

# **Energy Development Potential: an analysis for the Brazilian states.**

Fernando Salgueiro Perobelli  
Caio César Calheiros de Oliveira

***TD. 002/2012***  
***Programa de Pos-Graduação em Economia***  
***Aplicada - FE/UFJF***

Juiz de Fora

2012

# Energy Development Potential: an analysis for the Brazilian states \*

Fernando Salgueiro Perobelli  
PhD in Economics  
Full Professor at PPGEA/FE/UFJF – Brazil  
CNPq Scholar  
[fernando.perobelli@ufjf.edu.br](mailto:fernando.perobelli@ufjf.edu.br)  
+553232371429

Caio César Calheiros de Oliveira  
Research Assistant  
CNPq Scholar

## ABSTRACT

This paper elaborates an indicator of the Energy Development Potential (EDP) for 27 Brazilian states based on the data referring to infrastructure, energy supply and demand. Using factor analysis it is possible to identify three factors that together explained 95.27% of the total variance of the model, they are: energy demand, renewable energy supply and non-renewable energy supply. From there, we classified the Brazilian states according to their indicators of EDP and performed an exploratory spatial data analysis (ESDA) through the indicators of Moran and the local indicators of spatial association (LISA). With these results we conclude that there are indicators of EDP below the average for most of the units of the Federation and identifies the formation of a high-high cluster (HH) for the states located in South-Southeast and a low-low (LL) for the states located in the region north of the country.

**Key-words:** Energy; Factor Analysis; Spatial Analysis.

**JEL CODE:** R12, R15.

## 1. Introduction

Modern economies rely heavily on energy consumption. Thus, energy is an important and vital input in the development process of a country and plays a major role in achieving economic, social and technological progress. It is expected that population growth, the level of economic activity, the sector development and industrialization process must necessarily imply the increase of energy demand and the need for an increase in the supply of this input. Therefore, diversification in the energy matrix of a region makes it less susceptible to crises and able to meet, without damage, the growing

---

\* The authors would like to thanks CNPq to financial support to development this research.

worldwide demand for this input [Fan and Xia (2011); Chen and Lin (2008)].

Junior Pinto (2007) affirm that energy is an essential input for economic and social organization of all countries, and the process of energy production and consumption have impacts on the economic and social development and the environment. Therefore, the energy occupies a prominent role in the definitions of business strategy and government policy agenda.

There are several studies dealing with the energy issue. We want to shed light to the diversity of themes in this literature. There are some studies that focus on the demand side of the problem. In this group is possible to mention (Rapanos and Polemis, 2006) that seeks to estimate the determinants of residential energy demand in Greece. The residential consumption of energy depends on characteristics of the countries and is perhaps related to their climatic conditions. This paper enables us to better understand some of the determinants of energy demand. According to the authors these determinants are: real gross domestic product (GDP); the real price of energy; and the number of heating degree days. According to the literature, the energy demand/consumption issue are also related to final demand of non-energy sectors [Kagawa and Inamura (2001); Kagawa and Inamura (2004)]; the dynamic changes in expected socio-economic structure and temperature [Hekkenberg et al., (2009)]; technology effects [Zhang, (2011)]; structural changes [Kagawa and Inamura (2001); Kagawa and Inamura (2004)]; past consumption trends [Ito et al. (2010)]; development stage [Fan and Xia (2011)]; and energy use patterns and economic structure [Kahrl, (2009)].

Besides the development of sectors with high consumption of energy, which

demonstrates the importance of economic activity on energy consumption, Junior Pinto (2007) claims that since the 1960s was already recognized a relationship of proportionality between the consumption of energy and national income and that several subsequent studies have come to prove the relevance of this variable in determining the demand for energy. Therefore, national income, economic activity, industrialization and population growth present themselves as some of the determinants of demand for energy.

In terms of energy supply in a specific region it is possible to point the following variables as relevant: variables related to the infrastructure of production, transport, storage and distribution as well as geographical factors, for example, the presence of oil reserves, natural gas, coal and river flows. Besides these, Sola *et al* (2006) add as determinants of energy supply: the actions of state and government who seek not only to funding for the sector, but also encourage alternative sources of energy, and technology, mainly through investments in Research and Development (R&D), not only in energy infrastructure, but also in the use of technology devices, which convert the final energy into useful energy.

Recently in the literature it is possible to verify that the incorporation of spatial dimension into applied economic modeling is increasing. This phenomenon is occurring for the analysis of questions where the spatial aspects make a crucial difference in the analysis and policy implications. Most of the spatial data analysis was implemented in the fields of regional economics and economic geography. Nowadays we note that the spatial models are now being used for a wide range of applied economic topics (Burnett, 2011).

In this same vein we are beginning to see an increase in spatial analysis incorporated into energy economics models. According to Burnett (2011, pp.01), *“this makes sense given that energy resources, energy consumption, and energy production are defined over time and space, and therefore have spatial dimensions. Examples include patterns of energy use across space, spatial linkages between energy and the environment, spatial spillovers in regional energy consumption, spatial clustering in fossil fuel exploration activities, spatial structuring of electricity prices, etc.*

Thus, this work is motivated by economic growth observed in Brazil, which leads to an increase in domestic demand for energy and the heterogeneous spatial structure of production and consumption of energy in its various sources. Therefore, it is necessary to evaluate the energy issue in the context of units of the Federation, *i.e.*, as a spatial phenomenon.

The fact that there are few studies that address the energy sector in Brazil in its spatial dimension increases the importance of this article and differentiates it from others. This paper proposes a study that includes different sources of energy for all states. This will contribute to the literature has been shown to be of great importance today since the recent Brazilian economic development is heterogeneous among the space and is marked by large discrepancies between their regions. Thus, it may help the discussion about differentiated policies for the energy sector among the Brazilian states.

Therefore, to contribute to a regional understanding of the phenomenon this work builds an indicator of Energy Development Potential (EDP) for the 27 states of the Brazilian Federation by analyzing the supply capacity and energy consumption of each of these

units. In addition, we seek to assess how this potential is distributed in the country.

We use in this study multivariate analysis, more precisely factor analysis, to construct the Energy Development Potential indicator of each Brazilian Federation and Exploratory Spatial Data Analysis (ESDA) by observing the Moran indicators and local indicators of spatial association (LISA), for the territorial analysis of the results.

This paper is organized as follows. The second section presents a brief characterization of the Brazilian energy structure. The third section presents the database used and describes the methodology. The fourth section presents the results and finally we make some conclusions.

## **2. Brazilian Energy Profile**

Using the National Energy Balance (BEN) published by the Ministry of Mines and Energy we can observe the increase in energy demand and supply in Brazil. The final demand for energy present an increase of approximately 73.5% between 1990 and 2009. For the same period there is a 72% increase in domestic energy supply, and the supply of non-renewable energy accounts for 55% of this growth and renewable energy for 45%.

Regarding the sources of energy used internally we can highlight: oil and its derivatives, which since 1973 is the main source of supply and demand in the country; natural gas, which was the source who revealed greater percentage growth in supply and demand from 1990, and products derived from sugar cane, which grew 156% in its demand and 134% in its supply in the period 1990-2009.

Since 1995 electricity became the second source of energy most offered in the country. The national transmission system is almost totally integrated. This allows the exchange of energy among the Brazilian regions and a better administration of the system.

Analyzing data from the regional supply of non-renewable energy in Brazil, through BRASIL (2010), we can see a predominance of the Southeast region in oil production (89.8%) and natural gas (54.7%) and the South in the production of coal (100%). In terms of renewable energy supply we see that the South and Southeast regions are responsible for 26.3% and 36.6%, respectively, of the entire country's electricity generation.

