Online AMR Domestic Load Profile Characteristic Change Monitor to Support Ancillary Demand Services

Bruce Stephen, *Member IEEE*, Fridrik Isleifsson, Stuart Galloway, Graeme Burt, *Member IEEE* and Henrik W. Bindner

Abstract— With conventional generation capacity being constrained on environmental grounds and renewable alternatives carrying capacity uncertainties, increasingly accurate forecasts of demand are likely to be required in future power systems: highly distributed renewable generation penetrating low voltage networks must be matched to small dynamic loads, while spinning reserves of conventional generation that are required to maintain security of supply, must be reduced to more efficient margins. Domestic loads, likely to form significant proportions of the loads on islanded power systems such as those in remote rural communities, are currently modeled with homogenous and coarse load profiles developed from aggregated data. An objective of AMR deployment is to clarify the nature and variability of the residential LV customer. In this paper, an algorithm for tracking the consistency of the behavior of small loads is presented. This would allow them to be assessed for their availability to provide demand services to the grid. In the method presented, significant changes in behavior are detected using Bayesian changepoint analysis which tracks a multivariate Gaussian representation of a residential load profile on a day to day basis. A hypothetical single phase feeder, representative of an islanded rural power system, is used to illustrate the detected heterogeneity of load behavior consistency.

Index Terms— Automatic meter reading (AMR), LV Network, Demand Characterization, Bayesian Statistics

I. INTRODUCTION

DOMESTIC loads can be subject to environmental and lifestyle induced variability which can make modeling loads on the LV network difficult. The currently used assumption of multiples of a homogenous load profile on a LV feeder can result in spare capacity being wrongly estimated – for example, [1] notes that the base load for identical domestic properties can vary by as much as 300-400%. Current practice is to generate load profiles at the macroscopic or national level from exemplar or trial data and employ averaging. For example, the UK uses a profile class system, a set of 9 exemplar load profiles, of which domestic customers comprise the first two [2]. Rather than continue to assume a single,

highly averaged load profile, the introduction of Automated Meter Reading (AMR) has the potential to permit load profiles to be generated dynamically from recorded load. The availability of higher resolution load data through AMR deployment opens up the possibility of data driven profiles to be created according to typical customer behavior which can be updated as that behavior changes. Small power systems such as islanded low voltage networks [3, 4] will have smaller numbers of loads to serve lacking the population size to average out anomalies and behavioral artifacts [5] in energy use as can be done at the national scale. Although AMR provides the means to measure this, detecting such a change in the presence of variability in daily routine presents a challenge both to the choice of load profile representation and the generalization capabilities encoded within it.

A model to assess the consistency of residential loads is presented in this paper as a means of assessing availability to provide ancillary demand services. Characterizing predictability of behavior allows lump load availability to be better assessed for demand or frequency response purposes. This could be of benefit either in an islanded power system or on a smaller scale, a single phase rural feeder with high renewable penetrations where upgrades are not feasible for economic or access reasons. Notable practical examples are given in [6, 7] where economic necessity drove innovative solutions to addressing security of supply. In other cases, it is renewable obligation that drives the need for network investment; [8] notes that in a German case study, three main issues stemmed from increased PV installation: the reverse power flow scenario, the additional power flows that may result and the voltage and frequency control issues (that can result from imbalance generation and load). Automated or online tap changers (OLTC) can deal with the bidirectional power flows that result from demand failing to absorb PV generation, but as [9] notes, the least costly route would be – in planning upgrades to a power system with increasing penetrations of renewable generation such as PV, costs can be lowered if there is self consumption; correct identification of the likelihood of this would prevent needless expenditure. Assessing the potential characteristics of a given load has been difficult up until now owing to the dearth of metering on LV networks; AMR or Smart Metering has been seen as the key technology in addressing this [10], but these are specified primarily for retail business purposes and owing to the

granularity of AMR data, at best, 30 minute advances (the amount of energy by which the meter reading 'advances' over the given period), the ancillary service provision can only be addressed for supplemental reserve and load shaping. To be able to plan on the invocation of ancillary services from anticipation of demand, it is desirable to have an indication of the consistency of either consumers with a particular load profile or consumers in a specific part of the network. Frequency excursions could, for example, be addressed by suitably sized loads [11] and the indication of the availability of such loads to do this would be valuable for the reliable operation of the power system. Section II reviews the ancillary services that may be available to a network operator at the LV level. Section III reviews previous work on modeling disaggregated LV network load profiles from AMR data in a way that encodes their variability and expected values throughout a day; Section IV reviews Bayesian change point analysis, a technique for incrementally building statistical models from observed data and detecting when observational changes necessitate the creation of a new rather than an update of an existing, model. Section V shows how AMR data is used to profile individual residential customers and how this model is embedded into the changepoint detector. Section VI gives an overview of the network model used to test the algorithms. Section VII demonstrates how customers can be identified to fulfill particular demand service requirements.

