v. 2, n. 7, 2022

Journal of
Engineering
Research

BIG DATA AND SYSTEMIC GEOINTELLIGENCE: PLATFORM FOR PROPOSING AND ANALYZING OPERATIONAL BASE COMPOSITION SCENARIOS[1]

Eduardo de Rezende Francisco
Professor and Head of the Department of Tech-nology and Data Science at FGV EAESP. Consultant at GisBI, a geoinformation think tank and Big Data Analytics

André Insardi
Professor and master in Consumer Behavior at ESPM. Consultant at GisBI, a geoinformation think tank and Big Data Analytics.

Rubens de Almeida
Engineer, journalist and master's student in Management and Public Policy at the FGV. Consultant at GisBI, a geoinformation think tank and Big Data Analytics
EAESP

Ricardo Maciel Gazoni
PhD in Intelligence Technologies and Digital Design from PUC-SP. Consultant at GisBI, a geoinformation think tank and Big Data Analytics

1. This work was developed within the scope of the Research and Techno-logical Development Program for the Electric Energy Sector regulated by ANEEL and is included in the Annals of the IX Congress of Technological Innovation in Electric Energy (IX CITENEL). This study was awarded as the 2nd best project in the Utilities category (Energy, Sanitation, Communica-tion) of the World Congress GEO#Connect 2016.

 DOI 10.22533/at.ed.317272210056

Abstract: How will public services be operated in 3 years? And 30 years from now? How many operational bases, in which locations? This article presents the study of the influence of geographic variables and optimization models on the key performance indicators (KPIs) of AES Eletropaulo's operation in answering emergency calls. Within an innovative geographic Big Data perspective, a systemic geointelligence platform was created, and four informational perspectives were used: urban dynamics (traffic, micro-climate, trees), real estate dynamics, density of occurrences and history of displacement of the teams in the field. . Four algorithms were developed (brute force, Hillclimbing, genetic and K-means) from a computational modeling to allow the simulation of scenarios of base composition and the impact on the company's indicators (TMA in particular). The results and the analytical platform allow the realization of evaluations to become immediate and support better decisions for the distributors.
Keywords: Big Data, Operational Base Efficiency, Systemic Geointelligence, Asset Management, Geographic Optimization.

INTRODUCTION

The present work, result of the R&D project ANEEL nº 0390-1076/2014 – "Study for Proposition and Analysis of Scenarios of Composition of Operational Bases through Systemic GeoIntelligence Platform" studies the influence of geographic variables on key indicators of performance (KPI's – Key Performance Indicators) of AES Eletropaulo's operational area in answering emergency calls in the region covered by its energy distribution concession in the Metropolitan Region of São Paulo.

The focus of the study is to identify the influence of the location of the operational bases in the operation of emergency attendance and to verify the consequences of eventual alterations of address of these bases cause in their performance indicators.

It is important to highlight that the current location of the bases is the natural consequence of the historical evolution of the provision of energy distribution services in the city of São Paulo, and, evidently, did not take into account the present-day metropolitan geographic situation, but the intersection of opportunities for real estate (land) to exist. available and the sensitivity of the operation planning area.

The analysis and conclusions of this work took into account the extensive database made available by AES Eletropaulo and relied on the vision of the concessionaire's professionals to guide the company's modus operandi. Statistical models for prediction, evaluation, analysis and validation were created, which served as a basis for the construction of a software that will be available to AES Eletropaulo for future consultations and support for strategic decisions. The innovation is that after making this software available, the concessionaire's planning area now has an unprecedented geospatial view of relevant variables in the concession area, precisely to verify the consequences of urban changes on the performance of its teams.

The most surprising conclusions of the work are the decisive influence of traffic variables on operational performance and the relationship between the geographic location of the bases and the road network. In particular, there are possible immediate gains in the significant improvement in the performance of the service time of the first occurrence of the day, which is the occurrence impacted by the location of the base, for each team, through the simple redefinition of the coverage area of each current base.

In the long term, as the software is used by the planning teams, the tool developed will enable the drastic adjustment of the number of operational bases, making available important fixed assets of the company, being able to even modify logistic procedures of positioning of company vehicles and equipment.

