Analysis of Diversified Residential Demand in London using Smart Meter and Demographic Data

Mingyang Sun, Ioannis Konstantelos, Goran Strbac
Electrical and Electronic Engineering
Imperial College London
London, UK
{mingyang.sun11, i.konstantelos, g.strbac}@imperial.ac.uk

***Abstract*—In the interest of economic efficiency, design of distribution networks should be taillored to the demonstrated needs of its consumers. However, in the absence of detailed knowledge related to the characteristics of electricity consumption, planning has traditionally been carried out on the basis of empirical metrics; conservative estimates of individual peak consumption levels and of demand diversification across multiple consumers. Although such practices have served the industry well, the advent of smart metering opens up the possibility for gaining valuable insights on demand patterns, resulting in enhanced planning capabilities. This paper is motivated by the collection of demand measurements across 2,639 households in London, as part of Low Carbon London project's smart-metering trial. Demand diversity and other metrics of interest are quantified for the entire dataset as well as across different customer classes, investigating the degree to which occupancy level and wealth can be used to infer peak demand behavior.**

***Index Terms*--After diversity maximum demand, demand diversity, distribution network planning, smart meter.**

I. Introduction

When designing a distribution substation and sizing assets such as transformers and cables, the peak coincident demand across the house-holds that each asset should serve is a key consideration. Underestimating the coincident peak demand, which in Europe typically occurs during low-temperature winter days, will result in undersized assets and an inability to service load during some periods. On the other hand, over-estimating the peak coincident demand can lead to substantially cost-inefficient decisions, given that the same level of reliability could be provided with less expensive assets of reduced rating. Given that Distribution Network Operators (DNOs) have to meet strict reliability targets in combination with the inability to accurately infer the coincident peak demand of a group of consumers due to lack of instrumentation at the household level, DNOs have historically undertaken network design based on some pre-determined empirical metrics. These metrics are typically conservative to allow for worst case scenarios and have little customization related to the specific planning case at hand. The use of these general-purpose metrics is made possible because of the fact that the coincident demand of a large number of consumers exhibits reduced sensitivity to the attributes of individual consumers, which may otherwise vary wildly when viewed in isolation. This effect is known as demand diversification and is a very important concept that pertains to electricity consumption. Demand diversity exists because the use of individual appliances in different households occurs at different times due to consumers' different schedules and preferences. It follows that accurately quantifying the effect of demand diversity is a quintessential aspect of efficient network design.

In order to measure demand diversity, several metrics have been proposed in the literature. After-diversity maximum demand (ADMD) is one of the most commonly used metrics defined as the coincident peak demand attributed to each customer, as the number of customers connected to the network approaches infinity [1]. This metric can be applied to distribution network design in a straightforward manner in order to estimate the peak coincident load for a large number of consumers [2]-[4] and sizing network assets [5]-[6]. Another metric of interest is the coincidence factor (CF), defined as the ratio of the coincident peak demand of a group of customers to the sum of individual peak demands in this group [7]. CF has also been widely used in the past for planning studies as in [8] and [9]. Other metrics such as the diversity factor and conversion factor have also been proposed in the past for planning purposes as showcased in [10]-[12].

In the UK, ADMD is the metric of choice, having been fully incorporated in DNOs' planning guidelines for asset sizing. In particular, for properties with up to 4 bedrooms and gas heating, an ADMD value of 2kW is used [12]. In case of no gas heating being available on the property, ADMD is increased to 3kW. In addition, the recommendation states that ADMD should be increased by 0.5kW for each additional bedroom. The effect of demand diversification is further taken into account by subsequently multiplying by a scaling factor to compute the coincident peak demand across n households (denoted by D_C^n), as shown in equation (1).

$$D_C^n = 0.7 \times n \times ADMD \times \left(1 + \frac{12}{ADMD \times n}\right) \quad (1)$$

978-1-5090-4168-8/16/$31.00 ©2016 IEEE

Accordingly, assuming up to four-bedroom households without electric heating, the effective ADMD value defined as $ADMD^n = D_C^n/n$ follows the curve shown in Fig. 1.