II. Ancillary Services for Smart Grid

High penetration of renewable generation in a power system depletes the traditional demand services provided by conventional generation plant. The output variability of renewables can make demand services all the more vital to robust operation of the system if network investments, such as installation of Onload Tap Changers (OLTC) to distribution transformers, are to be avoided [9]. Load or demand cannot be realistically considered as negative generation as it has significantly different operating dynamics [12]: loads are not tailored to the power system at design time, 'ownership' is across several different 'operators' and in the case of domestic loads, the effect is fragmented and reliant on participation. Definition of the ancillary services that relate to demand are identified by [13] can be summarized as:

- Voltage control services require a response time in the order of 1 to 10 minutes; these are used to when changes in consumption or embedded generation cause fluctuations in local voltage.
- Spinning reserve services require a response time of from less than 1 second to 10 minutes; while not a generating unit, reduction of demand can be used to stabilize system frequency for example in the event of the loss of a generating unit. Because of the immediacy of such events, a shorter term response is preferred to the functionally equivalent service that balances power.
- Load shaping services respond in 30/60 to 120 minutes and are used either to mitigate network congestion or to reduce the load factor; speed of response is not a

primary concern, rather the flexibility of the load to have its operation time moved by up to several hours.
- Power balancing services respond in 1 to 10 minutes and can be used to address mismatch between power generation and demand during normal operating conditions with deferral times being in the order of tens of minutes.

Reducing the load factor or decongesting the network by reshaping load profiles can be addressed over as much as a 2 hour window so with both services, identifying suitable customers on the LV/Residential network who will be willing and consistent participants is a prerequisite to robust operation of the service. The identification process would necessitate examination of typical usage habits, which smart metering data could provide.

For operators of LV networks or islanded power systems, tracking the behavior changes of load profiles or parts of load profiles from specific (e.g. weekend) days to the next and what aspect of them changes such as magnitude, time of use, and variability will assist in informing system management and upgrade activities. Seasonal variation will play some part in this as certain loads such as lighting or heating will gain or lose flexibility according to weather effects. Detection of this cannot be assessed solely through outdoor conditions as variations in building fabric and behavioral routine will also influence the effect these have on the occupant.

III. Modeling Loads from AMR data

While metering only reports the aggregated load at the premises level, the levels of variability throughout the day can be used to infer non-critical or time shift-able loads. Capturing how much of this is variability is prone to significant changes is more of a challenge. In [14], a stratification of load profiles is proposed to abstract usage levels on a daily basis; occurrences of bandings of load levels or stratifications were accumulated over a day thus reducing profiles to collections of counts; disappearances of different strata populations and supplanted strata was often invoked by seasonal change. This idea was extended to a multivariate case to incorporate time of use in [15], with advance times modeled as separate variables; this approach differed from earlier MV network customer profiling by [16] for example in that it incorporated temporal variance. Communications methods proposed or already in place for Smart Meters do not permit real time relaying of load data so these will be of most use retrospectively generating advice for selecting participants and quantifying the consumption for demand response.

Residential LV network loads are noted for their variability [1, 14, 15]; residential customers have a greater deal of flexibility in the way they carry out their day to day tasks compared with the industrial customers connected to the MV network [16, 17]. This variability can stem from occupancy, building characteristics, heat sources as well as the aggregated combinations of appliance usage. A representation used to accommodate this variability in behavior adopted in [15] entailed a multivariate Gaussian mixture distribution: each dimension corresponded to one of the 48 half hourly meter

advance times throughout the day in terms of its mean and variance. Owing to the non-stationary and distinctly non-Gaussian nature of loads of this size [18], a mixture of Gaussian distributions that accommodates these features as multimodality was also used. This assumption of several underlying modes of behavior is a key part in capturing how consistently loads behave. Differences in the expected value of a load, represented by the Gaussian mean, and the confidence with which it can be expressed, represented by the variance parameters, will represent different, or changes in, domestic routine. The mixture distribution uses this to abstract a daily load into a single label that is indicative of the mean and variability of the load profile advances – but how long a customer retains that label is not captured by this model which would have implications on how well this customer would serve a demand reduction program or a load shifting one.

IV. BAYESIAN CHANGE-POINT DETECTION

Detecting abrupt changes in an observed measurement or tracking its evolution can be achieved by modeling how it is statistically distributed; the Kalman filter [19] is an example of this for continuous real valued observations. The act of inferring the distribution of a measurement is more robust to noise than say thresholding its observations which would be prone to changes incurred by outliers. The Kalman filter has its limitations, in particular Gaussianity and the assumed linear relation between successive observations. In situations where this cannot be assumed it is preferable to adopt no assumption as to the form of the dependency. The approach to changepoint detection proposed in [20] amounted to an online estimation of a predictive distribution over the number of observations for which that distribution was valid (the 'run length'). A further desirable property is the online operation of the algorithm, which alleviates the need for storing operational examples for training or selecting model cardinality such as the number of clusters.
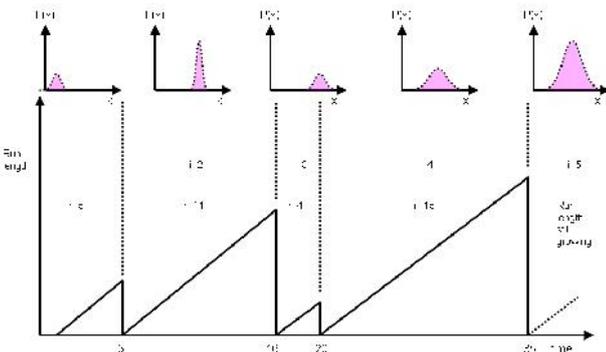

Fig 1 Change-point analysis over a time series for a hypothetical variable x: x is distributed differently at various points in the sequence, with the parameterization of the running estimate of its distribution reflecting this.