As the urban dynamics of cities involves evolutionary phenomena – occupation profile of neighborhoods, volume of new real estate developments, increase in the number of vehicles, opening of new roads, among others - the developed software is able to incorporate in its calculations other magnitudes that represent the urban complexity and can, therefore, be used to, over time, verify if such changes are capable of influencing the flow of traffic, generating unforeseen difficulties in meeting the service times.The results achieved also make it possible to verify whether the number of call centers and their positioning in the urban fabric are the most adequate at each moment in the evolution of the urban characteristics of the Metropolitan Region of São Paulo. Such an equation is complex enough to suggest that the conclusions are exposed not only by the simple identification of the best addresses to position the Service Bases, but also by the generation of estimates of service times, in order to offer elements for decision making. and when the concessionaire must review this position, due to possible changes.

To make this possible, all the work of analyzing variables, developing evaluation criteria and exposing results uses a platform of digital maps, introducing the concepts of geographic intelligence and Spatial Statistics in the decision process. Digital maps help analysts to better understand the evolution of changes in operational conditions in the urban space and offer the possibility to quickly perform complex spatial calculations, which consider situations of streets and roads with different traffic capacities, displacement speeds in different conditions of time and days of the week. Combined with climate change, local afforestation density and real estate value of the regions, the software uses geo-graphic algorithms that indicate, in each situation, the best locations to position the operational bases, today and in the future.

The developed software allows the realization of evaluations and decision making for the repositioning of the operational bases to become immediate and common, as long as access to the databases for monitoring traffic, climate, afforestation characteristics and real estate value used for the construction of the geographic-digital model of analysis.

RESEARCH DEVELOPMENT

II.1. Project Objectives and Product

In general terms, the objective and product of the project are highlighted in Figure 1, below.

The integration of existing KPIs traditionally monitored by AES Eletropaulo, as well as the adaptation and eventual creation of new indicators, linked to real estate performance and its idiosyncrasies, is an intrinsic part of this work. The use of this platform will be destined to the analysis of Predictive Models of the KPIs in a flexible model and capable of dynamic and continuous insertion of variables to monitor the growth of the demand for assets and the need for operational bases that satisfy the various conditions defined in the energy distribution rules.

Goal	Develop a methodology based on statistical models, operational research and data analysis techniques that that allows proposing an optimized configuration of AES assets without interfering with the operational performance of technical and commercial services in the field
Project Product	Analytical Platform equipped with tools for exploring various scenarios composed of the possible structuring of operational bases (quantity/location/dimension) and the offer of field services, which will allow the concessionaire to replicate the analyzes carried out systematically, optimizing performance geographic location of its operational bases

Table 1. Project Objectives and Product

II.2. Methodology

Figure 1 presents the macro-steps involved in the development of the methodology. The following subtopics describe each step in greater detail.

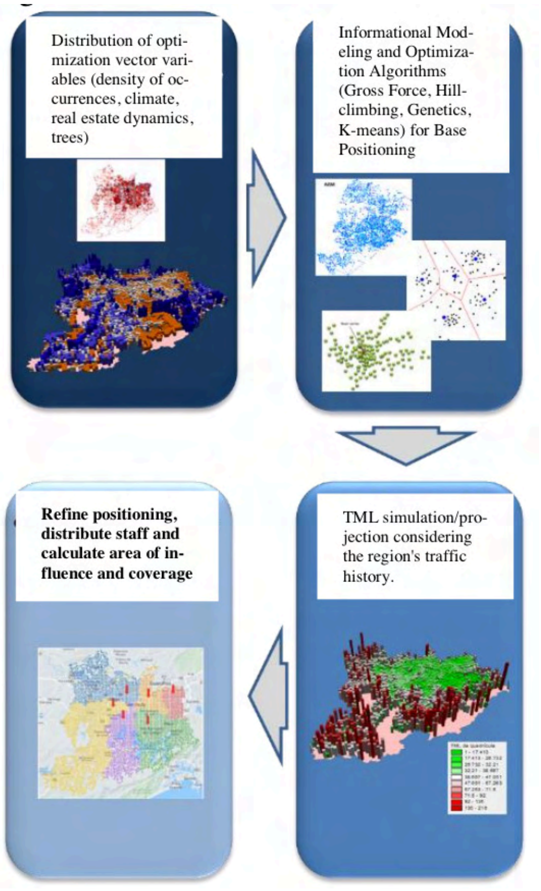


Figure 1. Stages of Big Data Methodology Development and Systemic Geointelligence

