MATHEMATICAL SOLUTIONS FOR ELECTRICITY NETWORKS IN A LOW CARBON FUTURE

Stephen HABEN
University of Reading – UK
s.a.haben@reading.ac.uk

Matthew ROWE
University of Reading – UK
m.rowe@pgr.reading.ac.uk

Danica Vukadinovic GREETHAM
University of Reading – UK
nv902387@reading.ac.uk

Peter GRINDROD
University of Reading – UK
p.grindrod@reading.ac.uk

William HOLDERBAUM
University of Reading– UK
w.holderbaum@reading.ac.uk

Ben POTTER
University of Reading – UK
b.a.potter@reading.ac.uk

Colin SINGLETON
CountingLab Ltd. – UK
c.p.singleton@reading.ac.uk

ABSTRACT

Smart meters are becoming more ubiquitous as governments aim to reduce the risks to the energy supply as the world moves toward a low carbon economy. The data they provide could create a wealth of information to better understand customer behaviour. However at the household, and even the low voltage (LV) substation level, energy demand is extremely volatile, irregular and noisy compared to the demand at the high voltage (HV) substation level.
Novel analytical methods will be required in order to optimise the use of household level data. In this paper we briefly outline some mathematical techniques which will play a key role in better understanding the customer's behaviour and create solutions for supporting the network at the LV substation level.

INTRODUCTION

With the electrification of transport and heating, domestic electricity demand is expected to increase, threatening the quality and reliability of the future networks, in particular at the low voltage (LV) network level. Conventional reinforcement such as increasing capacity is effective but expensive and other solutions will be necessary.

As part of the UK's low carbon agenda the government aims to have a smart meter in every home across the country by 2019 [9]. The increased amount of information that half hourly smart meters provide will enable Energy Suppliers and Distributed Network Operators (DNOs) to understand and forecast customer behaviour at a greater granularity than previously possible and provide better opportunities to manage and plan the networks.

Many techniques, such as load forecasting, have been developed in the past to predict and analyse the regional and national energy demand (for instance see [1], [2]). However, the data from smart meters, and even the LV substation level is very volatile and irregular and therefore these techniques turn out to be inadequate for analysing and forecasting at the household to LV substation level. New forecasting methods, analysis and validation techniques will be required to better interpret and optimise the use of smart meter data.

The University of Reading is analysing household and LV substation level data as part of the Scottish and Southern Energy Power Distribution led Low Carbon Network Fund project the *New Thames Valley Vision* [1](NTVV). The aim of this paper is to give an introduction to the more important mathematical and control theory approaches which will be necessary to better understand, anticipate and support the current and future demand on the LV networks. In the next section we give a brief overview of the research which is currently being conducted as part of the NTVV, including behavioural based segmentations, forecasts and smart control algorithms for smart storage. Finally we finish with a summary of the work. Data presented in this report comes from the *Energy Demand Research Project* (EDRP)[2] run by Scottish and Southern Energy.

MATHEMATICAL METHODS

Figure 1 compares the half hourly demand of 50 domestic households (solid line) and the UK total gross system demand (TGSD). It is clear that demand, even at the LV level, is very volatile, irregular and noisy and therefore methods developed for High voltage (HV) or even Medium voltage (MV) level may not be appropriate. New methods and new analysis of the LV networks may be required. Segmentation, forecasting and smart control are three very important areas of research which will play a vital role in getting the most out of smart meter data. In this section we consider these methods, their importance, the challenges and some brief results of each.

1 See http://www.thamesvalleyvision.co.uk/
2 See http://www.ofgem.gov.uk/sustainability/edrp/

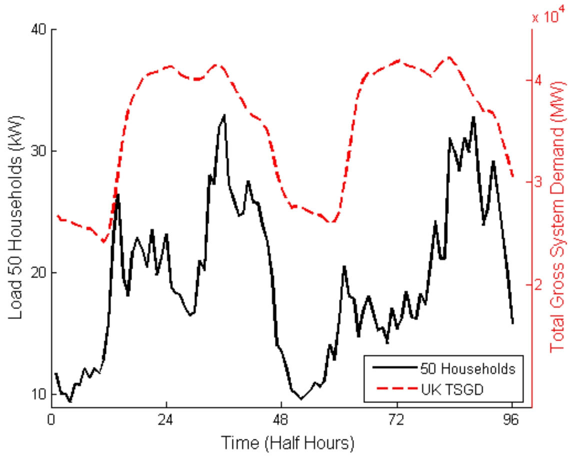

Figure 1: Aggregated demand of 50 households (solid line) and plot of Total Gross System Demand (Dotted line). TGSD data taken from [8].

