

Statistical analysis of driving factors of residential energy demand in the greater Sydney region, Australia

H. Fan^a, I.F. MacGill^b, A.B. Sproul^a

^a School of Photovoltaic and Renewable Energy Engineering,
University of New South Wales, Sydney, NSW, 2052, Australia

^b Centre for Energy and Environmental Markets and School of Electrical Engineering and Telecommunications,
University of New South Wales, Sydney, NSW, 2052, Australia

Corresponding author:

A.B. Sproul

School of Photovoltaic and Renewable Energy Engineering, University of New South Wales, Sydney, NSW 2052, Australia

Phone: +61 2 9385-4039

Email: a.sproul@unsw.edu.au

HIGHLIGHTS

- Study of drivers of residential electricity consumption in the Sydney region.
- Based on detailed demographic and housing survey of over 3400 households.
- Linear regression model shows good fit with aggregated household consumption.
- Findings highlight key contribution of ducted air-conditioning to consumption.
- Pools also a major contributor to average household energy consumption.

ABSTRACT

The residential sector represents some 30% of global electricity consumption but the underlying composition and drivers are still only poorly understood. The drivers are many, varied, and complex, including local climate, household demographics, household behaviour, building stock and the type and number of appliances. There is considerable variation across households and, until recently, often a lack of good data. This study draws upon a detailed household dataset from the Australian Smart Grid Smart City project to build a household electricity consumption model. A statistical linear regression model for household energy demand was established and tested for both individual households and regional aggregations of households. The model showed only reasonable performance in forecasting the consumption of individual households – highlighting the influence of factors beyond those surveyed – but good performance for aggregated household consumption. Models such as this would seem highly useful for a range of stakeholders including individual households trying to understand the potential implications of different choices, utilities looking to better forecast the impact of different possible residential trends and policy makers seeking to assist households in improving their energy efficiency through targeted policies and programs.

I. INTRODUCTION

Residential electricity consumption is responsible for approximately 30% of global electricity consumption [1]. Hence the residential energy sector will play a critical role in the future of the electricity industry, especially given the increasing global demand for affordable, secure household electricity services, as well as the urgent need to reduce climate change emissions from the electricity sector. The characteristics of residential electricity demand are very context specific with complex drivers including climate, demographics, housing stock, building types, household appliances and behavioural aspects. The respective influence of these is not well understood. There has also been considerable change in these elements over recent decades. In particular, more energy efficient technologies for lighting, communications, space heating and cooling, cooking, refrigeration and water heating have advanced rapidly in the last decade. Along with more energy efficient building standards and other energy efficiency oriented policy efforts, these developments seem likely to have contributed to falling residential electricity demand in a number of locations over recent years [2]. Australia provides a notable example with, for the first time in over a century, decreasing electricity demand in the residential sector since 2010 [3]. Beyond energy efficiency advances, this reduction has also been driven by the

increasing adoption of rooftop PV systems (now present on around 15% of Australian houses) and considerable increases in electricity prices [3].

A better understanding of how various factors influence residential electricity demand can assist in understanding possible future developments in the sector, as well as assisting in identifying opportunities to improve outcomes through targeted household and broader policy efforts. For example, such information can provide guidance to policy makers on the impact of different housing and household trends on local residential electricity demand and assist in forecasting the potential impacts of planning changes, housing retrofits and use of new energy efficient appliances under different possible government policy measures. Electricity utilities could use such insights to improve their planning and operational processes, while households could also benefit in better managing their electricity costs through an improved understanding of how decisions about what housing and appliances they choose can impact on their electricity bills, and what opportunities they might have to reduce consumption.

However, achieving an improved understanding of the nature of household electricity consumption is challenging, due to the heterogeneity of the residential sector, the complexity of the underlying drivers and the lack of comprehensive data. They include household demographics, including occupant numbers, age distributions, and income; household behaviour such as how often occupants use certain appliances and the interest and effort that they devote towards energy conservation; building types, appliances, such as the type of dwelling (free standing or unit), different appliance ownership and access to alternatives to electricity for some services such as gas for hot water, heating and cooking; and the climate zone of the households as well as the daily weather conditions. The wide variation seen across all of these drivers leads to considerable differences in households' electricity consumption. Furthermore, there has generally been only limited electricity consumption data available. For example, it has been common for households to have their electricity consumption recorded with simple accumulation meters that may only be read four times a year. Finally, this consumption has not been matched with detailed information regarding the houses and households. Where detailed data is available including household surveys covering the drivers above, and even appliance level metering, this has generally been for only a small number of households.

These challenges of complex drivers and poor data have posed significant challenges for reliable and useful residential electricity demand modelling. Using aggregated or partial data consisting of either social economic information or appliance ownership to model residential electricity consumption is insufficient for many of the potential applications of such models. Hence a more sophisticated modelling approach combining all aspects of household characteristics can add considerable value.

In this paper we report on a demand modelling study that could be used to forecast average daily electricity demand based on data from Australia's first large- scale smart grid project – the Smart Grid Smart City (SGSC) project. Amongst other activities, the project collected a year of half-hour demand data for more than 9000 households, and where available, solar generation and air temperature data. Around 40% of these households were also surveyed to establish house type and appliances, and household demographics [4]. We use this data to establish a bottom-up statistical model of residential electricity demand for these households.

More than 1.9 billion records from more than 9000 households over a one-year period have been sampled and analysed in our study [4]. This set of data has several novel elements in comparison with existing data sets in Australia. For example, the smart metered half hour electricity consumption readings from households are matched with half-hourly weather data at the same location (weather data from the closest weather station to the household). In addition there is a broad range of survey data for houses involved in the trial such as the dwelling type, number of occupants and their age groups, details of household appliances including those running on, gas, number of refrigerators, presence of air-conditioning and its type, presence of clothes dryers, and some self-identified behavioural indicators such as the household's frequency of clothes dryer usage and their intentions regarding saving electricity. As such, this survey data covers many of the driving factors of demand previously identified, including at least some aspects of household behaviour, demographics, building type, infrastructure and appliances, as well as climate.

The temporal resolution, large sample size and detailed household characteristics of this dataset present a unique opportunity for better understanding Australian household electricity consumption. The study presented in this paper aims to provide new insights into how a range of driving factors impact on residential

electricity demand. The rest of this paper is organised as follows. Section 2 delivers a brief literature review on residential electricity modelling in different regions as well as its techniques. Section 3 introduces the data available from the SGSC project, followed with data treatment methods. Section 4 presents some preliminary, high level, analysis of the data, while section 5 details the development of the detailed statistical model. The results of the aggregated model fitting and validation are illustrated in section 6. Discussion and suggested future work are given in section 7 while section 8 presents the conclusions of this paper.

II. PREVIOUS WORK

To date, there are three major modelling approaches for residential electricity consumption. Notably, they are all critically limited based on the availability of data. These three approaches are, respectively, the top-down approach, which focuses on the interaction between electricity consumption and economic metrics at a high level scale using aggregated socio-economic data; the bottom-up approach, which statistically analyses household survey data and electricity consumption readings; and the physical model approach, which models physically measured data on specific dwellings, appliances and technologies [5]. All three approaches have their strengths and weaknesses, due to the differing nature of their input data and assessment capability. Top down models are mostly high level studies and analyse highly aggregated data at a national scale. The majority of papers focus on analysing the socio-economic impacts of the electricity sector [6] [7]. Alternately, bottom up modelling utilises disaggregated data to estimate the impact of various factors on electricity consumption [8]. Some bottom up approaches use samples of houses' building physics to represent larger housing stock [9], combining building electricity calculations with statistical methods.

A considerable number of international studies have focussed on better understanding household electricity demand. As such the review presented here can only select a few sample studies and these are listed by the modelling approach used in the section below.

A. Top down:

A study in Pakistan [10] used a top down approach to forecast annual residential electricity consumption using a multiple linear regression model. The study utilised a correlation matrix to analyse the sensitivities and relationships between aggregated power demand and a set of demographic variables. It utilised a univariate time series in conjunction with econometric models to estimate the electricity demand of Pakistan for the next 15 years. The analysis of this model found that GDP, income per capita and population have a key impact on electricity demand. Validation was done by comparative analysis with historical real data and projections of other regression models. A study in Italy, [11] tested various regression models based on available data from 1970-2007 regarding national electricity consumption. A regression model was proposed based on the elasticity analysis of the different demographics' effect on domestic and non-domestic electricity consumption. Another study, [12] provided evidence of the impact of population aging on electricity consumption in Italy using results obtained from a calibrated overlapping generations' general equilibrium model. The study found that electricity consumption is particularly sensitive to, and increases with, an aging population.

B. Bottom up:

In Canada, a study used a bottom up approach utilising the Canadian Hybrid Residential End Use Energy and Greenhouse Gas Emission Model (CHREM) – a statistical method to analyse national residential energy consumption [13]. The model takes into consideration appliances, lighting, and domestic hot water to forecast national annual energy consumption. This study used the national Survey of Household Energy Use (SHEU) undertaken in 1993. Through the use of artificial neural networks, this study estimated the energy consumption from domestic hot water, appliances and lighting. However, this study was limited due to the use of high level data (national level), and the reliance of the model on national survey data from 1993, which raises the question of the validity of the model due to the age of the data.

In a Chinese study, energy simulations were conducted to predict the future electricity demand of urban residential buildings in Chongqing [14]. However, the study sampled primary household electricity demand by a structured questionnaire and subsequent energy intensities of various drivers were collected only from the literature concerning simulation of annual energy consumption. Electricity consumption of residential buildings in various cities were also analysed to investigate the driving factors for summer and winter [15] residential electricity demand. The data for this study covered more than 6 different cities and 5 different climate zones. Household characteristics, possession and utilization of energy appliances, and the climate

zone were collected through questionnaires, and electricity and gas consumption data were collected on a monthly basis. An electricity consumption comparison for old and new residential buildings in Shanghai has also been investigated [16] through statistical analysis. In Hangzhou [17], a residential energy consumption survey study was analysed and revealed that household socio-economic status and behaviour could explain 29% of the variation in heating and cooling related electricity demand for different households.

In the USA, a study [18] investigated the impact of residential occupants behaviour on electricity consumption using hourly appliance-level electricity consumption data for 124 apartments over 24 months. This research showed that 25-58% of the variation is due to behaviour induced electricity consumption for different households.

