CIRED 2020 Berlin Workshop (CIRED 2020)

22-23 September 2020

Theme 3: Flexibility Platforms and the Role of future DSO's

Models for characterising the final electricity demand

Journals
The Institution of Engineering and Technology

ISSN 2515-0855
doi: 10.1049/oap-cired.2021.0236
www.ietdl.org

Giuditta Pisano ✉, *Fabrizio Pilo, Matteo Troncia*

Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy
✉ *E-mail: giuditta.pisano@unica.it*

Abstract: Nowadays the consumption and generation profile estimation is of the greatest importance. New loads characterized by coincident peak of consumption (e.g., home charging of electric vehicles) or by high absorption peaks (heat pumps) are increasingly frequent. The presence of such loads must be carefully considered for network investments and for the optimization of asset management. Moreover, the massive diffusion of non-programmable renewable sources gives a leading role to the flexibility of demand, which is crucial for the success of the energy transition. The variety and difference of the electrical behaviour of LV customers, even nominally homogeneous, need stochastic methods for estimating the load profile on the LV/MV interfaces for the planning and the operation of distribution network, and for estimating the flexibility potential of demand. In this paper different techniques for modelling the demand composition are compared to evaluate the quality of the DSO models on real customers. In particular, the power peak of a given network section is calculated as key indicator for estimating the risk of overloading of lines and secondary substation transformers. Different methods of calculation have been applied on a dataset gathered with a recent measurement campaign in Italy by considering real LV distribution networks.

1 Introduction

The knowledge of the electrical load is essential for many applications and studies on the distribution electrical system.

The electrical load influences the market and, ultimately, the price of energy because its precise estimation impacts the bulk market-clearing price. The extent of system services also depends on estimates of electricity consumption and the need to continuously adjust production to demand. During the network operation goals as energy losses reduction, detection of the sources of non-technical losses, improvement of voltage profile, unbalances reduction, etc. can be pursued with studies whose accuracy depends on the quality of the models for the prediction of demand and production.

For these reasons, conventional and emerging stakeholders of the distribution systems, transmission and distribution systems operators (DSOs), regulators, market players such as Balance Responsible Parties, Aggregators, etc. are interested to accurately model the behaviour of customers with load profiles (LPs). LPs are the pattern of electricity load consumption of a customer or a group of customers over a given period that has been extensively used for many years. Generally, LPs are obtained from historical data or measurement campaigns, suitably elaborated for defining representative or typical consumption shapes. In the Literature, several techniques have been proposed for obtaining LPs. Each technique is specialised for the goal of the specific study or on the group of customers represented [1–3]. LPs are crucial for designing networks and stations. The quality of state estimation and system operation also depend on LPs since they model the behaviour of the end users if accurately built and regularly updated [3–5]. Unfortunately, the frequency of updating for LPs is often not adequate. Often, LPs refer to customers that no longer exist. Furthermore, the LPs are often based on measurement campaigns that involve a small number of customers not significant as a statistical sample of a defined ambit (e.g. a nation). Such LPs can capture only a portion of the end users and cannot represent modern consumers, particularly the new prosumers. Finally, LPs often associate to homogeneous groups of end-users one profile per season and/or the day (weekday, weekend, pre-weekend, holiday). These models consider daily curves relevant to large consumer groups often based on the economic activity [i.e. residential (RES), agricultural, industrial or tertiary] [6]. Unfortunately, the assumption of homogeneity is acceptable for all economic activities besides the RES consumers. The RES users exhibit a great variance of LPs that depends on several exogenous factors (i.e. household size, the number of persons living in a household, net income and employment status, level of education, etc.) [1, 5, 7]. Thus, models that associate the same day profile to a category of customers are inaccurate, particularly if the category is the RES one. Hiding single customers behind the average behaviour of a large and non-homogeneous group leads to significant errors in distribution studies. One typical example is represented by the time coincidence of the demand in RES neighbours that is considerably smaller than the one that can be achieved by superimposing the same LP several times.