II.2.1 Survey of Variables, Analysis and Use of Data

The entire development of linked ideas and solutions was based on the availability, by AES Eletropaulo, of an operational database with more than 1.2 million records of emergency care provided by Eletropaulo's teams in its service region, among the period from December 2012 to November 2015, provided to the team of R&D in data tables that configured a total of 12 million data (variables and observations) to be analyzed, with the following details: Emergency location, Service request registration time, Departure time of the team, Time of arrival of the team at the place of service, Responsible team, Equipment, and Type of occurrence.

Several other service processes were incorporated into these data, in an initial phase called "Proof of Concept", in which occurrences of other services, especially Reading and Delivery, Fraud Inspection and Technical Notes, were analyzed.

The recorded data of the occurrences in the field, in particular the times of registration, departure and arrival, were compared with consistent traffic data from the external database GisBI/Here. The verification of the consistency of the data collected/annotated by the field teams was then carried out by reading the displacement times of vehicles (light and heavy) in the urban region, consolidated by traffic monitoring systems in the metropolitan region of São Paulo. The work used as reference data the information processed and worked by GisBI from the data of Here North America, LLC, in its service called "Traffic Patterns", originated from the systematic monitoring of vehicles equipped with GPS and other types of traffic controls. displacements, whose use license during the R&D development period was integrated into the contracted values.

The traffic base used in this project has a minimum history of 3 years, constantly updated (every 15 minutes), as input for the identification (and updating) of the typical hourly traffic of each road (from and to), with hourly averages for each typical day of the week configure 13,171 different patterns of average hourly commuting by day of the week and categories of functional pathways.

For the preliminary comparison, the displacements between the operational bases and the first occurrences registered for each team in a period of one year (from December 2012 to November 2013) were considered, in a total of 111,912 successes, and for each Comparison was calculated by two criteria (shortest way and fastest way), available in the GisBI/Here database.

II.2.2 Variable Enrichment and Conceptual Modeling

Based on the conclusions of the debates and methodological definitions of the previous stage, and to make possible the aggregation of the several possible variables to the model to be proposed for the location of the operational bases, the team started to dedicate itself to the detailing and enrichment of the databases, work that demanded a special dedication, so that all the bases to be integrated in the methodology kept geographic spatial coherence with the original data and, at the same time, allowed their mathematical integration of different magnitudes, conveniently described in the stage of the effective development of the computational model of geospatial analysis.

Based on previous studies ([1], [2]), the feasibility of concentrating large volumes of data in squares was verified as a possible simplification of the geographic distribution of databases of different magnitudes and meanings.

With this possibility of simplification on screen, the researchers considered, however, that the problem of locating AES Eletropaulo's operational bases in the defined concession region involved not only the issue of the distribution of variables by the territory (and its geography) , but by observing the travel times along the real roads (streets and avenues), characterized by traffic data in the region, coming from the GisBI/Here database. In other words, it required the characterization of each (and all) of the squares as the origin and destination of displacements from all the other squares.

This combination of information and the necessary iteration of data from different sources and metrics, they realized, made solving the problem an example of "NP - Complete", which in computational complexity theory means a special level of difficulty to solve. -nar the problem in a polynomial time.

This motivated the group's initiative to seek aggregated representations and computable algorithms to solve the search for the optimal positioning of the bases, within the grid of grids, in order to generate conditions for decision-making under operational conditions, by the planning teams of the AES Eletropaulo.

II.2.3 Development of the Spatial Analysis Algorithm

The title adopted for this R&D proposes the expression "systemic geo-intelligence" (Study for Proposition and Analysis of Scenarios of Composition of Operational Bases through Systemic Geo-intelligence Platform) as an innovation of the process of analysis and operational planning within a energy distribution concessionaire [3].

Innovation does not only aim at modifying and/or increasing the theoretical and practical instruments of professionals involved in these tasks within energy distribution companies with just one more geographic technology.