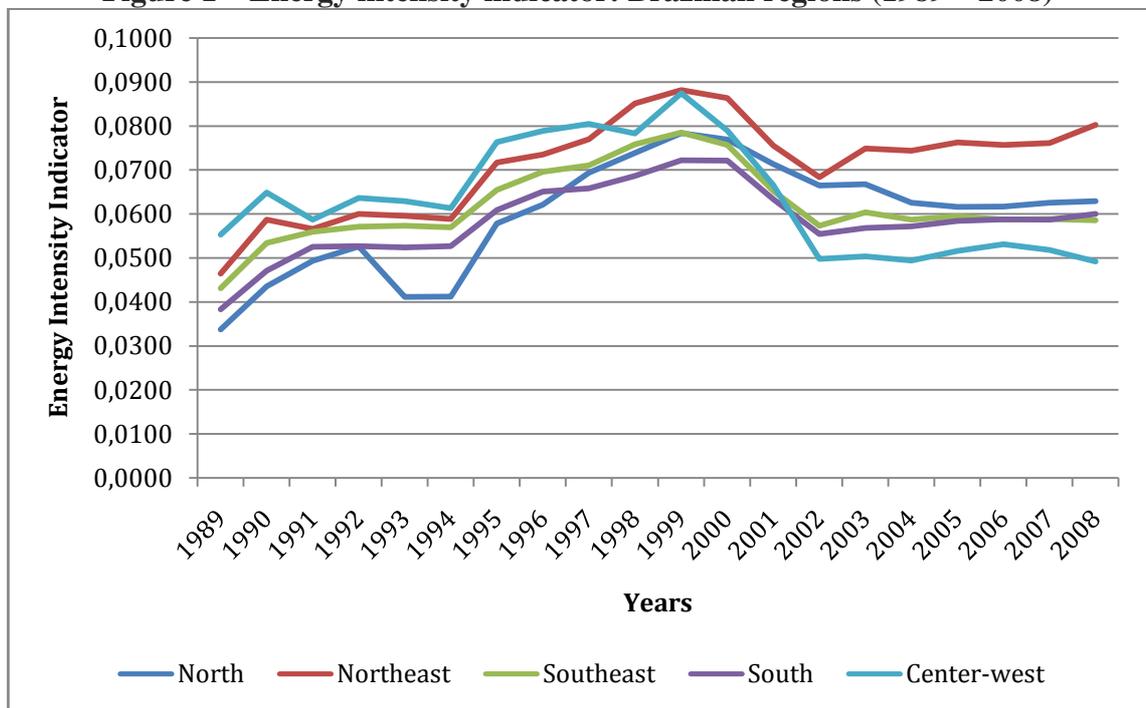
Regarding the evolution in the regional demand for energy for the period 1990-2009 there is a percentage growth in residential consumption of electricity from 170% to the North, 158.4% for the Northeast, 145.7% for the Centre West, 120.6% and 88.2% for South and Southeast, respectively. In the residential consumption of liquefied gas and oil region North had the biggest percentage increase in the period, 71.9%, followed by the Northeast (36.4%) and Centre west (30%). Despite these increases the Southeast region still has the highest energy consumption from all sources offered in the country.

When analyzing the industry demand for energy is possible to notice an intensive consumption of industrial sector and transportation sector, which in 2005-2009 had an average participation of 40% and 29%, respectively, in the total energy consumption.

In order to have a complete picture of the energy intensification of use among the Brazilian regions we calculate an energy intensity indicator for the 27 Brazilian states.

The indicator is the ratio between the residential consumption of electricity and gross domestic product (GDP). Figure 1 enables us to analyze the behavior of the energy intensity indicator. We can observe that from 1989 to 1999 there is a tendency of increase in this indicator. This occurs in the Brazilian economy and in all five regions in consideration. After 2000 we observe an inverse path, except for the Northeast. This region presents a decrease in the indicator for 2000 – 2003 and an increase in the rest of the period.

**Figure 1 – Energy intensity indicator: Brazilian regions (1989 – 2008)**



Source: Elaborate by the authors

Observing Figure 1 it is possible to shed light for the following aspect: the regions Northeast and Center-west have a greater indicator than the Brazilian economy for the period 1989 to 2000. In the next period we verify that the Northeast is more energy intense than the Brazilian economy, but the Center-west is less energy intensive. It is important to highlight that for the end of the period the Center-west region presents the smallest result for the indicator.

On the other hand is important to note the path of the indicator for the North region. At the initial periods the region is less intense than Brazil, but at the end the region became more energy intense than Brazil.

The analysis of the Figure 1 corroborates, in this aspect, – energy intensive use - the idea of spatial heterogeneity of the energy sector in Brazil and shed light to the necessity for going deep on the spatial analysis of this study.

### **3 Methodology**

The main purpose of this section is to describe the methods used in preparing the Energy Development Potential (EDP) of the Brazilian states. In this work we used factor analysis to develop the EPD and subsequently the exploratory spatial data analysis (ESDA) to analyze the results in the spatial dimension.

#### ***3.1 Factor Analysis***

Factor analysis (FA) is a multivariate technique used in order to identify a small number of factors formed by a set of observed variables. These factors express patterns of similar characteristics between the variables, so these are highly correlated with each other, which allows the grouping of them in the respective factors. Therefore, Johson and Wichern (2002) define factor analysis as *"a set of statistical techniques intended to represent or describe a number of initial variables from a smaller number of hypothetical variables."*

Therefore, this type of analysis can be characterized as an exploratory technique which

summarizes the information from a set of variables into a smaller set of factors, each factor is composed of variables that correlate highly with each other and interact weakly with the variables present in the other factors.

According Johson and Wichern (2002) factor analysis is used in the literature for the following purposes:

- a) to get through a large number of variables selected (the original data matrix) is obtained a smaller number of factors, without loss of information;
- b) reproduce patterns of relations between the separate groups of variables obtained through the factors, and;
- c) get conditions to interpret the relationships between variables in a logical way.

The attainment of these factors occurs through the analysis of covariance and correlation between different observed variables, and a portion of covariance (or correlation) is explained by common factors while the other portion is called an unexplained single factor. Thus, according Norusis (1994) their way of writing a model for a general variable " $X_i$ " standardized is as follows:

$$X_i = A_{i1}F_1 + A_{i2}F_2 + \dots + A_{ik}F_k + U_i + E \quad (1)$$

where:

- $A_{ik}$  represents the load factor that is used to combine linearly the common factors and points the intensity of the correlation between  $X_i$  and  $F_k$ ;
- $F_k$  are the common factors, with  $1 \leq k \leq N$ , where  $N$  is the number of variables;
- $U_i$  single factor, with  $i = 1, 2, \dots, N$ ;

-  $E_i$  is the error factor.

Haddad (1989) divides the process of factor analysis in four steps. They are:

- a) calculation of the correlation matrix of all variables to be analyzed, which allows investigation of associations between these variables. During this step you can check the appropriateness of using the factor analysis in the intended model. For this Kaiser-Meyer-Olkin (KMO) and Bartlett test are performed;
- b) the second step is performed to extract the most significant factors. The factor loading of each variable represents the correlation between this variable and its respective factor;
- c) after obtaining the initial factor, rotation of the factors is performed, which allows a better understanding of themselves and a better definition of the correlation between them and the variables;
- d) in the last step the matrices of factor coefficients and normalized original data are multiplied in order to generate the factor scores. These are the scores that are used in several tests.

### ***3.2 Exploratory Spatial Data Analysis***

In order to analyze the correlation of the results obtained from spatial form we use Exploratory Spatial Data Analysis (ESDA). In this paper we will use the Moran's I autocorrelation index and the Moran scatter plot of local indicators of spatial association (LISA).

According to Anselin (1998) this method is used to: describe the spatial distribution and patterns of spatial association (spatial clusters), identify the existence of different spatial

regimes or other forms of spatial instability and verify atypical observations.

### 3.2.1 Dispersion diagram and the Moran's I

The calculation of Moran's I is undertaken only in order to identify the overall trend of grouping data, *ie*, the global spatial autocorrelation, and is not efficient to identify local trends of association. The formula used to calculate the Moran's I is:

$$I_t = \left( \frac{n}{S_o} \right) \left( \frac{z_t' W z_t}{z_t' z_t} \right) \quad t = 1, \dots, n \quad (2)$$

where:

W is the matrix of spatial weights  $w_{ij}$  and indicate the form of connection between regions  $i$  and  $j$ .  $z_t$  is a vector of  $n$  observations in the form of deviation from the average in year  $t$  and  $S_o$  is the sum of all elements of  $W$ .

To identify local trends of association the Moran scatterplot is used. In this diagram it is possible to observe trends not only local but also the general trend of association.

The general trend of association is presented as a straight line and indicates how the data fit between the values of the variable lagged spatially ( $Wz$ ) and the observed values for each spatial unit ( $z$ ). The local trends are represented by the points contained in the diagram and the analysis of these points enables us to identify the outliers. (Anselin, 1995)

The Moran scatter plot is divided into four quadrants, each of which indicates a pattern

of association between the local spatial units and their neighbors.

The first quadrant, located in the upper right, is made up of regions with high values for the variable of analysis, surrounded by regions that also have high values for this variable. This quadrant is called high-high (HH);

The second quadrant, located at the top left of the diagram, is classified as low-high (LH), it indicates the regions with low values surrounded by neighbors with high values;

The third quadrant, located in the lower left corner, is formed by regions with low values surrounded by neighbors who also have this situation. This quadrant is classified as low-low (LL);

The fourth quadrant of the diagram is located in the lower right corner and consists of regions with high values for the variables analyzed and are surrounded by regions with low values for these variables in the analysis. This quadrant is classified as high-low (HL).

Therefore, the regions that are in quadrant HH and LL have positive spatial autocorrelation with its neighbors. While the regions located in the quadrants LH and HL have negative spatial autocorrelation.

### *3.2.2 Local Indicators of Spatial Association (LISA)*

The local indicators of spatial association (LISA) are another way to identify local trends of association, and indicate the degree of significance of spatial grouping. To Anselin (1996) the sum of the LISA statistics for all regions is proportional to the indicator of global spatial autocorrelation.