Nowadays, new and more accurate measurements can be gathered from the advanced metering infrastructure (AMI) that uses second-generation smart meters. New LPs of both active and reactive power, more realistic than in the past, can be produced. These models should be capable to capture the different behaviour of customers by using further information, that cannot be necessarily too detailed or do not adversely affect the privacy rights of the customers by exploiting *ad hoc* techniques able to deal with a huge amount of data [2, 4, 5]. As an example, new LPs can be obtained for a better representation of the typical behaviour of specified classes of users by filtering field data with geographic information to find correlations with the climate conditions (the use of electricity is different in the northern, colder, areas than in the southern ones), with socio-demographic characteristics (e.g. income, education level, social status, etc.), etc. For producing an accurate forecast of the demand, the models for network planning generally use a probabilistic approach, by exploiting probabilistic load flow (PLF) calculation algorithms, Monte Carlo methods or analytical methods [7, 8]. The PLF input variables (load and generation) are normally represented by suitable probability density functions (PDFs). However, it is not straightforward representing each load and generator with a proper PDF. The dynamic nature of consumer behaviour is heavy time dependent and to ensure a good characterisation of load and generation correlations and time dependencies should be considered simultaneously [8]. Other studies devoted, for instance, to elaborate market balancing mechanism in competitive frameworks, or devising marketing strategies, exploit predictive analyses, as the multi-linear regression analysis in order to avoid the huge investment of putting half-hourly metering into every market customer and to calculate the profile coefficients of several customer classes [9]. In this case, the accurate profiling of the customer behaviour can be used as an effective tool for tariff rate formulation.

Bottom-up methods can be performed for (smart) home energy systems design. They aim at building LPs or PDFs of specific electrical devices or specific households, by using detailed information on devices' usage, and elaborate and combine these profiles for building profiles that aspire to be representative of other households and areas [1, 5, 8]. Obviously, the use of any probabilistic or predictive method, by seeking a compromise between computational burden and accuracy, produces results affected by a certain, sometimes measurable, level of uncertainty.

In this paper, a recent and updated LP set is used for the specific task of evaluating the power peak that flows across the medium voltage (MV)/low voltage (LV) transformers that supply many Italian secondary substations (SSs). The calculation is compared with the real measures of the consumptions of the users, derived from a recent yearly measurement campaign, and with an estimation of the same quantity calculated according to a unified procedure currently adopted by the biggest Italian DSO.

2 DSO' approach: the Italian case

A standardised procedure of load modelling has been used currently used by the biggest Italian DSO for planning purposes since 1980 [10–12]. The procedure aims at facilitating the statistical composition of the demand for sizing MV/LV transformers in SSs or to verify the exploitation of the transformer that supplies an existing network. The main interesting quantities to be taken into account are the maximum power (peak) and the annual energy delivered.

Given a network section r that supplies a number N of users u, each of them (the ith) characterised by a rated power P_{Mi}, the power P_r flowing through this section can be modelled with the relevant PDF. The relevant peak power P_{rp} is a value that can be exceeded with a low probability. Similarly, the power absorbed by the ith customer supplied by r has a low probability to exceed his rated power P_{Mi}.

By hypothesising that the power of the ith customer P_{ui} and the P_r can be represented by a Gaussian PDF and that the peak power of all coincident and, as a consequence, simultaneous with the peak of the network section r, if P_{Mi} and P_{rp} have the same probability to be exceeded, the peak power P_{Mi} and P_{rp} can be written as in (1) and (2), respectively.

$$P_{Mi} = \mu_i(P_{ui}) + \alpha \cdot \sigma_i(P_{ui}) \tag{1}$$

$$P_{\mathrm{rp}} = \mu_r(P_r) + \alpha \cdot \sigma_r(P_r) \tag{2}$$

where $\mu_i(P_{ui})$ and $\mu_r(P_r)$ are the mean values and $\sigma_i(P_{ui})$ and $\sigma_r(P_r)$ are the standard deviations of the Gaussian PDF P_{ui} and P_r, respectively; α is the probability (the risk) that P_{Mi} and P_{rp} could be exceeded. By considering the N users supplied by the network section r, P_{rp} can be expressed as function of the mean values and standard deviations of the N users, as mentioned below

$$P_{\mathrm{rp}} = \sum_{i=1}^{N} \mu_i(P_{ui}) + \alpha \cdot \sqrt{\sum_{i=1}^{N} \sigma_i^2(P_{ui})} \tag{3}$$