But transforming all the complexity of spatial analysis into an instrument simple enough to be adopted in the daily practice of operational planning of this type of public service, incorporating the analytical possibilities of geographic platforms, made accessible very recently, thanks to the rapid development of maps. digital devices and their widespread use by all people, notably through applications linked to the planning of routes in city traffic, among others ([4], [5], [6], [3]).

Companies that already used maps and graphics in their more sophisticated strategic analysis now have the chance to develop new planning routines based on the intensive use of spatial reasoning on simplified platforms that do not require training. specific training in GIS software. The assumption is that from this work it becomes possible to disseminate the geospatial culture among analysis and operational planning professionals, with easy-to-use interfaces. ([7], [8]).

II.2.4 Urban Conditions

The initial hypothesis was that some urban characteristics would need to be considered in the formulation of a mathematical-statistical analysis algorithm capable of recognizing, at any time, the influence of changes in the concession area, in particular, of the conditions in:

- Traffic on the displacement routes of emergency response vehicles
- Times of identification and access to occurrences
- Local afforestation pattern (interference with aerial cabling)
- Intensity of atmospheric phenomena (rain, wind and electrical discharges)
- Real estate vectors (value of the land of the Operational Bases and urban occupation trends)

In terms of traffic, it was necessary to observe more slowly the times recorded by the field teams for the displacement of their vehicles to the places of the occurrences. This work was carried out in the preliminary phase of this work, when a small portion of inconsistencies was detected from a sample of 111,912 displacements of teams, certainly derived from the poor completion of the operational forms, which is fully understandable, although undesirable for performance checks, given the emergency conditions and sometimes urgency of these calls.

II.2.5 Data Volume versus Processing Capacity

After this initial cleaning of the base of real travel times - between the base address and the address of the occurrence, the challenge was to promote the verification of travel times, in these same routes, offered by the GisBI/Here platform, whose data provide traffic patterns that impact travel times and that, in theory, could be measured every 15 minutes, 7 days a week.

This work of comparisons of both travel times, for each specific emergency care, was carried out in the preliminary phase and indicated a strong correlation, offering tranquility to adopt the GisBI/Here bases as a reference for the production of al- final decision support algorithm and planning for the best location of operational bases.

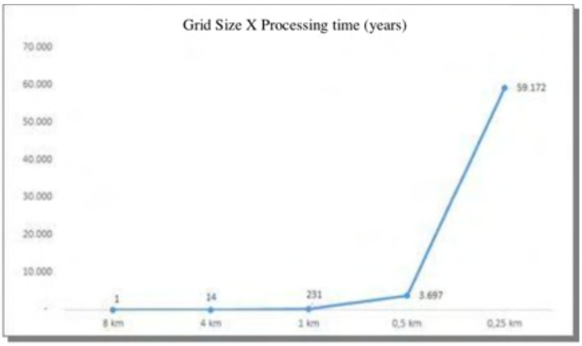
One of the biggest challenges identified by the R&D team was to find a way to simplify the issue of geolocation of eventual occurrences, in order to enable the volume of information to be considered in the calculation of travel times from the operational base to them, could, at the same time, respect the strength of this correlation, without, however, overloading the algorithm to the point of not allowing its processing by common microcomputers or, at worst, supercomputers.

II.2.6 Address Representation Boxes

The solution adopted started from the premise that it would be necessary to reduce the number of geographic variables in the matrix of possibilities for addressing the occurrences and bases, in order to make possible a computational solution accessible to the concessionaire's team of professionals. That is, it was necessary to verify the hypothesis of representing a series of real addresses in the city from a random division of the territory of the concession area into squares of1.000 meters from one side, converging all addresses in that grid to the address closest to its centroid ([1], [8], [10]).

As a result, an indefinite (and certainly enormous) number of possibilities for addressing occurrences began to be controlled, in a smaller matrix of squares, each one representing a portion of the territory of the concession area. The centroid addresses then became the origin or destination of displacements, whose travel times would be offered by the GisBI/Here transit routes, made available and standardized every 15 minutes of the day and night, for the 7 days a week in the concession area.

