Issue Paper Proposal on Using Geospatial Data for Educational Planning

SCAALE: a geospatial workflow for distributing schools to provide prospective populations access to primary and secondary education (Students Can Access All Levels of Education)

Alexander Tedeschi and Dejan Popovic 23.08.2020

BACKGROUND

One of the fundamental tenets of the Universal Declaration of Human Rights - that everyone has a right to an education - has long been unanimously recognized by the international community. Many states have achieved universal primary education, while others continue to strive to reach this ideal. According to the UNESCO Institute for Statistics (UIS), sub-Saharan Africa has the highest rates of educational exclusion with over 20% of children between the ages of 6 and 11 out of school[1]. Dropout rate correlates with age, with almost 60% of youths between the ages of 15 and 17 out of school in the same region[2].

A multitude of barriers may exclude children from accessing education: living in a region plagued by conflict, school fees, poor nutrition, a lack of classrooms or teachers, a lack of hygienic facilities, and gender disparities that arise from forced marriage. Another major obstacle that is often overlooked is geographical proximity. In a watershed study published by the World Bank[3], it was found that *proximity was the single most important determinant of primary school enrollment*. This evidence was reinforced by subsequent studies, such as a randomized evaluation conducted by researchers in Afghanistan in 2009 that indicated

---

[1] New Methodology Shows that 258 Million Children, Adolescents and Youth Are Out of School. UNESCO UIS Fact Sheet No. 56. September 2019. (URL)
[2] UNESCO UIS. Education & Literacy: Education in Africa (lURL)
[3] Improving Primary Education in Developing Countries, by D. Block and A. Verspoor, 1991. Oxford University Press (URL)

geographic proximity had a dramatic effect on children's academic participation and performance, with enrollment falling by 16 percentage points for every mile that children must travel to school[4]. According to the Africa-America Institute, seven out of 10 rural youth in Africa have never attended school, and even if students do manage to graduate from primary school, secondary schools can accommodate only 36% of qualifying secondary students[5]. In light of the persisting digital divide in developing countries, thrown into stark relief by recent lockdowns, it is clear that distance education and online learning are not sufficient to close the gap. While physically getting children to school is only one part of the complex equation to improve educational access, it is by far the most important.

Population momentum further exacerbates the school supply problem. This is particularly intractable in the context of most African states, where population is predicted to double by 2050 and where 60% of the population is under 25[6] . There is no doubt that this demographic surge will profoundly shape Africa's future, and that it has serious implications on educational planning and provision on the continent. The graphs in Figure 1 and Figure 2 illustrate this demographic arc.

[4] "The Effect of Proximity on School Enrollment: Evidence from a RCT in Afghanistan", by D. Burde and L. Linden. 2009 (URL)
[5] "State of Education in Africa Report 2015. The Africa-America Institute. 2015 (URL)
[6] "More than half of the world's population growth will be in Africa by 2050", by Yomi Kazeem. Quartz Africa. June 19th, 2017. (URL)

