LOSSES FORECAST USING ECONOMETRIC MODELS

André MEFFE
Daimon – Brazil
andre.meffe@daimon.com.br

Carlos OLIVEIRA
Daimon – Brazil
barioni@daimon.com.br

Alden ANTUNES
Daimon – Brazil
alden@daimon.com.br

Paulo BAUMANN
Daimon – Brazil
paulo@daimon.com.br

Fernando LANGE
Daimon – Brazil
lange@daimon.com.br

Anderson UYEKITA
Daimon – Brazil
anderson.uyekita@daimon.com.br

Armando ROCHA
CEMIG D – Brazil
afrocha@cemig.com.br

Geraldo GUSMÃO
Axxiom – Brazil
geraldo.gusmao@axxiom.com.br

ABSTRACT

Forecasting the energy consumption for the next 5 years is a crucial task that has great influence in the power purchase agreements of distribution companies. However, that information is not enough to determine the amount of energy to be purchased for the upcoming years. The overall losses (technical and non-technical losses) must be added to the forecasted energy consumption so to determine a more accurate energy value that needs to be purchased. Therefore, forecasting technical and non-technical losses is a big problem that has to be addressed. The aim of this work is to present a methodology to forecast technical losses using econometric models.

INTRODUCTION

One of the main tasks of the Measurement and Losses Management Department at CEMIG D – a distribution company that supplies energy to over 8 million clients in the Southeast of Brazil – is the calculation and proposal of mitigating actions for the technical losses. The goal is to provide information regarding technical losses forecast for the short and medium terms to the company. Such information is extremely important so the Market Department can predict the power purchase, which has to be as close to the actual demand as possible. The distribution company is subject to penalties if it does not comply with the power purchase agreement, wherein there is a small error margin.

In order to obtain a reasonable efficiency in the power purchase prediction, one of the main issues to be tackled is the technical losses forecast. Those predictions are complex and normally connected to internal and external factors regarding the control held by the company. Among the internal factors one can highlight the expansion planning, which directly affects the technical losses forecast.

This work aims at presenting a methodology capable of forecasting the technical losses for the entire company, and for each region, power transformer, and feeder in an annual basis by means of non-data-demanding models. The proposed methodology was developed within a research & development project currently in progress, which main product will be a computational system for technical and non-technical losses forecasting.

In order to accomplish that goal a technical losses forecast model for each segment of the distribution system was developed. In the case of low voltage (LV) and medium voltage (MV) networks, econometric models were developed, wherein a very detailed analysis of explanatory variables was conducted, such as network length, conductor resistance, average current, load centre/radius ratio (parameter that explains the load distribution along the feeder, wherein the radius is the distance from the substation to the farthest distribution transformer), etc., and their influence in the technical losses. To obtain the econometric models the losses were determined through a power flow calculation and to make the technical losses forecast in each segment possible, the energy consumption forecast per substation is used. Besides the load growth impact, the forecasting model developed allows to evaluate the impact of the works plan and the switching plan in the technical losses.

Preliminary results show that the econometric models developed produce results with reasonable accuracy when compared to those obtained through power flow calculation.

TECHNICAL LOSSES FORECAST

In order to achieve the main goal of obtaining a technical losses forecasting model, this work was divided into three steps. The first step aims to define some explanatory variables for the losses in each segment of the distribution system. In the second step the aim is to define econometric models to obtain a correlation between the chosen variables and the technical losses in each segment. Finally, the third step aims at improving the econometric models by including the load forecast in the technical losses forecasting model. To do so, the following segments were considered: distribution substation, MV network, distribution transformer, LV network, customer connection and energy meter.

The existing simplified calculation models for transformers, customer connections and energy meters are very simple and depend on a few data. In the case of the energy meters, only the quantity of meters and the demand loss per energy meter are necessary. In the case of the transformers, only the rated losses (iron and copper losses) and the loading are required. Such simplicity is not observed in LV and MV networks calculation models. The wide variety in the distribution networks' characteristics combined to a random load makes the problem of obtaining simplified calculation models for these segments a complex task.