This decision implied the verification of other possibilities of proposition of squares smaller than 1,000 meters of edge to make the calculation even more precise and representative, but a simulation of the number of mathematical interactions in each of the hypotheses indicated that the square of 1,000 meters of edge was a good choice, due to the extension of the concession territory, to make the calculations possible by the algorithm, since all available addresses (of any centroid) must be checked in terms of travel time for all the other addresses (other centroids) [9].



Graph 1 - Relationship between Grid Size and Processing Time for the Computational Solution adopted by the study

Graph 1 shows the estimated processing time – a function of the number of iterations between the various addresses (the closest to the centroids of the squares) – according to the number of squares. The graph makes evident the physical limitation of processing an "all addresses for all addresses" data matrix. And it demonstrates that the simplification for squares of 1,000 meters of edge with representative addresses closer to the centroids of the squares is at the limit of computational feasibility for the concession region of AES Eletropaulo.

II.2.7 Cost versus Time Matrix

After simplifying the address matrix, it was necessary to verify the behavior of the data available in the various databases chosen for the development of the work. Each of the squares, represented by the address closest to its centroid, then received all geolocated information belonging to that square.

Thus, the databases of: occurrences (AES Eletropaulo); real estate value; climate and afforestation (characteristic taken from the description of AES Eletropaulo occurrences, due to the lack of territorial coverage of this information from city halls). And, obviously, the displacement times between the squares, two by two (transit through the GisBI/Here base), concentrating them on the address close to the centroid.

Any type of information that can be expressed by its geographic characteristics of location could be attributed to the grids in the same way, which gives this investigation carried out by the R&D team a universal character, certainly capable of being reproduced in other analyses. logistical or not, in any part of the territory, in any Brazilian city or in the world.

The combination of the different factors (databases) inserted in the grid, weighted according to the interest of the analysis, received, in this work, the nomenclature of "cost matrix" of the analyzed system, whose values can be evaluated from the KPIs of reference.

In the analyzes presented below, depending on ANEEL's performance requirements, this cost matrix will always offer, as an output, the travel times between each particular square and the squares that represent the location of the operational bases best placed on the territory, based on the table of occurrences weighted by the variables of the system's cost matrix.

The interpretation of the territory as squares allows a computational mathematical modeling and therefore considers:

- The adequacy of the positioning of AES's operational bases depends, to a reasonable degree, and among other factors, on the travel time between the bases and the occurrences.
- The road network is treated as a graph composed of nodes and vertices that connect them, as shown in Figure 2 below.
- To enable the computation of the results in today's commercial equipment, some simplifications were made, such as dividing the area into squares of 1 km on each side.

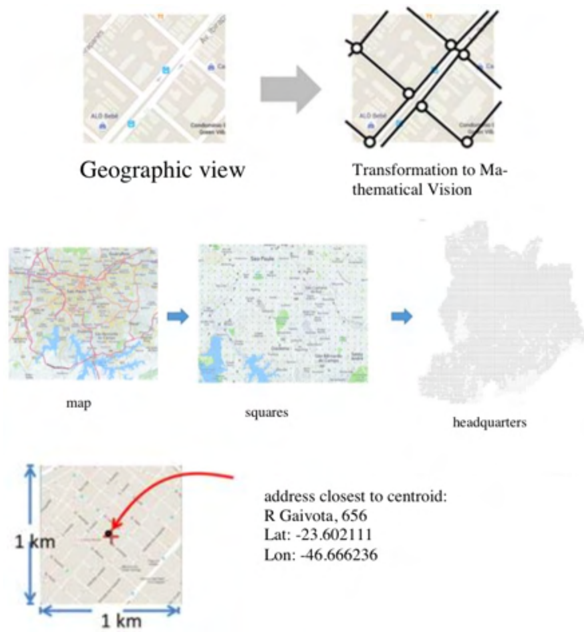


Figure 2 – Transformation of Geographical View into Mathematical View (matrix representation) for computational feasibility of the algorithms developed in the project

II.2.8 Calculation of the best location of the bases

The simplification introduced by the squares allowed the understanding of the calculation processes, but did not change the nature of the computational problem to be solved. all squares them.