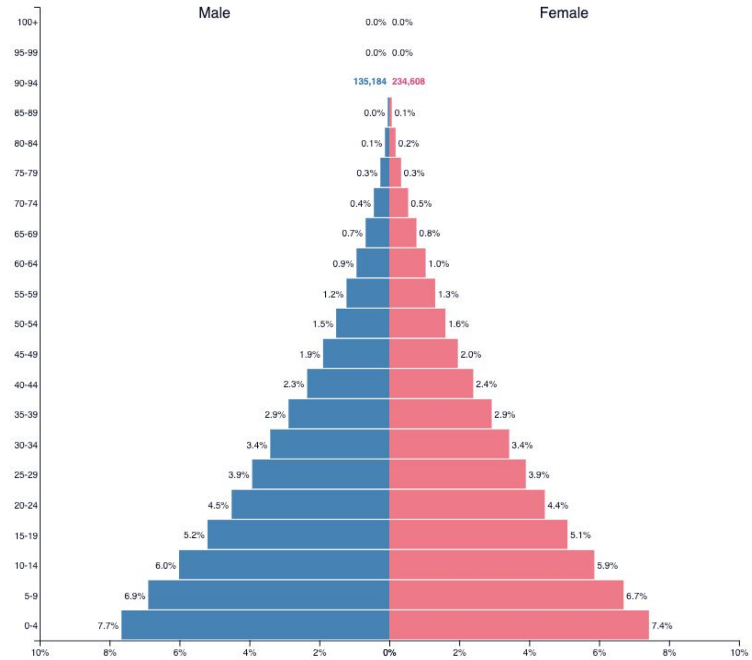


Figure 1. Africa’s population pyramid in 2019

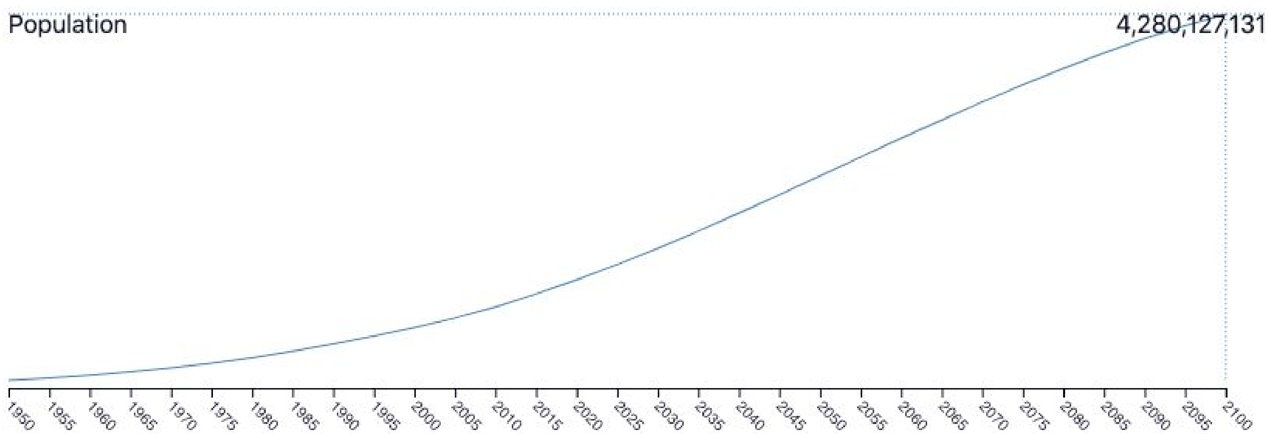


Figure 2. A Projected population growth in Africa to 2100[7]

Institutional shortage is as much a consequence of slow economic growth as it is of spatial inequality. For example, in Nampula, Mozambique, a provincial capital in the northeast of the country, there were only 2,232 primary and secondary schools (including 50 under construction) serving a school-age population of just under 2

[7] Population Pyramids of the World from 1950 to 2100 (URL)

million students in 2015, according to the national government portal[8]. Even without taking into account the spatial context of accessibility (catchment areas, etc.), this means the ratio of schools to students in the entire province is 1:893. Nampula serves as a microcosm, an archetypal example of the educational planning challenges at the provincial level facing countries with population momentum. For the purpose of the current proposal, Nampula will be referenced as an example to illustrate how the SCAALE methodology can practically be applied for proactive micro-planning.

OBJECTIVE

Defining the problem

To distribute or locate schools such that children have, at minimum, the first two basic educational levels accessible to them - primary and secondary.

The SCAALE method will be universally applicable to national, subnational, and district levels of educational planning, in particular where demand outstrips supply. The proposed methodology is intended to be useful to inform the decision-making process of *where* to build new schools or *where* to redistribute existing ones.

