Estimation of Losses in LV Distribution Grids - Combined Impact of Load Variability and Phase Imbalance

Alexandre M. V. Gouveia

Abstract—The purpose of this work is to evaluate the impact of load variability and phase imbalance on low-voltage (LV) grid losses, validating and extending previous work done on this subject. Validation will be achieved by running power flow simulations on reference LV grids, for different phase imbalance scenarios and meter resolutions, using real consumption data. The results obtained were in line with expectation and with the theoretical predictions established beforehand. These results are then extended for higher meter resolutions, where real-world data is lacking. A regression model is created to predict losses based on consumers' load variance and grid structure, including imbalance. An analysis of residential load variance behaviour for higher meter resolutions is also performed. These models combined estimated an increase of 13.38% of the total LV system losses, caused by higher resolution measuring and an improved modeling of three-phase load imbalance. Implications and limitations of the results obtained are considered and further research and data collection suggestions are discussed.

Index Terms—Grid losses, load variance, phase imbalance, linear regression, losses estimation, meter resolution.

I. INTRODUCTION

The ability to accurately estimate low-voltage (LV) distribution losses can be of great value, particularly to distribution network operators (DNOs). Being able to distinguish between technical losses (Joule losses caused by heating of conductors) and non-technical losses (accounting or metering errors, meter tampering or outright energy theft), DNOs can determine where money is being lost and where they should allocate resources in an effort to increase system-wide efficiency and, with it, the DNO's profits. Also, the reduction of losses has been emphasized by regulatory agencies and performance awards are given to stimulate this practice.

It remains, however, quite difficult to estimate losses on LV grids, where the existence of non-technical losses is predominant. In contrast with high-voltage (HV) and medium-voltage (MV) grids, LV grids are usually not very well characterized. Due to their size, number, and the cost of adequate metering devices, grid layouts and consumers' load profiles are not accurately recorded. This imposes a problem when one tries to estimate technical losses on these grids. As most of the total energy losses occur in LV grids, it is imperative to tackle this loss estimation problem. In this paper, and in an extension of what has been done in [1], a study will be conducted on the influence of load variability, phase imbalance and their combined effect on LV losses, focusing especially on higher meter resolutions (i.e., higher than 15 minutes).

Past work has been done on the problem of accurately estimating technical losses on LV grids. Attempts were made to correct the loss factors currently in use as rules of thumb, which were proved inadequate. Developing a new loss factor that would reflect the grid's layout, combined with an accurate measurement of per-phase currents at the transformer, yielded somewhat reliable results when compared with losses computed through a net energetic balance of a grid [2]. In [3], losses are estimated by clustering regions according to intensity of consumption and consumer density. Then, using well characterized feeders, loss functions are defined for each cluster and losses are estimated from the load carried by each feeder.

Other authors have studied the impacts of phase imbalance and load variability separately. The effects of imbalance were analyzed through a complex unbalance factor, whose magnitude and argument could be used as indicators for line losses [4]. In a similar fashion to [3], through the clustering of grids and the use of well defined, representative grids, an attempt to convey the impact of phase imbalance on distribution losses was made [5]. With regards to time variability, the relationship between metering resolution and errors in estimated losses was investigated using 1 minute resolution data and the underestimation of losses caused by low resolution of meters was demonstrated [6]. New methods for computing loss factors that take into account load variability - and the subsequent instantaneous imbalance caused by said variability - were developed as well [7]. However valid and important past contributions to this area of study have been, they all lack a consideration for the combined effect and interdependence of phase imbalance and load variability on distribution losses.

As stated before, this thesis will serve as a continuation to the work done in [1]. In that paper, the difficulties in modelling phase imbalance, load time variability and, more importantly, their combined effect are detailed. It was explained that losses are dependant on per-phase loading, since Joule losses are quadratic w.r.t. phase currents and assuming balanced (averaged) currents underestimates losses. Also, imbalance induces neutral current, which can be nearly as high as phase currents and its contribution to losses is non-negligible [7], [8]. With regards to time variance, it was established that the time resolution used in measurement of loads would impact its observed magnitude and that would affect any estimation on losses made with these measurements. Since

A.M.V. Gouveia is with Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal (e-mail: alexandre.venancio.gouveia@tecnico.ulisboa.pt)
most metering devices integrate instantaneous power in order to obtain the consumed energy in a given time interval, only an average power on that time interval can be obtained from these measurements and the power oscillations within said interval are lost. The loss of the spiky characteristic of the consumers’ load curve can lead to an underestimation in loss estimation, even when loads are aggregated [6]. These two phenomena combined have a particular effect on losses, since they also interfere with each other (i.e., time variance increases the importance of imbalance and vice-versa). This added difficulty makes any attempt to model these two factors independently, as is done in [4]–[7], incomplete.