To overcome this difficulty, the modeling had to incorporate different algorithms of approximation calculations, known for the solution of NP-complete problems. Such algorithms were then used to define the best positioning of the bases, when the volume of calculations became computationally intractable. For these approximation calculations, the algorithms Brute Force (an attempt to solve the NP-complete problem), Hillclimbing, Genetic and K-means, explained below, were applied.

For traffic data, the Here API was used, which adopts a traffic pattern – the result of studies that take into account historical data from a few years – and provides the time and distance between two points of the road network in the concession region. , in this case, the addresses closest to the centroids of each square.

As the collection and preparation of traffic data takes several days of processing, for the simulations presented below, traffic data was extracted for December 15, 2015 (a Tuesday), at 8:30 am in the morning.

II.2.9 The Geographic Problem Transformed into a Matrix

The geographic problem of determining the best position for the bases was then converted into the mathematical problem of determining which positions cause the lowest weighted cost of locating the bases in a matrix in which each cell represents a grid, as shown in Figure 2, previous.

To calculate the best location, a cost matrix was taken into account that determines the cost (time) of travel between each origin and each possible destination, and a weighting of the square that determines the weight of this square as a destination of travel.

Thus, the algorithm can be used to determine the best position of the bases according to any criterion. For example, it is possible to adopt the travel time between the squares as a cost matrix and the number of occurrences of each square as a weighting.

This way, it was possible to determine the location of the operational bases that allow the lowest average travel time to the occurrences, although this mathematical/ matrix approach allows calculating the positioning according to any other criterion that is desired.

II.3. Development of Optimization Algorithms

As already mentioned, the work analyzed four iterative calculation methodologies, whose characteristics will be presented in the sequence (items II.3.1 and II.3.2). The methodologies for producing iterations were implemented within the algorithm by a program in C++ (standard C++11), to be processed in common microcomputers.

The basic calculation of the concession area, in these conditions, is developed considering the costs of displacements between all the squares, two by two, with the square destined to receive an operational base the one that presents the lowest cost in the cost matrix. The cost of the configuration is given by the sum of the multiplications of the weight of each square with the cost of moving from the operational base to it. The average cost per occurrence, in each operating base location condition, is the total cost divided by the weight:

$$\text{Cost}_{\text{configuration}} = \sum_{i=1}^{N_{\text{Basis}}} \sum_{j=1}^{M_{\text{base squares}}} \text{Cost}_{ij} * \text{Weighting}_j$$

$$\text{Cost}_{\text{Medium}} = \frac{\sum_{i=1}^{N_{\text{Basis}}} \sum_{j=1}^{M_{\text{base squares}}} \text{Cost}_{ij} * \text{Weighting}_j}{\sum_{k=1}^{P_{\text{squares}}} \text{Weighting}_k}$$

II.3.1 The "Brute Force" method – Implementation (not feasible from 3 bases) of the NP-Complete problem

It is the most intuitive method: it tests the displacement costs from operational bases placed on each of the squares, to all the other squares.

The operational base will be placed on the grid that presents the lowest total cost.

- The method is not feasible for configurations with many grids;
- Therefore, for the purposes of this work, the maximum number of iterations was set at 10 million, that is, just over an hour of processing on good-performing microcomputers used in the project;
- This time corresponds to the calculation of the position of two bases, if taking into account, the data from the AES concession area;

Journal of Engineering Research ISSN 2764-1317 DOI 10.22533/at.ed.317272210056

• It can be used to measure the results of other methods;

• Can be used in areas smaller than the concession area.

It brings the exact result, although the massive processing torne inviável para o caso em estudo, que possui um número de bases operacionais maior

Para verificar a viabilidade de iterações (e processamento) according to the number of bases, the processing time was estimated from the necessary number of iterations, considering from one to 10 simultaneous operational bases. The processing time becomes computationally unfeasible (years of machine processing) from the optimization of 4 operational bases.

II.3.2 Hillclimbing, Genetic and K-means methods

The chart in Figure 3 highlights the concepts, characteristics and limitations of each optimization method developed in the project.