LITERATURE REVIEW

A review of existing literature on educational planning has revealed a lack of methodologies that take into account planning school locations for the first two basic levels of education. In *Rethinking micro-planning*, a number of useful techniques were summarized, but they tended to narrowly focus on locating one *type* of school - either primary, secondary, or postsecondary[9]. Examples included using gravity models to find university locations[10], finding catchment areas for

[8] "Nampula voi construir 50 novas escolas" Portal do Governo de Mocambique. 2015
[9] Rethinking micro-planning, Amelie A. Gagnon et al, UNESCO 2020
[10] Bruno, G. & Improta G. 2008

secondary schools[11] or locating elementary schools in fast expanding areas[12]. What also differs from other projects cited in the document is that our methodology relies on publicly available datasets, such as Gridded Population of the World (GPW), roads and points of interest from OpenStreetMap, and land cover maps from NASA and the European Space Agency. While inputs from educational planners are also expected given their intimate knowledge of local environments, the current methodology offers a basic, replicable workflow at every geographical scale.

Network analysis methods based on the literature cited in *Rethinking micro-planning* have been implemented for different applications (finding routes, distances, time distances) in order to determine either catchment areas or isochrones and aim to maximize the coverage of a client population. Geospatial methods have also been used for school mapping. The methods belonging to network analysis[13] and the 2SFCA[14] method drew our attention because they can calculate the ratio of supply-to-demand (schools- to-students) provide the weighted coefficients of distances[15]. The distance weights will be extracted as an input covariate for an implementation of AHP[16] model in the proposal. A Multi-Criteria Decision Analysis method such as AHP in combination with GIS was chosen for final site selection. Due to the simplicity of replication and execution of the workflow, AHP was chosen as the final step in our method.

The current issue paper proposes to combine established methods from a few of the common topics in educational planning to create a replicable methodology with the help of free, open-sourced geospatial analytical tools such as QGIS and R to be used by planners at the national, sub-national, and local levels.

---

[11] Singleton A. et al, Estimating secondary school catchment areas and the spatial equity of access 2010
[12] Menezes C. R. & Pizzolato N D, 2014 Brazilian Operations Research Society
[13] Maximum Covering Location Problem, Location Set Covering Problem, Capacitated Model
[14] Two-step floating catchment area
[15] Zhuolin T. et al BMC Health Services Research 2018
[16] Analytical hierarchical model

METHODOLOGY

CASE STUDY: Nampula, Mozambique

Locating primary & secondary schools

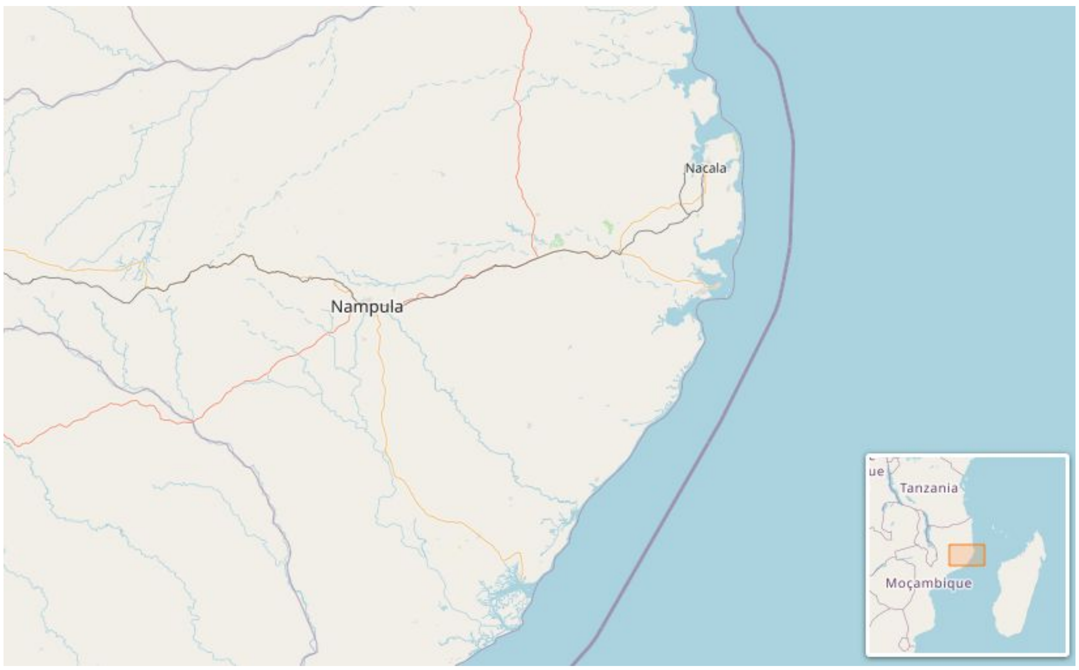


Figure 3. Nampula (Source: OpenStreetMap[17])