C. Combined top down and bottom up:

In the United Arab Emirates, a study [19] presented a mid-term forecast model for urban electricity demand in Abu Dhabi using a combined top-down and bottom up approach. The study aimed to isolate the impact of air-conditioning through a time series hourly regression model, then used the urban planning council's load constituent survey data to estimate the urban cooling load.

D. Potential strength and weaknesses of these approaches

The top-down approach provides a focused analysis that considers the links between the economy and the energy sector at a high level, but generally lacks technological detail and therefore cannot examine technology related impacts. The bottom-up approach generally provides a good understanding of the technological drivers of electricity consumption, however it requires a large sample size and typically relies on reliable historical consumption data, which is not always available.

E. By modelling methods

Various traditional and emerging modelling methods have been extensively utilised for demand management and electricity demand forecasting. Models used include time series models, regression models, neural networks models and complex hybrid models.

Time series models have been developed for jurisdictions including Israel [20], Northern Spain [21], Turkey [22] [23], the UK [24], and India [25]. Time series models also include those dedicated to short term forecasting [26], [27], [28], together with medium term forecasting models [29], [30], and long term predictions [31], [32], [33].

Regression models are one of the simplest and most traditional methods, which have been used both for short term forecasts [34] and long term forecasts [35], as well as peak load forecasting [36].

Stochastic models are good options for more complex systems, some bottom-up approach have been used to forecast domestic electricity demand [37] [38] [39], and demonstrated its ability for modelling high resolution energy demand.

Machine learning model [40] [41] and neural networks have been a popular approach in the last few decades for electricity forecasting, [42] [43], [44]. However, such analysis commonly acts somewhat as a "black box" and fails to provide information about potential cause-effect relationships between explanatory factors.

Others use hybrid models such as the hybrid approach of machine learning models [45], a neural network approach integrating a regression model [46].

F. In Australia

In Australia, a recent study [47] presented a statistical model to analyse residential Census Collection District level electricity consumption at the district level across 8,883 records for the year 2006 in NSW. The model combined local demographic information as well as climate zone data to analyse their impacts on energy consumption. Its main contribution and application is targeted to the prediction of electricity consumption variations under two future scenarios: shifting climate zones and growing population. Another top-down approach [48] forecasts Australian national level electricity demand using semi-parametric additive models, however, due to the nature of its high level modelling, it could only focus on calendar, temperature, and lagging effects, where limited its contribution to inform policy makers and households appliances purchasing decisions.

As a part of the findings in CSIRO's evaluation of 5 star energy efficiency standard residential buildings [49], it highlights the difficulty to draw robust conclusions regarding differences in electricity usage between high rating housing and low rating housing. The reason is mainly due to the dynamic range of household characteristics, as well as the lack of sample size on electricity consumption during different weather conditions, household type, occupancy, behaviour, etc. Also, a study in Australia [50] simulated scenarios of future climate, and indicated such changes could significantly increase heating and cooling electricity consumption in the Australian residential sector.

G. Limitations and Gaps

As discussed above a common limitation in existing research is the availability of data. Most studies try to utilise a bottom up approach at a very generic level using aggregated electricity consumption for a district or even a whole country with low time resolution (e.g. a year). Previously, a common practice from both top down and bottom up approaches was to analyse electricity consumption data together with disconnected survey information [13] [14]. As a consequence, electricity and survey data were not always matched for individual households nor were the time frames of the data sets matching. This introduces significant uncertainty and unreliability into the models and hence clouds resulting insights of how different drivers actually impact residential electricity demand. To date, most of the researchers (e.g. [15], [17], [48], [50]) focus on climate impact on residential electricity demand, but limited understanding has been achieved for the drivers of residential electricity demand from the household characteristics area.

In Australia, the energy market is entering a dynamic situation, where residential electricity consumption is expected to fall [51], and the recent emergence of smart meters brings a potential conservation benefit [52]. Under this dynamic energy market in Australia, It is even more important for us to understand what drives electricity consumption, so that policy makers and other stakeholders can make informed decisions. There have been data limitations that have restricted opportunities for researchers to undertake such bottom-up modelling, which enable the projection of future scenarios. The comprehensiveness and resolution of the SGSC data provides a unique opportunity to allow statistical modelling of key driving factors for residential electricity demand, and hence, implications for residential electricity consumption can be investigated.

III. DATA COLLECTION AND TREATMENT

A description of the underlying dataset used for this study is provided in the Smart Grid Smart City reports [53] [4]. Survey data and half-hour interval electricity readings were collected from 9903 households participating in the study, from 2010 until 2014. The data is gathered from households in 6 major towns or regions (2019 households in Lake Macquarie, 1980 in Newcastle, 1689 in Ku-Ring-Gai, 1577 in the Sydney CBD, 1439 in Auburn, and 469 in Cessnock) and 730 smaller towns scattered across NSW. They are located across four regions, namely the Upper Hunter, Hunter, Central Coast, and Sydney as illustrated in Figure 1.

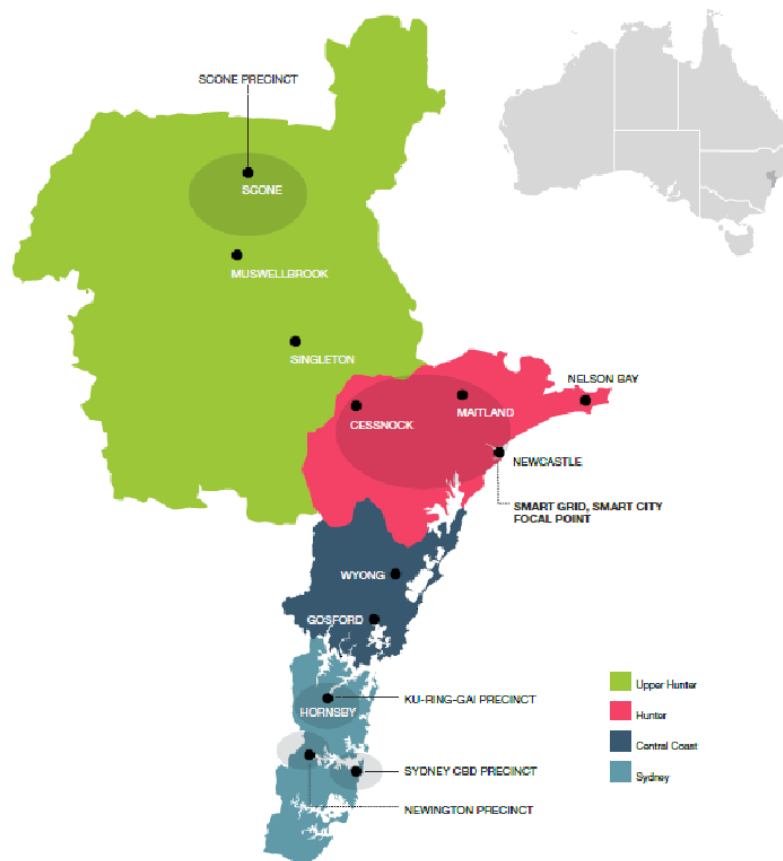


Figure 1: SGSC Project geographical coverage [54]

The households involved in this study cover a wide range of demographics, as well as being geographically diversified. Participating households were selected from an initial call for ‘expressions of interest’ in these targeted areas. However, there was no formal process for establishing a statistically representative sample of NSW households. Households who participated in this trial had no obligation for any of the cost of the new metering devices that were required.

The Smart Grid Smart City (SGSC) trial contains two main data sets. The first dataset contains all the half hour interval readings including general electrical supply, controlled load (e.g. off peak hot water), local air temperature, local apparent air temperature and photovoltaic gross generation (when a solar system was present) for each household. Note that households with net metered PV Solar systems were excluded since their measured electricity consumption is influenced by self-consumed PV generation. The second dataset includes a wide range of demographic and other household information collected during the interviews of the participants.

For this study, the two datasets are sampled for the same time period and are matched for each household to ensure alignment. Looking at the first dataset alone, from July 2012 until July 2013, 200 million plus readings were collected on a monthly basis. Since the quality and accuracy of electricity modelling is heavily dependent on the validity and reliability of the household information data and household electricity readings, only data for the financial year of 2013 (FY2013) has been utilised. Earlier years had less, and less consistent, data given the time taken for the trial to be implemented across nearly 10,000 houses.

Among the 9000 metered households, less than half 4000 of the households were interviewed. Only interviewed households with continuous data sets covering this period, totalling 3446 households were included for analysis and investigation.

During the trial period, a list of more than 30 pre-determined questions was asked of each interviewed household to identify corresponding household factors such as: income, appliance usage, historical electricity consumption, and building type. Some surveyed factors were not considered relevant to this study. For those

remaining, we tested their significance (p value as explained further in Section 5) as drivers for household electricity demand. The complete list of factors that were considered and their relevant definitions for the purposes of the SGSC project are tabulated in Table 1.

Table 1: Factor Description [4]

Factors	Descriptions	Distribution
Home ownership	Whether the dwelling is rented or owned by the occupying household	79% owned the home they lived in, either outright (40%) or mortgaged (39%)
Air-conditioning type	Air-conditioning installation type. E.g. Ducted, Split System, Other (any air-conditioned system that is not Split or Ducted such as a portable or window system) and None	19% Ducted, 50% Split-System, 3% Other and 28% None.
Clothes Dryer Usage	A relative measure of how much the clothes-dryer is used. (Infrequently - NONE, Monthly - LOW, Weekly - MED, Daily - HI, and NULL indicates the question was not asked, or was not answered)	42% used infrequently, 27% monthly, 25% weekly, and 6% for daily.
Dwelling type	Unit, semi-detached, separate house.	13% unit, 2% semi-detached, 85% separate house.
Natural gas connection	An indicator if the household has natural gas connected.	53% have natural gas connected
Gas hot water	An indicator if a household has a gas hot water system.	53% have gas hot water
Gas heating	An indicator if a household has a gas heating system.	30% have gas heating devices.
Gas other	An indicator if a household has other gas appliances.	5% have other gas devices.
Effort to save energy	How much effort households self-reported to save electricity	51% a lot, 36% a little, 8% not much, and 5% none.
Internet	An indicator if a household has access to the internet or not.	79% have Internet access
Pool pump	An indicator if the household has a pool and hence pool pump installed.	23% have a pool pump
Solar	An indicator if the household has solar panel installed (photovoltaic or solar hot water)	10% have solar devices such as solar hot water or solar panels
Income	An indicator for the household weekly income category (<\$800 (Low), \$800-\$1200 (Med), >\$1200 (High)).	28% Low, 34% Med, and 38% High.
Occupants 70plus	Number of occupants aged beyond 70 years old normally resident at the household	17% have people older than 70 years old
Children 0_10	Number of children aged from 0 to 10 years old normally resident at the household	11% have young children
Children 11_17	Number of children aged from 11 to 17 years old normally resident at the household	27% households have older children
Occupants	Total number of occupants normally resident at the household location	The average size of trial households was 2.8 people. The largest proportion of respondents came from two-person households (35%), followed by four- person and then three-person households (21% and 17%) respectively
Refrigerators	Total number of refrigerators	53% have only 1 fridge and 35% have 2 fridges with 10% with 3 fridges and 3% had more than 4.

