ASSOCIATIONS BETWEEN DENGUE AND SOCIO-ENVIRONMENTAL VARIABLES IN CAPITALS OF THE BRAZILIAN NORTHEAST BY CLUSTER ANALYSIS

JULIO CESAR BARRETO DA SILVA[1]
CARLOS JOSÉ SALDANHA MACHADO[2]

1 Introduction

Currently, Dengue is considered the most important arboviruses in the world. About 2.5 billion people are at risk of becoming infected, particularly in tropical and subtropical regions, where climatic, social and economic conditions are favorable to mosquito-borne proliferation. The large and rapid urban-rural population flow observed in the Americas has caused a disorderly growth of the cities due to the lack of satisfactory conditions of housing and basic sanitation for a large part of the population. Especially in Brazil, more than 85% of the population live in urban areas; and about 20% in slums, huts, squat and tenements; places these where the access to housing, water supply and regular garbage collection are poor (VALLE, PIMENTA and CUNHA, 2015).

The infestation by *Aedes aegypti* has grown considerably in recent years due to uncontrolled population growth and lack of urban planning, with poor basic sanitation conditions, creating favorable environments for mosquito breeding sites and their spread throughout the country. The high larval density of this vector has led to the planning of public policies due to the concern about the spread of other diseases in the country, such as Chikungunya Fever (BRAZIL, 2014). The strategies to combat Dengue, focused on the chemical control of mosquito-borne, have not reduced the index of their infestation, since these actions are not universal and, besides being discontinuous, do not include basic sanitation and environmental education measures (FEITOSA, SOBRAL and JESUS, 2015).

Basic sanitation is defined as a set of services, infrastructures and operational facilities of: 1) public supply of potable water, from its catchment to the network connections in the building, as well as its measuring instruments; 2) adequate collection, transportation, treatment and final disposal of sanitary sewage, from the building network

1. Doutorando do Programa de Pós-Graduação em Meio Ambiente. Universidade do Estado do Rio de Janeiro (PPG-MA/UERJ); barretojcs@gmail.com
2. Pesquisador Titular em Saúde Pública da Fundação Oswaldo Cruz e Professor do PPG-MA/UERJ; saldanha@fiocruz.br

connections to its final release into the environment; 3) collecting, transporting, transshipment, treatment and final disposal of household waste and garbage from the sweeping and cleaning of public places and streets (urban cleaning and solid waste management); and 4) urban drainage of rainwater, transportation, holding or retention for the damping of flood flows, treatment and final disposal of rainwater drained in urban areas (BRAZIL, 2007, p. 100). According to Machado (2014, p. 131), Brazil (seventh largest economy in the World) occupies the 112th place among 200 nations in terms of improvement of the basic sanitation sector in the last 12 years.

One of the tools designed to strengthen actions to prevent and combat Dengue is the use of indicators, especially those that have a qualitative-quantitative approach, as they allow to portray the landscape of the disease by relating it to different factors (historical, economic, social, environmental and health), as well as they provide a deeper understanding of the processes that promote favorable conditions for the development of *Aedes aegypti* (ARACAJU, 2014). According to Machado (2013), as urgent as political and tax reform, the debate on urban reform has risen to the top of the agenda during serious public health crisis caused by the infestation of the *Aedes aegypti* mosquito, transmitter of Dengue and Yellow Fever, besides Chikungunya and Zika. A look through public health is necessary to understand why we have been living with this mosquito for decades.

In 2001, public policies[i] emerged through guidance from the World Health Organization to eradicate *Aedes aegypti*; however, this option did not become viable over time. Currently, the national guidance is for mosquito-borne control (ZARA et al., 2016). As for the water supply, according to Sardão (2016), the worst problem in the fight against Dengue is the irregularity of the service, such as lack or intermittence of water, because it takes the population to use water boxes, pots and barrels. Without caps or poorly capped, these reservoirs become ideal mosquitoes breeding places.

The hypothesis of this study is that there is a correlation between Dengue, temperature and water. This can be observed through direct association with climatic factors, or indirectly, through analysis with indicators of basic sanitation, since these are essential factors for the maintenance of the cycle of the mosquito-borne of disease (FEITOSA, SOBRAL and JESUS, 2015; GIL et al, 2015; VALLE, PIMENTA and CUNHA, 2015). Regarding urban solid waste (RSU), it is expected to observe a relationship between Dengue and the disposal of materials such as metals, glass and plastic, as they are suitable sites for accumulation of water and food, ideal for the proliferation of infectious agents (SAN PEDRO et al. 2009; ELMEC, BATAIERO and CRUZ, 2016). Regarding social aspects, it is suggested that the lack of access to information in illiterate or low educated individuals could lead to the reduction of preventive measures of control to the reproduction of the mosquito-borne, increasing the epidemics of DF.

This work aims to investigate the sociodemographic, climatic and environmental (water, sanitary sewage and urban solid waste) aspects in relation to the DF cases, in the capitals of Brazilian Northeast, henceforth NEB, in order, to identify the correlation between Dengue and the variables of the study.

2 Materials and methods

They were used in this study data relating to the nine capitals of the NEB Region, corresponding to a 12-year collection (2001-2012), from average annual values of the variables, as follows:

a) Dengue cases notifications of individuals hospitalized in the Unified Health System – SUS, in the first month of symptoms, with theirs related associated social information, from the Department of Informatics of the SUS - DATASUS (BRAZIL, 2017a):
   a.1) Schooling (years of study completed): not applicable (children not yet training), 1 to 3 years of training, 4-7, 8-11, 12 or+, illiterate individuals and ignored record (IGN);
   a.2) Age range (years): child under 1 year old, 1 to 4 years old, 5-9, 10-14, 15-19, 20-39, 40-59, 60-64, 65-69, 70-79, 80 or + and IGN;
   a.3) Gender: female individuals, male individuals and IGN; and
   a.4) Ethnicity[ii]: white, black, brown, yellow, indigenous and IGN.
b) Climatic data: precipitation (PRP); relative humidity (RH); minimum temperature (T_MEAN); mean temperature (T_MEAN) and maximum temperature (T_MAX); all from the National Institute of Meteorology (BRAZIL, 2017b). From these variables collected, we calculated:
   b.1) Vapour pressure deficit (VPD), saturated vapour pressure deficit (SVPD) and potential evapotranspiration (ETP), all according to Allen *et al.* (1998);
   b.2) Evapotranspiration of reference (ETO), according to Thornthwaite (1948);
   b.3) Heat index (ICA), according to Steadman (1979); and
   b.4) Human comfort index (ICH), according to Rosenberg (1983).
c) Demographic data: urban and total (urban + rural resident people) population size, respectively, URBAN_POP and TOT_POP; both measured by resident inhabitants (inhab.) of each capital for each year of study; from the Brazilian Institute of Geography and Statistics – IBGE (BRAZIL, 2017c); and
d) Fifty-six indicators of water and sewage or AE; and eighty-nine indicators of urban solid waste or RSU (considering domestic solid waste – RDO; and public solid waste - RPU); from the National Information System on Sanitation - SNIS (BRAZIL, 2017d). Information (definitions, calculations and constitutions) referring to the codes of these indicators here used are available on the SNIS homepage (http://www.snis.gov.br/glossarios).

To the absent data, the filling-in-imputation technique was applied from the XL-STAT[iii] program.

