Executive Summary of Appendices

Attached as separate files are the following appendices, each referenced in the text.

Appendix #	Subject	Reference in proforma
Appendix 1	**University of Bath letter of engagement and support, and their data analysis methodology**	**Boxes 1, 18**
Appendix 2	Memorandum of Understanding between WPD and npower	Box 25
Appendix 3	Customer communications pack	Box 2
Appendix 4	Installing monitoring in LV substations	Box 4
Appendix 5	Support letters – University of Bristol / WAG	Boxes 18, 23
Appendix 6	Sample extract of Arbed data (anonymised)	Box 23
Appendix 7	RFQ issued to meter installers re the installation of voltage sensors at end of feeders / Provisional design of end-of-feeder voltage monitors	Box 25
Appendix 8	An overview on the concept of network templates	Box 1

Appendix 1: University of Bath letter of engagement and support, and their data analysis methodology

UNIVERSITY OF BATH

Professor Jane Millar
Pro-Vice-Chancellor (Research)
University of Bath
Claverton Down
Bath
BA2 7AY
United Kingdom
Tel. +44 (0)1225 386141
Fax +44 (0)1225 386626
Email: J.I.Millar@bath.ac.uk

10th August 2010

RE: Western Power Distribution Low Carbon Networks Fund Tier 2 Project – South Wales 2010

Dear Mr West

I would like to thank you for your invitation to participate in your LCNF project. The proposed research is of significant interest to us for a number of reasons. The University is strongly committed to working with industry partners to enhance the outreach and socioeconomic impact of its research as well as generating high quality publications. The cross-faculty nature of the work we are proposing to undertake with you strongly aligns with Bath's research strategy promoting interdisciplinary research. The area of electricity networks and future energy supply are key research areas within the University's Institute for Sustainable Energy and the Environment (I-SEE)

Having discussed the project with Dr Li and Dr Shaddick, I understand that the overall project will produce a large and comprehensive data set regarding network utilisation, electricity demand, voltage and load curves. Access to real-time network data such as this is very important in developing and validating models and using in undergraduate and postgraduate teaching to help expose engineers to real-world information needs and problems. I understand from our academics that there are a number of interesting research avenues that could be pursued in parallel with, and complementary to, your LCNF project. Consequently, if your proposal is successful, the University of Bath will contribute 50% funding (value **£56,000** at full economic cost) towards a 3 year University Research Studentship to develop a complementary interdisciplinary research project with Western Power Distribution utilising the data set being gathered.

I wish you every success with your application and look forward to the University further developing our strong research links with Western Power Distribution.

Yours sincerely,

Jane Millar

Professor Jane Millar
Pro-Vice-Chancellor Research

The below was written by Dr Li and Dr Shaddick of Bath University. It establishes what data they expect to receive from WPD's project, their data analysis methodology, and the learning and knowledge dissemination activities.

The major challenges facing the DNOs network development and operation are the increase in intermittent low carbon generation sources and the expected increase in demand for electricity as a result of transport and heat electrification. Work done by others, notably the DECC / Ofgem ENSG and Imperial College has pointed to the role of demand side management and voltage optimization in partial mitigation of the network impacts. There are however, significant gaps at present in the understanding of the ability of different combinations of network topographies and customer types to accommodate these new low carbon stresses.

WPD will deploy monitoring equipment at some 1050 HV/ LV substations and over 7,000 ends of LV feeders collecting network performance data. This monitoring will provide detailed information about the thermal and voltage performance and time of day margins of specific circuits and transformers. This information will better inform the planning and operation of a smarter network and identify areas where distributed generation (DG), electric vehicles (EVs), heat pumps (HPs), smart appliances (SAs) and other low carbon devices can be connected without triggering reinforcement. WPD has the potential to collect performance data both before and after the connection of the specific generation and responsive demand. It is expected that the data collection will be on a half-hour interval but can be selected to be as short as 10 seconds if required. It is also understood that demand side response could play two different roles that could shape demand into very different patterns in a low carbon electricity system, one is for power balancing to better integrate intermittent generation, the other is for peak shaving.