A unique key code is used to identify and distinguish each household in both datasets mentioned above. Due to the structural differences of the two datasets, various treatment processes need to be carried out before using the data for further modelling and analysis, as outlined in Table 2.

Table 2: Data treatment process

Aggregation	<p>All 3446 interviewed households analysed.</p> <p>INTERVAL_READING Data is first filtered to a specific time span only covering FY2013 (1 July 2012 to 30 June 2013). Then the data for each household has been aggregated into an average household electricity demand (general supply and controlled load) over FY2013 (48x365 intervals) on a daily basis (kWh/day).</p>
Relation	<p>Aggregated data consists of average household electricity demand. Household information Data (TRIAL_PARTICIPANT) is used to combine with existing aggregated data using Household IDs.</p> <p>Yearly heating and cooling degrees half hours are aggregated at the same time by treating the half hourly temperature value for each household. Each half hourly temperature value is compared with 24 and 18 degrees (this value is selected based on the comfort level temperature based on Australian Bureau of Meteorology practice [55]). The number of half hour degrees lower than 18 and higher than 24 degrees are summed for each household to generate Cooling and Heating factors respectively for the investigated year.</p>
Indicator Variables	<p>Each household factor was assigned appropriate indicator variables. E.g. INCOME was assigned 3 indicators:</p> <p>INCOME_Low, INCOME_Med, INCOME_High.</p> <p>Each has value 1 or 0 for each indicator.</p> <p>This data set is used particularly for the forecasting analysis of aggregated data at a zone level.</p>

IV. PRELIMINARY DATA ANALYSIS

Preliminary data analysis was carried out in order to ascertain which surveyed factors appear to significantly influence average annual household electricity demand. Of the 3446 households studied, the household with the maximum annual demand consumed 45.74 MWh per year, or approximately 125 kWh per day. The household with the minimum electricity demand consumed 364.5 kWh per year or approximately 1 kWh per day (this may be caused by the house being empty for some period of the year). Figure 2 presents a probability density histogram plot of annual average daily, total electricity demand of the households in this study. The mode, median and the average on a daily basis are 11.5 kWh, 16.8 kWh and 19.2 kWh respectively. The data is asymmetrically distributed with a tail at high electricity demand; hence, a log-normal curve is used to fit the data.

Based on reported data from IPART [56], and information from two major network operators in NSW (EnergyAustralia and Integral Energy), annual average residential electricity demand in NSW is 18.9 kWh per day. This is reasonably close to the average (19.2 kWh per day) of the studied households suggesting they may represent a reasonable distribution of typical households in the greater Sydney area where the majority of people in the state of NSW live.

The first step of this preliminary data analysis was to plot the probability density distributions of average household demand according to some of the surveyed factors as outlined in Table 1 - for example, household income (Low, Med and High), number of occupants and the presence of air-conditioning (Split-system, Ducted, Other and None). Figure 2 presents the probability density distribution of the average annual daily electricity demand for the surveyed households.

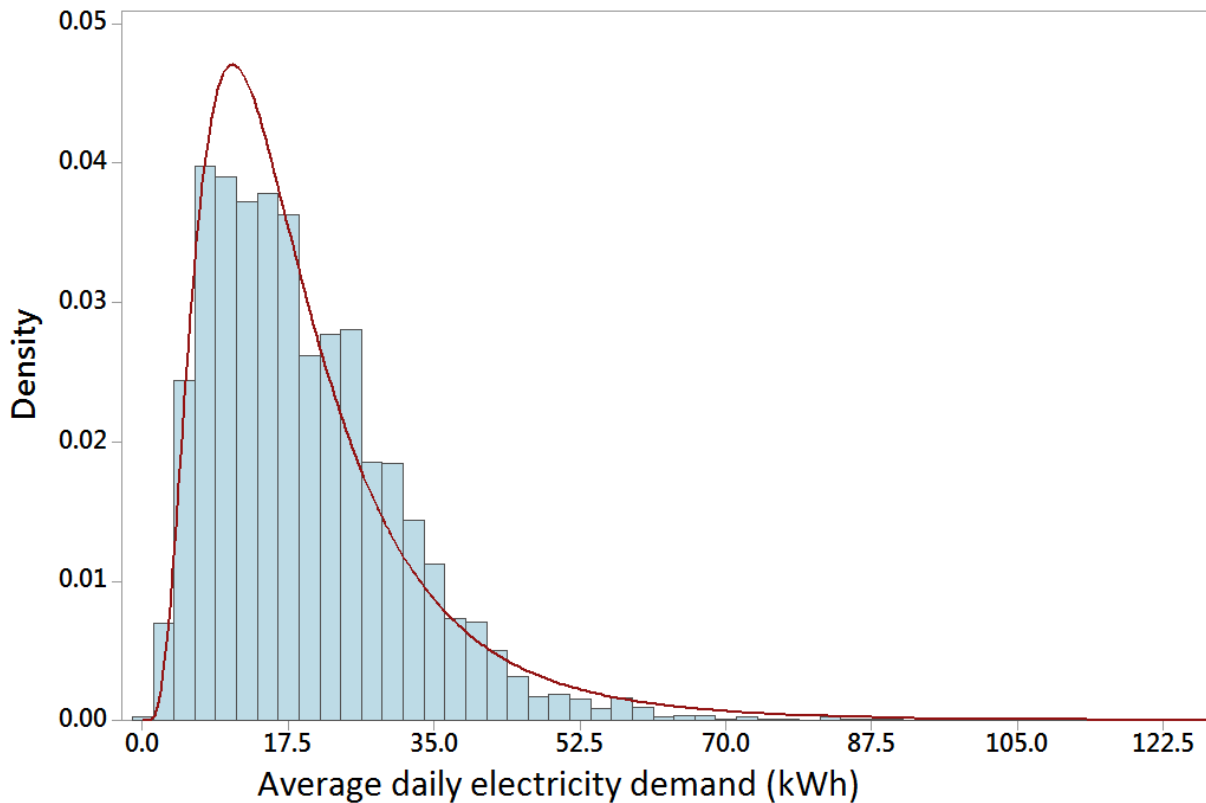


Figure 2: Probability density distribution of the average annual daily electricity demand for the surveyed households.

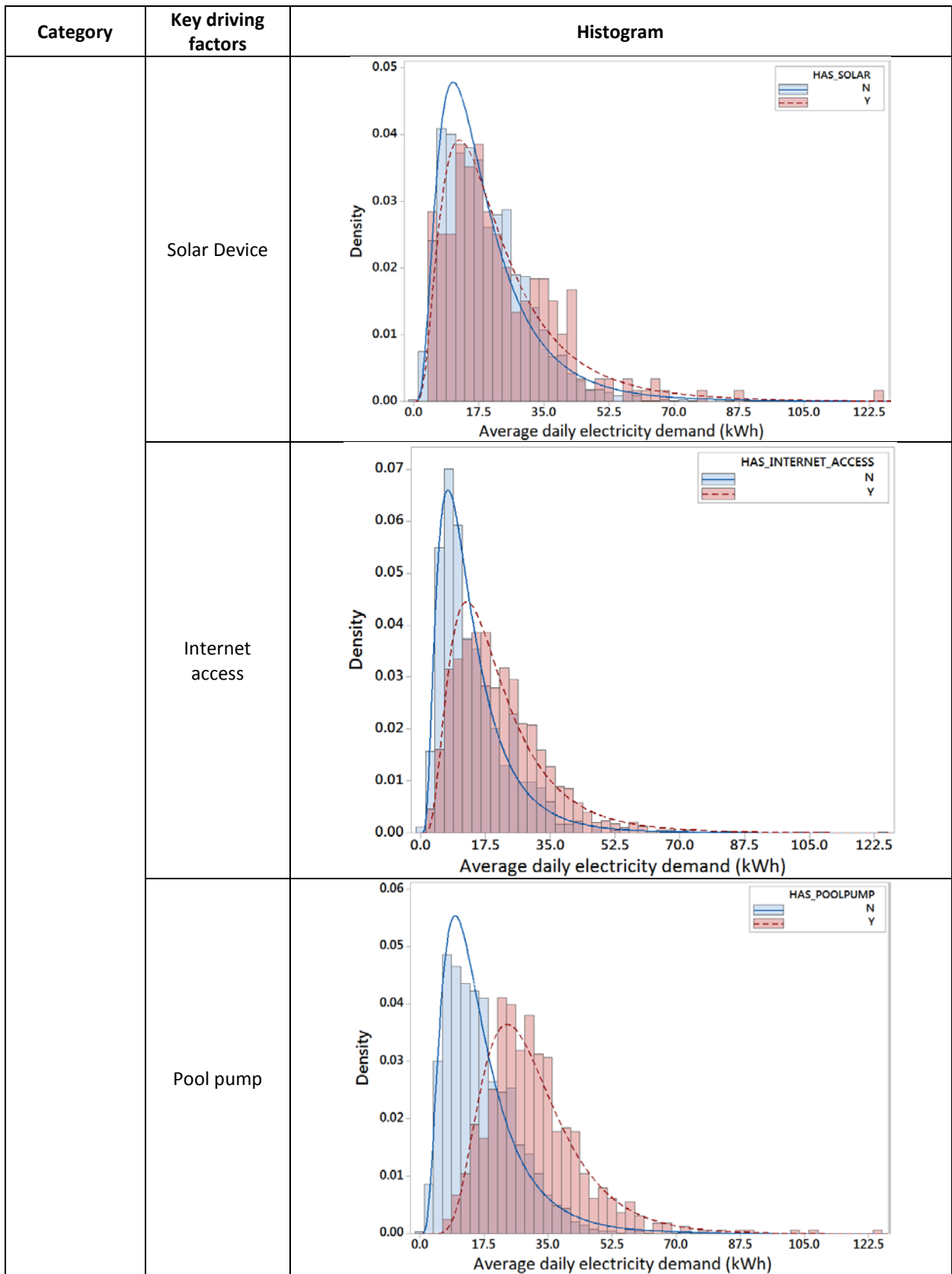
The annual average daily electricity usage over 3446 sample households, as a function of various key factors is shown in Table 3. The graphs in the table present normalised probability densities (bar height x bar width gives the probability for each range) against average annual household electricity consumption. Fitted log-normal distributions for these probability density functions are also shown. Some of the figures with multiple overlapping categories have density bars removed and only present the best log-normal fit to better illustrate the data.