As illustrated in Figure 1, over a sequence of observations of variable x of length T, there will be a set of I changepoints for which the ith distribution will be valid at a given time. The number of observations over which an observation distribution $P(x)$ remains valid is its run length r, as the series progresses the run length increases until a changepoint is encountered at which point it is truncated to zero. Observations, x are drawn from some distribution with parameter set , which is used in two competing measures, the first is the run length growth probability:

$$P(r_t = 1 + r_{t-1}, x_{1:t}) = \sum_{r_{t-1}} P(r_{t-1}, x_{1:t-1}) P(x_t | \mathbf{y}_t^{(r)}) P(r_t | r_{t-1}) \tag{1}$$

The second is the changepoint probability or the probability that the run length is zero:

$$P(r_t = 0, x_{1:t}) = \sum_{r_{t-1}} P(r_{t-1}, x_{1:t-1}) P(x_t | \mathbf{y}_t^{(r)}) P(r_t | r_{t-1}) \tag{2}$$

A change is detected if (2) is greater than (1). At time t there will be a run length r_t estimated from observations $x_{1:t}$, from which the posterior distribution of the current run length can be calculated as:

$$P(r_t | x_{1:t}) = \frac{P(r_t, x_{1:t})}{P(x_{1:t})} \tag{3}$$

This relies on the observation likelihood, sometimes referred to as the 'evidence' which can be obtained by marginalizing over the joint likelihood of run length and observation:

$$P(x_{1:t}) = \sum_{r_t} P(r_t, x_{1:t}) \tag{4}$$

This requires only a sum over the current run length and not all observed data points, thus removing the need to store a monotonically increasing sized archive of data points. The changepoint prior distribution, $P(r_t/r_{t-1})$, can have a hazard function encoded into it to make certain run lengths unfeasibly long or short. Where there is no prior knowledge in this regard, [20] proposed this be left constant. For the observation distribution, an exponential family distribution is assumed and can be expressed in the following form, known as its natural parameterization:

$$P(x | \mathbf{y}) = h(x) \exp\{\mathbf{y}^T U(x) - A(\mathbf{y})\} \tag{5}$$

Where:

- is the natural parameter
- U is the sufficient statistic
- A is the log partition function or normalization constant
- h is the non-negative base measure which can be constant

Members of the exponential family of distributions include the

Gaussian, Gamma, Poisson, Dirichlet, Beta, Multinomial and Bernoulli distributions. A useful property of these distributions is that they exhibit what is known as conjugacy; that is, the posterior distribution takes the same form as the prior distribution [21, 22]. Bayesian statistics differs from its Frequentist counterpart in that it expresses model uncertainty entirely through probability distributions [23, 24]. Where a Frequentist would seek to estimate likelihood from whatever data was available, a Bayesian would infer a posterior distribution from the evidence (the distribution of variables) and assumptions available. This results in greater capabilities for generalization [25]. For example, a Gaussian distribution has a Gaussian prior over its mean and a Gamma prior over its precision (precision is the inverse variance) which leads to a Gaussian-Gamma prior distribution over the hyper-parameters and also a posterior distribution in the form of a Gaussian-Gamma. An exponential distribution parameterized by   can be expressed generally in the form:

$$P(y|\epsilon, t) = f(\epsilon, t) g(y)^{\epsilon} \exp\{\epsilon y^T t\} \qquad (6)$$

The variables   and   are hyperparameters, known as the sufficient statistics that describe . This, through conjugacy, permits a posterior distribution over an observation x of the form:

$$P(y|X, \epsilon, t)$$
$$\propto g(y)^{\epsilon + N} \exp\left\{ y^T \left(\sum_{n=1}^{N} u(x_n) + \epsilon t \right) \right\} \qquad (7)$$

This conveniently allows the sufficient statistics to be updated as each observation is received as follows:

$$\epsilon' = \epsilon + 1 \qquad (8)$$

$$t' = t + u(x_t) \qquad (9)$$

In the case that the run length is reset, then these are reset to their prior values assigned at the start of the sequence. These will be updated for the observation distribution as new data is received unless the run length is reset. In order to lessen the computational requirements of this procedure, a zero run length results in the whole sequence being truncated thus allowing a smaller memory footprint.

