Forecasting Demand Flexibility of Aggregated Residential Load Using Smart Meter Data

Jelena Ponoćko, *Student Member*, *IEEE*, Jovica V. Milanović, *Fellow*, *IEEE*

Abstract— **Demand decomposition (disaggregation) presents the process of assessing time-varying participation of different load categories within the total active or reactive load. Information on load composition is highly beneficial for different demand side management applications. In order to decompose total forecasted load of aggregated households where only some are monitored by smart meters (SMs) with sub-metering capabilities, a two-level methodology is proposed. At the first level, load disaggregation of the load monitored by SMs is done based on measurements of power consumed by each home appliance, and at the second, the disaggregation of the total forecasted load is performed using artificial neural network. This paper investigates the required percentage of users in an aggregation which should be equipped with a SM with sub-metering capabilities in order to forecast (within certain confidence level) the load composition of the overall aggregated demand. The methodology was first tested on a UK statistics-based load model, and then validated on a real pilot site's consumption dataset. The results show that even with 5% SM coverage, one can forecast, with high confidence, the composition of the load at the substation (aggregation point). In other words, there is no techno-economic justification for sub-metering technologies to be installed at every user's premise; a limited installation of such devices would suffice.**

Index Terms-- **Demand-side management, demand response, distribution network, neural networks, smart meter**

I. INTRODUCTION

WITH the evolution of the smart grid and electricity markets, the responsibility of balancing between generation and demand is shifting towards the demand side. Demand response (DR) has been recognized as one of potentially cost-effective options for operating the power network [1]. The system operator in the UK, for example, is planning to obtain 30-50% of its balancing services through DR [2]. Since the power system is operated in real-time, the traditional load monitoring systems (electricity meters) have become obsolete for the demand side management (DSM) requirements. Increasing number of installed smart meters (SMs) in residential districts around the world will enable better observability of the end-users' behavior and their potential to participate in the network daily operation. Higher granularity of low-level consumption data will enable more

advanced profiling of consumers at different levels of aggregation. The widely deployed SMs, however, do not provide any sub-metering data, i.e. data about the use of individual appliances or categories of appliances which may be participating in DR programs. In other words, there is no time and space-based observability of the controllable loads in the residential and commercial load sectors.

The effectiveness of DR actions largely depends on the flexibility of the demand side, i.e. the size of its controllable loads. Until now, mostly large industrial users have been included in DR programs. The transmission system operator in the UK uses DR only from large industrial users and estimates it in advance, so that the balancing between generation and load can be done 2-60 minutes in advance of real time [3]. At the same time, there is a significant, though mainly untapped potential for DR in residential area. In the US, it was estimated that residential customers' participation in DR might bring up to half of the total peak reduction [4]. In the case of UK, residential sector is the largest final user of electrical energy, presenting around 30% of the overall consumption [5]. DR potential (flexibility) of an aggregated group of residential users, just as the flexibility of any other distributed energy resource, should therefore be quantified and aggregated in order to be appropriately exploited in network daily operation [6].

Forecasting the size of controllable load within the total load reduces the uncertainty of the actual (available) flexibility of the demand side, even before a DR signal is sent to the customers by their electricity supplier. Furthermore, flexibility profile of aggregated customers is more predictive than the profile of individual customers, which is generally highly random. The demand profiling can be performed in two dimensions: *i)* Time: By observing the changes in the size of controllable load within the total load over a day or season; *ii)* Space: By observing the size of controllable load over a power network; in this case, different distribution/transmission network buses will have different flexibility potential, depending on their load mix (namely residential, industrial or commercial), as well as socio-geographical factors.

Furthermore, load can be disaggregated into load categories, such as resistive loads, induction motors, lighting, etc., in order to obtain a more detailed insight into the types of load utilized on a daily or seasonal basis. In previous work [7], a methodology was developed for probabilistic load disaggregation based on the time series of total active and reactive load at a bulk point using artificial neural network (ANN) and statistical data about typical load composition in residential, commercial and industrial sectors. This work proved the validity of the approach showing reasonably good

accuracy in the estimation of load composition without any information on actual consumption by different types and categories of users. Both training and testing data were generated in probabilistic manner, using Monte Carlo simulations.

Similarly, demand decomposition based on non-intrusive load monitoring (NILM) methods has been analyzed in literature [8-12]. A common drawback of the load signature methods is that they require a library of high-resolution measurements of each home appliance's parameters needed to disaggregate the total load of an end-user, including current waveform, harmonics, switching transient waveform, etc., which are not always easily accessible, especially in case of a large number of users.

Disaggregation of load in terms of different categories of users, e.g. in [13] is commonly used for long-term demand forecasting and network planning purposes, but it does not provide information about demand flexibility.

The main aim of this paper is to extend the approach presented in [7] by incorporation of realistic consumption data based on the recorded end-users' behavior in the past, and such provide the information to e.g., the DSO or an aggregator about the time based aggregated demand flexibility, both in close to real time and as forecasted (e.g. day-ahead) demand flexibility, in order to support DR planning in network daily operation. It does not focus though on detecting the operation mode of individual appliances, but on examining the DR potential of aggregated demand, regardless of the means of shifting or curtailing this demand. An interesting approach to the centralized hierarchical control of aggregated thermostatically controlled loads was given in [14].

This paper introduces a load disaggregation methodology based on time series consumption data coming from SMs with sub-metering capabilities, regardless of the technology involved. The research is motivated by the existence of numerous, yet individual, trial sites performing load measurements on appliance level, mainly in residential sector [4, 9, 15, 16]. It assumes that only a number of consumers supplied from the same substation are monitored by this type of SMs, while the others have electricity meters with no sub-metering, and tries to answer the question of how many end-users need to regularly provide detailed information about their electricity consumption (per appliance or load category) in order to accurately decompose the forecasted electricity consumption of all the users in the aggregation. At the first level, the load flexibility assessment of the monitored users is performed based on the actual monitoring of the end-users' consumption of different appliances. The aggregated (both monitored and non-monitored) users' flexibility is then forecasted (e.g. day ahead) using total load forecast at the substation and ANN trained using the SM data (with sub-metering) recorded during the previous week – hence, it does not require large historical data. Furthermore, since the consumption data is aggregated at a data concentrator point (the substation), the size of the data sent to the upstream network, e.g. to the DSO or an aggregator, is also reduced.

The studies reported in the paper have been performed off-line, though the approach facilitates an on-line mode as well. In such case the monitored load data would be arriving at certain time-steps, and used for on-line prediction (day-ahead) of the demand composition. The methodology is applicable to any type of appliance-based monitoring, as long as active and reactive power of each appliance can be reported to the data concentrator at regular time intervals. (One minute is chosen here as the optimal granularity of data required for network operation purposes.)

The main contributions compared to the past work include: *i)* A methodology for time varying decomposition of forecasted real and reactive power demand, using sub-metering enabled SM data and pre-trained ANN, into six load categories and controllable/uncontrollable load in residential district. The individual and aggregated load curves (which in reality would be coming from SMs), are generated using a UK statistics-driven electricity consumption model, the so called CREST model, adopted from [15]; *ii)* This methodology, unlike [7], takes into account both, higher (1 min instead of 30 min) sampling rate of data (based on requirements of modern DR programs [3]), and analyses the effect of missing data and data streams with different sampling step on the accuracy of aggregated data; *iii)* The probabilistic derivation of reactive power data for an aggregation of users based on available real power measurements by SMs, as in many cases SMs measure only active load; *iv)* The assessment of the required coverage (in %) of a residential area with smart metering system with sub-metering capabilities in order to obtain the desired confidence level of the forecasted load decomposition.

