Estimating electricity impact profiles for building characteristics using smart meter data and mixed models

*Cameron Roach**

Abstract

Understanding the impact of building characteristics on electricity demand is important for policy and management decision making. Certain building characteristics and equipment may increase or decrease electricity consumption. Due to different operating practices, these impacts on electricity consumption may vary both across the day and across seasons. Quantifying the magnitude and statistical significance of these impacts will help managers and policy makers make better informed decisions. Here we present a mixed effects model to assess the importance of several variables on building electricity consumption. We use smart meter and building attribute data for 129 commercial office buildings. Our building attribute data includes information on installed equipment and meter characteristics of each building. To account for uncertainty in both variable significance and model selection we follow a multimodel inference approach. Demand impact profiles that show the expected change in electricity demand when a characteristic is absent or present are produced for each season. A discussion of the commercial office building characteristics we use and their impact on the daily profile of electricity demand is presented. Our approach has the advantage of only requiring building level demand and characteristic data. No equipment level sub-metering is required. Furthermore, our approach can also be used to quantify changes in electricity consumption caused by other factors that do not directly draw electricity from the grid, such as management decisions or occupant behaviour. We conclude with a discussion of applications for our methodology and future research directions.

1 Introduction

There is an increasing need to focus on the composition of electricity demand. Whereas in the past aggregated demand was sufficient for decision making, we are now often required to delve deeper and understand what underlying factors influence demand. Doing so can give a clearer picture of which building characteristics, occupant behaviours and policies are best able to reduce electricity consumption. For example, at a state or country level we may be interested in measuring the impact of solar or battery power on a typical power demand profile. At a building level facility managers may be interested in which equipment or building characteristics improve energy efficiency and at what points of the day they typically draw demand. Policies with the greatest efficacy can be identified and further promoted. Equipment and building management practices that are most efficient can help guide management and retrofitting decisions.

In this paper we focus on estimating the impact of several building characteristics on daily electricity demand. Electricity demand behaviour is typically presented in the form of power demand profiles which show the expected demand across a day. We define *demand impact profiles* as the observed change in power demand profiles when a particular characteristic is or isn't present. Demand impact profiles show how a characteristic increases or decreases demand over an entire day. Separate profiles are produced for each season due to the strong seasonality of some characteristics (e.g. heating and cooling equipment). Our models are trained on smart meter data from commercial office buildings across Australia and building characteristic data which describes which characteristics are present for each building. As we only require time

*Corresponding author. E-mail address: cameron.roach@monash.edu (C. Roach).

series data and metadata on individual buildings our methodology can be repeated for any scenario where similar data is available. All code and a walk-through have been made available online to allow for easy implementation (see Section 3.5).

Several approaches to disaggregating smart meter data and quantifying the consumption of appliances exist. Most of these approaches differ to ours. They attempt to reconstruct time series for the integrants making up total metered demand whereas we attempt to understand the average impact across the day during a particular season. Dinesh et al. (2016) use Karhunen Loeve expansion to decompose low frequency smart meter measurements into the appliance level electricity demand. Kalluri et al. (2016) examine time series subsequences to study appliance loads. Reinhardt et al. (2012) and Weiss et al. (2012) use several classification approaches that focus on identifying different appliances at a residential level. Disaggregation of household smart meter data to air-conditioning loads is explored by Perez et al. (2014) and validated on 19 buildings. Load disaggregation using supervised classification techniques are explored in Chahine et al. (2011) and Liao et al. (2014). Guo, Wang, and Kashani (2015) propose an algorithm for modelling appliance level consumption when only aggregated data is present and validate on a synthetic dataset. A load disaggregation approach for commercial buildings is proposed by Norford and Leeb (1996).

A drawback of these approaches is they can only assess the impact of equipment that draw electricity directly from the grid. They cannot assess other factors such as solar energy generation or the impact of building facade properties such as glazing. Behavioural factors can not be quantified either. Furthermore, uncertainty can not be assessed whereas we construct confidence intervals for our coefficient estimates.

A key advantage of other approaches is that some can create demand time series for certain loads, but we do not focus on achieving the same goal with our method. Our goal is to produce power demand profiles for characteristics and equipment of interest. We feel this is a reasonable aim. A decision maker is not interested in individual time series from a large number of buildings. Rather, they require a summary that clearly communicates how a characteristic of interest influences demand. Sub-metering may play a role in automating systems and diagnosing problems with individual pieces of equipment, but a statistical approach allows us to circumnavigate the time and costs associated with sub-metering to arrive at the same destination - a summary of how electricity usage is typically affected. In fact, the statistical approach goes further by allowing us to assign a degree of uncertainty to our findings. We can obtain an estimate for the expected change in electricity demand and a measure of confidence with which we can accept or reject our findings. In cases where inference is to be conducted we propose our multimodel inference approach using mixed models over traditional frequency analysis approaches.

Some papers have focused on calculating power demand profiles for building equipment. Gunay et al. (2016) and Mahdavi, Tahmasebi, and Kayalar (2016) used a data driven approach to calculate power demand profiles for office equipment. Plug loads were recorded for several pieces of office equipment and a predictive model was then trained off this metered data. A similar approach was used by Christiansen et al. (2015) to assess the energy consumption of medical equipment in hospitals. Menezes et al. (2014) offered two approaches to creating power demand profiles. The first relies on sampling from a database of monitored loads for equipment of interest. While it does allow for the calculation of confidence intervals, a drawback is its dependence on the quality of the metered data for each piece of equipment. Another matter is that it can only assess plug loads that have previously been recorded in a database. A second approach that did not rely on a database of metered data was also proposed, but was dependent on knowing or assuming the operational schedules of small office equipment which can sometimes be difficult. Our methodology has a similar aim to these studies, but can produce estimated power demand profiles without relying on a database of metered data for equipment or assuming operational schedules of equipment.

There are many building characteristics and items of equipment that can affect electricity demand. To avoid a naive data dredging analysis we approached building engineers to identify factors they suspected were influencing demand or were of interest to them. Our data included building attributes such as the type of electric equipment installed, building use and building meter characteristics (did the metered demand contain tenant usage?). The characteristics they identified are discussed in more detail in Section 2.2. We then used multimodel inference to test if these were statistically significant predictors and estimate how they influenced electricity demand over the day.

Multimodel inference is an information theoretic approach to model selection that relies on fitting multiple models with different combinations of predictors and then averaging the best performing models based on a suitable weight metric (Burnham and Anderson 2003). It is commonly used in ecology (Symonds and Moussalli 2011) but is not often used in the energy sector. To our knowledge the only example is So and Richman (2016) which used multimodel inference to create candidate models for the disaggregation of combined meter data for university campus buildings. It is a standard approach when there is no model specification that clearly outperforms others or when there is some uncertainty around which predictors to include. In our case it was necessary as using an all subset approach failed to produce a clear best model. Furthermore, simply picking the best performing model after an all subset analysis fails to take into account model selection uncertainty and often leads to inflated p-values (Burnham and Anderson 2003). Multimodel inference allows us to account for both model selection and parameter estimation uncertainty, thereby giving more reliable estimates of variable importance.

Electricity meter data is correlated within buildings. Some buildings will consistently have high demand, and others low. Mixed effects models allow us to account for this *within-subject* correlation by treating each building as a random effect. The random effect size indicates how much the mean of each individual's response variable differs from the sample population's mean. By modelling each building as a random effect, each building is treated as a random sample from a population of buildings with a specific distribution. Instead of model residuals being the only random component of our model, the random building selection is also taken into account. Mixed effects models also include fixed effects which are non-random quantities. In our case, fixed effects are the different building characteristics that we wish to create demand impact profiles for.

The main contribution of this paper is a methodology for generating demand impact profiles of various building characteristics at different times of the year. In addition to estimating the conditional mean we show how to calculate confidence intervals for our coefficient estimates that include both model and parameter selection uncertainty. We apply our approach to a real world dataset consisting of 129 commercial office buildings. In summary, the key advantages of our proposed methodology are:

- Only smart meter and building characteristic data are required to create demand impact profiles.
- Confidence intervals that include both parameter estimation and model selection uncertainty are produced.
- The impact of building characteristics other than plug loads can be quantified.

The paper is structured as follows. Section 2 discusses the time series data and metadata that motivates our research. Our mixed effects model and the multimodel inference approach are introduced in Section 3. Section 4 presents our estimates for building characteristic demand impact profiles and discusses applications and future research. Concluding remarks are made in Section 5.

2 Data

Data has been provided by Buildings Alive. Metered electricity consumption and building characteristic data is available for 129 buildings across Australia. Several years of data are available for each building. In general, data is available from 1 March, 2015 to 1 March, 2018 for most buildings, although some buildings have slightly less data. Only working days are included when fitting our model as electricity usage is dramatically different on non-working days (weekends and public-holidays). Electricity consumption tends to be significantly lower on non-working days due to equipment not being in use.

2.1 Time series data

We have 15-minute electricity consumption data available for each building. Time series plots of the raw data is shown for several buildings in Figure 3. In our analysis we divide the 15-minute electricity consumption values by each building's net lettable area to obtain *normalised* electricity consumption. These values are always positive and right-skewed (Figure 1) suggesting the use of a log-transform when modelling.

2.1.1 Hourly grouping

In our analysis we group the 15-minute interval readings into hours. While this reduces the granularity of the profile somewhat, it still allows us to assess the overall behaviour across a day while reducing the amount of variance in our coefficient estimates. It also means that each model can be fit to more data than would be available if 15-minute models were used. Working with 15-minute models or other temporal groupings, such as business and non-business hours, is also possible and can be chosen based on an analyst's needs. For our paper we prefer hourly models as they provided useful demand profile visualisations while also allowing sufficient training data for each model.

2.1.2 Outlier filtering

Smart meter data from buildings can be very noisy and typically contain outliers as seen in Figures 2 and 3. To avoid having our results adversely impacted by outliers we removed them from our analysis using a simple approach. For each hour, season and building the bottom (1st) quantile and top (99th) quantile were trimmed before fitting models. While this may have resulted in some valid values being excluded, it offered a quick way to remove the worst outliers.

2.2 Building attributes

We were motivated in our research by engineers that wished to statistically assess the relevance and importance of different building characteristics on electricity demand. Several characteristics were available but, to avoid overfitting and missing value issues, we limited ourselves to a subset that was of interest to their company and had high data quality. Our main research goal was to understand how each of these selected characteristics affected electricity demand. This is one of the contributions of our methodology - it allows us to assess the statistical significance and effect size of whatever characteristics we are presented with using only smart meter data and attribute data.

