

## Smarter Business processes resulting from Smart Data

Goudarz POURSHARIF  
University of Sheffield – UK  
g.poursharif@sheffield.ac.uk

Andrew BRINT  
University of Sheffield – UK  
a.brint@sheffield.ac.uk

John HOLLIDAY  
University of Sheffield – UK  
j.d.holliday@sheffield.ac.uk

Mary BLACK  
Northern Powergrid – UK  
Mary.Black@northernpowergrid.com

Mark MARSHALL  
Northern Powergrid – UK  
Mark.Marshall@northernpowergrid.com

### ABSTRACT

Potentially by 2020, a wealth of new knowledge of LV networks will be available to Distribution Network Operators (DNOs) through the deployment of Smart Meters and LV network monitors. Integration of these new sources of Smart Data into the business processes of DNOs will be a major challenge. This paper presents a model which utilizes the combination of smart meter data at varying granularities and real time network monitor data to improve distribution network processes and applications such as Load Allocation (LA), network planning, and fault management.

### INTRODUCTION

The introduction of Smart Meters and low carbon technologies into distribution networks is radically transforming the types of data that is available on Low Voltage (LV) networks. As these technological advances become widespread, a wealth of new data will be available to Distribution Network Operators (DNOs), whose business applications and processes have traditionally been mainly reliant on the one-way flow of energy and information. These new kinds of Smart Data could vary in terms of availability, frequency, and functionality. Therefore, network operators urgently need to understand the benefits that can be derived from the Smart Data in order to enhance their existing processes as they make strides towards a smarter grid operation (Figure 1). The challenge is how to sort, combine, utilize, visualize, present, and integrate the Smart Data to enhance the workflows within the DNOs applications such as network planning. This paper describes the major contributions that could result from combining real-time information from monitors at the start and mid-way along one of the UK DNOs, Northern Powergrid's low voltage feeders, with aggregated smart meter data and knowledge of customer types and connection points.

In this paper, particular attention is given to home Smart Meter Data and major benefits that can be drawn from having the customer load data at various granularities ranging from half hourly to two-hourly averages. Since it is highly likely that the network operators in the UK will be provided with the customer Smart Data in aggregated and anonymized formats due to costs, unavailability and

privacy concerns, it is of major significance to investigate the extent to which major network operation applications are affected by these changes in the way in which real time data are fed to them. Customer Smart Meter Data and smart network monitoring data can be combined to provide better current estimates on the LV network which can be used both as a strategy tool to allow cost effective examination of new ideas, e.g. examining correlation patterns, and as a filter to highlight potentially problematic network areas.

Another process that could benefit is the planning of network outages – increased confidence from better information about network configuration and customer phasing can benefit network operators and customers by reducing the scale of outages.

Load Allocation (LA) methods, as ways of predicting customers' load in absence of real-time data, are also investigated to identify the ways in which they can be improved by varying granularities of the customer load data.