Table 3: Distribution of key factors in daily household demand

Category	Key driving factors	Histogram
Household Behaviour	Dryer Usage	

Category	Key driving factors	Histogram
Household Demographic	Income	
	Occupants	
Building & Appliances	Gas hot water	

Category	Key driving factors	Histogram
	Dwelling type	<p>DWELLING_TYPE_CD SeparateHouse Unit</p>
	Air-conditioning	<p>AIRCON TYPE_CD Ducted SplitSystem None</p>
	Gas	<p>HAS_GAS N Y</p>



From the surveyed SGSC household data, there are clearly significant electricity consumption differences between households according to some of the factors. For clothes dryer usage there is no significant average electricity demand difference between households with “LOW” (20.4 kWh \square = 11.3 kWh) and “MED” “Dryer

Usage" (22.1 kWh σ = 14.1 kWh), however households with "HI" usage (28.3 kWh σ = 13.5 kWh) have an average consumption approaching double that of households that do not have an electrical dryer (15.6 kWh σ = 9.5 kWh). With regards to "Income", mean values for each income group are: "HI" (23.6 kWh σ = 14.2 kWh), "MED" (18.9 kWh σ = 9.6 kWh), and "LOW" (14.1 kWh σ = 8.6 kWh). As would be expected, electricity consumption increases with the number of occupants. Ownership of gas hot water shows a decrease of electricity demand as is the case for ownership of gas in general. With regards to " Dwelling Type", "Separate Houses" have on average twice as much electricity demand (22.3 kWh σ = 12.0) than "Units" (11.2 kWh σ = 6.5). How much of this relates to the typically greater floor space of separate houses and how much to other factors is, however, not possible to ascertain from the available data. Almost three quarters of the surveyed household have "Air-conditioning" systems installed in their home, and primarily are smaller split systems (70%) or ducted systems (26%). There is a clear difference in daily electricity demand between households with and without air-conditioning systems, and, for those that do, by type of system: households with "Ducted" air conditioning (27.3 kWh σ = 16.3 kWh) use on average 79% more electricity than households with no air-conditioning – "NONE" (15.2 kWh σ = 10.1 kWh); while households with a "Split System" (20.4 kWh σ = 10.4 kWh) on average consume 34% more electricity than households with no air-conditioning. From Table 3, it is evident there is an increasing trend of household electricity demand as household occupant numbers increase. Houses with more than 5 occupants are grouped together given their relatively small sample size, and hence challenges for statistical analysis. The category "solar device" in the data indicates households who have either gross metered solar PV or solar hot water, as the SGSC project survey does not distinguish between the two. Note that households with gross metered solar PV; their electricity consumption is metered separately from their solar PV generation. Therefore, not much difference in electricity consumption is seen from this factor. Houses with internet access (21.1 kWh σ = 12.2 kWh) show 51% higher electricity demand than houses without internet access (13.9 kWh σ = 9.1 kWh).

"Pool pump" appears to be a significant indicator of household electricity demand as shown in Table 3. About 15% of sampled households had a pool (and hence pool pump) and the annual average daily electricity demand of these households (31.2 kWh σ = 13.6 kWh) was 93% higher than those without (16.9 kWh σ = 9.8 kWh). In part, of course, this highlights the significant power draw and running hours of pool pumps. However, it also highlights a complexity in the analysis. Is the higher load for houses with pools related, in part at least, to other factors such as income levels, and the number of occupants (particularly children)? In Figure 3, household electricity demand is shown for households with and without a pool pump for different household income levels (Low, Med, and Hi). It can be seen that households with a pool pump consistently have higher electricity demand than households without, regardless of income level. As noted earlier, however, income levels certainly also influence demand. Driving factors for household electricity demand are multifaceted and this illustrates the importance of analysing their impact on household demand collectively. The regression model we apply towards this aim is described in the next section.

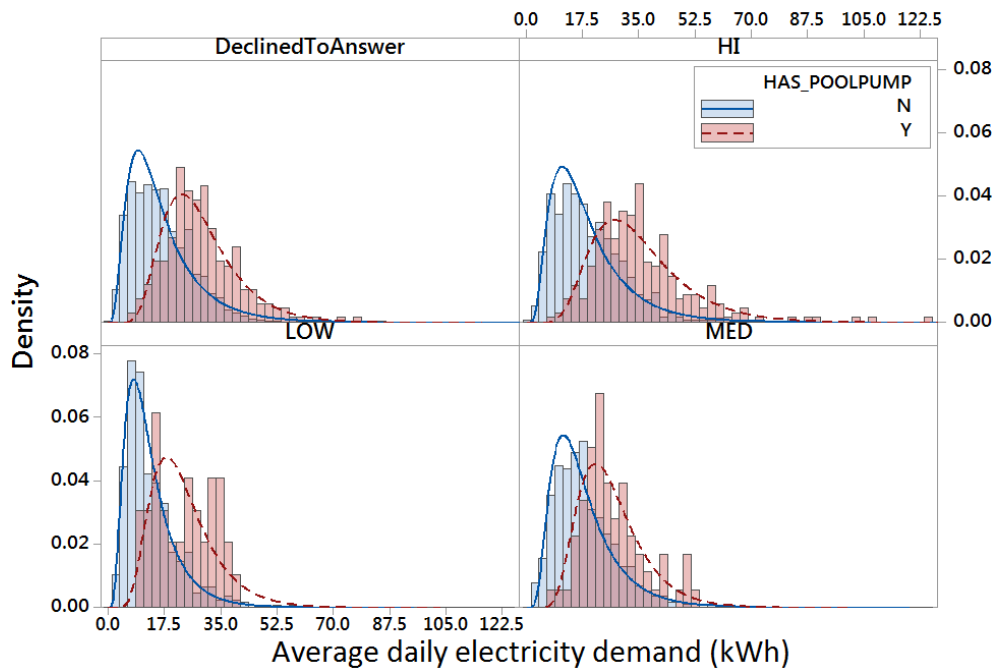


Figure 3: Household Income group (Declined to answer, High, Med, Low) and Pool pump impact on annual average daily electricity demand (Has pool pump – Red and dashed line; does not have pool pump – blue and solid line)

Finally, temperature is also a significant driver of household electricity consumption. The SGSC data set provides local temperature on a half hour basis for each household (derived from the nearest weather station). With comfort zone set-points of 18 and 24 degrees (heating and cooling respectively), the average daily CDD (Cooling Degree Days) and HDD (Heating Degree Days) across all the surveyed houses are calculated for FY2013 and shown in Figure 4. Heating and cooling degree days are based on the average daily temperature, which is calculated as follows: $0.5 \times (\text{maximum daily temperature} + \text{minimum daily temperature})$ [55]. The annual household cooling degree hour data ranges from 22 to 1556 degree hours, while annual household heating degree hour data ranges from 251 to 5514 degree hours across the surveyed houses. This highlights the rather different micro-climates experienced across the surveyed households with inland households experiencing colder winters and hotter summers than those households located near the coast.

Figure 4 shows that the average CDD and HDD profile remarkably resembles the daily average electricity demand trend. Note that average annual heating and cooling degree days are calculated by summing heating and cooling degree days for FY 2013, which are 685 HDD and 39 CDD respectively for all sampled households. This is caused by the average ambient temperature in winter being further from comfort levels, than in summer where the average ambient temperature is close to comfort levels with only occasional extreme heat driving up electricity demand.

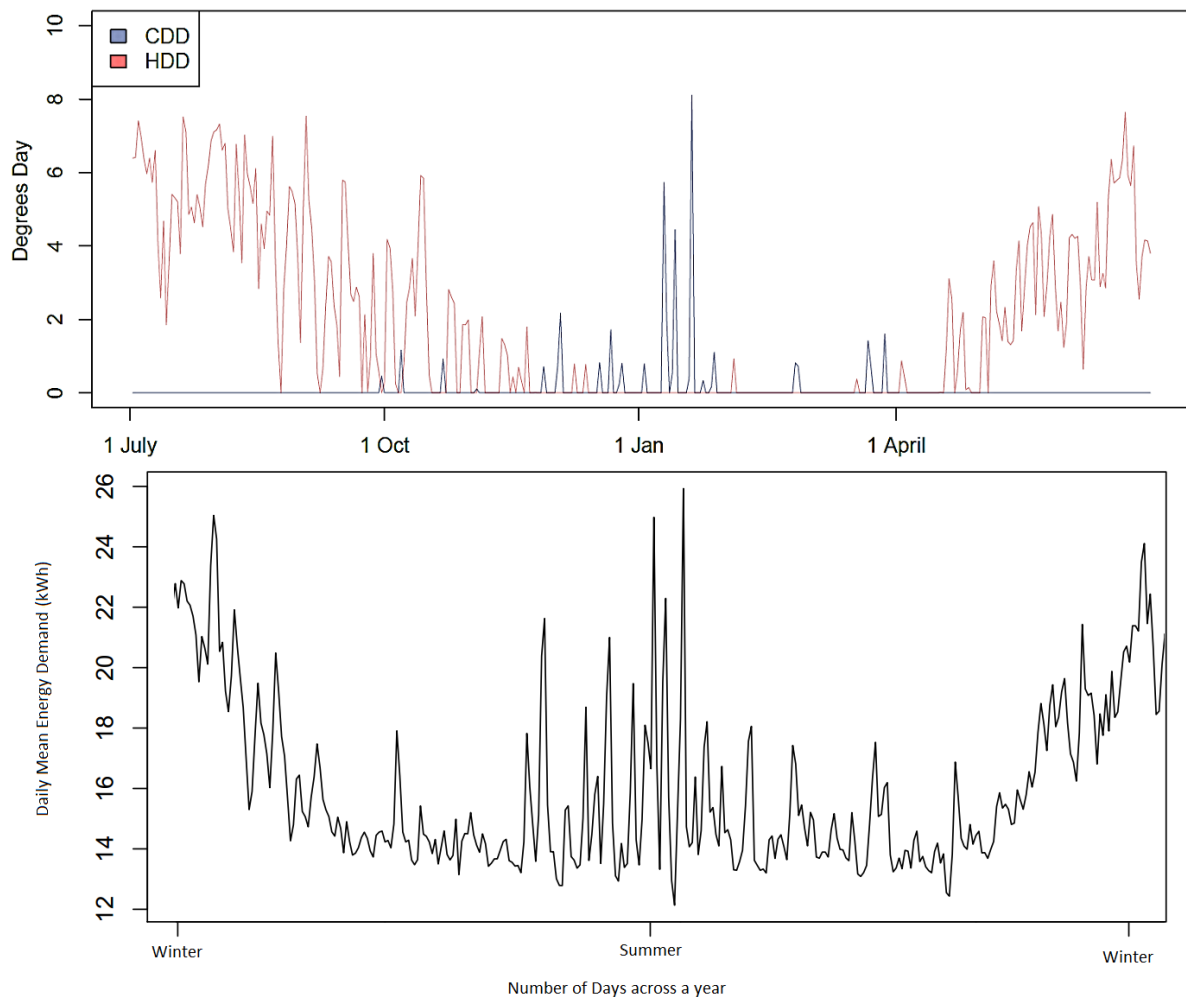


Figure 4: Average Daily CDD, HDD for sampled households in FY2013 (top) and Daily average household electricity demand for all across the year of analysis (bottom)

