A METHODOLOGY TO REFINE THE TECHNICAL LOSSES CALCULATION FROM ESTIMATES OF NON-TECHNICAL LOSSES

Clainer DONADEL, João ANICIO, Marco FREDES
Energias do Brasil – Brazil
clainer.bravin@enbr.com.br, bosco.anicio@enbr.com.br, marco.fredes@enbr.com.br

Flávio VAREJÃO, Giovanni COMARELA, Gabriela PERIM
Computer Science Department of Federal University of Espírito Santo – Brazil
fvarejao@inf.ufes.br, giovannicomarela@gmail.com, gabrielaatp@gmail.com

ABSTRACT
This work proposes a methodology to improve the calculation of technical losses in power distribution networks using estimates of non-technical losses. The estimation procedure of non-technical losses is based on the history of energy consumption of consumers and on the results of field inspections. As motivation, we have the fact that the main methodologies to compute technical losses proposed in the literature do not consider the influence of non-technical losses, and hence are not very precise. The methodology here developed is very flexible, robust and can be applied even when the irregularities history maintained by the utility company is small. The clustering and sampling techniques used allow an estimation of technical losses closer to the real one.

INTRODUCTION
The high costs necessary to build new power-generating units, the preoccupation with the depletion of natural resources and, above all, the importance of energetic efficiency lead the utility companies to try to minimize as much as possible the energy losses in distribution power systems. Energy losses may be technical or non-technical.

Technical loss is the amount of energy consumed by Joule effect during the energy distribution process, caused by internal resistance of conduits and transmission/distribution equipments. Non-technical loss is the amount of energy bought by the power distribution company and not paid by its customers due to irregular connections to the network, measurement mistakes or measurement devices faults.

To reduce losses is necessary to know how much and where the energy is lost. For this reason, many works ([4], [5], [10] and [11]) have been developed for calculating technical losses in distributions systems. Oliveira et al. [8] claims that these methodologies do not consider all segments that form the distribution system. In this context, they propose a new methodology for the calculation of technical losses that consider all the segments. It is based on typical load curves to represent the customers load. However, the methodology described in [8] does not consider technical losses arising of energy circulation related to non-technical losses. Therefore, there is a difference between the sum of consumers measured energy with the technical losses calculated and the energy effectively distributed on the substations. Suriyamongkol [12] experimentally shows that the occurrence of non-technical losses may increase the technical losses by more than 20%, even though the former represents an increment of just 8% in the load. This difference is even greater in regions where irregular connections to the distribution system are a common practice. In these regions, the magnitude of non-technical losses is really more significant.

Seeking to improve the methodology defined in [8], the work presented in [9] proposes the calculation of an adjustment factor that allows correcting the loads circulating in the network, in order to consider the non-technical losses and get an adjusted value to the technical losses. However, this method of adjustment equally distributes the non-technical losses between the LT (Low Tension) and MT (Medium Tension) loads, and also ignores the fact that the amount of irregular connections and the profile consumption of different geographic regions may be very distinct.

This work proposes a method to improve the calculation of technical losses defined in [8] based on estimations of non-technical losses of a power distribution substation. These estimations are used for adjusting the measured energy of low and medium tension costumers so that the calculation of technical losses also takes into account the energy that circulated in the network due to non-technical losses. The adjustment performed is not uniform and considers features like geographic location and consumption profiles, and the activity of each costumer, bringing the methodology closer to the situations that occur in practice.