II. REPORTED APPLICATIONS OF DEMAND DECOMPOSITION

The flexibility of the demand side varies in time, therefore the prediction/forecasting (e.g. minutes or day-ahead) of load composition can facilitate timely DR actions by showing whether the DR potential will be sufficient for different needs of the DSO. A distribution network, under appropriate regulatory framework, could provide ancillary services to the transmission network by using flexibility of the demand, distributed generators (DGs) and battery storage. In that respect, remotely controllable loads could participate in ancillary services market, and uncontrollable loads (those that cannot be controlled from a control center, but by customers only), storage units and DGs could participate in energy market [17].

Authors in [18] observed controllable load as a DER. They presented an energy management system in which the aggregator provides information about the forecasted available flexibility to the DSO, who then makes the optimal load schedule, also taking into account the needs of the upper grid. In another example [19], the individual appliance scheduling was done in two steps: first day-ahead, following the hourly price forecast and expected users' behavior; the second optimization step was performed in real time (in minutes), following the actual prices and users' comfort. In this respect, prediction of load composition of aggregated demand could facilitate more efficient scheduling of aggregated load, and such increase the effect of, to some extent controllable, changes in users' behavior on the network daily operation.

There are several types of DR that could benefit from demand decomposition. For example, [20] proved that, in case of a conservation voltage reduction program (voltage-based DR), the demand regulation potential depends on the initial size of load, as well as on load composition, namely the participation of load that can be modelled as constant impedance. Similarly, methodology for assessment of voltage-based DR potential in UK, described in [21], raised the need for information on the load composition at distribution system buses.

Price-based and incentive-based DR programs have been extensively analyzed in the literature, e.g. in [22-24]. Some authors [25], suggested that incentive-based programs could bring higher value to the reduction of peak demand. The approach presented in [26] ensures that the end-users receive a reward for their load shifting participation, as well as for the voltage improvement in the supply feeder. Following these, prediction of available demand flexibility would allow adjustment of the level of incentives or tariffs needed to increase involvement of the users in DR actions. In direct load control programs, described in [27], load composition could enable the system operator to have a more accurate overview of the size of available controllable loads whose consumption can be remotely shifted.

Most of the DR programs focus on reducing the total cost of power generation or curtailing the daily peaks in demand by shifting certain amount of total load from one time period (peak time) to another (off-peak time). On the other hand, shifting of load will change load composition at both affected hours, and thereby potentially the dynamic response of aggregated demand in case of a small or large network disturbance (e.g. voltage step change due to transformer tap changes, system faults, etc.), which might affect the angular and voltage stability of the power system [7], especially at transmission system levels. Information about the load composition can be very useful in these cases. It has been long recognized that the major sources of load dynamics are induction motors, thermostatic loads and energy efficient devices [28], which is why their effect on the aggregated load dynamic response should be further analyzed. Knowing the effect of different load compositions on the overall system dynamic response, and ultimately system stability, one may take measures, i.e., appropriate DR actions, to prevent those compositions of load that might lead to unfavorable dynamic responses and provoke instabilities in the power system.

III. METHODOLOGY

The output of the demand decomposition process is given as time-varying load shares (in per unit or percentage) of different load categories within the (time-varying) total active or reactive load, as it will be illustrated in Section V. Following methodology discussed in [29], load categories in this paper are defined as groups of appliances with similar voltage-dependent steady-state and dynamic load characteristics. Furthermore, load categories are divided into controllable and uncontrollable based on their potential to be shifted in time. The controllability of some loads is disputable, as in case of lighting loads - although, generally, they are considered to be uncontrollable, some of them can be dimmable and therefore controllable. Thus, the given

classification should be taken as illustrative only, as it could vary to a certain extent for different applications. This study considers as controllable all the appliances that may be a part of direct load control (e.g. fridges, water heaters) or incentive-based DR programs (e.g. tumble dryer), i.e., appliances that can be controlled/shifted automatically or by the users.

According to the most commonly used appliances in residential sector in the UK, six categories are recognized in this methodology and presented in Table I as: single-phase constant torque induction motors (CTIM1), single-phase quadratic torque induction motors (QTIM1), controllable resistive loads (RC), uncontrollable resistive loads (RUC), switch-mode power supply (SMPS) loads and Lighting. The full list of appliances, apart from heating, ventilation and air conditioning (HVAC) units, was adopted from the residential load model given in [15]. The same model was used in the case studies to generate individual daily load profiles of the end users, which served as a realistic representation of data streams coming from the SMs. Controllable loads mainly consist of thermostatically controlled loads (space and water heaters, fridges, freezers), which are the most common candidates for DR programs, as they do not affect customers' comfort drastically [30].

TABLE I
LOAD CATEGORIES AND CORRESPONDING TYPES OF APPLIANCES

Load controllability	Load categories	Residential appliances
Controllable	CTIM1	HVAC, dish washer, tumble dryer, washing machine, washer-dryer, vacuum cleaner
	QTIM1	Chest freezer, fridge-freezer, fridge, upright freezer
	R_C	Water heater, electrical shower, storage heater, electrical space heating
Uncontrollable	R_{UC}	Iron, hob, oven
	SMPS	Answer machine, CD player, Clock, telephone, high fidelity (HiFi) appliances, Fax machine, PC, printer, TV, VCR-DVD, receiver, microwave
	Lighting	Lighting

The diagram in Fig. 1 presents the methodology for demand decomposition in a smart metering system where only some users have per-appliance monitoring, as a fairly realistic scenario in the future smart distribution grid. Two assumptions are made in this respect: *i)* SMs can record real power of individual appliances, while reactive load is derived probabilistically; *ii)* Forecast of the total consumption (real and reactive power) at the substation (block {5}) is already available. As an initial step before the demand decomposition process, the SM data is pre-processed and aggregated at data concentrator point (block {1}). Following the first assumption, the part of the consumption which has sub-metering can be decomposed into categories or controllable/uncontrollable load by simply aggregating consumption of appliances belonging to the same category (block {2}), as detailed in Table I. The aggregated demand composition of these users can be presented as shown in Fig. 7 and 11 of the Section V. (Note: A number of 1000 is adopted as the total number of households/end-users supplied from the substation, including those with and without sub-metering.)
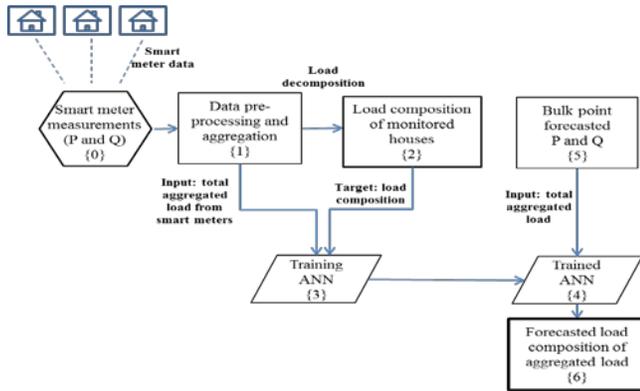
Fig. 1. Flow chart for load disaggregation in case of smart metering system with partial coverage

In the next step, the ANN is trained with this sub-metering data in order to be able to "recognize" the load composition based only on aggregated active and reactive load curve of the monitored users. Once trained, the ANN (block {4}) uses forecast of the total active and reactive load at the concentrator point (block {5}) as the input, and gives corresponding load composition, i.e. weighting factors of each load category, as the output (block {6}). The two main steps of the methodology are discussed in the following sections.

