# Spatial analysis applied to indicators of urban road network and population distribution for the identification of Functional Urban Regions in Angola

**Aldino Miguel Francisco**

PhD student in Civil and Environmental Engineering, UNESP-FEB, Brazil,
aldino.francisco@unesp.br.
Assistant Professor, Department of Environmental Engineering, UNINBE-FET, Angola,
aldino.francisco@uninbe.ao


**Gustavo Garcia Manzato**

Associate Professor, Department of Civil and Environmental Engineering, UNESP-FEB, Brazil.
gustavo.manzato@unesp.br

**ABSTRACT**

This research aimed to carry out an exploratory analysis on Functional Urban Regions (FURs) in Angola. Thus, two groups of data were used, road network and population, which were analyzed through Exploratory Spatial Data Analysis (ESDA). Road network data were extracted from the *OpenStreetMap* project and population data from the National Institute of Statistics (INE, 2020). Computational tools were used, namely: QGIS and GeoDa for the analyses. The results were presented by means of thematic maps, *box maps* and Moran index and p-value. There was greater spatial homogeneity for the Road Network Population Density (DPV) indicator, showing a good coincidence between the municipalities classified in the quadrant HH and the five most populous provinces in Angola.

**KEYWORDS:** Exploratory analysis of spatial data. Functional Urban Regions. Road infrastructure.

## 1. INTRODUCTION

In Angola, there have been several space-time transformations since the colonial and post-colonial period, followed by a long civil war, which resulted in a strong population migration in search of safer regions (such as Luanda, Benguela and Huíla). For FRANCISCO (2013), this spontaneous migration put pressure on the available infrastructure and services, which were no longer able to respond to the population demands, becoming a territorial problem. CASTRO and RESCHILIAN (2020) state that the process of metropolization of Luanda is the result of an accelerated urbanization process, intensified demographic growth, conurbations, the intensification of socioeconomic exchanges, the complexities of mobility systems based on modes of circulation and informal transport. The settlement of this population allowed for urban expansion, disordered urban agglomeration, the emergence of *musseques* (shantytowns, *ghettos*, slums), that is, many informal neighborhoods.

The loss of urban resilience in cities as a result of activities such as environmental degradation, transport, industrialization and accelerated population growth introduce new challenges to quality urban planning. The process of occupation of the territory often ends up in developing adjacent and continuous regions leading to the formation of large urban agglomerations, bringing challenges to urban planners and administrators, especially in the characterization of these large urbanized areas, here generically named (MANZATO AND RODRIGUES DA SILVA, 2010) as Functional Urban Regions (FURs).

In this context, the question about the identification of FURs has its complexity related to the fact that the definition and criteria are not uniform, as can be read in the world literature. By the term FURs, for example, metropolitan regions, urban agglomerations, conurbations, among others, are considered. Levels of industrialization, demographic indexes, road infrastructure, economic system, development models, among other criteria indicate that the FURs can be a complex urban system, leading to the idea that the urban phenomenon does not occur in the same way around the world in developed, developing and underdeveloped countries.

According to HAISH and MULLER (2013) urban or metropolitan regions have become increasingly important as nodes in the economy, observing the exchange for products, people and knowledge. They are usually the growth poles of their countries, where the main economic and social changes begin. Metropolitan areas are, by definition, attractive: they attract economic activities necessary to produce high-end exports and services.

There has been a growth in the treatment of population data to measure the dynamics of urban expansion, as summarized by AGUIAR ET AL. (2020). In another example, taking the city of Milan (Italy) as a case study, GUASTELLA and PAREGLIO (2016) used population information for spatial clusters of areas of high population density. In view of the delimitation of

concentrated urban spaces, other authors have combined population data with road infrastructure data (MANZATO AND RODRIGUES DA SILVA, 2010; PEREIRA AND RODRIGUES DA SILVA, 2010; AJAUSKAS, MANZATO AND RODRIGUES DA SILVA, 2012; RODRIGUES DA SILVA, MANZATO AND PEREIRA, 2014; OLIVEIRA JUNIOR, MATIOLLI AND MANZATO, 2017).

It is also noteworthy that most of the studies involving the theme use resources from Geographic Information Systems (GIS). For VALENTIM (2008), the growth and evolution of GIS introduced the possibility of its use as a tool to aid spatial analysis. Regarding the use of GIS to define FURs, NEVES DA SILVA (2011) explains that the use of Remote Detention, through analysis of satellite images, and Spatial Analysis, in particular neighborhood analysis, multicriteria analysis and modeling, have demonstrated the relevance and effectiveness of these methods to analyze the diversity of forms and the evolution of urban areas boundaries.