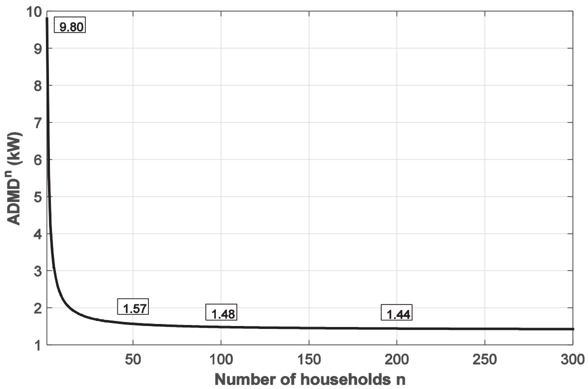

Fig. 1. ADMD as a function of households as suggested by current UK planning methodology. For visual inspection, boxed numbers are provided for 1,50, 100 and 200 households.

As can be seen in Fig.1 above, the maximum demand for a single housheold is set at 9.8kW; ADMD subsequently follows an exponential decrease pattern, reducing considerably in the range of 1-50 households and thereafter achieving a steady-state value that converges to 1.4kW per household. One crucial aspect to be highlighted in the overall calculation methodology is that there is little consideration of the particular details of individual households, (beyond heating type and an approximation of size) leaving little room for differentiation between areas of fundamentally different demographic makeup. It follows that this may be leading to highly conservative metrics to ensure the most adverse of possibilities are well covered, potentially resulting in asset overinvestment. The largest barrier to a transition to more custom and cost-efficient design methodologies has been the unavailability of data to experimentally calculate and validate diversified demand across different consumer groups. This reality is now being changed by the advent of smart meters creating new information streams and raising the opportunity for in-depth exploration and quantification of residential demand. To this end, data measurements from smart meter trials setup in anticipation of the UK-wide rollout target of 2020 constitute a source of valuable information.

The Low Carbon London (LCL) smart meter trial project [13] was setup as part of the UK regulator's Low Carbon Network Fund to enable DNOs to improve the understanding of the electricity consumption in the UK. Along with the collection of demand measurements, an integral part of the trial was the classification of all participating households in terms of occupancy and wealth level, allowing for a further analysis of how consumption is affected by economic factors and the number of people residing in a given property. The idea of making use of demographic data has already been proposed in the past based on the fact that different consumer categories exhibit significantly different levels of demand diversification and electric consumption intensity. The availability of this experimental data presents an unprecedented opportunity to evaluate the applicability of traditional approaches to the modern reality of electricity consumption habits and to design novel distribution network design methodologies on the basis of new information derived from high-resolution smart meter measurements and demographic data.

The paper structure is as follows: Section II introduces the the proposed methodology for quantifying demand diversity from experimental data. Section III presents the LCL dataset consisting of smart meter and demographic data; fundamental demand analyses are performed to show the main patterns of non-diversified and diversified demand. In addition, we calculate ADMD for the segregated subsets and investigate the relation wealth classes and occupancy levels. Section IV summarises and concludes the analysis, highlighting future potential uses of the derived metrics and insights.

