IDENTIFICATION AND EVALUATION OF ENERGY THEFT USING THE STATE ESTIMATOR IN MV AND LV GRIDS WITH EXOGENOUS PARAMETERS FOR PLANNING EXPANSION

Daniel DUARTE
Sinapsis – Brazil
daniel.duarte@sinapsisenergia.com

Daniela KONDO
Sinapsis – Brazil
daniela.kondo@sinapsisenergia.com

Felipe MATSUZAKI
Sinapsis – Brazil
felipe.matsuzaki@sinapsisenergia.com

João GUARALDO
Sinapsis – Brazil
joao.guaraldo@sinapsisenergia.com

Mariana SOUZA
Sinapsis – Brazil
mariana.souza@sinapsisenergia.com

Hugo SILVA
Eletrobras Alagoas – Brazil
hugo.silva@eletrobrasalagoas.com

Marcel FERREIRA
Eletrobras Alagoas – Brazil
marcel.ferreira@ceal.com.br

Rodrigo SILVA
Eletrobras Rondônia – Brazil
rodrigo.silva@eletrobrasrondonia.com

Luis BRITO
CEPEL – Brazil
grillo@cepel.br

Ricardo ROSS
CEPEL – Brazil
rpdr@cepel.br

Nelson KAGAN
USP – Brazil
nelson@pea.usp.br

ABSTRACT

Planning studies require knowledge about the load flow in installations. The calculation tools use registration data for load and grid characterizing. They also adopt typical load curves and operating measurements to make adjustments in order to mitigate any differences inherent to the adopted representation model. In general, this adjustment considers the load difference between the calculation with typical curves and operational measurement. This difference is distributed proportionally to the load registered along the feeder. When it is a small difference, this assumption is appropriate. However, in systems with high levels of energy theft, the difference between calculation and measurement assumes high values. Therefore, it makes the above premise a source of relevant inaccuracies in electrical calculations. The objective of this session is to present a methodology for conducting demand adjustment to better represent the load and allocate the energy theft in a sectored way, using energy balance, tacit knowledge and state estimation techniques.

INTRODUCTION

Planning studies of distribution systems use electric simulation of the present and future system condition. It considers interventions for grid expansion and improvement.
The tools ready for these studies have electrical modelling based on the information from the company register system. The definition of the grid loads are calculated using values of monthly billed energy for each consumer. The conversion of energy values into demand curves is performed through the association of typical load curves, stratified by consumer type and class, as required in the measurement campaign for Brazilian regulation purposes.
Finally, the system is represented with all the equipment and consumer electrical characterization through a daily demand curve in at least twenty-four hours schedules.
Obviously, the resulting calculation of the exposed modelling incurs inaccuracies inherent to the use of typical curves. In order to mitigate this problem, it is common to allow adjustment based on operational measured, approaching the model to curves observed in the field. This is usually made by rising or reducing loads uniformly for all consumers connected to the grid.
However, in grids where the energy theft is considerably high, the discrepancy between the power flow using monthly consumption with typical curves and the values obtained from the measurement can be significant. In these cases, the adjustment of demand with the increase of loads homogeneously over the feeder will certainly lead to relevant inaccuracies, especially in technical loss calculation studies and voltage levels evaluation. This is because energy theft niches are, in general, correlated to exogenous parameters, being able to be concentrated in a certain point of the feeder.
For this reason, the methodology and tool presented below are intended to enable, through cadastral data, measurement and maps that capture exogenous factors, the allocation with more precision of the amounts of losses due to energy theft. The distribution system and consequent adjustment of demand more closely to reality, make a greater accuracy in defining the actual grid loading conditions.
This paper is a final result of a R & D project of ANEEL (Agência Nacional de Energia Elétrica - Brazilian regulator) program, financed by Eletrobras Alagoas and Eletrobras Rondonia.

DEVELOPMENT

There are four steps that make up a complete and coherent study to the ultimate goal:

1. Load curve base definition for the grid
2. Alignment between consumers billing date and the measuring month
3. Survey maps with exogenous factors to help identify sites with probability of energy theft
4. State estimator to correlate measurements with the power flow of billed consumers and their typical curves

After accomplishing all these phases that will be better described in this topic, the user can identify or draw polygons with higher probability of energy theft. In this way, the software, finally, allocates the demands in an unequal way downstream of a meter.

Load curve base definition for the grid

The first step in the methodology consists in a software to analyze the demand curve from feeders' measurement data. It incorporated the characteristics of the company, for the measurement treatment, expurgating outliers, referring to measurement errors, operative maneuvers or protection action. This methodology produces the real hourly demand curve, for of load studies purpose, considering seasonality. The figure 1 shows the flowchart of the developed software.