	Description of the algorithm	Features and Limitations
Hillclimbing	This method consists of calculation iterations, considering the following phases: An initial configuration is drawn. One base at a time is shifted to each of the squares adjacent to it. The cost of each of the nine possible configurations is calculated, and the base is shifted to the position where the configuration has the lowest cost. The procedure is repeated for all bases, successively, until no displacement of any base results in a decrease in cost.	Result depends on the initial draw, i.e. the "initial seed" for the initial settings draws Eventually the starting position is such that a final position is reached which is a "local minimum", not representing the best possible situation. The calculation module allows determining the number of initial draws of the method, which is very fast and allows thousands of attempts to be calculated. It is not known which number of draws brings the best result, as this depends on the heterogeneity of the cost matrix and weightings. It can be said that, in general, the higher
Genetic	It consists of the following steps: 1. An initial population of configurations is drawn. 2. Each configuration has its cost calculated. A number of survivors is chosen, the ones with the lowest cost. 3. Each survivor generates a number of descendants equal to the number of bases, as follows: each base of the configuration is replaced, one at a time, by another randomly selected base. 4. The new population, together with the descendent configurations, is compared with a list containing all tested configurations, to ensure that survivors from the previous generation are not repeated, have their cost calculated, and the new survivors are chosen. 5. The process is repeated until the desired number of generations is reached.	It is not limited to local minima. The quality of the solution depends on the number of generations. There is no guarantee that the solution will be optimal or the best of all.
K-means	1. An initial configuration is drawn. 2. Calculate the coverage of each base in the configuration. 3. In each coverage area, calculate the best base placement. 4. Repeat the two previous steps (ie, no new draw is made) until the base positions are stabilized.	It depends on the initial draw. The quality of the solution depends on the number of repetitions. There is no reason to believe that stability is achieved at the lowest cost point.

Figure 3 – Description of Algorithms implemented in the project.

II.4. Implementation and Performance Tests

Tests were performed considering 1 to 2 bases with the simulation of the 4 algorithms and verification of the location of the optimal bases in comparison to the Brute Force (ideal solution). For 3 or more bases, the tests consisted of evaluating the performance of the solutions.

RESULTS

Among the main findings, it is highlighted that traffic has a great influence on the positioning of the bases, in a non-trivial way.

Under the observed conditions, the Hillclimbing method converges faster to the best solution; but the intermediate results do not consistently approach the best result. The Genetic method requires more iterations, but consistently approaches the best result: the more generations, the better.

These results influenced the decision to make available a fourth method, genetic with refinement, which we called "GENETIC + HILLCLIMBING". This method performs the genetic method as predicted, and takes the best solution to apply the Hillclimbing method only once. The idea is to take the best possible generation and optimize it in a way that the Genetic method cannot on its own. It is necessary to take into account, the performance with all the data to evaluate the best uses for the algorithm, in each case. The K-means method apparently converges to a point that is not the least expensive.

Figure 4 shows the arrival and departure costs of each square in the AES Eletropaulo concession area (as described in 2.3). Figure 5 shows an example of the final configuration for the 5-base simulation. It is estimated the general maintenance of current service indicators (in particular, average service and travel times) and cost reductions with the reduction of the total built-up area of the bases, which reach tens of millions of reais.

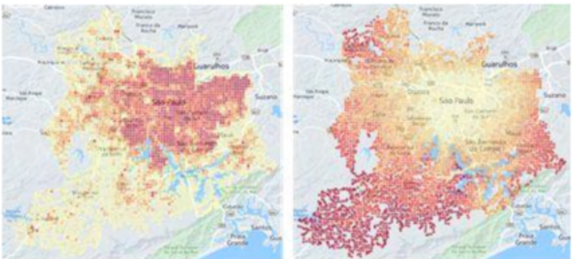


Figure 4 – CArrival and Departure costs of the grids of the AES Eletropaulo concession area.

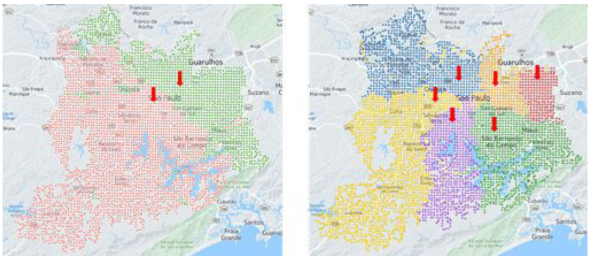


Figure 5 – Example of optimizing the composition of 2 (left) and 5 (right) operational bases.