II. Diversified Demand Quantification

Estimating the maximum coincident demand of a group of consumers is at the heart of designing efficient distribution networks. To this end, as already mentioned, ADMD is a standard metric used for planning purposes in the UK and other jurisdictions worldwide. Formally, it is defined as the coincident peak electrical demand per customer as the number of customers approaches infinity. Given that the number of customers connected to a distribution asset can vary considerably, we extend the concept of ADMD to be a function of customers connected to the network. As a result, $ADMD^n$ denotes the After Diversity Maximum Demand for n customers. It is constructive to note that, as shown in equation (1) and also in Fig. 1, ADMD typically reaches a steady state value after a large number of households is analyzed; we denote this steady state value by $ADMD^\infty$. Its value and the number of households at which it is reached are of great interest and one of the important outputs of the present research, along with analyzing how $ADMD^n$ evolves as a function of connected households. Given a data matrix $\boldsymbol{D}$ consisting of M observations of electricity demand measurements across N consumers (size denoted $[M \times N]$), the coincident maximum demand D_C^n across n consumers is calculated according to equation (2). Subsequently, $ADMD^n$ is calculated by dividing the coincident peak demand by the number of customers, as shown in (3).

$$D_C^n = \max_{m=1,2,\ldots M} \left\{ \sum_{i=1}^{n} D_{m,i} \right\} \tag{2}$$

$$ADMD^n = D_C^{n\,max}/n \tag{3}$$

In the cases that $n \in \{1, N\}$, the evaluation of equation (3) is trivial. However, in the cases that $1 < n < N$, the large number of possible ways to choose n households among a total of N gives rise to a distribution of $ADMD^n$ values instead of a single value; in fact, there exist $n!/n!\,(N-n)!$ possible household combinations. Naturally, which houses are actually sampled gives rise to some variability; we communicate this variability through the explicit calculation of upper and lower bounds. To this end, in this analysis we are interested in computing the mean, minimum and maximum values, denoted by $\overline{ADMD^n}$, $ADMD_{min}^n$ and $ADMD_{max}^n$ respectively. Naturally, the maximum value is of most importance for practical planning purposes, but average and minimum values serve in gaining further insight towards diversity's degree of variability and convergence behaviour. Furthermore, given that the number of combinations grows extremely fast (for example there are over 10 billion ways to choose 10 out of 50 households), a statistical approximation is necessary for its computation. To this end, random sampling techniques with replacement such as bootstrapping can be used to approximate the values of interest. In essence, instead of exhaustive enumeration we can perform the relevant calculations while only considering a smaller subset of $k \ll n!/n!\,(N-n)!$ combinations, where k is the number of sampled combinations. In order to compute the distribution of $ADMD^n$, a sampling algorithm is proposed as follows.

Step 1. Randomly select n households from the whole dataset $\boldsymbol{D}$. Repeat this process k times with replacement to construct k subsets $\boldsymbol{D}_C^{n(1)}, \boldsymbol{D}_C^{n(2)}, \ldots, \boldsymbol{D}_C^{n(k)}$.

Step 2. Calculate the maximum coincident demand of n household for each subset and construct a new vector $D_C^n = [D_C^{n\,max}{}_1, D_C^{n\,max}{}_2 \ldots, D_C^{n\,max}{}_k]$ representing the sampled population of coincident peak demand levels.

Step 3. Compute the expected, minimum and maximum value of $ADMD^n$.

The above procedure is shown in detail in Fig. 2.

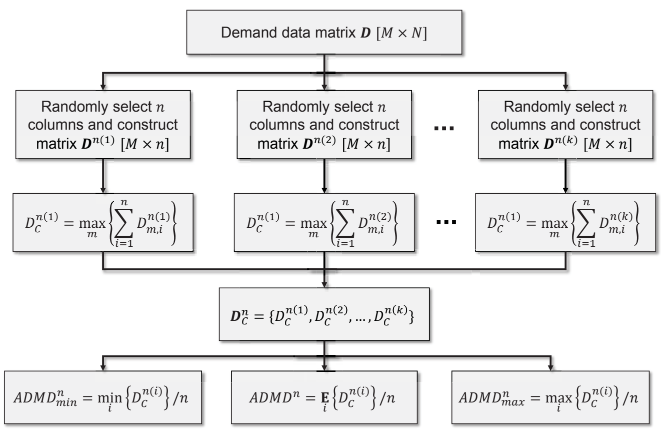

Fig. 2. Sampling methodology for calculating the maximum, minimum and average $ADMD^n$ values from a large dataset. E{} represents the expectation operator.