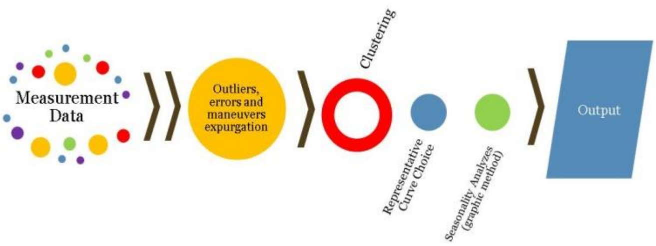


Figure1. Software for basic load definition

First, the software directly addresses the meters mass memory, composing the database with available data.
After importing them into the internal database, it is important to purge outliers, as these can damage the analysis. The purges considered are: data absence or null values, level change (falls greater than 90% or elevations greater than 10 times the previous value), negative data and operational maneuvers (coming from external file of occurrences record).
This software considers the k-means clustering technique with that target: grouping demand curves with similar characteristics. The k-means method is used to minimize the sum of squared distances between all the points and the center of the cluster.
In order to determine the seasonality, the measure of dissimilarity used was the sum of the distances between demands and the centroid of each hour (in 24-hour total). In this way, it is possible to obtain a great number of clusters for the definition of curves groups.
Finally, before running this program it is necessary to identify what role the study is playing. The result depends on some user choices, summarized below:

a) Calculate the base load curve of a specific substation or feeder;
b) Which years, months, and days of the week are interesting to consider in the base load module;
c) Which electrical quantities is to be generated;
d) Choose between imposing a specific number of clusters or execute the optimized method
e) Consider a mean curve or a representative curve (group that corresponds to a curves percentage, as can be observed in Figure 2);

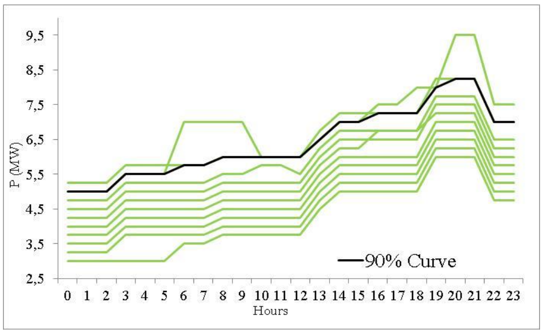


Figure2. Representative curves example

For example, on the above image, the representative curve corresponds to 90% probability of the measurement shows values equal or less than the data selected for each hour.

f) Which cluster is to be exported as base load.

The software provides bases for the decision maker, such as charts and data lists that results in a base load curve and helps in identifying and process errors. Furthermore, as mentioned, the methodology captures seasonality between the input measurements and variations between the days of the week. Finally, the outputs are:

1. Purged Curves, allowing the user to understand what measurement errors were detected;
2. Graphs and view lists of clusters;
3. Distribution of months and days of week in each cluster. Facilitating, therefore, the cluster choice to export, in order to determine seasonality or not. This is because clustering separates into different groups, dissimilar measurements corresponding to different seasons of the year;

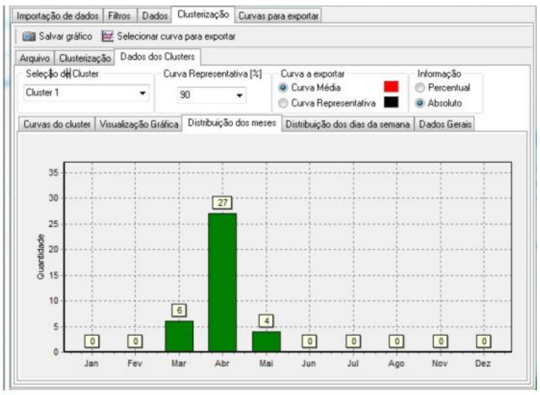


Figure3. Month distribution for a cluster

4. A .txt file that can be read by SINAPgrid (the calculation tool used by Eletrobras) with hourly data of the chosen cluster, curve selected, distribution of months and days corresponding to that cluster;

Consequently, following the planning process, the user can use these files in the .txt format to adjust the load in SINAPgrid, allowing power flow simulation under the previously chosen conditions.