The plot clearly reveals that the average electricity demand peaks at summer, but only spikes for a few days and the rest of the time demand stays relatively low. During winter, by contrast, most of the days have continuous high average electricity demand. Comparing average CDD, HDD and daily electricity demand, it is clear that heating related electricity demand dominates over cooling demand on an annual basis.

Due to the fact that heating and cooling requirements have a high impact on residential electricity demand, it is also important to pay attention to air-conditioning systems. There are, however, some complexities to consider as air-conditioning is the key load for cooling (the survey did not ask about fans), yet many of these units can operate in the reverse cycle as well to provide heating. They are indeed highly efficient in this role. However, heating can also be provided with conventional electrical bar and fan heaters (which are far less efficient than typical reverse cycle air-conditioners) as well as through gas heating. These issues are explored further in the Discussion section.

V. MODEL DEVELOPMENT

Bottom up statistical methods were applied to analyse the significance of each identified factor towards total annual electricity demand of the surveyed households. Then based on selected predictors, the model equation was developed. Furthermore, in the model we use, factors are considered to be independent. Interactions between factors were considered in the initial stages of model development, and we tested this for a number of factors. However, the interactions had low statistical significance, and hence were removed in the final model. Of course, the use of linear regression techniques requires a number of significant assumptions. Even when these assumptions don't all hold, the technique can still provide a useful model.

H. Factor selection

All surveyed household characteristics were initially analysed using *R project* [57] and *Minitab 17* [58] to assess if they were significant driving factors using two separate methods. For the first method, all suggested factors are added into a general linear regression model (GLM), and factors that do not pass a 95% confidence interval significance level test (that is, there is at least a 95% likelihood that the factor influences total (general supply and controlled load) household demand) are removed one by one from the model, until all factors satisfy the P-value test. In the second approach, a stepwise regression is performed to incorporate factors one by one and automatically remove insignificant factors while retaining those that are significant. Amongst all of the available factors sampled from the SGSC project, 14 different factors were finally chosen as important predictors for the modelling. A P-value test validates that all factors selected are significant with a 95% confidence interval as shown in Table 4.

Table 4: P-value of all selected factors

Source	P-Value
OCCUPANTS	0.000
CHILDREN_0_10	0.000
CHILDREN_11_17	0.022
REFRIGERATORS	0.000
Heating	0.000
DWELLING_TYPE	0.000
INCOME	0.000
DRYER_USAGE	0.000
AIRCON_TYPE	0.000
GAS	0.001
GAS_HOT_WATER	0.000
SOLAR DEVICE	0.044
POOLPUMP	0.000
INTERNET_ACCESS	0.000

I. Model equation

The final model has 4 components: 1) household demographics, 2) household behaviour, 3) building and appliances, and 4) climate impact. Each component of the model has associated factors, as illustrated in Figure 5.

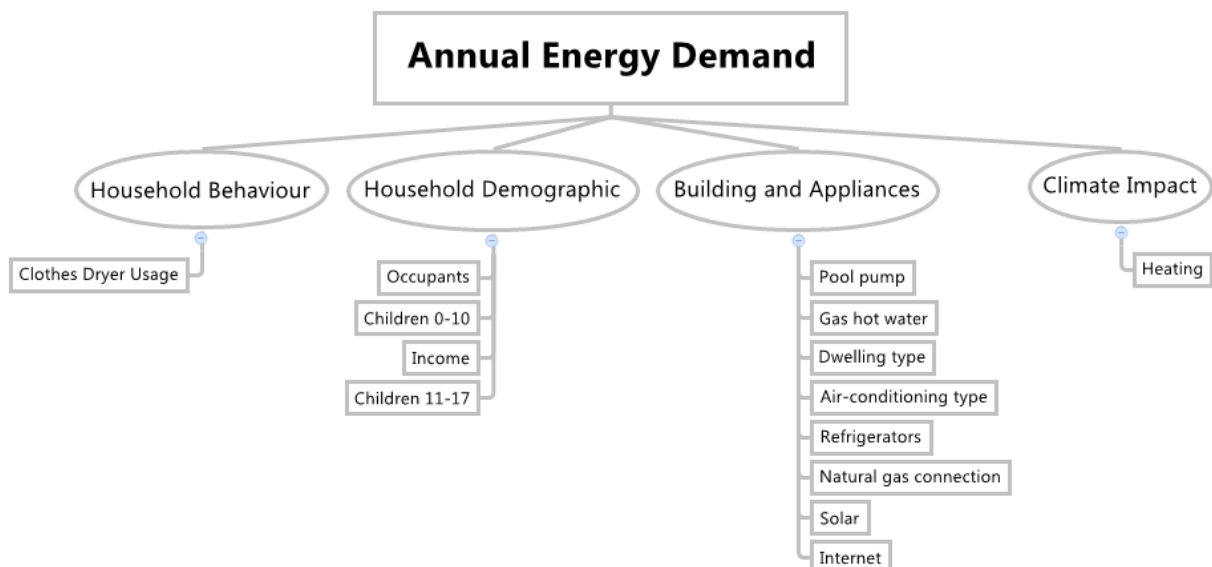


Figure 5: Model Structure

In order to model the average household electricity demand over FY2013 on a daily basis, the following equation was used:

$$Y_i = B_i + D_i + A_i + C_i + n \quad (1)$$

where Y_i is the average daily total electricity consumption (general supply and controlled load), over a year, for the i^{th} household. The term B_i represents Household Behaviour effects, D_i represents the Household Demographic effects, A_i represents the Building and Appliances effects, C_i represents the climate effect and n represents the error of this model. For each term in Equation 1 that describes a component of the model (B_i , D_i , A_i , C_i), there are a number of associated terms representing factors that influence electricity demand. For the household behaviour the term B_i is defined by:

$$B_i = \alpha_1 \cdot DU \quad (2)$$

where α_1 is the coefficient vector associated with dryer usage, and DU denotes *Clothes dryer usage* (none, low, med and high) for each household - a relative measure of how much the clothes dryer is used.

For household demographics the term D_i is defined by:

$$D_i = \beta_1 \cdot I + \beta_2 \cdot \text{NumC11_17} + \beta_3 \cdot \text{NumC0_10} + \beta_4 \cdot \text{NumOcc} \quad (3)$$

where β denotes the coefficient vector corresponding to various impact factors. The variable I denotes the *Income* of each household, NumC11_17 denotes how many children of ages 11-17 are in each household, NumC0_10 denotes how many children of ages 0-10 are in each household, and NumOcc denotes the number of occupants in each household.

For building and appliances the term A_i is defined by:

$$A_i = \gamma_1 \cdot \text{GH} + \gamma_2 \cdot \text{DT} + \gamma_4 \cdot \text{AT} + \gamma_5 \cdot \text{NR} + \gamma_6 \cdot \text{HP} + \gamma_7 \cdot \text{HG} + \gamma_8 \cdot \text{HS} + \gamma_9 \cdot \text{HInt} \quad (4)$$

where γ denotes the coefficient vector corresponding to various impact factors. The variable GH denotes *Gas hot water* for each household (an indicator if the household has a natural gas hot water system - values are Y/N), DT denotes *Dwelling type* of each household (unit, separate house), AT denotes *Air-conditioning type* of each household (Ducted, Split System and None), NR denotes *Number of refrigerators* for each household, HP denotes if a household has a *Pool pump*, HG denotes if there is a *Natural gas connection* at a household, HS denotes if there is *Solar Device* at a household, HInt denotes if there is *Internet* access at a household.

For the climate effect the term C_i is defined by:

$$C_i = \delta_1 \cdot \text{Heating} \quad (5)$$

where δ_1 denotes the coefficient vector corresponding to heating, and *Heating* denotes the number of "heating degree-half hours" relative to 18 degrees.

VI. MODEL FITTING AND RESULTS

a. INDIVIDUAL HOUSEHOLD

Using the treated datasets according to Table 2 with indicator variables containing all 3446 sampled individual households, and using a general linear regression model (GLM), the model has an adjusted R^2 value of 55%, when assessed by 5-fold cross validation. This is a reasonable model in terms of capturing the general relationship of household electricity demand to selected factors, but its prediction capability is relatively poor for forecasting annual electricity demand for any individual household on an annual basis. Details of the residuals are illustrated in Figure 6, with residuals scaled up to annual average daily household electricity consumption for better visualisation.

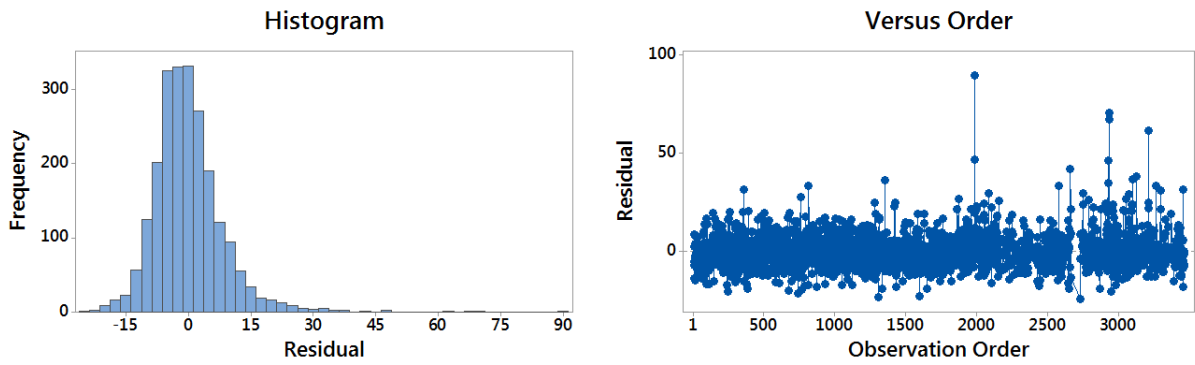


Figure 6: Residual plot of household model (residual in average daily electricity consumption kWh).