Multi-modal 2SFCA

The first part of our methodology is to implement a multi-modal 2SFCA method in order to measure the accessibility of existing schools via multiple forms of transportation (walking, driving, bicycling, public transport, etc.). Existing school locations and estimated number of students per school must be provided by the Ministry of Education or local educational authorities. In areas where school locations are not known, the UNICEF Office of Innovation could be leveraged as a

[17] https://www.openstreetmap.org/

potential source of data, given their extensive body of work combining machine learning and crowdsourcing[18].

2SFCA requires supply nodes and demand nodes. Our supply nodes will be the set of existing schools while the demand nodes will be the centroids of a population density raster; each centroid corresponds to the population value of that raster.

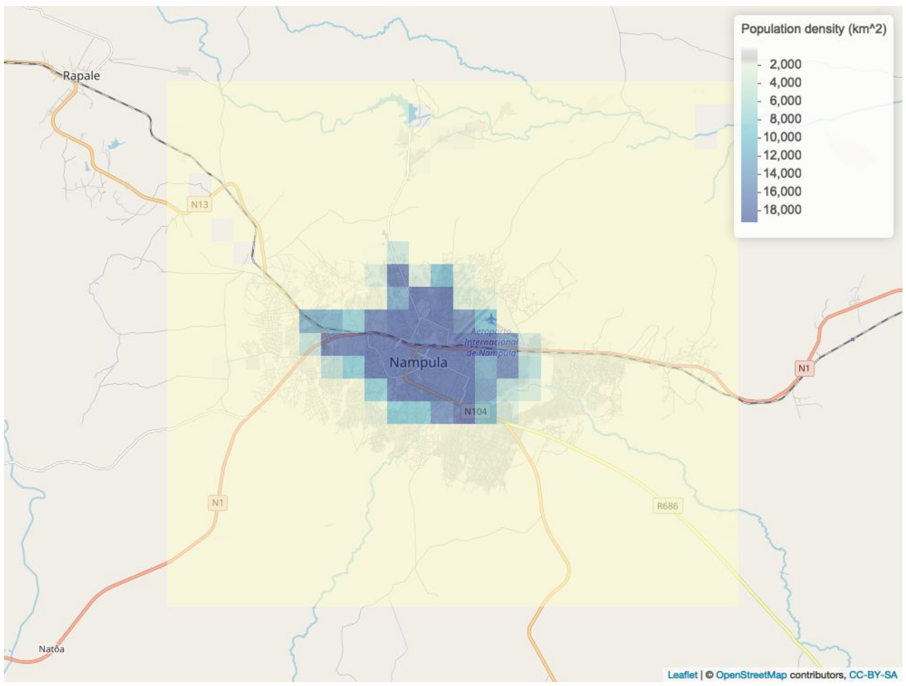


Figure 4. 1x1 km population raster derived from Gridded Population of the World[19]

[18] https://www.unicef.org/innovation/school-mapping
[19] NASA Socioeconomic Data and Applications Center. 2020 (URL)