The NEB Region is one of the five constituents of Brazilian territory. It encompasses nine states and their respective capitals: Alagoas – Maceió, Bahia – Salvador, Ceará – Fortaleza, Maranhão – São Luís, Paraíba – João Pessoa, Pernambuco - Recife, Piauí - Teresina, Rio Grande do Norte – Natal and Sergipe – Aracaju. The NEB territorial extension is from 1,554,257.0 Km^2, approximately 18.2% of the area of the national territory, with predominantly urban population of approximately 56,915,936.0 of inhabitants, equivalent to 73.0% of the NEB or 27.62% of the Brazilian population (BRAZIL, 2016b).

Regarding ethnicity, the NEB Region is composed by 62.5% of brown; 29.2% of white; 7.8% of blacks and 0.5% of indigenous people. In addition, about 55.0% of its households do not have environmental sanitation; and the life expectancy in this Region is the lowest in Brazil in the last 70 years: 1940-2010 (BRAZIL, 2010). In the period from 1980 to 2014, its population increased by 62.0%. Already, the number of extreme poverties fell sharply from about 15 million to about 6 million, in the period 2001-2013, with a largest number of people in this situation in the state of the Bahia, followed by Maranhão and Ceará, in 2013 (CAMPELLO and NERI, 2013; LIMA and BARRETO, 2015). According to Oliveira (2011), the NEB always had major development problems, especially with drought, which affects it periodically.

The NEB presents three types of climate: tropical humid coastline, from the coast of Bahia to Rio Grande do Norte States; tropical, in areas of Bahia, Ceará, Maranhão and Piauí States; and semi-arid tropical, in the backwoods. This diversity of climates is due to the performance of several physical mechanisms that interact. These are responsible for the distribution of precipitation in the region, with marked annual variability between 300.0 and 2,000.0 mm, with extremely dry years and others extremely rainy, one of the main regions of South America where the signs of intra-seasonal variability are more evident. As for temperature, it presents high values with annual average ranging from 20.0°C to 28.0°C (MOLION and BERNARDO, 2002; CAVALCANTI et al., 2009).

The cluster analysis comprises a variety of techniques and algorithms whose objective is to find and separate objects in similar groups, which is the determination of the coefficient of similarity. Thus, given a sample of *n* objects, each one of them measured according to *p* variables, we search for a classification scheme that groups the objects into *g* groups, from the existence of homogeneous or heterogeneous classes or groups (BUSSAB, MIAZAKI and ANDRADE, 1990). The similarity between two samples (S_1 and S_2) can be expressed as a function of the distance between two representative points of these samples in *n*-dimensional space. In this study, Euclidean distance, one of the measures of dissimilarity among groups most used in practice (GAUCH, 1982), was used to measure the similarity or dissimilarity between the data samples, S_1 and S_2. The classifier compares the Euclidean distance of the annual mean data samples grouped by *n* variables in relation to the mean of each grouping or cluster, according to:

$$d(S_1,S_2) = \left|\sum p_i \left(S_{ij(1)} - S_{ij(2)}\right)^2\right|^{1/2} \quad (2.01)$$

where S_{1j} is a given variable, grouped two to two, S_1 and S_2 (in this study, DF cases and socio-environmental variables). According to Brower and Zar (1977), the lower the Euclidean distance between two samples, the closer they are, in terms of quantitative parameters by classes; therefore, the lower the Euclidean distance, the greater the efficiency of the procedure.

According to Mardia *et al.* (1979), the usual way to represent graphically successive merges of subgroups in a hierarchical classification method is by means of an algorithm or group tree called dendrogram. A cut in this, at any level of agglomeration, produces a classification into k subgroups ($1 \leq k \leq n$), where n is the total number of samples (KRZANOWSKI, 1988) and, in this work, *it* represents the total number of variables analyzed along with DF cases. For the hierarchization of groups, we use the Ward's linkage method in which the measure of similarity is calculated as the sum of squares between two clusters that present the smallest value in the global sum of squares within the clusters and which tend to connect to other clusters of similar sizes due to minimization of their internal variation (HAIR et al, 2005).

To validate the clustering, we used the Calinski criterion (CALINSKI and HARABASZ, 1974; OKSANEN et al. 2016). According to Vendramin, Campello and Hruschka (2010), such technique is an adjustment measure that specifies how many groups to use next to a given cluster analysis. For each value of k (number of groups), cascadeKM compares the *k-means* partitions and calculates the *Calinski value*, defined by:

$$\text{Calinski} = \frac{SSB/(k-1)}{SSW/(n-k)} \tag{2.03}$$

where n is the number of samples; k the number of classes; SSW the sum of the squares within the clusters; and SSB is the sum of the squares between the clusters, analogous to ANOVA.

To validate the clustering, we also used Spearman's (1904) rank correlation (r_s), a nonparametric version of the Pearson product-moment correlation that allows to measure the strength and direction of association between two ranked variables. To calculate r_s, one converts each variable to ranks, assigning equal ranks to any tied scores, and then uses the usual formula for r, according to Ruscio (2008):

$$r_s = 1 - \frac{6 \sum d_i^2}{N(N^2-1)} \tag{2.04}$$

where d_i are the differences in the ranked scores of corresponding variables, and N is the sample size or number of observations. In terms of the strength of relationship, the value of the correlation coefficient varies between +1 and -1. A perfect Spearman correlation of +1 or −1 occurs when each of the variables is a perfect monotone function of the other. As the value of the correlation coefficient shifts to 0, the relationship between the two variables will be weaker. The sign of the coefficient indicates the direction of the relationship (SPEARMAN, 1904; DANIEL, 1990).

Additionally, we qualitatively analyse the studied variables; RSU, specifically, from 2001 to 2015, in fact, by the availability of these data collected.

All statistical analyses were conducted using the R-Project[iv] Free Software, Version 3.0.3.

3 Results

According to DATASUS (BRAZIL, 2017a), there were 551,206 total cases of Dengue in the NEB capitals during the study period. The highest occurrences were observed in Fortaleza and Natal, respectively 33.3% (183,618 cases) and 16.5% (91,062 cases) of that total (Figure 1A). As for gender, 56.2% of those cases were predominantly female (Figure 1B). Most of the DF occurrences was in individuals between the ages of 20-39 years old (41.2%) and between 40-59 years old (19.3%); however, occurrences in individuals over 59 years old amounted to only 5.1% of cases (Figure 1C). Regarding schooling, most of this information was ignored (67.1%); however, of the records obtained, there was a DF predominance on individuals with 8 to 11 years of educational training and in the non-applicable group, which is composed by children not yet literate (Figure 1D). Most information on ethnicity was also ignored (62.9% of the total); and, from the records made (204,399), there was predominance of the disease in brown individuals (25.9%), Figure 1E.

In Figure 2, in relation to mean annual DF incidence (DF cases/100,000 inhab.), the highest rates were observed in Natal (986.1 cases/100.000 inhab.); while in São Luis the lowest rates were recorded (152.1 cases/100.000 inhab.). About basic sanitation, high indexes of water urban service (IN023) were found in the cities studied, especially in Aracaju (98.6%); however, low indexes of sewage urban service (IN024) were found, especially in Teresina and Maceió, with respectively 15.8% and 31.0%. It is noteworthy that Salvador, João Pessoa, São Luis and Fortaleza were the capitals that present the better urban sanitation service (IN024); and, except for Fortaleza, those were the capitals with the lowest DF incidences.