A. Data conditioning and aggregation

In order to present as realistically as possible the future smart metering system, two assumptions are made: i) there are missing samples in the data streams of SMs; ii) different SMs have different sampling steps (with one, ten, thirty or sixty minute granularity). The second assumption is based on [31] where it was reported that the active and reactive consumption could be measured over periods from 1 to 60 minutes. One minute sampling rate is taken as the reference value here, as it avoids under-estimation of electrical consumption and provides sufficient data for detailed modelling of distribution networks [15].

Step 1 (Blocks {0} and {1}): Smart meter data aggregation and pre-processing

As the proposed methodology uses aggregated SM data, the first step is to pre-process "raw" data streams, i.e., restore the missing data and adjust the granularity of all the streams to minute-based samples. According to [32], up to 20% of active load measurements at substation points are inaccurate. Thus, it is assumed that there is 20% of missing data in the overall data set coming from the available SMs, due to some kind of sensor fault. In addition, there will be also samples missing in the data streams obtained with sampling steps larger than one minute. In the preliminary studies, the noise had been added to some data streams using Gaussian White noise with relatively low signal-to-noise ratio. The noise was then filtered using locally weighted polynomial regression method [33]. This however, had negligible effect on the accuracy of the results and hence the noise was not taken into account in further studies, including these.

The missing samples, which were missing at random, were restored using simple linear interpolation (LI) between the available (recorded) samples. In spite of the existence of various sophisticated data restoration methods [34-36], some of which were tested and compared in preliminary studies [36], LI is chosen as a computationally efficient method, showing acceptable (considering all other uncertainties involved) accuracy of the results. Performance of the approach is assessed via the total relative error e_i calculated in each time step i using (1).

$$e_i = \frac{|P_i - P_i^0|}{P_i^0} * 100\% \qquad (1)$$

P_i^0 is the actual (with no missing data) total active power of the aggregation, and P_i is the total active power obtained after data restoration. Table II shows the 90th percentile of the error over a range of aggregation levels for LI method and for the case when missing data were not restored (i.e. missing data replaced with zero). As expected, the errors are higher at lower aggregation levels, and range between 78% and 98% when there is no data pre-processing, and between 9% and 35% when LI is used. (Note: The amount of missing data was always 20%, so the number of missing samples grew proportionally to the aggregation level.)

TABLE II
90TH PERCENTILE OF THE TOTAL RELATIVE ERROR FOR ACTIVE POWER DATA RESTORATION

Aggregation level	90th percentile of the total relative error	
	Replacing by zero	Linear interpolation
1000 households	78%	9%
800 households	85%	9%
500 households	88%	9%
200 households	93%	18%
100 households	96%	29%
50 households	98%	35%

It is possible that using advanced data mining methods such as optimisation [34] or Expectation Maximisation [35] algorithms, could enhance this accuracy further, in particular when larger blocks of missing data are present in data streams. The weight adjusted k-nearest neighbour (WAkNN) [36] method, for example, was tested within this study, however it showed poor perfomance on individual data streams with missing data. In conclusion, it was deemed unnecessary to use more advanced data restoration methods, as the overall effect of missing data restoration on the accuracy of load disaggregation was minor/not significant.

B. Demand Decomposition

After the data streams coming from available SMs have been pre-processed, the data streams are aggregated and further used for demand decomposition.

Step 2 (Blocks {3} and {4}): ANN training

A two-layer (with one hidden layer) feed-forward ANN with Bayesian Regulation Backpropagation, similar to the one introduced in [37], was used for total forecasted demand decomposition. The transfer functions of the hidden layer and output layer are log-sigmoid and tan-sigmoid, respectively. Other machine learning methods that could be used for load decomposition, such as support vector machine, for example, were not analysed in this paper, as the ANN was recommended for this purpose based on comparison of

different approaches performed in [37, 38]. In addition, a comparison with a time series forecasting method, namely autoregressive integrated moving average (ARIMA [39]), was performed. It showed to be unsuitable for load decomposition due to the limited observability of the end-users' per-appliance consumption, which is the main assumption in this study. The ARIMA output for individual load category had to be scaled up to the total number of users, which resulted in errors larger (about 20% larger error in case of the most comparable results) than with the methodology proposed in the paper.

Data pre-processing for the ANN included:

- Restoration of missing data by linear interpolation, as described in part *A* of this Section;
- Data scaling – in order to set all the input values in a comparable range, active and reactive load data was scaled to the range {0,1} taking maximum monthly active load of the aggregation as the base value.

In the approach introduced in this paper, the aggregated sub-metering data is used for training the ANN, using total active and reactive power (of the aggregated monitored users) as input data and calculated participation (per unit contribution – weighting factors in total demand at given point in time) of the six categories as the target data (block {3} in Fig. 1). Target data, defined as W, presents the participation (share) of each load category in every time step. If in a time step i, active load of category j equals P_{ji}, then the participation w_{ji}^P (in p.u.) of that category is given as:

$$w_{ji}^P = \frac{P_{ji}}{P_i}, \tag{2}$$

where P_i is the total active load at a time step i. At each time step the following condition has to be fulfilled:

$$\sum_{j=1}^{6} w_{ji}^P = 1, \tag{3}$$

where 1 refers to the value of the total load at that time step.

The ANN training process is performed over 7 days data, which includes minute-based aggregated real (P) and reactive (Q) power measurements, giving 7*1440 minute-based samples for each of the variables, presented in a matrix form as follows:

$$Input = \begin{bmatrix} P_1 & \cdots & P_i & \cdots & P_{7*1440} \\ Q_1 & \cdots & Q_i & \cdots & Q_{7*1440} \end{bmatrix} \tag{4}$$

Active and reactive load are given in per unit (p.u.) values, based on the maximum monthly active load (at the substation). (Note: cases with and without voltage measurements as part of the input data were compared in [37] and it was concluded that the absence of voltage data did not affect the accuracy notably.) The target data can then be presented in a matrix form as follows:

$$Target = \begin{bmatrix} w_{1,1}^P & \cdots & w_{1,7*1440}^P \\ w_{2,1}^P & \cdots & w_{2,7*1440}^P \\ \vdots & \cdots & \vdots \\ w_{6,1}^P & \cdots & w_{6,7*1440}^P \end{bmatrix} \tag{5}$$