The finding highlights that it is difficult to predict a particular single household electricity demand as there is too much unknown information and too much variability in behaviour. However, the analysis enables a very specific understanding of the relationships between each factor and average household electricity demand.

a. AGGREGATED ZONE

This section examines the fitting of the model to aggregated electricity demand of residential households on a zone basis. Each household is flagged by their geographic location and 13 different zones are created (each zone consists of about 200-300 households). The 3446 household dataset then is aggregated according to the zone indicator by averaging every single predictor to get the mean value of electricity demand and the mean value of various factor levels (ranging between 0 – 1) for each 13 zones.

The same household electricity demand model was then used to perform the fitting against aggregated zone level data to investigate the zone level average daily electricity demand profile for the year. This model illustrates good accuracy in its fit (Figure 7).

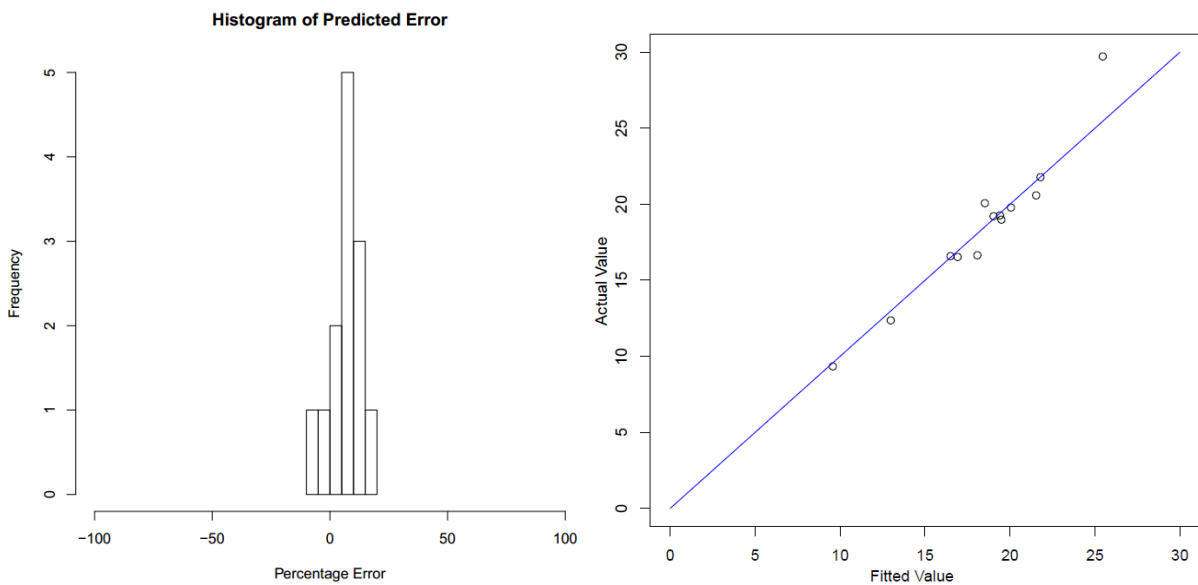


Figure 7: Percentage Error of Predicted value and Actual Value vs. Fitted Value (annual average daily kWh electricity demand) plot

Various prediction errors for this model have been evaluated using *R project* [57]. They are MAPE (Mean Absolute Percentage Error), MAE (Mean Absolute Error), CVRMSE (coefficient of variation of the root-mean-squared error), and MBE (Mean Bias Error) [59] [60] [61]. The statistical results for this model are shown in Table 5.

Table 5: Statistical Results

Model	GLM [57]
MAPE	3.9%
MAE	0.81

MSE	1.87
RMSE	1.37
Mean Electricity demand	18.84 kWh / day
Standard deviation	4.82 kWh / day
MBE	-0.63%
CVRMSE	7.4%

By applying the household model and utilising the model coefficients to fit zone level aggregated data, the forecasting capability has been significantly enhanced with a 3.9% MAPE.

VII. DISCUSSION AND FUTURE WORK

For the zone level model fit, an analysis is performed to illustrate quantitatively how each factor contributes towards household total electricity demand. As shown in Table 6, the aggregation of household behaviour, demographic, appliance and climate data has been performed to generate average household information for the 3446 households. This results in an *average factor value* for each factor for the investigated population. Each individual household has an information indicator variable which is either 0 or 1 to represent its True / False status or an integer (e.g. number of occupants). These aggregated information indicators are tabulated in Table 6 as *average factor values*, which represent the proportion of this characteristic for all sampled households. For example, the average number of household occupants is 2.9, and 15% of households have pool pumps. For each factor in the model, the *daily average electricity demand* is calculated by multiplying *average factor values* with the corresponding coefficient.

Table 6: Model Factors, coefficients, average factor value, and average electricity demand (daily)

Factor	Coef.	Average factor value	Daily average electricity demand
Constant	0.57		0.57
OCCUPANTS	3.15	2.90	9.15
CHILDREN_0_10	-1.06	0.40	-0.48
REFRIGERATORS	1.25	1.40	1.78
Heating	0.00	8127	3.12
POOLPUMP	8.77	0.15	1.34
HAS_GAS_HOT_WATER	-3.72	0.37	-1.36
DWELLING_TYPE Separate House	3.46	0.71	2.47
DWELLING_TYPE Unit	-0.97	0.28	-0.27
DRYER_USAGE HI	6.81	0.05	0.32
DRYER_USAGE LOW	0.48	0.36	0.17
DRYER_USAGE MED	3.53	0.21	0.74
AIRCON_TYPE Ducted	6.67	0.13	0.86
AIRCON_TYPE SplitSystem	1.45	0.49	0.72
HAS_GAS	-2.30	0.54	-1.24
HHOLD_INCOME_GROUP HI	1.47	0.27	0.39
HHOLD_INCOME_GROUP MED	-0.64	0.16	-0.11
HHOLD_INCOME_GROUP LOW	-0.94	0.19	-0.18
AIRCON_TYPE NONE	0.22	0.37	0.08
HAS_INTERNET_ACCESS	1.12	0.77	0.87
CHILDREN_11_17	-0.38	0.26	-0.10
HAS_SOLAR	-0.12	0.06	-0.01
Total			18.84

The coefficient of each factor can be used to indicate how each factor contributes towards the electricity demand of households in this study. The daily average electricity demand estimated in this study of 18.84 kWh is close to the NSW average of 18.9 kWh [56].

From this model, how average household electricity demand will respond when making certain decisions can be simulated. By aggregate all 3446 household data, and substitute into the model equation. It enables the model to evaluate each variable's contribution towards predicted total average electricity demand for the investigated sample (the entire 3446 households). Different scenarios are modelled by changing certain

average factor values to be either 0 or 1 to simulate corresponding daily electricity demand when a household chooses to have or not to have certain characteristics. For example, to simulate a household having and not having a pool pump, the average factor value is assigned to 1 and 0. Then, by calculating the daily average electricity demand for these two cases, the model then could predict that the average household would increase their daily electricity consumption compared to household without a pool pump by 8.7 kWh on an average annual basis. By choosing to have ducted air-conditioning, a household would increase their daily electricity consumption compared with a household without air-conditioning by 6.5 kWh on an average annual basis; however, by choosing to have a Split system air-conditioning, a household would only increase their daily electricity consumption compared to a household without air-conditioning by 1.2 kWh on an average annual basis; by choosing to have a gas hot water system, a household would decrease their daily electricity consumption compared with a household without a gas hot water system by 3.72 kWh on an average annual basis. It should be noted it is not the actual devices themselves that necessarily cause all of the increase or decrease in average electricity demand. Instead household ownership of particular devices is the indicator of change in electricity demand. The aforementioned key driving factors impacts are shown in Figure 8.

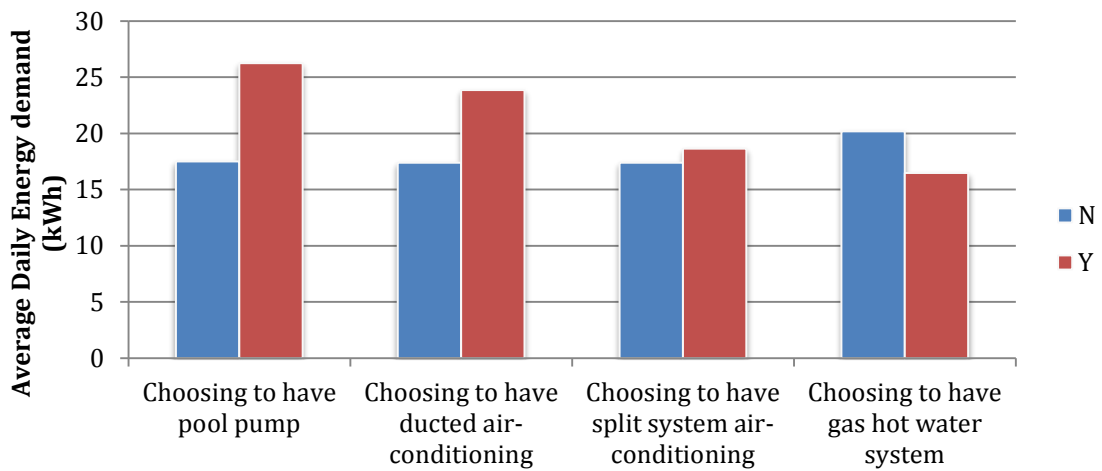


Figure 8: Impact of key driving factors

A bar chart summarising how each variable positively or negatively contributes towards total average electricity demand is shown in Figure 9.