V. LOAD CHARACTERISTIC CHANGE MODELING

To specialize the changepoint detector to the load model, an observation distribution must be chosen to represent the evolving load profile; this in turn dictates the formulation of the distribution parameter updates. Following the approach used in [15], a multivariate Gaussian distributed load profile is

assumed. This captures not only the implied mean usage for every advance time but also the variability attached to it and the relationship between the advance sizes in terms of their correlation. This latter feature implies the time variation in adjacent advance periods. The multivariate Gaussian has a vector mean μ and a covariance matrix . Assuming conjugacy [22], the covariance matrix has a prior distribution over all plausible values that takes the form of an inverse Wishart distribution:

$$iW(\Sigma; \epsilon, S) = \frac{1}{Z} |\Sigma|^{\frac{-(\epsilon+k+1)}{2}} \exp\left\{ -\frac{1}{2} tr[S\Sigma^{-1}] \right\} \qquad (10)$$

The mean μ is assumed to be Gaussian distributed and so has a prior of this form.

$$N(\sim; m, \ddagger) = \frac{1}{2} \exp\left\{ -\frac{1}{2}(m - \sim)^T \ddagger^{-1} (m - \sim) \right\} \qquad (11)$$

The covariance scale parameter can be integrated over; leaving a distribution that is a function only of its hyperparameters:

$$P(x) = \int_0^\infty N(x|\sim_t, uS_t) iW(u|\epsilon_t, S_t) du \qquad (12)$$

Integrating reduces this to the form of the multivariate Student-t distribution [26, 27]:

$$P(x; \hat{\ }, \sim, \Sigma)$$
$$= \frac{\Gamma((\hat{\ } + d)/2)}{\Gamma(\hat{\ }/2)(\hat{\ } f)^{d/2} |\Sigma|^{1/2}} \qquad (13)$$
$$\left(1 + \frac{1}{\hat{\ }}(x - \sim)^T \Sigma^{-1}(x - \sim) \right)^{\frac{-(\hat{\ }+d)}{2}}$$

The Gamma function is given by   and the dimension of the distribution is d. The Student-t distribution can then be parameterized as:

$$P\left(x; \epsilon_t - d + 1, \sim_t, \left(\frac{\ddagger_t (\epsilon_t - d + 1)}{(\ddagger_t + 1)} \Lambda_t \right)^{-1} \right) \qquad (14)$$

The sufficient statistics that correspond to (8) & (9) are as follows:

$$\epsilon_t = \epsilon_{t-1} + 1 \qquad (15)$$

$$\sim_t = \frac{\ddagger_{t-1} \sim_{t-1} + x_t}{\ddagger_{t-1} + 1} \qquad (16)$$

To update the covariance, the scatter matrix S must be

evaluated at each time step as well and the precision (inverse covariance) updated as follows:

$$S_t = S_{t-1} + (x_t - \sim_t)(x_t - \sim_t)^T \tag{17}$$

$$\Lambda_t = \Lambda_{t-1} + \frac{N}{N+1}(x_t - \sim_t)(x_t - \sim_t)^T \tag{18}$$

$$\ddagger_t = \ddagger_{t-1} + 1 \tag{19}$$

These are returned to a set of default values whenever the run length is reset. Default values are assumed to be the equivalent of a standard multivariate Gaussian distribution with zero mean and an identity matrix for the covariance matrix. To summarize, the algorithm proceeds as follows:

1. Set μ, , S, and to represent a multivariate standard Normal distribution
2. do
3.   Accept load profile x_t for day t
4.   Update load profile distribution parameters using expressions (15)-(19)
5.   Evaluate (1) and (2)
6. while expression (1) > expression (2)
7. Changepoint detected: reset the distribution by returning to step 1

In [15] entire daily load profiles are modeled which posed computational problems that stemmed from the sparseness of 48 dimensional data: as the number of dimensions increases, more data points are required to support the relations and variability in the covariance matrix [28]. For ancillary services, this is less of a problem as only short periods, possibly around peaks, are likely to see demand services invoked. Considering a short period of advances rather than an entire day, the dimensionality of the load profile of interest can be reduced. For example, peak times may require demand responses to be invoked between 16:30 and 19:30 which covers 6 half hourly advance periods. With the resolution of advances and variation in people's routines, tasks may overlap advance periods which will in turn be captured in the Gaussian covariance matrix as a correlation between variables representing each time period.

VI. TEST LV FEEDER CIRCUIT

There are very few fully metered LV circuit data sets which motivates the synthesis of a LV feeder from actual Smart Meter data being used to populate hypothetical properties on a single phase feeder. Load flow calculations are employed to evaluate the thermal and capacity constraints throughout the section of network including that at the substation. This model is shown in Figure 2 and comprises a single feeder with 32 customers attached by short lengths of cable. The cables are of three different types, reducing in capacity with distance from the substation. Each customer is represented by a point load at separate nodes in the network and the load magnitude is

