LONG TERM PLANNING BASED ON THE PREDICTION AND ANALYSIS OF SPATIAL LOAD

Marcus Rodrigo CARVALHO
Daimon Engineering – Brazil
marcus@daimon.com.br

Bruno Yukio ENOMOTO
Daimon Engineering – Brazil
Bruno.enomoto@daimon.com.br

Marcello PEK DI SALVO
Daimon Engineering – Brazil
marcello.salvo@daimon.com.br

André MÉFFE
Daimon Engineering – Brazil
andre.meffe@daimon.com.br

Luiz FERNANDEZ SILVA
CPFL – Brazil
fernandez@cpfl.com.br

Carlos Cesar Barioni OLIVEIRA
Daimon Engineering – Brazil
barioni@daimon.com.br

ABSTRACT

The main objective of this paper is to present a new strategic planning model based on spatial analyses of the load forecast inside the concession area of a power distribution company. The model is founded on the analyses of grids with pre-defined dimensions (such as 500 x 500 m, or 1,000 m x 1,000m), according to the planning engineer criterion. The customers' data are grouped together in only one load point inside every grid. Through this approach, one can generate theme maps to illustrate different scenarios of load density, load curves per consumption category and growing rates per region, helping to determine the optimal location for installing new power distribution substations.

INTRODUCTION

Due to the increasing regulation of the Brazilian electrical sector, with emphasis to the new Power Distribution Procedures – PRODIST [1], the submission of the Power Distribution System Works Plan to the regulatory agency (ANEEL) became mandatory. Such plan should indicate the expansion and construction of substations within a 10 years horizon. In order to design a technical and economically efficient plan, all alternatives for installation sites and nominal power of these new substations must be considered, the quality levels and limit values for technical indices established in the PRODIST have also to be followed, and the analysis of the costs needed by each alternative must be made. Through this perspective, the required investments will not overtax the customers.

The typical load curves of each customer category from the Brazilian power distribution companies are periodically during their tariff cycle. The representative curves for each customer category present in their concession area are obtained through a process based on suitable clustering techniques, namely load characterization. Using these representative curves and the monthly consumption information from each customer, it is possible to determine curves per end use (disaggregated in residential, commercial, industrial and others) for each grid.

The planning study is initialized through a load prediction model. In this work, a new methodology was developed, in which the growing rate for each grid is determined through the extrapolation of the consumption history of each category (residential, commercial, industrial and others). The rate for each grid is pondered by a global growing rate from the region or set of substations under analysis. The global rate is calculated through prediction models such as Box Jenkins or SARIMA (Seasonal Autoregressive Integrated Moving Average) [2], which is a multivariate prediction technique and that correlate the market growing with other explanatory attributes.

The results of this application allows one to diagnose the power network for the planning horizon, with respect to the need of increasing the capacity of the current substations (through the addition of new power transformers) or, even the need of building new substations (indicating their capacities where they should be built).

In order to satisfy this objective, different optimization methods were analyzed and improved, such as Hill Climbing method [3], Particle Swarm Optimization (PSO) Method [4] and Active-Set method [5].

Innovative methodologies were also developed, constituting in effective support tools for the strategic planning analysis, such as: theme maps of spatial load migration vectors, definition of idle areas and definition of demand attraction poles.

Before proposing the increase of the existent substations or the construction of new ones, a demand optimization study is carried out, in order to distribute the load evenly among the substations under analysis. Such study allows one to postpone large investments. Three algorithms were developed for determining the substations' influence area:

i. Minimizing the product kVA x km of each grid.
ii. Minimizing the product kVA x km of each grid with penalizations.
iii. Associating the shortest electrical path, considering the network from the study's initial year.

These methodologies supported the development of an algorithm and a strategic planning system capable to determine when the existing substation should be increased and, the case being, new ones should be built. Besides that, the system is also capable of indicating the areas where management actions are needed, considering the demand or the offer perspectives, constituting a financial and economical index.