Several works have already approached the losses calculation issue in LV and MV networks [1] [2] [3] [5] [6] [7] [8]. Certainly, the load flow calculation will always be the most accurate method. However, this solution is not viable for accomplishing the losses forecast. The forecasting of the technical losses' future behaviour in some years ahead requires other calculation techniques.

Several techniques like that have already been developed, but the focus has always been more generalist, aiming at minimize the model error when it is applied to a set of distribution networks from all over the country. This work focuses on a reduced universe: CEMIG D's distribution networks. Thus, simplified calculation models are intended to be obtained, which must be more accurate than the existing ones and that allow accomplishing the losses forecast.

To achieve that goal, the forecasting model devised shall accomplish the losses forecast in two steps. The first one is intended to forecast a time series of the variable that represents the load by using historical data. Unlike the network attributes, the load historical data are frequently maintained by companies, either through registered measurements at the MV feeders or at the consumers.

In the second step a regression model will be used so to explain losses from the network attributes and load. The objective of this step is to calculate losses for a certain network and load configuration, at the present time or in the future. The next items detail the whole methodology.

Econometric Models

The LV and MV network's calculation models were obtained through a multivariate linear regression. Because of the physical nature of the problem (the loss is a non-linear function) the Cobb-Douglas production function was adopted [4]. Consequently, the problem became linear after taking the natural logarithm of the endogenous and exogenous variables. Some models with different combinations of exogenous variables were estimated, but only those which all coefficients were statistically significant at 5% level were evaluated.

The regression models must observe the following assumptions [4]: (i) strict exogeneity (error has zero conditional mean); and (ii) lack of collinearity (regressors are linearly independent). In the case of the MV network, the following variables were tested to explain the average demand loss: (i) load density distribution; (ii) action angle [°]; (iii) lateral conductor's resistance [Ω/km]; (iv) main feeder conductor's resistance [Ω/km]; (v) lateral length [km]; (vi) main feeder length [km]; (vii) number of distribution transformers; (viii) load centre/radius ratio; and (ix) average current [A].

In order to obtain the regression models 1,399 MV feeders of CEMIG D have been utilized considering all the variables mentioned. The chosen variables were retrieved from the corporate database and the average demand loss was obtained through Pertec, the software for technical losses calculation used by CEMIG D. By using the data sample mentioned to obtain regression models, it was verified that the lateral conductor's resistance, action angle and load density distribution were not significant variables. The regression model obtained through the remaining variables presented a determination coefficient of 0.918, complying with the assumptions previously mentioned. It also possesses the expected signal for its respective coefficients and, as a result, the average demand loss in kW can be calculated from equation (1):

$$loss_{MV} = 0.995 \cdot \exp\begin{bmatrix} -4.930 + 0.354 \cdot \ln(L_L) + 0.548 \cdot \ln(L_M) \\ +1.758 \cdot \ln(I_{avg}) - 0.166 \cdot \ln(N_{tr}) \\ +0.354 \cdot \ln(LCRR) + 0.601 \cdot \ln(R_M) \end{bmatrix} \quad (1)$$

wherein L_L is the lateral length [km]; L_M is the main feeder length [km]; I_{avg} is the average current [A]; N_{tr} is the number of distribution transformers; $LCRR$ is the load centre/radius ratio; and R_M is the main feeder conductor's resistance [Ω/km].

Another measure of accuracy, the ratio between mean of back transformed residuals and mean of observed losses (ROR), was calculated. This measure is calculated from the sum of the losses of all feeders (observed and predicted losses) and it represents the error in % of the losses in all observations. This measure was chosen because not only is important the accuracy of each individual prediction, but also the accuracy of the overall losses is important. For the MV network model the ROR is equal to 11.7%

In the case of the LV network, the following variables were tested to explain the average demand loss: (i) average current [A]; (ii) lateral length [km]; (iii) main circuit length [km]; (iv) LV network total length; (v) lateral conductor's resistance [Ω/km]; (vi) main circuit conductor's resistance [Ω/km]; (vii) load factor; (viii) transformer type (1-phase or 3-phase); (ix) LV network typology; (x) number of poles; (xi) number of clients; and (xii) loss coefficient [kW/(kVA/m^2)].