TECHNICAL AND NON-TECHNICAL LOSSES
Technical losses can be also classified by the segment that it occurs. According to the functionality that each distribution system component performs, [8] identifies eight segments: transmission and sub-transmission lines, substation power transformers, primary distribution lines, distribution transformers, secondary distribution lines, service drops and others. The others segment comprehends capacitors, voltage regulators, corona effect, connections, etc. Figure 1 shows these segments. The methodology to calculate technical losses proposed in [8] is based on losses calculation in each network segment. It is necessary a data base containing records of all customers, primary and secondary distribution lines, distribution transformers, substations and high tension system. The basic idea of this methodology consists of computing the energy loss of each segment considering their physical characteristics and load curves.
However, the methodology proposed in [8] does not consider the technical losses arising from the circulating energy related to non-technical losses. Méffe [9] proposes to apply a correction factor $k_c$ to the substation loads in order to adjust the value of technical losses. The correction factor is used to equalize the value of the measured energy in a substation to the estimated energy consumption $E_m$. The estimated energy consumption $E_m$ is expressed by the sum of the measured energy of MT and LT customers ($E_{MT}$ and $E_{LT}$) plus the variable and fixed losses ($E_{var}$ and $E_{fix}$):

$$E_{st} = E_{MT} + E_{LT} + E_{var} + E_{fix}.$$  

In order to equalize the measured $E_m$ and estimated energy $E_m$, Méffe proposes to apply the correction factor $k_c$ as follow:

$$E_{m} = k_c \times (E_{MT} + E_{LT}) + k^2_{c} \times E_{var} + E_{fix}.$$  

The correction factor may be obtained by solving equation (2). Then, the measured energy ($E_{MT}$ and $E_{LT}$) are multiplied by this factor and the value of technical losses is recalculated with the adjusted measured energy ($E_{MT}'$ and $E_{LT}'$).

Once the non-technical losses may be more concentrated on some geographic regions, Méffe recognizes the need of distributing the difference between the measured and estimated energy in a non-uniform way between the MT and LT loads.

This work proposes to compute distinct correction factors for the customers measured energy according to estimations of non-technical losses. These estimations are obtained from two types of historic data contained in the company data warehouse:

1. Results of consumers' inspections: the company regularly selects a large number of consumers for being inspected on site. The proportion of irregular consumers identified on these inspections in a specific geographic region is used as an estimation of the amount of consumers causing non technical losses in that region.

2. The energy consumption of the irregular consumers some months before and after the inspection date that identified the irregular connection: The difference between the average energy consumption before and after the inspection is used as an estimation of the amount of non technical energy loss caused by an irregular consumer.

THE METHODOLOGY

The methodology is compound by three steps. These steps are applied to the substation data in order to quantify the non technical losses during a specific period (for instance, a month):

1. Inspections Sampling;
2. Consumers Selection;
3. Consumption Adjustment.

Next sections explain these steps. It is important to emphasize that these procedures must be performed separately for low and medium tension consumers.

**Inspections Sampling**

Aiming to have reliable estimations of the amount of irregular consumers in a specific region we rely on statistical sample theory. Assuming that inspections were performed independently and randomly following the same probability distribution, the difference between the estimated proportion of irregularities $\hat{p}$ and the real proportion of irregularities $p$ should be smaller than an admissible error $\varepsilon$ with confidence greater than $\gamma$, i.e.:

$$P(|\hat{p} - p| < \varepsilon) > \gamma,$$

where $P$ is a probability measure.

From expression (3) it is possible to get the formula to compute the minimal number of inspections that ensures a good estimate of the proportion of irregularities in a specific region [1]:

$$n \geq \left[ \frac{1}{N} + 4 \left( \frac{\varepsilon}{z_{\gamma}} \right)^2 \right]^{-1}$$  

where:

- $n$ is the minimal number of inspections;
- $N$ is the number of customers in the region; and
- $z_{\gamma}$ is the quantile of normal standard distribution, which is computed from $\gamma$.

We assumed that there is similarity of energy consumption between consumers supplied by the same tension in a particular region. Therefore, inspections performed on consumers served by the same transformer may be grouped together for constituting a statistical sample of that region. However, just using this strategy was not sufficient because some groups could not have the minimal number of inspections ($n$) required for a reliable estimative.