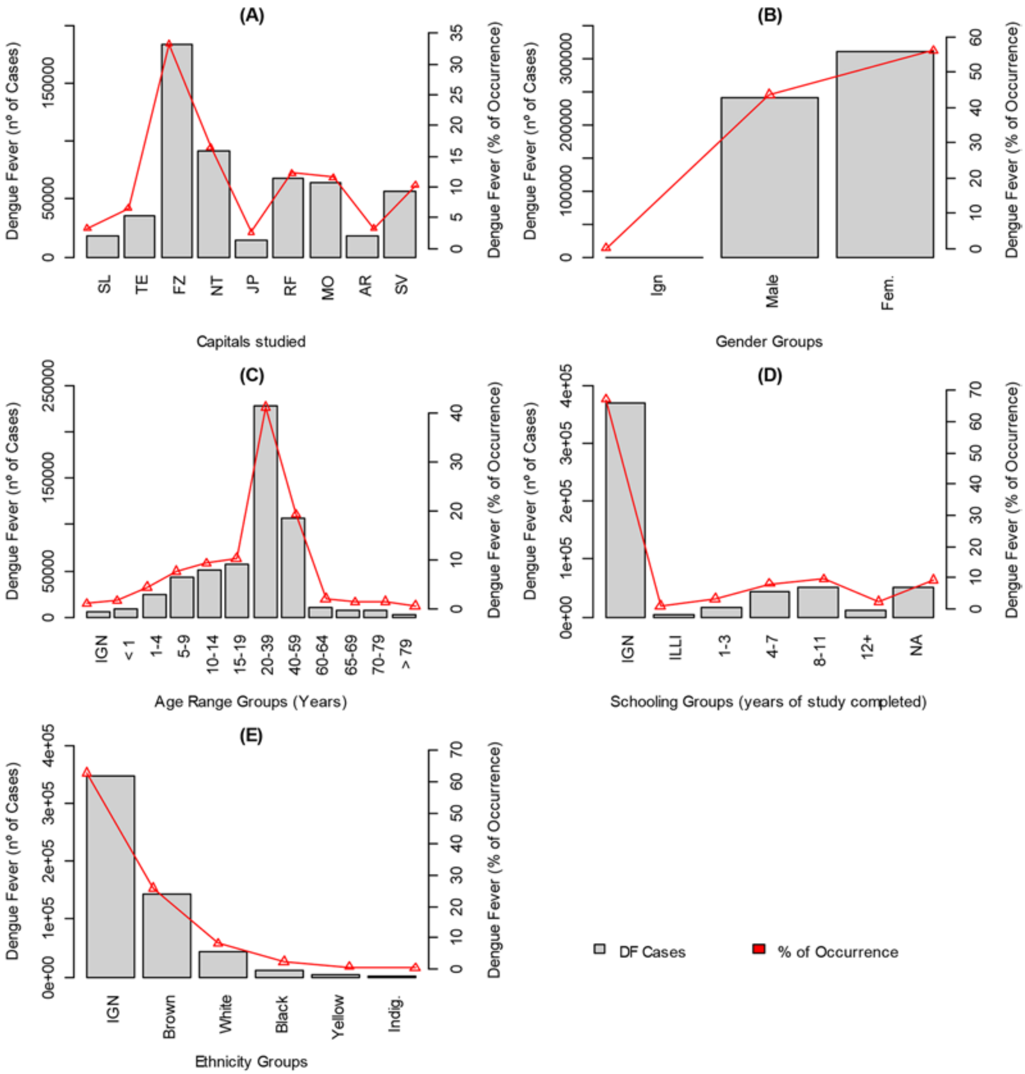


Figure 1. Distribution and percentage of occurrence of Dengue Fever (Figure 1A), and its social information: gender, age, schooling and ethnicity (respectively, Figures 1B, 1C, 1D and 1E); in the capitals of the NEB, from 2001 to 2012.

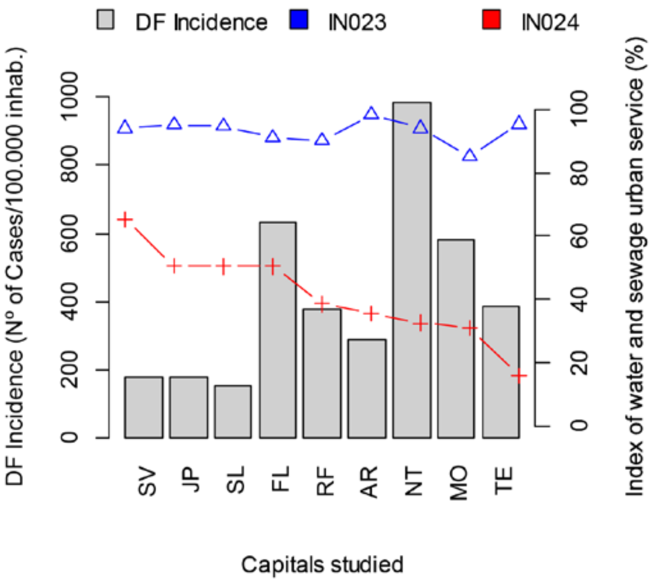


Figure 2. Association between mean annual DF incidences and indexes of sanitation urban service: water (IN023) and sewage in places already supplied with water (IN024), in the period from 2001 to 2012.

Figure 3 presents the result of the Calinski criterion, by desirable clusters distributed by K-means partitions, for different sets of variables, initially established for the dendrograms analysis of this study, according to Figure 4. The different colours shown in each group of analysed variables represent the k-means partitions found by this method used from the *n*-objects tested. In Figure 3A, we can observe the 90 class objects (DF cases + 89 indicators of RSU) distributed by the partitions, referring to the time series of the capitals distributed by the 12 years of the study. Calinski classifies the sample into two large clusters, among the eight presented, according to the red point in the graph. In Figure 3B, 70 objects (demographic, climatic variables and water and sewage indicators) were analysed by K-means, and Calinski also classified them into two larger clusters, as well as Figure 3D, with 14 objects (DF cases, and climatic and demographic variables). Already in Figure 3C, with 42 objects (DF cases with their associated social informations, climatic and demographic variables); besides the option of classification with two clusters, Calinski provides another option with four clusters, according to yellow point.

Figure 4 shows the dendrograms formed by the "ward" method, for each of the four datasets tested. The matrix indicates the distribution of similarity between each pair of objects relative to the variables contained in each set of groupings. By cutting each of the dendrograms to the desired level, the clusters formed are observed, and the red lines indicate the significant classes already classified by Calinski, according to Figure 3. In addition, it is possible to observe how the variables are related to each other, especially the DF cases in relation to the other variables of the study.

Figure 4A shows the dendrogram formed between DF cases and RSU. It can be observed that the DF cases are closely related to the expense (in R$/Km) of the public agents with the sweeping service (FN212) and to the total annual extension (Km) of the area swept by the public agent (VA010). However, DF cases are secondarily related to the total annual amount (ton/year) of RPU collected by all agent types (CO115), specifically by private agents (CO113). This secondary group is also composed by the total annual amount (ton/year) of RDO collected by the same previously mentioned agents, respectively, CO111 and CO109; as well as, the indicator of RDO mass collected per capita in relation to the population served with collection service (IN022), in Kg/inhab/day.