METHODOLOGY

The definition of planning technical criteria is based on normative resolutions that determine the technical indices that should be considered by the utility companies, regarding the risk aversion of not satisfying the list criteria during contingency situations and in economic use of assets studies.

The planning methodology can be described through the following stages:

i. Characterization of the System under Analysis:

<u>Network Characterization:</u> Definition of the attributes needed for characterizing the current network, in order to

allow the performance analysis through technical planning criteria and eventual reinforcement propositions when there are transgressions limits.

Market Characterization: Definition of the market supplied by each substation, based on historical consumption data, typical load curves and spatial representation (through a geo-referenced platform).

Future Demand Projection: Consists on the projection of the demand growing rates, in order to orient the proposition of the reinforcements required for supplying the market expansion.

ii. Grids definition and their link with the substations under analysis:

Division of study area into spatial grids, in order to consider the particular characteristics of each region in the planning studies.

iii. Analysis of the system and the work proposition for the offer expansion

Once the grids are defined and their links with the substations under analysis are established, the network analysis is carried out. Thus, it is verified if the limits fixed for the planning criteria are being respect and the impact of the reinforcements proposed for supplying the market.

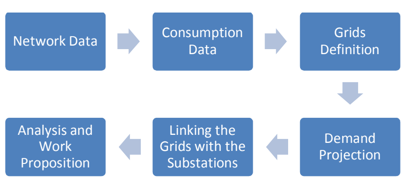

Fig. 1 – Planning Methodology Flowchart

Load Prediction Model

The market growing rates are defined to subsidize the energy purchase procedures. When such rates are applied in the existing networks, one can verify which substations will reach their capacity limits.

The developed model calculates two different rates. The first one, the global one, is calculated based on multivariate time series statistical models such as SARIMA. The historical consumption data from the whole region is used and may be co-integrated with other series such as the region's GNP, number of customers, etc. This model allows the identification of load seasonality within a short term periods (2 to 3 years) and the achievement of rate with a considerable assertiveness degree for long term periods.

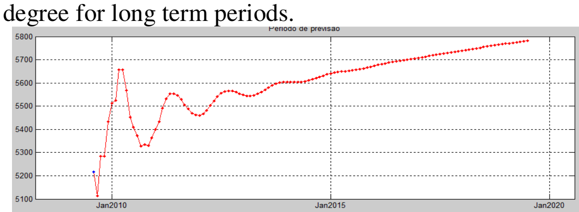

Fig. 2 – Example of a Load Prediction Model Using SARIMA

The SARIMA model developed is auto-regressive of (p) order and its main objective is to calculate (predict) the value for an instant t (Y_t), based on the combination of the past values with a random variable, the error, where (p) is the number of past values being considered to obtain (Y_t), as illustrated by Equation (1).

$$Y_t = c + \varphi_1 \cdot Y_{t-1} + \varphi_2 \cdot Y_{t-2} + ... + \varphi_p \cdot Y_{t-p} + \varepsilon_t \quad (1)$$

The moving average concept is also taken into consideration by this model, i.e., the series may be built through a combination of past errors, which are the differences between the actual real and the estimated one. The aim of a (q)-order moving average model is to calculate a value for the current instant t, through the past values for random variables. The (q)-order correspond to the quantity of random values considered for obtaining Y_t, as illustrated by Equation (2).

$$Y_t = c + \varepsilon_t + \phi_1 \cdot \varepsilon_{t-1} + \phi_2 \cdot \varepsilon_{t-2} + ... + \phi_q \cdot \varepsilon_{t-q} \quad (2)$$

The planner may co-integrate the main series with three other series, according to its will and the data available for the region under analysis. Another feature is the possibility of removing an atypical period from the series, so the prediction model may not be negatively affected.

