Load Decomposition and Profiling for “Smart Grid” Demand-Side Management Applications

Andreas Paisios¹ and Sasa Djokic¹

¹The University of Edinburgh, Edinburgh, Scotland, UK
a.paisios@sms.ed.ac.uk
sasa.djokic@ed.ac.uk

Abstract. Active and reactive power demands measured in Scottish Medium Voltage (MV) distribution networks are used in this paper to correlate variations in demands with the short-term and long-term variations of temperature. The methodology is illustrated on the example of the total network demands (sum of the demands at all monitored bulk load supply points), with the analysis presenting several metrics and indicators for quantifying and correlating temperature-demand dependencies during the course of one calendar year. The paper also presents a simple approach for decomposing aggregate network demands into the temperature-dependent loads (e.g. thermal heating or cooling loads) and temperature-independent loads, providing an important information for the application of the “smart grid” functionalities, such as demand-side management, or balancing of variable energy flows from renewable generation.

Keywords: Load modelling, decomposition and profiling, “smart grid”, demand-side management, temperature-demand correlation, correlation/regression coefficients, “sliding window” averaging, normalisation and “smoothening”.

1 Introduction

Future electricity supply systems (so called “smart grids”) will see significant increase of renewable-based distributed generation (DG), radical transformation of transmission and distribution networks and introduction of new highly efficient, intelligent and automated control, monitoring and communication infrastructures. This will be necessary to reduce CO₂ emissions and other drivers of climate change, while simultaneously maintaining high levels of security, sustainability and affordability of electricity supply. It is, however, widely recognised by all stakeholders that supply-side solutions alone will not be sufficient for the realisation of these challenging tasks. Additional strong support and contributions are expected to come from demand-side actions and measures, i.e. from a modified behaviour and physical demand for electricity through the consumer choice and evolved end-use of demand-manageable load and micro-generation. This will open major opportunities for a more direct and proactive system support, but will also result in profound changes in levels and nature of system-user interactions, shifting the actual system operating and loading conditions well outside the traditionally assumed ranges, limits and physical boundaries.
In spite of all anticipated changes, however, the simple fact is that a reliable and accurate assessment of the operation of both existing electricity networks and future “smart grids” cannot be performed if accurate models of aggregate system loads are not available. This is particularly true for the analysis of the effects and possible benefits of applying various demand-side management actions and schemes.

Demand-side management (DSM) is generally denoted as a set of measures, actions and interventions, initiated deliberately and with a specific purpose by end-users and/or network operators, or a third party (e.g. energy suppliers), aimed at changing, restructuring and/or rescheduling power demands of a group of loads, load sector(s), part of a system, or a whole system, in order to produce desired changes in the actual amounts and time patterns of power demands supplied at the dedicated point(s) of delivery for end-use consumption of electricity. Accordingly, before the specific DSM action or scheme should be applied, it is important to as accurately as possible identify the participation of the load targeted by the DSM in the total system demand and, equally, to estimate the potential effects of directly controlled DSM portion of the load on the improvement (or deterioration) of network performance.

Aggregate network demands are typically measured at primary distribution substations, representing bulk load supply points (BLPs) connected at medium voltage (MV). Identification of the demand-manageable portion of the load in the total demand, therefore, requires decomposition and profiling of the BLSP-measured aggregate demands. This paper presents initial results of the analysis aimed at identifying temperature dependency of the total network demands, obtained as the sum of all measured individual BLSP demands. The analysis includes several metrics and indicators for quantifying temperature-demand dependencies during the course of the year. These results are used as a basis for identifying contribution of the heating load to the total electricity demand during the winter season, as well as the participation of the cooling load in the total electricity demand during the summer season.

2 Overview of the Previous Work

There is an increasing amount of work and effort dedicated to the analysis of electricity demands, where particularly the relationships between the demands and various weather/meteorological conditions, as well as human behaviour and economic factors have been analysed. Ultimately, the results of this work should be utilised by electricity industry or transmission and distribution network operators for improved system operation and, in the recent years, as a means of devising and implementing optimal DSM actions and schemes, which are considered as an important part of "smart grid" functionalities.

Parameters that affect short-, medium- and long-term variations of electricity demand have been studied by a number of authors and include: the time/season of the year, day of the week, hour of the day, temperature, wind speed, relative humidity, public holidays, holiday seasons and geographic locations, as well as the factors relevant to a long-term analysis, such as economic development, climate change, technological innovations, population growth, urbanization, etc., [1], [2], [3] and [4].
Previous research included analysis of thermal electricity demands (e.g., for space and water heating loads), where contribution of thermal loads within a characteristic day/season of the year and for the characteristic hours of the day have been identified in [3] and [5]. These characteristic hours have been identified as the turning points in the daily power demands, either due to the group on/off switching of heating appliances, or as a consequence of habitual and socio-behavioural factors (e.g., differences in weekday-weekend working schedules and corresponding daily working hours).

