

Alvarez, Moises Lenyn Obaco; Royuela, Vicente; Vítores, Xavier

Conference Paper

Computing functional urban areas using a hierarchical travel time approach

56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Alvarez, Moises Lenyn Obaco; Royuela, Vicente; Vítores, Xavier (2016) : Computing functional urban areas using a hierarchical travel time approach, 56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/174640>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Computing Functional Urban Areas Using a Hierarchical Travel Time Approach: An Applied Case to Ecuador

[Last Version: July/2016]

Obaco A. Moisés, Vicente Royuela and Vítores Xavier

Abstract

This paper presents an applied case of how to shape integrated cities in developing countries, where usually there is no commuting data. We follow OECD's definition of Functional Urban Areas and we consider Ecuador as a case study. We use satellite imagery to overcome the problem of suitable administrative data. We identify urban cores by means of satellite-derived density and we use Open Street Maps to compute isochrones, which are used to build polycentric cities and to define the urban hinterland. We compared the results of our procedure with the ones resulting from using migration flows and from the commuting flows derived from others models. This procedure is close to standard techniques and arises as a good alternative for defining real cities in developing countries.

Keywords: Functional Urban Areas. Integrated cities. Travel time. Developing countries. Ecuador.

JEL: R12: R14: R52

Obaco A. Moisés and Vicente Royuela are at AQR-IREA, University of Barcelona, Barcelona, 08034, Spain. Telf: (+34) 934031907

Email: mobacoal7@ub.edu, vroyuela@ub.edu;

Vítores Xavier is an Independent Statistical Researcher.

E-mail: xavier.vs78@gmail.com

1. Introduction

Metropolitan agglomerations are generated by the urban and functional expansion of cities beyond their administrative boundaries. These areas concentrate most of the population and economic activity, and are the engines of their respective regions and countries.

There is a large diversity of names for such urban spaces (urban zones, conurbations, urban regions, large urban areas, metropolis, etc.) which illustrates the complexity of the phenomenon. Despite many country-specific definitions of urban areas, until recently there was no harmonised definition of a city.

One of the more ambitious developments in this regard has been the OECD cross-country analysis of cities. Together with the European Commission, they developed a new definition of a city and its commuting zone in 2011 under the label of Functional Urban Areas (FUAs). These increase international comparability and help to collect statistical data.

The methodology identified 1,179 urban areas of different size in 29 OECD countries, which gave as a further result the OECD metropolitan dataset, which considers 275 urban areas with 500,000 population and more¹.

The method applied by OECD is grounded on the use of population density to identify urban cores and of commuting flows to identify policentricity and urban hinterlands. The latter data is available in most (if not all) developed countries, but this is usually not the case in developing countries. Consequently, some additional work is needed in order to generalise such methodology to the rest of the world. Our paper takes the witness and approaches the OECD definition of urban structures in a country where there is a lack or that has poor socio-economic links to be used to connect cities. We do this by looking at geospatial information between spatial units, taking advantage of LandScan, Google Maps and Open Street Maps databases. LandScan stores information about the density of a country in grid cells of one squared kilometre. Google and Open Street Maps have information on the road system and cities.

We use these databases to identify urban settlements and define potential social-economic interactions, by assuming that urban locations follow a hierarchical pattern in space, being close enough in terms of travel time to have a potential socio-economic interaction. Our approach is a feasible and robust solution that can be applied in most developing countries.

¹ The data base is publicly available at <http://measuringurban.oecd.org/>

We test our approach in Ecuador, a ‘standard’ developing country in which there is no commuting data but there are several statistical sources allowing for completing robustness checks of the methodology. We use internal migration flows and commuting patterns derived from the gravity and radiation model. We show that by applying our proximity-based approach using travel time, we obtain results comparable to the ones resulting from alternative data.

The rest of the paper is structured as follows. Section 2 presents the background of the study. Section 3 shows the methodology while section 4 introduces the case of study and the used data. The results are displayed in section 5. Section 6 concludes summarizing the main outcomes of our work.

2. Functional Urban Areas

Administrative regions are “the expression of a political will: their limits are fixed according to the tasks allocated to the territorial communities, according to the sizes of population necessary to carry out these tasks efficiently and economically, and according to historical, cultural and other factors” (Eurostat, 1999, p.7). Even though they are not spatially random units, they are not the best spatial units to perform socio-economic analysis. One way to overcome the problems associated with administrative units is the identification and modification of political divisions in order to shape them in an existing social-economic relationship (Cörvers et al., 2009; Frey & Speare, 1992; Karlsson & Olsson, 2006). In this line, an FUA can be understood as the harmonized economic definition of “city”: a functional economic unit (OECD, 2013). It has preference over the political definitions when we aim at analysing, designing or considering urban policies, although this creates tensions and causes planning problems, since several local governments are responsible for planning, which calls for cooperation between agents within an integrated city.

We have to understand that urban agglomerations are the result of urbanisation processes, including the transformation of land cover and land use to categorize an area from non-developed to being developed (Pham et al., 2011; Weber, 2000). An urbanised space can be characterised by its population density or population size. Nevertheless, an urbanised area is not only dense, but also integrated. Connected urban zones define the new boundaries (also known as hinterland or fringe). The connection can be defined by considering many alternatives, the most common being daily interactions in the labour market (Casado-Díaz & Coombes, 2011; Duranton, 2015; Feria et al., 2015; Flórez-Revuelta et al., 2008; Klapka & Tonev, 2013; Smart, 1974).² Following Casado-Díaz &

² There are alternative approaches considered in the literature for considering spatial interactions, including services (Green, 1950), land prices (Bode, 2008), migration flows (Royuela & Vargas,

Coombes (2011), an FUA should meet some relevant characteristics such as partition, contiguity, autonomy, homogeneity, coherence and flexibility.

The process of delimiting of FUAs is usually applied in three identification steps (OECD, 2012, 2013; Puderer, 2008). Firstly, urban cores are identified, according to some density measure. All areas above some minimum threshold of population density are then characterised as potential urban cores. Such threshold may vary for every country. The OECD applied a threshold of 1,500 inhabitants for km², a sill that was lowered up to 1,000 for km² for US and Canada.³ Land cover using satellite imagery has been widely used in this identification step. Nowadays this information is available and easy to gather for most countries in the world (some recent examples of its use are Ferreira et al., 2010; Gisbert & Marti, 2014; Herold et al., 2003; OECD, 2013; Weng, 2012). However, the quality of such data will depend of the quality of the satellite images and the further recognition of density.

In this first step, a second condition must be fulfilled: areas need to contain a minimum of population size to be considered as an urban core. These minimum thresholds are established by the OECD at 50,000 inhabitants for Europe, US, Chile and Canada and 100,000 for Japan, Korea and Mexico. In addition, urban cores stay geographically fixed over time. It means that the urban cores identified now are the same as they were decades ago, and they are going to be the same in the next decades (Pham et al., 2011). As geographic areas usually do not coincide with administrative areas, some criteria are needed at this stage. According to this methodology, a municipality is part of an urban core if the majority of its population lives within the urban cluster (at least 50%).

The second identification step connects areas that may not be contiguous but that belong to the same integrated space. This way FUAs account for polycentric structures. Two areas are associated if they show some amount of connectivity. The OECD uses labour commuting data and poses that two urban cores are integrated and belong to the same FUA, if more than 15% of the residence population of any of the cores commutes to work in the other core.