The second rate, the grid rate, is calculated for each grid separately through a Simple Regression [2] of the consumption historical data for each customer category (residential, commercial, industrial and others). This model refers to any regression of a variable Y (dependent variable or the one to be predicted) over a variable X (independent variable or the explanatory one). The model considers a linear relationship between Y and X given by Equation (3).

$$Y_i = \alpha + \beta \cdot X_i + e_i \quad (3)$$

Where:

- α is the linear coefficient (point that the y-axis is crossed);
- β is the angular coefficient (line's slope);
- e_i is a random error within period i (corresponds to a deviation between the observation and the linear model).

The weighting between the global rate and the rate for each grid results in the effective growing rate for each consumption category. The effective rate is the one to be applied in the current demand of each grid to determine the future market, according to the Equation (4):

$$T_{effective} = T_{global} . D_{global} \big/ \sum D_i . \beta_i \quad (4)$$

Through this methodology it is possible to determine:

i. Spatial Load Migration Vectors:

The migration vectors are supposed to indicate the load "displacement" among the many areas considered by the study. It is possible to estimate the growth of each region (set of grids) through their historical data. This piece of information allows one to estimate the load migration vector, using the points with a lower growth (angular coefficient (β) lower than the global rate) as the origin point; and using the points with a higher growth (angular coefficient (β) higher than the global rate) as the reaching points. It is important to highlight that these vectors may be determined for each consumption category.

ii. Idle Areas Identification:

One possible methodology for verifying the idleness of a specific area is based on monitoring the angular coefficient from the linear regression for a specific period of time (around a few years). The areas where the angular coefficient (β) is close to zero are classified as the idle ones.

Linking the Grids with the Substations

Aiming to postpone the substations capacity increase, an optimization for the load supplied in each region by a specific set of substations is executed. Three different methodologies were developed for linking the grids to the substations:

Minimizing the Load*Distance Products

The objective of this method is to optimize the grids' linking. The summation of the products Load*Distance are minimized, using the Ford-Fulkerson [6] algorithm, which is a transport algorithm. The algorithm may be described as follows:

$$\min \quad C = \sum_{i=1}^{m} \sum_{j=1}^{n} C_{ij} . Z_{ij} \quad (5)$$

$$s.t. \quad \sum_{j=1}^{n} Z_{ij} = S_i, \quad i = 1, 2, \ldots, m \quad (6)$$

$$\sum_{i=1}^{m} Z_{ij} = D_j, \quad j = 1, 2, \ldots, n \quad (7)$$

Where:

- m is the number of substations;
- S_i is the capacity of substation i;
- n is the number of loads;
- D_j is the demand of load j;
- C_{ij} is the distance between substation i and load j;
- Z_{ij} is the demand D_j supplied by substation i.

Through this algorithm, the "cost" is the distance between the substation and the load. The network topology is not taken into consideration.

Shortest Geographic Distance with Penalization

This formulation executes the linking for each grid, in a way that minimizes its geographic distance with the substation. The substations' loading is taken into consideration, in order to avoid the overload of any of them. The algorithm may be described through the following steps:

1. Initially, one verifies if the total power from all substations is higher than or equal to the load, in order to guarantee the algorithm's convergence.
2. It is assumed that all substations have an unlimited capacity. Then, each grid is associated to the closest substation.
3. It is verified is the substation's capacity was exceeded, according to its rated power.
4. For overloaded substations, the algorithm penalizes the distances by increasing the geographical elevation where the substation is located (Z-Axis). Thus, the grids that are located in the boundaries of its influence area become more distant from the substation.
5. The grids are associated again to the closest substations.
6. Stages 3-5 are repeated until all substations reach permissible loading levels.

Shortest Network Path

This methodology considers the concept of electric distance, i.e., for the grids that are supplied by feeders from more than one substation, one calculates the distance between the grid and every substation through the network topology. One may also choose to weight the distance by the resistance from the feeders' conductors. The algorithm may be described through the following steps:

1. All switches from the network are considered to be closed.
2. The feeders that pass though each grid are determined.
3. The distances between the substations and the grids are determined through the network topology.
4. Each grid is associated with the substation that presents the shortest distance among the ones previously determined.