Three characteristic types are considered. *Direct* characteristics consume electricity and include items such as cooling equipment and electric heating equipment. *Behavioural* characteristics
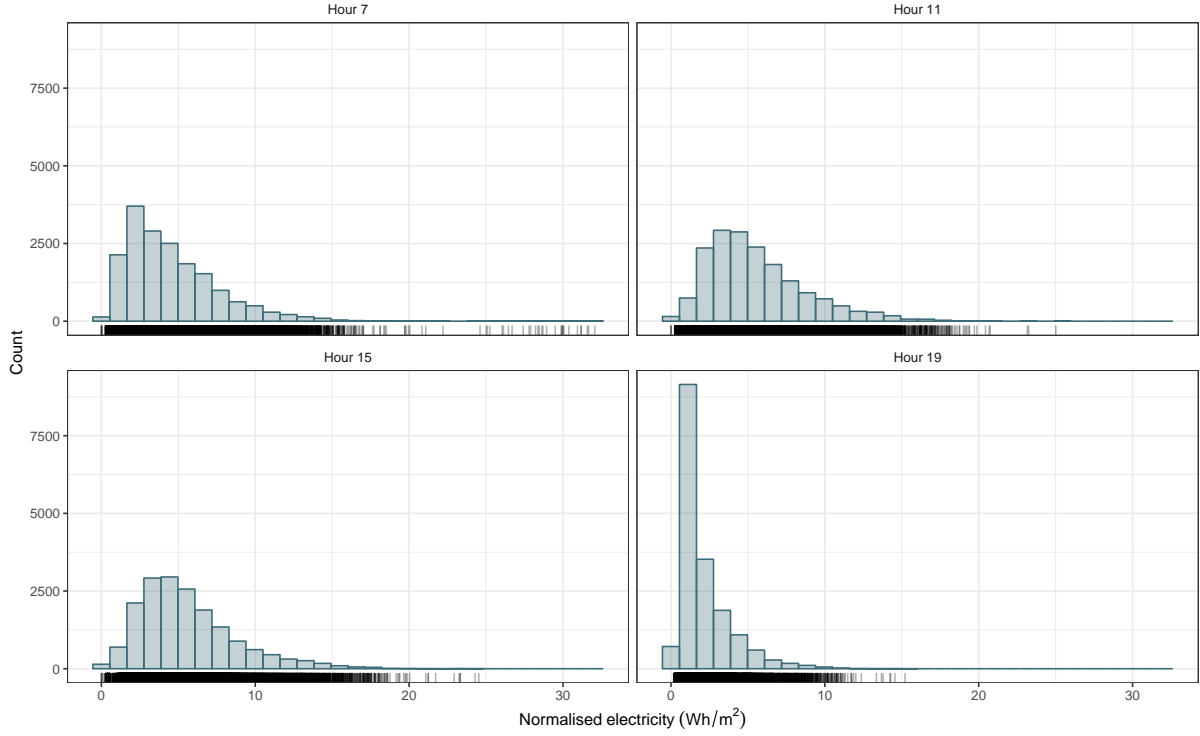
Figure 1: Histogram of normalised electricity consumption (Wh/m^2) for all commercial office buildings. A sample of 5% of meter readings at four different times of the day are shown. The readings are positive and right-skewed. Similar distributions occur for the remaining hours of the day.
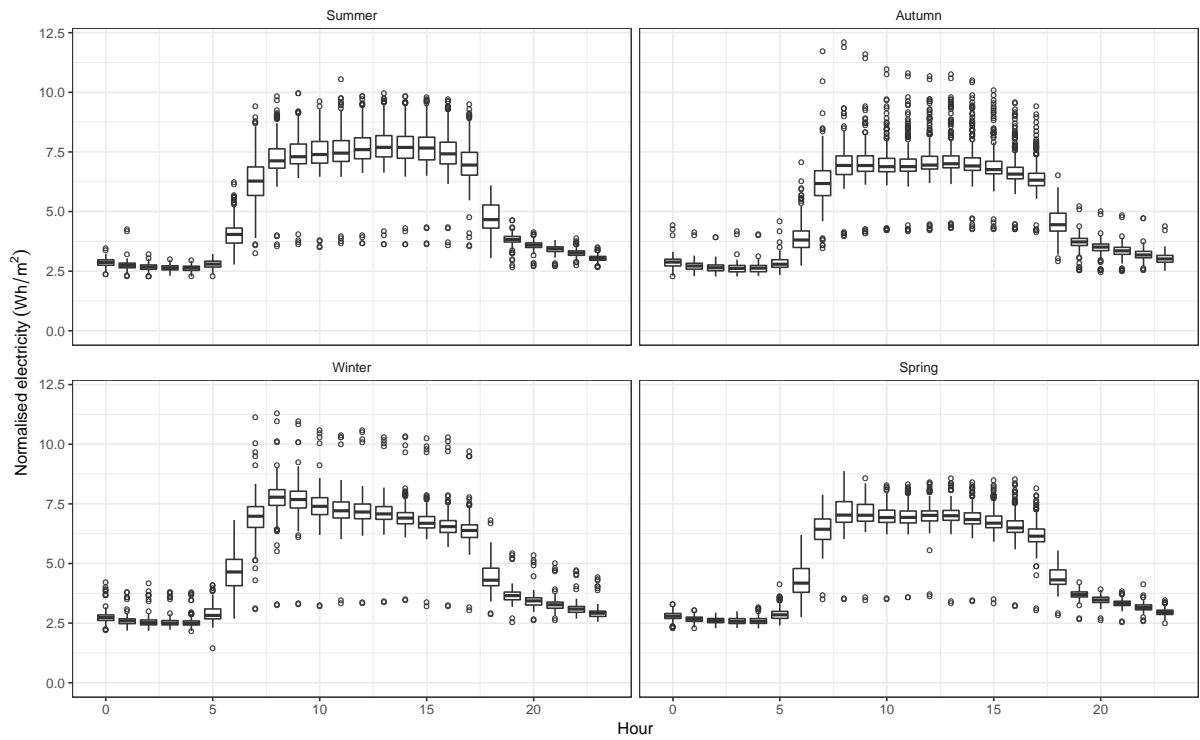

Figure 2: Hourly boxplots of normalised electricity consumption for building BID0212. Outliers are shown by circles. We see that there have been periods of low occupancy during business hours which result in outliers.

include tenant and management practices. *Indirect* characteristics are those that affect electricity demand but do not themselves use electricity. Examples can include glazing, insulation and gas heating equipment.

The variables we consider for modelling are described below.

Tenant feed. *(behavioural)* Some metered data only included electricity demand from the base building and excluded tenant usage. This variable identified if tenant consumption was included in the metered demand. Tenant data may include electricity demand from plug loads (e.g. computers, air conditioning) and lighting.

Water cooled condenser. *(direct)* A water cooled condenser discharges heat by transferring stored heat energy from a refrigerant to running water. The heated water may then be cooled in a cooling tower.

DX system. *(direct)* Direct expansion (DX) systems cool air. We were initially presented with two variables, DX system and chiller system, that had strong negative correlation. In other words, buildings would typically be equipped with one or the other. DX systems cool air whereas chiller systems cool water. We chose to work with DX system as one of our variables, though the choice was somewhat arbitrary as the main goal was to avoid multicollinearity. Reverse DX was another available variable, however this was omitted as only five buildings had reverse DX and there was obviously strong correlation with the DX system variable.

Gas fired boiler. *(indirect)* Gas fired boilers provide heating. They are fuelled by natural gas or propane and do not directly impact electricity demand. If a gas boiler is installed there will be no need to use electric heating to warm the building. Hence, they have an indirect effect as they offset the electricity that would otherwise have been required by an electric system.

Electric element heating. *(direct)* These heating systems are powered by an electrical source. A current is passed through metallic heating elements causing the elements to heat due to their resistance.

Centralised distribution. *(direct)* We were initially provided with *centralised distribution* and *per-floor distribution* variables. These had strong negative correlation (buildings typically had one or the other) and so we chose to use only centralised distribution in our model. Centralised distribution systems generate all cooling from one location and rely on ductwork for air distribution whereas per-floor distribution has cooling units on each floor.

Examples of the building attributes for six buildings are shown in Table 1. It is difficult to observe a relationship between each of these attributes and normalised consumption by simply comparing them to the electricity consumption plots in Figure 3. High volatility in each time series, coupled with the large number of buildings and the variation between buildings makes the attribute and demand relationship unclear. This motivates our research into finding an appropriate model to quantify their impact on demand.

3   Methodology

Figure 4 shows a flowchart of our proposed methodology. All of the available electricity and attribute datasets for each building are combined into one and then split by hour of day. Multimodel inference, parameter estimation and the estimation of proportional impacts is carried out separately for each hour (and season). The final step involves combining these hourly proportional impact estimates to produce a demand profile showing the estimated change in demand when an attribute is or isn't present. A simulation study that examines the capability of our proposed methodology is provided in Appendix B.
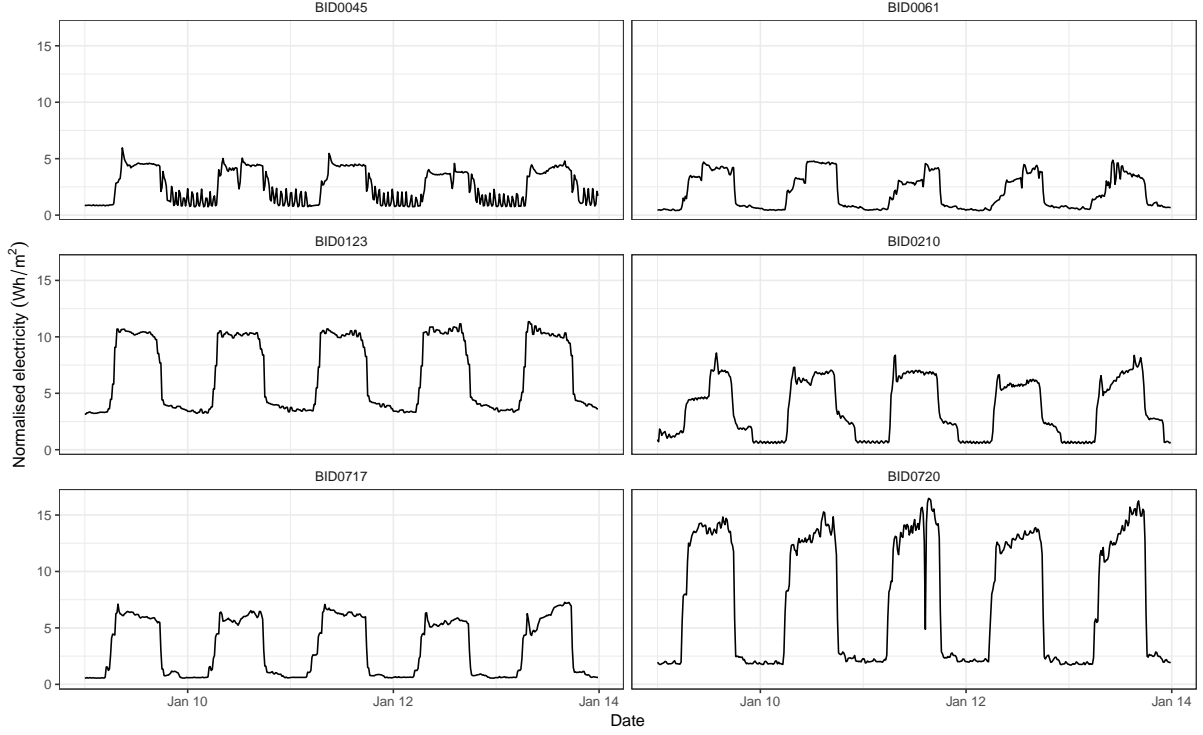
Figure 3: Time series plots of commercial office building smart-meter data for six of the 129 buildings. Each building has distinct business hour and non-business hour behaviour. The shape and volatility of the readings differ between each building. It is difficult to observe a clear relationship between the attributes listed in Table 1 and electricity demand using these visualisations alone. This motivates the development of our proposed approach that allows us to statistically test and quantify the relationship of each attribute and electricity demand.
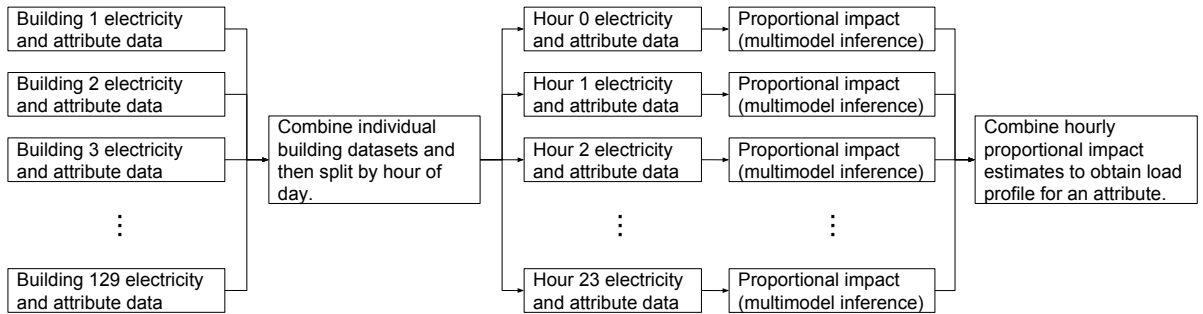