The third and final step of the methodology defines the worker catchment area, this is, the area of influence of the urban cores, again considering such influence in terms of connectivity, materialised in labour commuting. The OECD defines this hinterland as all municipalities with at least 15% of their employed residents working in a certain urban

2009), person-environment interactions (Murray et al., 2005; Van de Voorde 2011), quality of life aspects (Royuela et al., 2009), transportation or even, the vegetation-impervious surface-soil where urban land use can be modeled by the percentage of vegetation, surface and soil to understand urban interactions (Wu, 2004).

³ Recently (OECD, 2015) applied such methodology for identifying Chinese cities and lowered the threshold to 550 inhabitants for km².

core. A strong condition of contiguity is finally applied, by including municipalities surrounded by a single functional urban area and dropping non-contiguous municipalities.

In the developing world, the scarcity of data is a huge problem for developing a suitable identification of these existing relationships in space. In turn, it becomes a very difficult task to carry out any kind of analysis related to urban policies, planning or socio-economic analysis. Hence, this part of the world is hidden in most applied socio-economic analysis. Coombes (2004) proposed some alternative approaches to the use of commuting data to integrate urban system. He proposed the use of internal migration flows, concentration indexes or cluster analysis. Internal migration requires again a good range of data and it presents some problems. The biggest one is the fact migration not only takes place within urban areas, what can be interpreted as a substitute of commuting, but also between them⁴. Concentration indexes require again so much detailed information that in general is not available. Finally, cluster analyses do not consider integration links, which makes it a poor proxy.

In order to overcome the lack of commuting data, the gravity approach is a common option in territorial studies, including migration and trade (Ahlfeldt & Wendland, 2016; Cohen et al., 2008; Wang & Guldmann, 1996). The simplest expression derives flows as a result of a limited amount of data, including masses of population and distance between units (Goh et al., 2012). The gravity approach has been used to study commuting patterns: Vries et al., (2009) analyze the distance-decay function for commuting, while Persyn & Torfs (2015) derive a theoretical model for commuting and use count data models to overcome the problems of zeros in a large commuting matrix.

Recently, the radiation model has been used to estimate flows such as commuting or migration. Such models appeared first in physics to study the process of energetic particles or waves travel through vacuum. The model is parameter free, which makes it suitable for fitting models when no data is available for commuting parameters in a models such as the gravitational model (Masucci et al., 2013; Simini et al., 2012).

Some authors have performed the task of identifying FUAs in developing countries. In a few cases has been commuting data available. Duranton (2015) uses commuting census to define local labour markets in Colombia. The OECD used road network availability and gradient density to identify FUAs in China (OECD, 2015). Rodrigues da Silva et al., (2014) use a cluster analysis and a road supply index and to identify FUAs in the Brazilian region of Bahia; while Gajovic (2013) uses artificial neural networks, isochrones and cluster analysis in Serbia. As some Arsanjani et al., (2014) propose new techniques for FUA

⁴ Jones (2002) and Royuela & Vargas (2009) use migration flows to define Housing Market Areas. The level of self-containment is substantially lowered compared to the one used with commuting algorithms.

identification should be: easy to apply, requiring few data, and able to predict urban boundaries precisely.

In our work, we propose to identify connectivity between urban cores and their worker catchment areas by substituting commuting data by a proximity approach based on travel time information that is derived from transportation networks. This simple alternative has been already considered in other multinational experiences, such as in the ESPON project “Study on Urban Functions”, (ESPON, 2005), where isochrones were fixed at 45 minutes to determine the boundaries. Thus, our work connects previous experiences and links them to the standardized procedure based on the OECD definition of FUA.

3. Methodology

We describe our approach in three similar steps:

1. *Identifying urban cores*: We firstly identify high population density areas. We use of satellite data reporting grid cells classified in terms of inhabitants for km². An area is categorised as high density if it is beyond a minimum threshold. We identify clusters of contiguous grid cells of high population density according to the majority rule.⁵ The resulting high-density area is required to have a minimum population size to be considered an urban core. Finally, an administrative unit, e.g. a municipality, is part of an urban core if at least 50% of its population lives within the urban cluster

2. *Connecting non-contiguous urban cores that belong to the same functional area*: As described above, two non-contiguous urban cores belong to the same FUA if they are connected. This step allows for poly-centricity in FUAs. We estimated travel time to connect urban cores and to infer if they are close enough to have social-economic interactions. Next, we introduce the assumption that urban cores follow a hierarchical pattern in space, in which some areas are playing a superior role than others. Then we apply a clustering algorithm in which we first sort each urban core using the hierarchical variable, population size. Next, we test iteratively if any urban core is within a time threshold t , defined as the travel time from centroid to centroid of each urban core. If two urban cores are within such threshold, they are clustered, such that the one with the lower hierarchy is assigned to the one with the higher hierarchy. This procedure repeats until there are no possible additional merges.

3. *Identifying the hinterlands or fringe*: We identify the worker catchment area by applying a new threshold, defined as travel time t from the centroid of each urban core to

⁵ To fill gaps in a high-density cluster we use iteratively the majority rule. Following the OECD (2013) procedure, the majority rule means that if at least five out of the eight cells surrounding a cell belong to the same high-density cluster, the lower-density cell will be added. This procedure is repeated until no more cells are merged.

surrounded political divisions that are not covered by urban cores. We follow Ahlfeldt & Wendland, (2016) and consider that the internal commuting time of a location is proportional to its area. Consequently, we derive a city-specific hinterland, related to the dimension of each urban core by means of the following formula $t_i = 2/3 \sqrt{A_i \pi} / 22,5 * 60$; where, t_i is the internal travel time of FUA i , A_i is the area of each urban core in squared kilometres and π is the constant π . It allows for the creation of homogenized areas and it is more realistic than the use of fixed thresholds for all FUAs⁶. Finally, if one area is linked to two urban cores, it will be associated to the largest FUA, as it represents the highest position in the urban hierarchy.

4. Case of study: FUAs in Ecuador

In order to apply this methodology, we use Ecuador as a case study. It is a South American country located on the equator with a total population at around 16 million of people in 2014. The total territorial extension is 283,560 Km², close Great Britain or Italy, although each of these two countries have about 60 million inhabitants. The Ecuadorean urbanization rate is around 65 percent, lower than the average of Latin America, which is around 70 percent.

Ecuador is a good case of study for several reasons. a) it is representative of many developing countries in the world in terms of population size;⁷ b) its geographic, urbanization rate and population size characteristics allow analysing changes in minimum thresholds in order to be compared with developed countries. c) there are not railways or alternative important modes of transportation other than road, what makes implies that working only with the road network is the best option; and, d) it has not been previously analysed and consequently it expands present knowledge in the applied literature.

For analysing the Ecuadorean case, we use land cover information, roadways network, and demographic information at the lowest political division, parishes, which are the closest to municipality level. The LandScan (2013)TM dataset, developed by Oak Ridge National Laboratory, provides the land cover information based on Satellite Imagery.⁸ It uses approximately 1 Km² resolution (30" x 30") and represents an ambient population (average over 24 hours). It is practically RASS information vectorized into SHP format. The roadways information comes from Google maps and Open Street databases⁹. 2013 political

⁶ We assume an average speed of 22,5 km/h, that represents half of the Standard speed in open street maps. This average is below the figure in Ahlfeldt & Wendland (2016), as they consider 30 km/h.