For MOURA and LIRA (2011) and ALMEIDA (2012), the *Exploratory Spatial Data Analysis* (ESDA) technique comprises the techniques that make it possible to visualize and describe spatial distributions, identify patterns of spatial association (spatial clusters or clusters), identify atypical observations (extreme values or *outliers*) or the existence of spatial instabilities (non-stationarity). For the same authors, the ESDA technique is based on spatial autocorrelation and can be applied when numerical attributes (observed data) are associated with spatial areas (georeferenced data). This tool generates global and local indices as results that provide a measure of spatial association, in addition to auxiliary graphs and maps.

Considering the above, the research proposes to carry out an exploratory analysis of FURs in Angola, based on combined information from population data indicators and the supply of urban road infrastructure.

## 2 METHODOLOGY

### 2.1 Case Study

The case study was the territory of the Republic of Angola. The country is located on the African continent, on the west coast of Southern Africa. The Constitution of the Republic of Angola (CRA) points to the provinces/states as first-level administrative subdivisions, which are 18, and each province is divided into smaller parts that are the municipalities, in total 164 (Figure 1), and these can divide into even smaller parts that are communes and urban districts (CRA, 2010; Law n. 18/16, Political-Administrative Division Law).

The are 518 communes which may have one or more towns and villages within them. There are 44 urban districts, which can be organized into neighborhoods, these into zones and the zones into blocks.

**Figure 1** – Political Administrative Division of Angola (provinces and municipalities)
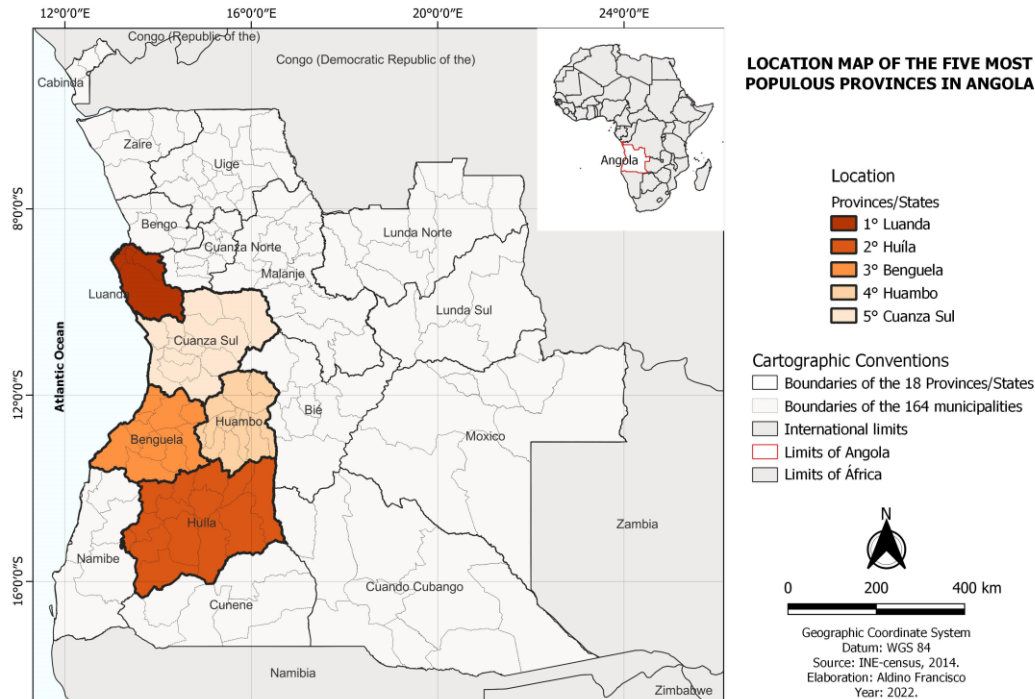


Source: AUTHORS (2022)

The country has an area of 1,246,700 km², an extension of the Atlantic Coast of 1,650 km, Land Borders of 4,837 km, it is located on the parallels 4º22' south latitude to the north and 18°02' south latitude, in a latitudinal amplitude of 13º40', and the meridians of 11º41' east longitude to the west and 24º05' longitude East to east, with a longitudinal amplitude of 12º24'. It has as bordering countries to the North: Republic of Congo and the Democratic Republic of Congo, to the East: Democratic Republic of Congo and Republic of Zambia, to the South: Republic of Namibia and West: Atlantic Ocean.