Figure 4: Flowchart of the proposed methodology for a given season.

7

Table 1: Building attributes for six of the 129 buildings.

Attribute	Building					
	BID0045	BID0061	BID0123	BID0210	BID0717	BID0720
Central Dist	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE
DXSystem	FALSE	FALSE	FALSE	TRUE	FALSE	TRUE
Electric Element Heating	FALSE	FALSE	TRUE	FALSE	TRUE	TRUE
Gas Fired Boiler	TRUE	TRUE	FALSE	TRUE	FALSE	FALSE
Tenant Feed	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE
Water Cooled Condenser	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE

3.1 Mixed effects model

We use a linear mixed effects model to describe the relationship between building attributes and electricity consumption. Mixed effects models can capture within-subject correlation which allows us to estimate parameters common to an entire population (fixed effects) and subject-specific parameters (random effects) (Pinheiro and D. M. Bates 1978). If we consider each building as a subject we have within-subject correlation in the electricity meter readings. Hence, we treat each building as a random effect. Building characteristics can be interpreted as population parameters and so they are modelled as fixed effects.

We create separate models for each season and hour of the day to allow for changes in building behaviour that typically occur over the course of the day and year. For instance, non-business hours will often have significantly less consumption than business hours. Peaks can occur in the morning due to pre-heating or pre-cooling. Heating equipment will be more important in winter and cooling equipment in summer. Distinct models for each hour and season allow for a comparison of their estimated fixed effect coefficients at different times. Plotting these estimates and their confidence intervals provide a clear overview of how each attribute's impact on demand evolves.

For a particular season and hour of the day the electricity consumption over 15-minute intervals, y_{ij}, for building j and observation $i \in \{1, 2, \ldots n_j\}$ is given below. The Boolean variables x_{hij} are equal to one if the hth building attribute is present and zero otherwise. We use a log-transform on the consumption data as it is positive and right-skewed (see Figure 1). The mixed model is

$$
\begin{aligned}
\log y_{ij} &= \beta_0 + \sum_{h=1}^{p} \beta_h x_{hij} + u_{0j} + u_{1j} t_{ij} + \epsilon_{ij}, \\
\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} &\sim N(0, \Omega_u), \quad \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{bmatrix}, \\
\epsilon_{ij} &\sim N(0, \sigma_\epsilon^2),
\end{aligned}
\tag{1}
$$

where β_h is the coefficient for attribute indicator x_{hij}, u_{0j} is the random intercept for building j, u_{1j} is the random slope coefficient for year t_{ij} and ϵ_{ij} are the residuals. This linear mixed effects model allows the intercept to vary with each building. A random slope for year has been added to capture any trends in building performance. We include a covariance term σ_{u01} in the variance-covariance matrix to allow for correlation between the random intercept and random slope. This is chosen as we tend to observe negative correlation between the random intercept and slope.

We also attempted to model residuals using an autoregressive correlation structure of order 1. While this did reduce autocorrelation in the standardised residuals (Pinheiro and D. M. Bates

1978) it had almost no impact on our final fixed effect coefficient estimates and confidence intervals. Furthermore, the $AR(1)$ correlation structure resulted in a considerable increase in computation time and occasionally convergence issues. Given this, we chose to model residuals as in equation (1).

3.2 AIC for mixed effects models

A means to assess each model for goodness of fit and complexity is required when conducting model selection. Complicated models may fit data better, but fail to generalise to new datasets. This is indicative of over-fitting rather than a well specified model. Information criteria such as the Akaike information criteria (AIC; Akaike (1998)) and Bayesian information criteria (BIC; Schwarz (1978)) score models on how well they fit data while also penalising them for complexity. Choosing a suitable information criteria for a linear mixed effects model is typically more complicated than in linear regression due to issues arising from the selection of covariance structures and positive semidefinite constraints on the covariance matrix (Müller, Scealy, and Welsh 2013).

For our model selection we use the marginal AIC (mAIC) which is the most widely used information criteria for mixed effects models. Vaida and Blanchard (2005) define the mAIC as

$$mAIC = -2\ell\left(\hat{\beta}\right) + 2(p+q), \tag{2}$$

where $\ell\left(\hat{\beta}\right)$ is our log-likelihood function, p is the number of fixed effects, q is the number of random effects. We choose this criterion as it is both simple to understand and has been used in many studies (Müller, Scealy, and Welsh 2013).

3.3 Multimodel inference

Despite the effort taken to carefully determine appropriate building characteristics for our model there is still uncertainty regarding the importance of each proposed variable. Multimodel inference allows us to incorporate this model selection uncertainty into our parameter estimates. Instead of simply conducting stepwise variable selection or best subset selection and then choosing the best model, we instead use a candidate set of models on which we base our inference. This model averaging approach has merit as sometimes the best model will only offer a small improvement over other models based on a quality score such as AIC. Had a different dataset sample been present it may have resulted in another model being selected as the best (Burnham and Anderson 2003). Furthermore, different models will sometimes show the same variable being significant or insignificant (Symonds and Moussalli 2011). The overall conclusion is that model selection uncertainty needs to be taken into account when conducting inference in situations such as ours. If we only focus on parameter estimation uncertainty without considering model selection uncertainty we will likely underestimate the size of our confidence intervals.

Multimodel inference focuses on selecting a subset of models on which to base inference. Note that effort should be taken to avoid data dredging (where the computer is left to select the best variables with no prior hypothesis on the researcher's part). Burnham and Anderson (2003) suggest using prior knowledge of the situation to determine suitable predictors. Fitting only a subset of all possible models is a sensible approach because fitting all possible models can quickly result in a large computational burden if the number of predictors is large[1]. As discussed

---

[1]Given p predictors we would require 2^p models to be fitted.

9

in Section 2.2 we have narrowed all available predictors down to a reasonable subset based on advice from domain experts.

Given our subset of predictors, multimodel inference fits a separate mixed effects model to every combination of predictor variables. This gives 2^p models each with an $mAIC$ score. The best performing models form our candidate set on which we conduct inference.

3.3.1 Candidate sets

To construct our candidate set we must first determine the probability of each candidate model. To do so we use Akaike weights. Given R candidate models the Akaike weight for model g_i is

$$w_i = \frac{\mathcal{L}\left(g_i|\mathbf{x}\right)}{\sum_{r=1}^{R} \mathcal{L}\left(g_r|\mathbf{x}\right)} = \frac{e^{-\frac{1}{2}\Delta_i}}{\sum_{r=1}^{R} e^{-\frac{1}{2}\Delta_r}}, \tag{3}$$

where $\mathcal{L}\left(g_i|\mathbf{x}\right)$ is the likelihood of model g_i given data $\mathbf{x}$ and $\Delta_i = mAIC_i - mAIC_{min}$ is referred to as the AIC difference. We use $mAIC$ in our analysis as we are dealing with mixed effects models. Other information criteria for mixed effects models such as the conditional AIC (Greven and Kneib 2010; Vaida and Blanchard 2005) may also be used.

We can use the AIC differences and Akaike weights to select a subset of models for inference. We refer to this subset as the candidate set. For our analysis we use the common practice of selecting models with the highest weights such that their cumulative sum is just above 95%. As the weights serve as model probabilities we can refer to this as a 95% confidence set.

3.3.2 Parameter estimates

Parameters are estimated by "averaging" the models in our confidence set. There are two common approaches to model averaging. *Natural-model averaging* averages over all candidate models where a parameter of interest occurs. *Full-model averaging* considers all candidate models. If a variable is not selected in one of the candidate models, full-model averaging sets its estimate to zero. Hence, full-model averaging takes into account when a variable has not been selected whereas natural-model averaging does not. Furthermore, simulation studies have found that full-model averaging can help to reduce problems caused by model selection bias towards overly complex models (Lukacs, Burnham, and Anderson 2010). We use full-model averaging for this reason.

Given a candidate set of R models our full-model averaged coefficients β_h are estimated by

$$\hat{\bar{\beta}}_h = \sum_{i=1}^{R} w_i \hat{\beta}_{hi}, \tag{4}$$

where $\hat{\beta}_{hi}$ is the estimate of β_h based on model g_i. If β_h is not chosen in model g_i then $\hat{\beta}_{hi}$ is defined to equal zero in the above formula.

3.3.3 Unconditional confidence intervals

Once a set of candidate models has been identified we can construct confidence intervals that reflect both parameter and model selection uncertainty. The $(1-\alpha)\,100\%$ unconditional confidence intervals for the a model averaged coefficient $\hat{\bar{\beta}}_h$ is given by

10

$$\bar{\hat{\beta}}_h \pm z_{1-\alpha/2}\widehat{\text{ase}}\left(\bar{\hat{\beta}}_h\right), \tag{5}$$

where $\widehat{\text{ase}}\left(\bar{\hat{\beta}}_h\right)$ is the adjusted standard error from Burnham and White (2002). It is given by

$$\widehat{\text{ase}}\left(\bar{\hat{\beta}}_h\right) = \sum_{i=1}^{R} w_i \sqrt{\left(\frac{t_{\text{df}_i,1-\alpha/2}}{z_{1-\alpha/2}}\right)^2 \widehat{\text{var}}\left(\hat{\beta}_{hi}|g_i\right) + \left(\hat{\beta}_{hi} - \bar{\hat{\beta}}_h\right)^2}, \tag{6}$$

where $\bar{\beta}_h$ is the model averaged estimator of β_h, $\widehat{\text{var}}\left(\hat{\beta}_{hi}|g_i\right)$ is the estimated variance of parameter β_{hi} in model g_i, and w_i are weights. The calculation of $\widehat{\text{var}}\left(\hat{\beta}_{hi}|g_i\right)$ for mixed effects models is reasonably complex and is omitted (see D. Bates, Mächler, et al. (2015) for a discussion).