Further in [1], the effects on losses of load aspect are considered. It was demonstrated that by defining relative load imbalance, $u$, as

$$u(p) = \left( \frac{p \cdot p}{3 p^2} - 1 \right) / 2,$$  \(1\)

where $p$ is a vector containing per-phase currents, the ratio between losses in imbalanced and balanced cases (i.e., relative losses) would be linear w.r.t. $u$:

$$l^u(u) = 1 + 5 \ u.$$  \(2\)

As for time variability, the different meter resolutions were mimicked by creating a phase current time-sequence, $s$, and averaging elements of $s$ within integration periods, $\tau$, which were defined by the number of time-stamps in each interval. It was shown that, for a perfectly balanced case, relative losses would be linear w.r.t. the inverse of $\tau$: the relative frequency, $f$:

$$l^f(f) = 1 + \sigma^2_R \ f$$  \(3\)

with a constant slope given by the relative variance of $s$, $\sigma^2_R = \sigma^2 \ E[s]^2$. With per-phase independent sampling of currents, the balance of phases is disrupted with the appearance of the emergent imbalance, caused by each phase-current’s independent and non-synchronous variability. This type of unbalance is instantaneous and vanishes when averaged currents are considered, in opposition to structural imbalance, that is observed even at total integration of the currents. To account for the interaction between variability and imbalance, (3) is changed to

$$l^f(f) = 1 + \sigma^2_R \left[ 1 + \frac{1}{(1 + 5 u_{1/T})} \right] f$$  \(4\)

where $u_{1/T}$ is the structural relative unbalance.

In these assessments, it was assumed that each value of the load’s time-sequence was an independent and identically distributed (i.i.d.) random variable, sampled from a beta probability density function (PDF). The final combined effect of load and phase variability on losses for a single load is given by the product of (2) and (4)

$$l^{uf}(u, f) = (1 + 5 \ u) \left( 1 + \sigma^2_R \left[ 1 + \frac{1}{(1 + 5 u_{1/T})} \right] f \right)$$  \(5\)

Consider multiple loads distributed in space, it is noted that if the loads were uniformly distributed along a feeder, $F$, and were i.i.d., then the reduction of variance of phase-currents, caused by load aggregation, was not enough to ignore high frequency variations (and their contribution to losses) when the loads were imbalanced. An analytic expression is derived for the relative losses in feeder $F$ w.r.t. to the relative frequency and the variance and imbalance of the loads present in $F$ and gives insight into what this combined impact might look like in more complex and generic grid configurations:

$$l^{uf}(u, f) = 1 + \frac{5 u + 2 \sigma_R^2 f}{2 M + 1} + \frac{5 u \times \sigma^2_R f}{(M + 1)(2 M + 1)},$$  \(6\)

in which $M$ is the number of loads in feeder $F$.

After running power-flow simulations on a 400kVA reference grid, using both PDF and Markov Chain generated load profiles [9], [10], it is concluded that the increase in losses at the 15 minute resolution is still significant, albeit at a slower pace than observed for lower time resolutions, and warrants further investigation into what these losses might look like for higher time resolutions. This is what will be attempted here.

This paper seeks to extend the curves obtained beyond the 15 minute resolution and confirm that, despite slowing down, the relative losses’ increase is still considerable at finer time resolutions. In Section II, the results obtained in [1] are validated by simulating real consumption data onto a typical LV grid; in Section III the difficulties in modelling complex grids are detailed and a simplified approach is suggested; in Section IV a method to estimate the variance increase caused by a change in meter resolution is presented; in Section V the increase in relative losses is predicted based on load variances, in specific imbalance scenarios; in Section VI the estimations made for a selection of reference grids are extrapolated to the entire LV system; in Section VII the paper is concluded, with the results obtained and their implications being discussed.

II. SIMULATING LOADS WITH LOW METER RESOLUTION

In order to validate the theory developed and preliminary results obtained in [1], power flow simulations were run on typical LV grids using real-world consumption data, with 15 minute resolution, recorded and provided by EDP-Distribuição (portuguese DNO). From this consumption data-set, it was extracted and attributed, to each customer in each grid, one day’s worth of data, assuring that the contracted power of the virtual and real customers would be the same. In essence, a time-sequence of active power, $P_t$, was created for each load. Then, each load profile had its values averaged in groups of 2, 4 and 96 to mimic the results that would have been obtained if different meter resolutions were used (in this case, half-hourly, hourly and daily). The number of time stamps used in each group in this averaging process was defined in [1] as the relative integration period, $\tau$, with its inverse being the relative frequency, $f$. This attribution process was repeated 30 times to ensure some statistical significance from the results.

To encapsulate the effects of structural imbalance, present even on average and not caused by load volatility, three imbalance scenarios were created, similar to those in [1]. They are:
TABLE I
CUSTOMER SIZES AND QUANTITIES

<table>
<thead>
<tr>
<th>$S_c$ [kVA]</th>
<th>1.15</th>
<th>2.3</th>
<th>3.45</th>
<th>4.6</th>
<th>5.75</th>
<th>6.9</th>
<th>10.35</th>
<th>13.8</th>
<th>17.25</th>
<th>27.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>5</td>
<td>3</td>
<td>48</td>
<td>8</td>
<td>3</td>
<td>38</td>
<td>15</td>
<td>6</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

TABLE II
CONDUCTOR CROSS-SECTIONS AND LENGTHS

Underground Cables

<table>
<thead>
<tr>
<th>Cross Section [mm$^2$]</th>
<th>185</th>
<th>95</th>
<th>50</th>
<th>35</th>
<th>25</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length [m]</td>
<td>92</td>
<td>206</td>
<td>302</td>
<td>212</td>
<td>154</td>
<td>330</td>
</tr>
<tr>
<td>Length [%]</td>
<td>4.5</td>
<td>10.1</td>
<td>14.9</td>
<td>10.4</td>
<td>7.6</td>
<td>16.3</td>
</tr>
</tbody>
</table>