## 2.2 Indicators Explored

The premise of using this methodology is due to the need to understand that the development process of FURs in Angola, not only due to population growth, but also due to the implementation of road infrastructure. The urban road network database was extracted from the *OpenStreetMap* project and the population data from the National Institute of Statistics (INE, 2014 census; INE, 2020). According to the reference presented by MANZATO; RODRIGUES DA SILVA (2010) and LEMES; MANZATO (2020), with adaptations, five indicators were explored: 1- the extension of the urban road network by municipality; 2- the spatial density of the urban road network by municipality; 3- the population density of the road network by municipality; 4- the population density by municipality and 5- the total population by municipality.

### 2.2.1 Road extension (EV)

This indicator is one of the simplest as it is based only on road length, its formula (Equation 1) consists of the sum of the length of all roads within the area of each feature of the study region.

$$EV_X = \sum_{1}^{n} r_i; r \in x \qquad (1)$$

Where:

    *ri*:    extension of route i belonging to x;

    *x*:    study region.

### 2.2.2 Spatial Density of the Road Network (DEV)

It seeks to explain the ratio of road length by area of the municipality (Equation 2). This indicator is widely used when evaluating the supply by management institutions.

$$DEV_X = \frac{\sum_{1}^{n} r_i}{A_x}; r \in x \qquad (2)$$

Where:

    *ri*:    extension of route i belonging to x;

    *x*:    study region;

    *Ax*:    area of region x.

### 2.2.3 Population Density of the Road Network (DPV)

This indicator is similar to the DEV, when relating the road length with the population of the municipality (Equation 3).

$$DPV_X = \frac{\sum_{1}^{n} r_i}{P_x}; r \in x \qquad (3)$$

Where:

    *ri*:    extension of route i belonging to x;

    *x*:    study region;

    *Px*:    population of the region *x.*

### 2.2.4 Population density (DP)

Also presenting itself as one of the simplest indicators, the DP will be expressed by the ratio of the population contingent by area of the municipality, as seen in Equation 4.

$$DPx = \frac{P_x}{A_x} \qquad (4)$$

Wherein:

    *Ax*:    region area *x*;

    *Px:*    population of the region *x.*

### 2.2.5 Total Population by Municipality (PTM)

This is the simplest indicator, regarding to the total population in the municipality (Equation 5).

$$PTM_x = P_x \in A_x \qquad (5)$$

Wherein:

    *Ax*:    region area *x*;

    *Px:*    population of the region *x.*

**2.3 Techniques and Software Used**

Data pre-processing was done using QGIS v. 3.22.4, so that the indicators were identified and subsequently transferred to GeoDa v. 1.20 where the ESDA was developed, preparing the Box Maps. The first order of neighbors (immediate neighbors) was adopted, with "Queen Contiguity", in which a point of contact between two adjacent areas is enough to consider them contiguous. Finally, the Box Maps were created in QGIS in order to improve their representation.

In the case of ESDA, Moran's index can be applied to provide a general measure of spatial association within a dataset. In addition, it can be used to test whether the areas under analysis present a certain degree of similarity in relation to one or more reference attributes than those observed with randomly distributed values.

In the Univariate ESDA technique, two parameters are assigned to each area. The first, $Z$, is a function of the difference between the value of a given variable in an area and its global average. The second, $W_z$, is a function of the difference between the mean value of the variable in adjacent zones (those bordering the area) and the global mean of that variable. With these parameters it is possible to determine the Moran index (I), which provides a general measure of the existing spatial association in a data set, where each value represents an area. This index can be calculated in a matrix form as given by Equation 6.

$$I = \frac{(Z^t . W_z)}{Z^t . Z} \tag{6}$$

Where:

I: Spatial autocorrelation index (Moran index);

$W_z$: vector with the difference between the aveage of the variable in adjacent zones and its global average;

$Z^t$: vector of deviations, where $t$ indicates the transposed vector of the data set.

For normalized values, the Moran index varies between -1 and +1 and is interpreted as follows: values close to zero (0) indicate the absence of significant spatial autocorrelation between the values of each area and its neighbors; positive (+) values indicate that the attribute value in an area tends to be similar to the values of its neighbors, while negative values (-) indicate that the variable value in an area tends to be different from the values of its neighbors. The term "neighbors" usually refers to those bordering, that is, immediate neighbors, which border the region in question. However, it is possible to consider different levels of neighborhood (neighbors of immediate neighbors, for example).