provided by a time series of actual residential AMR data. UK AMR typically measure 30 minute advances [10] which results in a load profile with 48 point resolution. No assumptions are made regarding the similarity of the dwellings in terms of size, utilization, construction or occupancy – this is a relevant assumption especially for networks in rural areas in the UK where properties tend to be heterogeneous in their size, age and palette of construction materials.
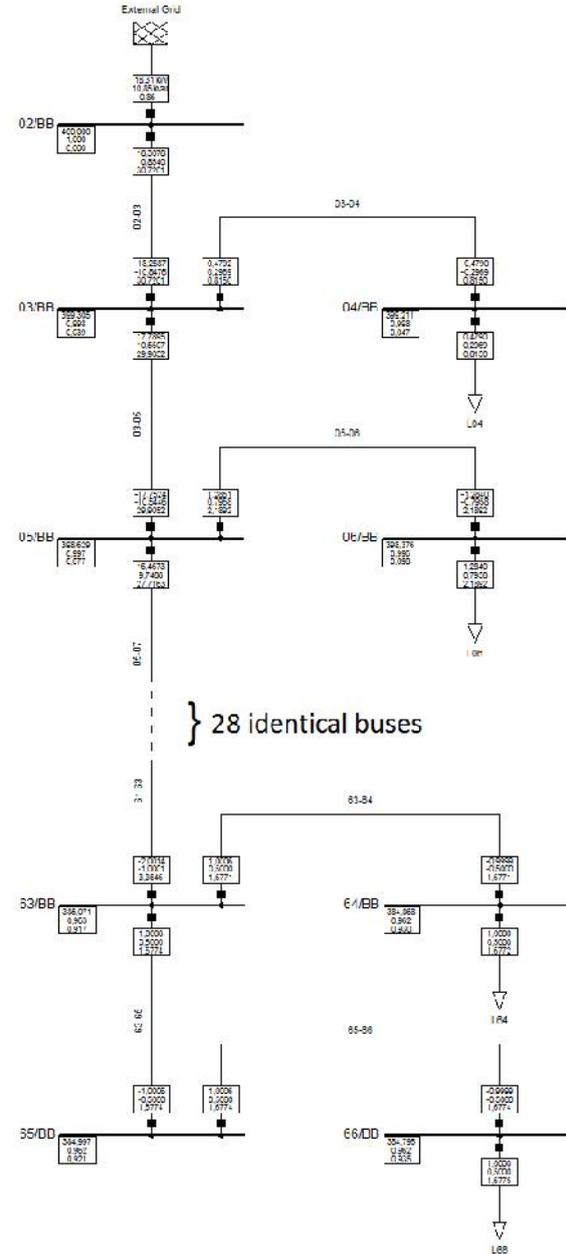

Fig. 2. Simulation network used with the models presented.

VII. SERVICE VIABILITY ANALYSIS

Providing ancillary services requires participant consistency to be effective. To evaluate this, the test network is run for a period of T=180 days, which takes in quarter 1 and quarter 2 for the year, concluding in early July. There are 32 customers

on the network, each is metered half hourly for the duration of the test and each is given its own dedicated changepoint analyzers: for morning weekday, evening weekday morning weekend and evening weekend usage. Peak periods are the most critical operating times for all networks; in the UK and much of northern Europe, there are two of these in the daily load profile: one in the morning and one in the evening. At the end of each day, load profiles are formed by time ordering the meter readings for each customer into a 48-dimensional vector. To account for variability in the start times of household routines, 7-dimensional sub-profiles centered on the 08:00 and 18:30 peak times are extracted from the 48-dimensional daily load profile vector.
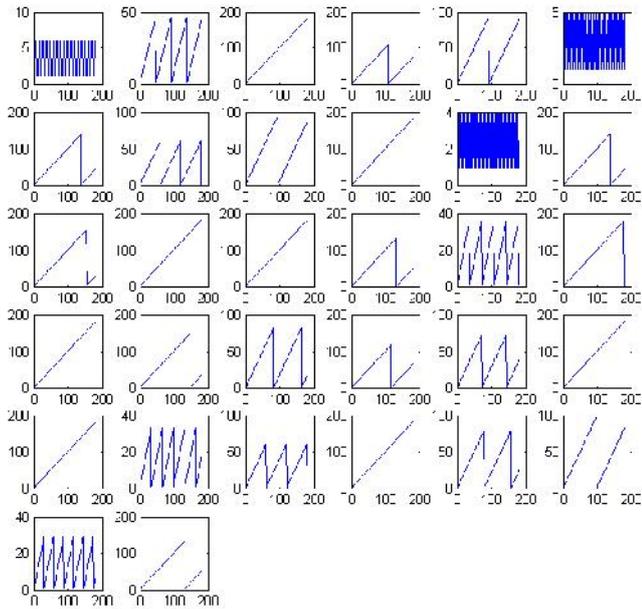

Fig 3 Run length estimates demonstrating change-points as discontinuities for weekday peak time distributions over the whole 32 node network for the 6 month duration.

Weekends and weekdays will invoke distinctly different behaviors, so separate run length estimators are employed accordingly. These load profiles are fed into their respective changepoint analyzer at the end of each day which will yield a distribution of the load profile and an indication of whether or not a changepoint has taken place for each residential customer. At the end of the 180 day period, each customer will have a list of when changepoints occurred in their load profile distributions from each of their changepoint analyzers. Statistics are then calculated at the end of the monitoring period for illustrative purposes. Figure 3 shows the resulting run lengths over the trial period for the evening peak time distribution for all 32 participants. The visualization of this data allows the variability of residential load behavior to become clear: some loads remain in effectively the same distribution for the entire trial, while others can change as often as every week (run lengths of approximately 5 days) and some remain for greater than a month's worth of week days indicating possible seasonal influence. Table I summarizes the temporal characteristics of the loads on the test network and shows that in terms of consistency, weekday morning peak