We continue with applying the presented calculation procedure to the analysis of the LCL dataset, where a k parameter of k^* =100,000 has been used for the cases where $n!/n!\,(N-n)! > k^*$.

III. Analysis of Smart-metering Data

In this section, we present in detail the electricity consumption data measurements recorded by the smart meters and demonstrate how the customers have been split into categories based on property size and wealth. The peak demand distribution for all customer categories is shown and analyzed for the purpose of investigating wealth and occupancy level impact on electricity consumption in the absence of diversification. Subsequently, two analysis of ADMD are undertaken, one utilizing the entire LCL dataset and one focusing on segregated subsets, investigating the evolution of diversified demand as a function of connected households and customer classification.

A. Demand Dataset

The Low Carbon London smart meter trials aim to characterize the residential consumer demand of London and to evaluate the benefits from exploiting smart metering for distribution network design. Within the scope of the LCL project, Landis and Gyr (L+G) E470 electricity meters were installed in 2,639 residential homes across the Mayor of London's Low Carbon Zones (LCZ) and the London Power Networks (LPN) distribution network license area operated by UK Power Networks [14]. In particular, the Engineering Instrumentation Zones (EIZs) of the LCL trial included the areas of Brixton, Merton and Queen's Park. The LCL demand dataset consists of half-hourly load consumption data for a full calendar year from 1st January 2013 to 31st December 2013. As such, the dataset contains 17,520 half-hourly measurements of demand across 2,639 customers in kW. In addition, various socio-economic conditions of the participated household were recorded. In this paper we focus on the data pertaining to household occupancy and wealth level. The former relates to the number of people living in the property. The latter has been drawn on the basis of mapping all participating households to ACORN groups [15]. Subsequently, three wealth classes have been defined: *Adverse*, *Comfortable* and *Affluent* in increasing order. A customer category or class is defined as a combination of occupancy and wealth level. The number of customers belonging to each of the nine categories is shown in Table I.

TABLE I. Number of Participating Households Across Customer Categories

	1 occupant	2 occupants	3+ occupants
Adverse	315	278	234
Comfortable	240	304	214
Affluent	431	400	223

As can be seen above, the LCL smart-metering trials have a good representation across all household classes with several hundred customers in each of the nine categories.

B. Household peak demand analysis

The LCL dataset is summarized in Fig.3 where we present how the peak demand varies within each of the nine categories. We plot the probability density of peak demand of individual households in order to extract some characteristics of interest as well as explore the variability of consumption patterns among households belonging to the same class.

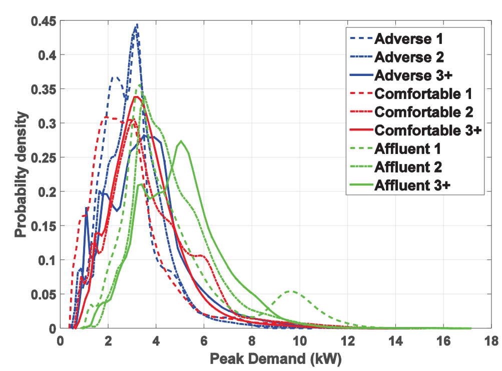

Fig. 3. Peak demand probability distribution across the nine customer categories.