3.4 Estimating power demand profiles

In equation (1) we have modelled the log-transform of electricity consumption over 15-minute intervals as our response. We use the estimator proposed by Kennedy (1981) to calculate the proportional impact, p_h, of X_{hij} on the dependent variable Y_{ij}. Kennedy's estimator is consistent and almost unbiased (Giles 2011). For a dummy variable changing from zero to one the estimator is given by

$$\hat{p}_h = e^{\bar{\hat{\beta}}_h - 0.5\widehat{\text{ase}}\left(\bar{\hat{\beta}}_h\right)^2} - 1, \tag{7}$$

where we have replaced the coefficient estimate and variance of Kennedy's original estimator with our full-model averaged counterparts. As this is the proportional impact we can rewrite the above expression in terms of our response variable to obtain

$$Y_{ij}^* = e^{\bar{\hat{\beta}}_h - 0.5\widehat{\text{ase}}\left(\bar{\hat{\beta}}_h\right)^2} \mathbb{E}\left[Y_{ij}|X_{hij} = 0\right], \tag{8}$$

where Y_{ij}^* is the new consumption value after our Boolean variable has changed from false to true. Calculating Y_{ij}^* using this formula for each hour of the day gives our power demand profiles. Taking the difference between Y_{ij}^* and $\mathbb{E}\left[Y_{ij}|X_{hij} = 0\right]$ gives our estimate of the demand impact profiles. Note that we have chosen to base our profiles off $\mathbb{E}\left[Y_{ij}|X_{hij} = 0\right]$ rather than an unconditional mean or median because we are working with proportional impacts. Since we show the impact of a variable switching from false to true it seems reasonable to apply the proportional impact to the mean demand that we observe when an attribute isn't present, hence the use of the conditional expectation. Simply using the mean or median may exaggerate the demand impact profiles.

3.5 Fitting models

We fit our models using the R statistical programming language (R Core Team 2018). The `lme4` package is used to fit our mixed effects model and calculate mAIC scores (D. Bates, Maechler, et al. 2019). Multimodel inference is carried out using the `MuMIn` package (Bartoń 2019).

All code used to produce this analysis has been made available at https://github.com/camroach87/1801-MMME.

4   Discussion

In this section we apply our methodology to assess the impact of building characteristics on electricity demand across the day. We also examine how well our models fit the data and comment on the limitations of our statistical methodology. Several future research directions are put forward.

To validate our approach a simulation study was also conducted using simulated time series data designed to mimic electricity demand in commercial buildings. We found that after simulating electricity consumption for 129 buildings our methodology was able to estimate the coefficients acceptably, with most estimates falling within the 90% confidence intervals. Details of the simulation study are provided in Appendix B.

4.1   Goodness of fit

Prior to analysis of our results it is important to assess if our models actually fit the data acceptably. We use marginal and conditional R^2 values for each season and hourly model to assess the goodness of fit. Adapting the specification of Nakagawa and Schielzeth (2013) to our case, the conditional R^2 calculates the proportion of the variance explained by both fixed and random effects and is given by

$$R_c^2 = \frac{\sigma_f^2 + \sigma_u^2}{\sigma_f^2 + \sigma_u^2 + \sigma_\epsilon^2},\tag{9}$$

where $\sigma_f^2 = \text{var}\left(\sum_{h=1}^p \beta_h x_{hij}\right)$ is the variance of the fixed effects, σ_u^2 is the variance of random effects and σ_ϵ^2 is the residual variance. As we are working with a random slope model the random effect variance is calculated as described in Johnson (2014). Marginal R^2 considers only fixed effects and is defined as

$$R_m^2 = \frac{\sigma_f^2}{\sigma_f^2 + \sigma_u^2 + \sigma_\epsilon^2}.\tag{10}$$

The goodness of fit statistics are shown in Table 2 and Figure 5. Overall, it appears that our model formulation gives reasonable fits with the conditional R^2 values averaging close to 0.9. Models close to the start and end of business hours show the worst fits due to the noisiness of data during these times. This likely reflects the different operating schedules for different buildings at these times of the day. Another point to note is our marginal R^2 values are consistently higher during winter business hours compared to other seasons. This indicates that the chosen building attributes are better at modelling heating loads than cooling loads.

Marginal R^2 only includes the impact of fixed effects allowing us to assess how much of the variance in electricity consumption is explained by them. This lets us examine the goodness of fit when using building attribute variables only. If fixed effects do not explain any variation in the data our marginal R^2 values will be close to zero. In our case we see marginal R^2 values range between 0.25 and 0.5 whereas the conditional R^2 values range between 0.8 and 0.95. To be able to explain 25-50% of the variance in the data with only a handful of attributes is encouraging and shows that at least some of the building characteristics we are working with have explanatory power. For reference, 0.25-0.5 is not an unreasonable range[2] when compared with ecology studies (Nakagawa and Schielzeth 2013). Given the building attributes alone are

---

[2]More studies using mixed effects models on smart-meter data need to be conducted before it is fair to conclude that the values presented here are low or high.

Table 2: Conditional and marginal R^2 values for each hour and season's best model after all-subsets selection.

Hour	Summer R_c^2	Summer R_m^2	Autumn R_c^2	Autumn R_m^2	Winter R_c^2	Winter R_m^2	Spring R_c^2	Spring R_m^2
0	0.94	0.44	0.95	0.44	0.96	0.45	0.96	0.46
1	0.94	0.45	0.95	0.45	0.96	0.46	0.96	0.47
2	0.94	0.45	0.95	0.43	0.96	0.46	0.96	0.47
3	0.94	0.45	0.95	0.45	0.96	0.46	0.96	0.47
4	0.92	0.47	0.95	0.45	0.95	0.48	0.95	0.51
5	0.84	0.44	0.90	0.47	0.91	0.50	0.91	0.53
6	0.74	0.34	0.80	0.41	0.87	0.42	0.83	0.45
7	0.75	0.25	0.76	0.33	0.91	0.42	0.82	0.37
8	0.80	0.21	0.79	0.31	0.96	0.42	0.85	0.34
9	0.82	0.22	0.80	0.30	0.96	0.42	0.85	0.33
10	0.82	0.22	0.80	0.29	0.96	0.43	0.84	0.31
11	0.83	0.22	0.80	0.28	0.96	0.42	0.84	0.30
12	0.83	0.21	0.79	0.26	0.96	0.40	0.85	0.29
13	0.84	0.21	0.80	0.26	0.96	0.40	0.84	0.29
14	0.81	0.21	0.79	0.26	0.96	0.40	0.84	0.29
15	0.81	0.20	0.78	0.25	0.96	0.41	0.84	0.29
16	0.83	0.21	0.77	0.25	0.96	0.40	0.83	0.29
17	0.77	0.15	0.77	0.26	0.95	0.39	0.82	0.27
18	0.81	0.21	0.85	0.30	0.92	0.39	0.85	0.31
19	0.86	0.30	0.90	0.36	0.95	0.43	0.91	0.39
20	0.89	0.35	0.89	0.36	0.93	0.42	0.92	0.41
21	0.90	0.38	0.92	0.37	0.95	0.44	0.94	0.42
22	0.92	0.41	0.93	0.39	0.96	0.44	0.95	0.44
23	0.93	0.43	0.95	0.43	0.96	0.45	0.96	0.45

able to account for some of the variation in the data suggests that they do improve the model fit. Once inter-building differences are accounted for by including the random effects we see subjectively good fits based on the conditional R^2. It is recommended that future smart-meter studies using mixed effects models should include both marginal and conditional R^2 values to allow for a discussion on how well their fixed effects model the data. Being able to examine how much variance the fixed effects capture is important to consider and provides interesting information (Nakagawa and Schielzeth 2013). Should marginal R^2 values ever be close to zero applying our methodology will not prove particularly useful even if high conditional R^2 values are observed, and so it is important to include this information.

We also examined quantile-quantile (QQ) plots to check if model residuals were normally distributed. For the most part the residuals did appear to be normally distributed, though there was some evidence of heavy tails. This was largely due to some remaining outliers in the time series data - sometimes caused by erratic spikes and unexpected drops in demand. Extreme weather and unmarked local holidays in the data could each be responsible for some of the noise. Including the relationship between temperature and electricity demand would likely improve our fits. However, due to the added complexity[3] and a desire to keep our mixed effects model clear for explanatory purposes we leave this for future research. Considering the noisiness of the

---

[3]The electricity temperature relationship in each building can be quite nonlinear and varies across the day.
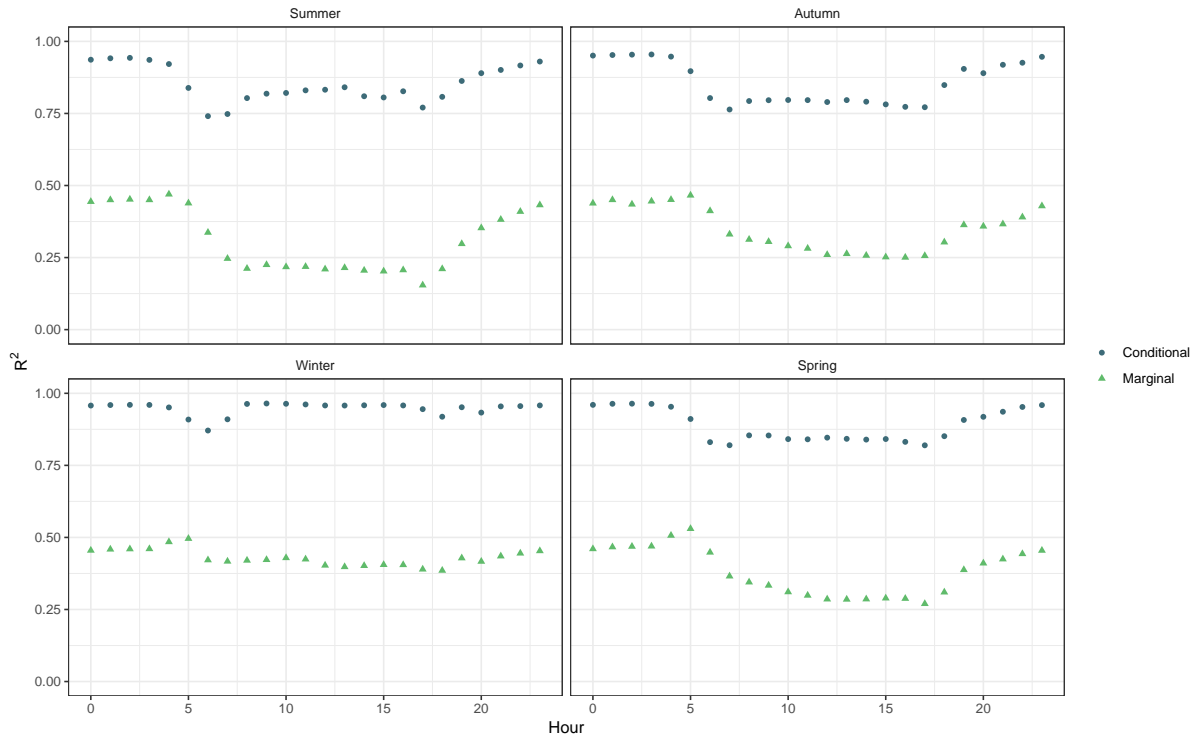
Figure 5: Conditional and marginal R^2 values for each hour and season's best model after all-subsets selection. The marginal R^2 value represents the variance explained by fixed effects whereas the conditional R^2 value is the variance explained by both fixed and random effects. When we plot these goodness of fit values we see a drop during the working hours which is expected due to the more volatile nature of demand during these times.

data we were satisfied that the residuals were close enough to being normally distributed for our analysis.