As the number of the training samples (*N*=7*1440) was much larger than the number of input variables (*d*=2), the number of neurons in the hidden layer (*n*) was calculated as follows [38]:

$$n = \sqrt{\frac{N}{d\,lnN}} \tag{6}$$

Fig. 2 illustrates the input and output for the ANN, while Fig. 3 represents the architecture of the ANN used in Matlab [40], where *w* and *b* correspond to network's weights and biases, respectively, assigned to the inputs in hidden and output layer. The number of hidden layer neurons is 44 in this case, as the values for *N* and *d* in Equation (6) are 7*1440 and 2, respectively.
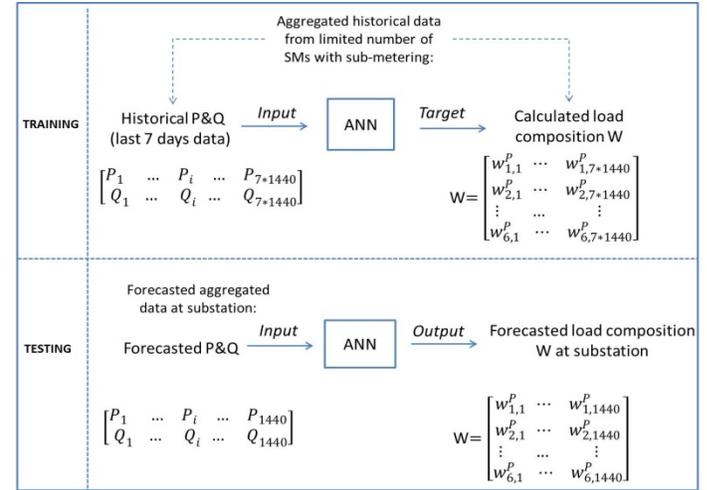

Fig. 2 Detailed presentation of ANN input and output, during training and testing process
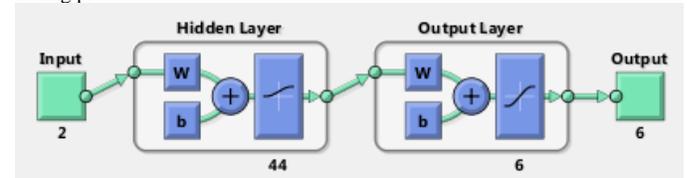

Fig. 3 Architecture of the feedforward network used in Matlab, adopted from [40]

In case of the reactive load, the participation of each category is given as:

$$w_{ji}^Q = \frac{Q_{ji}}{Q_i} = \frac{P_{ji}\tan(\varphi_{ji})}{P_i\tan(\varphi_i)} = w_{ji}^P \frac{\tan(\varphi_{ji})}{\tan(\varphi_i)}, \tag{7}$$

where φ_{ji} and φ_i are phase angles of category j and the total load at time step i, respectively.

In order to test the accuracy of the ANN, the network is tested using one day (the eighth day) data from the historical dataset, which served as the day-ahead load forecast. It should be noted that the inaccuracy of the load forecast was not taken into account in this study. The training of the ANN with 7 days data took between 3 and 8 minutes using a PC with the 64-bit operating system and 3.40 GHz processor. Training with larger sets of data, e.g. 28 days, was also performed but it did not result in higher accuracy of results. Once the ANN is trained, the forecasted active and reactive load curves at the substation (block {5} in Fig. 1) are used as the input to the trained ANN (block {4} in Fig. 1). The output is presented in a form of decomposed forecasted active and reactive load curves for all aggregated end-users (monitored and non-monitored ones), as it will be illustrated in Section V. It is important to note that the training dataset should not include days with activated DR programs, as this data will not show the actual DR potential, but the loading curve and demand composition after load shifting or curtailment.

IV. PROBABILISTIC DERIVATION OF REACTIVE LOAD

For a complete profile of the aggregated load in an area, both active and reactive load curves are needed. Reactive load is also necessary as the input variable for ANN training. In cases where SMs record only active demand, reactive demand can be derived probabilistically. A bottom-up approach is therefore taken in this paper, by considering the possible ranges of PFs for different home appliances, adopted from manufacturers' websites. Assuming that the active demand of individual appliances is monitored, reactive demand of the monitored users is derived probabilistically for each appliance and in every time step by running Monte Carlo simulations over the most common values of PF for each type of residential appliance. PF value for each appliance in each time step is sampled 100 times using randomization with uniformly distributed samples within the considered range to account for PFs of different devices and possible variability of this PF from one operating condition to the other. Then, the probabilistic reactive load values $Q_{k,j,i}$ are calculated for each appliance j in each time step i based on the (deterministic) active load of the appliance P_{ji} and the corresponding probabilistic values of the $PF_{k,j,i}$ ($k=1\div100$ in each time step), as follows:

$$Q_{k,j,i} = P_{ji} \cdot \sqrt{1 - PF_{k,j,i}{}^2} / PF_{k,j,i} \qquad (8)$$

In order to validate the approach, real load data from a 15kV substation was chosen for testing. The available data had half hour based measurements of voltage and current and yearly information on typical values of the PF based on the period of the day (48 samples), type of the day (working day, Saturday or Sunday) and season of the year (winter/summer). Although the PF is not measured, but given by the electric utility, it is adopted as the correct value used for calculating active and reactive load. As there was no sub-metering in the dataset, the demand was decomposed using probabilistic approach (i.e. ANN trained with probabilistically generated data), as described in [7].

In the next step, after active demand composition was obtained with respect to load categories, a range of min/max PF values for each category had to be derived as there was no per-appliance data for this pilot site. This was done using the CREST model [15] for 1000 end-users and extracting two probabilistic values for each appliance in each time step (relying on (8)), namely, the mean and the most probable value of reactive load. Finally, by aggregating corresponding appliances' reactive load into categories, two sets of PFs per load category were derived, based on mean and most probable values, and correspondingly, two ranges of PFs (min/max value) for each load category.

In the following step, PFs were randomized (following uniform distribution) 1000 times for each load category and in each time step over the two obtained ranges; accordingly, two decomposed reactive load curves were derived based on the test site's active load measurements and the probabilistic PF values. The obtained total reactive load curves, as the sums of all 6 individual categories, are presented in Fig. 4 and compared with the original reactive load curve. The grey area shows the range between maximum and minimum possible reactive load, based on minimum and maximum PF values, respectively.
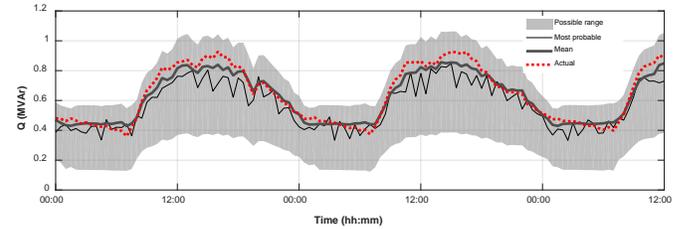

Fig. 4. Derived and actual reactive load curves

It can be seen that the reactive curve built up based on mean values of the PF is closer to the actual one. The mean absolute percentage error (MAPE [29]) was used to assess the accuracy, giving 12.9% error for the curve based on the most probable PFs, and 5.3% for the curve based on mean values of the PF. Another measure of accuracy is the mean square error (MSE [41]) whose value was 0.0099 MVAr^2 for the curve based on the most probable PFs, and 0.0017 MVAr^2 for the curve based on mean values of the PF. Therefore, in equation (8) of the methodology, it was decided to use the mean PF value of the probabilistic set of values to derive reactive load data.

