Velasco, J.A, Amaris, H., Alonso, M. y Miguelez, M. (2018). Stochastic Technical Losses Analysis of Smart Grids under Uncertain Demand. In *53rd International Universities Power Engineering Conference (UPEC).*

DOI: https://doi.org/10.1109/UPEC.2018.8542099

Stochastic Technical Losses Analysis of Smart Grids under Uncertain Demand

J-A. Velasco
Electrical Engineering Department Universidad Carlos III de Madrid
Madrid, Spain
jovelasc@pa.uc3m.es

H.Amaris
Electrical Engineering Department Universidad Carlos III de Madrid
Madrid, Spain
hamaris@ing.uc3m.es

M. Alonso
Electrical Engineering Department Universidad Carlos III de Madrid
Madrid, Spain
moalonso@ing.uc3m.es

M. Miguelez
Gas Natural Fenosa Madrid, Spain
mmiguelez@gasnatural.com

***Abstract*—Technical losses of smart grids can be computed using the customer's smart meter measurements (active and reactive energy) and the energy measurement registered by the Low Voltage (LV) supervisor deployed at secondary substations. However, in some LV networks, some customers do not provide information regarding the energy consumed and produced in real time. This fact complicates the calculation of technical losses because this information is necessary for estimating the load demand for this subset of customers. In this paper, a stochastic approach is proposed for the estimation of technical losses in smart grids under uncertain load demands (e.g., non-telemetered customers and uncertain smart meters readings). Load demand estimation of non-metered customers was performed by means of a top-down approach. Intra-hour load demand profiles of customers were synthetically generated by applying a Markov process. The data and network used in this process corresponded to the roll-out deployed by the Spanish Research and Development (R&D) demonstration project OSIRIS.**

Keywords— Markov Process, Non-Linear Programming, Kernel Density Estimation, Advanced Metering Infrastructure

NOTATION

$\Omega_c{}'$	Telemetered Customers
$\Omega_c{}''$	Non-telemetered Customers
Ω_c	All customers
Ω_d	Set of Hours of the period of time considered
$P_{(i,h,d)}$	Hourly load demand for customer i, hour h, day $d \in \Omega_d$
$\hat{P}_{(i,h,d)}$	Hourly load demand Measurement for customer i , hour h in day $d \in \Omega_d$
$\hat{E}_{(i,d)}$	Daily energy Measurement for customer i in day $d \in \Omega_d$
$\hat{P}^s_{(h,d)}$	Hourly load demand Measurement registered by the supervisor in hour h in the day $d \in \Omega_d$
$\hat{E}^s_{(d)}$	Daily energy Measurement registered by the supervisor in day $d \in \Omega_d$
$X_t = i$	Markov chain in time discrete t with state i
$p^{(t)}{}_{i,j}$	One-step probability transition from state i to state j in discrete time t
$P_{k,k}$	Transitional probability matrix of k states
$f(x)$	Probability Density Function (PDF) of the random variable x
$F(x)$	Cumulative Distribution Function (CDF) of the random variable x
$\hat{f}(\tau)$	Estimated density function of the variable τ
$\hat{h}$	Smoothing parameter/Bandwidth
$\varphi(\tau)$	Gaussian normal density function
$\hat{\sigma}$	Standard deviation of the measurements
L_p	Active Technical Losses in percentage

This work has been partly funded by the Spanish Ministry of Economy and Competitiveness through the National Program for Research Aimed at the Challenges of Society under the project OSIRIS (RTC-2014-1556-3).

I. INTRODUCTION

One of the European Union Targets is to replace at least 80% of all energy meters with smart meters by 2020 [1]. This large-scale roll-out of smart meters would not only allow service providers to gain information about the energy consumed and produced by each customer in real time but it would also allow them to compute network Technical Losses (TLs) at any given time. TLs of smart grids can be computed using customers' measurements (active and reactive energy) and the energy measurement registered by the Low Voltage (LV) supervisor deployed at secondary substations. The more accurate the load demand data are, the more precisely TLs can be calculated. However, in some LV networks, there are non-telemetered customers, who do not provide information about the energy they consume or produce in real time. This complicates TL calculations because this information is necessary for estimating the load demand of non-telemetered customers. Moreover, in some instances, smart meter measurements do not have the required accuracy, the data they provide can be anomalous (null or extremely high) or the device can even be out of service. These challenges make it necessary to estimate load demand in order to perform any real-time power analyses and, especially, TL calculations.