⁷ This is close to the average size of a country in the world once we exclude the 10 largest and 10 smallest countries.

⁸ We follow OECD, as they also used the LandScan database. For information can be accessed at <http://web.ornl.gov/sci/landscan/index.shtml>.

⁹ The OpenStreetMap database can be accessed at <http://download.geofabrik.de>

division at the parish level comes from INEC (National Institute of Statistics and Census of Ecuador).¹⁰

Ecuador has four natural regions: the coastal plain (Costa), inter-Andean central highlands (Sierra), Eastern jungle (Oriente), and the Galapagos Islands (Insular). The final Landscan dataset considers 122,544 valid grid cells of 1 km² of population density. These are mainly concentrated at Coastal plain and inter-Andean central highlands regions (see figure A.1 in the Appendix) in two specific urban poles, one located at the Coastal plain region (Guayaquil) and the other at the inter-Andean central highlands region (Quito).

In 2013, our reference period of analysis, there are 1,046 parishes in Ecuador. The mean of population density is around 120 inhabitants per km² and the median is around 35 inhabitants per km². Hence, line with other countries, the distribution of population over municipalities follows a very lumpy and concentrated distribution. In addition, they are largely heterogeneous, being especially large in Eastern jungle and Coastal plain regions, and smaller and more fragmented in inter-Andean central highlands region. The use of LandScan database allows us to identify the location of each cell inside every parish.

In order to perform further robustness analysis, we also consider the Survey of Households' Living Conditions of 2014. Even though this survey is not designed to map the commuting pattern of the whole country, it reports information of this variable for a large sample of individuals. We will use this source to report the average commuting time in Ecuador. Finally, we also use the Ecuadorean National Census of Population 2010 in order to perform additional robustness checks based of the analysis of internal migration flows patterns, and the computation of migration flows based on the gravity and radiation model.

5. Results

In order to identify the urban cores, we firstly decide the minimum thresholds of population density and minimum urban size. Our decision has to allow for capturing the maximum presence of urban settlements in the whole country, including the Eastern jungle region, which is less populated and urbanized, but has representative urban settlements.

We set the minimum threshold for density in 500 inhabitants per km², accounting for 3,699 satellite grid cells, which represent 3% of the highest percentile of density distribution of the whole Landscan dataset. This low threshold is half of the one used in

¹⁰ We have considered the use of official data. However, there are other international database with the same available information e.g. <http://www.gadm.org/country> , <http://www.diva-gis.org/gdata> and <http://www.statsilk.com/maps/download-free-shapefile-maps> .

most developed countries and allows us to identify urban settlements in most parts of the country, where otherwise, cities would be invisible to the satellite. Figure 1 displays the map for Ecuador with such cells together with a higher detail for the example of the largest city in the country, Guayaquil. Once the threshold has been set, we apply the majority rule and merge contiguous grid cell of high population to identify urban cores. Finally, we set the minimum population size at 25,000 inhabitants.

Our decisions are consistent with our objective of maximising the number of FUAs in Ecuador. We assume that any technique and any threshold are somehow arbitrary (Adams et al., 1999). Nevertheless, our decisions are not far from other experiences. ESPON (2007) uses 650 inhab./km² at the NUTS-5 level (municipalities) to identify level urban areas in Europe. OECD (2015) applies a minimum threshold of 550 inhab./km² in China. Even, there has been also considered by authorities an urban density of 400 inhab./km² (Demographia, 2015). In the same vein, the minimum size threshold is somehow flexible: Toribio (2008) argues that the typical population size to delimit a municipality as central core inside of a Metropolitan Area is 50,000 inhabitants. However, he used a minimum of population size of 100,000 inhabitants because he considered that Spain is a big country in demographic terms. The OECD used 50,000 for Europe. Instead, Gisbert & Marti (2014) followed for Spain the minimum threshold of 1,500 inhab./km² and 50,000 inhabitants to consider urban centers. Later on, we will analyse the sensitivity of our procedure to alternative thresholds.

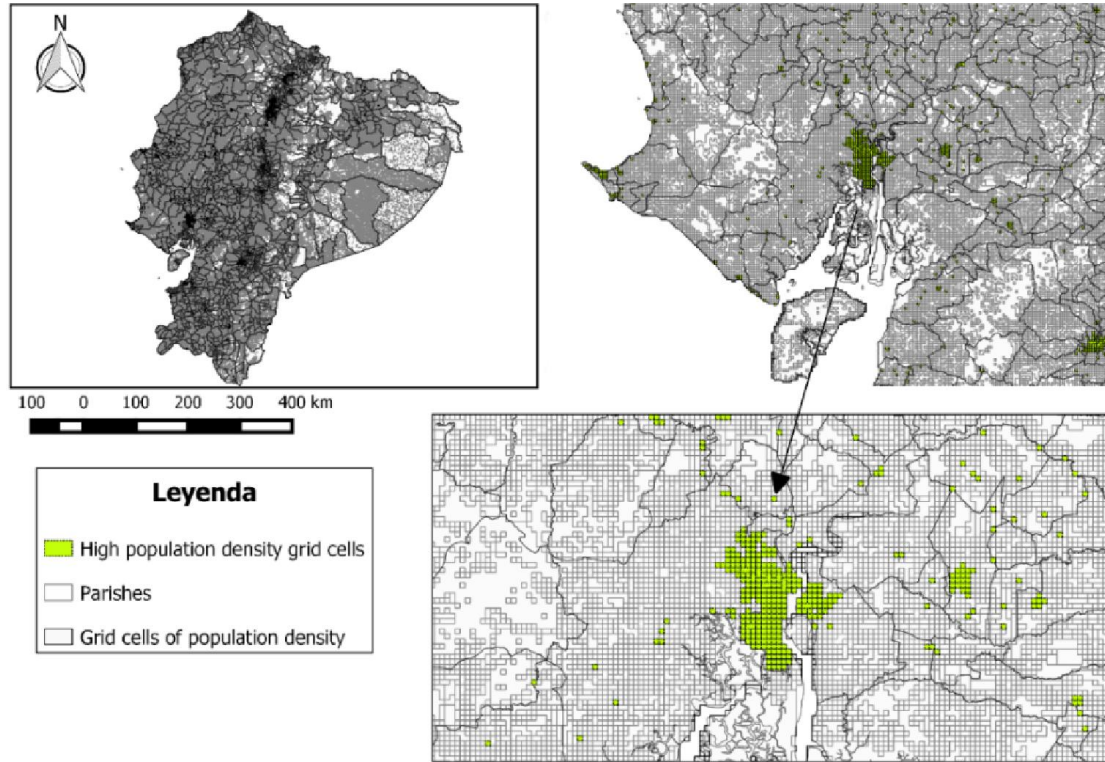
Using these thresholds, we identified 34 urban cores, which are contained around 50% of total population and around 80% of total urban population in the country. Table A.1 in the Appendix report some descriptive statistics of those urban cores. Given its specific characteristics, we treated the Galapagos Island as a special case. There the considered density threshold was 200 inhabitants for km² and a minimum population size of 10,000 inhabitants. We describe this special case in detail in the Appendix.