Allocation and Sizing for New Substations

When the substations' capacity is exhausted due to the load supply, it is necessary to determine the increase of the existing substation or to propose new substations (location and sizing). In order to satisfy such objectives, different optimization methods were analyzed and improved, such as the Hill Climbing Method, the Particle Swarm Optimization (PSO) e Active-Set Method.

The algorithm is based on the following steps:

1. The grid with the largest distance with respect to its own substation is identified.
2. The grid position is used as the initial position for a new substation.
3. The optimization algorithm is executed the value for the objective function is retained.
4. Steps 1-3 are repeated. One new substation is added in the next iteration, if the objective function continues to decrease.
5. The solution that supplies the load demand with the smallest objective function value is returned.

Objective Function:

$$\min C = A \sum_{i=1}^{I} d_{ij} . D_i^2 + B \sum_{i=1}^{I} d_{ij} + \sum_{j=n+1}^{N} CC_j + \sum_{j=1}^{N} T_{ij} . CT_j \quad (8)$$

Where:

- d_{ij} is the distance between grid i and substation j;
- D_i is the demand from grid i;
- A is the weighting factor for the losses cost;
- B is the weighting factor for the feeders' construction cost;
- I is the total number of grids;
- N is the total number of substations;
- CC_j is the construction cost for substation j;
- T_j is the quantity of type i transformers needed in substation j; and
- CT_i is the cost of a type i transformer.

Optimization Algorithms

Active-Set Algorithm

The solution procedure is composed by two stages. The first stage involves the calculation of a feasible point (if there is one). The second phase involves the generation of an iterative sequence of feasible points that converge to the optimal solution. Given a point x in the feasible region, a constraint $g_i(x) \geq 0$ is called **active** at x if $g_i(x) = 0$ and **inactive** at x if $g_i(x) > 0$. Equality constraints are always active. The active set at x is made up of those constraints $g_i(x)$ that are active at the current point. In this method an active set is maintained that is an estimate of the active constraints (i.e., those that are on the constraint boundaries) at the solution point.

The Active-Set is updated at each iteration k. This is used to form a basis for a search direction. The search direction is calculated and minimizes the objective function while remaining on any active constraint boundaries. When n independent constraints are included in the active set, without locating the minimum, Lagrange multipliers, λ_k, are calculated to satisfy the nonsingular set of linear equations. If all elements of λ_k are positive, x^k is the optimal solution of Quadratic Programming. However, if any component of λ_k is negative, and the component does not correspond to an equality constraint, then the corresponding element is deleted from the active set and a new iterate is sought.

Hill Climbing Algorithm

Hill Climbing is a local searching method that considers an iterative improvement procedure. The strategy is applied in a singular point x (candidate solution) in the search space. It standard (or simple) algorithm may be described as follows:

- Initialize (randomly) the point x in the problem's feasible region.
- At each iteration, a new point x' is selected though a disturbance in the current point, i.e., a x' point in the neighborhood of x is selected, where x'∈ N(x).
- If the new point presents a better value for the evaluation function, the new point becomes the current one.
- After a few steps, if a better value is not found, the method reduces the disturbance magnitude and starts the searching procedure again.
- The method stops when the disturbance magnitude or the variation in the evaluation function reaches a minimum value.