For the purpose of smoothening input data into regression/correlation models and, in that way, compensating for random fluctuations by providing overall characteristic trends over several time series of demand and temperature, “moving average”, denoted in this paper as a “sliding window (SW) smoothening” technique has been commonly used, e.g., in [16]. As the smoothened data provide only the average characteristic variations for a given period, the study of the residuals, or differences between the actual data and smoothened data, has also been a subject of research in [6].

Demand-side participation is an area which is recently receiving perhaps the most of the attention in the field of electricity demand analysis and optimization, [7]. For the implementation of various DSM actions and schemes, load identification in the residential, commercial and industrial sectors is of particular importance, and is typically aimed at a detailed profiling of variations of various load types (e.g., heating loads, power electronic loads, lighting loads, etc.) and their associations with weather and socio-economic factors. The examples include development of thermostatically controlled load models or frequency controlled system reserve loads, [8].

Similarly, there is a number of models developed for electricity demand forecasting, including time series analysis, linear, nonlinear and multiple regression or correlation analysis, artificial intelligence and neural networks, Gray-based approaches, expert systems and others, [1], [9], [10]-[13]. Long-term forecasting techniques are important for the system development and upgrading, particularly planning of transmission and distribution networks, [14]. These researches focused on a long-term peak electricity demand forecasting, where, among the other methodologies, seasonal bootstrapping and variable blocks methods and non-linear analysis of demand variations have been developed, [7], [9]. In the context of time series analysis, Box-Jenkins methodology has also been used, including autoregressive moving average and autoregressive integrated moving average (ARMA and ARIMA) models in [11] and [15].

Finally, artificial neural network approaches are used, requiring a lot of prediction rules and sufficient amount of accurate input data for the development/learning phase of the neural networks [12]. Unlike the regression models, these approaches use different forms of "black box" models, which generally do not allow for the more obvious understanding of the learning phase of neural networks, [1].

Building on some of the previous work, this paper provides initial analysis of decomposition of electricity demands in Scotland, UK. The approach is aimed at a more detailed analysis of the dependency of electricity demand on temperature variations from a number of different temporal perspectives (annual, seasonal, daily and hourly). This allows for the distinction between dependencies on different time scales, which, in turn, will enable to study these effects in the context of short to medium to long term variations of electricity demand and related DSM applications in “smart grids”.


3 Description of Available Measurement Datasets

Active power (measured in MW), reactive power (measured in MVAR) and temperature (measured in degrees Celsius) were monitored for the period of one whole calendar year, from 01/04/2009 to 31/03/2010. Two power demand datasets consists of total active and total reactive power from 84 BLSPs, covering a geographical area of southern Scotland, UK. Temperature is taken from the measurements available at the University of Edinburgh weather station.

3.1 Annual Variations

Figure 1 presents annual variations of per-unitised values of actually measured active power, reactive power and temperature (recalculated from originally measured Celsius degrees to Kelvins to prevent occurrence of negative values, which may mask correlation between temperature and demands). The normalisation is done using (1).

\[
V_{pu_i} = \frac{V_{ui}}{V_{max}}
\]

where: \(V_{pu_i}\) is the \(i\)-th per-unit value of either active power, reactive power or temperature dataset, \(V_{ui}\) is the actually measured \(i\)-th value and \(V_{max}\) is the maximum value of each dataset (i.e. maximum yearly active power, reactive power and temperature).

The graphs also include the “sliding window (SW) smoothed values”, which were calculated by (2), using a window of +/- 14 days around each actual value.

\[
V_{SW_i} = \frac{\sum_{j=-14}^{14} V_{pu_{i+j}}}{29}
\]

where: \(V_{SW_i}\) is the SW-smoothed \(i\)-th value in each dataset, calculated by the summation of the per-unit values for a window length of 29 days around the actual per-unit value \(V_{pu_i}\) and divided by the total number of values in the summation series (29).

For each of the three variables, Figs. 1a, 1b and 1c illustrate annual variations of per-unit daily mean values and sliding window daily mean values, daily maximum values and sliding window daily maximum values, as well as daily minimum values and sliding window daily minimum values. Original measurements for each day of the year include 48 half-hourly measurements. These were first per-unitised with respect to the absolute maximum in each dataset and then daily mean, maximum and minimum values were calculated.