The second step connects non-contiguous urban cores that belong to the same functional area. Every urban core identified in the previous step is shaped into a polygon, for which we identify the centroid¹¹. We then define the isochrones by computing the time distance from centroid to centroid. In order to define the time threshold, we have analysed the 2014 Survey of Household's Living Conditions. The survey contains 109,694 observations and around 50,000 are workers. We discard people younger than 15 years old. 6,763 workers commute to another city, and 3,917 do it by bus. The time of the median commuter using the more popular transportation mode is 60 minutes by bus¹².

¹¹ One alternative is the use of the coordinates of the highest populated grid cell as the center of an urban core. We did not find significant changes by using both options.

¹² The mean of commuting by bus is 83 minutes. The median and mean of all commuters are 46 and 68 minutes respectively. The average is about one hour of travel time and it is supported with the

Figure 1. Grid cells of high population density. Detail for Guayaquil.



Next, we use the Google maps of buses service to account for travel time by road. We notice that in Ecuador Google maps does not report actual travel time by bus but the one by private car assuming roads in good conditions and fluent traffic. We believe that developing countries usually have poor quality roads, congested traffic and buses networks with improvable efficiency. We concluded that 30 minutes by private car mode reported by Google maps is equivalent to 1 hour by bus mode¹³. Once we set such threshold, we apply our algorithm based on a hierarchical travel time approach¹⁴. By applying the clustering algorithm with such thresholds, we merge four high-density urban cores and we finally identify 30 FUAs, some of them being polycentric.

Marchetti's constants that fixes the average amount of commuting time in approximately one hour (Marchetti, 1994).

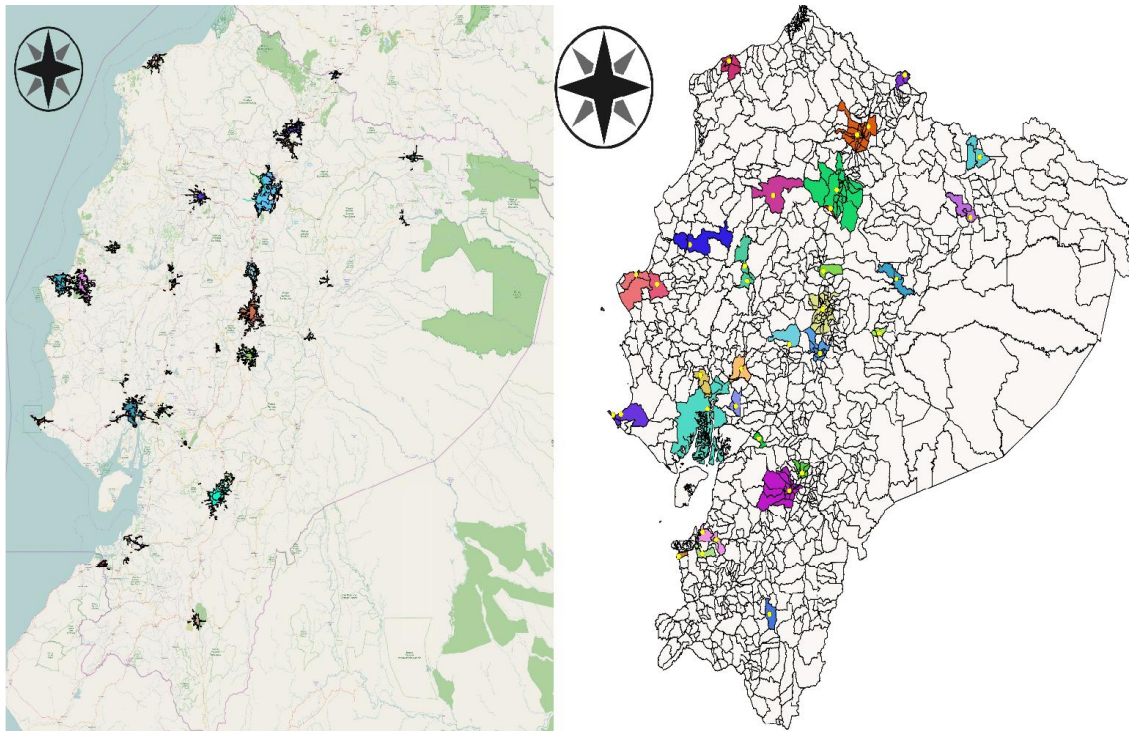
¹³ This approximation was done, using the Survey of Household's Living Conditions 2014, as an interpolation the median of time between commuters who take bus, and commuters who take private car. Additionally, as we know the origin and destination of such workers, we corroborated this by regressing informed time of commuting against travel time by car of google maps and mode of transportation (we excluded in this regression the marginal transportation modes, such as ride on animals, boats, airplanes, planes and those usual for short distances, such as walking and biking). Some descriptive statistics is presented in Table A2 and Figure A3 in the Appendix.

¹⁴ The first and third steps have been done using QGIS software. As this kind of algorithm is not available in QGIS, the second step has been programmed in Stata. It is available to control for different minimum thresholds of travel time and population size at the same time, and it is available upon request.

The final step delimitates the hinterlands of every FUA. As described above, we define city specific isochrones proportional to its size. Any municipality within both belonging to the urban core but at lower distance of the threshold is set to be part of the FUA¹⁵.

Left panel of Figure 2 shows the centroids of the urban cores and the main road structure, while the right panel displays the final shape of the FUAs (everyone associated with a different colour) in terms of administrative boundaries.

Figure 2: Results using the minimum threshold of at least 500 inhab./km², 25,000 of population size and ½ hour of travel time



5.1 Sensitivity analysis

We explore the changes in our results for alternative thresholds. Table 1 reports the number of urban cores increasing the minimum threshold for the density, minimum population size and travel time. As expected, increasing such thresholds imply a reduction in the total amount of urban cores. No definition should be preferred a priori. However, we understand that in a country where urbanization is still taking place, the identification of the maximum number of FUAs is preferred. This is the case of Amazon region, for instance, as there would not be any representative urban core if we apply the higher minimum thresholds.

¹⁵ For instance, the threshold for Quito, the urban core with the largest area, above 500 km² is set at 72 minutes and for the smallest FUA, San Jacinto de Buena Fe, with just 5 km², the threshold is set at 7 minutes. We apply the half of those values.

In addition, we have to be careful with an interesting effect that we can observe in Table 1. The highest minimum threshold of population density (1,500 inhab./km²) combined with at least 25,000 inhabitants fragmented large urban cores, and using the explained criteria to identify urban cores were treated as independent urban cores. Consequently, the lowest minimum threshold of population density was more representative of urban cores across the country than higher minimum threshold of population density.

Table 1: Sensitivity test of urban cores based on travel time

Threshold	Grid cells	Threshold	Initial	Results / FUAs (travel time)			
			Cores	1/2 h	1h	1h30	2h
500 inhab./km²	3,699 (3%)	25,000	34	30	23	16	13
		50,000	21	20	16	14	12
		100,000	16	15	13	12	11
1,000 inhab./km ²	2,114 (1.75%)	25,000	29	28	22	15	13
		50,000	20	20	16	14	12
		100,000	16	15	13	12	11
1,500 inhab./km ²	1,532 (1.25%)	25,000	33	31	22	15	14
		50,000	21	20	16	14	12
		100,000	16	15	13	12	11

6.-Robustness checks

6.1 Commuting patterns

As we report above, the Survey of Household's Living Conditions 2014 contains some information on worker's commuting patterns. However, this survey is devoted neither to commuting analysis nor to representative at local levels. Consequently, we report the application of an algorithm using these commuting flows as an additional exercise.