Overhead Lines

<table>
<thead>
<tr>
<th>Cross Section [mm$^2$]</th>
<th>70</th>
<th>50</th>
<th>35</th>
<th>25</th>
<th>16</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length [m]</td>
<td>1</td>
<td>151</td>
<td>288</td>
<td>30</td>
<td>151</td>
<td>115</td>
</tr>
<tr>
<td>Length [%]</td>
<td>0.0</td>
<td>7.4</td>
<td>14.2</td>
<td>1.5</td>
<td>7.4</td>
<td>5.7</td>
</tr>
</tbody>
</table>

1) All customer loads are considered to be balanced and are distributed evenly among the three phases;
2) Customer loads with $S_c < 10.35$ kVA are considered single-phase connected and the corresponding loads are randomly assigned to a specific phase. The remaining loads are considered balanced and are distributed evenly among the three phases;
3) All customer loads are considered imbalanced and are assigned randomly to a specific phase.

After assigning a phase to each customer (or making them three-phase), power flow simulations are run on each scenario and for each of the 30 sets, and the total Joule losses are computed. This process is repeated for the different $\tau$ values mentioned above (except for the balanced case, where only $\tau = 96$ is simulated and whose results will be used as reference). With the values of losses in absolute values, it is possible to compute the relative losses by dividing each value obtained by its corresponding reference value (the one obtained when simulating that set of load profiles with $\tau = 96$ in scenario 1).

The grids used for simulation were five reference grids, representative of the LV grids in the entire network and defined by their transformer’s nominal power (50kVA, 100/160kVA, 250kVA, 400kVA and 630kVA) [11]. In this work, the results of the 400kVA grid are shown in graphical form as an example and its geographic view is displayed in Figure 1. The list of contracted powers and the number of customers for each power is present in Table I and information regarding cable cross-sections can be seen in Table II.

In Figure 2 the results for relative losses are plotted onto relative frequency, and it can be observed the average trend as well as the distribution of results for each frequency value. It can be observed that the trends are similar to those obtained when simulating the grid with Markov Chain generated load profiles and that structural imbalance plays an important role in aggravating the effects of load variance and vice-versa.

Looking only at imbalance it would have been concluded that average losses increased by 17.91% and by 68.38% when the imbalance changed from scenario 1 to 2 and from 1 to 3, respectively, when considering the flat profiles. Looking at load variances, the increase in resolution from daily to quarter-hourly would increase losses by 12.37% in scenario 2 and by 26.87% in scenario 3. Analyzing the combined effect however, reveals that losses increase by 32.50% and by 113.62%, when resolution is increased to 15 minutes and imbalance changes from scenario 1 to 2 and from 1 to 3, respectively. Note that neither of these scenarios is realistic; it is unlikely that all loads in a grid are single-phase, especially some of the larger ones that are present in this grid, and that three-phase loads are always balanced throughout time. They give us, however, an upper and lower bound for relative losses on this grid, as a realistic result would be somewhere between these two plots.

As stated before, these simulations were run on several reference LV grids. The results obtained through the previously described process, applied to all grids, are summarized in Tables III and IV.

In the next sections an attempt will be made to extend these results for higher resolutions and explain the difficulties.
Fig. 2. Boxplots and mean curves for the relative losses obtained in 30 power flow simulations, for imbalance scenario 2 (above) and 3 (below). Mean curves were not computed by averaging relative losses, but by computing the average absolute valued losses on a given scenario and dividing it by the average absolute valued losses for daily resolution, on the reference scenario (1).

TABLE IV

<table>
<thead>
<tr>
<th>Grids [kVA]</th>
<th>Some Imbalanced</th>
<th>All Imbalanced</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.189</td>
<td>4.400</td>
</tr>
<tr>
<td>100</td>
<td>1.429</td>
<td>1.841</td>
</tr>
<tr>
<td>250</td>
<td>1.242</td>
<td>1.575</td>
</tr>
<tr>
<td>400</td>
<td>1.143</td>
<td>1.325</td>
</tr>
<tr>
<td>630</td>
<td>1.135</td>
<td>1.225</td>
</tr>
</tbody>
</table>

associated with this attempt.

III. MODELLING LOSSES FOR NON-TRIVIAL GRIDS

In [1], it was postulated that relative losses would be linear w.r.t. relative imbalance and relative frequency, even when loads were aggregated in a linear feeder. This fact stands on several assumptions that, while valid and useful on an initial approach, might not be realistically upheld in reality. Consumers’ loads rarely display the same variance and this variance does not grow linearly with frequency. In Figure 3 the evolution of the average relative variance (σ²_R) w.r.t. frequency is displayed. These values were obtained using 1 minute resolution load data measured at 22 residential consumers, for two years [12]. It can clearly be seen that linearity does not hold. Therefore, since no clear relationship between variance and relative frequency can be established and it is variance that increases losses, not necessarily frequency, the influence of load variance on losses will be analyzed directly, without considering which frequency might have generated loads with that variance.

For i.i.d. loads spread uniformly across a linear feeder, it is expected that relative losses increase linearly with the loads’ relative variance, which would be the same across every load [1]. As mentioned above, this is an ideal and theoretical scenario and in reality loads have diverse probabilistic distributions and LV grids in use are more complex than linear and uniform feeders. In Figure 4, it can be observed that, even though relative losses do indeed increase w.r.t. average relative variance, this increase is not exactly linear and no grid-wise relation can be made, as similar values of average relative variance can lead to disparate values of losses.