The LV network typology is an integer value ranging from 1 to 5 that refers to a standard network configuration according to Figure 1 [1]. One of those 5 configurations must be assigned to every single LV network.

The loss coefficient is constant for a given circuit, and it reflects the circuit's topology, conductor's resistance and voltage level in a convenient and compact way [9].

In order to obtain the regression models 235,950 LV circuits of CEMIG D have been utilized considering all the variables mentioned. The chosen variables were retrieved from the corporate database and the average demand loss was calculated through Pertec as well.

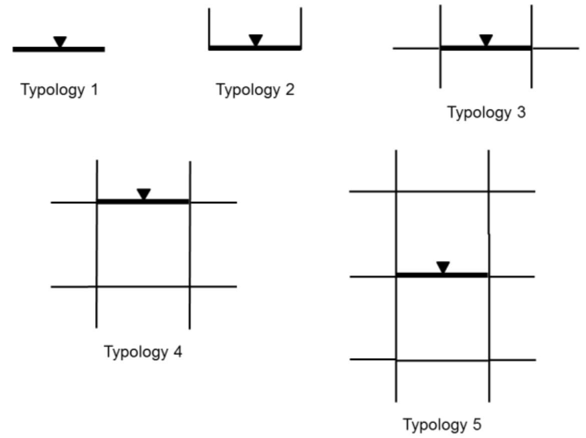

Figure 1. LV network typologies

In this case, two basic regression models were studied. The first model uses the average current, the main circuit length, and the lateral length. In the second model the LV network total length replaces the main circuit length and the lateral length. From the two basic models, all the possible combinations including the remaining variables were tested (512 models were generated). The models that violate some assumption or that present at least one coefficient signal contrary to expectations were discarded. Hence, 223 valid models were obtained. An analysis of the obtained results allowed choosing the model with the best trade-off between number of extra variables and deviation, which is the model 155. Figure 2 depicts a chart wherein the ROR regarding the valid models is plotted as a function of the determination coefficient. Each dot corresponds to a model and the assigned colour refers to the number of extra variables. It can be seen that model 155 does not have the highest determination coefficient, but it does have the lowest ROR, which is more important, since all models have determination coefficients that are similar.

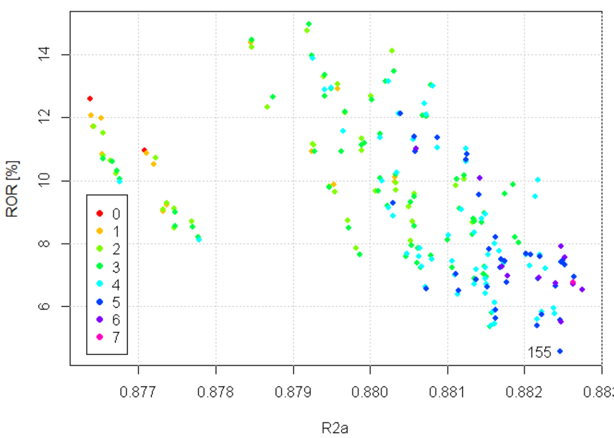

Figure 2. ROR as a function of the determination coefficient

The chosen model presented a determination coefficient of 0.882 and, as a result, the average demand loss in kW can be calculated from equation (2):

$$loss_{LV} = 1.118 \cdot \exp\begin{bmatrix} -6.538 + 0.417 \cdot \ln(L) + 1.614 \cdot \ln(I_{avg}) \\ +0.356 \cdot \ln(R_L) + 0.247 \cdot \ln(LC) \\ -0.308 \cdot \ln(N_P) + 0.286 \cdot \delta_{TT=3} \\ -0.010 \cdot \delta_{typ=2} - 0.022 \cdot \delta_{typ=3} \\ -0.091 \cdot \delta_{typ=4} - 0.193 \cdot \delta_{typ=4} \end{bmatrix} \quad (2)$$