4.2 Demand impact profiles

Plots of the power demand profiles and demand impact profiles for tenant consumption, electric element heating and DX systems are shown in Figures 6, 7 and 8, respectively. Power demand profiles with and without the attribute in question are shown in the top panels; whereas the bottom panels show the demand impact profiles. These three plots show examples of behaviour impact, caused by occupants; and equipment impact, caused by equipment that directly draws electricity demand. Note that normalised consumption over 15-minutes intervals was used as the response variable in our mixed effects model. To aid in interpretation of results the fixed effect estimates are converted from units of energy to units of power[4] when plotting.

As we have data for buildings with and without tenant consumption data we can estimate the expected behaviour of tenants (Figure 6). This is an example of assessing occupant behaviour with our model. As expected we see a large contribution to electricity demand by tenants, ranging from 6 to 15 W/m^2 across all seasons. Of particular interest is the large morning spike we see at 6:00 am in summer, possibly caused by tenants attempting to pre-cool their offices to ensure occupant comfort over the course of the day. Observing tenant behaviour allows us to quantify and view when demand from tenants is occurring. This provides some guidance on how much potential savings there might be (as an upper bound), or how the shape of the profile might be modified through intervention. One useful application of this is to assess if a demand management initiative is successful. Including it as a fixed effect will allow for demand impact profiles to be created. Decision makers may then assess statistically if said initiative was successful or not.

Profiles for electric element heating are shown in Figure 7. Here we assess a building characteristic that directly uses electricity and where the behaviour is known. As would be expected with heating equipment we see only a small change in demand during summer. Slightly higher demand is seen in autumn and spring, likely due to cold days during the months that border winter. Winter sees a large increase in electricity demand due to heating required during business hours.

As a final example, the subplots in Figure 8 highlight that buildings with DX systems tend to use more electricity than those without. Given there was a strong negative correlation with chiller systems (see Section 2.2) it seems reasonable to conclude that chiller systems offers a more energy efficient alternative based on this dataset. However, this conclusion should be tempered by inspecting the confidence intervals in Figure 9. The coefficient estimates are, for the most part, only significant ($\alpha = 0.1$) during winter. It is possible that this result captures the impact of reverse DX systems being used for heating during winter months.

Centralised distribution and gas fired boilers did not show statistically significant ($\alpha = 0.1$) effects at any time of the year. Water cooled condensers were, for the most part, not statistically significant. Only one model for water cooled condensers (autumn at 5:00 pm) had a statistically significant impact, but did not provide any useful or interesting conclusions. For reference, a discussion of the coefficients and confidence intervals of all attributes and seasons is provided in Appendix A.

In summary, for our building stock we can conclude that:

- We do not observe a statistically significant difference between centralised and per-floor

---

[4]This is a simple matter of dimensional analysis where 1 Wh of energy consumed over a duration of 15-minutes is equivalent to 4 W: $\frac{1 \text{ Wh}}{15 \text{ min}} = \frac{1 \text{ Wh}}{1/4 \text{ h}} = 4$ W.

distribution systems. Hence, we can not conclude that one is more or less efficient than the other based on this dataset and controlling for the variables we have chosen to model.

- We observe DX systems use more energy than chiller systems during winter, though this is possibly due to reverse DX systems being used for heating.
- Electric element heating increases the expected electricity demand by as much as 6 W/m^2 during the winter. The largest draw occurs in the morning between 7:00 am and 9:00 am.
- Gas fired boilers did not make a statistically significant impact on electricity demand. There was an expectation that their presence may result in offsetting electricity demand, but this did not show in the final results. This could perhaps indicate issues with data or building operation.
- Unsurprisingly, tenant feed had a significant impact for all seasons and at all times of the year. Large increases in tenant demand of approximately 8 W/m^2 are seen between the hours of 6:00 am to 4:00 pm. Tenant demand gently tapers off after 4:00 pm during autumn, winter and spring, but has a sharp drop after 4:00 pm in summer. There is a large spike in tenant electricity demand at 6:00 am in summer which indicates that there may be opportunities within our building stock to spread out demand at this time to reduce peaks.
- The presence of water cooled condensers in buildings does not appear to have a significant impact on electricity demand.

Note that these findings only apply to our building stock when controlling for the variables we have used in our model. Each conclusion is subject to idiosyncrasies in the data or perhaps be indicative of unusual building behaviour.

Determining the expected consumption of tenants in buildings may allow for targeted energy saving measures. For example, if there are large peaks at certain times of the day, facility managers may make tenants aware of the need to reduce demand during these times. The early morning spike we observe in summer could be deemed one such time. This peak possibly corresponds to early morning cooling to ensure indoor environments are comfortable when tenants arrive. If this was automated to spread cooling across a longer period the peak could be reduced. While this wouldn't necessarily decrease the overall consumption, the lower peak may reduce costs for peak power consumption[5] and be useful in grid management.

Using these results, the company that provided data were able to determine which attributes were important and how heavily they should be weighted when assessing building similarity for electricity demand. This allowed for benchmarking and target setting for facility managers based not only on a single building's historical demand, but also on other buildings with similar characteristics.

4.3   Modelling limitations

As with any statistical approach there are certain limitations. While we may find that some variables are statistically significant we can not necessarily comment on causality. Confounding variables and idiosyncrasies of a given dataset may be responsible for some results. Furthermore, if a large number of fixed effects are being explored then it is important to have a large number of buildings in the dataset to avoid the curse of dimensionality. Despite these matters, it is still useful to conduct analyses such as this to better understand underlying behaviour and identify possible anomalies that are worthy of further investigation.

One difficulty of drawing conclusions from mixed effects models is that technical data describing installed equipment is conflated with operational practice. For instance, our centralised distribution variable merely indicated if a building had centralised distribution and not the model, size

---

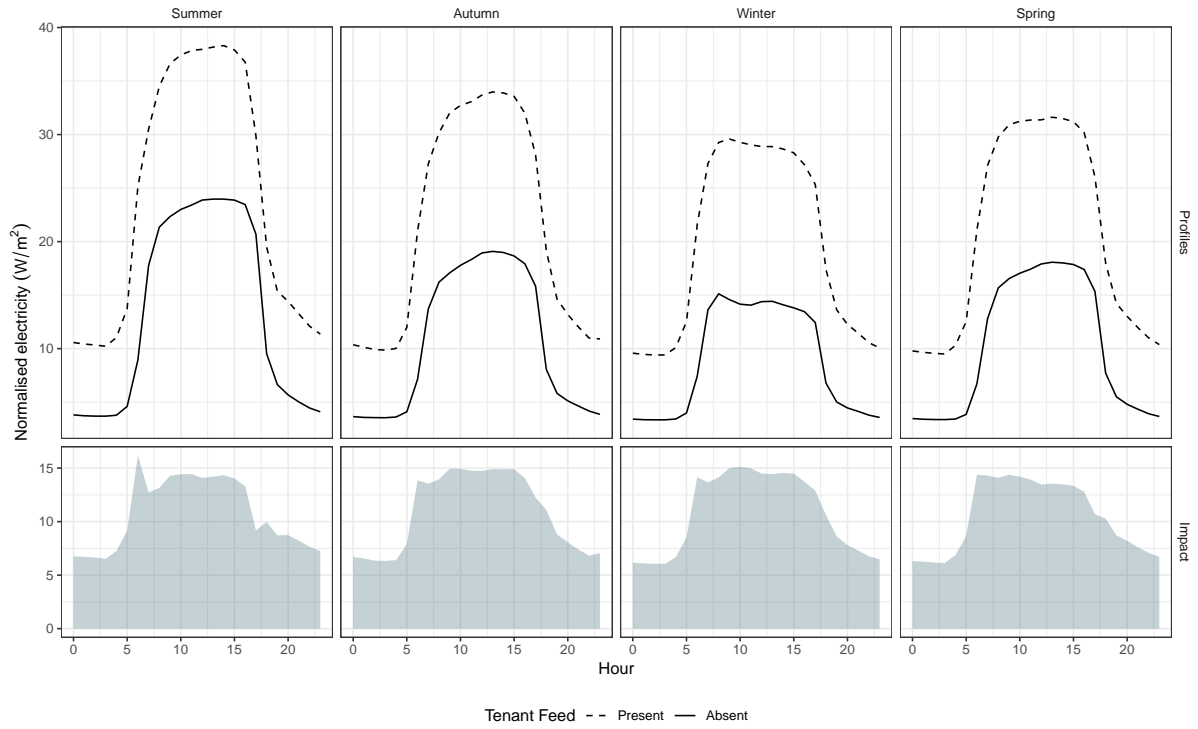[5]Especially if time of use tariffs are in place and we observe peaks during expensive times of the day.