978-1-5090-4168-8/16/$31.00 ©2016 IEEE

The maximum demand measurement for each customer class, equivalent to $ADMD^1$ is shown in Table II. As can be seen below, the maximum demand recorded throughout the entire dataset is of 15.10kW form an *Affluent 3+* household, while the group exhibiting the lowest peak demand of 6.26 kW is *Adverse 1*. It is imperative to highlight that these experimentally derived values do not agree with the current planning guidelines in UK, where the recommended sizing estimations for a single household of up to four bedrooms are 9.80kW and 10.50kW for properties with and without electrical heating respectively. More specifically, in the cases of *Adverse 2* and all categories with 3+ occupants the data gathered from the LCL trails suggest that the guideline values are above the suggested numbers, meaning there is a real risk of asset undersizing in some cases. This may be indicating that consumer habits are changing over time, increasing the number of electrical appliances in households. In general, the patterns of maximum peak demand are as expected, with larger and wealthier households displaying increased values. Notably, there are exceptions in the case of *Comfortable 2*, *Comfortable 3+* and *Affluent 2* households, where they are slightly lower than their *Adverse* counterparts. These differences are relatively small and can be attributed to a single marginal case in the *Adverse 2* class. These reversals indicate that worst case scenarios in the absence of diversification may be difficult to infer on the basis of wealth data. On the other hand, household size is confirmed to be clearly correlated with maximum peak demand and can be a useful proxy for sizing connections of individual households. This is especially important in light of the ease with which household data can be obtained compared to wealth level information as well as their longer-term persistence.

TABLE II. MAXIMUM PEAK DEMAND ($ADMD^1$) FOR DIFFERENT WEALTH AND OCCUPANCY CLASSES

	1 occupant	2 occupants	3+ occupants
Adverse	6.26 kW	10.53 kW	12.60 kW
Comfortable	8.98 kW	9.17 kW	11.28 kW
Affluent	9.06 kW	10.26 kW	15.10 kW

Regarding the population distributions, as can be seen in Fig. 3, the peak demand curves exhibit a unimodal concentration towards larger values for the customers in the categories characterized by increased occupancy and wealth. There is a notable exception for 'Affluent 1' households, where two distinct peaks are observed. This in combination with the large range of values occupied highlights the fact that households belonging to this classification have significantly more variability among them and can potentially exhibit very high consumption levels. Further information obtained by observing the probability density plot is that the peak demands are mostly concentrated in relatively lower values ranging from 2 kW to 8 kW. Notably, there is a very low probability that an extreme high value of peak demand occurs (e.g. above 12 kW) across all categories. It is imperative to highlight that the above analysis only takes into account the peak demand of a single customer, while most of networks are designed to supply multiple customers. In this case the diversification effect becomes paramount, as shown in the following sections.

C. ADMD analysis across all participating households

In the previous subsection, we analyzed how the individual peak demand varies with occupancy and economic factors. In this section we investigate the diversified peak demand behavior as a function of connected customers. We compute $ADMD^n$ across all customers participating in the trial according to the methodology outlined in Section II and show the results in Fig. 4. As mentioned earlier, up to 100,000 combinations have been sampled and three curves are provided. The black curve shows the maximum $ADMD^n$ observed across the sampled combinations, denoted $ADMD^n_{max}$. Dark grey and light grey curves show the average and minimum values respectively. Note that in order to facilitate visual inspection, a plot inset has been provided zooming in the range of 1 to 500 households. This lower range can be of particular interest to distribution planners, enabling to better understand the effect of diversification when the number of customers is low and $ADMD^n$ is far from convergence. In addition, the values of $ADMD^n_{max}$ are shown in boxes for n =1, 50, 100, 200, 300, 400 (inset), 500, 1000, 1500, 2000 and 2500 households.

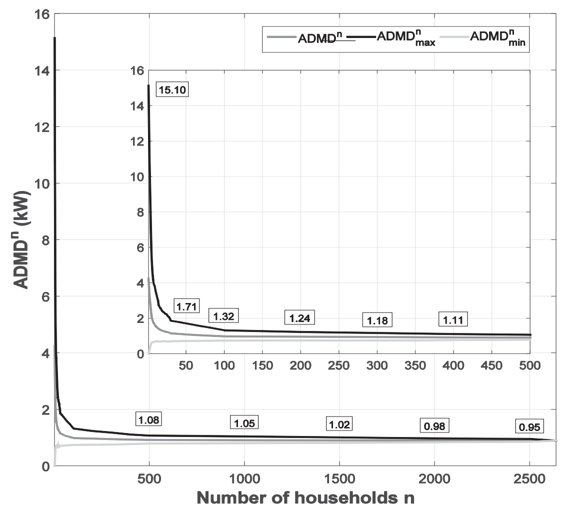

Fig. 4. $\boldsymbol{ADMD^n}$ across all participating households.