The calculation of the LISA statistic has the following form:

$$I_{i,t} = \frac{(x_{i,t} - \mu_t)}{m_o} \sum_j w_{ij} (x_{j,t} - \mu_t) \quad \text{with } m_o = \frac{(x_{i,t} - \mu_t)^2}{n} \quad (3)$$

where  $x_{i,t}$  corresponds to the observation of a variable analyzed in the region  $i$  in year  $t$ .  $\mu_t$  indicates the average of observations among regions in year  $t$ , for the calculation of  $j$ , only the values for the nearest neighbors are considered.

So when  $I_{i,t}$  has a positive value means that there are spatial clusters with values (high or low) for the region  $i$  and its neighbors, and when presenting negative values indicate the existence of spatial clusters with different values (high-low or low- high) for the regions and their neighbors.

### **3.3 Database**

The database used in this paper was obtained from the following Brazilian ministries and statistical offices: Institute of Applied Economic Research (IPEA), Brazilian Institute of Geography and Statistics (IBGE), Ministry of Mines and Energy (through the National Energy Balance - BEN) and Ministry of Transport.[Brasil (2005, 2008, 2010); IBGE (2011) and IPEA (2011)].

All data are collected within the period of 2004-2009. The choice of variable and each year has been based on the availability and on the possibility of a wide characterization of the energy sector.

After being collected, the data were transformed into relative values in order to avoid mistakes in interpreting the results that could arise if variables were used in absolute values, since these can be correlated to the size of population or to the extension of the study area.

The choice of variables in the construction of the energy development potential aimed to pose questions such as: demand side; supply side; economic activity; sectoral development; industrialization process; GDP; national income; economic structure and presence of reserves [Fan and Xia (2011); Chen and Lin (2008); Rapanos and Polemis (2006); Kagawa and Inamura (2004); Karl (2009) and Sola *et al* (2006). [Table 1 and Figure 2].

**Table 1. Variable Description**

CODE	DESCRIPTION	SOURCE
<b>var001</b>	Share in total residential electricity consumption (2004)	B.E.N
<b>var002</b>	Share in total residential electricity consumption (2009)	B.E.N
<b>var003</b>	Share in total residential LPG consumption (2004)	B.E.N
<b>var004</b>	Share in total residential LPG consumption (2009)	B.E.N
<b>var005</b>	Share in total number of vehicles (2004)	ANTT
<b>var006</b>	Share in total number of vehicles (2009)	ANTT
<b>var007</b>	Share in total electricity consumption in the commercial sector (2004)	B.E.N
<b>var008</b>	Share in total electricity consumption in the industrial sector (2004)	B.E.N
<b>var009</b>	Share in total electricity consumption in other sectors (2004)	B.E.N
<b>var010</b>	Share in national GDP (2004)	IBGE
<b>var011</b>	Share in national GDP (2004)	IBGE

<b>var012</b>	Share in total power generation GWh (2004)	B.E.N
<b>var013</b>	Share in total power generation GWh (2009)	B.E.N
<b>var014</b>	Share in total hydraulic potential in MWh (2005)	B.E.N
<b>var015</b>	Share in total hydraulic potential in MWh (2009)	B.E.N
<b>var016</b>	Share in total oil production in $10^3 \text{ m}^3$ (2004)	B.E.N
<b>var017</b>	Share in total oil production in $10^3 \text{ m}^3$ (2009)	B.E.N
<b>var018</b>	Share in total natural gas production $10^6 \text{ m}^3$ (2004)	B.E.N
<b>var019</b>	Share in total natural gas production $10^6 \text{ m}^3$ (2009)	B.E.N
<b>var020</b>	Share in total proved reserves of oil and natural gas $10^6 \text{ m}^3$ (2005)	B.E.N
<b>var021</b>	Share in total proved reserves of oil and natural gas $10^6 \text{ m}^3$ (2009)	B.E.N

**Figure 2. Synthetic Framework of the variables used in the final model**

<b>Author</b>	<b>Main idea of the paper</b>	<b>Factors</b>	<b>Variables</b>
Fan and Xia (2011) Chen and Lin (2008) Junior Pinto et al (2007)	<ul style="list-style-type: none"> <li>✓ Level of economic activity</li> <li>✓ Industrialization</li> <li>✓ Income</li> </ul>	F1	Var 010; Var 011; Var 012; Var 003 e Var 004.
Rapanos and Polemis (2006)	<ul style="list-style-type: none"> <li>✓ Residential energy demand</li> <li>✓ Gross Domestic Product</li> </ul>	F1	Var 001; Var 002; Var 004; Var 005; Var 017; Var 018;
Kahrl (2009)	<ul style="list-style-type: none"> <li>✓ Economic structure</li> </ul>	F1	Var012; Var 011; Var 012
Sola <i>et al</i> (2006)	<ul style="list-style-type: none"> <li>✓ Infrastructure of generation</li> </ul>	F2; F3	Var 023; Var 026; Var 032; Var 033; Var 040; Var 041; Var 042; Var 043
Sola <i>et al</i> (2006)	<ul style="list-style-type: none"> <li>✓ Reserves</li> </ul>	F3	Var 055; var 056

Source: Elaborated by the authors

## 4 Description of the results

### 4.1 Factor Analysis

To analyze the generation capacity and consumption of the 27 Brazilian states was used

initially the factor analysis in order to construct an indicator of energy potential development in these regions. Initially, we built a database formed by 66 variables relating to different forms of energy supply and demand.

The final model was constructed for 21 variables. The reduction in the number of variables used was due to the impossibility of invert the correlation matrix with a determinant equal to zero. The removal of each variable in the model followed these criteria: lower level of correlation with other variables, lower commonality of the variable and lower factor load of these variables with their respective factors.

To verify the consistency of the original data (21 variables for the 27 Brazilian states) tests were performed. The Kaiser-Meyer-Olkin index test had stood at 0.54 that is an acceptable range to a factor analysis. With the Bartlett's test of sphericity (BTS) was verified the impossibility of the correlation matrix was an identity due to the high value in the BTS (1937.056) and a significance level equal to zero.

**Table 2 - Commonality, eigenvalue and explained variance**

<b>Factor</b>	<b>Eigenvalue</b>	<b>% of Variance</b>	<b>% Cumulative</b>
1	11,257	53,605	53,605
2	5,797	27,604	81,209
3	2,954	14,066	95,275
<b>Variable</b>	<b>Commonality</b>	<b>Variable</b>	<b>Commonality</b>
Var0001	0,993	Var0012	0,787
Var0002	0,990	Var0013	0,862
Var0003	0,977	Var0014	0,915
Var0004	0,977	Var0015	0,944
Var0005	0,994	Var0016	0,929
Var0006	0,995	Var0017	0,924
Var0007	0,984	Var0018	0,915
Var0008	0,945	Var0019	0,978
Var0009	0,978	Var0020	0,979
Var0010	0,988	Var0021	0,966
Var0011	0,989		

Source: Elaborated by the authors based on the results of Factor Analysis

It was used three factors in this paper. The number of factors was predetermined so that there was a better explanation of them. Each of these factors had an eigenvalue greater than one, and together was able to explain 95.27% of the total variance of the model (Table 2).

In analyzing the results we see that for all variables the value of commonality is greater than 0.75, as shown in Table 2. This means that over 75% of the variance of the variable is reported by common factors.

**Table 3 - Matrix of rotated components**

Variables	Component		
	1	2	3
var001	,975		
var002	,979		
var003	,964		
var004	,967		
var005	,990		
var006	,989		
var007	,968		
var008	,933		
var009	,953		
var010	,979		
var011	,978		
var012	,590		,661
var013	,627		,683
var014			,952
var015			,962
var016		,959	
var017		,955	
var018		,956	
var019		,988	
var020		,969	
var021		,944	

Source: Elaborated by the authors based on the results of Factor Analysis

In the initial factor extraction was used principal components method. All values of the factor loadings were above 0.65 and in 19 variables were above 0.9 (Table 3). This represents a high positive correlation between the variables and their respective factors.

In order to rotate the factors we used the VARIMAX method, which minimizes the number of variables that have a higher weight factor. This rotation was performed in order to better define the relationship between the variables and factors.

With the relationship between variables and factors established we rank each of the three factors and determine the variables that comprised each one:

1. Factor 1 represents 53.6% of the total variance of the variables and could be classified as being the **energy demand factor**. This factor consists of the following variables: Share in total residential electricity consumption in 2004 and 2009 (Var001 and Var002) Share in total residential LPG consumption in 2004 and 2009 (Var003 and Var004) Share in total number of vehicles 2004 and 2009 (Var005 and Var006); Share in total electricity consumption in the commercial sector (2004) (Var007), industrial (Var008) and other industries (Var009) for the year 2004, and share in the national GDP in 2004 and 2009 (Var010 and Var011).
2. The second factor, which represents 27.6% of the total variance was identified as a *factor of supply of non-renewable energy*. The following variables make up this factor: Share in total oil production in  $10^3 \text{ m}^3$  - 2004 and 2009 (Var016 and Var017), Share in total natural gas production  $10^6 \text{ m}^3$  - 2004 and 2009 (Var018 and Var019), and Share in total proved reserves of oil and natural gas  $10^6 \text{ m}^3$  - 2005 and 2009 (Var020 and Var021).
3. The third factor is responsible for explaining 14.1% of the total variance and was determined to be the *factor of renewable energy supply*. The four variables in this factor are: Share in total power generation GWh - 2004 and 2009 (Var012 and Var013) and Share in total hydraulic potential in MWh (2005 and 2009 (Var014 and Var015).