Different approaches have been applied to the technical losses estimation [2]-[6]. Nevertheless, all of them are hardly applicable to large distribution system areas, taking into account the uncertainty in topology and demand data. In this paper, a top-down approach for intra-hour load demand data estimation was developed, and a stochastic loss analysis was performed.

II. Hourly Demand Estimation for Customers Without Measurements

In LV smart grids there are telemetered customers for whom smart meters provide hourly measurements. There are also non-telemetered customers, who have monthly energy meters. Moreover, in some situations, hourly smart meter data from telemetered customers can go missing. Consequently, the hourly load demand from non-telemetered and telemetered customers with missing measurements has to be estimated. This procedure is based on a top-down approach to obtain hourly load demand estimation for customers lacking hourly measurements.

Non-telemetered customers have energy meters that provide measurements of the energy consumed during the most recent months. For each non-telemetered customer, historical monthly energy measurements are used to infer the hourly energy consumption by applying a top-down approach using three levels of resolution:

- Upper level: Estimation of the current monthly energy consumption
- Middle level: Daily energy estimation
- Lower level: Hourly energy estimation

A. Upper-level: Estimation of the current monthly energy consumption

Available data from non-telemetered customers consist of historical monthly energy values measured during the most recent months. This information for energy consumption during the current month is typically lacking because it is collected once the month ends. To estimate the current month's energy consumption, an Energy Consumption Tendency curve (ECT curve) is deduced using the historical monthly measurements. The monthly energy consumed from the previous month (last available measurement) to the current month (still unknown) is inferred through an interpolation process based on polynomial functions (splines) [7]. The ECT for the energy consumption of a non-telemetered customer is shown Fig. 1.

B. Middle-level: Daily energy estimation

Once the energy estimation for the current month is obtained, it is possible to estimate each week's consumption. This weekly estimation process is necessary in order to differentiate the energy consumption of the working days from holidays. The weekly profile is deduced from the weekly consumption registered by the smart meter located at the secondary substation. This secondary substation weekly profile is considered the reference Weekly Energy Consumption (WEC) profile for the LV network. In Fig. 2, the estimated WEC profile for a non-telemetered customer is shown. The first three points correspond to holydays, the next five points correspond to the work days and the last two points correspond to Saturday and Sunday. The WEC therefore provides the daily energy consumption for each day of the week.

C. Lower level: Hourly energy estimation

At this level, the complete set of hourly values that are missing is estimated using an optimization process. The procedure proposed is formulated as a Non-Linear (NLP) optimization problem as described below.

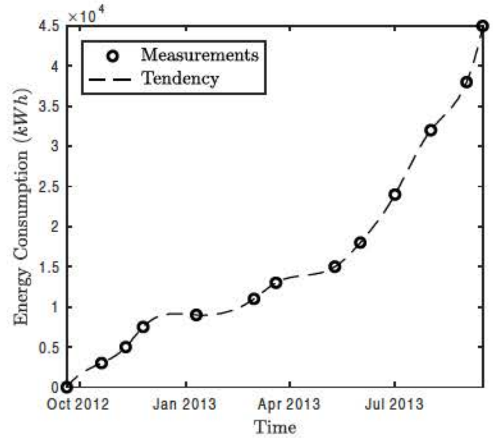

Fig. 1. ECT curve for the historical measurements of a non-telemetered customer from September 2012 to September 2013.