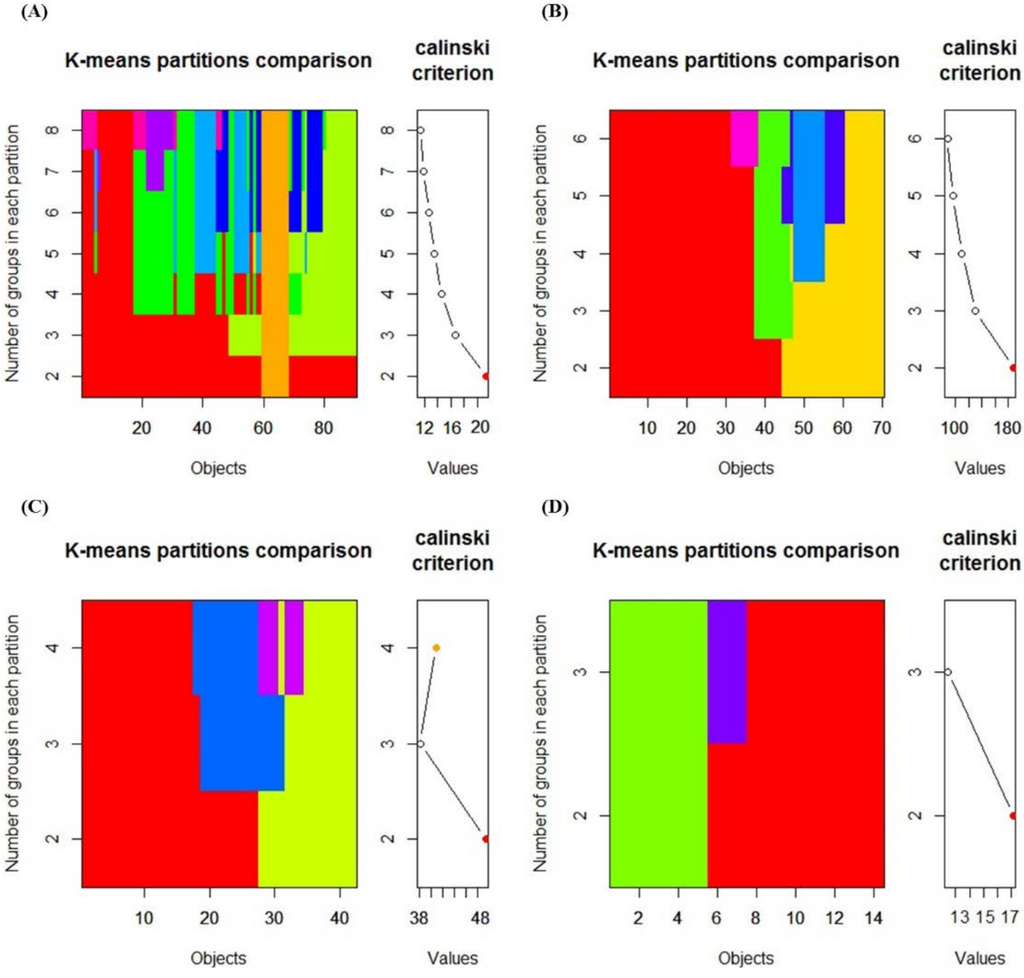


Figure 3. Distribution of groups by k-means partitions and classification by the Calinski criterion, obtained from the hierarchical cluster analysis between DF cases and variables: (A) of RSU; (B) demographic, climatic and water and sewage; (C) socio-demographic and climatic; and (D) demographic and climatic.

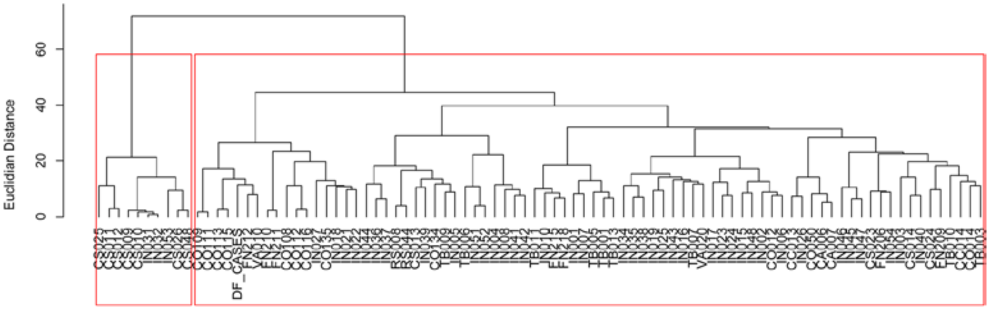
Also, in Figure 4A, in relation to group non-similar to DF cases, this cluster is isolated with ten elements. These are the indicators of selective collection (CS), specifically to recyclable materials (MR): recovery rate in relation to the total amount of RDO and RPU collected (IN031), MR per capita recovered mass in relation to urban population (IN032) and MR rate collected by selective collection in relation to the total amount collected from RDO (IN053). The lower indexes presented by these indicators for all capitals suggests disuse or lack of practice by the municipal authorities in managing or providing of these services to the community.

In Figure 4B, it can be observed that the DF cases correlate primarily with three sanitation variables: average sewage tariff (IN006), in R$/m^3; indexes of hydrometric and sewage treatment (IN009 and IN016, respectively), both in percentage. It is still observed, the correlation of DF with the extension of the sewage network per connection (IN021), in meters; density of water savings[v] per connection (IN001); sewage collection index (IN015), in percentage; total number of own employees of the prefecture in the sanitation service (FN026) and indirect operating income from water and sewage (FN004), in R$/year.

In Figure 4C, DF cases are exclusively related to social information, specifically to female gender individuals and in the age group between 15 and 19 years old. Secondarily, DF cases are related to children from 10 to 14 years old and older than 79 years old. Also, climatic factors are observed in another larger cluster, together with population size and social information about illiterate individuals and with unknown age. Already in Figure 4D, regarding the relation of the climatic and demographic factors to the DF cases, there is an intimate relation of these with the total size of urban population and, secondarily, to the vapour pressure deficit, minimum temperature, relative humidity and precipitation.

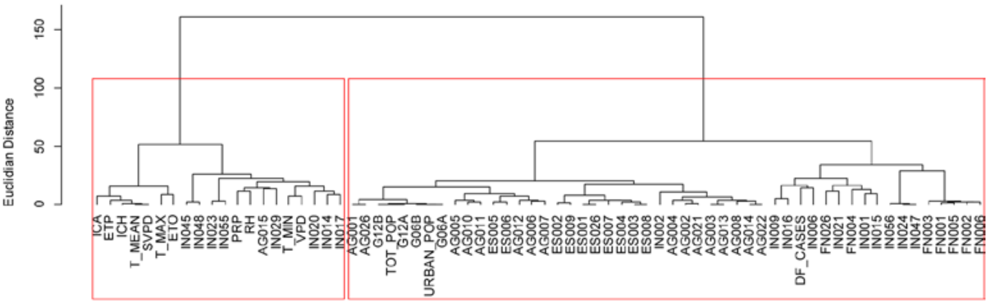
Table 1 shows the Spearman correlation coefficient (%) between DF cases and the study variables to those most related by Clustering. Social informations are the most correlated those cases, all at a significance level of 1% and, particularly, the information on the female gender, with a correlation of 99.86%. As for water and sewage, only two of eight indicators who presented association to DF have any level of significance of correlation with this one: hydrometric index (IN009), with 19.51% of correlation and p-value of 0.0431; and the extension of the sewage network per connection (IN021), with negative correlation of 23.04% and p-value of 0.0165.