Finally was performed matrix multiplication. The factor coefficient matrix was multiplied by the normalized original data. This multiplication results in the factor scores for each observation of the model. Through the multiplication of factor scores with the variance of each factor was created a classification of the Brazilian states and its potential energy demand, supply of non-renewable energy and renewable energy supply.

The Energy Development Potential (EDP) was established for each Brazilian state as follows: weighting the factor scores of each of the three factors of the corresponding factor variance. Mathematically we have:

$$PDE = \sigma_i F1 + \sigma_i F2 + \sigma_i F3 \quad (4)$$

Where:  $\sigma_i$  is the variance of each factor. The subscript "i" indicates the factor.

## ***4.2 Spatial Analysis of Results***

### *4.2.1 Descriptive Analysis of the factors and the Energy Development Potential (EDP)*

In order to analyze the Energy Development Potential (EDP) and the factors constructed in the factor analysis for the 27 Brazilian states was held exploratory spatial data analysis (ESDA). This analysis identifies clusters between these units, and is possible to describe the spatial distribution of each of the three factors and the EDP.

Since the ranking of the Brazilian states for each of the factors and for the Energy

Development Potential, it was possible to perform a grouping of these states into five levels: very high, high, medium, low and very low. For the identification of the Brazilian states belonging to each level was used the following criteria:

Very High (VH):  $X \geq \text{mean} + 0.5 \text{ SD}$

High (H):  $\text{mean} \pm \text{SD} 0.5 > X \geq \text{mean} + 0.25 \text{ SD}$

Medium (M):  $\text{mean} + 0.25 \text{ SD} > X \geq \text{mean} - 0.25 \text{ SD}$

Low (L):  $\text{mean} - 0.25 \text{ SD} > X \geq \text{mean} - 0.5 \text{ SD}$

Very Low (VL):  $X < \text{mean} - 0.5 \text{ SD}$

Where: X represents the result of each spatial unit for a particular factor or for the EDP and SD is the standard deviation of these factors.

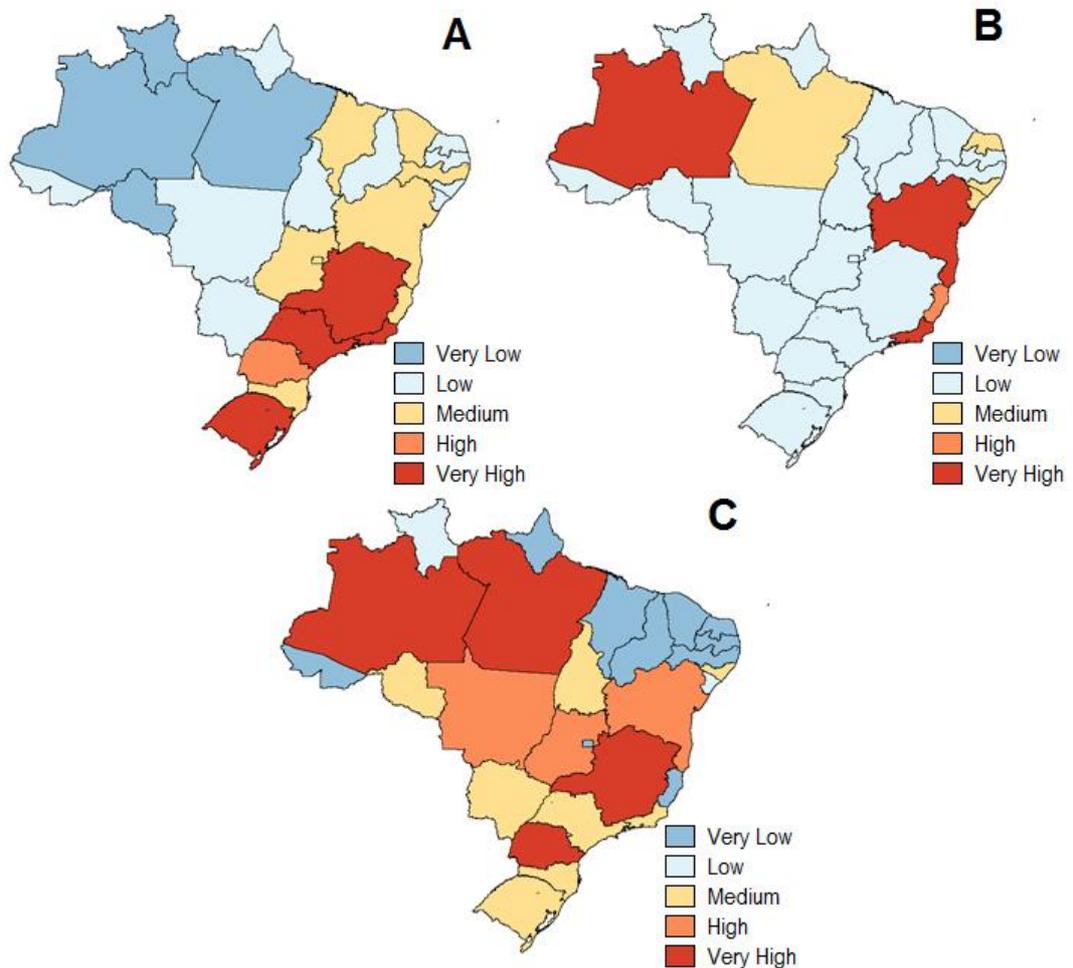
After this classification was possible, through the software GeoDa<sup>TM</sup>, perform a spatial analysis of the results.

For factor 1, which is the factor related to electricity demand, has been represented in the classification map in Figure 3. In this analysis we can identify a very high level (VH) of energy demand for the states of São Paulo, Minas Gerais, Rio Grande do Sul and Rio de Janeiro. On the other hand states like Pará, Amazonas, Roraima and Rondônia had very low level (VL). These results are consistent in that there is a strong direct relationship between Gross Domestic Product (GDP) and energy demand. Therefore, one would expect that states with higher GDP's and higher industrialization levels present higher demands for energy, as seen in map A.

On the map B (Fig. 3) was exposed the classification of factor 2, which represents the

supply of non-renewable energy. In it was possible to see that the states where most of reserves of natural gas and oil are located, and consequently where the amount produced is high are classified as very high level (HL). Examples are the states of Rio de Janeiro, Bahia and Amazonas. While states where the production of non-renewable energy resources are scarce, for example, Sao Paulo, Minas Gerais and Rio Grande do Sul, the level of supply of non-renewable energy was classified as low (L).

**Figure 3 - Maps of energy demand and supply of non-renewable energy and renewable Brazilian states**



Source: Elaborated by the authors

Regarding the supply of renewable energy (factor 3) 51.85% of Brazilian states have levels greater than or equal to the average (A). This follows from the fact that at these states there are high levels of water power and electricity generation, such as the State

of Parana, which takes advantage of the high hydraulic potential of the Iguazu River and the Itaipu dam, and the State of Minas Gerais, where are located large amounts of hydroelectric power plants. Both States had a very high level (HL) for this factor. This analysis was more detailed and the map C (Figure 3).

With the result obtained for each factor was possible to calculate the energy potential development for each unit of the Federation. The result of this calculation is shown in Table 4 and by means of spatial analysis in Figure 4.

**Table 4 – Energy Development Potential**

<b>State</b>	<b>EDP</b>	<b>State</b>	<b>EDP</b>
Acre	0	Mato Grosso	7,9821
Amapá	0,2743	Pernambuco	8,9809
Roraima	0,6527	Espírito Santo	11,5756
Piauí	1,5363	Goiás	13,7963
Paraíba	2,4240	Santa Catarina	14,9394
Tocantins	2,7458	Amazonas	18,5674
Rondônia	3,9529	Pará	21,6785
Maranhão	4,3683	Rio G. do Sul	24,0600
Distrito Federal	4,7788	Bahia	25,9859
Sergipe	5,5461	Paraná	35,0299
Rio G. do Norte	6,2801	Minas Gerais	40,5942
Mato G. do Sul	6,4974	Rio de Janeiro	74,7200
Ceará	6,9903	São Paulo	100,000
Alagoas	7,3304		

Source: Elaborated by the authors

Analyzing the results of the EDP could make the following inferences:

- i. 59.26% of Brazilian states showed low Energy Development Potential (L) or very low (VL), ie, the vast majority.
- ii. The State of São Paulo had the highest potential of energy development across all units of the Federation. This is due in part to the fact that State has the highest degree of industrialization and is the most populous among the Brazilian states, making it a very

high demanding for energy.

iii. The Rio de Janeiro presented high EDP produced from the large amount of non-renewable sources of energy such as natural gas and oil in the state. The state is considered the major producer of these sources. In addition, the State presented high levels of demand for energy due to its high industrialization and high participation in the national GDP.