A. A Trivial Case

In order to obtain a better understanding of the behaviour of losses for more diverse sets of loads, a simple two load uniform feeder is studied, as the one shown in Figure 5. Since the data available is concerning power consumption, losses will be estimated based on power instead of currents, which are interchangeable if a constant power factor and unitary elasticity of loads w.r.t. voltage is assumed. Obviously there is no power flow in the neutral conductor, but in this example, and for simplicity’s sake, the neutral current will be defined by the vectorial sum of the power demands that would have caused each phase current at nominal voltage. As Joule losses...
Fig. 4. Curves displaying the relative losses w.r.t. average relative variance, obtained by simulating each set of demand time-sequences for imbalance scenario 2 (above) and 3 (below).

are quadratic w.r.t. current, these losses can be defined as:

$$L(P(t)) = \sum_i C P_i(t)^2$$  \hspace{1cm} (7)

where $P_i(t)$ is the power flowing in section $i$ at time $t$ and $C$ is a constant that captures the line segment’s resistance and the relationship between current and power demand, which is assumed to be equal for every segment of the feeder. To observe the effects of load variability, relative losses are defined as:

$$l^P(P) = \frac{T/\tau}{T L(E[P_i])} \sum_i L(P_i)$$  \hspace{1cm} (8)

where $E[\cdot]$ is the expected value of the time sequence.

With the assumptions made and the conditions of this problem, it can be shown that (8) can be expanded upon and generalized for $N$ loads, resulting in:

$$l^P(P) = 1 + \frac{2 \sum_{k=1}^{N} \sum_{n,m=k}^{N} \text{cov}[P_{n}^{\phi_{n}}, P_{m}^{\phi_{m}}]}{2 \sum_{k=1}^{N} \sum_{n,m=k}^{N} E[P_{n}^{\phi_{n}}] E[P_{m}^{\phi_{m}}]} - \sum_{k=1}^{N} \sum_{n,m=k}^{N} \text{cov}[P_{n}^{\phi_{n}}, P_{m}^{\phi_{m}}]$$

$$- \sum_{k=1}^{N} \sum_{n,m=k}^{N} E[P_{n}^{\phi_{n}}] E[P_{m}^{\phi_{m}}]$$

where $\text{cov}[\cdot, \cdot]$ is the covariance function. The proof for deriving (9) from (8) is in Section A. In the first half of (9), the losses on each segment $k$ of the feeder are computed by adding the covariances of every pair of power demand time-sequences of loads that are downstream of segment $k$ and that are on the same phase. The losses on each segment are then summed. As for the neutral losses, these are obtained by doubling the phase losses and subtracting the covariance pairs of downstream loads that are connected on different phases, which represent the degree of balance and thus reduces the neutral current, that, by consequence, reduces the neutral losses.

B. Generalizing into a Model

With (9), a better understanding of how relative losses are affected by load variability is obtained and, by observing its structure, it can be used to estimate losses for higher resolutions, for which complete power demand time-sequences are not available. To make this estimation, a linear model based on the loads’ variances and the covariances of the time-sequences of total per-phase loads is developed, which will encapsulate the effects of the first and second parts of (9), respectively. This can be explained by noting that in the specific case of LV loads’ time-sequences, variances have generally a higher value than any covariance pair possible, especially as relative frequency is increased. Even though aggregated load profiles display known intra-day dynamics, such as peak hours for instance, and thus individual loads are correlated with each other, this correlation decreases as time resolution is increased, as the likelihood of coordinated activity decreases when analyzed with finer detail (e.g., most people might have their morning routine between 7 and 8AM, and therefore their consumption peak will be within that period, each individual’s peak will be randomly distributed in smaller time intervals in that one hour period). The reduction in correlation and increase in variance makes covariance stay somewhat stable w.r.t. relative frequency, since, using Pearson’s correlation coefficient:

$$\rho_{P_n, P_m} = \frac{\text{cov}[P_n^{\phi_n}, P_m^{\phi_m}]}{\sigma_{P_n} \sigma_{P_m}} \Rightarrow \text{cov}[P_n^{\phi_n}, P_m^{\phi_m}] = \rho_{P_n, P_m} \sigma_{P_n} \sigma_{P_m}.$$  \hspace{1cm} (10)
With this in mind, the \( \sum_{k=1}^{N} \sum_{n,m=k}^{N} \text{cov}[P_{n}^{\phi_{n}}, P_{m}^{\phi_{m}}] \) term in (9) can be reduced to the terms where \( n = m \), which results in:
\[
\sum_{k=1}^{N} \sum_{n,m=k}^{N} \text{cov}[P_{n}^{\phi_{n}}, P_{m}^{\phi_{m}}] \approx \sum_{k=1}^{N} k \sigma_{\phi_{k}}^{2}.
\]
(11)

The term \( \sum_{k=1}^{N} \sum_{n,m=k}^{N} \text{cov}[P_{n}^{\phi_{n}}, P_{m}^{\phi_{m}}] \) represents, as mentioned above, the level of unbalance observed in the grid in question. The accurate computation of this term will scale badly in terms of computation time and memory required. In an attempt to represent the same effect, this term will be replaced by:
\[
\sum_{k=1}^{N} \sum_{n,m=k}^{N} \text{cov}[P_{n}^{\phi_{n}}, P_{m}^{\phi_{m}}] \approx \text{cov}[\Sigma P^{A}, \Sigma P^{B}] + \text{cov}[\Sigma P^{B}, \Sigma P^{C}] + \text{cov}[\Sigma P^{C}, \Sigma P^{A}]
\]
(12)
in which the covariances between the total loads connected to each phase are computed.