In order to deal with this problem, we use a clustering...
algorithm ([3], [6] and [7]) for grouping data of inspections between geographically near distribution transformers. We have used a graph based divisive algorithm which guarantees that each cluster has enough number of inspections. The algorithm starts with the construction of a Minimal Spanning Tree (MST) [2] of the graph. This task is performed using the classical algorithms of Prim described in [2]. Then, the longest edges (those formed by the farthest vertexes) of the MST are removed, while the connected components (clusters) have enough inspections. Figure 2 shows an illustrative example of how a cluster is partitioned. If the sub graphs $S_1$ and $S_2$ have enough inspections, the original sub graph is divided and the procedure is recursively repeated for $S_1$ and $S_2$.

![Image](image.png)

**Figure 2:** A graph based clustering approach.

**Consumer Selection**

This step aims to select the group of consumers of the substation that must have their consumption adjusted. The selection is performed individually for each cluster identified on the Inspection Sampling step in order to compute more specific correction factors. We consider each consumer belonging to the cluster that contains its transformer.

Since we do not really know the specific consumers irregularly connected to the distribution network, we took the approximate approach of selecting the consumers which energy consumption are closer to the energy consumption mean of the cluster consumers. This approach avoids randomness in the adjustment process and also the influence of outliers.

The proportion of irregularities $I_j$ found in a cluster $G_j$ is given by (5)

$$I_j = \frac{\sum_{i=1}^{T_j} Irregularities_i}{\sum_{i=1}^{T_j} Inspections_i},$$

where $Inspections_i$ is the number of performed inspections in transformers $t_i$ and $Irregularities_i$ is the number of irregularities identified on these inspections. $T_j$ is the number of transformers belonging to cluster $G_j$.

The number of consumers $F_j$ selected in each cluster is given by (6)

$$F_j = \left[ I_j \times Total \ of \ Consumers \ in \ G_j \right].$$

**Consumption Adjustment**

The energy consumption of each consumer selected in the former step is adjusted by a correction factor. We compute this factor based on the energy consumption history of the inspected clients identified as irregularly connected. We can verify an increase on the energy bills of these clients after the inspection. We assume this increase is the amount of energy consumed without being registered that causes the non-technical losses. The following information is used on this step:

- The *class* (residential, commercial, industrial and rural), *service class* (single-phase, two-phase and three-phase) and *location* (feeder, municipal district, state region and state) of the consumers selected;
- The consumption history of consumers identified as irregularly connected by the inspections. Our approach requires the energy consumptions of $d_b$ months before and $d_a$ months after the inspection.

The energy absolute variation ($\Delta$) and relative variation ($\delta$) of a consumer identified as irregularly connected by an inspection are given by following equations:

$$\Delta = \frac{1}{d_a - d_b} \sum_{i=d_b}^{d_a} c_{+i} - \frac{1}{d_b} \sum_{i=1}^{d_b} c_{-i} \quad \text{and} \quad (7)$$

$$\delta = \frac{\Delta}{\frac{1}{d_b} \sum_{i=1}^{d_b} c_{-i}}, \quad (8)$$

where:
- $c_{+i}$ is the registered energy consumption $i$ months after the inspection;
- $c_{-i}$ is the registered energy consumption $i$ months before the inspection;
- $d_b$ is the number of months after the inspection that must be disregarded in the adjustment model (some energy bills still keep residues from earlier bills some months after the inspection).

The correction factor of each consumer selected is the consumption variation mean of irregular consumers of the substation with the same *class, service class* and *location* of the consumer. Whereas the correction factor $f_i$ for clients with zero (or close to zero) consumption is the mean of the absolute variation $\Delta$, the correction factor $f_a$ of the other clients is the mean of the relative variation $\delta$.

We use these correction factors for computing the adjusted consumption $\text{cons}_{\text{new}}$ in the following way ($\text{cons}_{\text{old}}$ is the measured consumer consumption):

$$\text{cons}_{\text{new}} = \begin{cases} (1 + f_a) \times \text{cons}_{\text{old}} & \text{if } \text{cons}_{\text{old}} \neq 0; \\ f_a + \text{cons}_{\text{old}}, & \text{otherwise}. \end{cases} \quad (9)$$

After these three steps, the consumption of some consumers LT and MT of the substation were adjusted and these new values are used for computing the non-
technical losses according to the method proposed in [8].