Pro-Rata Methodology

The consumer billing does not occur on the same day for them all attending by the same feeder. There is a program of the reading routes, what causes a mismatch between reading dates between these consumers. Considering that, a software was developed that standardizes, through a calculation model, these consumers billing dates. In this way, all consumers of the grid segment are aligned to a calendar month. This methodology was called "pro-rata" and the calculated consumption should be performed as follows:

$$\text{New Consumption} = \frac{((D2-Di)+1*C2}{D2-D1} + \frac{(Df-D2)*C3}{D3-D2}$$

Where:
D1 = Billing date in the month prior to the chosen month
D2 = Billing date in the selected month
D3 = Billing date in the month after the chosen month
Di = First day of the chosen month (01 / month chosen)
Df = Last day of the chosen month
C2 = Consumption in the chosen month
C3 = Consumption in the month after the chosen month
Considering the load standardization file based on pro-rata rule, the software adjusts the demand of each grid consumer. In order to carry out this adjustment, it considers the energy of the georeferenced system (present in SINAPgrid) and the energy calculated by the pro-rata methodology. For each consumer, it realizes an energy correction without changing the profile of the typical curve, using only a multiplier that raises or lowers the energy to match to that calculated by the pro-rata methodology.

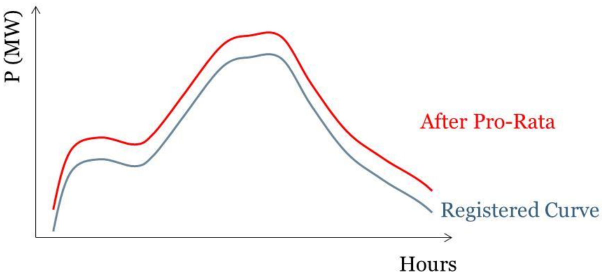


Figure4. Energy correction performed after pro-rata

Maps with exogenous parameters

Red Alerts

The red alert map highlights consumers who have already been disconnected due to irregularity, according to the commercial system. Thus, with an open grid in SINAPgrid, it is identified by red markings, where these irregular consumers are located as in Figure 5.

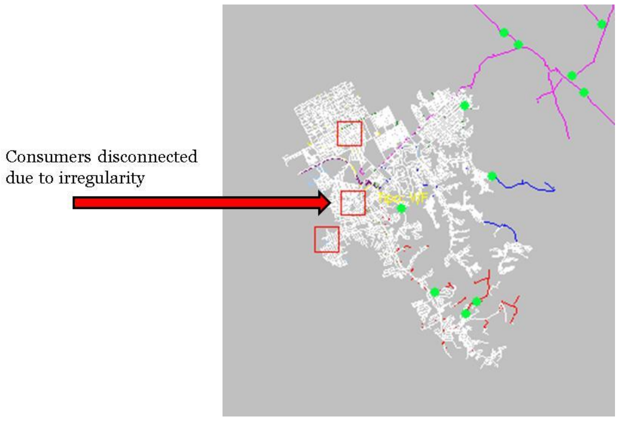


Figure5. Map of consumers disconnected by irregularity

IBGE Maps

Three IBGE (Brazilian Institute of Geography and Statistics) maps were included to compose the library of exogenous factors most likely to have energy theft: subnormal agglomerations (slums), income per capita and meters number per inhabitant (%). An example of the SINAPgrid screen with the IBGE map and the electricity grid can be observed in Figure 6.

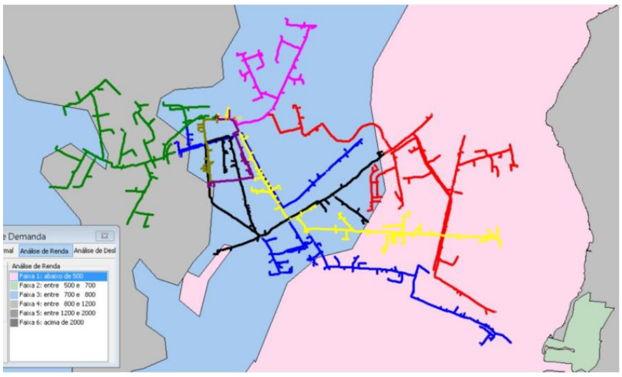


Figure6. Income per capita map with the electric grid

State Estimator

Methodology

The state estimation allows, from a set of measurements, in an electric system, to estimate the state of the system more consistent with its configuration and its electrical characteristics. From this, there is a better adjustment where it is measured and it estimates other electric quantities in grid points, such as equipment loads, voltage levels and technical losses.
In the case of application to high voltage distribution systems, or subtransmission systems, the methodology application is straightforward, since the measurements in the frontier bars, in the substations and large consumers are known.
However, to apply the methodology in medium and low voltage grids (MV and LV), some adaptations were made. Such adaptations are based on typical load curves.