Another way of interpreting and presenting data is the Moran scatterplot: a Cartesian plane composed of the lines Z (X-axis) and Wz (Y-axis), which configures four quadrants that allow classifying the results as follows:

➢ Q1 (or HH, for *High-High*): quadrant where the areas in which the attribute (variable) has a value greater than the global average are represented and the average value of this attribute in adjacent areas is also found higher than the global average;

➢ Q2 (or LL, for *Low-Low*): in this quadrant, unlike Q1, the areas in which the attribute has a value lower than the global average are represented, as well as the average value of this attribute in the adjacent areas, which is also lower than the global average;

➢ Q3 (or LH, for *Low-High*): here are the areas in which the attribute has a value lower than the global average, however the average value of this attribute in adjacent areas is higher than the global average;

➢ Q4 (or HL, for *High-Low*): finally, in this quadrant are represented the areas in which the attribute has a value higher than the global average, but in the adjacent areas its average value is lower than the overall average.

The data resulting from the Moran scatter plot can also be represented in thematic maps, called *Box Maps*, in which the areas are classified in one of the quadrants Q1 (HH), Q2 (LL), Q3 (LH) or Q4 (HL) and visualized over the territory, assigning a color to each quadrant, as shown in Figure 2.

**Figure 2 –** Example of Moran scatterplot and Box Map
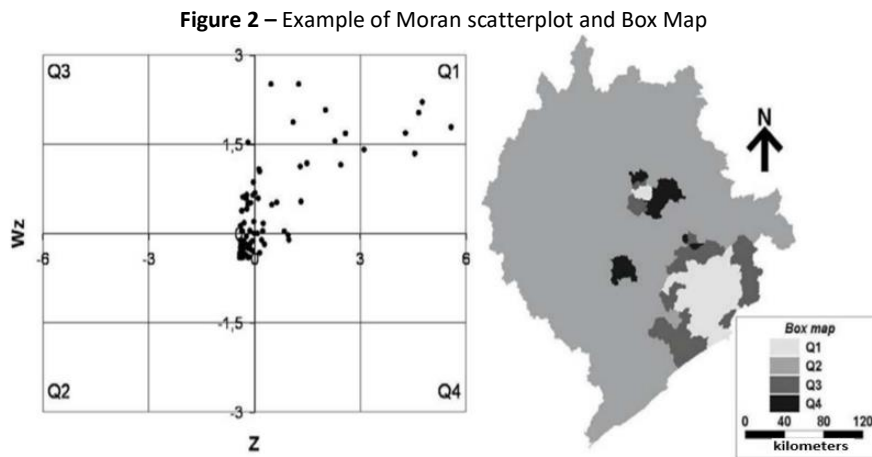


Source: MANZATO (2007)

## 3 RESULTS

Table 1 presents the results obtained from the Moran Global Univariate Index of the five studied indicators EV, DEV, DPV, DP and PTM. It was observed that the DPV indicator resulted in the highest value for the Moran index, 0.538, while the DEV resulted in the lowest value.

Table 1 - Moran index for the studied indicators

| Indicators | Global Univariate Moran Index | pseudo p-value |
|---|---|---|
| EV | 0.352 | 0,001 |
| DEV | 0.073 | 0,002 |
| DPV | 0.538 | 0,001 |
| DP | 0.230 | 0,003 |
| PTM | 0.373 | 0,002 |

Source: Own elaboration (2022)

Figures 3 to 7 present, respectively, the *Box Maps* obtained for the EV, DEV, DPV, DP and PTM indicators. When analyzing Figures 3 and 4, on the EV and DEV indicators, which tested

the road supply, the EV indicator showed numerous municipalities classified as HH (45 in total), with some significance in the FURs, while for DEV most municipalities are classified in the LL quadrant, with the exception of the Metropolitan Region of Luanda (RML) which obtained HH and LH. It became clear that the road supply alone was not enough to conclude that there is a high correlation, if we consider the FURs of the five most populous provinces in Angola.

A similar behavior is observed in the PTM and DP indicators, which test the population variable (Figures 6 and 7). The PTM obtained some municipalities classified as HH and HL, while the DP shows a large part of the municipalities that were classified as LL. Only RML obtained the classifications in the HH, HL and LH quadrants. The indicators (EV, DEV, PTM and DP) also had low values of the global Moran index, demonstrating low spatial autocorrelation.