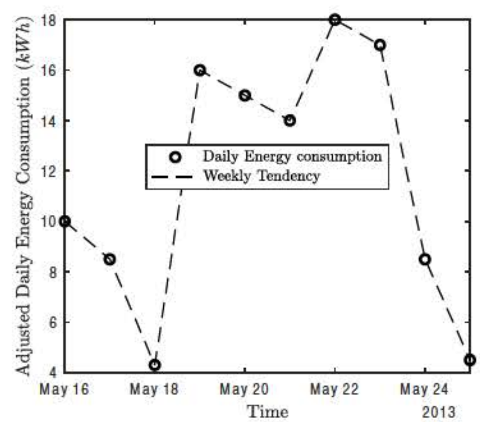

Fig. 2. WEC profile of a non-telemetered customer for a week in May 2013

Minimize:

$$\sum_{i\in\Omega_c'}\sum_{d\in\Omega_d}\sum_{h=1}^{24}\left(P_{(i,h,d)}-\hat{P}_{(i,h,d)}\right)^2 + \sum_{i\in\Omega_c''}\sum_{d\in\Omega_d}\left(E_{(i,d)}-\hat{E}_{(i,d)}\right)^2 + \sum_{d\in\Omega_d}\sum_{h=1}^{24}\left(\sum_{i=\Omega_c}P_{(i,h,d)}-\hat{P}^{s}_{(h,d)}\right)^2 \quad \text{(1a)}$$

Subject to:

$$\sum_{i\in\Omega_c}\sum_{h=1}^{24}P_{(i,h,d)} = \sum\nolimits_{i\in\Omega_c}E_{(i,d)}\ \forall\, d\in\Omega_d \quad \text{(1b)}$$

$$\sum_{i\in\Omega_c}\sum_{h=1}^{24}P_{(i,h,d)} \le \sum_{h=1}^{24}P_{(s,h,d)};\ \forall\, d\in\Omega_d \quad \text{(1c)}$$

$$\sum_{i\in\Omega_c}E_{(i,d)} \le \hat{E}^{s}_{(d)};\ \forall\, d\in\Omega_d \quad \text{(1d)}$$

The objective function (1a) is composed of three terms. The first term represents the total square error between the real hourly demand $P_{(i,h,d)}$ and the hourly measurement $\hat{P}_{(i,h,d)}$ for the 24 hours of the day *d*. The second term refers to the square error between the real daily energy consumption $E_{(i,d)}$ and the measured daily value $\hat{E}_{(i,d)}$. The third term in the equation refers to the square error related to the daily energy demanded from the complete set of network customers $\sum_{i=\Omega_c} P_{(i,h,d)}$ and the daily energy registered at the secondary substation $\hat{P}^{s}_{(h,d)}$.

Constraints (1b) equalize the 24-hour power consumption with the daily energy consumption by each customer. Constraints (1c) establish that the 24-hour power consumption registered at the secondary substation has to be greater than the power demanded by the network customers. Constraints (1d) establish the energy constraint for consumption throughout the day.

III. INTRA-HOUR DEMAND ESTIMATION

The estimation of intra-hour load demand for network customers is modelled as a stochastic Markov process [8] Therefore, high-resolution load demand customer profiles are modelled as a discrete time Markov chain X_t, where t is the parameter running over an index set T that corresponds to discrete units of time $T = (t_0, t_n)$ (where t_0 starts the hour and t_n finishes the hour), and the value that takes X_t is the state of the Markov process [8]. The Markov chains are characterized by the property (2), which defines the probability of shifting from the state $X_t = i$ to the state $X_{t+1} = j$ in t discrete units of time, which only depends on the previous state.

$$\Pr\{X_{t+1} = j | X_0 = i_0, \dots, X_{t-1} = i_{t-1}, X_t = \mathrm{i}\} \quad (2)$$
$$= \Pr\{X_{t+1} = j |\, X_t = \mathrm{i}\} = p^{(t)}{}_{\mathrm{i},j}$$