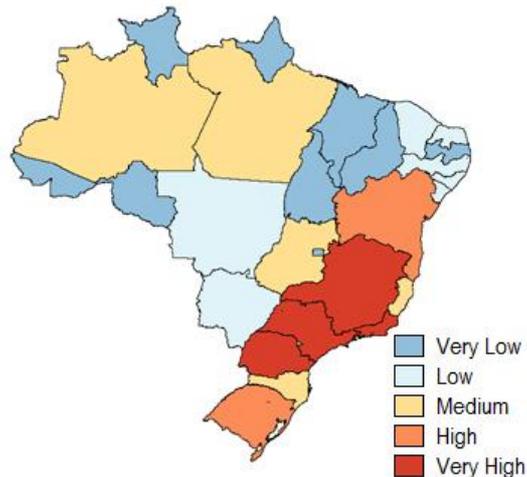
iv. States such as Minas Gerais and Paraná showed high values for the EDP as a result of large amounts of hydroelectric power plants in those states and high standards of hydroelectric potential.

v. The State of Amazonas had a medium level (M) is the EDP. This level is directly linked to the large energy supply of the State in relation to both renewable and nonrenewable sources, due to the great hydroelectric potential of rivers in the state and the production of natural gas and oil.

vi. States of Acre, Amapá, Roraima, Piauí and Paraíba, for example, had very low levels of EDP. This may be closely linked to low income in these states. In addition, these states offer very little renewable energy and produce almost of non-renewable sources.

vii. It was found higher rates of EDP for the states of the south-Southeast, a region with higher income and a few northeastern states. In contrast states with lower incomes, as is the case for most states in the region north-northeast, had the lowest rates of EDP.

**Figure 4 - Map of the Energy Potential Development of the Brazilian States**



Source: Elaborated by the authors

#### 4.2.2 Univariate Spatial Analysis

In order to identify a spatial correlation of the data obtained were used in this work two tools derived from Exploratory Spatial Data Analysis (ESDA). These were: the Moran scatter plot and local indicators of spatial association (LISA). Both indicators were used for  $k = 3$  neighbors.

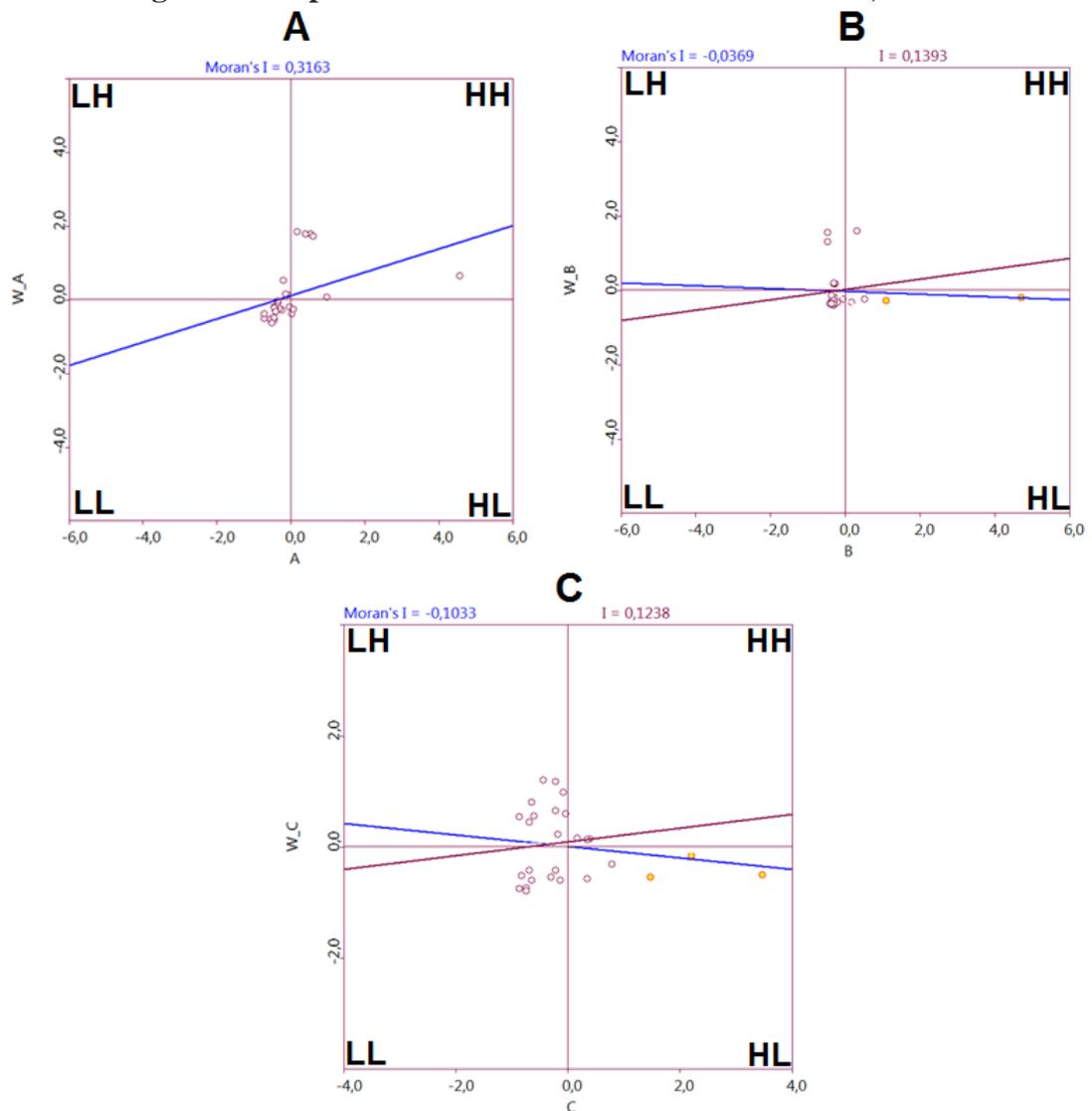
##### 4.2.2.1 Moran's $I$

Using the Moran scatter plot was possible to identify trends of association between each spatial unit by the results obtained for each of three specific factors and the EDP. Besides this association, was identified as discrepant values held in relation to the others (outliers).

The factor 1 with  $k = 3$  neighbors is represented in the graph of figure 5 in the Moran scatter plot

The results show a strong concentration in energy demand among the states. The high positive value found for Moran's I (0.3163) showed a high positive spatial correlation between the states. At this factor the majority of the states is located at the third quadrant (LL). This means that a large number of Brazilian states have lower standards of energy demand and that are surrounded by neighbors who also have low values of this factor.

**Figure 5 - Dispersion of univariate Moran for factors 1, 2 and 3**



Source: Elaborated by the authors

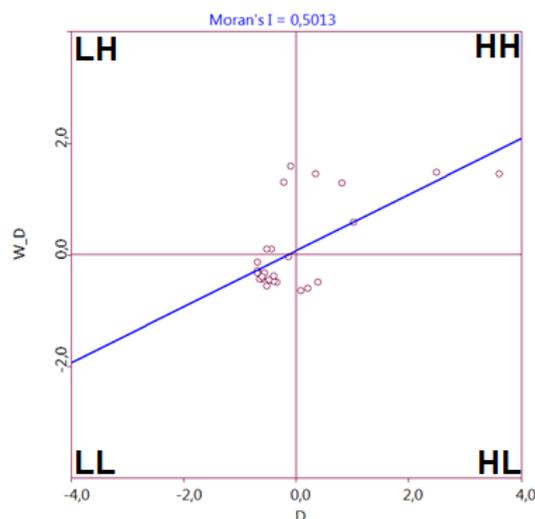
The graph B of Figure 5 also presents the Moran scatter plot for the supply of non-renewable energy factor (factor 2). In it was possible to identify a Moran's I equal to -

0.0369, which indicates a negative spatial correlation between units of the Federation with respect to the provision of non-renewable energy. But by eliminating outliers was possible to obtain a value for the Moran's I equals 0.1082, which represents a strong spatial concentration of this factor.

For factor 3 (supply of renewable energy) was performed the same process of removing outliers mentioned in factor 2. This withdrawal took place in order to obtain a positive value for Moran's I, so that the graph C (Figure 5) are located two values found (before and after removal of outliers). In this chart was still possible to notice a large number of units of the Federation in the second quadrant (LH), which indicates that there is an excessive number of states with low supply of renewable energy surrounded by neighbors who have a high supply of this factor.

The results of the scatter plot for the EDP are shown in Figure 6.

**Figure 6 - Scatterplot univariate Moran for the EDP**



Source: Elaborated by the authors

With respect to the results we conclude that the Moran's I for the Energy Development

Potential has value 0.5013 (Figure 6), which indicates a strong positive spatial correlation between the states. It was also possible to identify that much of the Brazilian state is located at the third quadrant (LL), which shows a high number of states with low EDP and have neighbors who have the same situation.