The DPV indicator jointly tested the road supply and the population, which presented the best spatial coincidence with the five most populous provinces, as shown in Figure 5. There is, in fact, a good relationship of the municipalities classified in the quadrant HH and some neighboring municipalities in the quadrants HH, HL and LH located within or close to the FURs of the municipalities in the five highlighted provinces. Since the behavior of the DPV indicator is not random, as it is explained with the Moran Index and the p-value, I=0.538 and p=0.001, respectively. This indicator consistently highlights some of the main municipalities that make up the RML, the main FUR in Angola.
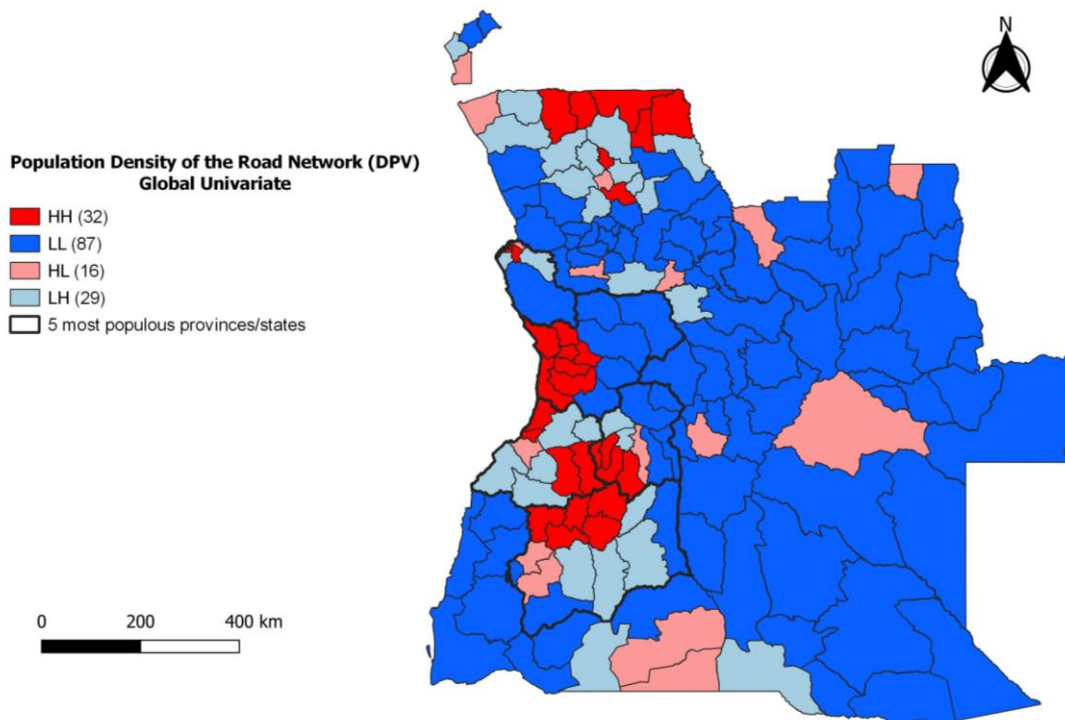
**Figure 3 -** *Box Map* referring to the Road Extension (EV)
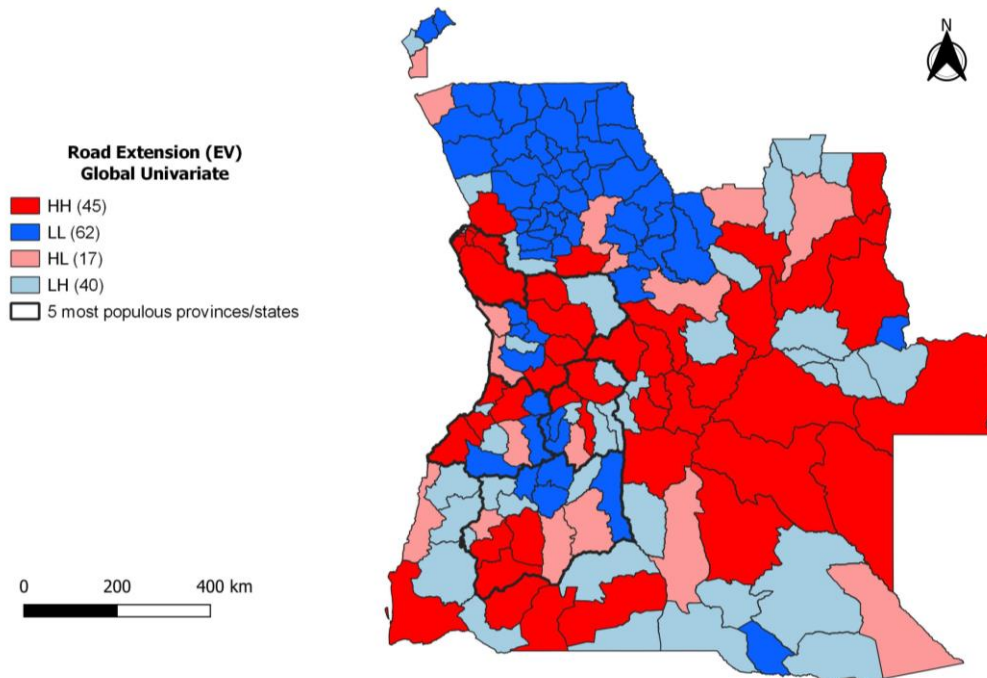