Combining these alterations, (9) can be written in the approximated form:
\[
l^2(P) \approx 1 + \sum_{k=1}^{N} \sum_{\phi_{n} \neq \phi_{m}} \left( \text{cov}[\Sigma P^{A}, \Sigma P^{B}] + \text{cov}[\Sigma P^{B}, \Sigma P^{C}] + \text{cov}[\Sigma P^{C}, \Sigma P^{A}] \right)
\]
(13)

Since the expected value of a time-sequence is invariant to relative frequency, (13) can be generalized into a linear model that can be used in non-trivial grids. This model is defined as:
\[
l^{u}(\tau, u) = \alpha_{0} + \sum_{\tau}^{T} \alpha_{\tau}^{u} + \sum_{\tau}^{T} \alpha_{\tau u}^{\phi-\phi} + \alpha_{cov}^{u}
\]
(14)

where \( \alpha_{\tau}^{u} \) and \( \alpha_{cov}^{u} \) are the coefficient vectors and whose elements are the weight of each load’s variance and each phase-to-phase covariance on relative losses, defined by how the loads are connected in the grid and how imbalanced the phases are, denoted by index \( u \). \( \alpha_{0} \) is the intercept and is expected to be 1. To obtain these coefficients, the results of the previous simulations using 15 minute resolution data and lower will be used, creating a set of coefficients for each imbalance scenario. The losses, the variances of the power time-sequences and the phase-to-phase covariances are aggregated into a vector and into matrices, in such a way that each row represents the simulation of one set of power demand time-sequences, on a given resolution. After aggregating them, a new equation can be written based on a system of (14) like equations:
\[
l^{u}(u) = \alpha_{0} + \sum_{\tau}^{T} \alpha_{\tau}^{u} + \text{Cov}_{u}^{\phi-\phi} \alpha_{cov}^{u}
\]
(15)

where
\[
\text{Cov}_{u}^{\phi-\phi} = \begin{bmatrix}
\sigma_{10} & \sigma_{20} & \ldots & \sigma_{N0} \\
\sigma_{10} & \sigma_{20} & \ldots & \sigma_{N0} \\
\sigma_{10} & \sigma_{20} & \ldots & \sigma_{N0} \\
\sigma_{10} & \sigma_{20} & \ldots & \sigma_{N0} \\
\end{bmatrix}
\]

In these matrices the rightmost subscript denotes the power time-sequence set and the rightmost superscript refers to the time resolution used, in minutes. For the variances, the leftmost subscript indicates the load for which the variance was calculated. Since the effects of variability are the ones being analyzed, for this estimation the values of \( l^{u} \) will be referred to the losses for daily resolution on that imbalance scenario, not on the balanced case as was done before. After the estimation is performed, the values will be referred back to the balanced case so that they can be compared with the simulations done with real data.

The vectors and matrices defined above will be used to compute the coefficients in (15). This computation will be made using an ordinary least squares (OLS) linear regression model. The cost function to be minimized is defined as:
\[
J_{l}(u) = ||l^{u} - X_{u} \alpha^{u}||^{2}
\]
(16)
in which \( \alpha^{u} \) is the vertical concatenation of \( \alpha_{0}, \alpha_{\tau}^{u}, \) and \( \alpha_{cov}^{u} \), and \( X_{u} \) is the horizontal concatenation of a vector with ones, \( \Sigma^{2} \) and \( \text{Cov}_{u}^{\phi-\phi} \). This minimization is performed per grid and per phase imbalance scenario, as variances and phase-to-phase covariances are weighted according to grid layout and phase connection.

With the coefficients computed, it is possible to estimate relative losses using only information pertaining to load variance and phase-to-phase covariance.

IV. MODELING VARIANCE AND COVARIANCE INCREASE

Since the data used in the simulations in Section II is limited in time-resolution at 15 minutes, there is a need to predict the variances of loads and the phase-to-phase covariances in this data set for higher time-resolutions.

A. Predicting Variance

Using the data collected in [12], one can attempt to predict the behaviour of loads’ variances, when time-resolution is changed. A proposed solution here is to develop a linear regression model to predict variances for 5 and 1 minute resolutions, sampling data from [12], computing the variance
for each of the time resolutions (1, 5, 15, 30 and 60 minute) and fitting the increase in variance caused by the higher resolutions to the previous increases in lower resolutions. With this attempt it is assumed that load variance maintains its trend when its measuring switches from low to high time-resolution (i.e., a load that sharply increases its variability from 60 to 30 and from 30 to 15 minutes will also sharply increase from 15 to 5 and from 5 to 1 minutes, or a load that has steadied its variance will maintain itself steady). This is translated into:

\[
\Delta \sigma^2_{15 \rightarrow 5} = \beta_1 \Delta \sigma^2_{1440 \rightarrow 60} + \beta_2 \Delta \sigma^2_{60 \rightarrow 30} + \beta_3 \Delta \sigma^2_{30 \rightarrow 15} \tag{17}
\]

\[
\Delta \sigma^2_{5 \rightarrow 1} = \gamma_1 \Delta \sigma^2_{1440 \rightarrow 60} + \gamma_2 \Delta \sigma^2_{60 \rightarrow 30} + \gamma_3 \Delta \sigma^2_{30 \rightarrow 15} + \gamma_4 \Delta \sigma^2_{15 \rightarrow 5} \tag{18}
\]

in which \( \beta \)'s and \( \gamma \)'s are coefficients used to relate increases in variance (\( \Delta \sigma^2 \)) from low to high time-resolutions. The subscripts indicate for which time-resolution transition the increase is made.