Crowning the implementation, the systemic platform (software) was developed completely web-based, to allow access via browser to AES Eletro-paulo analysts to the definition of simulation variables (traffic, real estate dynamics, data of climate and trees), definition of the weights of these variables, selection of the number of bases, optimization method and specific parameters, simulation of scenarios and measurement of KPIss.

CONCLUSIONS

The introduction of the concepts and possibilities of geotechnologies and geointelligence in the operational activities of public service concessionaires can be considered as an inexorable trend of methodological advances and business practice, notably of the company dedicated to activities distributed throughout the territory, whether in urban or rural areas.

The project developed within the scope of R&D demonstrates that this introduction of concepts, considerations and methodologies of geostatistical calculations to the daily business can be carried out without the specialists in planning and operational management having to learn or incorporate all the complexity and nuances of the GIS (Geographic Information Systems) software available on the market.

The developed system (software) aims to offer an instrument for the automated implementation of the geographic calculation methodology for the relocation of the operational bases at AES Eletropaulo, and is the main result of the work carried out. The system contemplates the objectives of simplicity for the introduction of the variables to be considered in the analyzes and reasonable computational times with the possibility of studies in multiple scenarios of modification of the urban and climatic conditions of the concession region.

The same system can be used in other regions, provided that local variables are made available and introduced, in combination with the operational data of any energy distribution company in the country or even in other countries. The resources of the digital maps are integrated into the developed tool and, therefore, the results of this R&D are not limited to the geographic area of the AES Eletropaulo concession, but may also serve other types of public service concessionaires.

For the concessionaires' operational planning and management teams, the results of this R&D offer the possibility of immediately incorporating geotechnologies into their professional activity, without requiring in-depth knowledge of spatial statistics. This possibility may lead to a new level of operational performance, as well as evidence of significant improvements in the provision of public energy distribution services, constituting, for that very reason, a great methodological success for those who adopt the computational system for reallocation of operational bases developed.

It is, therefore, a Big Data project with unprecedented Geographical Optimization in the Brazilian electricity sector.

REFERENCES

1. FRANCISCO, E. R: Indicadores de Renda Baseados em Consumo De Energia Elétrica: Abordagens Domiciliar e Regional na Perspectiva da Estatística Espacial. Tese de Doutorado - Fundação Getúlio Vargas - EAESP. São Paulo: 2010.

2. KIRSCHNER, P: A study of vision as a measurement of decision-making behavior: 1971. Thesis (Ph.D.) - Dept. of Industrial Engineering, Stanford University.

3. LACERDA, L. Armazenagem Estratégica: Analisando Novos Conceitos. Mar. 2000. On-line. Disponível em <http://www.ilos.com.br>

4. BALLOU, R. H. Gerenciamento da Cadeia de Suprimentos/Logística Empresarial. 5a Edição. Porto Alegre: Bookman,2006.

5. COELIS, E. L. Logística Empresarial. Belo Horizonte: 2006. 10p. Instituto de Educação Tecnológica. On-line. Disponível em <http://www.techoje.com.br/site/techoje/categoria/detalhe_artigo/507>

6. BARROS, M. C. Warehouse Management System (WMS): Conceitos teóricos e Implementação em um Centro de Distribuição. Rio de Janeiro: 2005. 132p. Dissertação de Mestrado - Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

7. FRANCISCO, E. R. . RAE-eletrônica: Exploração do Acervo à Luz da Bibliometria, Geoanálise e Redes Sociais. ERA (Impresso), v. 51, p. 280-306, 2011.

8. WANGENHEIM, A. Análise de Agrupamentos 2006. On-line. Disponível em <http://www.inf.ufsc.br/~patrec/agrupamentos.html>

9. LLOYD, C. Spatial Data Analysis, Oxford University Press, 2010.

10. WANKE, P . Aspectos Fundamentais do Problema de Localização de Instalações em Redes Logísticas. 2001. On-line. Disponível em http://www.centrodelogistica.org/new/fr-aspec.htm