(A)



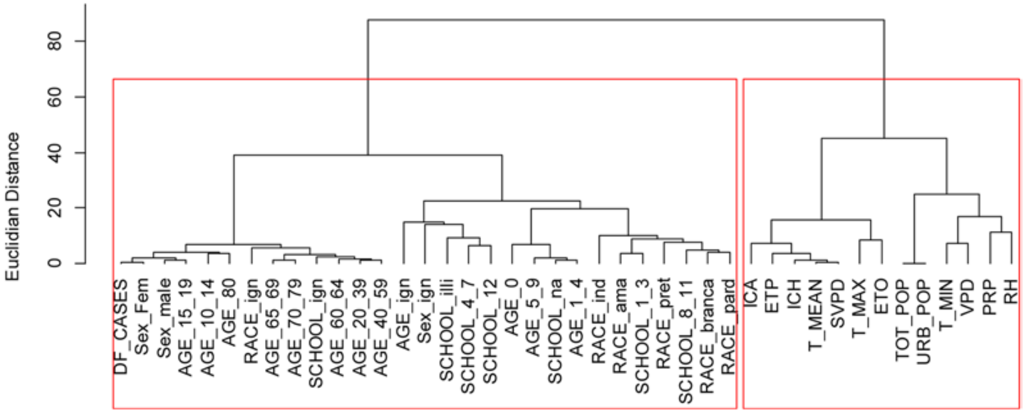
Dengue Fever Cases and Urban Solid Waste (RSU)
Ward Linkage Method

(B)



Dengue Fever Cases and Climatic, Demographic and Water and Sewage Variables
Ward Linkage Method

(C)



Dengue Fever Cases, Climatic and Sociodemographic Variables
Ward Linkage Method

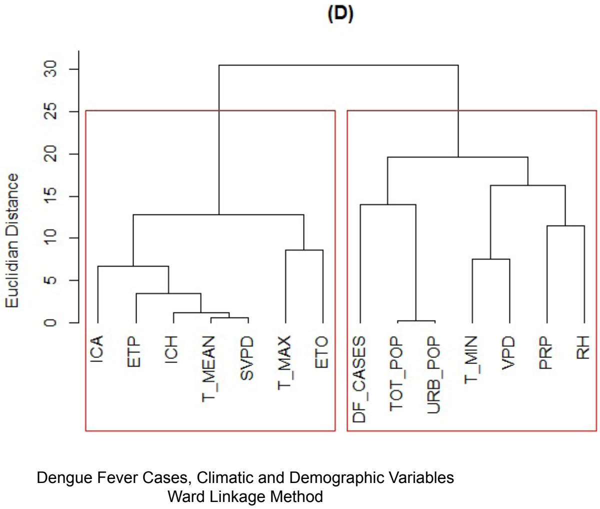


Figure 4. Dendrograms obtained from the hierarchical analysis of clusters using simple Euclidean distance and "ward" method on DF cases and variables: (A) of RSU; (B) demographic, climatic and water and sewage; (C) socio-demographic and climatic; and (D) demographic and climatic.

Still in Table 1, in relation to RSU, except for the per capita mass of RDO collected in relation to the population served by collection service (IN022), all other indicators tested were correlated and they presented a certain level of significance to DF cases, especially the quantity total annual RPU collected by private agents (CO113), with 51.35% of correlation, at a p-value ≤ 0.0001. The size of urban and total population is also correlated to DF (33.71% and 33.46%, respectively) at a high significance level. However, the climatic variables did not present a level of significance with DF, possibly because they were used here with annual values. Monthly climatic data tested in another study reiterate the correlation with DF at a high level of significance, proving that seasonality is a determining factor for epidemiological cycles.

Table 1. Spearman's correlation (percentage) with respective lower and upper confidence interval values at 95% level, in relation to the variables associated with DF cases.

Variables	Indicators	Spearman (%)	95 % CI lower	95 % CI upper	p-value	
AE	IN009	0.1951	0.0063	0.3704	0.0431	
AE	IN016	0.1100	-0.0807	0.2928	0.2573	
AE	IN001	0.1024	-0.0883	0.2858	0.2917	
AE	FN004	0.0602	-0.1303	0.2464	0.5360	
AE	IN006	0.0578	-0.1326	0.2441	0.5525	
AE	IN015	0.0190	-0.1706	0.2072	0.8455	
AE	FN026	-0.0095	-0.1982	0.1798	0.9220	
AE	IN021	-0.2304	-0.4019	-0.0433	0.0165	
RSU	CO113	0.5135	0.3594	0.6403	0.0000	*
RSU	CO115	0.4823	0.3228	0.6152	0.0001	*
RSU	VA010	0.4782	0.3179	0.6119	0.0002	
RSU	FN212	0.3980	0.2260	0.5459	0.0200	
RSU	CO111	0.2735	0.0892	0.4398	0.0042	
RSU	CO109	0.2700	0.0854	0.4367	0.0047	
RSU	IN022	0.0264	-0.1634	0.2143	0.7866	
Social	gender female	0.9986	0.9980	0.9991	0.0000	*
Social	gender male	0.9977	0.9967	0.9984	0.0000	*
Social	AGE 15-19	0.9889	0.9838	0.9924	0.0000	*
Social	AGE 10-14	0.9596	0.9414	0.9723	0.0000	*
Social	AGE > 79	0.9568	0.9373	0.9703	0.0000	*
demographic	URBAN POP	0.3371	0.1582	0.4946	0.0004	
demographic	TOT POP	0.3346	0.1554	0.4924	0.0004	
climatic	RH	0.0299	-0.1600	0.2177	0.7586	
climatic	VPD	0.0194	-0.1702	0.2076	0.8422	
climatic	T_MIN	0.0117	-0.1777	0.2002	0.9046	
climatic	PRP	-0.0600	-0.2461	0.1305	0.5376	

AE - water and sewage; * Significance level of 1%

Information from the Policy and Municipal Sanitation Plan were qualitatively analysed, in the period 2001-2015, according to SNIS (BRAZIL, 2017d). São Luis was the only capital that: 1) adopted parameters for the essential guarantee to public health; 2) defined mechanisms for participation and social control; 3) created the council or collegiate body (all items with occurrence in 2012-2013); 4) established obligations to users of basic sanitation services; 5) implemented the municipal sanitation information system (both in 2012); as well as, 6) elaborated the Municipal Sanitation Basic Plan, under the terms established in Law nº. 11,455/2007 (in 2012).

Already, in relation to the Solid Waste Management Plan, according to Law nº. 12,305/2010, which deals with the National Solid Waste Policy, the only cities to comply with this measure were Natal (adopted on June 1, 2012), Fortaleza and São Luís (both without date of approval), all individualized by each county (not inter-municipal or regional character).

It was also verified that, until 2015, no capital studied was part of an Inter-municipal Consortium, regulated by Law nº. 11,107/2005, which has, among its specific attributions, the management or provision of one or more RSU collection services (collection of household or public waste, collection of waste from health services, landfill operations, etc.). However, all these capitals declared to have a process in preparation, between the years of 2009 and 2012; but without declaring an initial procedural date, according to SNIS (BRAZIL, 2017d).

4 Discussion

As for the associations found here between Dengue and social information, da Costa et al. (2011) analysed the DF associations to gender and age group, in Amazonas State, and observed that females accounted for the majority of cases in 2008 (55.7%) and 2009 (52.3%). In relation to the age, those authors found higher DF occurrence in individuals in the age groups between 10 and 49 years old, for the study period.