It can be seen from Figs. 1a and 1c that single-value daily temperature and active power demand curves exhibit a higher frequency short-term variations (representing changes in the actual per-unit mean, maximum and minimum daily values), which are imposed on a lower frequency long-term sinusoid-like curves with a period of one year (represented by the corresponding SW-smoothed per-unit mean, maximum and minimum daily values). As the measured demands correspond to a cold-climate geographic area of Scotland, for which electricity demand of the thermal loads is predominantly determined by the mostly resistive heating load in winter (i.e., there is a smaller contribution of non-resistive cooling load during the summer), SW-smoothed per-
unit curves for temperature and active power demand are mostly opposite in phase, indicating that the maximum network active power demand occurs around the minimum temperature and, *vice versa*, the minimum network active power demand occurs around the maximum temperature. These results clearly suggest that temperature and active power demand are correlated (with a strong negative or inversely proportional relationship), with active power demand additionally influenced by the differences between weekdays and weekends. Fig. 1b, however, shows that the reactive network power demand curves exhibit much bigger short-term variations and have almost no long-term periodical variations, which is due to the application of the network reactive power compensation and control schemes. In the next section, the above measurement data are further processed, to quantify differences between the different seasons, as well as between the weekdays and weekends.

![Graph](image_url)

**Fig. 1.** Annual variations of total network active and reactive power demands (sums of all 84 BLSP measurements) and temperature without distinction between the weekdays and weekends.
3.2 Seasonal Variations

While previous section presents active power, reactive power and temperature variations over a course of the year by only three pairs of values for each day (mean, maximum and minimum per-unit and SW values), this section shows variations of active/reactive power demands as “seasonal average days”, calculated as mean half-hourly values for each season (Spring: March to May, Summer: June to August, Autumn: September to November, and Winter: December to February), Fig. 2. Here, a main distinction is made between weekday and weekend active/reactive power demands, indicating considerable variations between the corresponding datasets.

![Seasonal Variations Chart](image)

**Fig. 2.** Seasonal variations represented as “average days” (daily load curves), clearly showing differences between the demands on weekdays (WDays) and weekends (WEnds)

The results in Fig. 2 are again per-unitised with respect to the maximum value of each dataset (MW or MVAr), to preserve consistency and allow for a direct comparison between the weekdays and weekends, as well as between the four seasons. None of the values reaches 1pu, since each plotted half-hourly value is the mean value through all days (weekdays or weekends) of each season. These results indicate that active and reactive power demands on weekdays and weekends should be clearly distinguished, as they represent variations due to an important socio-behavioural factor. For example, morning increase of demands on weekends occurs later and is with a lower gradient than on weekdays, while weekdays are characterised by a higher overall per-unit demand than the weekends.

3.3 Variations for Selected Characteristic Hours

This section presents the annual variations of active power, reactive power and temperature for two selected characteristic hours of the day: 3am (representing increased thermal heating demand during winter) and 5pm (representing increased thermal cooling demand during summer), Figs. 3a and 3b. In Section 5, load decomposition and profiling of the thermal load components (heating and cooling loads) of the total demand will be performed for these characteristic hours.
The results in Figs. 3a and 3b are presented as per-unit values (where maximum value for each variable throughout the year is used for the normalisation), as well as the corresponding SW-smoothed values for each variable at each characteristic hour. Additionally, Figs. 3a and 3b also present plots of the “residuals” for each variable, given by (3), with the corresponding values oscillating around the 1pu:

\[ V_{\text{res}_i} = V_{\text{pu}_i} / V_{\text{SW}_i}, \quad (3) \]

where \( V_{\text{pu}_i} \) is the actual per-unit value divided by the corresponding sliding window value \( V_{\text{SW}_i} \) for the \( i \)-th value in the datasets. Plots of the residuals are used to extract higher frequency variations of the actual measured per-unit values from the lower frequency variations of SW-smoothed values, acting as a kind of a “low-pass” filter. These results are of importance for the decomposition of the total demand curves/envelopes in the temperature-dependent and temperature-independent loads.

**Fig. 3. Annual variations for two selected characteristic hours**
The results in Fig. 3 indicate that there is a much stronger negative (i.e. inversely proportional) correlation between the active power demands and temperature at 3am (Fig. 3a) than at 5pm (Fig. 3b) during the autumn-spring-winter seasons, while active power demand at 5pm (Fig. 3b) exhibits a stronger positive correlation with (i.e. it is proportional to) the changes in the temperature during the summer season than at 3am (Fig. 3a). These results indicate a higher contribution of thermal heating load to the total demand at 3am than at 5pm, while the presence of thermal cooling load is indicated much stronger at 5pm than at 3am during the summer.