Figure 6: Profile plots of tenant feed impact. This is an example of a behavioural aspect of electricity demand. Here we can directly assess how tenants typically use electricity during each of the seasons. One notable feature of the plots is the large spike at 6:00 a.m. during summer mornings, possibly caused by precooling of offices. We can not know this for sure without collecting further attribute data on tenant cooling equipment, but does serve to motivate further investigation.
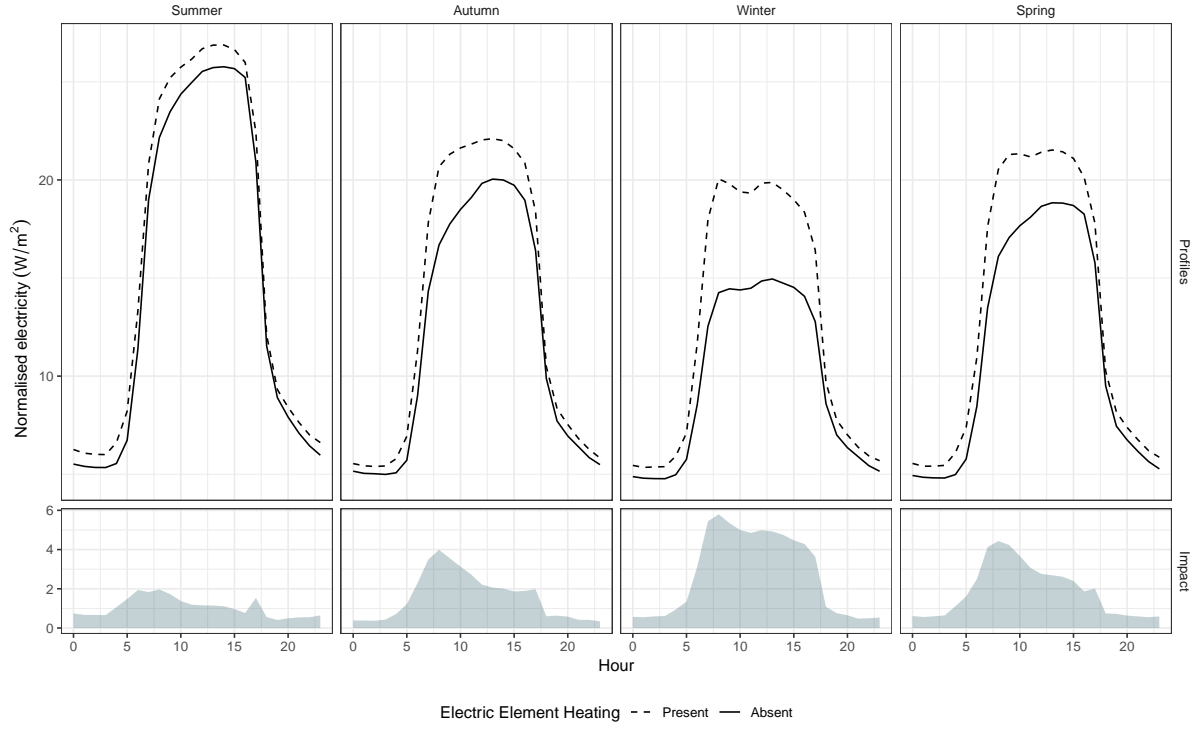
Figure 7: Profile plots of electric element heating impact. As expected the heating demand mainly plays a role in winter. The demand in autumn and spring is likely caused by cold days during the shoulder months of winter.
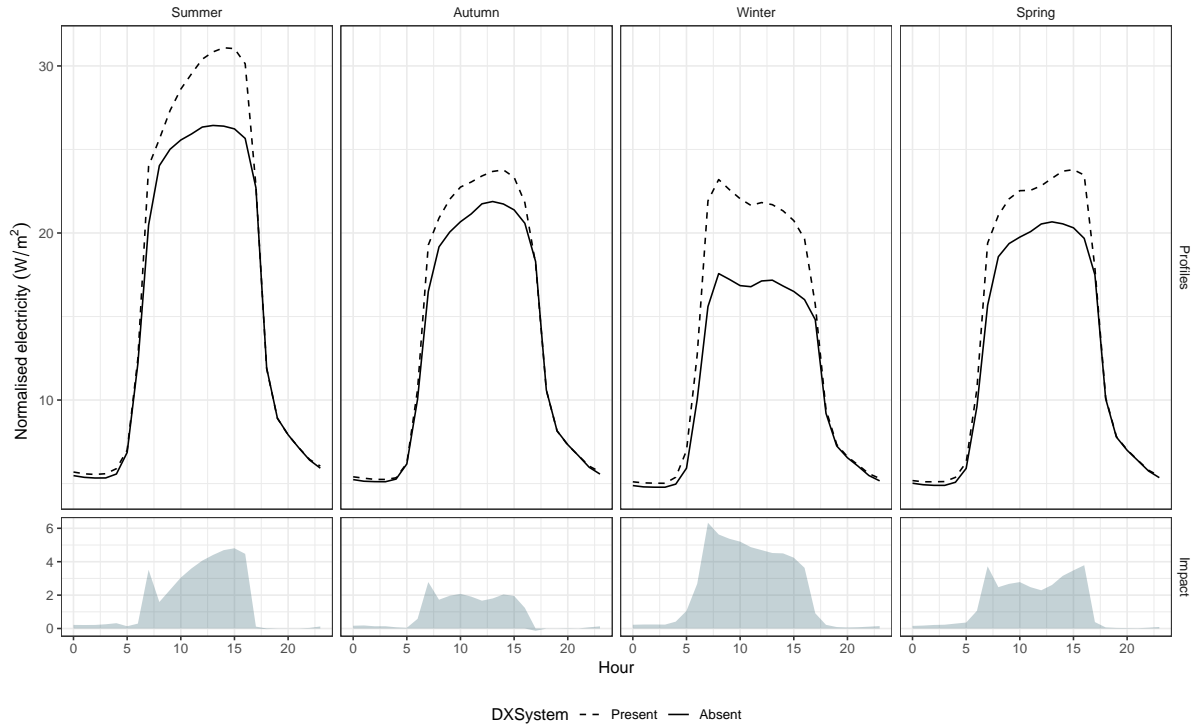
Figure 8: Profile plots of DX system impact. Buildings fitted with DX systems use more electricity during winter months, which may be representative of reverse DX systems being used for heating. Confidence intervals for the estimated coefficients show the winter results are statistically significant during business hours.
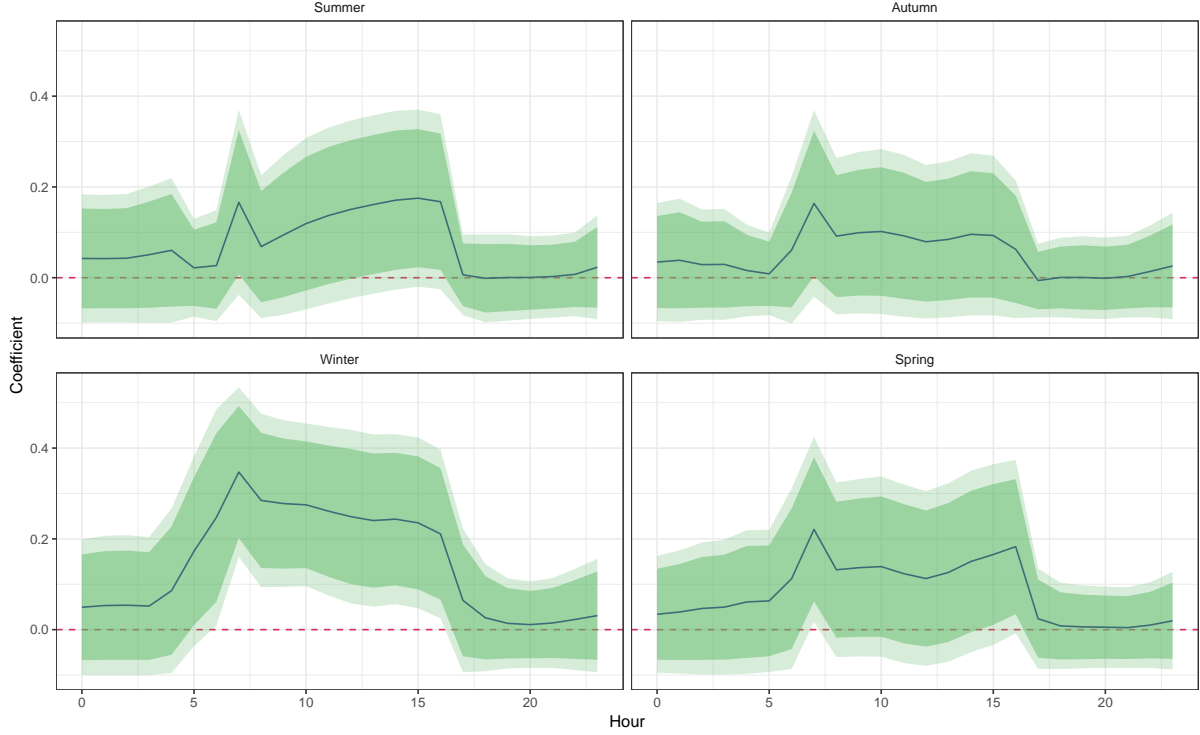
Figure 9: Coefficient profile plots for DX system coefficients. The solid line shows the estimated coefficients, β_h, for each hourly model described in equation (1). The 80% and 90% confidence intervals are indicated by the shaded ribbons. Coefficient estimates for winter business hours are significant.

or how it is typically operated. Larger systems tend to be more efficient than smaller ones and centralised systems tend to have lower cooling loads then per-floor systems as they cannot be influenced (easily) by tenants. While it is theoretically possible to work around this by adding new attribute variables, it quickly increases the dimensionality of the model which then requires more buildings when fitting. More building data and detailed attribute descriptions are required to facilitate fine-grain analysis such as this.

When using statistical models any analysis is limited by the quality and breadth of the available data. Survey results or building characteristics are needed for multiple buildings before any models can be fit. It is important to be aware that when statistical findings do not match with expected behaviour further investigation should be carried out.

4.4 Further research

Including more buildings will allow for more confident conclusions to be reached and more characteristics to be investigated. Of particular interest could be applying our approach to the residential sector to quantify the impact of solar generation, batteries or other household items. Given the prevalence of smart meter data available to both government and distributors, only a survey of household goods would be required. One interesting research direction involves assessing the impact of policies on electricity consumption. If separate models are fit for each year, we can observe how the demand impact profiles are changing between each yearly model. If a policy is introduced to reduce demand for a particular building attribute, we should see a decrease in the profiles between the yearly models. This may help us to determine if polices are proving to be effective or not.

Our mixed effects approach may also be adapted to estimate demand profiles for other building portfolios or energy sources. As long as appropriate goodness of fit measures are calculated and the statistical significance of variables are tested, most modelling pitfalls and erroneous conclusions should be avoided. Different settings will require consideration of which building characteristics to include and possible modifications to the formula we presented. However, the key idea of splitting data into buckets across the day (hours in our case) and fitting mixed effects models to each should still be applicable. Doing so is beyond the scope of this paper and is left as a future research direction.

Including the relationship between demand and weather variables may offer further improvements to our mixed effects modelling. Despite pursuing a simple formulation throughout our analysis and only adding a trend term to capture yearly changes in climate, we do expect including weather variables would further improve model fits. Doing so also removes them as potential confounding variables.

5 Conclusion

This paper presents an approach to estimating the impact different building characteristics may have on electricity demand of commercial office buildings. Our approach only requires that smart meter data and attribute data is available for several buildings. The impact that each building characteristic has on energy demand can be presented in the form of demand impact profiles which show the expected change in demand should a characteristic be absent or present. Key advantages of this approach are that it does not require equipment sub-metering and that it can be used to estimate the impact of indirect and behavioural factors.

In addition to providing demand impact profiles, our methodology also produces confidence intervals based on a statistical framework. This allows us to assess the statistical significance of each building characteristic. Mixed effects models are used to account for the correlation within each building's meter readings. A multimodel inference approach, which fits multiple models and weights coefficient estimates by the strength of each model fit, allows us to take both parameter estimation and model selection uncertainty into account when calculating confidence intervals. Hence, we can produce estimates of the effect size and the statistical significance of each characteristic's effect on electricity demand.

To justify the validity of our approach we have conducted a simulation study which shows our model estimates latent variables well. Furthermore, we also describe a case study using 129 commercial office buildings from across Australia. We were provided with several building attributes that engineers wished to assess statistically. Applying our methodology allowed us to understand how tenant behaviour and equipment behaviour affected electricity demand across the day at different times of the year. Furthermore, it also highlighted that some characteristics did not appear to have a statistically significant impact on energy demand.