The problem illustration is shown in Figure 7.

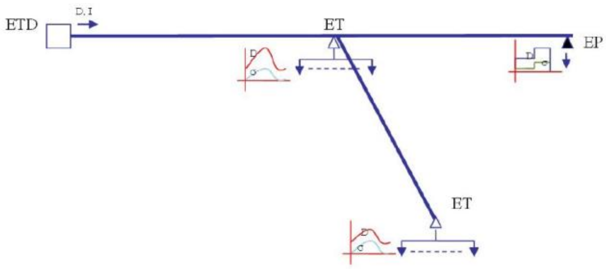


Figure7. Primary feeder with ETs (Transforming Station) and EP (Primary Station)

In the illustration in Figure 7, we note three types of measurement:

a) Measurement at the beginning of the feeder (ETD), where the electrical quantities, current and, in general, the values of active and reactive power. For these measurements, the standard deviation to measurement errors is known.

b) Energy measurement in low voltage consumers: this consists in the monthly consumption in KWh that is realized for the billing and later transferred to the company technical systems. The typical curves in p.u. establish the demand and standard deviation values, based on the average feeder demand. A form of consumers aggregation in ET is made by the sum of average demands and variances, i.e.:

$$\mathrm{D}_{\mathrm{ET}} = \sum_{j=1}^{N\_cons}(Dj), \quad \partial^2{}_{\mathrm{ET}} = \sum_{j=1}^{N\_cons}(\partial^2 \mathrm{j})$$

c) Demand measurement in private consumers, primary consumers (in MV). In general, for these consumers, the demand and power factor values are known. Also, in this case, it can be evaluated, based on the measurement and system, the maximum errors and corresponding standard deviation.

Considering that, from input measurements that may have an error, the main purpose of the state estimation is to obtain a grid state (described by the module and the voltage angle at each node) in order to minimize some global criterion. The objective function implemented minimizes the total value of the quadratic deviations between the measured values and estimated values. The estimated value of a measurement is the data that would be provided by a perfect meter that knows the grid state (i.e. a meter that perfectly reflects the physical laws of grid operation - Kirchhoff and Ohm Laws).

Energy Balance Map

The energy balance map confronts two values along the feeders:

1. The value calculated by SINAPgrid using the energies billing (after the pro-rata calculation) together with the typical curves of the measurement campaign - "Flow_Value".
2. The value calculated by the state estimator from input measurements (from the base load software). The measurements can be located anywhere in the grid and are considered concomitantly - "State_Estimator_Value".

The difference is calculated as follows:

$$\text{Difference}(\%) = \frac{State\_Estimator\_Value - Flow\_Value}{State\_Estimator\_Value}$$

Finally, it is present an example of a map relative to the difference mentioned that can be observed in Figure 8.

Figure8. Energy balance

This results map is called "Energy Balance". It can also be used as an "exogenous parameter" to identify energy theft.

Polygons

Based on the exogenous parameters maps and the energy balance, the last software discussed on this article allows the user to draw regions, or take advantage of existing IBGE polygons to specify the distribution of the amount of losses due to energy theft. This mapping will be used to perform the demand adjustment at the end of the procedure.

The criteria for valuation of non-technical losses were classified into three qualitative categories (Categories of Energy Theft or CET). The risk of energy theft is a value between 0% and 100%. Initially, the following values will be displayed as default (according to the probability of energy theft):

· Category 1: Low Probability - 20% - green color
· Category 2: Medium Probability - 30% - yellow color
· Category 3: High Probability - 40% - red color

For all loads that do not belong to any polygon, it adopts 10% (Category 0 and gray color).

It is important to note that the percentages must necessarily add up to 100% (10 + 20 + 30 + 40). Then the software guarantees that even when the user changes some value in the advanced note settings, the others suit to add 100%. On the other hand, for a demand adjustment where, among the loads to be adjusted, there are not one or more categories, it is standardized as follows:

$$\text{New category value} = \frac{Old\ Category\ Value}{Existing\ Categories\ Sum}$$

In Figure 9, it is shown an example of qualitative values of polygons and their respective colors in SINAPgrid.