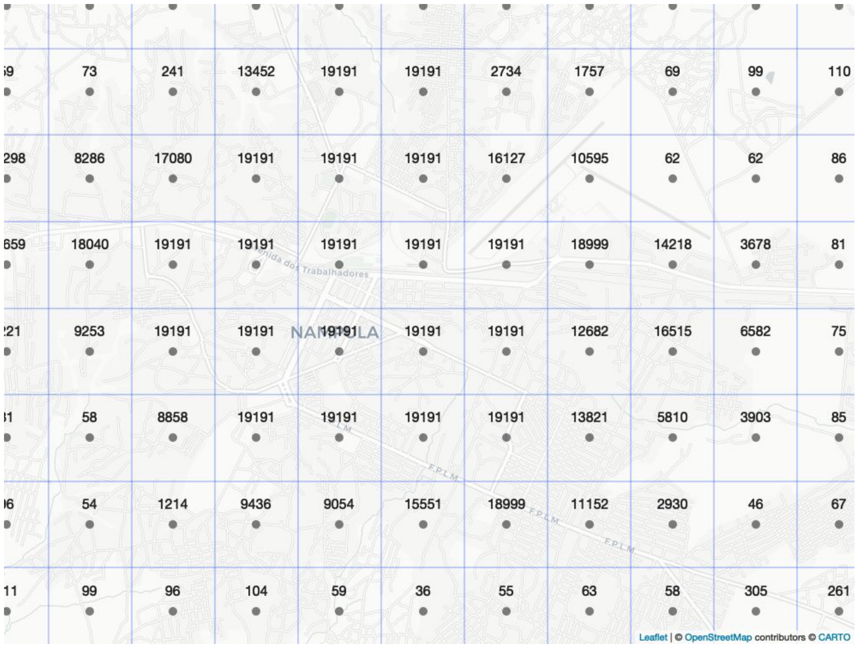


Figure 5. Demand nodes based on population values in Nampula

The method searches all demand nodes within the catchment area of each school, then calculates the supply-to-demand ratio for each facility (the average supply of schools per student). In the second step, it adds up the supply to demand ratios of all facilities located within the catchment area of the demand node. The sum of supply-to-demand ratios for each demand node is equal to the demand node's spatial accessibility score (SAS). At the end of 2SFCA, the output is a spatial accessibility score for each demand node, which will serve as an important input to the next step.

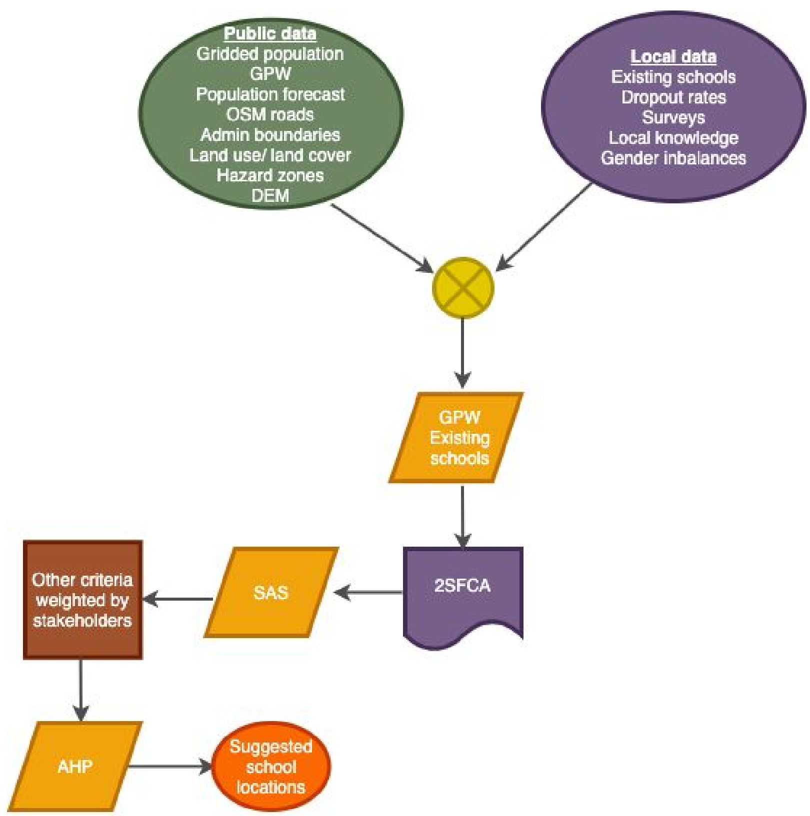
Criteria

Each demand node - or point - will also include other features joined through available datasets. These values include:

1. Spatial Accessibility Score (2SFCA)
2. Estimated school-age population*:* (float) a fraction of the population value *n* * proportion of school age children based population projections (see Figure 6)
3. Slope*:* (float) Derived from local Digital Elevation Maps (DEMs)
4. Built-up areas/settlements: (boolean) point's proximity to or within built-up area based on land cover dataset.[20]

[20] ESA LCLU 2018 (URL)

5. Farmland: (float) distance to nearest farmlands to ensure there is a food supply for students
6. Hazard areas: (float) distance to known conflict affected areas (if spatial data on violent events or flashpoints are available), or known flooding areas
7. Critical public facilities such as hospitals, police stations and parks: (float)
8. Water: (float) proximity to reservoirs, wells, or other sources of water
9. Survey data (where possible): local surveys should provide qualitative data about local perceptions



AHP

Criteria should be weighted based on local needs, and therefore requires the participation of the educational planners. In some regions, for example, food insecurity means that proximity to productive farms should be weighted higher, whereas in other regions, the progression of conflict means that weighting proximity to a conflict zone should be high.

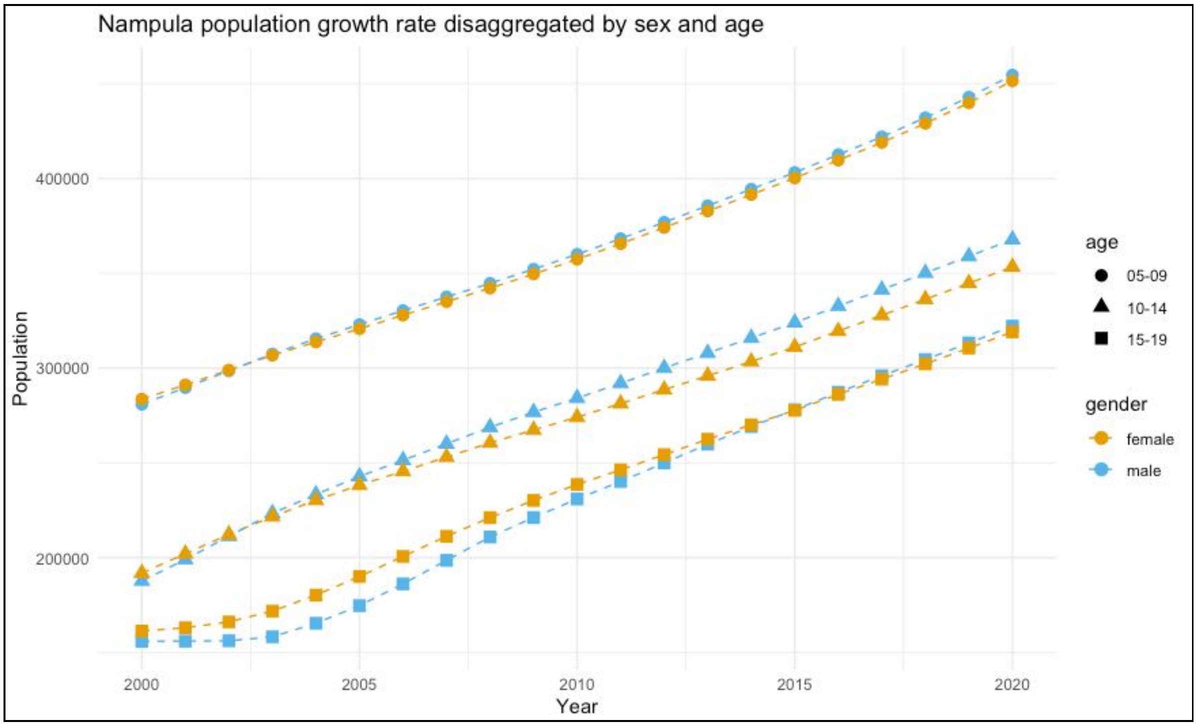


Figure 6. Population growth rate of school-aged demographic (2000-2020)

Suggestion for final locations

Once the final rasters with the number of schools are output, the final step is to suggest locations *within* raster cells (see Figure 7.) Using OSM road networks and the polygonize tool, city blocks were created, which can be used as a proposal for final school locations.