It is our hope that this approach will open the door to further research in this area. Combining smart meter data from multiple buildings with additional descriptive data sources will allow us to quantify electricity consumption patterns. This has potential applications in benchmarking and assessing the efficacy of demand management programs. Such research can play a role in helping us understand the potential of energy efficient practices by highlighting which equipment or design choices result in lower demand. This in turn may help managers and policy makers to enact better, data-driven decisions.

Acknowledgements

A   Profiles of all attributes

Hourly coefficient profiles for each attribute discussed in Section 2.2 are shown in Figure 10. These are produced using equation (1) and the multimodel inference approach. The coefficients represent the proportional change in demand when an attribute is or isn't present. Unconditional confidence intervals are also shown which allow us to gauge how statistically significant each attribute's impact is. Note that these coefficients and confidence intervals are calculated separately for each hour of the day to take the changing building dynamics into account.

We can draw several conclusions from Figure 10. For example, some of the attributes only have a statistically significant impact at certain times of the year. For example, electric element heating is not significant during summer, but does have a statistically significant impact during business hours in winter. A similar result is observed for DX systems. Interestingly, centralised distribution does not appear to be statistically significant at any time of the year. This indicates that in the buildings used in our analysis, the normalised demand does not appear to be affected by whether centralised or per-floor distribution is present. Water cooled condenser and gas fired boiler variables do not appear to play any statistically significant role in energy demand during any of the seasons. We can always see significant coefficient estimates for tenant feed, which is unsurprising given the amount of energy tenants are using.

B   Simulation results

To validate our approach we conduct a simulation with known fixed effect coefficient values. If we assume that each of the estimated coefficients shown in Figure 10 are correct, we can simulate time series for multiple buildings and test if our proposed methodology can correctly estimate these effect sizes.

Using the estimated winter coefficients shown in Figure 10 we simulate demand for building j at time t using

$$\log y_{tj} = \sum_a X_{ja} A_{ah} + \epsilon_{tj}, \quad \epsilon_{tj} \sim ARMA(1,1), \tag{11}$$

where

- $A_{ah} \sim \mathrm{N}(a_{ah}, \sigma_a)$ is the assumed known impact of attribute a during hour h,
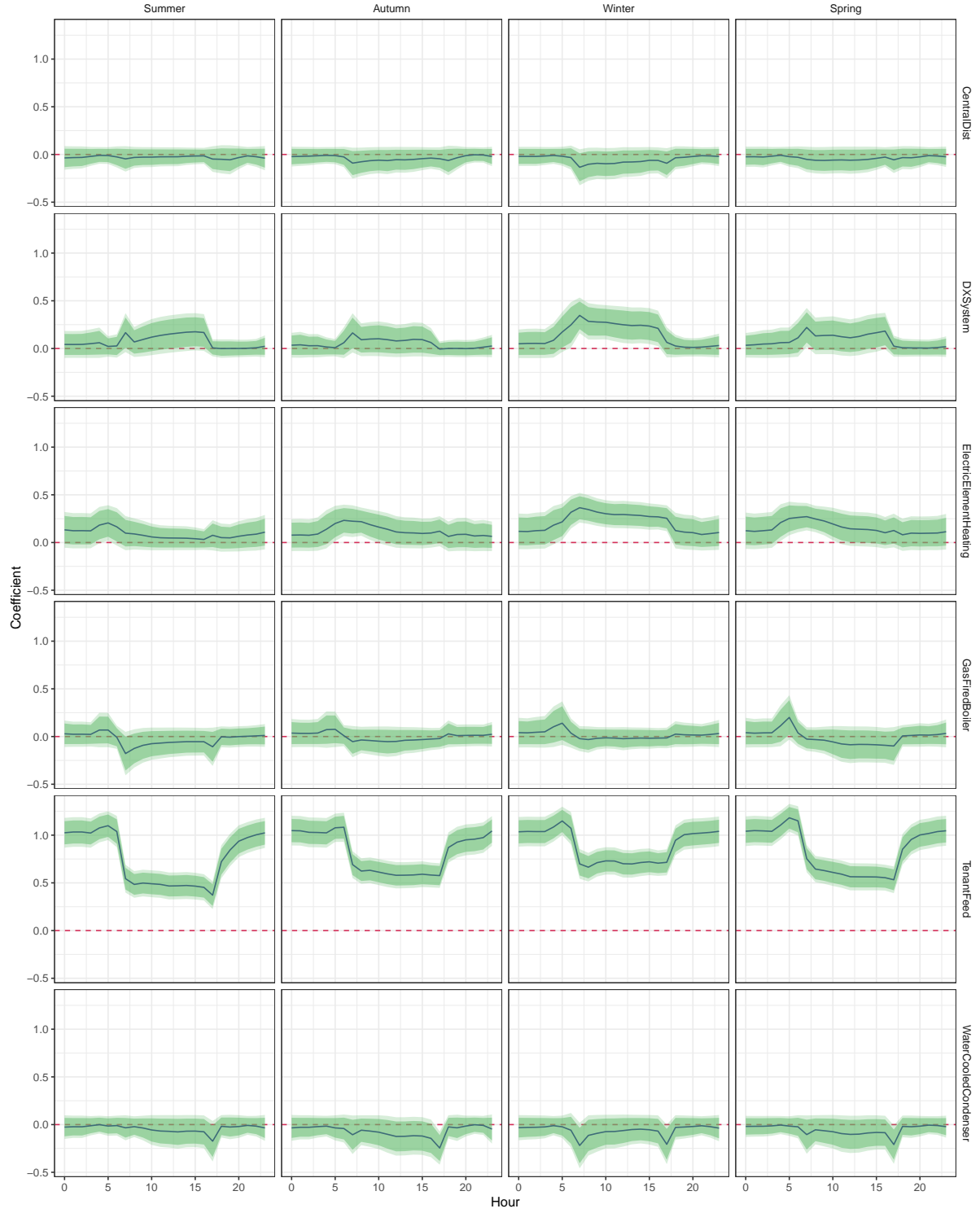- a_{ah} is the estimated coefficient,

Figure 10: Coefficient profile plots for all attribute coefficients. The solid line shows the estimated coefficients, β_h, for each hourly model described in equation (1). The 80% and 90% confidence intervals are indicated by the shaded ribbons.
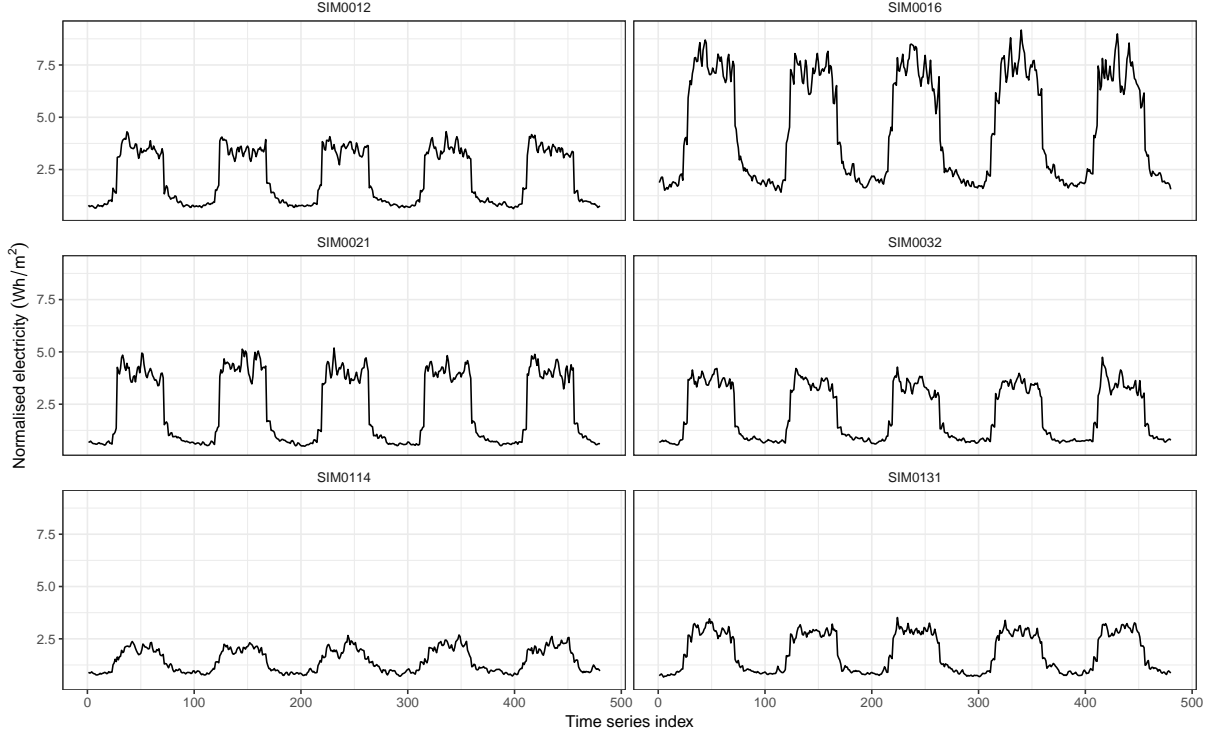
Figure 11: Time series plots of simulated smart-meter data for six of the 150 simulations. As observed with the real data, each building has distinct business hour and non-business hour behaviour and the load profiles differ between each building.

- σ_a is the standard error of the fixed effect coefficients, and
- $X_{ja} \sim \text{Bernoulli}(p_a)$ is a random variable indicating if attribute a is present for building j.

The probabilities that an attribute was present for a building, p_a, were chosen based on the relative frequency of attributes in the 129 commercial office buildings. The expected impact of an attribute is a_{ah} and is assumed to be equal to the estimates we produced in our analysis using real data. The variance σ_a allows for the impact of attribute a to vary between buildings and is chosen to be close to the estimated standard error for the fixed effects. Most fixed effects had standard error close to 0.3 and so for simplicity we chose $\sigma_a = 0.3$ for all attributes. As we are working with time series data we allow for autocorrelation in the model errors by treating them as an $ARMA(1,1)$ process with $\phi_1 = 0.6$, $\theta_1 = 0.6$ and $\sigma_\epsilon = 0.05$. Doing so produced residuals with similar standard deviation to those observed in the real data. As a final check, we plotted the simulated time series as shown in Figure 11 to ensure that they looked reasonable. The simulated time series appear to be realistic when compared to the time series in Figure 3.

We simulated 90 days of 15-minute demand data for 150 buildings using the estimated profiles from winter. Applying our mixed effects multimodel inference approach to this data produced the fixed effect estimates and confidence intervals shown in Figure 12. We see that our estimates and confidence intervals perform quite well in estimating the actual values.