TABLE I
CHANGEPOINT STATISTICS FOR 32 CUSTOMERS OVER 180 DAYS

Case	Load Model	Number of Changes	Average Run Length	Earliest Reset	Latest Reset
Morning Peak	Weekday	24	122	90	180
	Weekend	5	41	31	52
Evening Peak	Weekday	154	97	4	180
	Weekend	35	31	5	52

behavior is fairly consistent with no run length truncations seen until 90 days into the trial and a total of only 24 peak load profile distribution changes seen throughout the trial. In recognizing this sort of consistency of behavior, a useful participant in demand response would have been identified. Weekend morning peaks are more consistent still, with only 5 changes in peak load profile distribution and an average run length of 41 days (equivalent to 20 weeks since only weekend behavior is taken into account here). Evening peak load profiles show greater variability, with a reset in run length observed 4-5 weeks from the start of the trial. Some residential loads, as elaborated upon in Fig 3, exhibit almost a weekly change in their load profile distribution while there are others that maintain the same load profile shape for the entire trial period. For this latter case, the time period over which groups of demand response households participated would be shorter, and those households forming the group reselected more often.
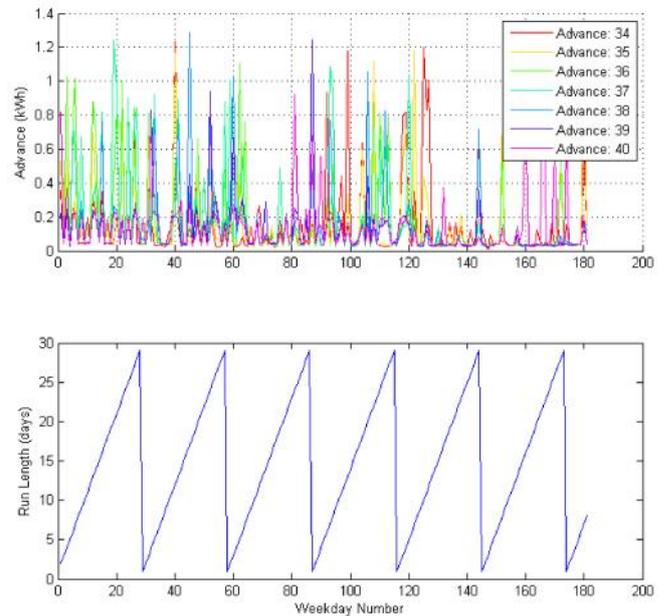

Fig 4 Node 31 weekday behavior over the monitoring time period in terms of advances (a - upper) and (detail of Fig 3) overall load profile probability distribution run length (b - lower).

A total of 154 run length resets are observed over the entire trial with an average run length of 97 days; run lengths range from 4 to 180 days, highlighting how the variability of behavior has skewed this average. To consider this in more detail, Figures 4, 5 and 6 show an example household over the trial period. Figure 4a shows the 7 advances of periods 34 to

40 (i.e. the load profile between 17:00 to 20:00 inclusive) over the whole trial period while Figure 4b shows the run length of the multivariate distribution that corresponds to this. The various resets are temporally aligned with the advance time series. Periods 36 and 37 in particular can be seen from Figure 4a as having very changeable variances – large up to day 60, then varying again after day 80, although to a lesser extent. After day 120, both periods 36 and 37 exhibit less variance with only intermittent peaks. Summarizing the delineations in this multivariate time series, the associated mean load profiles and covariance matrices captured prior to each reset are those shown in figures 5 and 6 respectively.
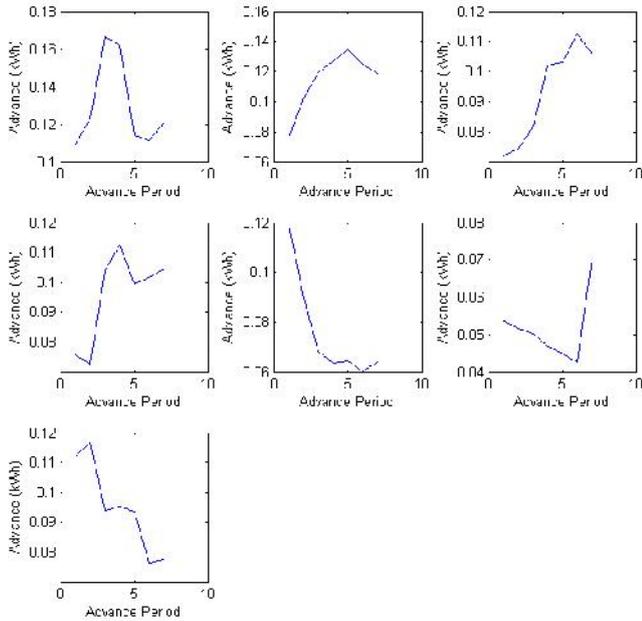

Fig 5 Means of the load profile distributions captured over each run length at network node 31 for weekday evening peak-time profiles. These are presented left to right, top to bottom as they evolve over time.

