo-
rate the conditions of selection. These selected divi-
sions will be called "candidate political division" $\Gamma_{Hk}$.

$$\Gamma_{Hk} \in \Gamma_H :$$  \hspace{1cm} (9)

_Step 10:_ Location of the centroid for each of the Can-
didate political divisions.

$$\forall \Gamma_{Hk} \in \Gamma_H \ni s3_k$$  \hspace{1cm} (10)

Where "$s3_k$" is the centroid of the divisions of each
political candidate $\Gamma_{Hk}$. And Where "$s3_k$" is a new
generator.

**Step 11:** With the Buffer tool around the generator
obtained in step 10, it was generated the area of in-
fluence of each candidate centroid "$s3_k$". One charac-
teristic is that there must be as many areas generated by
the buffer tool as the number of candidate centroids
"$s3_k$" that was generated in step 10.

$$B_{s3,ik}(S, generators, distance) = B_s(v_{ik}, s3_k, d)$$  \hspace{1cm} (11)

**Step 12:** the goal of this step is to reduce the number
of areas generated for each candidate centroid, from
step 11, with those.

$$B_{ik} \in B_{s3,ik}$$  \hspace{1cm} (12)

Where $B_{ik}$, are the reduced candidate areas corre-
sponds to a reduced set of $B_{s3,ik}$

**Step 13:** An intersection between the candidate areas
reduced by the area corresponding to the roads ob-
tained in step 6. $B_{ik}$ is carried out, with the area of
influence of step 8.

$$I_{iki} = B_{ik} \cap B_{ii}$$  \hspace{1cm} (13)

### 2.2 Case of Study

The following case study was carried out in Ecuado-
rian territory. In this way, it was essential to ho-
menize the geographic data for proper geoprocessing.
Thus, the geographic reference system was deter-
mined in EPSG / SRID code 32717. This identifier corre-
sponds to Datum WGS84, UTM zone 17S. The Post-
greSQL Database Management System was used with
its extension for GIS, PostGIS, to carry out this geo-
processing. Steps 1 and 2 are shown in Figure 2.

**Step 1:** In the Ecuadorian context, the geographical
information on the positioning of educational institu-
tions is registered by the MINEDUC. The point type
geographic records are found in the geographic refer-
sence system corresponding to the EPSG / SRID code

**Step 2:** Before creating the Voronoi Polygons, it is
necessary to have geographic information of the poly-
gon type that describes the political boundary of the
country, whose function is to create the Voronoi con-
tour and segment the Ecuadorian territory. This op-
eration results in the Voronoi segmentation of educa-
tional institutions whose external contour is a quadri-
lateral by default. Here, using the intersection opera-
tion, only the Ecuadorian territory is outlined, which
is the region of this case study (Gutiérrez, 2006)
(Llario, 2013).
these two paths converge. It is possible through the \textit{ST\_Intersection} operation, which takes two geometries as parameters and returns the space intersect (see Figure 4) (Gutiérrez, 2006) (Llario, 2013).

**Figure 4:** Intersection of Voronoi polygons of each EMU with 20 km buffer.

**Step 5:** In Figure 5, the geographic information of the Ecuadorian road network can be found on the official IGM site and the “multiline” type geographic data set can be found in the EPSG/SRID 4326 reference system. For the conversion of geographic data to the EPSG/SRID 32717 reference system specified in step 1, the \textit{ST\_Transform} operation was used (Gutiérrez, 2006) (Llario, 2013).

**Figure 5:** Ecuadorian road network.

**Step 6:** The zone of influence of the road network can also be obtained with the \textit{ST\_Buffer} operation. In this case, being of type line, the parameter within the operation indicates the width of the buffer. For this, the relative distance of 2 km from the road’s edge was applied to each side. That distance was estimated considering that there are records of children who have to walk up to 30 minutes from their home to reach school transport on the nearest road. With this, we know how far the schoolboy can walk, but not the distance. However, if the ACSM is considered, which indicates that an average adult can reach a 5 km/h average walking pace, that is, the adult would walk 2.5 km every 30 min. If we extrapolate this value to the reality of a student, we can consider that in the best of cases, it would be expected that they could walk approximately 2 km (see Figure 6 (Gutiérrez, 2006) (Llario, 2013).

**Figure 6:** Ecuadorian road network buffer.

**Step 7:** In this step, the expected zone of influence of the methodology is determined, where the zone of influence of the educational institution (step 4) converges with the buffers that describe the Ecuadorian road network (previous step). Thus, finally, through the \textit{ST\_Intersection} operation, it can be seen in 7, the area of influence of each educational institution, including its road accesses (see Figure 7 (Gutiérrez, 2006) (Llario, 2013).

**Figure 7:** Coverage area of each EMU through its road accesses.

**Step 8:** The information that describes the level of poverty (according to the NBI indicator) in Ecuador, is available at the National Institute of Statistics and Censuses. The granularity in which this information is found is at the parish level, accompanied by polygon-type data, which describes its spatial position and limits as is shown in Figure 8.

In this way, we proceeded to intersect the area of influence from the previous step, contrasting the polygons that describe the parishes where each institution is located, to assign a percentage of these intersections impact (Gutiérrez, 2006) (Llario, 2013). From this step it was intended to obtain potential locations for new educational institutions. An update was made, going from 77 (used in the previous steps) to 97 institutions. Once the institutions are adjusted to the geographic reference system used, the study continues.