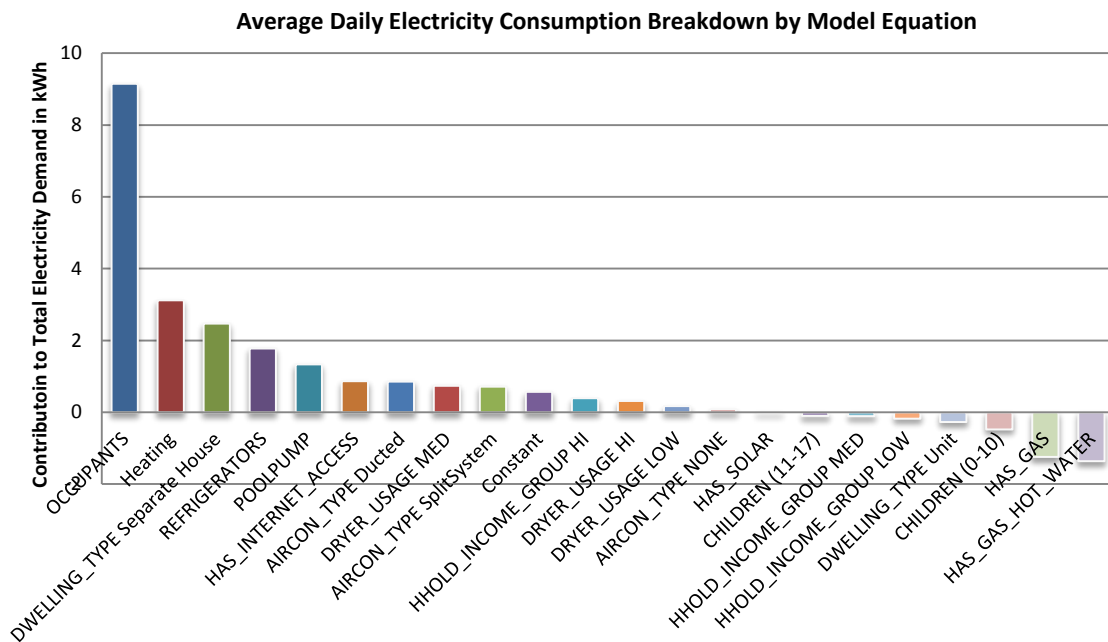


Figure 9: Annual average daily electricity demand breakdown by model equation

Clearly, the number of occupants contributes the most to household electricity demand. Also note that children (0-10 years) and children (11 – 17 years) are counted as occupants; however, the model shows that children consume less electricity than adults. In particular, children (0-10 years) consume considerably less than adults, while children (11 – 17 years) consume only slightly less than adults. The next largest contributor is then annual household heating degree half hours, followed by separate house dwelling type. However other factors such as access to gas and gas hot water act to reduce households’ electricity demand. Interestingly, as discussed previously cooling demand is not a significant contributor to household annual electricity demand; however, its impact on peak demand is very significant and will be examined in a separate work.

One of the applications for this developed model is to understand what drives residential electricity demand, and how factors actually influence it. Below are various quantitative intervention scenarios where key driving factors are manipulated to determine the effect on residential electricity demand. Selected key drivers are tabulated in Table 7 with the intervention tested.

Table 7: Impact of key factors on electricity demand

Intervention	Factors considered
Increase 10% of population of this characteristic	HAS_POOLPUMP, DRYER_USAGE CDHI, AIRCON_TYPE Ducted, DRYER_USAGE MED, DWELLING_TYPE , Separate House, HHOLD_INCOME_GROUP HI, AIRCON_TYPE SplitSystem, HAS_INTERNET_ACCESS, DRYER_USAGE LOW, AIRCON_TYPE_NONE, HAS_SOLAR, HHOLD_INCOME_GROUP MED, HHOLD_INCOME_GROUP LOW, DWELLING_TYPE Unit, HAS_GAS, HAS_GAS_HOT_WATER

The effect of these interventions on average daily electricity demand is also presented in Figure 10 as a percentage of 2013 annual average daily electricity demand.

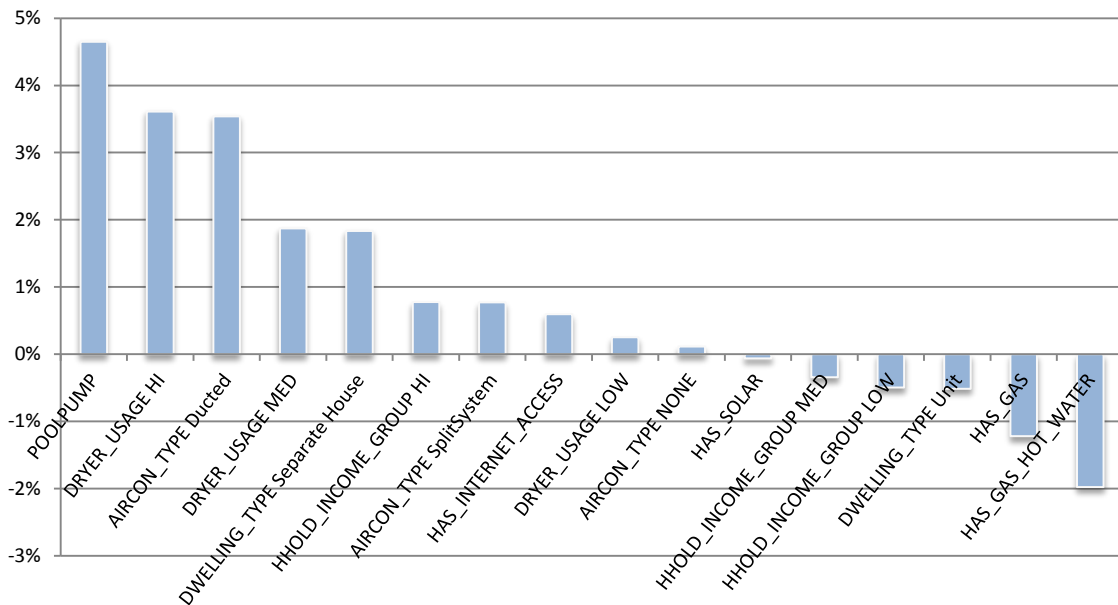


Figure 10: Percentage influence comparing with annual average daily electricity demand

Further analysis of the intervention will be addressed in future work. In addition, a separate work in the future will focus on half hour electricity demand modelling to better understand how each factor affects peak electricity demand and introduce scenarios of electricity efficiency measures, distributed renewable energy integration and distributed energy storage to reduce peak consumption and also achieve carbon abatement.

VIII. CONCLUSIONS

There have been many studies that have highlighted the potential electricity savings from energy efficiency, policy measures and simple demand side controls. However these seemingly “low hanging fruit”; have proved to be difficult to reach. Residential electricity demand is dynamic in nature. Making the right decisions for households to obtain electricity saving improvements is a complex issue. This study presents a way to identify opportunities to reduce electricity consumption by clearly identifying what and how different factors drive residential electricity demand, so as to inform both policy makers and end users on their decision making. Also, through bottom up statistical modelling, this study is capable of providing annual electricity prediction for the residential electricity sector at a local community zone level with a 3.9% MAPE. Such forecasting work is vital in electricity policy and decision making.

The main contribution of this paper is the development of an electricity demand model. The model distinguishes itself from traditional time series prediction methods. It excludes the influence of specific hours of the day, days of the week, holidays and other chronological impacts, focusing on demographic, household behaviour and building appliance influences on household electricity demand. Beyond this, the model statistically determines the answers to questions about how such factors drive residential electricity demand. While caution is required in interpreting the modelling outcomes, it is interesting to note that houses with a pool pump show an increase of daily electricity consumption by 8.7 kWh (a 50% increase compared with the average household without); houses with ducted air-conditioning show an increase of their daily electricity consumption by 6.5 kWh (a 37% increase) compared with houses without air-conditioning; houses with a gas hot water system show a decrease of their daily electricity consumption compared with houses without by 3.7 kWh (an 18% decrease). Direct comparisons of our findings with other work are restricted by the absence of similar studies in the Australian context. The Australian Department of Industry and Science has published relevant household electricity usage breakdown based on surveys, which broadly supports our findings in highlighting the key contribution that space heating and cooling and water heating play in overall consumption [62] [63]. With limited survey data, detailed measurements and specific end user statistics, previous work has rarely combined electricity readings with matched individual household survey data. These are huge barriers for policy makers and prevent effective energy efficiency improvements due to lack of accurate models. This paper utilises SGSC data to present an exploratory study aiming to systematically understand what and how various factors drive residential household electricity demand.

Acknowledgments: This work is supported by an Australian Postgraduate Award (APA), and additional scholarship support from the CRC for Low Carbon Living, SPREE UNSW. Special thanks go to Miss Zoe Hungerford for her editing and Jin Zhang for his statistics advice.