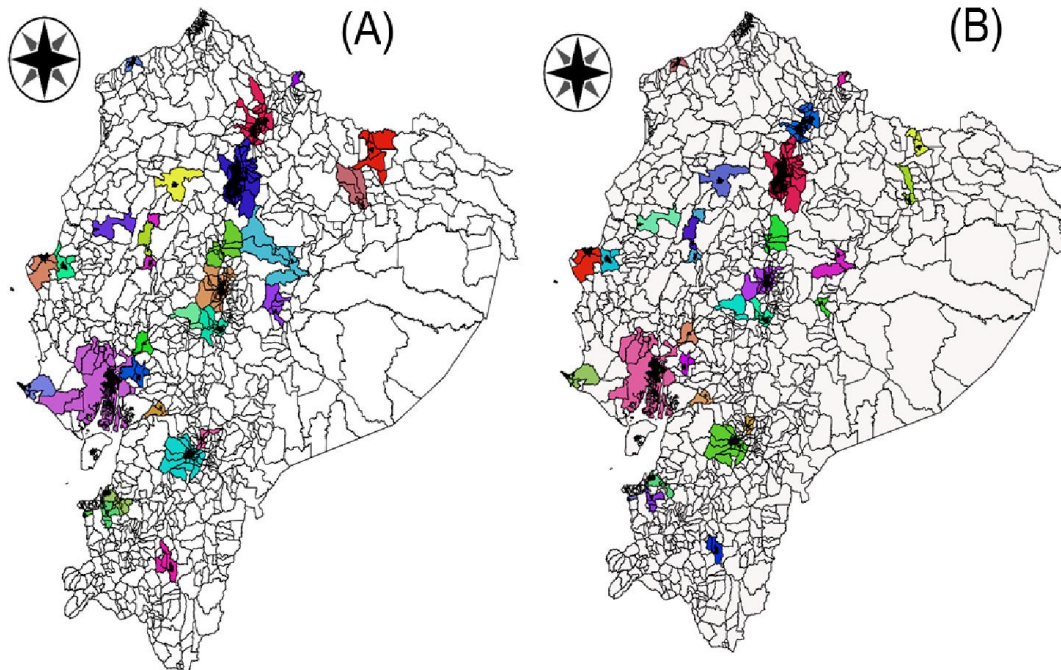
We use the information of 6,763 commuters and build a matrix of 641 parishes of origin by 540 of destination, although only 2,800 pairs of cells have non-zero observations. We consider a multi stage algorithm as in Duranton (2015), which works with percentage of commuting flow among parishes. We obtained a percentage of commuting flow from origin i to destination j , by summing up the total interviewed workers that travel from parish i to parish j , and then dividing by the total interviewed workers in parish i , $\sum W_{ij} / \sum W_i$. Table 2 shows the results of applying the algorithm between urban cores. Urban cores connected in commuting terms are exactly those that were relatively close in travel time terms. Therefore, it gives validation to our proposed based on proximity. A minimum threshold of at least 10% of commuting flow (the same as the preferred threshold for the Colombia case reported by Duranton) gives the closest approximation to our approach using travel time.

Table 2: Sensitivity test of urban cores based on commuting patterns

		Initial	Results / FUAs (% min. commuting flow)			
	Size	Cores	10%	15%	20%	25%
500 inhab./km2	25,000	34	31	32	32	32
	50,000	21	20	20	20	20
	100,000	16	16	16	16	16
1000 inhab./km2	25,000	29	27	28	28	28
	50,000	20	19	19	19	19
	100,000	16	16	16	16	16
1500 inhab./km2	25,000	33	28	29	29	30
	50,000	21	19	19	19	19
	100,000	16	16	16	16	16

Figure 3 reports the hinterland of at least 10% and 15% of commuting flow. Empty spaces between zones were covered in order to conserve contiguity criterion. Instead, we observe that in this case the hinterlands were very sensitive to the minimum threshold applied. Similar results of hinterlands, not equal, are obtained when we use a minimum threshold of at least 15% of commuting flow and half of one hour by private car of travel time. Consequently, both methodologies are similar but not identical, what can be expected given the poor quality of the commuting data.

Figure 3: Functional Urban Areas based on commuting patterns derived from the SHLC (A) 10% threshold of commuting (B) 15% threshold of commuting



6.2 Gravitational approach

We use the gravity equation under the idea of extending the commuting flow to the whole population matrix of pairs of origin and destination. Using the same survey, we forecast the total expected number of commuting flow according to whole population living in each area. In order to do that, we rescaled commuting flow, multiplying the share of commuters by population size.

Equation (1) shows the gravitational exponential decay function devoted to inter-urban mobility; where our dependent variable is the total rescaled commuting flow between origin and destination. The rescale was just obtained multiplying the share of commuters derived from the survey by population in origin. The masses in origin and destination are total economically active population (pea) or whole population (pop). Distance is measured as straight geographical distance in meters (dist)¹⁶.

$$Flow_{orig,dest} = Mass_{orig}^{\beta_1} Mass_{dest}^{\beta_2} e^{-\beta_3 Distance_{orig,dest}} \quad (1)$$

We perform a variation of this expression in order to account for the large number of zeros in our derived commuting matrix. We estimate a zero inflated negative binomial model¹⁷. The results of the gravitational equation are in Table B1 in the Appendice.

The forecasting of equation 1 was able to show a complex and decreasing in distance commuting flow pattern. We obtain a ratio of forecasted commuters from origin i over population of origin, $\sum F_{ij}/POP_i$ ¹⁸. However, the forecasted flows are not performing well, as minimum thresholds needed to be very low to have reasonable results from the algorithm.

Table 3 introduces the results of sensitivity test of urban cores. These results are similar to those presented using our travel time proposal and to the flows at the SHLC. Nevertheless, we assume that, again, this is a poor approximation of commuting flows.

¹⁶ We preferred using travel time distance because parishes were too large compared with urban settlements, and consequently Google maps or Open Street Maps were reporting incorrect estimates in too many occasions.

¹⁷ We also preformed OLS and Zero inflated Poisson estimations, and log-log, semi-log specifications. All of them led to overestimate the real size of commuters between distant areas. We opted for the ZINB by rejecting the null hypothesis in over dispersion and Young tests.