The correlation observed between DF cases and IN009 index, compared to the other components of AE here studied, can be justified because water availability is one of the determinants for the maintenance of the life cycle of the mosquito-borne of the disease (VALLE, PIMENTA and CUNHA, 2015). The water supply has been identified as a primary factor in the transmission of DF (CAPRARA et al., 2009; SAN PEDRO et al., 2009).

Regarding the DF correlation presented to the IN021 index, some authors (SILVA, 2007, GIL et al., 2015) verified the proliferation of *Aedes aegypti* in sewage septic tanks. Silva (2007) concluded that this is indicative of a public health problem, since these mosquitoes may have adapted to a new environment, with tolerance to cloudy water, rich in decomposing organic material and salts. Gil et al. (2015) point to the urgency of investments and the adequacy of efforts in public sanitation systems in developing countries to assist in the control of this important mosquito-borne.

The relationships between the sweeping services (through the VA010 and FN212 indicators) and the DF cases found here contemplate one of the effective actions to combat this disease by the municipal agencies in the capitals studied, since the collection of solid urban waste promotes the decrease of possible breeding sites of *Aedes aegypti*. In a study on sanitation and arboviruses, Elmec, Bataiero e Cruz (2016) identified risk situations in sanitary inspection actions for surveillance of *Aedes Aegypti* breeding sites, among the main ones: unusable open-air materials (cans, bottles, containers and other objects that accumulate water), with an index of 29.19% and external areas in general (vacant lands, gardens, backyards, etc.), with 11.7%. According to these authors, the data (in addition to expressing the record of actions to control arboviruses) demonstrate the importance of promoting the integration of regional and municipal health surveillance and vector control services.

Grillo, Araújo and Borja (2016) investigated DF and basic sanitation in the municipalities of the State of Bahia (BA), between 2009 and 2012, and observed that DF showed a significant correlation, although weak, with water supply variables, solid waste collection and urbanization rate, with Spearman's correlation of 18.3%, 22.41% and 21.77%, respectively. Also, according to these authors, the service with the greatest coverage was water supply, with percentages higher than the state index, of 80.31% (except for the municipality of Canarana, with 77.77%); while sanitary sewage represented the service with the most deficient coverage, with a state percentage of 55.99%. They also verified that in Irerê city there was one of the largest coverages of water supply service by network (98.25%) and of solid waste collection (88.72%); as well as, it was the municipality with the highest coverage of sanitary sewage services (28.28%), among the five studied with higher DF incidence. The results found by these authors corroborate with the values here presented between DF cases and those variables of sanitation correlated.

It is important here to emphasize the normative role of Directive SNCC nº. 3 (BRAZIL, 2016a), which guides the states and municipalities to promote permanent and emergency actions of basic sanitation that contribute to the elimination of *Aedes aegypti* breeding sites, ensuring, among others, the uninterrupted supply and adequate domestic storage of water, as well as, the regular collection and disposal of solid wastes. Feitosa, Sobral e Jesus (2015) propose that the selection of socio-environmental indicators can help in the analysis of the infestation rates of the DF mosquito-borne in the regions studied, serving as a subsidy in the implementation of public policies of prevention and control of the disease. Furthermore, the use of health planning indicators is an important tool for evaluating and monitoring public policies in progress, thus helping the work of management and giving transparency to their actions. According to the authors, public authorities should ensure substantial investments in environmental sanitation and community awareness, aiming for behavioural changes that can corroborate to eliminate breeding sites for larvae and mosquitoes from infectious diseases, specifically DF.

According to Valle, Pimenta and Cunha (2015), DF subsists, in the part that corresponds to people, for the absence of vector control behaviours among those who could participate in source reduction practices. Therefore, the refusal to participate in such practices is equivalent to the finding that certain preventive behaviour, which is considered necessary for the reduction of the source, do not occur. Above all, accepting the responsibility for such a reduction implies not blaming populations at risk of the disease, but producing behavioral change by creating and evaluating educational strategies and processes that encourage citizens to so.

The relationship found here between DF and water and sewage services (e.g., IN009 and IN021) is in part justified by San Pedro et al. (2009). According to these authors, the precarious provision of the water distribution service stimulates the adoption of storage practices, which, in turn, may promote the formation of conditions conducive to mosquito-borne breeding. When considering indicators that do not reflect the inequality of relationships between economic and housing resources (for example, the analysis of households connected to the general water distribution network), the tendency is to find no correlation between DF and environmental conditions. At the same time, the trend of inverse correlation may be a consequence of the possibility that social groups with greater economic power accumulate excess water and more permanently in places that present a precarious supply, or even for leisure activities such as untreated pools, which may be potential outbreaks of DF mosquito-borne.

5 Conclusion

The study of DF is still complex because it involves several associations; however, this paper served to elucidate the degree of relationship between this disease and the variables presented here, especially the indicators of basic sanitation and urban solid waste.

Among the capitals studied, Natal and Fortaleza stood out because they presented the highest incidences of the disease. DF mainly affects brown individuals of the female gender in the age group of 20 to 39 years old, but the cluster analysis showed association of the disease in individuals between 10 and 19 years old, as well as in individuals over 79 years old. The high level of significance (p-value < 0.0001) presented by Spearman confirms the strength of the correlation for these age group tested in relation to Dengue by clustering.

High correlation (p-value ≤ 0.0001) was observed between Dengue Cases and the annual total of urban public waste (RPU) collected by all types of agents, especially private agents; as well as in relation to the total extent of area swept by public agents (respectively, the indicators CO115, CO113 and VA010). This suggests the effective promotion of actions by the municipal agencies to improve the sanitation service, with consequent control and combat to vectors transmitting infectious diseases, since these services remove from the environment possible *Aedes aegypti* breeding sites. Therefore, studies should be done to verify these collection actions "*in situ*".

Correlations were also observed between dengue and basic sanitation indicators, especially by indexes of hydrometric and sewage treatment (IN009 and IN016, respectively). It is suggested that a greater distribution (or consumption) of water (and, consequently, of sewage) provide a higher occurrence of this disease. Appropriate studies are needed to verify whether these correlations are due to irregular or regular use of water.

As for climatic variables, due to the low p-value presented in correlations with Dengue, such associations were discarded for this study. It should be noted here that the annual time scale was used in this study, which would possibly justify these low associations.

As noted, the issue of water supply, sanitary sewage and urban solid waste collection is a reality present in the epidemics of dengue, according to the associations presented here.

São Luis was the only capital, among those studied, to present (until 2015) a solid waste management plan (Law nº. 12,305/2010), with an effective date of approval. The establishment of these public policies and actions in sanitation aim to guarantee and improve the quality and well-being of the population, which can promote the improvement of public health and, consequently, the reduction of infectious diseases. In fact, that is the capital of NEB with the lowest incidence of DF. In view of this, such public policies could be better studied to examine their potential benefits to the population, especially related to DF.

Finally, this study criticizes the precariousness of access to public data information (Law nº. 12,527/2011) practiced by the federal government (by its Agencies, Institutes and Ministries), in view of the amount of missing or incomplete data collected from SNIS, IBGE and DATASUS (without historical series from 2013; and with many of the existing ones classified as "ignored", such as, for example, ethnicity, schooling, information about pregnant women). Unfortunately, we are faced with a persistent structural problem throughout the history of information systems (MORAES and SANTOS, 1998) with no prospect of a short-term solution; which poses problems for the effectiveness of national public policies such as dengue prevention and control.