After sampling 10,000 days' worth of data, the coefficients are computed by an OLS, linear regression. The cost functions to minimize are:

\[
J_5(\beta) = ||\Delta \sigma^2_{15 \rightarrow 5} - \Delta \Sigma^2 \beta||^2_2 \tag{19}
\]

\[
J_1(\gamma) = ||\Delta \sigma^2_{5 \rightarrow 1} - \Delta \Sigma^2' \gamma||^2_2 \tag{20}
\]

where \( \Delta \Sigma^2 \) is the variance transition matrix, whose columns contain the various time-resolution transitions and rows contain each observation. In \( \Delta \Sigma^2' \) a column for \( \Delta \sigma^2_{15 \rightarrow 5} \) is added. This process resulted in:

\[
\beta = \begin{bmatrix} 0.04759 \\ 0.5313 \\ 0.7543 \end{bmatrix}, \quad \gamma = \begin{bmatrix} 0.06590 \\ 0.02777 \\ 0.1716 \\ 0.4803 \end{bmatrix}.
\]

One can observe that an increase in variance for a given resolution change is more heavily impacted by the previous increase than by those occurring at lower time resolutions.

Finally, an estimate for the variances of the loads used in the simulations of Section II, if they had been measured with a higher resolution, is made by computing:

\[
\sigma^2_{i,j}^{25} = \sigma^2_{i,j}^{15} + \beta_1 \Delta \sigma^2_{i,j}^{1440 \rightarrow 60} + \beta_2 \Delta \sigma^2_{i,j}^{60 \rightarrow 30} + \beta_3 \Delta \sigma^2_{i,j}^{30 \rightarrow 15} \tag{21}
\]

\[
\sigma^2_{i,j}^{21} = \sigma^2_{i,j}^{15} + \gamma_1 \Delta \sigma^2_{i,j}^{1440 \rightarrow 60} + \gamma_2 \Delta \sigma^2_{i,j}^{60 \rightarrow 30} + \gamma_3 \Delta \sigma^2_{i,j}^{30 \rightarrow 15} + \gamma_4 \Delta \sigma^2_{i,j}^{15 \rightarrow 5} \tag{22}
\]

with \( i \) being the load and \( j \) being the time-sequence set.

### B. Predicting Covariance

The phase-to-phase covariances can be estimated using the coefficients obtained above by considering the properties of variance, in particular the sum of the two variates:

\[
\sigma^2[X + Y] = \sigma^2[X] + 2 \text{cov}[X, Y] + \sigma^2[Y], \tag{23}
\]

in which \( X \) and \( Y \) are random variates. Applying this to the problem at hand results in:

\[
\text{cov}^\phi_{\tau,u} = \frac{1}{2} \left( \sigma^2[P^\phi_{\tau,u} + P^\phi_{\tau,u}] - \sigma^2[P^\phi_{\tau,u}] - \sigma^2[P^\phi_{\tau,u}] \right) \tag{24}
\]

where \( P^\phi_{\tau,u} \) is the total power flowing in phase \( \phi \) and imbalance scenario \( u \), using time resolution \( \tau \). If the variances of the total per-phase demands for low time resolutions are recorded, they can be predicted for high resolutions, using the previously described method for predicting variances of individual loads. After predicting these variances, an estimate for the phase-to-phase covariances is obtained by computing (24) for each load set and imbalance scenario, using the new predicted variances.

### V. Estimating Losses for High Meter Resolutions

With the results of Section IV it is possible to build new \( \Sigma^2 \) and \( \text{Cov}^\phi_{\tau,u} \) matrices, using the 5 and 1 minute predictions. Using these matrices, combined with the \( \alpha \)'s computed with the method described in Section III, the relative losses for higher time resolutions are estimated. In Figure 6, an extension of the plots in Figure 4 is presented, in which the effects of the estimation on the relationship between relative losses and average relative variance can be observed.
Fig. 7. Boxplots and mean curves for the relative losses obtained in 30 power flow simulations and subsequent estimated extensions, for imbalance scenario 2 (above) and 3 (below). Mean curves were computed in the same manner as in Figure 2.

In a similar manner to what was made in Figure 2 for the results of the simulations, the relative losses w.r.t. relative frequency are plotted, this time extending the frequency range to 1 minute and observing the estimated increase in losses caused by an increase in time resolution. The extensions to Figure 2 are displayed in Figure 7.

It can be observed that an increase in resolution from 15 to 5 minutes results in an average increase in relative losses of 3.52% and 6.11% in scenarios 2 and 3, respectively. Extending the resolution from 15 to 1 minute gives us a 6.18% and 10.68% increase in relative losses for the two imbalance scenarios. Comparing with the reference case (balanced phases and flat load profiles) it can be seen that for 5 minute resolution the losses would be 1.372 and 2.267 times higher in scenarios 2 and 3, while for 1 minute they would be 1.407 and 2.364 times higher.