Clustering

Before any real analysis can be performed on the network it is useful to understand the different types of energy usage behaviour that exist. Clustering can help with the following:

- Planning the networks: If we understand the possible types of behaviour then we can anticipate the networks stresses from new builds.
- Managing the networks: Different combinations of behaviours will have potentially different solutions for managing them.
- Creating bespoke forecasting methods. It is unlikely every customer can be forecasted with the same method.
- Identifying key influencers by comparison of different behaviour types with external factors such as household size and assets.

There are many different clustering methods that exist, but for our purposes they fall into two main categories: Hierarchical and Non-hierarchical. Hierarchical are either agglomerative in which each structure to be clustered begins as its own cluster [2]. Clusters are then combined in an "optimal" way depending on a chosen method (say minimum variance). However a disadvantage is that structures cannot change cluster at a later stage of the algorithm. Once assigned they stay in their assigned cluster even if a more efficient clustering exists. Of the non-hierarchical clustering, k-means is the most popular because of its simplicity [2]. However a disadvantage with non-hierarchical methods is that the number of groups in a clustering must be chosen a priori.

In power systems most attempts at clustering are at HV and MV level of the daily electrical demand profiles. However recently, as smart meter data has become available, attempts have been made to segment the household level daily demand profiles [3]. Distinct profiles can be found that are dependent on the day of the week and seasonality. We created a behavioural segmentation using a Gaussian finite mixture model [6]. The method was applied to half hourly demand data from over 4500 households from 02/03/2009 to 21/02/2010. Three key attributes were identified which contained most of the information for that week:

- Daily average usage over the week
- Variance of daily usage over the week
- An average daily smoothness over the week (given as a lag 1 correlation coefficient).

It was found that 10 descriptive groups could be used to segment the data. The usefulness in clustering on a weekly basis is that the dynamics of behaviour could be investigated. For instance, seasonality could be investigated and the probabilities of moving from one group to another could be calculated. This obviously has implications for the robustness of certain customers and deciding appropriate interventions for reducing usage/peak demand. Interestingly, it was found that 46% of customers stayed within their cluster for more than 60% of the year. This clearly has implications for forecasting in which people obviously have favourite modes of behaviour.

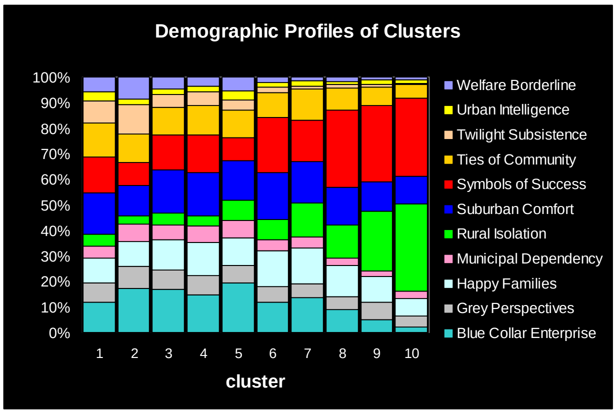

Figure 2: Breakdown of clusters by Mosaic classification.

Another important application of clustering is identifying the most and least important features in the data. For instance, geo-demographics have often been used within energy companies to manage their customers and predict their usage. However it was found that geo-demographics (In this case the 2004 Mosaic[3] classifications which is a popular postcode based segmentation) were a poor fit with our behaviour segmentation. Figure 2 shows the breakdown of each cluster (1-10 in increase average daily energy usage) in terms of the Mosaic classifications. It is clearly visible that, although there are some trends (such as a larger proportion of affluent "Symbols of Success" in the higher energy groups 9 and 10) there is weak correlation between Mosaic group and energy usage behaviour.