4 Correlation/Regression Analysis

There is no indication of any significant correlation between temperatures and reactive power demands. Therefore, the analysis presented in this section correlates only temperature and active power demand variations, using a number of coefficients (or indicators) to quantify dependencies. The first one is Pearson’s coefficient, [17]:

\[ r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \]

where: \( r_{xy} \) is the Pearson’s coefficient for two considered variables \( x \) and \( y \), \( x_i \) and \( y_i \) are the \( i \)-th value for each variable and \( \bar{x} \) and \( \bar{y} \) are their mean values.

The second is Spearman rank correlation coefficient, which is similar to the Pearson’s, but allows for ranking and non-parametric correlation of variables with a lower sensitivity to data outliers. To further strengthen the establishment (or rejection) of correlations, linear regression analysis was also used in this paper. In the context of linear regression, a “best fit line” was determined using the “least squares method”, for which the sum of the squared residuals of the model is minimised:

\[ \min S = \sum_{i=1}^{n} r_{res}^2 \]  

where: \( \min S \) indicates the minimum of the sum of squared residuals \( r_{res}^2 \). Residuals in (5) represent deviation of actual values from the linear best fit line and are, therefore, difference of the sampled values from the values predicted by the model, [18].

\[ r_{res} = y_i - f(x_i, \beta) \]

where: \( r_{res} \) residuals of sample \( i \), \( y_i \) predicted value \( i \) and \( f(x_i, \beta) \) sampled value \( i \).

Using this technique, linear regression produces a linear equation:

\[ y = \beta_1 x + \beta_0 + e \]

where: \( \beta_1 \) is the slope of the best-fit line, \( \beta_0 \) is the y-intercept, \( e \) is the random error and \( y \) and \( x \) are the depended and explanatory variables respectively (often referred to as response and predictor variable, used to make predictions by substitution).

Finally, correlation analysis includes coefficient of determination, \( R^2 \), which indicates the proportion of the variance of the dependent variable that can be attributed to the variability of the independent variable, or the proportion of variability in the da-
taset that can be explained by the model of the least squares fit, [18] and [19]. In the cases where an intercept is included (as for the linear regression technique used in this paper), $R^2$ is simply the squared value of the correlation coefficient. The coefficient of determination can also be expressed as the fraction of the sum of squares explained by regression, divided by the total sum of squares, taking values from 0 to 1:

$$R^2 = \frac{SS_{reg}}{SS_{tot}} = 1 - \frac{SS_{err}}{SS_{tot}}$$

(8)

where $R^2$ is coefficient of determination, $SS_{reg}$ sum of squares explained by regression and $SS_{tot}$ the total sum of squares.

4.1 Regression/Correlation Analysis for Characteristic Hours

Figures 4a and 4b show coefficients for two characteristic hours (3am and 5pm) throughout the year, calculated using a moving regression/correlation model for +/-10 weekdays within a sliding window of 2 weeks. In that way, the variability of the correlation through the year is captured to a greater detail compared to providing only one coefficient for the whole year, or one for each season/month of the year.

![Figure 4a](image1.png)

![Figure 4b](image2.png)

Fig. 4. Regression analysis for two characteristic hours

Figure 4 allows to identify and quantify annual correlation of active power demand and temperature on a per-hour basis. It can be seen, for example, that a strong nega-
tive correlation at 3am starts to change at the beginning of July and oscillates between positive and negative until the beginning of September, indicating periods for which heating load (significantly) reduces and cooling load is comparable to or higher than the heating load (Beta values change from negative to positive). For 5pm, the change is similar, but it starts one month earlier, at the beginning of May. The strongest positive correlation and highest positive values of Beta at 3am and 5pm coincide with the highest measured temperature, indicating period of minimum/base heating load. Longer periods of time of positive correlation between demand and temperature at 5pm than at 3am again suggest a higher contribution of cooling load to total demand at 5pm than at 3am. Easter and Christmas holiday seasons are clearly seen as the reductions in the total demand around mid-April and end of December, best indicated by reduced R² values. Few unusually hot days at beginning of February are responsible for a similar Beta/R² changes at 5pm, and a bit less at 3am. These results (mainly Beta values) are used to decompose the total demand into heating and cooling load.