The overall aim of the research will be to develop a number of common network templates and be able to classify each substation to a template with a statistical confidence. If statistical confidence for a substation classification is low, we will either develop an alternative template for the substation or indicate the necessity for on-line monitoring, depending on the nature of the statistical analysis results. Based on these templates, we will evaluate the available "headroom" in terms of thermal and voltage constraints, time of day and other network stresses as to whether further demand and DG capacity can be accommodated into these. The academic team at Bath will interrogate the extensive data using statistical approaches with an aim to: i) inform critical conditions in networks in relation to distributed generation and demand side responses (DSRs), allowing DNOs to have targeted interventions as required to maintain their quality of supply, ii) infer conclusions as to which of the two DSR roles is most beneficial to maintain an intact network.. Our statistical interrogation will develop new knowledge in three broad areas:

1) Basic templates to represent customer class and seasonal variations with different uptake levels of new generation and demand without DSR
    a. Classify and develop templates that give a more accurate picture of demand profiles
    b. Understand how templates could accommodate demand shift with different penetration levels of distributed generation, EVs (electric vehicles), HPs (heat pumps) and SAs (smart appliances) and classify and develop templates based on the sampled and inferred information
    c. Understand if there is the potential that the demand peak with DSR for balancing is higher than that under DSR for peak shaving

2) Templates for peak utilization of intermittent generation
    a. Understand the value of DSRs for power balancing, quantify the degree of utilisation of intermittent resources at the district, regional and remote areas from the three levels.
    b. If we are to unlock the full potential of DSR for power balancing, understand how demand might shift with different penetration levels of EVs (electric vehicles), HPs (heat pumps) and SAs (smart appliances), classify and develop templates based on the sampled and inferred information

3) Templates for peak shaving
    a. Understand the constraints in the present network that restrict the power balancing potential from DSRs
    b. Understand the full potential of DSRs in peak shaving and estimate how the basic templates might change with different penetration levels of EVs (electric vehicles), HPs (heat pumps) and SAs (smart appliances)

Creating templates

The analysis of the data will consist of two parts, starting with an initial data exploration in which possible structures in the series of data will be explored using time series techniques. Loads might be expected to exhibit patterns at different temporal resolutions, e.g. seasons, days and within-days. Of particular interest will be the identification of regular patterns in the data, for example peaks with days and longer-term seasonal patterns.

After this first step, 'cluster analysis' techniques will be used in order to group series of data, for example time series of loads from substations, into groups of substations that exhibit similar patterns at a chosen temporal resolution. These groups, or 'clusters' will form the basis of the templates that will be developed based on the data and will later be compared to those templates hypothesised by WPD.

Cluster analysis

Taking for example, daily patterns from the substations as the object of interest, the aim is to create clusters of substations within which the daily patterns are more similar than those in other clusters. In statistical terms, this means that the inter (or between)-cluster variation should dominate the intra (or within)-cluster variation.

The proposed method for this analysis is that of agglomerative hierarchical clustering. This works by grouping data objects, for example time series of loads within a day, into a tree of clusters. There are two possible approaches, depending on whether building the clusters is performed using a top-down or bottom-up strategy. In the former, initially all objects are considered to be in a single cluster and this is then split into smaller clusters iteratively based on measures of dissimilarity until, ultimately, each object forms its own cluster. In the latter, the initial set-up sees each object being classified as its own cluster and clusters then merged, according to some similarity measure, until ultimately there is again a single cluster. With larger datasets, such as those being considered here, it is often computationally more efficient to use the bottom-up approach.

Two crucial aspects of this approach are (i) the choice of dissimilarity measure, i.e. the measure of how close two objects are and (ii) the criteria which decides whether, based on this measure, two clusters are considered close enough to be merged together to create single cluster.

The precise nature of the dissimilarity measure will be determined to a large extent by the output in which in interest lies. For example, if the aim was to identify clusters according to peak usage then the measure might be based on differences between daily peaks. If very short-term trends were of primary interest, then the measure could be constructed to filter out longer-term patterns. The measures can also be constructed to reduce the effects of outliers and noise to a greater or lesser extent depending of the requirements of the analysis.

Starting with M clusters (the number of individual substations) each cluster would contain series of data from each substation. This might be, for example, data from time points within a single day, alternatively it could comprise data over a longer time period or multiple sets of data from smaller time periods (although it is noted that with large sets of data this may be computationally expensive). When the dissimilarity measure has been calculated between all clusters, the two clusters that are most similar are merged, leaving M-1 clusters. This is repeated until a there is a single cluster and the result is a binary tree of 2M-1 clusters, which can be represented in a ‘dendogram’.