The probability described in (2) is also called a one-step transition probability. To model a complete Markov process of k states, a transitional probability matrix *P* (3) had to be built, which constitutes a first-order Markov chain. This takes into account the current state and that the immediately preceding in order to calculate the probability of the next state. Since equivalent discrete time unit times are considered (e.g., 1 minute), the Markov chain is homogenous so that the entry $P_{i,j}$ of the matrix *P* provides the probability of shifting from state *i* to j.

$$P = \begin{bmatrix} p_{1,1} & \cdots & p_{1,k} \\ \vdots & \ddots & \vdots \\ p_{k,1} & \cdots & p_{k,k} \end{bmatrix} \quad (3)$$

The i^{th} row is the probability distribution of the values of X_{t+1} under the condition $X_t = i$. The number of rows of matrix *P* determines the number of states in the Markov process. Note that the sum of all values in a single row (a group of states) is the equal to the unit [8].

A first-order Markov process is set up using the hourly load demand probability functions for the hourly measurement, and the data are estimated from the LV network customers. A transitional probability matrix $P_{60,60}$ (with 1-minute resolution) is used to model intra-day and high-resolution load demand profiles. The first state of the Markov chain sequence is chosen by means of a random variable that selects values in the interval (0,1) [9]. Next, the matrix $P_{60,60}$ is applied to determine the next states.

A. Load demand probability functions

To determinate the entries $p_{i,j}$ for the transitional probability matrix *P*, the Probability Density Functions (PDFs) and the Cumulative Distribution Functions (CDFs) have to be determined. PDF and CDF are denoted with $f(x)$ and $F(x)$, respectively, and their definitions are provided in (4) and (5) (where X denotes the Markov chain). This describes the relationship between them [10].

$$Pr[i \le X \le j] = \int_i^j f(x)dx = F(i) - F(j) \quad (4)$$

$$Pr[X \le j] = \int_{-\infty}^j f(x)dx = F(j) \quad (5)$$

Due to the difficulty associated with selecting a theoretical parametric distribution (normal, log-normal, etc.), the Kernel Density Estimation (KDE) [10] is chosen to carry out the fitting procedure. Recently, in the scientific literature, other authors have used KDE to fit PDFs and CDFs to the continuous operational variables of the network, such as load demand in [11] or wind speed in [12]-[13]. Therefore, to avoid making assumptions about the statistical distribution of the load demand data, the KDE method is applied. The non-parametric representation of the PDF of the load demand variable S_d is defined by means of a smoothing function (6) as follows:

$$\hat{f}(\tau) = \frac{1}{n\hat{h}} \int_1^{24} \varphi(\tau) \left(\frac{\tau - S_d(t)dt}{\hat{h}} \right) \quad (6)$$

Where $\hat{f}(\tau)$ is the estimated density function, τ is the time in hours, $\varphi(\tau)$ is the Gaussian kernel (a standard normal probability density function [10]) and $\hat{h}$ is the bandwidth (a positive smoothing parameter). One of the criteria used to obtain the value for $\hat{h}$ is the Mean Integrated Squared Error (MISE) (7).

$$\mathrm{MISE}(\hat{h}) = E\left[\int \left(\mathrm{y}(\tau) - \mathrm{y}'(\tau) \right)^2 d\tau \right] \quad (7)$$

Additionally, $y'(\tau)$ is the unknown real density function. Due to the unknown $y'(\tau)$, the MISE formula cannot be applied in a straightforward fashion. Therefore, the most popular solution for using an appropriate value for $\hat{h}$ is to apply the Silverman's rule of thumb [14] by means of (8).

$$\hat{h} = \left(\frac{4\,\hat{\sigma}^5}{3n} \right)^{1/5} \approx 1.06\hat{\sigma}n^{-1/5} \quad (8)$$

Where $\hat{\sigma}$ is the standard deviation of the variable S_d.