3 http://www.experian.co.uk/business-strategies/mosaic-uk.html

Although data led analytics can be useful there are still a number of challenges. The choice of model is far from straight forward and informed decisions need to be made on the number of clusters, the method used, the attributes to cluster and what time scale to consider (hourly, daily, weekly). The choice of these variables will largely depend on balancing between accuracy (by increased model complexity) and practicality.

Forecasting

Forecasting at the household level will be vital for managing and planning current and future LV networks. The increased uptake of low carbon technologies and the higher penetration of micro-generation are likely to increase the strain on the LV network and create a potential two-way flow of electrical energy as consumers become suppliers.

In previous research, forecasts have been applied to the smoother and less volatile high voltage level [1]. They are usual split into short (a day up to a week ahead); medium (up to couple of years ahead) and longer (up to several years ahead) term forecasts. All three times periods will be considered as part of the NTVV project.

Many different methods have been developed and applied including regression, time series (e.g. ARMAX), artificial neural networks, fuzzy logic, similar day methods, Kalman filter, support vector machines and other artificial intelligence methods [1], [5]. However, the research is sparse for household level forecasts. This is expected to change as smart meters become more ubiquitous and as the advantages of forecasting at the LV level become apparent. In the next section we show how short term forecasts are vital for smart control algorithms of storage devices to efficiently reduce peak demand.

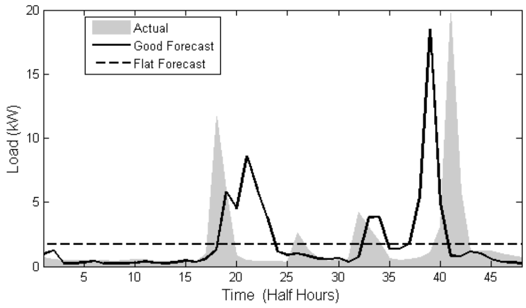

Figure 3: Plot of a poor, flat forecast (dashed) and a good forecast (solid line) together with actual load for a single household.

For household forecasts the size and position of the peaks of demand are most important. To develop accurate forecasts it is necessary to have correct validation and verification methods. Unfortunately standard point-wise metrics (such as root mean square error (RMSE)) are inappropriate for assessing the accuracy of forecasts. Standard metrics suffer from the 'double penalty' effect since missed peaks are penalised once for missing the peak and then again for predicting a peak at the wrong time. As a result subjectively better forecasts can score worse than poor forecasts. Additionally a forecast which only misses a peak by a small displacement in time may still have useful application in smart control algorithms (see next section). Figure 3 illustrates 2 forecasts to the actual (shaded). In this example the poor, flat forecast (dashed) scores a 2-norm error (RMSE) error of 22.8kW compared to the simple, better forecast (solid line) with error 31.2kW (See appendix in [4]). Early work has been to develop a measure which allows the subjectively better forecasts to have improved scores by allowing small restricted permutations in time (for more details see [4]). The natural irregularity and movement of typical peaks will need to be taken into account when creating forecasts of household level demand.

The demand on the network varies greatly over the year due to seasonality and trend effects. Medium term forecasts will help to identify network stress points throughout the year. Understanding the potential uptakes of different low carbon technologies will be vital for identifying strains on the network several years into the future. For instance, in a simple simulation of a low voltage network it can be shown that a high uptake (50%) of electric vehicles (EVs) can create a new, overnight, peak demand when the EVs are being charged. Hence both medium and long term, scenario driven forecasts are essential for managing and planning the networks of the future.

Smart Control of Storage

As part of the NTVV project 23 storage devices will be deployed in Bracknell, UK. The devices have several roles including keeping the network within voltage limits, grid frequency regulation, supporting phase balance and reducing LV network peak demand. Here we show some brief results on how smart control algorithms can be combined with storage to reduce peak demand.

It has been found that basic set point control of the storage device does not guarantee reduction in peak demand due to the limited capacity of the battery [7]. However, utilising household level forecasts (see previous section) a more optimal plan can be created for the behaviour of the storage device.

To create the smart control plan for the storage device first the half hourly forecasts for individual customers must be aggregated. Next, since the data is highly volatile an algorithm is used to reduce the effect of errors in the forecasts and gives the final plan for discharging and charging the storage device. If real time data is available a model predictive controller can be used to update the plan and reduce the errors between the forecasted and actual demand.