References

Akaike, Hirotogu (1998). "Information Theory and an Extension of the Maximum Likelihood Principle". en. In: *Selected Papers of Hirotugu Akaike*. Springer Series in Statistics. Springer,
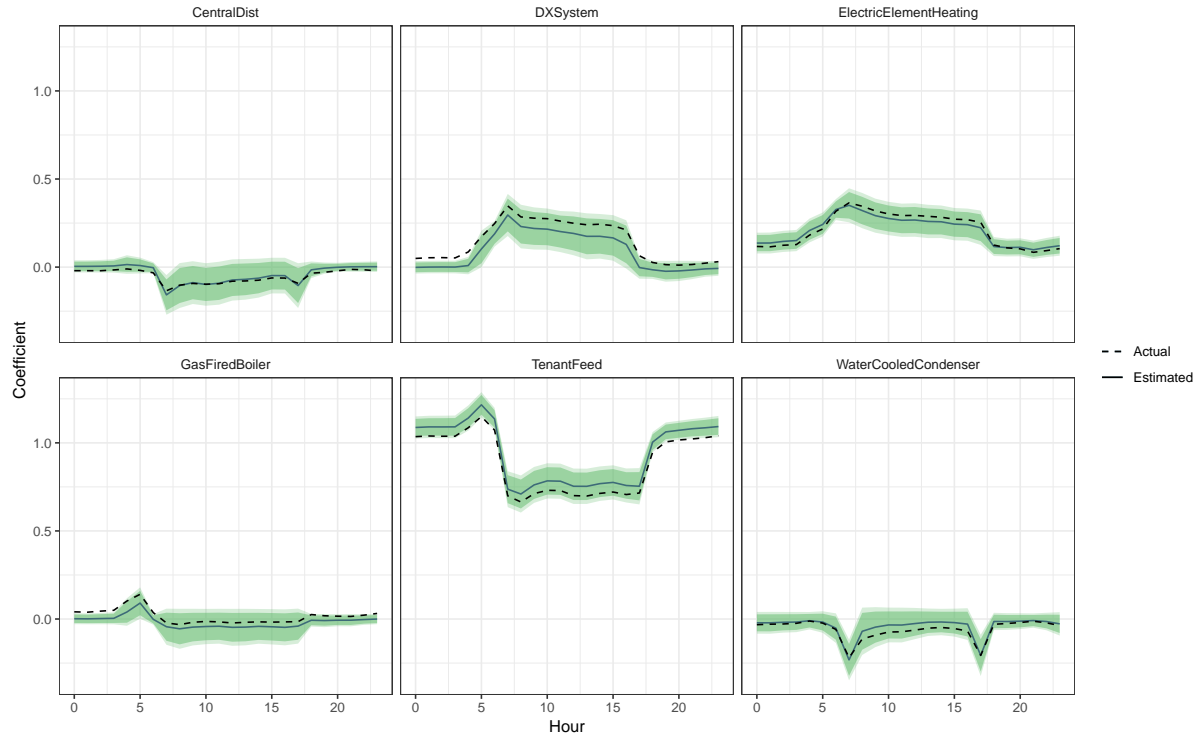
Figure 12: Coefficient profile plots based on simulated data for all attribute coefficients. The solid line shows the estimated coefficients for each hourly model and the dashed line shows the known fixed effect values that were chosen. 80% and 90% confidence intervals are indicated by the shaded ribbons. We see that our methodology is able to provide good estimates of the fixed effects.

618   New York, NY, pp. 199–213. URL: https://link.springer.com/chapter/10.1007/978-1-4612-
619   1694-0_15.

620   Bartoń, Kamil (2019). *MuMIn: Multi-Model Inference*. R package version 1.43.3/r450. URL:
621   https://R-Forge.R-project.org/projects/mumin/.

622   Bates, Douglas, Martin Mächler, et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4".
623   In: *Journal of Statistical Software, Articles* 67.1, pp. 1–48. URL: https://www.jstatsoft.org/
624   v067/i01.

625   Bates, Douglas, Martin Maechler, et al. (2019). *lme4: Linear Mixed-Effects Models using 'Eigen'*
626   *and S4*. R package version 1.1-21. URL: https://CRAN.R-project.org/package=lme4.

627   Burnham, Kenneth P and David R Anderson (Dec. 2003). *Model Selection and Multimodel*
628   *Inference: A Practical Information-Theoretic Approach*. en. Springer Science & Business
629   Media. URL: https://market.android.com/details?id=book-fT1Iu-h6E-oC.

630   Burnham, Kenneth P and Gary C White (Jan. 2002). "Evaluation of some random effects
631   methodology applicable to bird ringing data". In: *Journal of applied statistics* 29.1-4, pp. 245–
632   264. URL: https://doi.org/10.1080/02664760120108755.

633   Chahine, Khaled et al. (Jan. 2011). "Electric Load Disaggregation in Smart Metering Using
634   a Novel Feature Extraction Method and Supervised Classification". In: *Energy Procedia* 6,
635   pp. 627–632. URL: http://www.sciencedirect.com/science/article/pii/S1876610211014834.

636   Christiansen, Nils et al. (Nov. 2015). "Electricity consumption of medical plug loads in hospital lab-
637   oratories: Identification, evaluation, prediction and verification". In: *Energy and Buildings* 107,
638   pp. 392–406. URL: http://www.sciencedirect.com/science/article/pii/S0378778815302073.

639   Dinesh, C et al. (Nov. 2016). "Residential Appliance Identification Based on Spectral Information
640   of Low Frequency Smart Meter Measurements". In: *IEEE transactions on smart grid* 7.6,
641   pp. 2781–2792. URL: http://dx.doi.org/10.1109/TSG.2015.2484258.

642   Giles, David E (2011). "Interpreting dummy variables in semi-logarithmic regression models: exact
643   distributional results". In: *University of Victoria Department of Economics Econometrics*
644   *Working Paper* 1101. URL: http://web.uvic.ca/~dgiles/downloads/working_papers/ewp1101.
645   pdf.

646   Greven, Sonja and Thomas Kneib (Dec. 2010). "On the behaviour of marginal and conditional
647   AIC in linear mixed models". In: *Biometrika* 97.4, pp. 773–789. URL: https://academic.oup.
648   com/biomet/article-abstract/97/4/773/241321.

649   Gunay, H Burak et al. (June 2016). "Modeling plug-in equipment load patterns in private office
650   spaces". In: *Energy and Buildings* 121, pp. 234–249. URL: http://www.sciencedirect.com/
651   science/article/pii/S0378778816301268.

652   Guo, Z, Z J Wang, and A Kashani (Jan. 2015). "Home Appliance Load Modeling From Aggregated
653   Smart Meter Data". In: *IEEE Transactions on Power Systems* 30.1, pp. 254–262. URL: http:
654   //dx.doi.org/10.1109/TPWRS.2014.2327041.

655   Johnson, Paul Cd (Sept. 2014). "Extension of Nakagawa & Schielzeth's R2GLMM to random
656   slopes models". en. In: *Methods in ecology and evolution / British Ecological Society* 5.9,
657   pp. 944–946. URL: http://dx.doi.org/10.1111/2041-210X.12225.

658   Kalluri, Balaji et al. (Sept. 2016). "Applicability of using time series subsequences to study
659   office plug load appliances". In: *Energy and Buildings* 127, pp. 399–410. URL: http://www.
660   sciencedirect.com/science/article/pii/S0378778816304546.

661   Kennedy, Peter (1981). "Estimation with Correctly Interpreted Dummy Variables in Semiloga-
662   rithmic Equations". In: *The American economic review* 71.4. URL: https://EconPapers.repec.
663   org/RePEc:aea:aecrev:v:71:y:1981:i:4:p:801.

664   Liao, J et al. (Nov. 2014). "Non-intrusive appliance load monitoring using low-resolution smart
665   meter data". In: *2014 IEEE International Conference on Smart Grid Communications*
666   *(SmartGridComm)*, pp. 535–540. URL: http://dx.doi.org/10.1109/SmartGridComm.2014.
667   7007702.

Lukacs, Paul M, Kenneth P Burnham, and David R Anderson (2010). "Model selection bias and Freedman's paradox". In: *Annals of the Institute of Statistical Mathematics* 62.1, p. 117. URL: https://link.springer.com/article/10.1007/s10463-009-0234-4.

Mahdavi, Ardeshir, Farhang Tahmasebi, and Mine Kayalar (Oct. 2016). "Prediction of plug loads in office buildings: Simplified and probabilistic methods". In: *Energy and Buildings* 129, pp. 322–329. URL: http://www.sciencedirect.com/science/article/pii/S0378778816307071.

Menezes, A C et al. (June 2014). "Estimating the energy consumption and power demand of small power equipment in office buildings". In: *Energy and Buildings* 75, pp. 199–209. URL: http://www.sciencedirect.com/science/article/pii/S0378778814001224.

Müller, Samuel, J L Scealy, and A H Welsh (May 2013). "Model Selection in Linear Mixed Models". en. In: *Statistical science: a review journal of the Institute of Mathematical Statistics* 28.2, pp. 135–167. URL: https://projecteuclid.org/euclid.ss/1369147909.

Nakagawa, S and H Schielzeth (2013). "A general and simple method for obtaining R2 from generalized linear mixed-effects models". In: *Methods in ecology and evolution / British Ecological Society.* URL: http://onlinelibrary.wiley.com/doi/10.1111/j.2041-210x.2012.00261.x/full.

Norford, Leslie K and Steven B Leeb (1996). "Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms". In: *Energy and Buildings* 24, pp. 5l–64. URL: https://pdfs.semanticscholar.org/d6b1/43ad1c53a9c5263b62e5534ef899425a23bd.pdf.

Perez, Krystian X et al. (Oct. 2014). "Nonintrusive disaggregation of residential air-conditioning loads from sub-hourly smart meter data". In: *Energy and Buildings* 81, pp. 316–325. URL: http://www.sciencedirect.com/science/article/pii/S0378778814005131.

Pinheiro, José C and Douglas M Bates (1978). *Mixed-Effects Models in S and S-plus.* Statistics and Computing. Springer.

R Core Team (2018). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Reinhardt, A et al. (Oct. 2012). "On the accuracy of appliance identification based on distributed load metering data". In: *2012 Sustainable Internet and ICT for Sustainability (SustainIT)*, pp. 1–9. URL: https://ieeexplore.ieee.org/abstract/document/6388037/.

Schwarz, Gideon (Mar. 1978). "Estimating the Dimension of a Model". en. In: *Annals of statistics* 6.2, pp. 461–464. URL: https://projecteuclid.org/euclid.aos/1176344136.

So, Nicholas and Russell Richman (Jan. 2016). "A high level method to disaggregate electricity for cluster-metered buildings". In: *Energy and Buildings* 111, pp. 351–368. URL: http://www.sciencedirect.com/science/article/pii/S0378778815303844.

Symonds, Matthew R E and Adnan Moussalli (2011). "A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion". In: *Behavioral ecology and sociobiology* 65.1, pp. 13–21. URL: https://link.springer.com/article/10.1007/s00265-010-1037-6.

Vaida, Florin and Suzette Blanchard (June 2005). "Conditional Akaike information for mixed-effects models". In: *Biometrika* 92.2, pp. 351–370. URL: https://academic.oup.com/biomet/article-abstract/92/2/351/233128.

Weiss, M et al. (Mar. 2012). "Leveraging smart meter data to recognize home appliances". In: *2012 IEEE International Conference on Pervasive Computing and Communications.* ieeexplore.ieee.org, pp. 190–197. URL: http://dx.doi.org/10.1109/PerCom.2012.6199866.