Particle Swarm Optimization (PSO)

Particle Swarm Optimization is an evolutionary algorithm that simulates a social behavior. Instead of applying evolutionary operators in each individual, such as the Genetic Algorithms, in PSO each individual, namely particle, belongs to a set that "flies" over the solution space with a specific speed, which is dynamically adjusted according to its own flight experience and the flight experience from the other set members. Each particle may be represented through a position vector and a speed vector. They also retain the best position achieved in the past iterations. The swarm retains the best position of each particle. These data are used to update the speed of each particle through the following way:

$$v_{id} = w.v_{id} + c_1.rand().(x_{id} - x_{pbest}) + c_2.rand().(x_{id} - x_{gbest})$$
$$x_{id} = x_{id} + v_{id} \quad (9)$$

Where:

- v_{id}: speed in dimension i for particle d;
- x_{id}: position in dimension i for particle d;
- w: inertial weighting;
- c_1, c_2: positive weighting constants;
- x_{pbest}: best position achieved by the particle;
- x_{gbest}: best position in the set;
- *rand()*: random number between 0 and 1.

The disadvantage of this particular algorithm is due to the attraction of x_{gbest} to local minima, limiting the set's search capability. In order to overcome this disadvantage, the method uses the concept of mutation from the Genetic Algorithms. The mutation operator is applied in the particles' position x_{pbest}, in order to make them travel through a larger space during the search.

TESTS AND RESULTS

The presented algorithms were tested with the network from Campinas city. The system is composed by 17 substations with a total installed power of 729.45 MVA. There are 388,612 customers connected to the network, corresponding to 246.76 MVA.

In order to evaluate the algorithm's robustness, a high growing rate of 11.6% per year was considered. A substantial increase need was then verified. Table I presents a comparative between the results achieved by the optimization algorithms.

The Active-Set algorithm was faster than the others. Nevertheless, the results achieved were less reliable and, for a few cases, it reached a local minimum or did not converge towards the solution. Comparing with the other algorithms, one can say that the results achieved by the Hill Climbing method and the PSO method were similar for the studied cases, with similar execution times. However, the PSO presented a smaller variation for the execution time from one test to another if compared with the Hill Climbing. Such behavior indicates that the PSO is more suitable for an increase in the problem's complexity.

Table 1 – Comparative Results

	Method		
	Active-Set	*Hill Climbing*	*PSO*
# New Substations	1	1	1
# Transformers	4	4	4
Objective Function	1,958e+7	1,714e+7	1,714e+7
Execution Time (s)	45,4	399,2	386,5
Substation Position: X-Axis	287900	287300	287300
Substation Position: Y-Axis	7462300	7465700	7465700

CONCLUSIONS

These methodologies allowed the development of an algorithm and a strategic planning system that provides:

- Chronological indication of when new substations should be installed and when capacity from the existing ones should be increased;
- Indication of the areas where management actions regarding the demand or offer sides are needed.
- Indication of areas where induction actions are needed.

This system resulted from a R&D Project sponsored by RGE. RGE is a power utility that supplies 262 counties in Rio Grande do Sul state, Brazil. It has 1.2 million customers supplied by 62 distribution substations and 4 transmission substations.

The developed system was applied in RGE's networks during the development and testing stages. One tried to satisfy the current regulatory demands described in PRODIST's planning module in the most suitable way.

REFERENCES

[1] Procedimentos de Distribuição de Energia no Sistema Elétrico Nacional– PRODIST: http://www.aneel.gov.br/area.cfm?idArea=82

[2] Makridakis, Wheelwright and Hyndman , 1998, *Forecasting: Methods and applications*, John Wiley & Sons, New York, 3rd edition.

[3] Russell, Stuart J.; Norvig, Peter, 2003, *Artificial Intelligence*: A Modern Approach (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, pp. 111–114.

[4] Liu Zifa; Zhang Jianhua; *Optimal Planning of Substation of Locating and Sizing Based on GIS and Adaptive Mutation PSO Algorithm*, Power System Technology, 2006.

[5] Gill, P.E., W. Murray, M.A. Saunders, and M.H. Wright, 1984, *Procedures for Optimization Problems with a Mixture of Bounds and General Linear Constraints,* ACM Trans. Math. Software, Vol. 10, pp 282–298.

[6] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L.; Stein, Clifford (2001). *Introduction to Algorithms* (2nd ed.). MIT Press and McGraw–Hill. pp. 651–664.