wherein L is the LV network total length [km]; I_{avg} is the average current [A]; R_L is the lateral conductor's resistance [Ω/km]; LC is the loss coefficient [kW/(kVA/m^2)]; N_P is the number of poles; and the δ-variables are dummies that take the value 1.0 when its corresponding attribute (TT = transformer type; typ = typology) is equal to the value in the index, otherwise the dummies take the value 0.0. The model obtained for LV networks presents a ROR of 4.5%.

For the MV and LV networks' regression models the average current shall be calculated from the forecasted energy according to the procedure presented in the next item. In the case of the calculation classic models for the other segments, the energy obtained through that procedure shall be directly used when a variable for load is necessary.

One has to highlight that equations (1) and (2) allow calculating the average demand loss. However, the main goal is the determination of the energy loss, which is calculated from equation (3):

$$eloss = loss \cdot (1 + CV^2) \cdot T \quad (3)$$

wherein CV is the coefficient of variation (standard deviation/mean ratio), calculated from the load profile; and T is the time interval considered [h] (one day, month or year) [1].

It is important to notice that this work is currently in progress within a research & development project and the next steps consist of including the works plan and the switching plan in the calculation model of the MV network. Moreover, a computational system to calculate and forecast technical and non-technical losses will be developed.

Load Forecast

The regression models presented in the previous item combined to the calculation models of the other segments are responsible for the second step of the calculation of losses, which consists of calculating the energy losses for a certain month in the future. However, to make this calculation possible, a load forecast has to be performed because the losses depend on that.

Nowadays, CEMIG D has a methodology of its own to forecast the overall energy consumption as well as the energy consumption for each substation for the next 10 years. The available data are trustworthy and the best way to forecast the load aiming at the losses calculation, since the forecasted overall energy is already divided into its substations. However, it is necessary to match the forecasted energy per substation to the data that will be used in the losses calculation through the corresponding models. On the one hand, the annual energy per substation is available, but on the other hand, the monthly energy per feeder and per distribution transformer is required. Therefore, the annual energy per feeder must be calculated from the annual energy per substation. In a second step, the monthly energy per feeder and per distribution transformer must be calculated from the annual energy per feeder. At last, the average current required for the regression models can be calculated from the monthly energy in every single feeder and distribution transformer. Figure 3 portrays the calculation process.

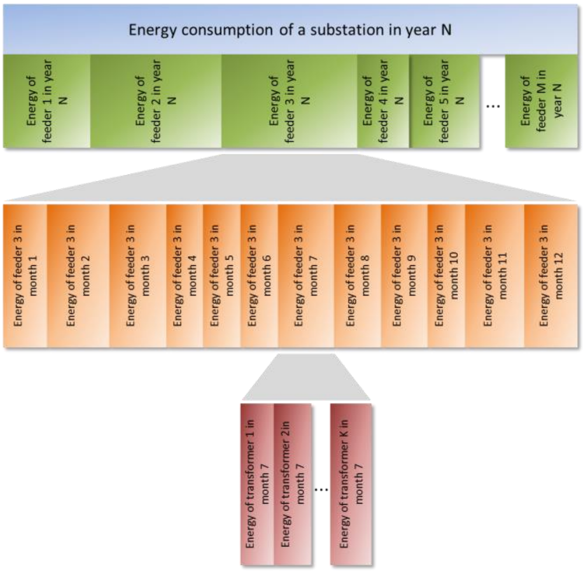

Figure 3. Transformation of the annual energy of a substation into monthly energy per feeder and per distribution transformer

In order to obtain the monthly energy per feeder and per distribution transformer a calculation in three steps is necessary. In the first step, the annual energy per feeder is calculated from the annual energy per substation. To do so, billing data per feeder are used so to establish a proportion of the annual energy per substation for each feeder and, therefore, the annual energy per feeder is obtained.