IV. Estimation of technical losses in a smart grid network: Case study

The data and network used in process correspond to the roll-out deployed for the Spanish Research and Development (R&D) demonstration project OSIRIS, which is an innovative project to develop knowledge, tools and new equipment for optimizing the supervision of LV Smart Grids [15]. The project scenario concerns a primary substation that supplies power to 25,849 customers in a region located in the south of Madrid (Spain). In Fig. 3 one of the LV distribution networks in the OSIRIS project is shown. The 680-kVA network has eight LV residential and commercial feeders with 32 load consumption points. There are 22 telemetered customers (individual contractual power ≤15 kW) and 10 non-telemetered customers
with a total contractual power of 442 kW.

A. Description of the smart grid communications architecture

In the OSIRIS demonstration project, the communication infrastructure between smart meters and the Meter Data Management System (MDMS) has been designed specifically for LV networks. It is based on PoweRline Intelligent Metering Evolution (PRIME) [16], (Fig. 4). Smart meters are associated with telemetered customers, whereas the concentrators are located at the secondary substations. Concentrators send their data up to the so-called gateway, which manages communications with the MDMS using the medium voltage infrastructure as a communication medium. A much more detailed explanation of the communication architecture deployed in the OSIRIS project can be found in [16].

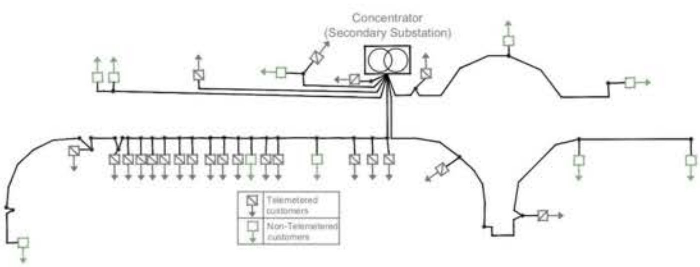

Fig. 3. OSIRIS network layout

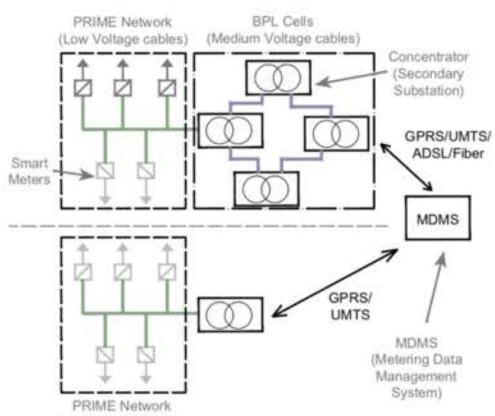

Fig. 4. Communication architecture

B. Data acquisition

Each telemetered customer has an individual smart meter that registers the active load energy (imported and exported), in addition to the reactive load demand (four quadrants). This provides hourly and daily measurements. A smart meter supervisor is located at the secondary substation. In this study, smart meter data from all the metered customers have been collected by the concentrator of one of the seven networks operated by the OSIRIS project from September 2013 to September 2014.

C. Statistical study of load demand data

A statistical study of the hourly demand consumption of the OSIRIS network is shown in Fig. 5. The histograms describe power demand across different daily time-slots. During the first hours of the day (01:00 and 06:00) and the last hours (20:00 and 22:00), load demand is concentrated around a certain value and it displays a clear skewness. Meanwhile, for the rest of the hours of the day (10:00 to 18:00), the load demand is spread in a more heterogeneous fashion.