Notes

i Public policy is "any systematic set of interrelated actions and procedures, publicly adopted by governmental authority for the purpose of dealing routinely with some specific subject" (MACHADO, 2014).

ii The Ministry of Health uses "race" to establish genetic difference between humans, when it is challenged by scientific research that demonstrates that there is no well-defined genetic difference between races, as occurs, for example, with chimpanzees (YUDELL et al, 2016).

iii Statistical software for data analysis add-on for Excel. Available in: <https://www.xlstat.com/> Access in: April 17, 2017.

iv Available in: <https://cran.r-project.org/bin/windows/base/old/3.0.3/>. Access in: January 05, 2017.

v This is the number of building installations, according to SNIS (BRAZIL, 2017d).

References

ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. (1998). Crop Evapotranspiration (guidelines for computing crop water requirements). **Fao Irrigation and Drainege**, 56, 297p.

ARACAJU (2014). Secretaria Municipal de Saúde. **Informe da Vigilância Epidemiológica**. Sergipe.

BRAZIL (2007). **Lei Federal de Saneamento Básico**. Lei nº. 11.445, de 05 de janeiro de 2007. Available in: <http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/lei/l11445>. Access in: April 27, 2017.

______. (2010). **Censo Demográfico de 2010**. Instituto Brasileiro de Geografia e Estatística, IBGE. Ministério do Planejamento, Desenvolvimento e Gestão.

______. (2014). **Prevenção e Resposta à Introdução do Vírus Chikungunya no Brasil**. Ministério da Saúde. Brasília/DF, p.100.

______. (2016a). Sala Nacional de Coordenação e Controle para o enfrentamento à Dengue, ao vírus Chikungunya e ao vírus Zika. **Diretriz SNCC nº. 3 – Saneamento**

Básico. Available in: <http://combateaedes.saude.gov.br/images/sala-de-situacao/diretriz--sncc-n-3-saneamento-basico.pdf>. Access in: April 27, 2017.

______. (2016b). **Estimativas de população brasileira**. Brasília, DF: Diário Oficial da União, 30 de agosto de 2016. IBGE. Ministério do Planejamento, Desenvolvimento e Gestão.

______. (2017a). **Sistema de Informação de Agravos de Notificação – SINAN**. Ministério da Saúde. Available in: <http://portalsinan.saude.gov.br/sinan-net/43-institucional>. Access in: February 01, 2017.

______. (2017b). **Dados históricos do Banco de Dados Meteorológicos para Ensino e Pesquisa – BDMEP**. Instituto Nacional de Meteorologia – INMET. Ministério da Agricultura, Pecuária e Abastecimento. Available in: <http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmep>. Access in: February 01, 2017.

______. (2017c). **Sistema IBGE de Recuperação Automática - SIDRA**. Instituto Brasileiro de Geografia e Estatística, IBGE. Available in: <https://sidra.ibge.gov.br/home/pnadcm>. Access in: February 01, 2017.

______. (2017d). **Sistema Nacional de Informações sobre Saneamento - SNIS**. Ministério das Cidades. Available in: <http://www.snis.gov.br/aplicacao-web-serie-historica>. Access in: February 01, 2017.

BROWER, J. E.; ZAR, J. H. (1977). **Field and laboratory methods for general ecology**. 2 ed. Dubuque: Wm. C. Brown Publishers, 226 p.

BUSSAB, W. de O.; MIAZAKI, S. E.; ANDRADE, D. F. (1990). **Introdução à análise de agrupamento**. Simpósio Brasileiro de Probabilidade e Estatística, 9, São Paulo.

CALINSKI, T.; HARABASZ, J. (1974). A dendrite method for cluster analysis. **Communications in Statistics-theory and Methods**, 3(1):1-27.

CAMPELLO, T.; NERI, M. C. (Orgs.) (2013). **Programa Bolsa Família:** uma década de inclusão e cidadania. Brasília: Ipea.

CAPRARA, A.; LIMA, J. W. de O.; MARINHO, A. C. P.; CALYASINA, P. G.; LANDIM, L. P.; SOMMERFELD, J. (2009). Irregular water supply, household usage and dengue: a bio-social study in the Brazilian Northeast. **Cadernos de Saúde Pública**, 25(Sup. 1):125-136.

CAVALCANTI, I. F. de A.; FERREIRA, N. J.; SILVA, M. G. A. J. da; DIAS, M. A. F. da S; Organizadores (2009). **Tempo e Clima no Brasil**. São Paulo, SP: Oficina de Textos.

DA COSTA, A. G.; DOS SANTOS, J. D.; DA CONCEIÇÃO, J. K. T.; ALECRIM, P. H.; CASSEB, A. A.; BATISTA, W. C.; HECKMANN, M. I. O. (2011). Dengue: aspectos epidemiológicos e o primeiro surto ocorrido na região do Médio Solimões, Coari, Estado do Amazonas, no período de 2008 a 2009. **Revista da Sociedade Brasileira de Medicina Tropical**, 44(4):471-474.

DANIEL, W. W. (1990). **Spearman rank correlation coefficient**. Applied Nonparametric Statistics (2nd ed.). Boston: PWS-Kent. pp. 358-365.

ELMEC, M. A.; BATAIERO, M. O.; CRUZ, M. G. B. da (2016). Saneamento do meio, arboviroses e as estratégias de Vigilância Sanitária para combate aos vetores no Estado de São Paulo. **BEPA**, 13(153-154):63-68.

FEITOSA, F. R. S.; SOBRAL, I. S.; JESUS, E. N. de (2015). Indicadores Socioambientais como subsídio à prevenção e controle da Dengue. **Revista Eletrônica em Gestão, Educação e Tecnologia Ambiental**, 19(3):351-368.

GAUCH, H. G. (1982). **Multivariate analysis in community ecology**. Cambridge: Cambridge University Press.

GIL, L. H. S.; KATSURAGAWA, T. H.; LIMA, A. A. de; TADA, M. S.; OZAKI, L. S.; JULIÃO, G. R. (2015). Rudimentary cesspits as breeding sites for *Aedes aegypti* in urban areas of Northern BRASIL. **Rev Pan-Amaz Saúde**, 6(3):73-80.

GRILLO, R. B. P.; ARAÚJO, L. M.; BORJA, P. C. (2016). **Fatores intervenientes na ocorrência de Dengue nos municípios da Bahia**. IV Congresso Baiano de Engenharia Sanitária e Ambiental.

HAIR, J. F., et al. (2005). **Análise multivariada de dados**. 5 eds. Porto Alegre: Bookman.

KRZANOWSKI, W. J. (1988). **Principles of Multivariate Analysis - A User's Perspective**. Oxford Science Publications [A user-friendly introduction to modern multivariate statistical analysis].

LIMA, S. R. R.; BARRETO, R. C. S. (2015). **O comportamento socioeconômico da Região Nordeste:** do meio século XX ao século XXI. XI Encontro de Economia Baiana.

MACHADO, C. J. S. (2013). **Animais na sociedade brasileira**. Rio de Janeiro, RJ: E--Papers.

______. (2014). **Desenvolvimento Sustentável para o Antropoceno**. Rio de Janeiro, RJ: E-Papers.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. (1979). **Multivariate Analysis**. Academic Press.