CONCLUSION

The methodology proposed here was implemented as a computational system developed for ESCELSA (Esperito Santo Centrais Elétricas S.A.), a brazilian energy distribution company that sponsored this research. We had run the computational system on fifteen distribution substations along three months for evaluating its results. Table 1 presents the difference (in percentage) on the technical losses calculation when the methodology proposed in [9] is applied without our approach and using it. Positive values mean an increase on the technical losses, while negative values mean a decrease.

Table 1: Technical Losses Differences

<table>
<thead>
<tr>
<th>Substation</th>
<th>jan/08</th>
<th>feb/08</th>
<th>mar/08</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.93%</td>
<td>-0.28%</td>
<td>-1.85%</td>
</tr>
<tr>
<td>2</td>
<td>-0.40%</td>
<td>-0.97%</td>
<td>-1.85%</td>
</tr>
<tr>
<td>3</td>
<td>-3.39%</td>
<td>-1.59%</td>
<td>-2.67%</td>
</tr>
<tr>
<td>4</td>
<td>-6.76%</td>
<td>-7.16%</td>
<td>-6.55%</td>
</tr>
<tr>
<td>5</td>
<td>0.99%</td>
<td>1.15%</td>
<td>1.07%</td>
</tr>
<tr>
<td>6</td>
<td>-7.39%</td>
<td>-6.73%</td>
<td>-6.83%</td>
</tr>
<tr>
<td>7</td>
<td>0.64%</td>
<td>0.05%</td>
<td>-0.55%</td>
</tr>
<tr>
<td>8</td>
<td>0.35%</td>
<td>0.67%</td>
<td>-1.05%</td>
</tr>
<tr>
<td>9</td>
<td>4.05%</td>
<td>2.74%</td>
<td>2.43%</td>
</tr>
<tr>
<td>10</td>
<td>-3.59%</td>
<td>-4.81%</td>
<td>-2.40%</td>
</tr>
<tr>
<td>11</td>
<td>-1.16%</td>
<td>-1.62%</td>
<td>-1.71%</td>
</tr>
<tr>
<td>12</td>
<td>-0.98%</td>
<td>-0.33%</td>
<td>-2.99%</td>
</tr>
<tr>
<td>13</td>
<td>3.22%</td>
<td>3.94%</td>
<td>3.67%</td>
</tr>
<tr>
<td>14</td>
<td>1.60%</td>
<td>2.02%</td>
<td>1.67%</td>
</tr>
<tr>
<td>15</td>
<td>0.97%</td>
<td>-0.21%</td>
<td>-0.40%</td>
</tr>
</tbody>
</table>

As it may be noticed, there were significant differences on the technical losses calculation (from -7.39% to +4.05%). These results confirm that it is important to consider where the non-technical losses occur in order to have a more accurate technical losses calculation. Therefore, our approach is an improvement to the methodologies of non-technical losses calculation described in [8] and [9].

We emphasize the following advantages of our method:

- Our approach considers the estimate of non-technical losses directly in technical losses calculation, distributing it in a non uniform way between LT and MT consumers located on different geographical regions. Instead of equally treating all consumers of a substation (such as made by [9]), we consider the type of consumer and their location to adjust its consumption;
- The inspections sampling step allows for applying our method to substations with a small number of inspections, while still guaranteeing good estimates. Indeed, the proposed procedure is very flexible because it is adaptable to the number of inspections available;
- The step of consumers selection is robust because is capable of discarding the influence of discrepant consumption consumers in each cluster;
- Although the usual methods of clustering and classification used in literature are randomized, all steps of the proposed methodology were carefully constructed to avoid randomness. Therefore, the same results are always achieved when the procedure is performed on the same data.

REFERENCES