5 Decomposition of Thermal Heating and Cooling Loads

This section uses statistical information on the UK electricity demands from [20] and [21] to estimate contributions of heating and cooling loads to total network demands at characteristic hours, listed in Table 1 (in%) for minimum/summer, average/spring/autumn and maximum/winter loading conditions. Heating load represents direct and storage hot water loads, as well as direct, storage and top-up space heating loads, while cooling load represents air-conditioner loads of different sizes and types, including large heating, ventilation and air-conditioning (HVAC) systems.

Table 1. Estimated percentage contributions of the thermal heating and cooling loads to the total network active power demands, [20] and [21].

<table>
<thead>
<tr>
<th>Hour</th>
<th>Minimum (Summer) Heating</th>
<th>Heating</th>
<th>Cooling</th>
<th>Average (Spring &amp; Autumn) Heating</th>
<th>Heating</th>
<th>Cooling</th>
<th>Maximum (Winter) Heating</th>
<th>Heating</th>
<th>Cooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>3am</td>
<td>6%</td>
<td>7%</td>
<td>18%</td>
<td>2%</td>
<td>29%</td>
<td>0%</td>
<td>2%</td>
<td>29%</td>
<td>0%</td>
</tr>
<tr>
<td>5pm</td>
<td>10%</td>
<td>12%</td>
<td>15%</td>
<td>4%</td>
<td>23%</td>
<td>0%</td>
<td>4%</td>
<td>23%</td>
<td>0%</td>
</tr>
</tbody>
</table>

The information from Table 1 generally correspond to SW-smoothed values, as the statistics in [20] and [21] is obtained from a large sample of surveyed customers, representing mean demands for selected hours and loading conditions. This information and results from Section 4 are used to specify analytical relationships for temperature-demand correlation of heating and cooling loads over the course of the year n (9):

\[ P_{\text{heat,hr}} = P_{\text{base,heat,hr}} + C_{\text{heat,hr}} \cdot \beta_{\text{hr}} \cdot (T_{\text{hr}} - T_{\text{max,hr}}) \]  
\[ P_{\text{cool,hr}} = P_{\text{max,cool,hr}} - C_{\text{cool,hr}} \cdot \beta_{\text{hr}} \cdot (T_{\text{hr}} - T_{\text{max,hr}}) \]

where: \( P_{\text{heat,hr}} \) and \( P_{\text{cool,hr}} \) are heating and cooling demands at hour \( (hr) \) 3am or 5pm on the \( i \)-th day, \( i = 1, \ldots, 365 \) (negative values are equalled to zero), \( P_{\text{base,heat,hr}} \) and \( P_{\text{max,cool,hr}} \) are base/minimum heating demand (2nd column in Table 1) and max-
imum cooling demand (3rd column in Table 1), $C_{\text{cool, hr}}$ and $C_{\text{cool, hr}}$ are adjustment coefficients (for adjusting values calculated in Section 4 with values from Table 1, $C_{\text{cool, 3am}} = 0.75$, $C_{\text{cool, 5pm}} = 0.7$, $C_{\text{heat, 3am}} = 0.6$, $C_{\text{heat, 5pm}} = 0.4$). Beta coefficient is value of Beta coefficient on the $i$-th day at hour $hr$, $T_{\text{hr}}$ is $i$-th value of SW-smoothed temperature at hour $hr$, and $T_{\text{max, hr}}$ is maximum SW-smoothed temperature at hour $hr$.

The values of decomposed thermal cooling and heating demands at 3am and 5pm calculated by (9) for the whole year are plotted in Figs. 5 and 6, respectively.

![Annual Active Power Variations at 3am (Per-unit)](image)

**Fig. 5.** SW-smoothed values of decomposed cooling and heating demands at 3am, compared with the total active power demands from Fig. 3

![Annual Active Power Variations at 5pm (Per-unit)](image)

**Fig. 6.** SW-smoothed values of decomposed cooling and heating demands at 3am, compared with the total active power demands from Fig. 3

6 Conclusions

Some of the preliminary results for the load decomposition of the thermal electricity demands in Scotland, UK are presented in this paper. The results provide a more detailed analysis of the dependency of heating and cooling demands on temperature variations over the different temporal scales (annual, seasonal, daily and hourly), allowing to study these effects in the context of “smart grid” DSM applications.

The results are of particular importance for the implementation of “smart grid” DSM functionalities in the residential and commercial sectors, as they may provide (close to) real-time estimation of the expected heating and cooling demands based on weather/temperature forecasts, or for the long term planning of the electricity sector.
7 References