The second step comprises the calculation of the monthly energy per feeder from the annual energy according to the seasonality of consumption. Measurements are used to obtain the seasonality of each feeder, from which the annual energy can be divided into monthly energy per feeder through seasonality indices. The seasonality indices must be calculated using a 12-month interval and the measurements history. At the completion of a year, such indices must be updated so to represent the seasonality of each feeder in a proper way. Such indices are applied to the annual energy per feeder in order to obtain the monthly energy per feeder considering the seasonality. Thus, the forecasted monthly energy per feeder is obtained for any year of the 10-year study period.

Finally, the third step involves the calculation of the monthly energy per distribution transformer from the monthly energy per feeder. An apportionment of the monthly energy of a feeder can be accomplished using the billing data of its transformers so to obtain the monthly energy per transformer for any year of the 10-year study period.

CONCLUSIONS

This paper presented a methodology to forecast energy losses using econometric models. The methodology consists of two steps. The first one is intended to obtain the forecasted energy consumption in a monthly basis for each feeder and distribution transformer from the energy consumption forecast of each substation. The second step aims at calculating the losses of MV and LV networks through econometric models using the forecasted energy. For the other segments, the well-known classical models are used.

It is important to highlight that this work is currently in progress and the impact of the works plan and switching plan still have to be computed in the econometric models. Preliminary results show the econometric models have good accuracy regarding the overall losses in MV and LV networks in comparisons to load flow calculation However, the accuracy of the models per region considering all distribution segments still has to be checked and some adjustments may be necessary.

At the present time, the computational system is being developed and it must be finished by September 2015. At the end of this project, CEMIG D will have a powerful tool to forecast technical and non-technical losses and, therefore, the amount of energy that needs to be purchased in the next years will be more accurate. Furthermore, the company will be aware of its possible losses for the upcoming years.

REFERENCES

[1] ANEEL, 2011, “Aprimoramento da metodologia utilizada pela ANEEL para o cálculo das perdas de energia elétrica na distribuição constante do Módulo 7 – Cálculo de perdas na distribuição – do PRODIST”. Nota Técnica nº 014/2011, Brasília, Brazil.

[2] ANEEL, 2007, “Perdas técnicas nas instalações de distribuição – Descrição detalhada da metodologia”. Nota Técnica nº 035/2007, Brasília, Brazil.

[3] CODI, 1996, “Método para determinação, análise e otimização das perdas técnicas em sistemas de distribuição”. Relatório CODI-3.2.19.34.0, Brazil.

[4] Gujarati, D. N.; Porter, D. C., 2008, *Econometria*

básica, McGraw-Hill, Porto Alegre, Brazil.

[5] Meffe, A.; Oliveira, C. C. B., 2009, "Technical loss calculation by distribution system segment with corrections from measurements", *Proceedings 20th CIRED*, Prague, Czech Republic.

[6] Oliveira, C. C. B.; Kagan, N.; Meffe, A.; Jonathan, S.; Caparroz, S. L.; Cavaretti, J. L., 2001, "A new method for the computation of technical losses in electrical power distribution systems", *Proceedings 16th CIRED*, Amsterdam, Neetherlands.

[7] Queiroz, L. M. O., 2010, *Estimação e análise das perdas técnicas na distribuição de energia elétrica*, Tese de Doutorado, Universidade de Campinas, Campinas, Brazil.

[8] Roselli, M. A.; Yatsu, R. K., 2007, "Cálculo de perdas técnicas de energia em sistemas de distribuição", *Proceedings Congreso Internacional sobre Alta Tensión y Aislamiento Eléctrico – ALTAE*, Cuernavaca, Mexico.

[9] Schmidt, H. P.; Oliveira, C. C. B.; Meffe, A.; Rosa, M. A., 2003, "Loss estimation in LV circuits using intelligent techniques – The RGE experience", *Proceedings 17th CIRED*, Barcelona, Spain.