This estimation process was also performed for the remainder of the reference grids. For the less populated grids, in particular the 50kVA one, the model predicted in some cases a decrease in losses. This can be attributed to the higher susceptibility the losses on these grids have on local unbalance, which this model did not portray. However, the average trend was not affected by these errors. In Tables V and VI a summary of the results of the estimation made for losses in the other reference grids using 5 and 1 minute resolutions is shown. For the values marked with *, negative values were not taken into account.

### Table V

<table>
<thead>
<tr>
<th>Grids [kVA]</th>
<th>Some Imbalanced [%]</th>
<th>All Imbalanced [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.73* 13.03 61.21</td>
<td>0.90* 12.74 52.88</td>
</tr>
<tr>
<td>100</td>
<td>2.44* 6.01 14.22</td>
<td>0.93* 6.54 15.94</td>
</tr>
<tr>
<td>250</td>
<td>1.81 3.56 5.38</td>
<td>3.00 5.11 10.37</td>
</tr>
<tr>
<td>400</td>
<td>2.16 3.52 5.75</td>
<td>3.31 6.11 12.51</td>
</tr>
<tr>
<td>630</td>
<td>1.78 2.57 4.21</td>
<td>2.68 3.90 7.68</td>
</tr>
</tbody>
</table>

### Table VI

<table>
<thead>
<tr>
<th>Grids [kVA]</th>
<th>Some Imbalanced [%]</th>
<th>All Imbalanced [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.88* 4.24 15.90</td>
<td>0.02* 5.85 16.77</td>
</tr>
<tr>
<td>100</td>
<td>0.70 4.55 10.63</td>
<td>0.92 5.02 10.90</td>
</tr>
<tr>
<td>250</td>
<td>1.12 2.54 3.75</td>
<td>2.27 3.95 7.39</td>
</tr>
<tr>
<td>400</td>
<td>1.66 2.66 4.17</td>
<td>2.25 4.57 8.51</td>
</tr>
<tr>
<td>630</td>
<td>1.34 1.95 3.13</td>
<td>2.09 3.00 5.70</td>
</tr>
</tbody>
</table>

Using the results obtained in previous estimations of losses in the LV system under EDP-D’s tutelage [13] and the estimated increase in losses for higher meter resolutions, a new estimation of annual losses for the entire LV system can be made. To achieve this, a new and more realistic imbalance scenario is defined (scenario 4), in which three-phase loads have a non-zero relative unbalance value (i.e., are not balanced). This value is obtained by computing the average instantaneous relative unbalance of a set of real per-phase load sample, which results in \( \hat{u}^{\phi} \approx 0.1987 \). With this value, and knowing that single-phase loads have \( u = 1 \) and balanced three-phase loads have \( u = 0 \), the average relative unbalances of the loads in each grid in scenarios 2, 3 and 4 are obtained. Since relative losses were shown to be linear w.r.t. relative unbalance [1], they can then be computed, for this new scenario, by interpolating between the losses obtained for scenarios 2 and 3.

After obtaining the relative losses for scenario 4, the results of the previous study can be adjusted. These results were obtained considering balanced three-phase loads (i.e., scenario 2) and using load profiles with 15 minute resolution, and therefore had to be corrected to account for three-phase load imbalance and a 1 minute resolution:

\[
\Delta l^u = \frac{\hat{l}^{u,\tau}(\text{real}, 15)}{\hat{l}^{u,\tau}(\text{some}, 15)} \Delta l^\tau = \frac{\hat{l}^{u,\tau}(\text{real}, 1)}{\hat{l}^{u,\tau}(\text{real}, 15)}. \tag{25}
\]

In Table VII, the values for the conductor losses gathered for each reference grid in the previous study are shown, together
with the increases caused by imbalance ($\Delta I_d^*$) and resolution ($\Delta I_k^*$), and the final results after applying said increases. Transformer losses and conductor losses were separated, as the simulations and estimations made throughout this process were concerning conductor losses only. Since the load flowing through the transformer is aggregated load, its copper losses will be relatively invariant to the loads’ time resolution [6], not to mention the iron losses, which are invariant to load entirely. As such, these losses, taken from the same report as before [13], can be considered constant and are added to each reference grid’s loss value. Thus, the new annual losses observed in each reference grid are obtained. Knowing that each reference grid represents a certain number of real grids of the LV system, the new grid losses can be multiplied their corresponding number of grids represented and then added to each other to obtain an estimate for the total LV losses of the system. The final result is 13.38% higher than the one that resulted from the previous study (1043.6 GWh), which is a significant difference and one that reveals that the impact of load variability and phase unbalance is observable not only on the local level or for isolated loads, but on the scale an entire LV system.

### VII. Conclusion

In this paper, past work on the effects of phase imbalance and load variability was discussed and their results were validated through power flow simulations on a typical LV grid, using real consumption data. The outcome of these simulations confirmed that load imbalance and variability had a combined positive effect (in the mathematical sense) on losses and that the slope of the relative losses curves w.r.t. relative frequency warranted further investigation into losses using higher time resolutions of meters.