V. CASE STUDIES

An aggregation of 1000 households is analyzed in this paper, illustrating a relatively high number of users. The load model given in [15] was used to generate individual load profiles (decomposed into home appliances) over one month, which also served as training and testing data sets for the ANN. The residential occupancy statistics for the UK (29% of households accommodate a single resident, 35% accommodate two, 16% have three residents and 20% have four) is adopted from [42] to generate appropriate load profiles of residential customers.

The three illustrative case studies analyzed in the paper are as follows:

- *Case A*: Smart metering system with all meters sending data every minute, with no missing data;
- *Case B*: Smart metering system with meters sending data every minute, with 20% missing data ("NaN" values);
- *Case C*: Smart metering system with different meters sending data at different time steps (1, 10, 30 and 60 minutes) and with 20% missing data.

In order to assess the required percentage of users with sub-metering data, five levels of SM coverage are investigated within each case study: 5% coverage (50 households out of 1000 have SMs), 10%, 20%, 50% and 80% SM coverage. The ANN is trained, correspondingly, with sub-metering data coming from 50, 100, 200, 500 or 800 households. The objective of these examples is to illustrate the effect of SM (with sub-metering functionality) coverage on the accuracy of demand decomposition in an aggregation of 1000 households. The accuracy is calculated compared to the composition of the aggregated load (during the eighth day) obtained from the actual values in the given dataset with 1000 households. Absolute weighting factor error (AWFE) is used for this

purpose, and calculated in each time step (minute) as follows:

$$AWFE_{cat} = \left| WF_{cat,ANN} - WF_{cat,real} \right| \quad (9)$$

where $WF_{cat,ANN}$ is the share of the load category obtained as the result of the ANN, and $WF_{cat,real}$ is the actual share of the category, both given in p.u. based on the average aggregated monthly active load.

Fig. 5 illustrates the way errors are accounted for, on the example of two load categories, namely controllable and uncontrollable load. If the total load at time t equals 0.7 p.u. (where 1 p.u. refers to the average monthly load at the aggregation point, which is in this study around 0.6 MW), and the estimated load shares of controllable and uncontrollable load are 0.3 p.u. and 0.4 p.u., respectively, then the real values of the shares are within the following ranges: P_c=0.3p.u. $\pm AWFE$ and P_{uc}=0.4p.u. $\pm AWFE$.
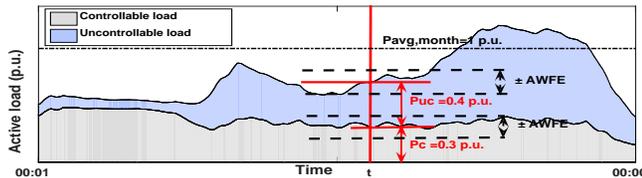

Fig. 5. Presentation of the confidence level on the example of controllable/uncontrollable load over one day (24 hours)

In all the cases (A to C), the errors are compared to those obtained with 100% SM coverage, as the reference case, revealing this way the size of the error coming from the ANN itself. The load forecasting error is not taken into account in this study. In addition, the errors are compared with those in case of 0% SM coverage (no SMs installed at the users' premises), where the ANN is trained with probabilistically derived data, originating from statistical data about the electricity usage in UK domestic sector [43]. According to these data, controllable load within the total daily load ranged between 15 and 50%. The training data was generated following approach similar to the one described in [7]. As the probabilistic training data for ANN is generated only once, this ANN (at 0% coverage) has the same accuracy for cases A, B and C.

The results are presented in the form of cumulative density functions (CDFs) of the AWFE over a range of SM coverage levels, including the reference (100%) and 0% coverage, for controllable load shares only. At the same time, the 90th percentile confidence level of the AWFE for different load categories and controllable load are presented in a form of bar plots over a range of SM coverage levels. The 90th percentile confidence level is chosen as it shows the maximum error for 90% of the observed time steps (here, 1296 out of 1440 time steps over a 24h period).

A. Active load decomposition

The accuracy of estimation of controllable load shares is very similar between cases A-C, which is why only case C is presented in Fig. 6. It can be seen from the figure that all SM coverage levels provide errors smaller than 0.1 p.u. (i.e. 10% of the average monthly load, which corresponds to around 60 kW) in 90% of the time steps. In cases where there is no sub-metering provided and the estimation can only be done probabilistically, the 90th percentile of the AWFE is 0.23 p.u,

which corresponds to around 140 kW. It can be also seen that for the SM coverage levels of 50% and higher, the accuracy remains the same. The calculated errors for the three cases (A to C) do not change notably, confirming that 20% missing data, together with different sampling steps, do not affect the accuracy significantly if the missing samples are restored. This also confirms that the use of simple data restoration method (here, LI) is fully justified.

Figure 7 illustrates the load composition in the base case and with 5% SM coverage. It should be noted that the total active demand is the same in the two figures, only the shares of the load categories differ.
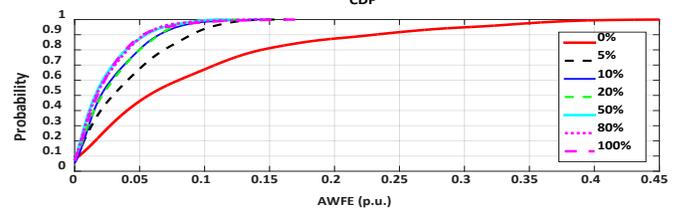

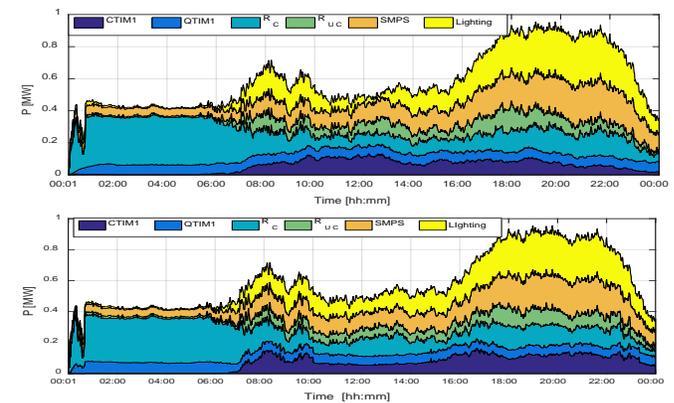Fig. 6 CDF of AWFE for the estimation of controllable active load, case C

Fig. 7. Demand composition of 1000 end-users' active load based on actual values (top) and based on the ANN trained with data from 50 users (bottom)