D. ADMD analysis across customer categories

In this section we investigate the diversified peak demand behavior as a function of connected customers. We compute $ADMD^n$ across the nine customer categories according to the methodology outlined in Section II and show the results in Fig. 5. As before, three curves are shown to describe the evolution of upper, lower and mean $ADMD^n$ values as a function of households; specific numbers are provided on the plots for n =1, 50, 100 and 200 households. Note that $ADMD^1$ values shown on the plot are same as in Table II. As before, the lower bound is the first to reach a steady state value, followed by the mean and then the upper bound. In addition, it is important to note that convergence of the upper bound to a steady-state value occurs after a few hundred households. The terminal $ADMD^\infty$ values for the nine customer categories are shown in Table III.

TABLE III. $ADMD^{\infty}$ FOR DIFFERENT WEALTH AND OCCUPANCY CLASSES BASED ON SMART-METERING MEASUREMENTS

	1 occupant	**2 occupants**	**3+ occupants**
Adverse	0.51 kW	0.81 kW	1.06 kW
Comfortable	0.57 kW	0.91 kW	1.28 kW
Affluent	0.75 kW	1.07 kW	1.72 kW

Comparing $ADMD^{\infty}$ across different wealth and occupancy levels, it is apparent that the pattern observed in Table II does not persist. Whereas in the case of maximum peak demand, there were instances of *Adverse* and *Affluent* households exhibiting higher peak consumption than their *Adverse* counterparts, it is evident that diversified demand is clearly correlated with customer category. Note that this trend reversal occurs in the range of <50 households. This is an important insight showing that coincident demand across an increased number of households exhibits reduced sensitivity to the attributes of individual consumers (that may vary wildly across a group of households) due to the effect of demand diversification. As a result, we can conclude that wealth and occupancy information can both be useful proxies in inferring diversified demand as long as the number of customers is above the order of a few tens of households. Another important observation is that houses of increasing wealth and occupancy level have a less pronounced diversification effect. This can be inferred from the curve shapes that continue to reduce considerably past the 100 households range, suggesting increased variability of consumption patterns.

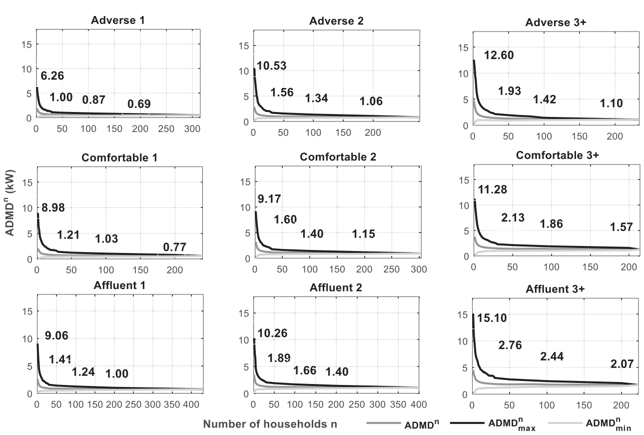

Fig. 5. $\boldsymbol{ADMD^n}$ for all nine customer categories.

IV. Conclusions

The LCL smart meter trial provides a vast number of high-resolution residential demand measurements across 2,639 customers. The analysis of this dataset has enables us to gain valuable insights towards the evolution of residential electricity consumption patterns in the UK and evlauate how they are affected by consumer demongraphic characterisics such as housheold occupancy and welath level. In addition, peak and diversified demand metrics have been analysed and quantified, aiming to inform future distribution planning processes.