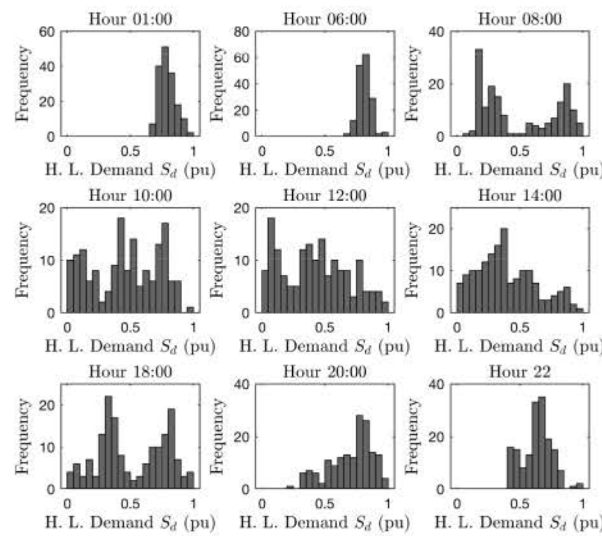

Fig. 5. Hourly Load (H. L.) demand histograms

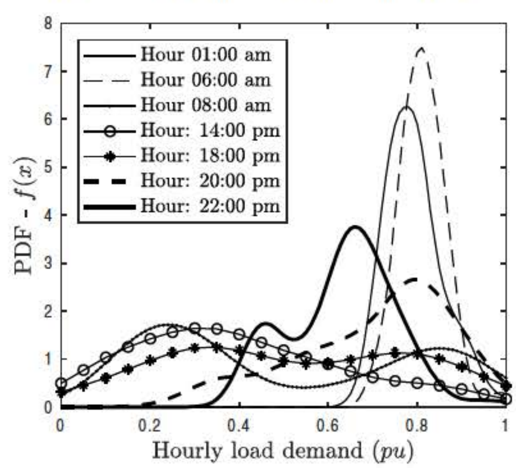

Fig. 6. PDFs obtained for hourly load demand.

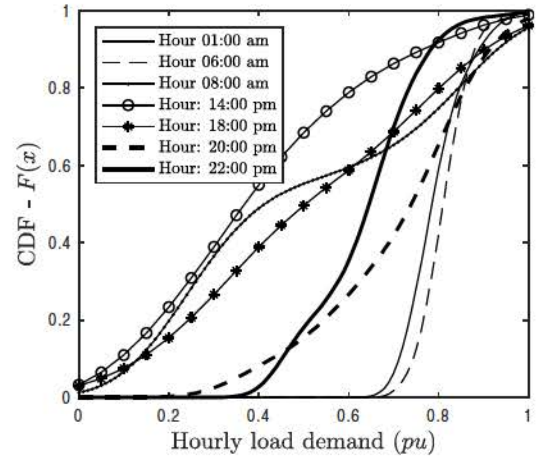

Fig. 7. CDFs obtained for hourly load demand.

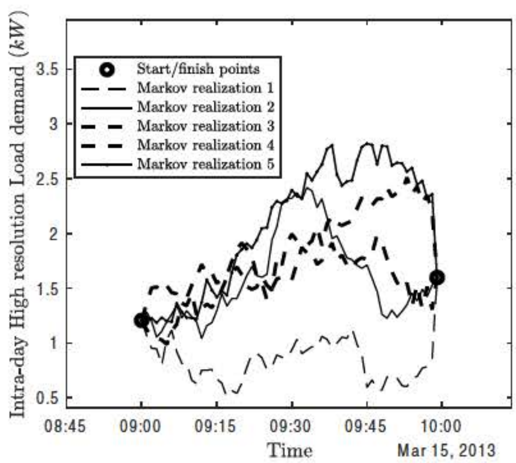

Fig. 8. Fifteen intra-hour high-resolution load demand realizations of the Markov process.

Fig. 6 and Fig. 7 show the PDF and CDF distributions obtained by the KDE method, respectively.

For each customer (telemetered and non-telemetered), a stochastic high-resolution load demand has been synthetically generated (Fig. 8). The initial point corresponds to 9:00 AM, and the final point corresponds to 10:00 AM. These points are known using the hourly load demand. Between these two hours, five Markov realizations are presented as an illustration of the Markov process carried out in this paper.

D. Stochastic TLs analysis

Hourly TLs have been calculated by three-phase load flow using the Back Forward Sweep (BFS) algorithm for every stochastic high-resolution load demand profile synthetically generated. TLs are expressed as a fraction of the load demand (relative TLs) in (9) where I_s is the current, r_s is the linear resistance (Ω/km) and l_s is length (m) of the segment of line s.

$$L_p = \frac{P_L}{S_d} = \sqrt{3}\frac{\sum_s I_s^2 r_s l_s}{V_n I_T} \tag{9}$$

Fig. 9 shows the boxplot representation of the stochastic TLs for each percentage level of loading capacity in the secondary substation.