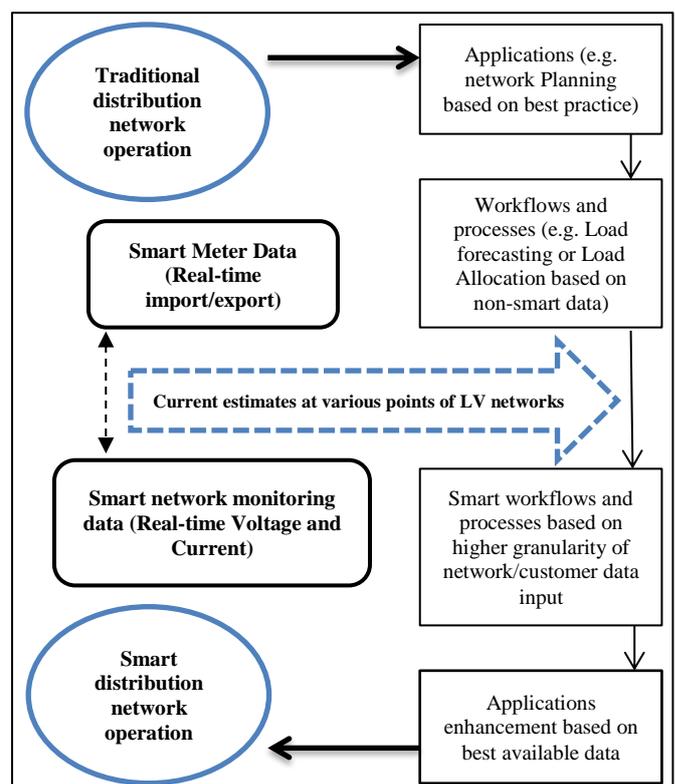


Figure 1: Integration of Smart Data and distribution network operation

## PREVIOUS LOAD ALLOCATION (LA) STUDIES

State Estimation (SE) algorithms rely on customer load data which are not normally available in real time on electricity distribution networks, hence state estimators require load models which are produced by methods generally known as Load Allocation (LA) [1]. Research has been focused on methods that can provide load demands at node points, mainly on MV distribution network scales and efforts have been made on moving away from traditional estimation of loads based only on peak loading conditions [2]. The more realistic the load modelling technique is, the more accurate are the results of SE and other advanced applications fed by SE results, which include estimation of power losses, power and reactive power and can also contribute to voltage optimization, voltage and reactive power control, feeder reconfiguration, and demand-side management [1-4]. Some have pointed out that even processes such as power flows, fault detection, and service restoration, benefit from data produced by LA methods [3]. The LA Methods have been improving to take into account various factors other than only peak loads based on transformer kVA and customers' peak consumption or predefined demand tables and voltage drops [1 and 3]. Over the past 20 years researchers have focused on improving the LA methods by incorporating more variables such as customers' monthly consumption values, customer types, time of day, weather conditions, ratio factors in relation to customer types, customer class curves, Customer type Load Modeling Factors (LMF) [1, 4, 5, 6]. From the turn of the century, investments were made by utility companies to gain more knowledge about the customers, which was not essential in the past. Since the introduction of Advanced Metering Infrastructure (AMI), it is evident that research has been shifted towards using Customers' load readings in LA models and consequently in geospatial models of the distribution network as well as modern Distribution Management Systems (DMS) used by DNOs [3, 5, 7, 8]. Some have highlighted that the traditional kVA method of LA can cause overestimation of demand and there is no direct relationship between the transformer rating and the number of consumers served by it [5]. In one study, results of four methods of LA are compared to the values from AMI to determine which method best reflects the true values obtained from AMI readings [5]. The four methods investigated are Daily kWh, Monthly, kWh, transformer kVA and REA, which are proved to be less accurate compared to using maximum diversified demands of the transformers obtained from 15 minute AMI data [5]. On the other hand, [8] point out that with the widespread presence of AMI in the future, different load shapes can be drawn for different customers and that information can be fed into network modelling and simulation programs. Although [8] argue that aggregated AMI load measurement can significantly improve the processes and State Estimation

(SE) studies of distribution network, LV networks may potentially be more susceptible to demand variations in comparison to MV networks, which have been the subject of most previous LA studies.

## DATA STREAMS

Our chosen Low Voltage (LV) network sample is operated by Northern Powergrid and it is equipped with Voltage and Current monitors at the substation and along the feeders. The monitors measure Voltage and Current every minute (Figure 2). However, the most complete data available to us was only for the Month of March 2014.

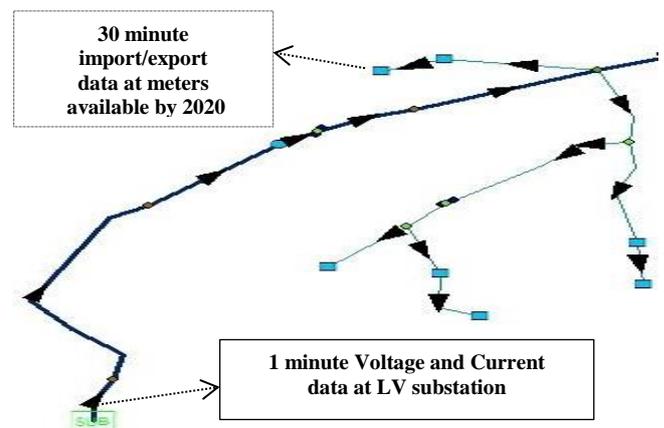


Figure 2: Potential future Smart Data available at LV level

The Smart Meter Data of households on this particular piece of network could not be obtained due to privacy concerns. Therefore, customer load data had to be acquired using data from trials before 2010. For the purpose of this study, home Smart Meter Data were obtained from a previous study carried out by Loughborough University on 22 different types of houses in 2008 and 2009. The datasets contained customer import readings of all 22 houses for March 2008. However, the 2009 data were incomplete for some houses. After initial analysis of the available customer data, a small LV model was built based on the fact that it will be very likely that the smart meter data provided to the DNOs are aggregated at some points along the LV network (e.g. A and B) (Figure 3). The characteristics of the 6 houses used in our model are summarized in table 1.