As N goes to infinite, the second term of (3) becomes negligible and (3) can be re-written as mentioned below

$$P_{\infty p} = \sum_{i=1}^{\infty} \mu_i(P_{ui}) \tag{4}$$

Disregarding the power losses, the energy E_r supplied by the network section r is the sum of the energies E_i ($i = 1 \ldots N$) delivered to the users. Such energies can be expressed as function of the relevant equivalent hours, H_r and h_i, calculated as ratio between the annual energy and the maximum power, as mentioned below

$$E_r = \sum_{i=1}^{N} E_i; \quad E_r = P_{\mathrm{rp}} \cdot H_r; \quad E_i = P_{Mi} \cdot h_i \tag{5}$$

By considering (4) and (5) the balance between the total energy delivered and the energy measured in an upstream section is defined as below:

$$P_{\infty p} \cdot H_\infty = \sum_{i=1}^{\infty} \mu_i(P_{ui}) \cdot H_\infty = \sum_{i=1}^{\infty} P_{Mi} \cdot h_i \tag{6}$$

Thus, by assuming $M_i = \mu_i(P_{ui})$ and because (6) has to be valid for any group of customers (7) can be written as

$$\mu_i(P_{ui}) \cdot H_\infty = P_{Mi} \cdot h_i; \quad M_i = \mu_i(P_{ui}) = \frac{h_i}{H_\infty} \cdot P_{Mi} \tag{7}$$

By substituting (7) in (1), it can be obtained the quantity L_i defined as.

$$L_i = \alpha \cdot \sigma_i(P_{ui}) = P_{Mi} \cdot \left(1 - \frac{h_i}{H_\infty}\right) \tag{8}$$

and, again, by considering (8) in (1), (1) and (4) we obtain

$$P_{Mi} = M_i + L_i \tag{9}$$

$$P_{\mathrm{rp}} = \sum_{i=1}^{N} M_i + \sqrt{\sum_{i=1}^{N} L_i^2} \tag{10}$$

Equation (10) represents the general formulation of the load composition for calculating the peak power of a network section r. Since the equivalent hours h_i of each single user are usually unknown, the standardised procedure provides an average value, derived by a specific study conducted on a given group of customers [12]. Obviously, M_i can be underestimated by using average values if the real equivalent hours of the users are greater than the estimated value.

In Table 1 the values of m_j and l_j coefficients for calculating M_j and L_j by the formula (11) are provided.

$$M_i = m_i \cdot P_{Mi}; \quad L_i = l_i \cdot P_{Mi} \tag{11}$$

Table 1 Coefficients assumed for the average customer

P_{Mi}, kW	m_i	l_i	P_{Mi}, kW	m_i	l_i
1.5	0.13	0.87	31 ÷ 40	0.22	0.78
3.0	0.15	0.85	41 ÷ 50	0.24	0.76
4.5	0.16	0.84	51 ÷ 75	0.25	0.75
6.0	0.17	0.83	76 ÷ 125	0.30	0.70
7 ÷ 20	0.18	0.82	126 ÷ 250	0.34	0.66
21 ÷ 30	0.20	0.80	—	—	—

All the above formulas are valid for modelling three-phase users. They become acceptable for balanced single-phase users provided that (10) is adjusted as

$$P_{\text{rp}} = \sum_{i=1}^{N} M_i + \sqrt{3 \cdot \sum_{i=1}^{N} L_i^2} \tag{12}$$

For a homogeneous group of n_j customer with rated power P_{Mj}, (12) can be rewritten as in (13).

$$\begin{aligned} P_{\text{rp}} &= n_j \cdot M_j + \sqrt{3 \cdot n_j \cdot L_i^2} \\ &= n_j \cdot m_j \cdot P_{Mj} + \sqrt{3 \cdot n_j \cdot l_i^2 \cdot P_{Mj}^2} \end{aligned} \tag{13}$$

By calculating (13) the P_{rp} of any network section r can be assessed starting from Table 1.

3 Load profiling

In this paper, a set of typical load profiles (TLPs), recently produced by the Authors through a bottom-up approach that identifies similarity in consumption patterns by exploiting clustering algorithms applied on a large dataset of time domain data, has been used. The details of the approach can be found in [4, 5]. The TLPs, patterns of 96 samples per day (i.e. every 15 min), are subdivided into customer category [i.e. domestic resident or not, commercial (COM), industrial and agricultural] and in 12 typical days (i.e. working days, Saturdays and Sundays and holidays). Each category of customers is represented by more than one TLPs, the number of which corresponds to the resulting number of clusters identified by the clustering algorithms. This permits to highlight the differences of customer behaviour even within the same consumer category.