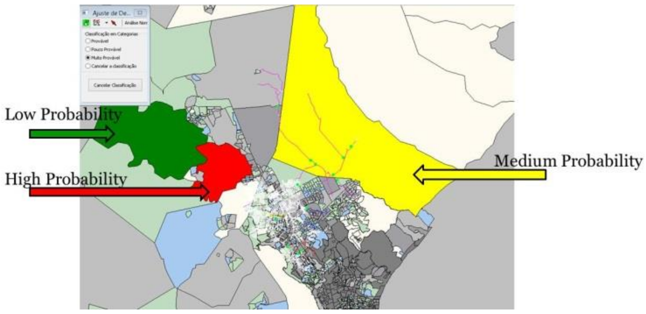


Figure9. Qualitative values of energy theft probability

Demand Adjustment

The planning of the electric system aims, in its primary function, to predict and plan the expansion and improve the grid needed to load attendance within the study horizon.

Thus, it is necessary to know the market currently served, represented by the electric demand of the installations that make up the distribution system throughout the 24 hours of the day.

Within this scope of approaching the electric calculations of the values measured by the concessionaire, the demand adjustment of SINAPgrid accepts measurements at any point in along the grid. These measurements may have been handled by base load software or can be inserted manually.

From external measurements, the following formula is used to calculate the correction factor of each load Individually (k). Note that k is the same for all loads that have the same "note" and are being considered in the adjustment:

$$kc = \frac{(CETc) * Difference + Dcategory}{Dcategory}$$

Where:

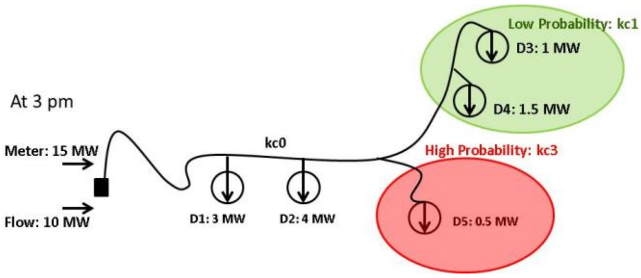
Kc = Load correction factor of a category

CETc = Risk of energy theft losses of a category

Difference = Hourly difference between measured total Demand (meter input data) and load Demand belonging to the region contemplated by this meter

Dcategory = Total category demand (power flow)

A simple example is presented to illustrate the problem:



$$C0 = \frac{10\%}{10\%+20\%+40\%} = 14\% \quad kc0 = \frac{(14\%)*5+(3+4)}{(3+4)} = 1.1$$

$$C1 = \frac{20\%}{10\%+20\%+40\%} = 29\% \quad kc1 = \frac{(29\%)*5+(1.5+1)}{(1.5+1)} = 1.6$$

$$C3 = \frac{40\%}{10\%+20\%+40\%} = 57\% \quad kc3 = \frac{(57\%)*5+(0.5)}{(0.5)} = 6.7$$

Figure10. Demand Adjustment Example

It is noted that the difference of 5 MW was divided unequally along the feeder, aiming to take advantage of tacit knowledge of the user through exogenous factors and other maps such as energy balance.

With this, the difference between the measurement and the power flow is located in greater weight where there is a high probability of energy theft.

CONCLUSION

The planner performs studies based on billing data and typical curves seeking to know the most real state of grid demand. Unfortunately, because of the high rates of non-technical losses, the results are very different within power flows and the operational measurements. Therefore, to make more reliable simulations of the Brazilian companies’ reality, it approaches the estimated values in the tools to the output measurements of feeders or even along them.

The methodology proposed in this paper is aimed at capturing the main differences in the simulation of electrical calculations between distributing the non-technical losses equally along the feeder or locating them more precisely. It is known that equal distribution can cause unrealistic rates of losses, quality, voltage and loading. Consequently, the defined process exposes the following tools:

·Development of a base load curve methodology

·Alignment of billing dates and month calendar (pro-rata)

· IBGE and auxiliary maps with exogenous factors

·State Estimation for comparison between measurements and billing values

·Design and identification of polygons with qualitative values of energy theft

·Differential demand adjustment according to the qualitative values of non-technical losses

Finally, a complete range of tools was provided and they can be used separately for many other utilities or completely used to approximate the demand representation of the electric grid to the reality from non-technical losses point of view.

REFERENCES

[1] J.A. Jardini, R.P. Casolari: Curvas de carga de consumidores e aplicações na engenharia de distribuição, São Paulo, 1999, ISBN 85-900839-1-8.

[2] J. A. Jardini, C.M.V. Tahan, M.R. Gouvêa, S.U. Ahn, F.M. Figueiredo: “Daily Load Profiles for Residencial, Commercial and Industrial Low Voltage Consumers”, IEEE Trans. Power Delivery, Vol. 15, pp. 375-380, 2000.