The estimated shares of individual load categories show similar accuracy (errors up to 0.1 p.u.) in most cases, as seen in Fig. 8, which represents the 90th percentile of AWFE read from the CDF plots for each load category in case C. The highest accuracy is seen with QTIM1, and the lowest with CTIM1. Some categories' shares, namely CTIM1 and QTIM1, are very accurately assessed even with 0% coverage (with errors around 0.1 and 0.05 p.u., respectively), which shows that in these cases only statistical data is sufficient for confident load decomposition. For all other categories, probabilistic approach introduces higher errors. The calculated range of errors for cases A and B is very similar to errors illustrated in Fig. 6 for case C. Different accuracy in prediction of participation of different load categories in total demand can be attributed to variation of particular load category during the observation period. The correlation between total active/reactive demand and shares of different load categories is studied using Spearman's rank coefficient of nonlinear correlation [44], as the correlation between the parameters is not linear. Category QTIM1 shows the highest Spearman's coefficient (Spearman's rho) at most aggregation levels, followed by lighting, hence the highest accuracy in prediction. Category CTIM1, on the other hand, has lower Spearman's coefficient than other appliances in general, when

comparing their correlation with total active and reactive load, leading to the lowest accuracy in prediction.
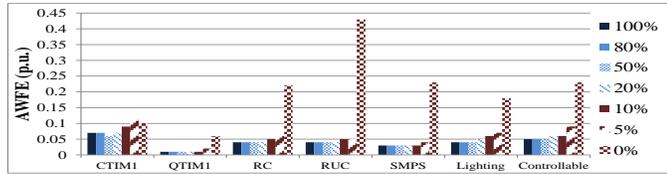

Fig. 8 AWFE with 90th percentile confidence level for real power, case C

B. Reactive load decomposition

The estimation of shares of controllable reactive load is also very accurate, between 0.04 p.u. and 0.08 p.u. over the range of SM coverage levels, as shown in Fig. 9. In case of 0% coverage, the 90th percentile of AWFE is 0.17 p.u., which corresponds to around 100 kVAr. Similarly to active load, there is only a minor deterioration in the accuracy in cases B and C at lower SM coverage levels, compared to case A.
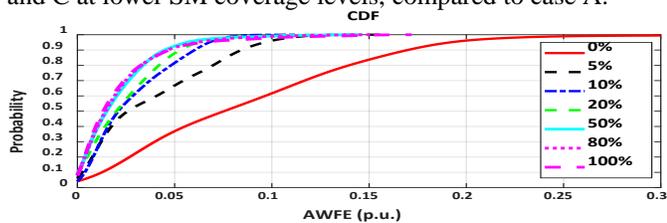

Fig. 9 CDF of AWFE for the estimation of controllable reactive load, case C

The estimation of load categories' shares shows AWFE between 0 and 0.08 p.u. for all the SM coverage levels, as shown in Fig. 10. The reactive power of the controllable resistive loads (R_C) equals zero in all time steps, which is why the 90th percentile of AWFE is around 0, except for the case of probabilistic approach. The same applies for uncontrollable resistive loads which are, due to some home appliances such is oven, modelled as imperfect resistors. Except for these two categories, the QTIM1 share is estimated with the highest accuracy and the CTIM1 and Lighting with the lowest.

Fig. 11 illustrates the load composition in the base case and with 5% SM coverage.
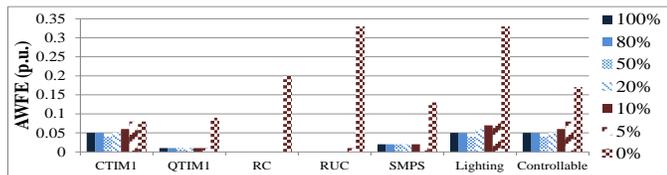

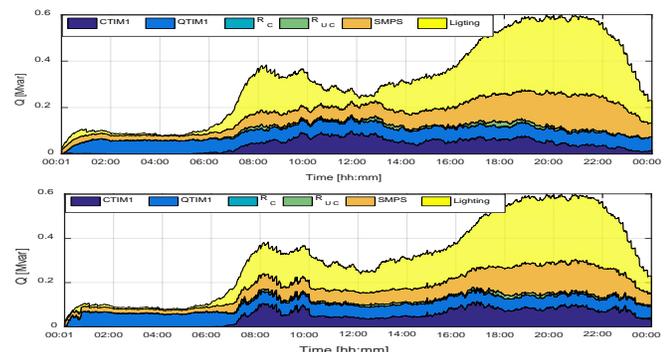Fig. 10 AWFE with 90th percentile confidence level for reactive power, case C

Fig. 11. Demand composition of 1000 end-users' reactive load based on actual values (top) and based on the ANN trained with data from 50 users (bottom)

C. Discussion

Based on the studies performed and illustrative results shown in the previous section, it was found that the overall accuracy of the assessment is not significantly affected by SM coverage, nor by missing data. The accuracy, though, changes more with the coverage level, than with missing data. Furthermore, the daily load shares for some load categories (CTIM1 and QTIM1), can be estimated very accurately only based on statistical data on electricity usage in the area, i.e., without any SMs. For other load categories, the utilization of ANN with probabilistically generated training data is justified only if the target application of demand decomposition does not require high accuracy. It was also observed that the estimation of reactive load composition with 0% SM coverage is more accurate than the estimation of active load.

Furthermore, for the DR programs relying on wet appliances (CTIM1) and cold appliances (QTIM1), only limited statistical data can be used to estimate their load shares. At lower SM coverage levels, AWFE values for these two categories are up to around 0.1 p.u. In case of voltage-based DR relying on resistive loads (e.g. water heaters), AWFE of the load share equals 0.06 p.u. at 5% SM coverage, while it is lower than 0.05 p.u. for other levels.

VI. VALIDATION OF THE METHODOLOGY

In order to validate the methodology on a real dataset and test the effect of weather data on the accuracy of results, another dataset was chosen, namely Pecan street electricity consumption data [16], which includes minute-based data from 200 residential users located in Austin, Texas, during June-August 2015. All of the observed time series had all samples available, i.e., there was no missing data, similarly to case A discussed in Section V. Therefore, only the influence of SM coverage was examined. Sub-metering data is available over a set of years, starting from 2013, and includes circuit-based measurements, e.g. total consumption of the living room, bathroom, or kitchen plugs, and in some cases individual appliances' measurements (washing machine, furnace, etc.). Therefore, some assumptions on the load categories used in different parts of the household were made to enable load classification into 6 categories. For example, it was assumed that half of the loading in the living room consisted of lights and half consisted of the electronic appliances (SMPS category). Weather data includes temperature, humidity and wind speed measurements from the corresponding period of the year, together with the day type (1 for working day and zero for non-working day). The dataset did not contain reactive load data, so the probabilistic approach was taken to derive this, following steps described in section IV. The ANN trained with probabilistic data only (with 0% SM coverage) was not taken into account in this study.

Figures 12-14 present the CDF of AWFE for controllable active load estimated for one day in August based on ANN trained with different size of historical data, namely two previous (most recent) months (Fig. 12), one previous (most recent) month (Fig. 13) and one previous (most recent) week

(Fig. 14). In each case, the methodology was tested with and without weather data. As can be seen from the figures, the longer historical data and the inclusion of weather data do not improve the accuracy – quite contrary, the use of longer historic data results in slightly reduced accuracy while the use of weather data results in higher variability in accuracy (both slight improvement and slight deterioration in accuracy can be observed) at different SM coverage levels. Furthermore, the range of the 90th percentile of AWFE is very similar to the ones in the test cases presented in Section V-A). Therefore, only the most recent historic data (past 7 days) can be used for training the ANN to obtain the accuracy presented in case studies in Section V and the inclusion of the weather data will not in general improve the accuracy of the results. The presented results confirm the validity of the methodology and the fact that the training data corresponding to the observed set of consumers yields high accuracy.