#### *4.2.2.2 Local Indicators of Spatial Association (LISA)*

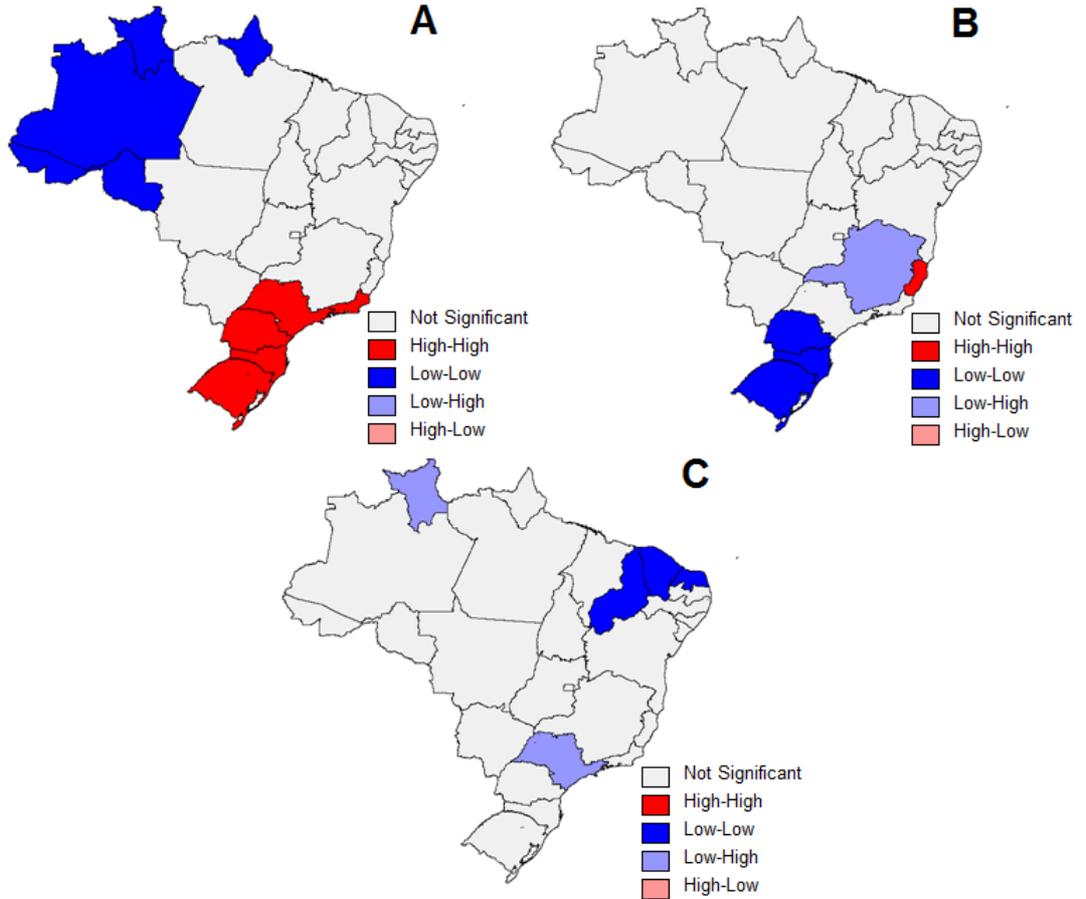
The measure of the degree of significance of spatial clustering was performed by the LISA index. In Figure 7 LISA indicators are represented for factors 1, 2 and 3. On the map, for the first factor, it is possible to identify the formation of a high-high cluster (HH) to Sao Paulo, Rio de Janeiro and the States of the South region, and a low-low cluster (LL) for the states of Acre, Amazonas, Rondônia, Roraima and Amapá.

For the second factor it was possible to observe the formation of a high-high cluster (HH) to Espírito Santo, a low-high (LH) to Minas Gerais and a low-low (LL) for the states located in South region. These clusters are represented on the map B.

The supply of renewable energy has been represented on the map C. It is possible to observe the formation of a low-low cluster (LL) for Piauí, Ceará and Rio Grande do Norte and a low-high cluster (LH) to Sao Paulo and Roraima.

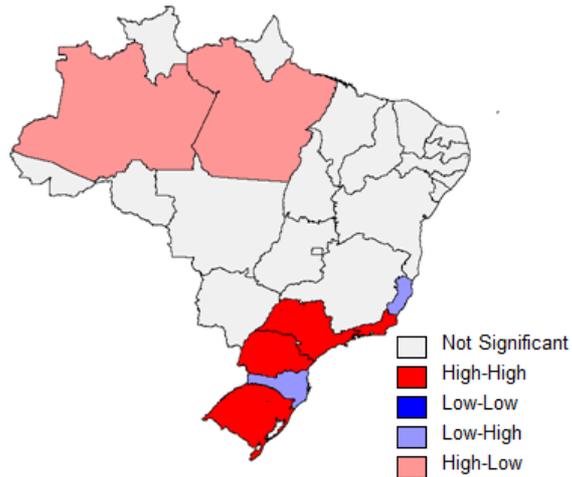
Regarding the Energy Development Potential (EDP) - Figure 8 - you can observe the formation of a high-high cluster (HH) to Sao Paulo, Rio de Janeiro, Paraná and Rio Grande do Sul, a low-high cluster (LH) to Santa Catarina and Espírito Santo, and a high-low cluster (HL) for Pará and Amazonas.

Figure 7 - Cluster Maps for factors 1, 2 and 3



Source: Elaborated by the authors

Figure 8 – Cluster Maps for the EDP



Source: Elaborated by the authors

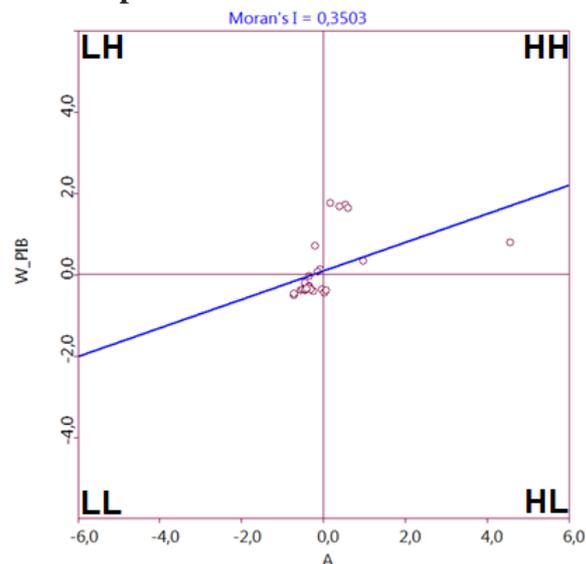
4.2.3 - Bivariate analysis

The bivariate analysis is performed through the analysis of correlation between gross domestic product and factor 1 (standard energy demand) and the correlation between the energy development potential and the gross value of industrial production (economic activity).

#### 4.2.3.1 Dispersion of bivariate Moran

In order to better understand the level of energy demand we decide to add the GDP in the study of the spatial correlation of the units of the Federation. It is possible to observe that there is a high correlation index and the observations are closer to the regression line. Moreover, the value of Moran's I rose to 0.3503, showing this effect, which is presented in Figure 9.

**Figure 9 - Scatterplot of bivariate Moran for Factor 1 x GDP**



Source: Elaborated by the authors

Figure 9 has on the horizontal axis the level of energy consumption, represented by A, and GDP on the vertical axis, represented by W\_PIB, so that states with higher energy consumption and GDP present neighbors in the same situation and are located in the first quadrant (high-high), and the states with lowest GDP's and energy demands are

neighbors of states with the same characteristics and are located in the third quadrant (low-low).

Another analysis made by Bivariate Moran scatter plot use the Energy Development Potential (EDP) and the gross value of industrial production (used as a proxy for economic activity) for industries with low, medium and high energy intensity in each unit of the Federation.