Source: Own elaboration (2022)

**Figura 4 –** *Box Map* referring to the Spatial Density of the Road Network (DEV)
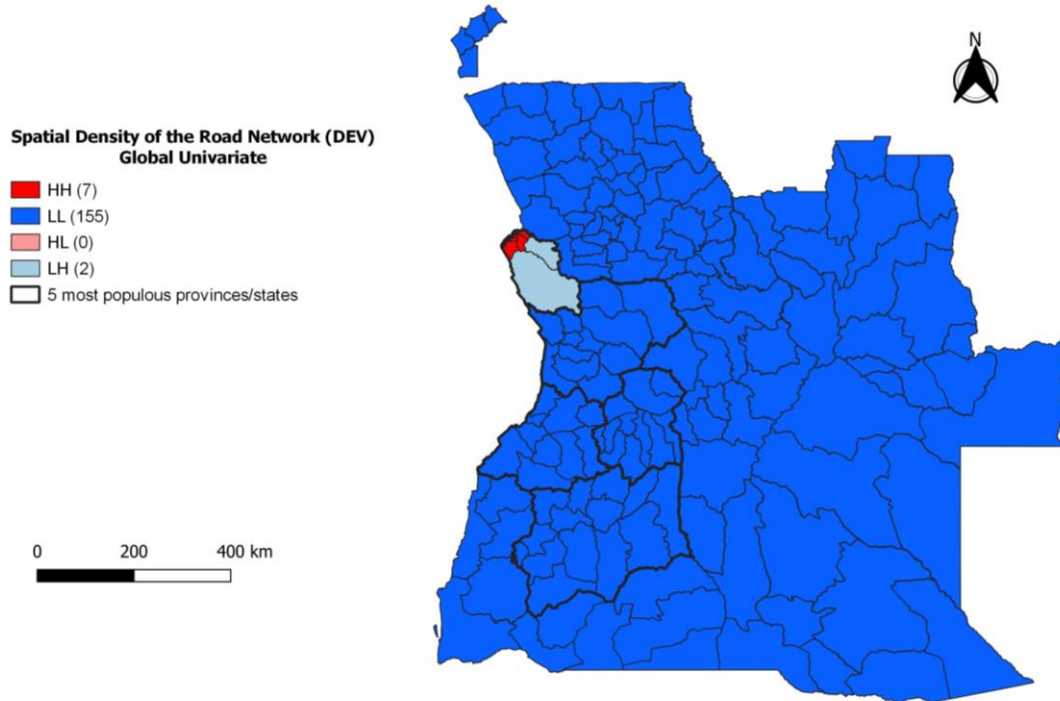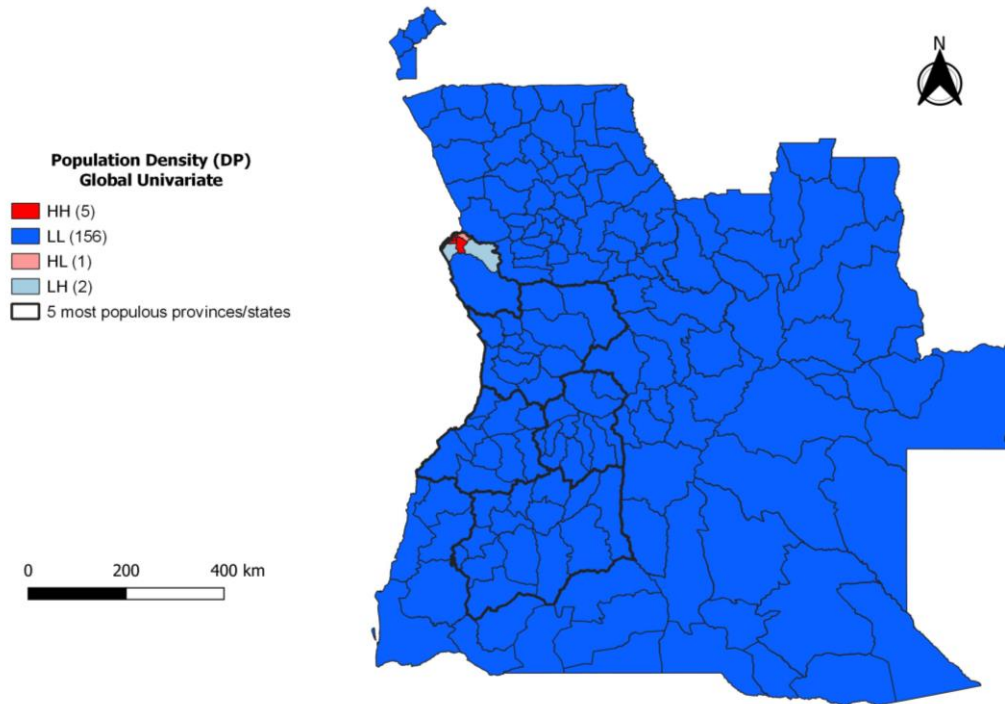