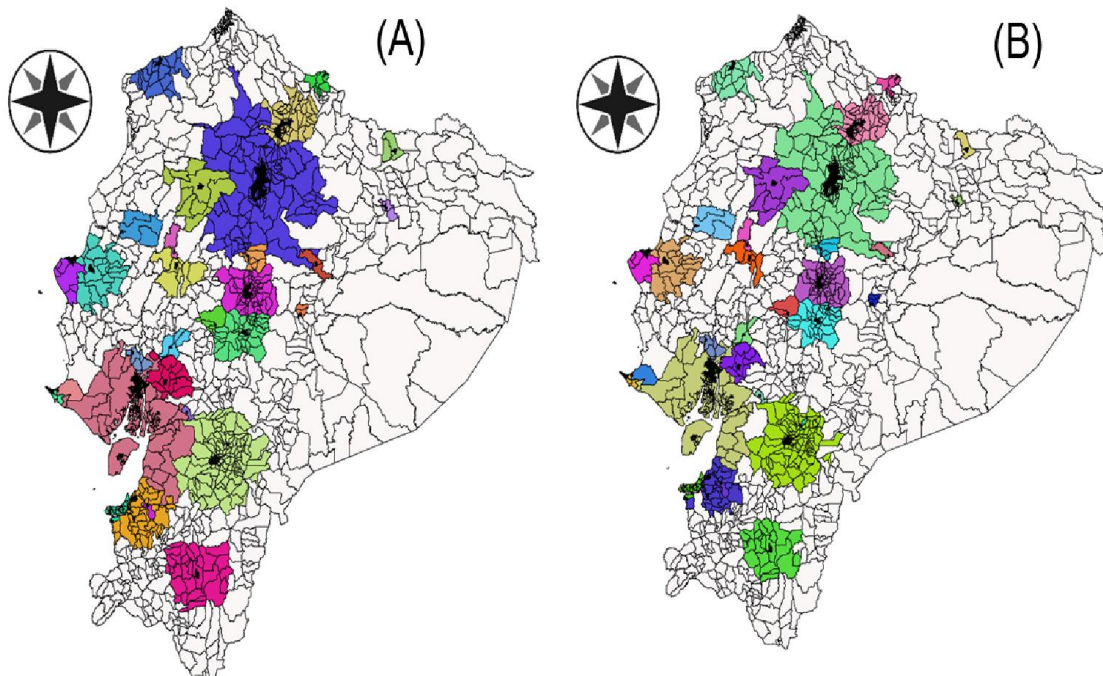
¹⁸ Some descriptive of the predicted commuting flows is in Table C1 reports in the Appendix.

Table 3. Sensitivity test of urban cores based on gravitational approach

		Initial	Results / FUAs (% min. commuting flow)			
	Size	urban cores	2%	2.5%	3%	4%
500 inhab./km ²	25,000	34	27	29	30	33
	50,000	21	20	20	20	21
	100,000	16	15	16	16	16
1000 inhab./km ²	25,000	29	24	29	29	29
	50,000	20	20	20	20	20
	100,000	16	15	16	16	16
1500 inhab./km ²	25,000	33	23	31	31	31
	50,000	21	17	20	20	20
	100,000	16	15	16	16	16

Figure 4 display the results considering hinterlands based thresholds 2.5% and 3% from commuting flows derived from the gravitational model. As we can see, the hinterlands were very sensitive to those minimum thresholds. Consequently, those boundaries are not justifiable based on the other boundaries obtained previously, probably because of the poor quality of the data.

Figure 4: Functional Urban Areas based on commuting patterns derived of the gravitational model (A) 2.5% threshold for commuting (B) 3% threshold for commuting



6.3 Radiation model approach

The radiation model for commuting is expressed in equation (2).

$$F_{ij} = F_i * \frac{Pop_i * Pop_j}{(Pop_i + w_{i,j}) (Pop_i + Pop_j + w_{i,j})} \quad (2)$$

Where;

F_{ij} is the forecasted commuters from origin i to destination j . F_i is the total outflow of commuters from origin i ; Pop_i and Pop_j are the total population in origin i and j destination respectively; and $w_{i,j}$ represents the population contained in a radius given by the distance between origin i and destination j , excluding the population contained in origin i and destination j , both. This approach does not need any parameter just the total population and the total outflow of commuters. To account for these magnitudes, we use the national census of Ecuador 2010. This census has a specific question that allows accounting for the proportion of workers commuting out of the parish. We obtained from the same census the population size. Next, we programmed an algorithm in Stata to build the matrix W_{ij} ¹⁹.

We use the forecasted commuters resulting from equation (2) as the source flow for the Duranton's algorithm. Table 5 reports the results and a sensitivity analysis for different thresholds. These outputs are pretty close to the ones derived from the travel time procedure, again at the 10% threshold of commuting.

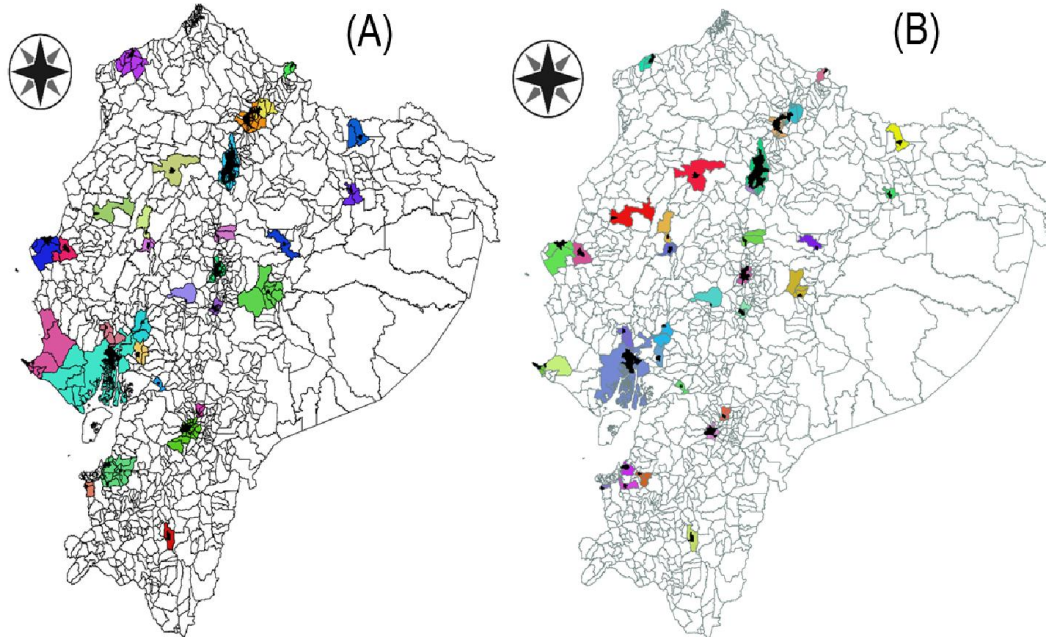
Table 5. Sensitivity test of urban cores based on radiation model

	Size	Initial	Results/FUAs (% min. commuting flow)			
		urban cores	5%	8%	10%	15%
500 inhab./km ²	25,000	34	29	31	32	34
	50,000	21	20	21	21	21
	100,000	16	15	16	16	16
1000 inhab./km ²	25,000	29	24	26	27	29
	50,000	20	19	20	20	20
	100,000	16	15	16	16	16
1500 inhab./km ²	25,000	33	27	31	32	33
	50,000	21	20	21	21	21
	100,000	16	15	16	16	16

Figure 5 displays the FUAs including the hinterlands computed using 10% and 15% of commuting flows derived from radiation model. As before, the hinterland is the most sensitive part of the analysis again.

¹⁹ As before, we use geographical distance because of the problems mentioned above.

Figure 5: Functional Urban Areas based on commuting patterns derived of the radiation model (A) 10% threshold for commuting (B) 15% threshold for commuting



6.4 Internal migration patterns

We use internal migration patterns, gathered from the national census of population 2010 of Ecuador again. There is information of internal migration between the years 2005 and 2010. The actual matrix is 1,149 parishes by 1,211 parishes, as there were several changes in the boundaries of some parishes.