The Gaussian distribution captures the evolution of the load profile over time through its mean parameter μ: Figure 5 shows the movement of a peak towards later in the evening for the first three run lengths, then moving back earlier in the day for the next four. Figure 6 shows the Gaussian covariance matrix parameter evolving with the means in figure 5, which attaches a level of certainty to the magnitudes of the expected peak loads indicated by the means. The first distribution (Figure 5, top left) can be seen to have relatively low variance around the peak of the load profile at advance 37-38 representing a confident prediction of the time of the peak. The advances shouldering this peak have greater variability as indicated by the change in color of the cell towards red. The off diagonal elements of the matrix represent the covariance or the linear dependence between advance times. Large, positive covariances mean that when a particular advance period load increases in magnitude, its corresponding advance load in the matrix does as well; covariances tending towards uniformity would result in the preservation of load profile shape with the increase in energy used, a change which could result from the increase in usage of electric space heating, for example. The

final reset shows very low covariance for all advance pairs representing little or no dependency between times of use.
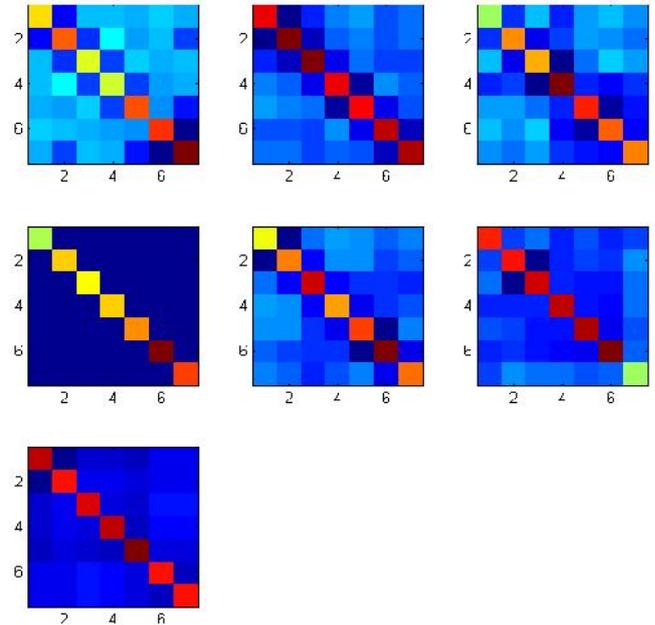

Fig 6 Covariance matrices of the load profile distributions captured over each run length at network node 31 for weekday evening peak-times. Dark areas represent small variance/covariance. These are presented left to right, top to bottom as they evolve over time.

The practical benefit from this understanding is in the realization of a consistent aggregated demand resource that could supplant the need for an artificial or industrial load on a small power system, Recognizing when particular loads are of a particular form could be a future function of substation computing devices which in turn could identify potential for seasonal or event driven reverse power flows in the presence of high penetrations of distributed generation.

VIII. CONCLUSIONS

This paper has shown how basic AMR meter readings can be used in conjunction with a novel but computationally simple changepoint detection algorithm to assess behavioral consistency in domestic loads. Behavioral consistency is critical to achieving desired demand response from groups of small loads: identifying the load sources that are the least unpredictable will be the best candidates for participation in demand response schemes. The algorithm has been shown to identify particular behaviors as probabilistic representations of load profiles and their variability over selected periods of interest, and durations in time over which these representations are valid. In algorithmic terms, the next stage of this work is to retain distributions after their run length has reset and identify their future recurrence. This way, sets of behaviors for a given load can be accrued over time allowing more advanced, higher level models to utilize these for future grid services and possibly control. In [15], finer grained residential load profiling was compared to an averaged load profile in line voltage and current calculations with performance benefits

being demonstrated in several cases. Network operators can use tools employing this algorithm to assess customer suitability for providing ancillary services and can plan microgeneration integration and capacity requirements with greater accuracy.

REFERENCES

[1] Gram-Hanssen, K. "Standby consumption in households analysed with a practice theory approach", Journal of Industrial Ecology Vol. 14, No. 1, 2009.

[2] "Load Profiles and their use in Electricity Settlement", Available: http://www.elexon.co.uk/documents/participating_in_the_market/market_guidance_-_industry_helpdesk_faqs/load_profiles.pdf

[3] Jorgensen, J.M.; Sorensen, S.H.; Behnke, K.; Eriksen, P.B.; , "EcoGrid EU — A prototype for European Smart Grids," Power and Energy Society General Meeting, 2011 IEEE , vol., no., pp.1-7, 24-29 July 2011

[4] Smart Grid Gotland – Electricity Network for the Future. Available at: http://www.smartgridgotland.se/eng/index.pab

[5] Mathieu, J.L., Price, P.N., Kilicote, S. & Piette, M.A. "Quantifying Changes in Building Electricity Use, With Application to Demand Response". IEEE Transactions on Smart Grid, Vol. 2, No. 3, September 2011.

[6] Seung Tae, C., Qiuwei, W. & Ostergaard, J., "A generic Danish distribution grid model for smart grid technology testing," Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on , vol., no., pp.1,6, 14-17 Oct. 2012

[7] Gill, S., Dolan, M., Frame, D. & Ault, G. (2011) The role of the electric heating and district heating networks in the integration of wind energy to Island networks. International Journal of Distributed Energy Resources, 7 (3). 245 - 263. ISSN 1614-7138

[8] von Appen, J., Braun, M., Stetz, T., Diwold, K. & Geibel, D. "Time in the Sun: The Challenge of High PV Penetration in the German Electric Grid", IEEE Power & Energy Magazine, pp. 55-64, March-April 2013

[9] Stetz, T.; Marten, F.; Braun, M., "Improved Low Voltage Grid-Integration of Photovoltaic Systems in Germany," Sustainable Energy, IEEE Transactions on , vol.4, no.2, pp.534-542, April 2013.