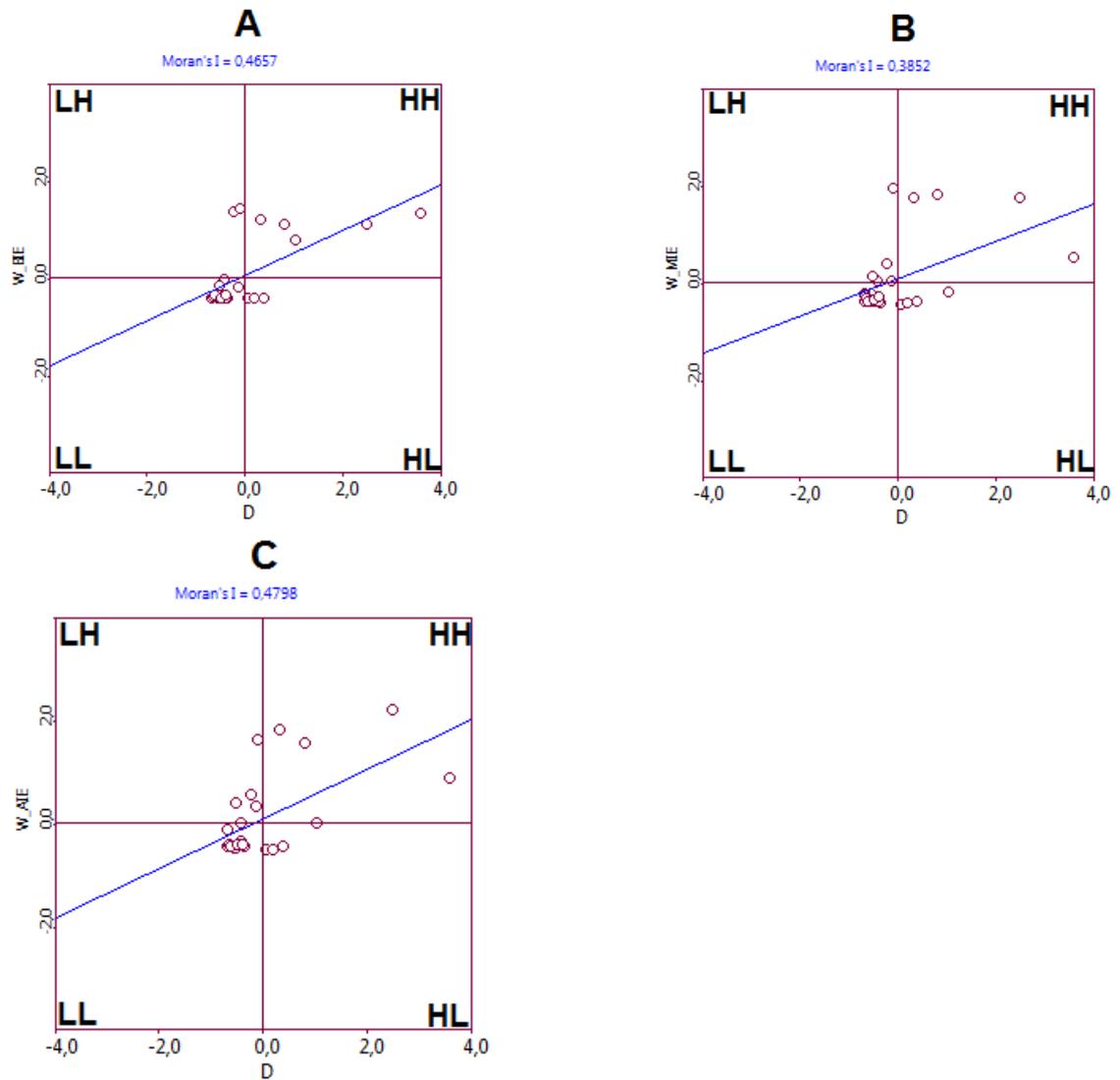
For it was possible to establish three different levels of energy intensity. We used the work of Santos (2010), which calculates the intensity to 30 Brazilian sectors. From these calculations we could classify the divisions of industrial activities found in the Annual Industrial Survey (PIA) at low, medium or high energy intensity. After the industrial activities are classified, we summed the gross production of each sector belonging to each of the three levels of energy intensity for all 27 states of the Federation, and then perform a bivariate analysis for this index and the EDP. Figure 10 shows the result obtained through the charts for the areas of low (Figure A), medium (Figure B) and high energy intensity (Figure C).

Higher positive values of Moran's I statistic to analyze the three levels of technological intensity indicate a strong positive global spatial autocorrelation, i.e., a strong interaction between the states. This indicates that states that have high values of the indicator of EDP have neighbors with high levels of economic activity, here represented by the gross value of production, or alternatively, states with low values of the indicator of EDP have neighbors with low levels of economic activity.

These results confirm the theory that the level of economic activity is a fundamental

variable in determining the potential of energy development in a region and the potential of its neighbors, since according to the National Energy Balance 2009, the industrial sector accounted for 40% of the average energy consumption in the period between the years 2005 and 2008.

**Figure 10 - Scatter plots of bivariate Moran for the EDP x Gross Value of Production in industries with low, medium and high energy intensity**

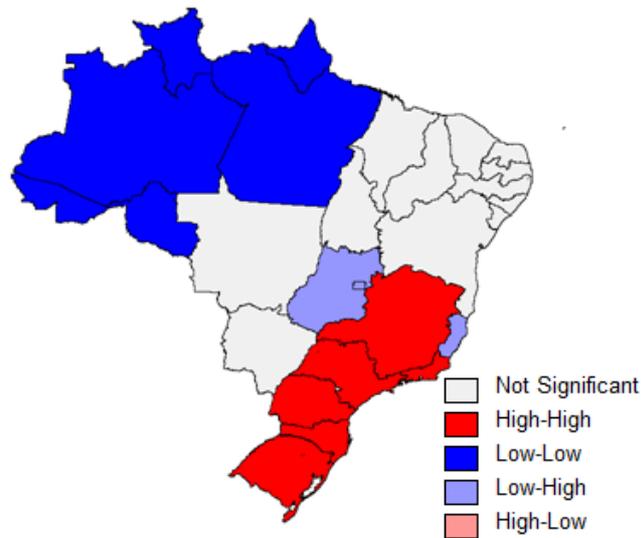


Source: Elaborated by the authors

#### 4.2.3.2 Multivariate LISA

Incorporating the variable GDP also in the analysis of LISA for factor 1 was possible to identify the emergence of a low-high cluster (LH) for Goiás and Distrito Federal Espirito Santo, a cluster of high-high (HH) to Minas Gerais and a low-low cluster (LL) to Para, and the clusters that existed prior to the inclusion of the variable GDP (Figure 11).

**Figure 11 - Clusters Maps- LISA Multivariate**

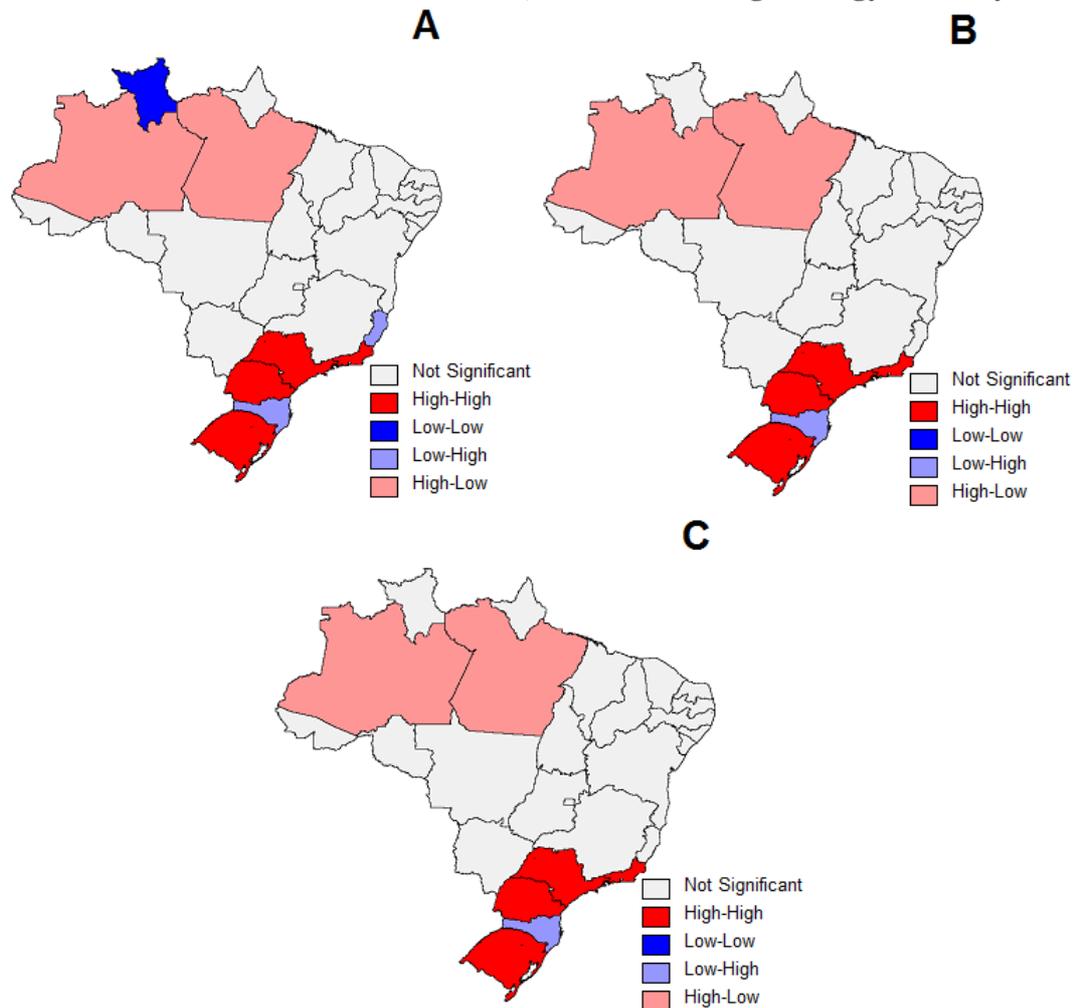


Source: Elaborated by the authors

Figure 11 shows the dichotomous nature of Brazilian development, i.e., there are two well-defined spatial patterns in this figure. The first cluster pattern formed by the HH which brings together the units of the Federation located in South and Southeast regions of the country and the second pattern formed by the cluster that brings together the LL of the Federation located in northern Brazil.