House number	House type	Number of Residents	Electric Shower
1	Semi-detached	5	Yes
2	Semi-detached	6	Yes
3	Detached	3	No
4	Detached	3	No
5	Terraced	4	No
6	Detached	4	No

Table 1: Characteristics of sample houses in the LV model

For house no.1 the data from 1st March 2009 to 6th March 2009, house no.2 data from 7th to 31st March, and house no.6 the data from 1st March 2009 to 6th March 2009 are missing, but data for houses no.3, no.4, and no.5 are complete for both March 2008 and 2009. In Figure 3, in theory the readings at the substation should be the sum of readings at A and B, i.e.  $C=A+B$ , but to what extent can the usage pattern be predicted by having the figures for the total customer import at the substation based on the figures for the same day in previous weeks or years? It is also of vital importance to understand how the results vary as the granularity of Smart Data are reduced (e.g. from half-hourly to hourly and two-hourly).

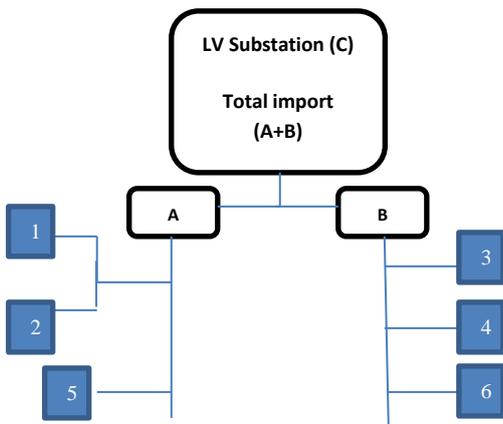


Figure 3: The LV model with data aggregation points (A and B)

### CONSUMPTION PROFILE PATTERNS

The graphs below (Figure 4 and 5) are plotted based on half hourly, hourly, and two-hourly import (kW) averages of the 6 households in the LV model (Figure 3). The final week of March 2008 has been removed, due to it coinciding with Easter holidays in 2008. In the interest of clarity, one working day (Wednesday) has been chosen for presenting the results. The aim of the analysis is to examine to what extent our predictions of unavailable customer import readings on a particular date (e.g. Wednesday 25.03.09) can be determined using previous real time data from the same kind of day on previous weeks and years. It is also significant to examine how the value of the real time data changes as granularity of data decreases, especially in relation to rising levels of embedded generation in LV networks of future.

Figure 4 demonstrates that if the load data for Wednesday 25<sup>th</sup> March were missing the data from the previous Wednesdays in the same month might be useful to the operators in predicting the behavior on the 25<sup>th</sup> of March. At first glance, the load shapes follow similar patterns to some extent. However, the times of day and the values at which demand levels peak are different. Therefore, it is of major importance to identify the best ways in which the value of load shapes acquired from historical Smart Data can be quantified. Figure 5 demonstrates the same load shapes using hourly averages and two-hourly averages of

loads. Higher granularity of data results in smoother graphs with fewer fluctuations in readings, which makes the graphs easier to understand, but will result in details being neglected. Also, it can be observed that peak demand values are reduced when the granularity of the data decreases from half hourly to hourly and two-hourly respectively (Figure 4 and 5).