VII. CONCLUSION

This paper presented methodology for aggregated demand decomposition based on sub-metering enabled SM measurements and ANN. The overall methodology results in an estimation of the shares of different load categories and overall controllable and non-controllable load within the total forecasted load, with a foreseen application in various DR programs. The two main assumptions are that the SMs can record active load of individual appliances and that only some of the end-users in the aggregation have this type of SM. An approach for obtaining probabilistic aggregated reactive load curve is also discussed as a solution to the lack of reactive load measurements at the end-users' point.

The results of the case studies have shown the effect of missing data and different SM coverage on the accuracy of forecasted active and reactive load composition. It can be concluded that the SM coverage level has higher impact on the accuracy than the missing data samples. The shares of all predefined load categories were assessed with high accuracy, some even in cases without SMs, in which case the ANN was trained with probabilistically derived data. For most of the load categories even the lowest SM coverage (5%) enables very high accuracy of identification of load composition. The results show that even with minimal investments in sub-metering technologies (for only 5% of the users) the desired accuracy of load decomposition forecast can be obtained, and notably improved compared to the probabilistic approach, when no sub-metering data is available. The minimum SM coverage and resulting accuracy of estimation, however, may be different for different applications, e.g., estimation of dynamic response of demand, and needs to be investigated further. The methodology was further validated on an actual pilot site's dataset, resulting with the similar accuracy and confirming that there is no need for weather data and large historical data to be included. Even though the methodology has been illustrated on residential load sector, it can be equally well applied to other load sectors, e.g., commercial, industrial or mixed. Finally, it is important to note that the proposed methodology is area-dependent. Therefore, the accuracy is

higher when the data used for training the ANN comes from the electrical or geographical "neighborhood" of the aggregated users. In addition, statistical data has to correspond to the type of the users under analysis, as the daily range of controllable load differs among industrial, commercial and domestic types of users. The proposed approach can be used for either real-time or forecasting applications, although the focus of this paper was on day-ahead forecasting.
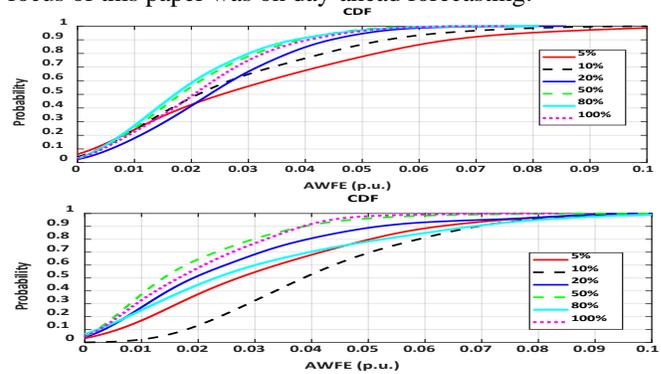

Fig. 12. Controllable load estimation for one day in August with 60 days training (June+July), without (top) and with (bottom) weather and day type
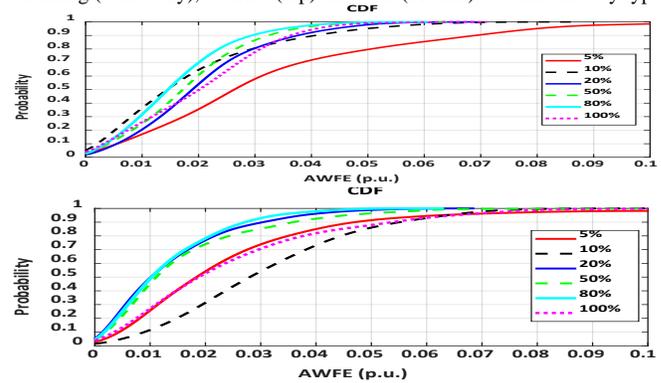

Fig. 13. Controllable load estimation for one day in August with one month training (July), without (top) and with (bottom) weather and day type
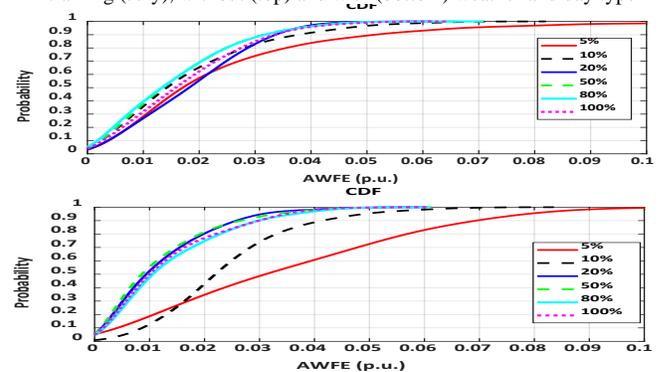

Fig. 14. Controllable load estimation for one day in August with 7 days training (August), without (top) and with (bottom) weather and day type

REFERENCES

[1]   M. H. Albadi and E. F. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research,* vol. 78, pp. 1989-1996, 2008.

[2]   C. C. C4.112, "Guidelines for Power Quality Monitoring - Measurement Locations, Processing and Presentation of Data ", ed, 2014.

[3]   K. Samarakoon, J. Ekanayake, and N. Jenkins, "Reporting Available Demand Response," *Smart Grid, IEEE Transactions on,* vol. 4, pp. 1842-1851, 2013.

[4]   M. Pipattanasomporn, M. Kuzlu, S. Rahman, and Y. Teklu, "Load Profiles of Selected Major Household Appliances and Their Demand Response Opportunities," *IEEE Transactions on Smart Grid,* vol. 5, pp. 742-750, 2014.

[5]   "Digest of United Kingdom Energy Statistics 2015," Department of Energy and Climate Change, [Online].Available:https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/450302/DUKES_2015.pdf.

[6]   F. L. Müller, J. Szabó, O. Sundström, and J. Lygeros, "Aggregation and Disaggregation of Energetic Flexibility from Distributed Energy Resources," *arXiv preprint arXiv:1705.02815,* 2017.

[7]   Y. Xu and J. V. Milanović, "Artificial-Intelligence-Based Methodology for Load Disaggregation at Bulk Supply Point," *IEEE Transactions on Power Systems,* vol. 30, pp. 795-803, 2015.

[8]   D. Srinivasan, W. S. Ng, and A. C. Liew, "Neural-network-based signature recognition for harmonic source identification," *IEEE Transactions on Power Delivery,* vol. 21, pp. 398-405, 2006.

[9]   J. Z. Kolter and M. J. Johnson, "REDD: A public data set for energy disaggregation research," in *Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA*, 2011, pp. 59-62.

[10]   A. Rahimpour, H. Qi, D. Fugate, and T. Kuruganti, "Non-Intrusive Energy Disaggregation Using Non-negative Matrix Factorization with Sum-to-k Constraint," *IEEE Transactions on Power Systems,* 2017.

[11]   W. Kong, Z. Y. Dong, J. Ma, D. Hill, J. Zhao, and F. Luo, "An Extensible Approach for Non-Intrusive Load Disaggregation with Smart Meter Data," *IEEE Transactions on Smart Grid,* 2016.