Here is also performing a multivariate analysis of LISA for the EDP and the gross value of production (a proxy for economic activity) for the three levels of industrial energy intensity. In Figure 12 are the maps of cluster analysis to the low level (map A), medium (map B) and high energy intensity (map C).

**Figure 12 - Cluster Maps - Multivariate LISA - to the PDE and the Gross Value of Production in industries with low, medium and high energy intensity**



Source: Elaborated by the authors

When the analysis is performed to sectors with medium and high energy intensity (Map B and Map C) is obtained maps of identical clusters. On these maps you can identify the formation of clusters high-low (HL) for the states of Para and Amazonas, high-high clusters (HH) to Rio Grande do Sul, Parana, Sao Paulo and Rio de Janeiro and cluster low-high (LH) for the State of Santa Catarina.

In the multivariate analysis for industries with low energy intensity beyond the clusters seen in the previous analysis also showed that the low-low cluster (LL) for the states of Roraima and Amapá and low-high cluster (LH) to Espírito Santo.

But once it was observed the characteristic of this division in the development of Brazil, now expressed by the indicators of EPD and the levels of economic activity of the Federation, indicated by the gross industrial output. This dichotomous character development gives rise to higher values of the indicators in the South-Southeast region and low values for the states located in the north.

## **5 Final Considerations**

The use of factor analysis as a way to get the scores needed in the preparation of Energy Development Potential has identified three factors that together explained 95.27% of the total variance of the model. Through analysis of these factors was possible to conclude there is a concentration of energy demand in the South-Southeast, showing the strong link between income level and energy consumption. It was also observed a concentration of supply of non-renewable energy in a few Brazilian states, for example, Rio de Janeiro and Amazonas, and a large number of states with high renewable energy supply.

The development of the indicator EDP allowed to check for 59.26% of the units of the Federation with below-average level of this index, which indicates a low capacity of most Brazilian states to meet adequately the growing needs of the country in the energy sector. In the ranking of states in its indicators of EDP was possible to identify very-high levels (HL) to Sao Paulo, most probably due to industrialization and population and, consequently, increased demand, Rio de Janeiro, which has a high energy supply non-renewable and high participation in the national GDP, and Minas Gerais and Parana, both with high renewable energy supply due to the large number of dams and

hydroelectric potential of high standards.

At the other extreme, states with low income, low supply of renewable energy production and almost no non-renewable energy, as is the case of Acre, Amapá, Roraima, Piauí, Paraíba, Tocantins, Rondonia, Maranhão and Distrito Federal, showed very-low levels (HL) for the indicator EDP.

These results were observed when done better spatial analysis. Through it you can add the first GDP variable in the multivariate analysis, and highlight the dichotomous nature of Brazilian development, through the existence of a high-high cluster of the states of South and Southeast and a low-low cluster aggregating the states of the Federation the Northern region of Brazil. Then we could confirm, by including the gross value of production in industries with low, medium and high energy level, the relevance of economic activity in determining the potential of energy development in a region.

This is important to state that although this work using data on supply, demand and energy infrastructure, seeking to serve the largest possible sphere of industry, it is not possible to make predictions or projections of the data, since it is not that the purpose of factor analysis. For this technique is only possible to provide an image of the current energy scenario and thus assist the provider in making public policy decisions best for this sector.

The absence of similar work, i.e., the absence of studies that seek to measure an indicator of potential energy development in the literature prevented the development of comparison between this study and others.

## REFERENCES

- ANSELIN, L. Local indicators of spatial association – LISA. *Geographical Analysis*. V 27 (2), April. p 93-115. 1995.
- ANSELIN, L. The Moran scatterplot as an ESDA tool to assess local instability in spatial association. Fisher, M, Scholten, H.J and Unwin, D W (eds). **Spatial analytical perspectives in GIS**. Taylor&Francis. London. p 111-125. 1996.
- ANSELIN, L. Interactive techniques and exploratory spatial data analysis. Longley P. A, Goodchild M.F, Maguire D.J and Wind D. W (eds). *Geographical information system: principles, techniques, management and applications*. Wiley: New York. p 253-365. 1998
- ANSELIN, L. *Spatial Econometrics: Methods and Models*. Boston: Kluwer Academic.
- ANSELIN, L. E BERA, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A. and Giles D. E. (eds.) *Handbook of Applied Economic Statistics*, Marcel Dekker, New York, p. 237-289, 1998.
- BRASIL, Ministério de Minas e Energia. **Balanco Energético Nacional 2005**: Ano Base 2004. Rio de Janeiro: Empresa de Pesquisa Energética (EPE), 2005.
- \_\_\_\_\_, \_\_\_\_\_. **Balanco Energético Nacional 2010**: Ano Base 2009. Rio de Janeiro: Empresa de Pesquisa Energética (EPE), 2010.
- \_\_\_\_\_, Agência Nacional de Energia Elétrica. **Atlas de energia elétrica do Brasil**. 3ªed. Brasília: ANEEL, 2008.
- BURNETT, J.W., *Spatial Modeling Techniques in the Analysis of Energy Economics*, 2011.
- CHEN, C-Z and LIN, Z-S., Multiple timescale analysis and factor analysis of energy ecological footprint growth in China 1953-2006, *Energy Policy*, 2008(36): 1666-1678.
- FAN, Y and CHIA, Y. Exploring energy consumption and demand in China, *Energy* 2011, 1-8.
- HADDAD, P.R. **Economia regional: teorias e métodos de análise**. Fortaleza, BNB-ETENE, 1989.
- HEKKENBERG, M; MOLL, HC and SCHOOT UITERKAMP, A,J,M. Dynamic temperature dependence patterns in future energy demand models in the context of climate change. *Energy* 2009;34:1797-1806.
- IBGE – Instituto Brasileiro de Geografia e Estatística. **Banco de dados – Estados**. Disponível em: <<http://www.ibge.gov.br/estadosat/>>. Acesso em: abril, 2011.
- IBGE – Instituto Brasileiro de Geografia e Estatística. **Pesquisa Industrial Anual 2008**. Disponível em: <<http://www.ibge.gov.br/home/estatistica/economia/industria/pia/empresas/2008/defaultempresa.shtml>>. Acesso em: junho, 2011.
- IPEA – Instituto de Pesquisa Econômica Aplicada. **Ipeadata Regional**. Disponível em: <<http://www.ipeadata.gov.br/Default.aspx>>. Acesso em: abril, 2011.
- ITO, T; CHEN, Y,Q; ITO, S; YAMAGUCHI, K. Prospect of the upper limit of the energy demand in China from regional aspects. *Energy* 2010;35:532-537.
- JOHSON, R.A and WICHERN, D.W., *Applied Multivariate Statistical Analysis*. Prentice Hall. Fifth Edition. 2002.
- KAGAWA, S and INAMURA, H. A structural decomposition of energy consumption based on a hybrid rectangular input-output framework: Japan's case. *Economic Systems Research* 2001;4(13):339e63.

- KAGAWA, S and INAMURA, H. A simple multi-regional input output account for waste analysis. *Economic Systems Research* 2004;16(1):3e22.
- KAHRL, F and ROLAND-HOLST, D. Growth and structure change in China's energy economy. *Energy* 2009;34:894-903.
- NORUSIS, M. J. **SPSS** - professional statistics 6.1. Chicago: SPSS, 1994.
- PINTO JUNIOR, H. Q. **Economia da energia**: fundamentos econômicos, evolução histórica e organização industrial. Rio de Janeiro: Elsevier, 343 p., 2007.
- RAPANOS, V.T and POLEMIS, L.M., The structure of residential energy demand in Greece, *Energy Policy*, 2006 (34):3137-3143.
- SANTOS, G. F. **Política Energética e Desigualdades Regionais na Economia Brasileira**. São Paulo, 2010. 180 p. Tese (Doutorado em Economia) – Faculdade de Economia, Administração e Contabilidade. Departamento de Economia, Universidade de São Paulo.
- SCHMIDT, C. A. J.; LIMA, M. A. M. A demanda por energia elétrica no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 58, n. 1, p. 67-98, 2004.
- SOLA, A. V. H; XAVIER, A.A.P; KOVALESKI, J.L and RESENDE, L.M., Análise dos fatores determinantes para Eficiência energética. *Revista Produção on line*, 2006(6)1:1676-1901.
- ZHANG, N; LIOR, N; JIN, H,G. The energy situation and its sustainable development strategy in China. *Energy* 2011;36: 39-49.