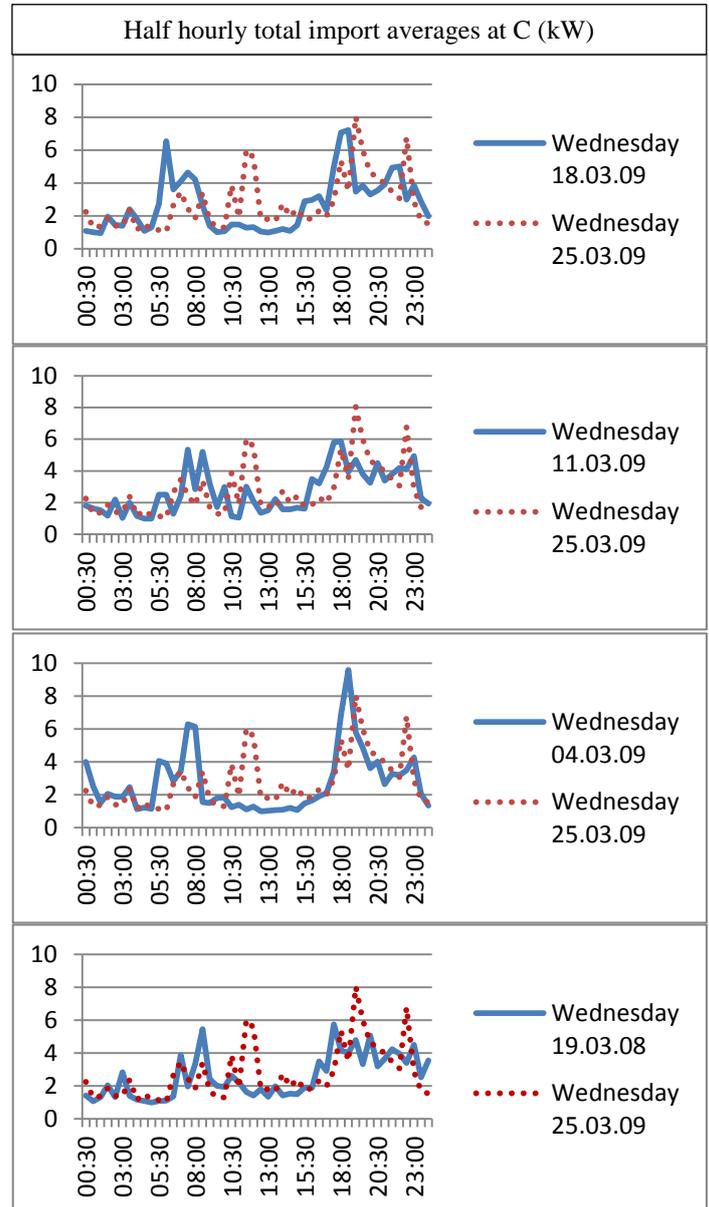


Figure 4: Half hourly import reading comparison of Wednesdays

Considering the previous research in Load Allocation (LA), it is very important to devise a method to evaluate the value of historical smart Data that will be available to the DNOs for load predictions as Smart Meters are rolled out and Smart Data are accumulated. Real-time and historical Smart Meter Data combined with Real-time Voltage and Current readings at the LV substation and feeders can provide higher visibility levels of LV networks to the operators.