Source: Own elaboration (2022)

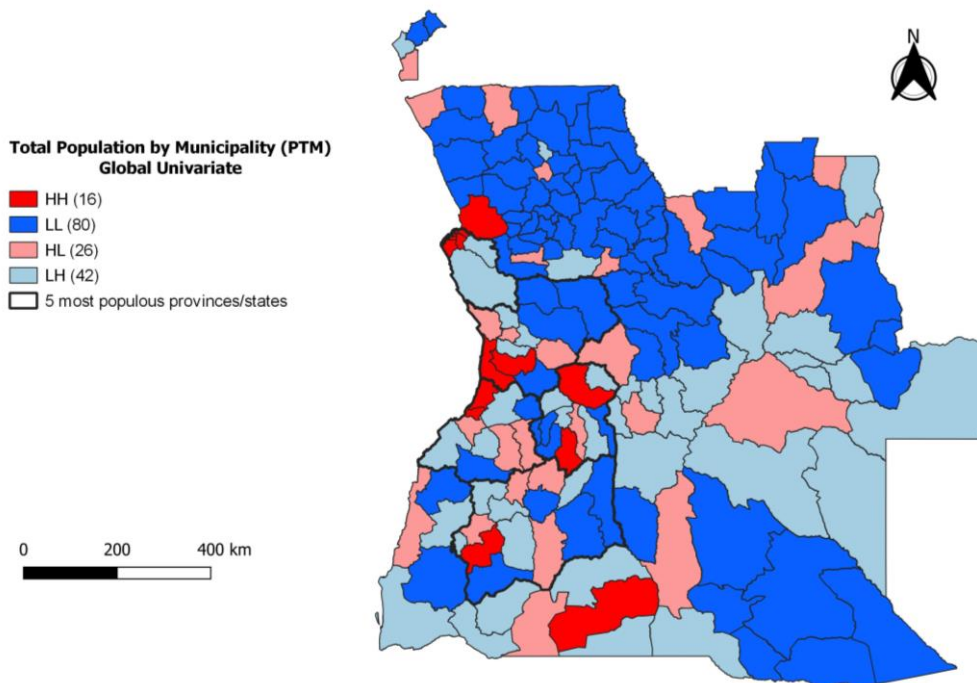**Figura 5 –** *Box Map* referring to the Population Density of the Road Network (DPV)



Source: Own elaboration (2022)

**Figura 6 –** *Box Map* referring to Population Density (DP)



Source: Own elaboration (2022)

**Figura 7 –** *Box Map* referring to the Total Population by Municipalities (PTM)



Source: Own elaboration (2022)

## 4 FINAL CONSIDERATIONS

The objective of the work was to carry out an exploratory analysis of FURs in Angola, based on combined information from population data and the supply of urban road infrastructure in the 164 municipalities of Angola. We applied the Global Univariate ESDA

technique, highlighting the most populous provinces (Luanda, Benguela, Huíla Huambo and Cuanza Sul).

The *Box Map* results showed that there was greater spatial homogeneity of the DPV indicator, with this indicator presenting the best results in the classification of the HH quadrant, showing coincidence with the municipalities of the five most populous provinces in Angola. The RML was consistent across all HH-rated indicators. This is due to the fact that the FUR of Luanda is the main region of the country, it is the largest urban agglomeration (with the largest population and demographic density), the largest industrial, commerce and employment center.

The next step to continue this research will be the application of the Bivariate ESDA technique, providing new approaches to the study. The research has its relevance for being a pioneer in the study of African FURs, through a methodology based on easily obtainable data and tools available in free software, with good accessibility of technical appropriation of the procedures.

**Acknowledgment**

**5 REFERENCES**

AGUIAR, L. L.; G. G. MANZATO E A. N. RODRIGUES DA SILVA. Combining travel and population data through a bivariate spatial analysis to define Functional Urban Regions. **Journal of Transport Geography**, Elsevier, v. 82, n. 102565. https://doi.org/10.1016/j.jtrangeo.2019.102565, jan. 2020.