We have identified large migration flows between the largest urban poles of the country. Consequently, we have opted for imposing a geographical distance restriction. This allows generating a correct identification of flows that can enter in the algorithm. We opt to use a hierarchical pattern and keep away those urban cores that are relatively far from each other. The restriction of distance was 34,765 meters, which according with Google maps is the distance by car with a half hour of travel time.

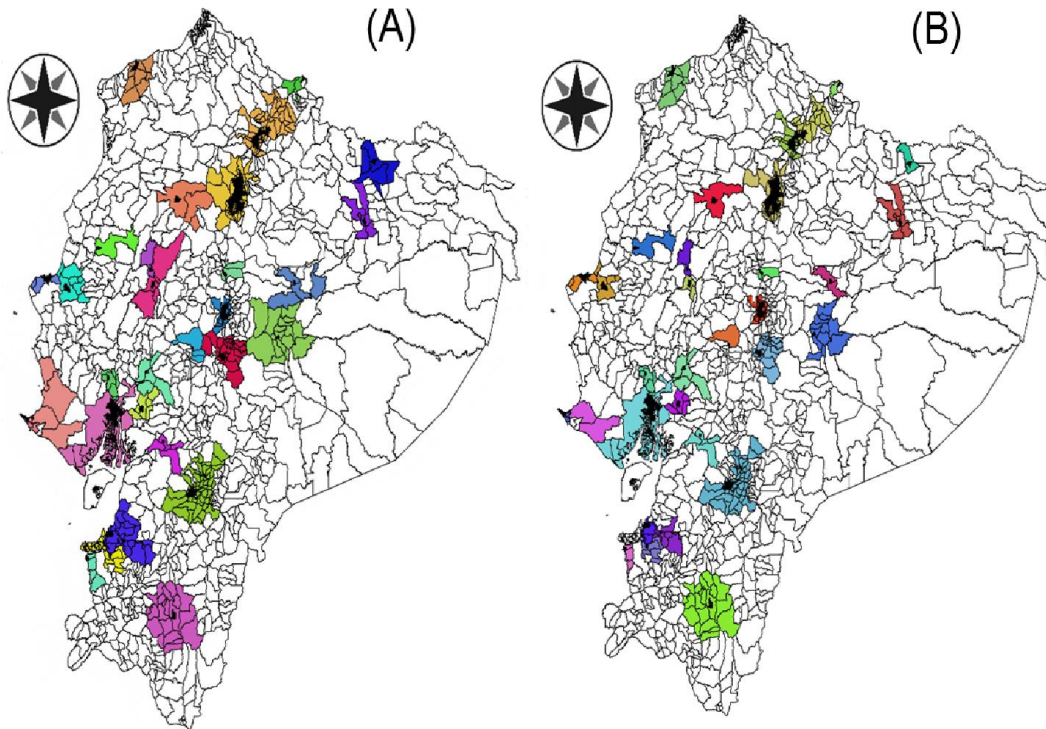
Table 6 shows the results of the algorithm for different thresholds. The algorithm was successful at connecting cities using a minimum threshold of internal migration, although the patterns are different to the results obtained from travel time and derived commuting flows. In this case, the closest approximation is obtained when using a threshold set at 15% of internal migration. As before, high minimum thresholds make the results more stable. Even if this is a good approach the results are not very similar to commuting patterns.

Table 6. Sensitivity test of urban cores based on internal migration

		Initial	Results / FUAs (% min. commuting flow)			
	Size	urban cores	10%	15%	20%	25%
500 inhab./km ²	25,000	34	27	29	33	33
	50,000	21	20	21	21	21
	100,000	16	15	16	16	16
1000 inhab./km ²	25,000	29	26	27	29	29
	50,000	20	19	20	20	20
	100,000	16	15	16	16	16
1500 inhab./km ²	25,000	33	27	29	32	32
	50,000	21	21	21	21	21
	100,000	16	15	16	16	16

We also present in Figure 6, the hinterlands of each FUA at least 15% and at least 20% of internal migration. The results are relatively similar. However, the hinterlands are also too sensitive as the others approaches introduced previously. In this case, the best approximation of the hinterland was using the minimum threshold of at least 20% of internal migration.

Figure 6: Functional Urban Areas based on migration patterns (A) 10% threshold for migration (B) 15% threshold for migration



7. Comparative analysis

We summarize the main results based on our preferred minimum thresholds for the travel time procedure: minimum density of 500 inhab./km², 25,000 inhabitants, 60 minutes of travel time (30 minutes in Google maps) to construct the urban cores. We compare this result with the ones resulting from all others methodologies to connect urban cores and define their hinterlands.

Table 7 displays the preferred results for every technique and some descriptive statistics of the population contained there in 2013. Column (1) shows the number of identified FUAs. Columns (2) to (6) present some descriptive statistics of population contained in the FUAs. Finally, Column (7) is the total population contained by the completely identified FUAs, and the percentage of total population of the country.

As the hinterlands are not very populated, the added hinterlands are practically not adding so much difference in the results, even though, the parishes could be very large in shape. On the contrary, the travel time and the commuting flows derived from the radiation model are the ones with lower standard deviation. The travel time method reports an average figure of the population living in FUAs that is the close to the estimated urbanization rate of the country in that year.

Table 7: Comparative analysis of results among all applied methodologies in terms of population contained in each FUA

	FUAs (1)	Min (2)	Max (3)	Mean (4)	Median (5)	St. Dev. (6)	TOTAL (7)
Travel time (30 minutes)	30	25,603	2,809,089	339,962	144,927	641,762	10,166,220 (64.5%)
Commuting SHLC (10 %)	31	25,603	3,015,640	354,595	184,454	679,463	10,992,445 (69.8%)
Commuting Gravitational (2.5 %)	32	42,198	3,148,618	423,357	184,932	746,583	13,547,424 (86%)
Commuting Radiation (10%)	32	25,603	2,933,851	332,330	168,060	640,351	10,302,230 (65.5%)
Migration (15 %)	29	32,510	2,869,727	364,940	190,031	650,940	10,583,260 (67.2%)

8. Conclusions

Throughout this paper, we have focused on how to identify integrated cities in developing countries when the researcher has no data on commuting flows to follow international standards. Here, we applied the OECD methodology to identify Functional Urban Areas using Ecuador as a case study. Our starting point is the use of satellite imagery to identify and connect integrated cities. Here, we use travel time in a hierarchical approach to define potential interaction between urban cores.

In addition, we test different minimum thresholds to identify cities. Low thresholds seem to better identify the largest amount of cities in a developing country that is under an urbanisation process.

We perform some robustness checks based on survey and administrative data that is available in that country. We consider commuting patterns directly derived from the SHLC, and commuting flows resulting from a gravitational approach and the radiation model. We also consider internal migration flow. All the methodologies derived similar results because of the concentration of the urban population. Our preferred options are the travel time approach and the use of commuting patterns derived from the radiation model.

As a conclusion, we identified 34 cities, 30 FUAs, two of them of metropolitan size and the remaining of small size.

Our contribution allows researchers, policy makers and planners to have a better perspective of the integrated cities in the developing world.

Acknowledgements

Many thanks to Landscan project for the database, Duranton Gilles for the algorithm and AGAUR (Agencia de Gestión de Ayudas Universitarias y de Investigación) for founding part of this paper. Vicente Royuela thanks the support of ECO2013-41022-R.