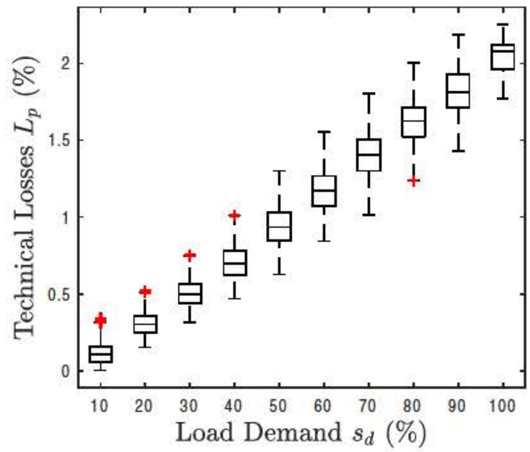

Fig. 9. Boxplot for stochastic TLs

The median values (middle line of the box) are distributed according to a linear evolution [10]. Thus, linear regression is proposed as a method to fit stochastic TLs by means of the ordinary least squares model (10), where $\overline{L_p}$ and $\overline{s_d}$ are the mean values of the relative TLs and load demand, respectively, $Cov[s_d, L_p]$ is the covariance of TLs L_p and load demand s_d; and $Var[s_d]$ is the variance of load demand s_d.

$$L'_p = \beta\, s_d + \alpha \tag{10a}$$

$$\beta = \frac{Cov[s_d, L_p]}{Var[s_d]} \tag{10b}$$

$$\alpha = \overline{L_p} - \beta\overline{s_d} \tag{10c}$$

Fig. 10 shows the linear regression calculated for relative TLs. Points of the relative TLs for every stochastic realization are concentrated in the middle of the cloud. Moreover, the bandwidth error represents the variability with respect to the regression model prediction. Finally, Fig. 11 shows the PDF and CDF functions through KDE for the stochastic TLs obtained by means of the Markov process. For this specific OSIRIS network, hourly TLs do not exceed 2% with a high density in the interval (0,1.5) %.