AJAUSKAS, R.; G. G. MANZATO AND A. N. RODRIGUES DA SILVA. The Definition of Functional Urban Regions: Validation of a Set of Spatial Models with Recent Census Data and Analysis of an Additional Model Specification. **Proceedings of CAMUSS, the International Symposium on Cellular Automata Modeling for Urban Spatial Systems**, p. 91-104. Porto, Portugal, 2012.

ALMEIDA, E. **Econometria espacial aplicada**. Editora Alínea, 2012.

ANGOLA. [Constituição (2010)]. **Constituição da República de Angola de 2010 (CRA)**. Luanda, Assembleia Nacional, de 5 Fevereiro, 2010.

ANGOLA. Divisão Política Administrativa. Lei n.º 18/16, de 17 de outubro. **Diário da República**. Luanda. 2016.

CASTRO, J. C.; RESCHILIAN, P. R. O impacto da informalidade na (re)estruturação das metrópoles periféricas contemporâneas: o caso de Luanda, Angola. **Scripta Nova, Revista Electrónica de Geografía y Ciencias Sociales**, Universitart de Barcelona, v. XXIV, n. 639, 1 jun. 2020.

FRANCISCO, Aldino Miguel. **Desafios do desenvolvimento e do ordenamento do território na província do Namibe (Angola).** 2013. Dissertação (Mestrado em Gestão do Território/Planeamento e Ordenamento do Território) - Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa, Lisboa, 2013.

GUASTELLA, G.; PAREGLIO, S. Urban spatial structure and land use fragmentation: the case of Milan FUA. **Aestimum**, v. 69, p. 153-164, 2016.

HAISH, T.; MULLER, U. Defining and measuring urban regions: A sensitivity analysis. **Papers in Regional Science**, v. 94, n.1, p. 219-226, Doi:10.1111/pirs.12043, 26 may ,2013.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE, Censo 2014). **Projeções da população por municípios 2015 a 204**. Luanda, Angola. 2020.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE, Censo 2014). **Resultados Definitivos do Recenseamento Geral da População e da Habitação em Angola**. Luanda, Angola, mar. 2016.

LEMES, T. J; MANZATO, G. G. Rede viária urbana combinada a indicadores de oferta de transportes visando a definição de regiões urbanas homogêneas. **34º Congresso de Pesquisa e Ensino em Transporte da ANPET**. 2020.

MANZATO, Gustavo Garcia. **Regiões Urbanas Homogêneas e Oferta de Transportes.** 2007. Dissertação (Mestrado em Engenharia Civil) –  Escola de Engenharia de São Carlos, USP, São Carlos, 2007.

MANZATO, G. G.; RODRIGUES DA SILVA, A. N. Spatial-Temporal Combination of Variables for Monitoring Changes in Metropolitan Areas, **Applied Spatial Analysis and Policy**, v. 3, n. 1, p. 25-44, 2010.

MOURA, R.; LIRA, S. A. Aplicação da análise exploratória espacial na identificação de configurações territoriais. **R. bras. Est. Pop.**, Rio de Janeiro, v. 28, n. 1, p. 153-168, jan./jun. 2011.

NEVES DA SILVA, Arménio. **Dinâmica socioespacial e produção habitacional na periferia de Maputo-Moçambique a partir da década de 1970:** destaque para bairros Polana Caniço "A" E "B". 2011. Dissertação (Mestrado em Geografia) - Universidade Federal de Santa Catarina, Florianópolis, 2011.

OLIVEIRA JUNIOR, M. A.; J. A. C. MATIOLLI; G. G. MANZATO. Incorporating recent census and transportation supply data into cellular automata-based models for the delimitation of Functional Urban Regions. **Proceedings of 96th Annual Meeting of the Transportation Research Board (TRB)**. Washington D.C., USA. 2017.

PEREIRA, H. T. S. AND A. N. RODRIGUES DA SILVA. Comparing spatial analysis methods for the definition of Functional Urban Regions - The case of Bahia, Brazil. **Proceedings of 10th International Conference on Design and Decision Support Systems in Architecture and Urban Planning**, Eindhoven, The Netherlands. 2010.

RODRIGUES DA SILVA, A. N.; G. G. MANZATO AND H. T. S. PEREIRA. Defining Functional Urban Regions in Bahia, Brazil, using roadway coverage and population density variables. **Journal of Transport Geography**, v. 36, p. 79-88. DOI:10.1016/j.jtrangeo.2014.03.001. 2014.

VALENTIM, Miguel Jorge Malveiro. **Identificação de áreas urbanas por analise espacial matricial**. 2008**.** Dissertação (Mestrado em Engenharia do Território) – Instituto Superior Técnico, Universidade Técnica de Lisboa, Lisboa, 2008.