References

- Adams, J. S., VanDrasek, B. J., & Phillips, E. G. (1999). Metropolitan Area Definition in the United States. *Urban Geography*, 20(8), pp. 695–726.
- Ahlfeldt, G. M., & Wendland, N. (2016). The spatial decay in commuting probabilities: Employment potential vs. commuting gravity. *Economics Letters*, 143, pp. 125–129.
- Arsanjani, J. J., Helbich, M., & Mousivand, A. J. (2014). A Morphological Approach to Predicting Urban Expansion. *Transactions in GIS*, 18(2), pp. 219–233.
- Bode, E. (2008). Delineating Metropolitan Areas Using Land Prices *. *Journal of Regional Science*, 48(1), pp. 131–163.
- Casado-Díaz, J. M., & Coombes, M. (2011). The delination of 21st century local labour market areas: a critical review and a research agenda. *Boletín de la Asociación de Geógrafos Españoles*, (N.57).

- Cohen, J. E., Roig, M., Reuman, D. C., & GoGwilt, C. (2008). International migration beyond gravity: a statistical model for use in population projections. *Proceedings of the National Academy of Sciences of the United States of America*, 105(40), pp. 15269–15274.
- Coombes, M. (2004). Multiple Dimensions of Settlement Systems: Coping With Complexity, in: *New forms of urbanization: beyond the urban-rural*, p. Chapter 16.
- Cörvers, F., Hensen, M., & Bongaerts, D. (2009). Delimitation and Coherence of Functional and Administrative Regions. *Regional Studies*, 43(1), pp. 19–31.
- Demographia. (2015). Demographia World Urban Areas & Population Projections. *Demographia*, (January), p. 132.
- Duranton, G. (2015). A proposal to delineate metropolitan areas in Colombia. *Desarrollo y Sociedad*, 6(75), pp. 223–264.
- Espón. (2005). *Potentials for polycentric development in Europe*.
- Feria, J. M., Casado-Díaz, J. M., & Martínez-Bernabéu, L. (2015). Inside the metropolis: the articulation of Spanish metropolitan areas into local labor markets. *Urban Geography*, 3638(October), pp. 1–24.
- Ferreira, J. A., Condessa, B., Castro e Almeida, J., & Pinto, P. (2010). Urban settlements delimitation in low-density areas-An application to the municipality of Tomar (Portugal). *Landscape and Urban Planning*, 97(3), pp. 156–167.
- Flórez-Revuelta, F., Casado-Díaz, J. M., & Martínez-Bernabeu, L. (2008). An evolutionary approach to the delineation of functional areas based on travel-to-work flows. *International Journal of Automation and Computing*, 5(1), pp. 10–21.
- Frey, W. H., & Speare, A. (1992). *Metropolitan areas as functional communities a proposal for a new definition. Research Report No. 92-245*.
- Gajovic, V. (2013). Comparative analysis of different methods and obtained results for delineation of functional urban areas. *Spatium*, (29), pp. 8–15.
- Gisbert, F. J. G., & Marti, I. C. (2014). El concepto europeo de ciudad : una aplicación para España. , 30, pp. 145–156.
- Goh, S., Lee, K., Park, J. S., & Choi, M. Y. (2012). Modification of the gravity model and application to the metropolitan Seoul subway system. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 86(2), pp. 1–6.
- Green, F. H. W. (1950). Urban Hinterlands in England and Wales : An Analysis of Bus Services. *The Geographical Journal*, 116(1), pp. 64–81.
- Herold, M., Goldstein, N. C., & Clarke, K. C. (2003). The spatiotemporal form of urban growth: Measurement, analysis and modeling. *Remote Sensing of Environment*, 86(3), pp. 286–302.
- Karlsson, C., & Olsson, M. (2006). The identification of functional regions: theory, methods, and applications. *The Annals of Regional Science*, 40(1), pp. 1–18.
- Klapka, P., & Tonev, P. (2013). FUNCTIONAL REGIONS OF THE CZECH REPUBLIC : COMPARISON OF SIMPLER AND MORE ADVANCED METHODS OF REGIONAL TAXONOMY. , 44(1), pp. 45–57.
- Marchetti. (1994). Anthropological Invariants in Travel Behavior. *Technological Forecasting and Social Change*, 47, pp. 75–88.
- Masucci, a. P., Serras, J., Johansson, A., & Batty, M. (2013). Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 88(2), pp. 1–9.
- Murray, G., Judd, F., Jackson, H., Fraser, C., Komiti, A., Hodgins, G., et al. (2005). The Five Factor Model and Accessibility/Remoteness: Novel evidence for person–environment interaction. *Personality and Individual Differences*, 39(4), pp. 715–725.
- OECD. (2012). *Redefining “Urban” - A New Way to Measure Metropolitan Areas*.
- OECD. (2013). *Definition of Functional Urban Areas (FUA) for the OECD metropolitan database*.

- OECD. (2015). *OECD Urban Policy Reviews: China 2015*.
- Persyn, D., & Torfs, W. (2015). A gravity equation for commuting with an application to estimating regional border effects in Belgium. *Journal of Economic Geography*, 16(March 2015), p. lbv003–.
- Pham, H. M., Yamaguchi, Y., & Bui, T. Q. (2011). A case study on the relation between city planning and urban growth using remote sensing and spatial metrics. *Landscape and Urban Planning*, 100(3), pp. 223–230.
- Puderer, H. (2008). *Defining and Measuring Metropolitan Areas: A Comparison Between Canada and the United States Geography*. Canada.
- Rodrigues da Silva, A. N., Manzato, G. G., & Pereira, H. T. S. (2014). Defining functional urban regions in Bahia, Brazil, using roadway coverage and population density variables. *Journal of Transport Geography*, 36(2014), pp. 79–88.
- Royuela, V., Romaní, J., & Artís, M. (2009). Using quality of life criteria to define urban areas in Catalonia. *Social Indicators Research*, 90(3), pp. 419–440.
- Royuela, V., & Vargas, M. a. (2009). Defining Housing Market Areas Using Commuting and Migration Algorithms: Catalonia (Spain) as a Case Study. *Urban Studies*, 46(11), pp. 2381–2398.
- Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. *Nature*, pp. 8–12.
- Smart, M. . (1974). Labour market areas: Uses and definition. *Progress in Planning*, 2(July), pp. 239–353.
- Toribio, J. F. (2008). Un ensayo metodológico de definición de las áreas metropolitanas en España a partir de la variable residencia-trabajo. *Investigaciones Geográficas*, 46, pp. 49–68.
- Van de Voorde, T., Jacquet, W., & Canters, F. (2011). Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landscape and Urban Planning*, 102(3), pp. 143–155.
- De Vries, J. J., Nijkamp, P., & Rietveld, P. (2009). Exponential or power distance-decay for commuting? An alternative specification. *Environment and Planning A*, 41(2), pp. 461–480.
- Wang, F., & Guldman, J. M. (1996). Simulating urban population density with a gravity-based model. *Socio-Economic Planning Sciences*, 30(4), pp. 245–256.
- Weber, C. (2000). Urban agglomeration delimitation using remote sensing data, in: *Remote sensing and urban analysis: GISDATA 9*, p. 304.
- Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. *Remote Sensing of Environment*, 117(August), pp. 34–49.
- Wu, C. (2004). Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment*, 93(4), pp. 480–492.