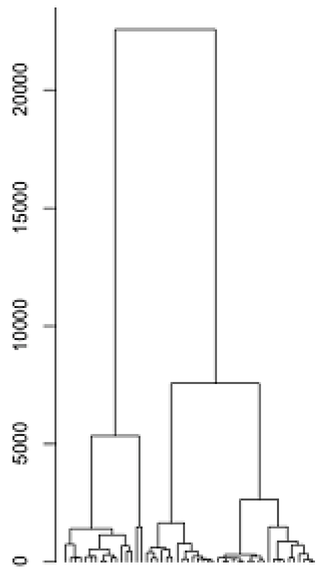

Figure: Example of a dendogram

The appropriate number of clusters can be assessed by the length of the vertical lines between mergers, with longer lines indicate that clusters that are less similar are being merged. In the example shown, it appears that there is a long break between the joining of the four clusters at about 5000 (on the y-axis, which represents the dissimilarity measure) compared with that when smaller clusters lower down the dendogram are merged.

Constructing customer templates

It is anticipated that certain clusters (of substations) might be associated with certain characteristics of the customer profiles associated with those stations. When a set of clusters have been chosen, summaries of such information together with the data used to produce the clusters themselves, will enable descriptions of the clusters to be developed. The trade-off between the number of clusters and the distinction between them can often determine how easy it is to allocate sensible descriptions to the resulting clusters. For example, a larger number of clusters may mean that a number of them are similar whilst a smaller number may mean that there is too much heterogeneity within the clusters to accurately describe them, i.e. there may still be different groups within the clusters.

Comparison with WPD templates

WPD have compiled a list of possible templates for the substations without reference to the data to be used in this analysis. It will be possible to construct a similarity matrix between these hypothesised templates and those obtained from the cluster analysis from which an assessment of how close the two schemes are in classifying the substations can be made. The objective being to assess whether or not a DNO, using data already in its possession, could identify, on a statistically sound basis, particular substations or groups of substations as fitting a particular "template" without the need to undertake the detailed monitoring employed to provide the input to this project

Assessing uncertainties

The approach described assumes to some extent that the final allocation of objects to clusters is correct. More realistically, there may be a probability that a particular object is a member of a number of different clusters, with the cluster being associated with the highest probability being the most likely. By performing the clustering within a Bayesian framework it is possible, instead of allocating each object to a single cluster, to obtain estimates of a 'posterior' probability distribution which indicates the likelihood that it belongs to each of the candidate clusters. The term posterior in this case refers to the fact that in the Bayesian statistical framework, 'prior' information is specified on the distribution of all parameters that is then combined with the observed data to produce 'posterior' probability distributions. In this case, for each object prior probabilities are assigned to its membership of each cluster (which might be equal for each cluster or informed by prior knowledge) which are then adjusted by patterns observed in the data to produce the posterior probabilities. It should be noted that due to the intractable nature of much of the mathematics required to implement such an approach, inference often has to be performed using simulation techniques. In Markov Chain Monte Carlo MCMC) simulation, samples are repeatedly drawn from combinations of distributions for the priors and the data in order to build up an accurate representation of the posterior distributions. Often such simulations have to be repeated a very large number of times in order to ensure convergence to the correct posterior distributions is achieved which can be very computationally expensive, especially when dealing with large datasets as in this case.

Dissemination

It is Bath's understanding that WPD would be keen on disseminating and publicising the research outputs and outcomes of the work being undertaken. Research impact is a key driver for all universities and Bath would seek to jointly publicise this work with WPD. There are a number of complementary routes that the University would seek to undertake.

Academic publications – The research team would seek to co-author academic journal papers and target these at the leading power engineering and statistical journals e.g. Econometrica, IEEE journals.

Dissemination conferences – The University proposes to host a conference at Bath (or London or Cardiff) to disseminate the outcomes of the project and other relevant research / case studies to an invited audience containing regulators, policymakers, Government, industry and academic groups. The timing of this event would be agreed with WPD.

Conference attendance – It is expected that members of the research team would attend at least one conference in the duration of the project to disseminate the project outputs. The nature of this particular project provides opportunities to publish both within the power engineering and statistics field increasing dissemination.

Press articles – The University would seek to develop articles for a more general audience for publication on its website, through its Research@Bath publication and possibly in local press.

Teaching and events – The University strongly engages with the local community around its research projects through for example its annual Bath Taps into Science Festival, Science Café talks and lectures at the Bath Royal Literary & Scientific Institute. The outputs of the research may also provide extremely valuable teaching materials for power engineers and statisticians alike.

Linked projects – The wealth of the data set being produced means that there are a significant number of statistical interrogations and linked demographic investigations that could be undertaken to complement the main project. The investigation of demand profiling in a low carbon system is critical to inform future research in the emerging space of smart(er) grid. In particularly, Bath is leading international efforts in developing smart pricing for a smart grid and is active in the planning and operation of smart grid and the optimal balance between demand response for power balancing and peak shaving.