IX. REFERENCE:

1. Swan, L.G. and V.I. Ugursal, *Modeling of end-use energy consumption in the residential sector: A review of modeling techniques*. Renewable and Sustainable Energy Reviews, 2009. **13**(8): p. 1819-1835.
2. IEA, *Energy technology Perspectives 2014*. 2014, International Energy Agency.
3. Saddler, H., *Power Down - Why is electricity consumption decreasing?* 2013, The Australia Institute.
4. Ausgrid. *Smart Grid, Smart City (SGSC) Information Clearing House (ICH)*. 2012; Available from: <https://ich.smartgridsmartcity.com.au/>.
5. Kavagic, M., et al., *A review of bottom-up building stock models for energy consumption in the residential sector*. Building and Environment, 2010. **45**(7): p. 1683-1697.
6. Wiesmann, D., et al., *Residential electricity consumption in Portugal: Findings from top-down and bottom-up models*. Energy Policy, 2011. **39**(5): p. 2772-2779.
7. O'Neal, D.L. and E. Hirst, *An energy use model of the residential sector*. IEEE Transactions on Systems, Man and Cybernetics, 1980. **10**(11): p. 749-755.
8. Rivers, N. and M. Jaccard, *Combining top-down and bottom-up approaches to energy-economy modeling using discrete choice methods*. Energy Journal, 2005. **26**(1): p. 83-106.
9. Aydinalp-Koksal, M. and V.I. Ugursal, *Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector*. Applied Energy, 2008. **85**(4): p. 271-296.
10. Gul, M., S.A. Qazi, and W.A. Qureshi. *Incorporating economic and demographic variables for forecasting electricity consumption in Pakistan*. 2011. Sharjah.
11. Bianco, V., O. Manca, and S. Nardini, *Electricity consumption forecasting in Italy using linear regression models*. Energy, 2009. **34**(9): p. 1413-1421.
12. Garau, G., P. Lecca, and G. Mandras, *The impact of population ageing on energy use: Evidence from Italy*. Economic Modelling, 2013.
13. Swan, L.G., V.I. Ugursal, and I. Beausoleil-Morrison, *Occupant related household energy consumption in Canada: Estimation using a bottom-up neural-network technique*. Energy and Buildings, 2011. **43**(2-3): p. 326-337.
14. Farzana, S., et al., *Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China*. Energy and Buildings, 2014. **81**(0): p. 161-169.
15. Chen, S., et al., *Statistical analyses on winter energy consumption characteristics of residential buildings in some cities of China*. Energy and Buildings, 2011. **43**(5): p. 1063-1070.
16. Chen, S., et al., *Contrastive analyses on annual energy consumption characteristics and the influence mechanism between new and old residential buildings in Shanghai, China, by the statistical methods*. Energy and Buildings, 2009. **41**(12): p. 1347-1359.
17. Chen, J., X. Wang, and K. Steemers, *A statistical analysis of a residential energy consumption survey study in Hangzhou, China*. Energy and Buildings, 2013. **66**(0): p. 193-202.
18. Chen, V., et al., *What Can We Learn from High Frequency Appliance Level Energy Metering? Results from a Field Experiment*. 2014, UC Center for Energy and Environmental Economics.
19. Friedrich, L., P. Armstrong, and A. Afshari, *Mid-term forecasting of urban electricity load to isolate air-conditioning impact*. Energy and Buildings, 2014. **80**(0): p. 72-80.
20. Bargur, J.J., *Energy consumption and economic growth in Israel: Trend analysis*, in *Proceedings of the Third International Conference on Energy Use Management*. 1981.
21. Gonzales Chavez, S., J. Xiberta Bernat, and H. Llana Coalla, *Forecasting of energy production and consumption in Asturias (northern Spain)*. Energy, 1999. **24**(3): p. 183-198.
22. Ediger, V.Ş. and H. Tatlidil, *Forecasting the primary energy demand in Turkey and analysis of cyclic patterns*. Energy Conversion and Management, 2002. **43**(4): p. 473-487.
23. Unakitan, G. and B. Türkekul, *Univariate modelling of energy consumption in Turkish agriculture*. Energy Sources, Part B: Economics, Planning and Policy, 2014. **9**(3): p. 284-290.
24. Hunt, L.C., G. Judge, and Y. Ninomiya, *Underlying trends and seasonality in UK energy demand: A sectoral analysis*. Energy Economics, 2003. **25**(1): p. 93.
25. Kumar, U. and V.K. Jain, *Time series models (Grey-Markov, Grey Model with rolling mechanism and singular spectrum analysis) to forecast energy consumption in India*. Energy, 2010. **35**(4): p. 1709-1716.
26. Hagan, M.T. and S.M. Behr, *TIME SERIES APPROACH TO SHORT TERM LOAD FORECASTING*. IEEE Transactions on Power Systems, 1987. **PWRS-2**(3): p. 785-791.
27. Nogales, F.J., et al., *Forecasting next-day electricity prices by time series models*. IEEE Transactions on Power Systems, 2002. **17**(2): p. 342-348.

28. Fan, J.Y. and J.D. McDonald, *Real-time implementation of short-term load forecasting for distribution power systems*. IEEE Transactions on Power Systems, 1994. **9**(2): p. 988-994.
29. Abdel-Aal, R.E. and A.Z. Al-Garni, *Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis*. Energy, 1997. **22**(11): p. 1059-169.
30. Jaramillo-Morán, M.A., E. González-Romera, and D. Carmona-Fernández, *Monthly electric demand forecasting with neural filters*. International Journal of Electrical Power and Energy Systems, 2013. **49**(1): p. 253-263.
31. Barakat, E.H., *Modeling of nonstationary time-series data. Part II. Dynamic periodic trends*. International Journal of Electrical Power and Energy System, 2001. **23**(1): p. 63-68.
32. Willis, H.L. and H.N. Tram, *LOAD FORECASTING FOR TRANSMISSION PLANNING*. IEEE transactions on power apparatus and systems, 1984. **PAS-103**(3): p. 561-568.
33. Filik, Ü.B., Ö.N. Gerek, and M. Kurban, *A novel modeling approach for hourly forecasting of long-term electric energy demand*. Energy Conversion and Management, 2011. **52**(1): p. 199-211.
34. Charytoniuk, W. and M.S. Chen, *Nonparametric regression based short-term load forecasting*. IEEE Transactions on Power Systems, 1998. **13**(3): p. 725-730.
35. Al-Hamadi, H.M. and S.A. Soliman, *Long-term/mid-term electric load forecasting based on short-term correlation and annual growth*. Electric Power Systems Research, 2005. **74**(3): p. 353-361.
36. Haida, T. and S. Muto, *Regression based peak load forecasting using a transformation technique*. IEEE Transactions on Power Systems, 1994. **9**(4): p. 1788-1794.
37. Paatero, J.V. and P.D. Lund, *A model for generating household electricity load profiles*. International journal of energy research, 2006. **30**(5): p. 273-290.
38. Fischer, D., A. Härtl, and B. Wille-Haussmann, *Model for electric load profiles with high time resolution for German households*. Energy and Buildings, 2015. **92**: p. 170-179.
39. Richardson, I., et al., *Domestic electricity use: A high-resolution energy demand model*. Energy and Buildings, 2010. **42**(10): p. 1878-1887.
40. Ali, A.B.M.S. and S. Azad, *Demand forecasting in smart grid*. 2013. p. 135-150.
41. Tsanas, A. and A. Xifara, *Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools*. Energy and Buildings, 2012. **49**(0): p. 560-567.
42. Aydinalp, M., V. Ismet Ugursal, and A.S. Fung, *Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks*. Applied Energy, 2002. **71**(2): p. 87-110.
43. Sözen, A., E. Arcaklioğlu, and M. Özkaymak, *Turkey's net energy consumption*. Applied Energy, 2005. **81**(2): p. 209-221.
44. Ermis, K., et al., *Artificial neural network analysis of world green energy use*. Energy Policy, 2007. **35**(3): p. 1731-1743.
45. Mostafavi, E.S., et al., *A novel machine learning approach for estimation of electricity demand: An empirical evidence from Thailand*. Energy Conversion and Management, 2013. **74**: p. 548-555.
46. Badran, S.M. *Neural network integrated with regression methods to forecast electrical load*. 2012. Birmingham.
47. Boulaire, F., et al., *Statistical modelling of district-level residential electricity use in NSW, Australia*. Sustainability Science, 2013: p. 1-12.
48. Fan, S. and R.J. Hyndman. *Forecasting electricity demand in Australian National Electricity Market*. in *IEEE Power and Energy Society General Meeting*. 2012.
49. Ambrose MD, et al., *The Evaluation of the 5 - Star Energy Efficiency Standard for Residential Buildings*. 2013, CSIRO, Australia.
50. Wang, X., D. Chen, and Z. Ren, *Assessment of climate change impact on residential building heating and cooling energy requirement in Australia*. Building and Environment, 2010. **45**(7): p. 1663-1682.
51. Pears, A.K., *Imagining Australia's energy services futures*. Futures, 2007. **39**(2-3): p. 253-271.
52. Giurco, D.P., S.B. White, and R.A. Stewart, *Smart metering and water end-use data: Conservation benefits and privacy risks*. Water, 2010. **2**(3): p. 461-467.
53. Langham, E., Downes, J., Brennan, T., Fyfe, J., Mohr, S., Rickwood, P. and White, S., *Smart Grid, Smart City, Customer Research Report*, in *Report prepared by the Institute for Sustainable Futures as part of the AEFI consortium for Ausgrid and EnergyAustralia*. 2014.
54. Sharp, R., et al., *Smart Grid, Smart City: Shaping Australia's Energy Future*. 2014, AEFI consulting consortium. p. 16.
55. BOM. *Average annual & monthly heating and cooling degree days*. 2012 [cited 2014 5 Nov 2014]; Available from: http://www.bom.gov.au/jsp/ncc/climate_averages/degree-

[days/index.jsp?maptype=1&period=an&product=hdd18#maps.](#)

56. Rod Sims, J.C., Sibylle Krieger, *Residential energy and water use in Sydney, the Blue Mountains and Illawarra*. 2010.
57. R-CoreTeam, *R: A language and environment for statistical computing*. 2014, R Foundation for Statistical Computing, Vienna, Austria.
58. QualityCompanion, *Minitab 17 Statistical Software*. 2010, State College, PA: Minitab, Inc. .
59. Bradford, L.W.J., *M&V Guidelines: Measurement and Verification for Federal Energy Projects Version 3.0*. 2007, U.S. Department of Energy Federal Energy Management Program.
60. Armstrong, J.S. and T.S. Overton, *Estimating nonresponse bias in mail surveys*. Journal of marketing research, 1977: p. 396-402.
61. Willmott, C.J. and K. Matsuura, *Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance*. Climate research, 2005. **30**(1): p. 79.
62. Milne, C.R.G. *Your Home: Australia's guide to environmentally sustainable homes, 5th edition*. 2013 [cited 2015 28/05/2015]; Available from: <http://www.yourhome.gov.au/energy/appliances>.
63. DEWHA, *ENERGY USE IN THE AUSTRALIAN RESIDENTIAL SECTOR 1986 – 2020*. 2008. p. 41.

X. APPENDIX

Table 8: Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
OCCUPANTS	1	7.683	7.683	275.840	0.000
POOLPUMP	1	4.748	4.748	170.480	0.000
GAS_HOT_WATER	1	2.986	2.986	107.220	0.000
DWELLING_TYPE	2	4.754	2.377	85.350	0.000
DRYER_USAGE	3	4.370	1.457	52.300	0.000
AIRCON_TYPE	2	2.602	1.301	46.720	0.000
REFRIGERATORS	1	1.213	1.213	43.540	0.000
Heating	1	0.406	0.406	14.560	0.000
CHILDREN_0_10	1	0.402	0.402	14.440	0.000
GAS	1	0.286	0.286	10.270	0.001
OCCUPANTS_70PLUS	1	0.240	0.240	8.620	0.003
CHILDREN_11_17	1	0.197	0.197	7.080	0.008
INCOME	3	0.350	0.117	4.190	0.006
SOLAR	1	0.104	0.104	3.740	0.053
Error	2244	62.50	0.03		
Lack-of-Fit	2203	61.23	0.03	0.90	0.71
Pure Error	41	1.27	0.03		
Total	2274	140.7			