The paper first introduces distribution planning practices currently applied in the UK, based on the concept of ADMD, highlighting the practical significance of accurate diversified demand metrics. Subsequently, the LCL dataset is presented and analysed. Major conclusions stemming from this analysis are that the ADMD value calculated across all customers suggests that currently adopted values may be overestimating residential demand while also overestimating the degree with which ADMD converges as a function of connected customers. In addition, household occupancy is shown to be a primary driver of peak demand of individual households. Finally, both household occupancy and wealth level are shown to drive diversified demand, suggesting that there may be merit in novel planning methods that consider customers' demographic makeup.

Further research should propose novel methods for using the information extracted from the smart-meter data to design and operate distribution network in the smart grid era. Potential applications include the assessment of new connections for different mixes of consumers based on the detailed data presented in this paper. In addition, further temporal analysis of the dataset can provide useful insights towards electricity consumption patterns across different calendar seasons, days and hours. The outputs of such an analysis can be used for the optimization of outage management processes to minimize service disruptions.

References

[1] J. Boggis, "Diversity, bias and balance," *Distrib. Elect.*, pp. 357–362, July 1953.

[2] C. Barteczko-Hibbert, After Diversity Maximum Demand (ADMD) Report. Report for the 'Customer-Led Network Revolution' project: Durham University, Feb, 2015.

[3] D. H. O. McQueen, P. R. Hyland, and S. J. Watson, "Monte Carlo simulation of residential electricity demand for forecasting maximum demand on distribution networks," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1685–1689, Aug. 2004.

[4] I. Richardson, M. Thomson, D. Infield, and C. Clifford, "Domestic electricity use: A high-resolution energy demand model," *Energy Buildings*, vol. 42, no. 10, pp. 1878–1886, 2010.

[5] A. Seneviratne, Information – Electrical Design for Distribution Networks : After Diversity Maximum Demand, HPC-3DC-07-0001-2012, 1–12.

[6] K. C. Schneider and R. F. Hoad, "Initial transformer sizing for single-phase residential load," *IEEE Trans. Power Del.*, vol. 7, no. 4, pp. 2074–2081, Oct. 1992.

[7] C. Bary, "Coincidence factor relationships of electric service load characteristics", *Trans AIEE*, vol. 64, September 1945.

[8] C. Ziser, ZY Dong, "Probabilistic Modelling of Demand Diversity and its Relationship with Electricity Market Outcomes", *Proc. IEE PES General Meeting* 2007.

[9] J. Nazarko, R. P. Broadwater, and N. I. Tawalbeh, "Identification of statistical properties of diversity and conversion factors from load research data," in *Proc. 9th Mediterranean Electrotechnical Conf. (MELECON 98),* vol. 1, 1998, pp. 217–220.

[10] Nabeel I. Tawalbeh, A.M. Al-Fahed Nuseirat " Peak and daily profile estimation for substations in rural networks"*Electrical Power and Energy Systems 30 (2008) 202-206.*

[11] R. P. Broadwater, A. Sargent, A. Yarali, H. E. Shaalan, and J. Nazarko, "Estimating substation peaks from load research data," *IEEE Trans. Power Delivery*, vol. 12, pp. 451–456, Jan. 1997.

[12] V. P. Chatlani, D. J. Tylavsky, D. C. Montgomery, and M. Dyer, "Statistical properties of diversity factors for probabilistic loading of distribution transformers," in *Proc. 39th North Amer. Power Symp., 2007*, pp. 555–561.

[13] E.ON, "Network Design Mannual," 2006.

[14] UK Power Networks, Low Carbon London Project. Available at: http://innovation.ukpowernetworks.co.uk/innovation/en/Projects/tier-2-projects/Low-Carbon-London-(LCL)

[15] UK Power Networks, "Use of smart meter information for network planning and operation," Report C1 for the "Low Carbon London" LCNF project, 2014.

978-1-5090-4168-8/16/$31.00 ©2016 IEEE