MOLION, L.C.B.; BERNARDO, S. O. (2002). Uma revisão da dinâmica das chuvas no Nordeste Brasileiro. **Revista Brasileira de Meteorologia**, 17(1):1-10.

MORAES, I. H. S.; SANTOS, S. R. F. R. (1998). Informação em Saúde: Os Desafios Continuam. **Ciência & Saúde Coletiva**, 3(1): 37-51.

OLIVEIRA, M. L. L. (2011). **O Brasil de JK, A invenção do Nordeste**. CPDOC FGV. Available in: <http://cpdoc.fgv.br/producao/dossies/JK/artigos/Economia/Nordeste>. Access in: April 27, 2017.

OKSANEN, J.; BLANCHET, F. G.; KINDT, R.; LEGENDRE, P.; MINCHIN, P. R.; O'HARA, R. B. et al. (2016). **Vegan:** Community Ecology Package. R package version 2.3-5.

ROSENBERG, N. J.; BLAND, B. L.; VERMA, S. B. (1983). **Microclimate:** The Biological Environment. New York: John Wiley & Sons, 467p.

RUSCIO, J. (2008). Constructing Confidence Intervals for Spearman's Rank Correlation with Ordinal Data: A Simulation Study Comparing Analytic and Bootstrap Methods. **Journal of Modern Applied Statistical Methods**, 7(2): 416-434.

SAN PEDRO, A.; SOUZA-SANTOS, R.; SABROZA, P. C.; OLIVEIRA, R, M. de. (2009). Condições particulares de produção e reprodução da dengue em nível local: estudo de Itaipu, Região Oceânica de Niterói, Rio de Janeiro, Brasil. **Cadernos de Saúde Pública**, 25(9):1.937-1.946.

SARDÃO, T. (2016). **Dengue e Zika: ABES alerta sobre a importância do saneamento básico no combate às doenças**. Associação Brasileira de Engenharia Sanitária e Ambiental – ABES. Available in: <https://abes-dn.org.br/?p=923>. Access in: April 28, 2017.

SILVA, E. (2007). **Reprodução do *Aedes aegypti* em fossas de esgotamento sanitário no bairro dos Pioneiros, zona urbana de Pimenta Bueno – Rondônia, Amazônia Ocidental.** Dissertação de Mestrado, Universidade de Brasília. Brasília, Brasil.

SPEARMAN, C. S. (1904). The proof and measurement of association between two things. **American Journal of Psychology**, 15:72-101.

STEADMAN, R. G. (1979). The assessment of sultriness: part I: A temperature-humidity index based on human physiology and clothing science. **J. Appl. Meteor.**, 18:861-884.

THORNTHWAITE, W. C. (1948). An approach toward a rational classification of climate. **Geographycal Review**, 38 (1): 55-94.

VALLE, D; PIMENTA, D. N.; CUNHA, R. V. da (2015). (orgs.) **Dengue:** teorias e práticas. Rio de Janeiro, RJ: Ed. Fiocruz.

VENDRAMIN, L., CAMPELLO, R. J. G. B.; HRUSCHKA, E. R. (2010). Relative Clustering Validity Criteria: A Comparative Overview. **Statistical Analysis and Data Mining**, 3:209-235.

YUDELL, M.; ROBERTS, D., DE SALLE, R.; TISHKOFF, S. (2016). Taking race out of human genetics. **Science**, 351(6273):564-565.

ZARA, A. L. de S. A; SANTOS, S. M. dos; FERNANDES-OLIVEIRA, E. S.; CARVALHO, R. G.; COELHO, G. E. (2016). Estratégias de controle do *Aedes aegypti*: uma revisão. **Epidemiol. Serv. Saúde**, 25(2):391-404.

Submitted on: 15/05/2017
Accepted on: 02/10/2018
http://dx.doi.org/10.1590/1809-4422asoc0133r2vu18L4TD
2018;21:e01332
Featured Topics

ASSOCIATIONS BETWEEN DENGUE AND SOCIO-ENVIRONMENTAL VARIABLES IN CAPITALS OF THE BRAZILIAN NORTHEAST BY CLUSTER ANALYSIS

***Resumo*:** Este artigo busca identificar, nas capitais do Nordeste brasileiro, o nível de similaridade entre dengue e variáveis climáticas, sociodemográficas e de saneamento, entre 2001-2012, através da análise de agrupamentos, uma técnica explanatória usada sobre dados multivariados para verificar o inter-relacionamento entre grupos formados pelas distâncias similares entre seus componentes. Os resultados, validados por Spearman, mostraram alta correlação (p-valor ≤ 0,0001) da dengue com: indivíduos do gênero feminino entre 10-19 anos e acima de 79 anos; o total anual de resíduos públicos urbanos coletados por todos os tipos de agentes, especialmente os agentes privados; a extensão total de área varrida por agentes públicos, sugerindo que um maior volume de resíduos propicie uma maior ocorrência de criadouros de mosquitos-transmissores da doença. Houve também certa correlação junto aos índices de tratamento de água e de esgoto, sugerindo estar relacionado à manutenção do ciclo de vida do mosquito-vetor devido à disponibilidade de água.

***Palavras-chave*:** *Aedes aegypti*; Análise de Agrupamentos; Distância Euclidiana; Esgotamento Sanitário; Resíduos sólidos urbanos

***Abstract*:** This article aims to identify the level of similarity between dengue and climatic, sociodemographic and sanitation variables in Brazilian Northeast capitals between 2001 and 2012, by cluster analysis, an explanatory technique used on multivariate data to verify the interrelationship between groups formed by the similar distances among its components. The results, validated by Spearman, showed high correlation (p-value ≤ 0.0001) of dengue with: female subjects between 10-19 years old and over 79 years old; the annual total of urban public waste collected by all types of agents, especially private agents; and the total extent of area swept by public agents, suggesting that a higher volume of residues would lead to higher occurrence of mosquito breeding sites transmitting the disease. There was also some correlation with the indexes of water and sewage treatment, suggesting that it is related to the maintenance of the mosquito-borne life cycle due to the water availability.

***Keywords*:** *Aedes aegypti*; Cluster Analysis; Euclidian Distance; Sanitary Sewage; Urban Solid Waste

***Resumen*:** Este artículo busca identificar, en las capitales del Nordeste brasileño, el nivel de similitud entre dengue y variables climáticas, sociodemográficas y de saneamiento, entre 2001-2012, por análisis de agrupamientos, una técnica usada sobre datos multivariados para verificar la interrelación entre grupos formados por distancias similares entre sus componentes. Los resultados, validados por Spearman, mostraron una alta correlación (p-valor $\leq$ 0,0001) del dengue con: individuos del género femenino entre 10-19 años y mayores de 79 años; el total anual de residuos públicos urbanos recogidos por todos los tipos de agentes, especialmente los agentes privados; y la extensión total de área barrida por agentes públicos, sugiriendo que un mayor volumen de residuos propicie mayor ocurrencia de criaderos de mosquitos-transmisores de la enfermedad. También hubo cierta correlación junto a los índices de tratamiento de agua y alcantarilla, sugiriendo estar relacionado al mantenimiento del ciclo de vida del mosquito-borne, debido la disponibilidad del agua.

***Palabras-clave*:** *Aedes aegypti*; Análisis de Agrupamientos; Distancia Euclidiana; Agotamiento Sanitario; Residuos sólidos urbanos