Figure 3 shows the effect of the control algorithm using a 25kWh storage device applied to an aggregation of 23 domestic customers. Using a day ahead forecast (see appendix in of [4]) a reduction of 18% is achieved of the aggregated peak demand with no real-time data available. This compares with a potential peak reduction of 30% given a perfect forecast. On 4 other days a reduction in peak demand of at least 10% is possible without real time data. This example shows that a peak demand reduction on the LV network is still possible even with reduced monitoring.

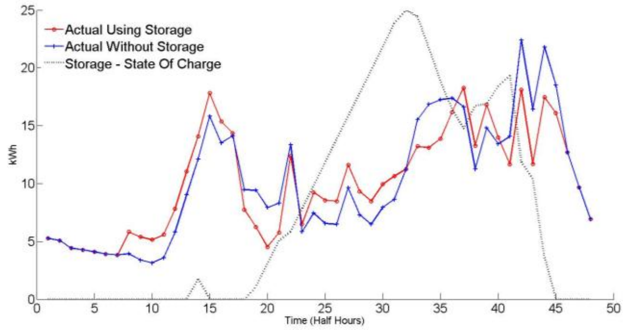

Fig 3: Simulation Results showing the day with and without storage and the SOC of the storage device.

Summary

The increased uptake of low carbon technologies as the UK moves towards a low carbon future will require that DNOs have a greater understanding of demand at the household to low voltage (LV) substation level. In particular, without understanding the stresses on the LV network, highly expensive interventions such as replacing assets may be required.

The anticipated roll-out of smart meters will provide an opportunity for more accurate analysis and greater visibility of demand on the network. At the LV substation level energy usage is extremely volatile, irregular and noisy. Conventional methods, used to forecast, categorise and control the smooth and regular regional and national demand level may not be appropriate at the LV substation level and therefore novel data analytics are being developed by the authors as part of the NTVV project.

In this short paper a brief introduction has been given into the possible mathematical approaches that are available in better understanding customer's energy behaviour through categorisation, anticipating changes in behaviour by forecasting and helping to support the network with smart control algorithms. This is by no means an exhaustive list and the authors encourage researchers to consider these and new methods as the complications of a low carbon future become realised. Energy usage is changing and new methods must be developed to ensure the security of the grid and the smooth transition for customers.

REFERENCES

[1] H.K. Alfares and M. Nazeeruddin, 2002, "Electric load forecasting: literature survey and classification of methods", *Int. J. of Systems Science*, vol. 33, 23-34.

[2] G. Chicco, R. Napoli and F. Piglione, 2006, "Comparisons Among Clustering Techniques for Electricity Customer Classification", *IEEE Trans. Power Sys.*, vol. 21, 933-940.

[3] C. Flath, D. Nicolay, T. Conte, C. Van Dinther and L. Filipova Neumann, 2012, "Cluster Analysis of Smart Metering Data – An Implementation in Practice", *Bus. & Info. Sys. Eng.*, vol. 4, 31-39.

[4] S. Haben, J. Ward, D. Vukadinovic-Greetham, P. Grindrod and C. Singleton, 2012, "A New Error Measure for Forecasts of Household-level, High Resolution Electrical Energy Consumption", *University of Reading Preprint Series.* Available at http://www.reading.ac.uk/web/FILES/maths/preprint_12_14_Haben_6Sept.pdf

[5] H. Hahn, S. Meye-Nieberg and S. Pickl, 2009, "Electric load forecasting methods: Tools for decision making", *European. J. Operational Research*, vol. 199, 902-907.

[6] G. McLachlan and D. Peel, 2000, *Finite Mixture Models*, Wiley.

[7] P.R. Thomas, *American Electric Power's Community Energy Storage*, EPRI 4th Int. Conf. Integration of Renewable and Distributed Energy Resources, 2010.

[8] Total Gross System Demand data taken from National grid. Available at (last accessed Jan 2013) : http://www.nationalgrid.com/uk/Electricity/Data/

[9] *Smart Meters Programme Plan*, Department of Energy and Climate Change, 2012

Acknowledgments

We wish to thank Scottish and Southern Energy Power Distribution (SSEPD) for their support and funding via the NTVV project (SSET203- New Thames Valley Vision) – funded through the Low Carbon Network Fund. We also wish to thank Scottish and Southern Energy for providing the EDRP data for use in this work.