**Step 9:** To obtain a candidate political division, the selection conditions were taken: poverty $\geq 90\%$ and population density $\geq 5$ inhabitants/km$^2$, where the population density data was obtained from the INEC. Resulting in 379 candidate political divisions (see Figure 9) (Gutiérrez, 2006) (Llario, 2013).
**Step 10:** With the results of the previous step, an ST\_Centroid operation was performed on each of the candidate political divisions.

**Step 11:** Once the centroids were obtained, the ST\_Buffer operation was performed to obtain their respective 20 km buffers, presenting intersections with the existing institutions. To eliminate the intersections, a differential analysis was performed using an ST\_Difference operation, obtaining 298 candidate political divisions.

**Step 12:** The buffers obtained in the previous step present overlapping and to simplify them, the following points are taken into consideration.

To assign the demand to the solution, the selection conditions of Candidate Political Divisions were taken (Step 9), based on the heuristic method. On the other hand, for the random substitution of candidate sites, the centroids that cover most of the Candidate Political Divisions and do not have an intersection of more than 40% between the existing institutions were taken.

Next, it seeks to choose the buffers that occupy the most territory of the Candidate Political Divisions, based on the Maximum Coverage Model (COBE-MAX), so that most of the potential users have access to the institutions.

Taking these considerations into account, a manual selection of these centroids with their corresponding buffers was carried out, obtaining a result of 90 Candidate Political Divisions (see Figure 11 (Gutiérrez, 2006) (Llario, 2013).

**Step 13:** At the time of the investigation, it was decided to carry out an intersection between the road map (Step 6) and the result of the previous step with the ST\_Intersection operation, in order to have a more applicable response. With the information obtained, it was possible to establish 49 possible institutions (see Figure 12 (Gutiérrez, 2006) (Llario, 2013).
3 RESULTS

Within the application of the methodology, it has been possible to categorize the EMUs according to the percentage of poverty that each unit covers, confirming what was stated by (Navas et al., 2019). On the fact that the units provide coverage to the poorest populations in Ecuador, as shown in Fig 15.

Among the institutions with a higher poverty level are EMU Chontapunta, EMU Cuyabeno, EMU Nuevo Rocafuerte, with 99.7%, 98.7%, and 98.2%, respectively, and those with the lowest impact are the replica of schools and schools. Emblematic, such as the case of the Réplica 28 de Mayo Educational Unit, the 24 de Mayo R éplica Educational Unit, and the Vicente Rocafuerte Réplica Educational Unit that reach a poverty level of 48.9%, 48.5%, and 47.5%, respectively, as shown in Fig 16.

The circular area of each of the units intersects with the polygons delimited by Voronoi, and with the buffer of the access roads and, the ease of accessing it through the tracks. As a result of the execution of steps 9 to 13, a coverage of 60.62% of the national territory was obtained between the existing institutions and the new ones that are proposed. In order to give a more tangible result, the first 10 institutions where the level of poverty are the highest were taken. The results obtained describe a name for the institution, parish, canton, province, percentage of poverty, population density (inhab/km²) and coordinates in UTM format zone 17 S with SRID 32717 with their respective X, Y axis, as can be seen in (Table 1).

The circular area of each of the units intersects with the polygons delimited by Voronoi, and with the buffer of the access roads and, the ease of accessing it through the tracks. As a result of the execution of steps 9 to 13, a coverage of 60.62% of the national territory was obtained between the existing institutions and the new ones that are proposed. In order to give a more tangible result, the first 10 institutions where the level of poverty are the highest were taken. The results obtained describe a name for the institution, parish, canton, province, percentage of poverty, population density (inhab/km²) and coordinates in UTM format zone 17 S with SRID 32717 with their respective X, Y axis, as can be seen in (Table 1).
4 CONCLUSIONS

This document proposes an evaluation methodology using Geospatial tools to address the optimal location of schools by taking factors such as poor population, distance, spatial distribution, and road access. In addition, a case study is analyzed with educational centers obtaining promising results. The feasibility and rationality of the method proposed in this document are verified, considering that it has been applied to the entire country and its road network. The results show that 90% of the academic units are in areas where more than 50% of the poor population is concentrated, and there is also a road axis for access. It should be noted that the methodology can be easily adapted to other analysis scenarios and other study areas such as health, transport, logistics, etc.

A base political division of the study territory included the analysis variable, which allowed obtaining the results. All of this could be applied to locations that offer a service other than those of the EMUs, which would work similarly. The main advantage of using an amalgamation of methodologies to obtain potential institutions over a conventional method based on existing functions is having several considerations for the locations. At the same time, social inclusion is taken into account, giving way to more humanly applicable results instead of obtaining an exclusive result for not having some GIS-based feature. The result can be adjusted to more convenient locations, being able to carry out the process as many times as necessary, to obtain a response that is geographically valid. These results, coming from a heuristic method, are optimal and conform to the concepts of Spatial Efficiency. These institutions are as close to the population and access roads. In addition, the concept of Spatial Justice could be considered, where the educational services of the institutions are distributed in such a way that the largest possible part of the Ecuadorian territory is covered. Additionally, as future work, other features can be explored, such as safety and environment, and other techniques like reverse Voronoi, to find optimal school locations.

REFERENCES