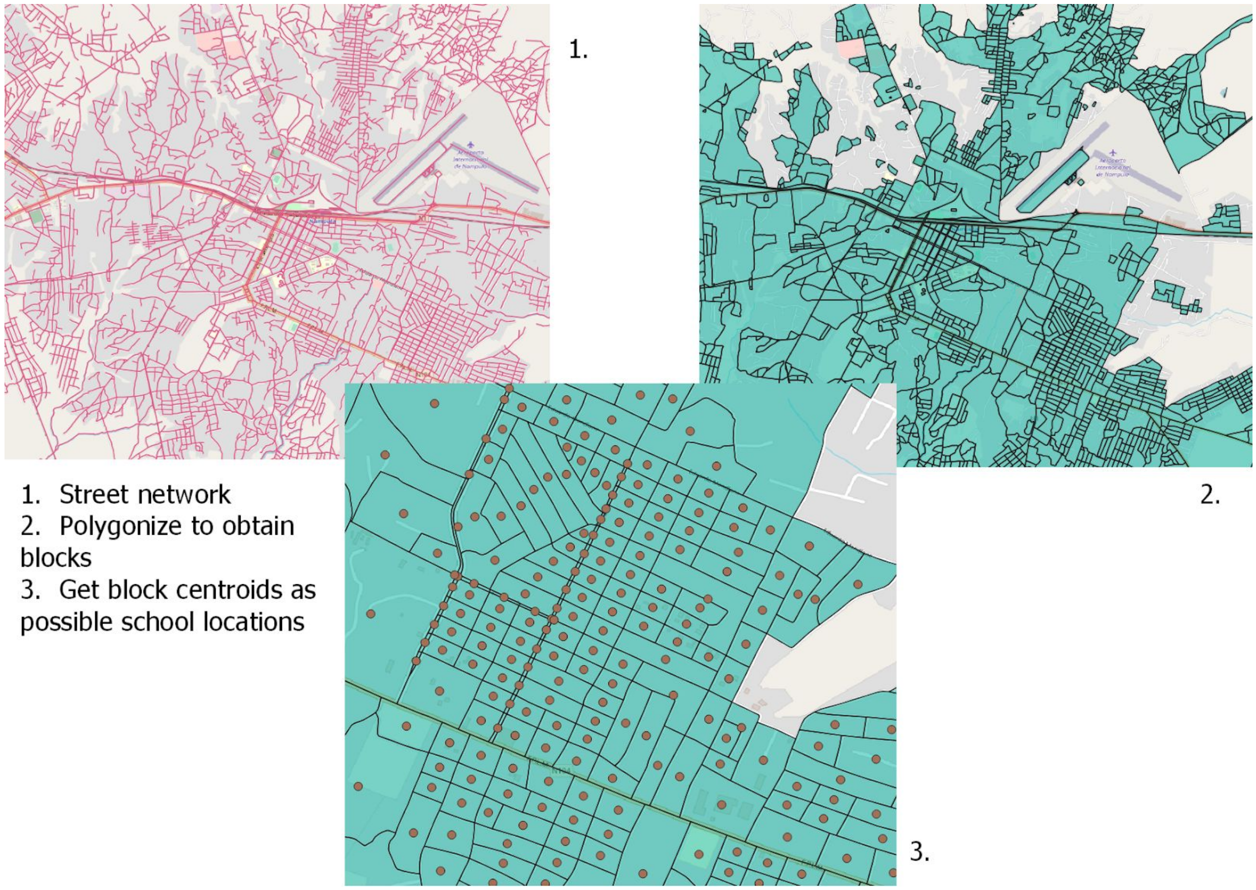


Figure 7. Suggestion for final school location based on one(or more) street block centroids