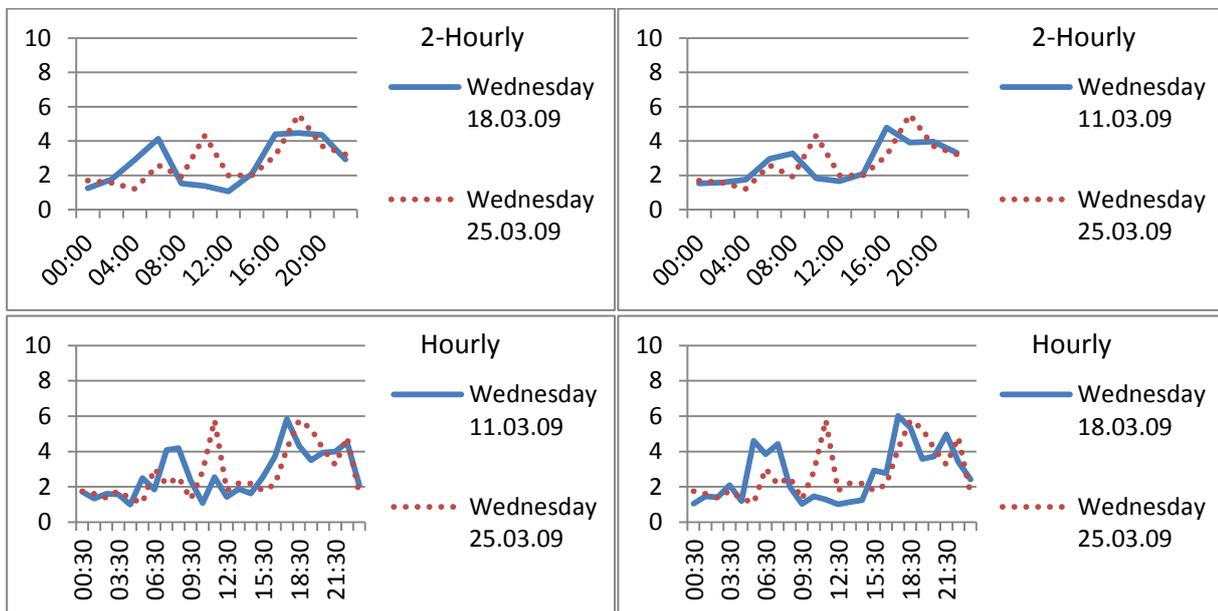


Figure 5: Total import averages (kW) at C with Hourly and 2Hourly granularities

## SMART METER LOAD ALLOCATION

The goal is to calculate the best estimate of the ratio of A and B's loads, given the real time load at C in Figure 3 along with the historical loads at A and B. This comes down to determining the closest analogues in the historical data to the present one. We use four "distance" measures to assess how similar a previous time was to the current one:

- The load at C
- The day of the week
- The number of weeks apart, mod 26.
- The time separation, mod 12 hours.

These differences were combined into an overall measure:

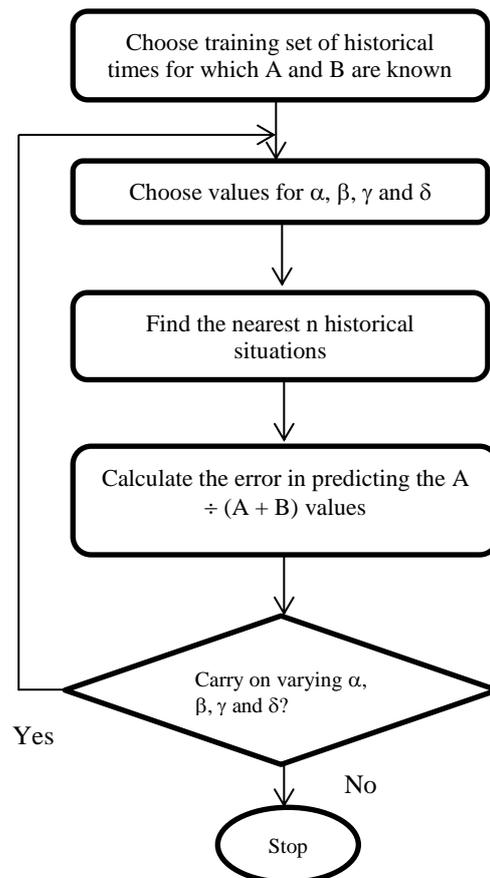
$$\text{load}^\alpha \times \text{day}^\beta \times \text{weeks}^\gamma \times \text{time}^\delta$$

Where  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  are constant parameters. The algorithm for choosing the best values of  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  is shown in Figure 6.

The number of years apart is another important distance measure but this was not employed as the data set was mainly restricted to a 12 month period.

Once the closest historical analogues have been identified, then a weighted average of  $A \div (A + B)$  for these analogues is formed. The weights are inversely proportional to the "distance" measure between the historical data and the current situation.

The approach was applied to the situation in Figure 3 and the predicted and actual load fractions on branch A are shown in Figure 7 for 26 test times. This approach outperformed load allocations using the annual peaks on branches A and B, and the average loads on the branches.