In Fig. 1 the TLPs of COM and RES customers during the summer working days are shown as an example. The profile shapes are pretty different within the same category of customers. Both domestic and COM customers exhibit evident differences in terms of peak hour, valley depth, etc.

4 Case study and results

The database (DB) is constituted by 3.260 real LV networks supplied by SS that deliver electrical energy to over 55,000 end users.

The number of customers served, their rated powers and contract, their type of connection (three-phase or single phase), and the annual active power profiles with 15 min sample rate is known for each substation with reference to 2017. Unfortunately, other interesting data as the size of the MV/LV transformer, the phase of connection of the single-phase customers, and the topology of the network supplied by each SS are unknown. Furthermore, the DB is constituted predominantly by SS that supply less than ten customers and many of them supply only one customer (i.e. about 1400 SS). Since the unified procedure described in Section 2 is meaningless for network sections that supply only one customer (the estimated peak is exactly the rated power of that customer), and, furthermore, it is not very reliable also for SS that deliver energy to less than ten customers, the available DB has been reduced to the SS that deliver energy to more than ten customers (about 450 LV networks).

The following methods have been applied to the reduced DB:

M1. The calculation of the peak power of each SS is performed by using the measured profiles of the customers; the result of this calculation gives the ‘true’ or ‘real’ peaks, as they are assessed by starting with accurate data provided by the AMI.
M2. The peak power *Prp_est* of each network section (i.e. the SS) is estimated by adopting the standardised procedure described in Section 2. These estimated values are the ones currently used by the DSO for estimating the exploitation of existing transformers.
M3. For evaluating the effectiveness of the theory of the statistical composition of the demand, the estimation of the peak power has been repeated by hypothesising a gaussian PDF for the power P_r that flows through the network section and using the mean value and standard deviation calculated by starting from the real power profiles. The calculated values *Prp_cal* are the values that, according to the probability theory, have a scarce percentage to be exceeded (i.e. by assuming $\alpha = 3$ in (2) the 99.7% are smaller than *Prp_cal*).
M4. The peak power of each network section is performed by assigning to the customers their TLPs, derived from the load profiling method described in Section 3.

In Fig. 2 the peak powers calculated with the four methods listed above are shown. The SS are sorted in ascending order starting from the one that exhibits the smallest real peak. From the results, it clearly arises that the standardised procedure adopted by the DSO overestimates the power peaks in the majority of the cases, but the computational effort is small. The absolute errors vary from a minimum of 4% to a maximum of over 2300%. M2 underestimates the real peak only in one case with an error of about −4%. M3 performances seem much better than those of M2 and the errors, in this case, vary in a smaller range (i.e. 0% ÷ 100%). In M4 the forecasted peaks have a trend similar to M3 with an increased range of errors (i.e. 1% ÷ 186%), but only 12% of the SS have peaks underestimated. However, it is worth noticing that M4 does not require the knowledge of the real profiles but needs only few high-level information about the customers. In particular, M4 uses only the category/contract of the customers and the annual energy consumption of the customers served by the given SS and only one information about the

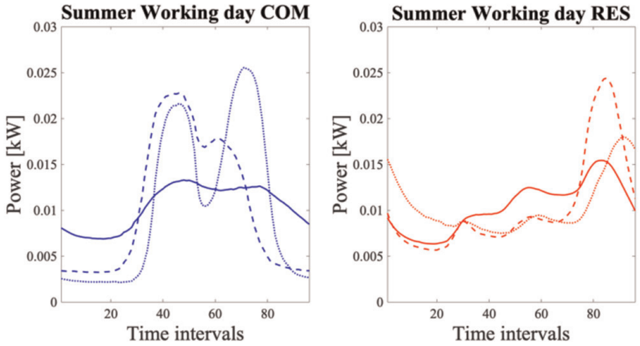

Fig. 1 *TLPs of the three resulting clusters of COM and RES customers during the summer working days*