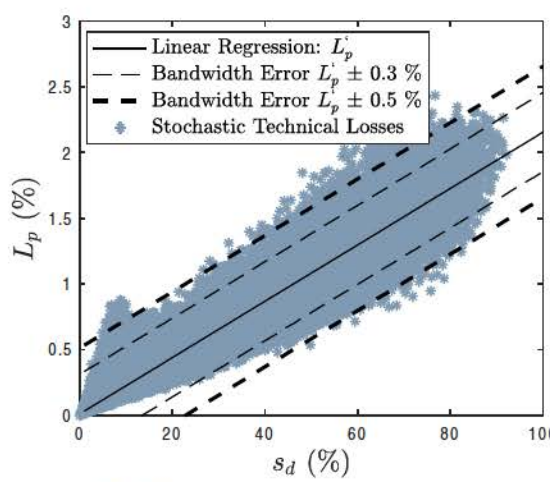

Fig. 10. Stochastic Technical Losses obtained

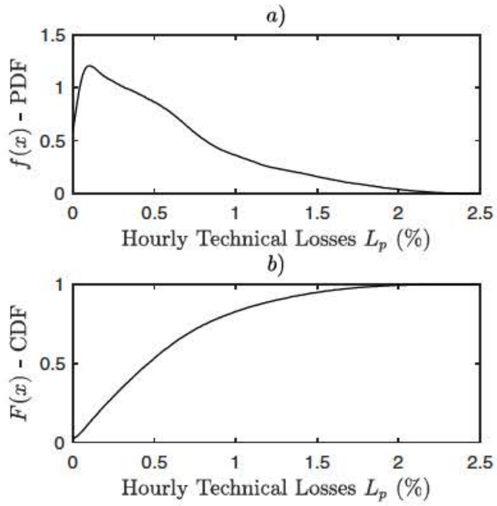

Fig. 11. PDF and CDF for hourly TLs obtained for the stochastic Markov process

V. CONCLUSIONS

In this paper, TLs of a smart grid demonstration project have been stochastically analysed. Load demand from the project network has been obtained through the communications infrastructure (AMI). The existence of non-telemetered customers as well as the missing load demands data have been taken into account in order to generate an estimate them by means of a top-down approach. Missing hourly load demand values have been estimated through an NLP optimization process. Additionally, intra-hour high-resolution load demand profiles have been synthetically generated by means of Markov chains. For every stochastic realization of the Markov process, a balanced three-phase load flow analysis has been carried out to obtain the network TLs. TL results have been statistically analysed using a linear regression fitting to calculate a maximum value for TLs in the network.

REFERENCES

[1] Directive 2009/72/EC of the European Parliament and of the Council of 13 July 2009 concerning common rules for the internal market in electricity and repealing Directive 2003/54/EC

[2] X. Fu, H. Chen, R. Cai, and P. Xuan, "Improved LSF method for loss estimation and its application in DG allocation," *IET Generation, Transmission and Distribution*, vol. 10, pp. 2512–2519, July 2016.

[3] K.A. Ibrahim, M. T. Au, C. K. Gan and H. J. Tang. "System wide MV distribution network technical losses estimation based on reference feeder and energy flow model," *International Journal of Electrical Power & Energy Systems* vol. 93, pp. 440-450, Dec. 2017.

[4] S. Wang, P. Dong, and Y. Tian, "A Novel Method of Statistical Line Loss Estimation for Distribution Feeders Based on Feeder Cluster and Modified XGBoost," *Energies*, vol. 10, no. 12, Dec. 2017.

[5] A. J. Urquhart and M. Thomson, "Impacts of Demand Data Time Resolution on Es- timates of Distribution System Energy Losses," *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1483-1491, May 2015.

[6] G. Poursharif, A. Brint, M. Black, and M. Marshall, "Using smart meters to estimate low-voltage losses," *IET Generation, Transmission & Distribution*, vol. 12, pp. 1206– 1212, March 2018.

[7] P. R. Massopust, "Interpolation and Approximation with Splines and Fractals," Oxford, Oxford University Press, 2010 .

[8] Howard M. Taylor and Samuel Karlin, "An introduction to Stochastic Modeling," London: Academic Press limited, 1984.

[9] McLoughlin, F., Duffy, A. and Conlon, M. "The Generation of Domestic Electricity Load Profiles through Markov Chain Modelling," *Euro-Asian Journal of Sustainable Energy Development Policy*, vol. 3, January-December 2010.

[10] D. C. Montgomery, G. C. Runger, "Applied Statistics and Probability for Engineers," New York. Wiley, 1999.

[11] Hyndman, R. J., and Fan, S. "Density forecasting for long-term peak electricity demand," *IEEE Transactions on Power Systems*, vol. 25, no. 2, pp. 1142-1153, May 2010.

[12] Xu, X., Yan, Z., and Xu, S. "Estimating wind speed probability distribution by diffusion-based kernel density method," *Electric Power Systems Research*, vol. 121, pp. 28-37, Apr. 2015.

[13] Miao, S., Xie, K., Yang, H., Karki, R., Tai, H. M., and Chen, T. "A mixture kernel density model for wind speed probability distribution estimation," *Energy Conversion and Management*, vol. 126, pp. 1066-1083, Oct. 2016.

[14] Silverman, B.W.,"Density Estimation for Statistics and Data Analysis," London: Chapman & Hall/CRC., 1986.

[15] OSIRIS project web site: http://www.proyecto-osiris.com/es.

[16] López, G., Moreno, J. I., Amarís, H., and Salazar, F. "Paving the road toward Smart Grids through large-scale advanced metering infrastructures," *Electric Power Systems Research*, vol. 120, pp. 194-205, March 2015.

.