An attempt to estimate the effect of increasing meter resolution is proposed and its results are demonstrated. The model on which the estimation was made is based on the observed behaviour of loads on a simple linear feeder. To compensate the lack of high resolution data, a variance prediction model was developed. These models combined provided with the predictions for losses that showed that using 15 minute data lead to an underestimation of technical losses, particularly when the grid was highly unbalanced. The use of 5 minute resolution yielded more satisfactory estimates.

The work done here serves as a stepping stone towards a better understanding of LV grid losses, w.r.t. imbalance and variability of loads. Data collection on load imbalance and high resolution load profiles must be performed in order to develop more accurate prediction models, as the lack of representative data impairs the accuracy of any estimation made.

### APPENDIX A

**Proof that (8) can be generalized as (9)**

We can start by noticing that the numerator of (8) can be written as:

$$\frac{\tau}{T} \sum_{P} L(P) = E[L(P)].$$  \hspace{1cm} (A1)

Combining (A1) with the denominator of (8) results in:

$$I^P = \frac{E[L(P)]}{L(E[P])}.$$ \hspace{1cm} (A2)

Knowing that the expected value of the sum is equal to the sum of the expected value, the numerator of (A1) can be divided into the expected value of losses in each section of each phase and of the neutral. The denominator will be the sum of losses in those sections using the expected of the power flowing in said section.

Before expanding (A2), the neutral current must be defined based on power demands alone. Defining the square of the RMS of the neutral current, $n n^*$, as $[P^A, P^B, P^C, a^2]^T$, where $a = e^{-j \frac{2\pi}{3}}$ and * is the conjugate operation, it can be written that:

$$n n^* = P^A + P^B + P^C + a^2 P^A P^B + a P^A P^B + a^2 P^B P^C + a P^B P^C + a^2 P^C P^A + a P^C P^A.$$  \hspace{1cm} (A3)

Assuming that phases are symmetrical, each phase, even if imbalanced, will retain a 2/3π angle separation and thus:

$$a^2 P^A P^B + a P^A P^B = -P^A P^B$$

$$a^2 P^B P^C + a P^B P^C = -P^B P^C$$

$$a^2 P^C P^A + a P^C P^A = -P^C P^A$$ \hspace{1cm} (A4)

### TABLE VII

**Previous Results of Losses for Each Reference Grid, Each Average Increase and the Resulting New Grid Losses**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>3.47</td>
<td>2.25</td>
<td>2.25</td>
<td>14.69</td>
<td>1.45</td>
<td>5.685</td>
</tr>
<tr>
<td>100</td>
<td>12.00</td>
<td>10.84</td>
<td>10.84</td>
<td>14.69</td>
<td>3.33</td>
<td>18.02</td>
</tr>
<tr>
<td>250</td>
<td>9.90</td>
<td>6.896</td>
<td>6.896</td>
<td>11.60</td>
<td>5.04</td>
<td>16.64</td>
</tr>
<tr>
<td>400</td>
<td>13.63</td>
<td>7.464</td>
<td>7.464</td>
<td>16.43</td>
<td>6.93</td>
<td>23.36</td>
</tr>
<tr>
<td>630</td>
<td>11.51</td>
<td>5.185</td>
<td>5.185</td>
<td>13.47</td>
<td>9.15</td>
<td>22.62</td>
</tr>
</tbody>
</table>

### TABLE VIII

**Total Losses Estimated for the LV System**

<table>
<thead>
<tr>
<th>Reference Grids [kVA]</th>
<th>New Grid Losses [MWh/year]</th>
<th>Number of Grids</th>
<th>Total LV System Losses [GWh/year]</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>5.685</td>
<td>9 008</td>
<td>51.21</td>
</tr>
<tr>
<td>100</td>
<td>18.02</td>
<td>21 311</td>
<td>384.0</td>
</tr>
<tr>
<td>250</td>
<td>16.64</td>
<td>11 184</td>
<td>186.1</td>
</tr>
<tr>
<td>400</td>
<td>23.36</td>
<td>9 798</td>
<td>228.9</td>
</tr>
<tr>
<td>630</td>
<td>22.62</td>
<td>14 721</td>
<td>333.0</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>66 022</td>
<td>1183.2</td>
</tr>
</tbody>
</table>
can be affirmed. So, the RMS neutral current squared can be defined as:

\[ n^* = \frac{a^2[P_A^2 + P_B^2 + P_C^2 - P_A^2 P_B - P_B^2 P_C - P_C^2 P_A]}{2} \]  

(A5)

With this information, (A2) can be expanded by looking at each section of each phase and the neutral conductor, and, for the denominator, add the expected value of the RMS current equivalent squared flowing through that section, and, for the numerator, the square of the expected values are added. In the end, for the example given, the result is:

\[ l^P(P) = 1 + \frac{2 \sum_{k=1}^{2} \sum_{n,m=k}^{2} \text{cov}[P_{\phi_n}, P_{\phi_m}]}{2 \sum_{k=1}^{2} \sum_{n,m=k}^{2} E[P_{\phi_n}^\phi P_{\phi_m}^\phi]} - \sum_{k=1}^{2} \sum_{n,m=k}^{2} \text{cov}[P_{\phi_n}^\phi, P_{\phi_m}^\phi] \]  

(A8)

Repeating this process with more loads, including single phase loads, results in a similarly shaped formula for losses, which can be generalized into equation (9).

REFERENCES