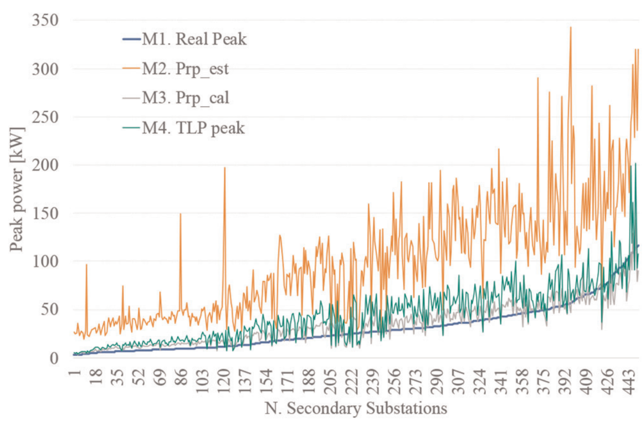

Fig. 2 *Peak powers of SS of the reduced DB calculated with the methods M1–M4*

location of the SS (i.e. its zone that corresponds to a very high-level subdivision of the Italian territory). On the contrary, for calculating mean values and standard deviations the M3 needs the knowledge of the whole customers load profiles, and, as consequence, with such deep information it could be possible assessing the real peaks.

5 Conclusion

In this paper, the comparison of different methods for composing the consumption profiles of groups of customers supplied by the same SS has been proposed. The evaluation of the power peak in MV/LV transformers has been performed by using the measured profiles obtained with a recent measurement campaign. Real values have been compared with the results of a profile composition method, currently used by the biggest Italian DSO. Furthermore, TLPs that aimed at accurately representing the diversity of the customer behaviours are used for the same goal. The results demonstrate that both the unified procedure and the use of TLPs need only a few information relevant to the customers. The methodology based on the use of TLP obtained with data analysis applied to real measurements outperforms traditional methodologies derived by the application of probabilistic laws on old data.

6 Acknowledgments

This work has been supported by the project ‘Planning and flexible operation of micro-grids with generation, storage and demand control as a support to sustainable and efficient electrical power systems: regulatory aspects, modelling and experimental validation’, funded by the Italian Ministry of Education, University and Research (MIUR) Progetti di Ricerca di Rilevante Interesse Nazionale (PRIN) Bando 2017 – grant no. 2017K4JZEE.

7 References

1 Grandjean, A., Adnot, J., Binet, G.: ‘A review and an analysis of the residential electric load curve models’, *Renew. Sust. Energy Rev.*, 2012, **16**, pp. 6539–6565
2 Varga, E.D., Beretka, S.F., Noce, C., *et al.*: ‘Robust real-time load profile encoding and classification framework for efficient power systems operation’, *IEEE Trans. Power Syst.*, 2015, **30**, (4), pp. 1897–1904
3 Quilumba, F.L., Wang, D.Y.: ‘Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities’, *IEEE Trans. Smart Grid*, 2015, **6**, (2), pp. 911–917
4 Pilo, F., Pisano, G., Troncia, M.: ‘Updated typical daily load profiles for LV distribution networks customers’. Proc. of SyNERGY MED 2019, Cagliari (I), 28–30 May 2019
5 Mocci, S., Pilo, F., Pisano, G., *et al.*: ‘Two-stage clustering for profiling residential customer demand’. IEEE (International Conference on Environment and Electrical Engineering (EEEIC) 2018), Palermo (I), 12–15 Jun 2018
6 Di Lembo, G., Petroni, P., Noce, C.: ‘Reduction of power losses and CO2 emissions: accurate network data to obtain good performances of DMS systems’. Proc. CIRED 2009, Prague, June 2009
7 Herman, R., Gaunt, C.T.: ‘A practical probabilistic design procedure for LV residential distribution systems’, *IEEE Trans. Power Deliv.*, 2008, **23**, (4), pp. 2247–2254
8 Machado, J.A.C., Carvalho, P.M.S., Ferreira, L.A.F.M.: ‘Building stochastic non-stationary daily load/generation profiles for distribution planning studies’, *IEEE Trans. Power Syst.*, 2018, **33**, (1), pp. 911–920
9 Elexon: ‘Load Profiles and their use in Electricity Settlement’, 2013. Available at: www.elexon.co.uk/reference/technical-operations/profiling
10 Unificazione Enel, DZ1230 La legge di composizione dei carichi per i calcoli di rete, November 1986
11 Unificazione Enel, DZ1200 Il carico, November 1986
12 Unificazione Enel, DZ1231 Procedure per l’applicazione della legge di composizione dei carichi per i calcoli di rete, November 1986