 Figure 6: Setting the values of  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$

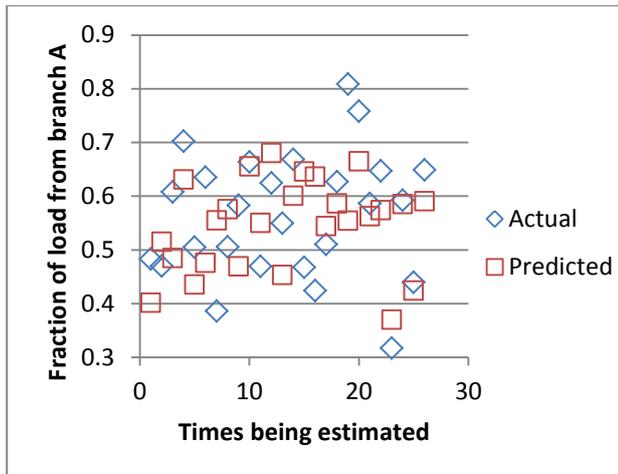


Figure 7: Predicted and actual values for  $A \div (A + B)$

These latter methods predict constant ratios for  $A \div (A + B)$  while using Smart Meter data allows the predicted ratio to vary at different times as shown in Figure 7. This led to 17% more accurate predictions than just basing the predicted ratio on the average load. Whether this performance difference will increase or decrease as the number of customers on a branch increases, needs to be investigated.

## CONCLUSIONS AND FURTHER WORK

In theory, the deployment of Smart Meters and the investment made by utility companies in network monitors at LV level will provide near real time Smart Data to the network operators. However, it is highly likely that the data will not be perfect and will be made available to DNOs in aggregated and anonymised formats with variations in terms of availability, frequency, and functionality. Our model demonstrates the ways in which DNOs can become more proactive in predicting customer loads and consumption patterns by having aggregated historical Smart Meter Data.

Our analysis demonstrates that although load shapes of the same days of the week in a specific month of the year to some extent follow the same pattern, the peak times and values can vary. Additional analysis also shows that as the granularity of the Smart Data decreases from half hourly to hourly and two-hourly averages, small load variations are lost and peak load values drop, which can be significant in a network containing embedded generation. Our paper also presents a model to evaluate the value of aggregated historical Smart Data in projecting the estimates of the ratio of the loads on different branches of an LV network.

The synthesised network model represents a simple network with a limited number of houses without any installed embedded generation.

This work will continue by expanding and developing the LV network further to match the sample Northern Powergrid LV network shown in Figure 2. Solar PV generation data will also be added to the analysis and the result will be plotted based on a typical future LV network containing embedded generation. As the penetration level of embedded generation in the LV network rises, the need for real time and geographically accurate data also increases. Therefore, the results of the analysis will be geospatially visualized and represented in meaningful ways which are relevant to the requirements of various distribution network applications such as network planning, asset management, and fault management, paving the way towards more proactive distribution grid operation.

## REFERENCES

- [1] A.K. Ghosh, D.L. Lubkeman, and R.H. Jones, 1997, "Load modelling for distribution circuit state estimation", *Transactions on Power Delivery*, IEEE vol.12, 999-1005
- [2] A.K. Ghosh, D.L. Lubkeman, M.J. Downey, and R.H. Jones, 1996, "Distribution circuit state estimation using a probabilistic approach", *Transactions on Power Systems*, IEEE, vol.12, 45-51
- [3] C. Carmona, E. Romero-Ramos, J. Riquelme, and A. Gomez-Exposito, 2007, "Distribution transformer load allocation from substation measurements and load patterns", *Innovative Smart Grid Technologies Europe*, IEEE, 1-8
- [4] R. Sharifian, E. Jafari, M. Rahimi, P. Ghaebi Panah, 2012, "State estimation for load allocation in distribution power systems", *IJTPE Journal*, vol.4, Issue 10, 86-90
- [5] W.H. Kersting and W.H. Philips, 2008, "Load allocation based upon automatic meter readings", *Transmission and distribution conference and exposition*, IEEE, 1-8
- [6] Y.C. Lee and M. Etezadi-Amoli, 1993, "An improved modelling technique for distribution feeders with incomplete information", *Transactions on Power Delivery*, IEEE, vol.8, 1966-1972
- [7] R. Brakken, T. Ait-Laoussine, S.J. Steffel, and D. Sikes, 2006, "Integrating customer load and GIS data for improve distribution planning and operations", *Transmission and distribution conference and exhibition*, IEEE PES, 8-14
- [8] R.F. Arritt, R.C. Dugan, and R.W. Uluski, 2012, "Investigating load estimation methods with the use of AMI metering for distribution system analysis", *Rural electric power conference*, IEEE, B3, 1-9