[10] Renner, S. et al., "European Smart Metering Landscape Report", Smart Regions Deliverable D2.1., 22 February 2011, available at http://www.smartregions.net/

[11] Hammerstrom, D.J. et al. Pacific Northwest Gridwise Test bed demonstration projects part 2: Grid Friendly Appliance Project. October 2007.

[12] Gehrke, O.: "Proposal for an information model for demand-side resources", online at: http://vea-e02.risoe.dk/proposal_infmodel_demand_side.pdf

[13] Hirst, E. & Kirby, B., "Electric Power Ancillary Services", Oak Ridge National Laboratory, Tech. Report, Oak Ridge, TN, 1997.

[14] Stephen, B. & Galloway, S. "Domestic Load Characterization through Smart Meter Advance Stratification", IEEE Transactions on Smart Grid Vol. 3, No. 3, September 2012.

[15] Stephen, B., Mutanen, A., Galloway, S., Burt, G. & Jarventausta, P. "Advanced Load Profiling for Residential Customers", IEEE Transactions on Power Delivery, In Press, March 2013.

[16] Chicco, G., Napoli, R. & Piglione, F. "Comparisons Among Clustering Techniques for Electricity Customer Classification", IEEE Transactions on Power Systems, Vol. 21, No. 2, pp. 933-940, May 2006.

[17] Mutanen, A., Ruska, M., Repo, S. & Jarventausta, P., "Customer Classification and Load Profiling Method for Distribution Systems", IEEE Transactions on Power Delivery, Vol. 26, No. 3, pp.1755-1763, July 2011.

[18] Carpaneto E. & Chicco G. Probabilistic characterization of the aggregated residential load patterns. IET Generation, Transmission & Distribution, 2007.

[19] Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME--Journal of Basic Engineering, Volume 82, Series D, pp. 35-45, 1960.

[20] Adams, R.P. & MacKay, D.J.C., "Bayesian Online Changepoint Detection". University of Cambridge Technical Report 2007.

[21] Ghahramani, Z. & Beal, M.J. Propagation Algorithms for Variational Bayesian Learning. In Advances in Neural Information Processing Systems 13:507-513, eds. T.K. Leen, T. Dietterich, V. Tresp, MIT Press, 2001.

[22] Fink, D., and Goodman, D. Using Conjugate Composition Priors to Compute Posterior Distributions. 1997.

[23] Lee, P.M. Bayesian Statistics: An Introduction. Oxford: Oxford University Press, Third Edition, 2003.

[24] Mackay, D.J.C. Information Theory, Inference and Learning Algorithms Cambridge Univ. Press, Cambridge, 2003.

[25] MacKay, D.J.C., Probable networks and plausible predictions—a review of practical Bayesian methods for supervised neural networks. Network: Computation in Neural Systems, 6:469–505, 1995.

[26] Peel, D. & McLachlan, G. J. Robust mixture modelling using the t-distribution. Statistics and Computing 10:4, pp 339 – 348. Springer, October 2000.

[27] Bishop, C. M. Pattern Recognition and Machine Learning, Springer (2006)

[28] Tipping, M. E. and C. M. Bishop. Mixtures of Probabilistic Principal Component Analysers. Neural Computation 11(2), pp. 443–482, 1999

Dr Bruce Stephen (M '09) currently holds the post of Senior Research Fellow within the Institute for Energy and Environment at the University of Strathclyde. He received his B.Sc. from Glasgow University and M.Sc. and PhD degrees from the University of Strathclyde and is a Chartered Engineer. His research interests include Distributed Information Systems, Machine Learning applications in Power System and Agricultural Condition Monitoring/Asset Management.

Fridrik Rafn Isleifsson received the M.Sc. degree in electrical engineering from the Technical University of Denmark in 2006 He is currently a Ph.D. student with the Intelligent Energy System Programme at the Technical University of Denmark. His research focuses on demand side participation in providing ancillary services to the power system and implementation of control strategies in laboratory demonstration.

Dr Stuart Galloway is a Senior Lecturer within the Institute for Energy and Environment. He obtained his MSc and PhD degrees in mathematics from the University of Edinburgh in 1994 and 1998 respectively. His research interests include power system optimization, numerical methods and simulation of novel electrical architectures.

Prof Graeme Burt received his BEng in Electrical and Electronic Engineering from the University of Strathclyde in 1988. He received his PhD also from the University of Strathclyde in 1992. He is currently a Professor at the University of Strathclyde, and is Director of the University Technology Centre in electrical power systems, sponsored by Rolls Royce. His research interests lie predominantly in the areas of power system modeling and simulation, power system protection, and network integration of distributed generation.

Henrik W. Bindner is a Senior Scientist at Intelligent Energy Systems Research Programme, Risø National Laboratory for Sustainable Energy. He has worked in the field of integration of wind power into power systems and island systems with a high penetration of renewable energy. Currently he is researching integration of renewable energy in the context of intelligent power systems with particular focus development of smart grids and supporting technologies such as application of batteries. He has been responsible for and project manager for establishing an experimental facility for distributed intelligent power systems, SYSLAB, that include wind turbines, PV plant, batteries as well as other equipment and an advanced distributed control infrastructure.