[12]   S. Singh and A. Majumdar, "Deep Sparse Coding for Non-Intrusive Load Monitoring," *IEEE Transactions on Smart Grid,* 2017.

[13]   A. Gerossier, T. Barbier, and R. Girard, "A novel method for decomposing electricity feeder load into elementary profiles from customer information," *Applied Energy,* vol. 203, pp. 752-760, 2017.

[14]   J. Hu, J. Cao, M. Z. Chen, J. Yu, J. Yao, S. Yang, and T. Yong, "Load following of multiple heterogeneous TCL aggregators by centralized control," *IEEE Transactions on Power Systems,* vol. 32, pp. 3157-3167, 2017.

[15]   I. Richardson, M. Thomson, D. Infield, and C. Clifford, "Domestic electricity use: A high-resolution energy demand model," *Energy and Buildings,* vol. 42, pp. 1878-1887, 2010.

[16]   Pecan Street Inc. Dataport 2017 [Online]. Available: http://www.pecanstreet.org/

[17]   J. Yao, W. Zhengyu, C. Jiang, and Y. Zhang, "Dispatch and bidding strategy of active distribution network in energy and ancillary services market," *Journal of Modern Power Systems and Clean Energy,* vol. 3, pp. 565-572, 2015.

[18]   Y. Xiang, L. Junyong, Y. Wei, and C. Huang, "Active energy management strategies for active distribution system," *Journal of Modern Power Systems and Clean Energy,* vol. 3, pp. 533-543, 2015.

[19]   P. Du and N. Lu, "Appliance Commitment for Household Load Scheduling," *IEEE Transactions on Smart Grid,* vol. 2, pp. 411-419, 2011.

[20]   B. P. Bhattarai, B. Bak-Jensen, P. Mahat, and J. R. Pillai, "Voltage controlled dynamic demand response," in *IEEE PES ISGT Europe 2013*, 2013, pp. 1-5.

[21]   A. Ballanti and L. F. Ochoa, "Initial assessment of voltage-led demand response from UK residential loads," in *Innovative Smart Grid Technologies Conference (ISGT), 2015 IEEE Power & Energy Society*, 2015, pp. 1-5.

[22]   H. Zhong, Q. Xia, C. Kang, M. Ding, J. Yao, and S. Yang, "An Efficient Decomposition Method for the Integrated Dispatch of Generation and Load," *IEEE Transactions on Power Systems,* vol. 30, pp. 2923-2933, 2015.

[23]   M. Muratori and G. Rizzoni, "Residential Demand Response: Dynamic Energy Management and Time-Varying Electricity Pricing," *IEEE Transactions on Power Systems,* vol. 31, pp. 1108-1117, 2016.

[24]   A. H. Mohsenian-Rad and A. Leon-Garcia, "Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments," *Smart Grid, IEEE Transactions on,* vol. 1, pp. 120-133, 2010.

[25]   S. Shao, M. Pipattanasomporn, and S. Rahman, "Grid Integration of Electric Vehicles and Demand Response With Customer Choice," *IEEE Transactions on Smart Grid,* vol. 3, pp. 543-550, 2012.

[26]   C. Vivekananthan, Y. Mishra, G. Ledwich, and L. Fangxing, "Demand Response for Residential Appliances via Customer Reward Scheme," *Smart Grid, IEEE Transactions on,* vol. 5, pp. 809-820, 2014.

[27]   C. Zhang, Y. Xu, Z. Y. Dong, and J. Ma, "Robust Operation of Microgrids via Two-Stage Coordinated Energy Storage and Direct Load Control," *IEEE Transactions on Power Systems,* vol. 32, pp. 2858-2868, 2017.

[28]   Y. V. Makarov, D. J. Hill, and J. V. Milanovic, "Effect of load uncertainty on small disturbance stability margins in open-access power systems," in *System Sciences, 1997, Proceedings of the Thirtieth Hawaii International Conference on*, 1997, pp. 648-657.

[29]   Y. Xu and J. V. Milanović, "Day-Ahead Prediction and Shaping of Dynamic Response of Demand at Bulk Supply Points," *IEEE Transactions on Power Systems,* vol. 31, pp. 3100-3108, 2016.

[30]   J. Hu, J. Cao, M. Z. Q. Chen, J. Yu, J. Yao, S. Yang, and T. Yong, "Load Following of Multiple Heterogeneous TCL Aggregators by Centralized Control," *IEEE Transactions on Power Systems,* vol. 32, pp. 3157-3167, 2017.

[31]   "D3.4 Smart meters architecture and data model analysis ", NOBEL GRID project, [Online].Available: http://nobelgrid.eu/deliverables/ , 2016.

[32]   X. Chen, C. Kang, X. Tong, Q. Xia, and J. Yang, "Improving the Accuracy of Bus Load Forecasting by a Two-Stage Bad Data Identification Method," *IEEE Transactions on Power Systems,* vol. 29, pp. 1634-1641, 2014.

[33]   W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American statistical association,* vol. 74, pp. 829-836, 1979.

[34]   R. Chiky, L. Decreusefond, and G. Hebrail, "Aggregation of asynchronous electric power consumption time series knowing the integral," presented at the Proceedings of the 13th International Conference on Extending Database Technology, Lausanne, Switzerland, 2010.

[35]   F. V. Nelwamondo, S. Mohamed, and T. Marwala, "Missing data: A comparison of neural network and expectation maximisation techniques," *arXiv preprint arXiv:0704.3474,* 2007.

[36]   E.-H. S. Han, G. Karypis, and V. Kumar, *Text categorization using weight adjusted k-nearest neighbor classification*: Springer, 2001.

[37]   Y. Xu and J. V. Milanović, "Accuracy of ANN based methodology for load composition forecasting at bulk supply buses," in *Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on*, 2014, pp. 1-6.

[38]   Y. Xu, "Probabilistic Estimation and Prediction of the Dynamic Response of the Demand at Bulk Supply Points," PhD thesis, School of Electrical and Electronic Engineering, University of Manchester, 2015.

[39]   G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*: John Wiley & Sons, 2015.

[40]   M. H. Beale, M. T. Hagan, and H. B. Demuth. (2017). *Matlab Neural Network Toolbox - User's Guide*.

[41]   R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting,* vol. 22, pp. 679-688, 2006.

[42]   F. Lamberti, D. Cuicai, V. Calderaro, and L. F. Ochoa, "Estimating the load response to voltage changes at UK primary substations," in *IEEE PES ISGT Europe 2013*, 2013, pp. 1-5.

[43]   "The impact of changing energy use patterns in buildings on peak electricity demand in the UK," Building Research Establishment Ltd, 2008.

[44]   M. M. Mukaka, "A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal,* vol. 24, pp. 69-71, 2012.

Jelena Ponoćko (S'15) received her BS and MSc degree in electrical engineering and computer science from the University of Belgrade, Belgrade, Serbia in 2011 and 2012, respectively. She is currently working towards the PhD degree at the School of Electrical and Electronic Engineering, University of Manchester, UK, in the areas of demand response and load profiling.

Jovica V. Milanović (M'95, SM'98, F'10) is a Professor of Electrical Power Engineering, Deputy Head of School and Director of External Affairs in the School of Electrical and Electronic Engineering at the University of Manchester, U.K.