Intra-day Electricity Demand and Temperature

James McCulloch[a] *and Katja Ignatieva*[b]

ABSTRACT

The objective of this paper is to explain the relationship between high frequency electricity demand, intra-day temperature variation and time. Using the Generalised Additive Model (GAM) framework we link high frequency (5-minute) aggregate electricity demand in Australia to the time of the day, time of the year and intra-day temperature. We document a strong relationship between high frequency electricity demand and intra-day temperature. We show a superior model fit when using Daylight Saving Time (DST), or clock time, instead of the standard (solar) time. We introduce the time weighted temperature model that captures instantaneous electricity demand sensitivity to temperature as a function of the human daily activity cycle, by assigning different temperature signal weighting based on the DST time. The results on DST and time weighted temperature modelling are novel in the literature and are important innovations in high frequency electricity demand forecasting.

Keywords: High Frequency, Electricity, Instantaneous Demand, Temperature, Generalised Additive Model (GAM)

https://doi.org/10.5547/01956574.41.3.jmcc

1. INTRODUCTION

This paper introduces a parsimonious model for modelling the relationship between high frequency electricity demand and intra-day temperature. Modelling frameworks developed in the literature for high frequency electricity data suggest to either use multi-equation modelling (Cottet and Smith, 2003; Cancelo et al., 2008; Soares and Medeiros, 2008) which treats each intra-day period as a different series, or univariate time series modelling (Darbellay and Slam, 2000; Smith, 2000; Taylor, 2003, 2010, 2012; Kim, 2013) that treats the entire time series as a single series. A comprehensive overview of electricity demand forecasting models is provided in Hahn et al. (2009). The major drawback of the first methodology is that it is a more complex model that can require a large number parameters to be estimated. A univariate time series model is parsimonious and, since our primary goal is to elucidate the relationship between temperature and high frequency demand, we choose a parsimonious univariate modelling approach for demand forecasting.

We use time of the day, time of the year and outside temperature as the main driving factors when predicting electricity demand at 5-minute frequency. To our knowledge, none of the existing research papers have incorporated these three variables simultaneously for demand forecasting at such high frequency. Temperature variations are documented to play a crucial role when forecasting electricity demand. Intuitively, during cold winter months electricity demand is expected to increase due to electrical heating, whereas during hot summer months the use of air conditioners

a Corresponding author. Kellerberrin, Australia. E-mail: james.duncan.mcculloch@gmail.com.

b UNSW Australia, Business School, School of Risk and Actuarial Studies, Sydney, NSW 2052, Australia. E-mail: k.ignatieva@unsw.edu.au.

and coolers also leads to increased electricity consumption. A non-exhaustive list of papers that deal with forecasting demand using low frequency data (typically, monthly or daily) includes Pardo et al. (2002), Moral-Carcedo and Vicens-Otero (2005), Bessec and Fouquau (2008), Hekkenberg et al. (2009), Lam et al. (2009), Tung et al. (2013), Moral-Carcedo and Perez-Garca (2015). None of these research works incorporates time of the day as an explanatory variable. Other weather variables, such as sunshine hours, rainfall, wind, humidity, cloudiness etc. are shown to have a much lower impact on electricity demand, see e.g. Basta and Helman (2013) and Moral-Carcedo and Perez-Garca (2015). Furthermore, focusing exclusively on the temperature allows us to avoid potential collinearity problems when simultaneously employing several weather variables as explanatory variables in the regression modelling[1], see e.g. Lam et al. (2009) and Moral-Carcedo and Perez-Garca (2015). Studies that have analysed intra-day (hourly) patterns in electricity demand based on the hour-of-the day and have documented the existence of this effect include Darbellay and Slam (2000), Taylor (2003), Mirasgedis et al. (2006), Soares and Medeiros (2008), Taylor (2010) and Kim (2013). The paper by Taylor (2010) also incorporates an intra-year effect.

This paper introduces a parsimonious Generalised Additive Model (GAM) forecasting model for high frequency intra-day (5-min) aggregate electricity demand. The parsimonious model allows us to focus exclusively on modelling the link between electricity demand and human activity cycle (modelled through the time of the day), the intra-day temperature variations and time of the year. High frequency data enables us to obtain interesting and novel insights into demand forecasting. Using yearly and seasonal demand models, we document a strong relationship between high frequency electricity demand and intra-day temperature. When examining intra-day demand using daylight saving time (DST), i.e. clock time and standard (astronomical) time, we show that using the DST (clock) time provides a significant improvement to the model fit. We explain how and why model fit improves even further when we introduce the *time weighted temperature model*, which assigns different temperature signal weighting based on the DST time. This relates the magnitude of the temperature demand signal to the human daily activity cycle. The motivation behind using the time weighted temperature model is the observation that electricity demand attributed to temperature variation away from the maximum comfort temperature (20.0 degrees Celsius[2]) is time sensitive: Electricity consumption is less sensitive to temperature variation away from the 'comfort' temperature late at night and early in the morning, which are time periods characterised by low human activity. At the same time, electricity is more sensitive to temperature variation away from the 'comfort' during periods of high human activity. We observe that the minimum morning sensitivity is at 4:00am, the morning maximum is reached at 9:00am and the night decline begins at 18:30pm. Our proposed methodology, which suggests to weight temperature demand signal depending on the DST time (daily activity cycle) is confirmed when using cross-sectional regressions estimated at each (5-minute) time interval, resulting in cross-sectional daily time dependent demand.

The contributions of this paper are as follows. Firstly, our results allow us to characterise the high frequency relationship between electricity consumption and temperature. To our knowledge, none of the papers in the existing literature model demand and temperature data at such high (5-minute) frequency. Secondly, we show that the sensitivity of electricity demand to temperature is a function of time of day. Daily times of low human activity such as 04:00am have a much lower demand/temperature sensitivity than a period of higher human activity as as 18:00pm. In addition, we demonstrate that when using DST (clock time) as an independent variable for electricity demand prediction, we obtain a significant improvement compared to using standard (astronomical) time.

1. For example, temperature is expected to be correlated with sunshine hours and cloudiness.
2. This threshold is chosen empirically to provide the optimal model fit.

We emphasise that the results on DST and time weighted temperature modelling are novel in the literature and are important innovations in high frequency electricity demand forecasting. Finally, this is the first study that predicts electricity demand in Australia using both the outside temperature and time of the day.

The parsimonious GAM model is accurate, with a MAPE (Mean Average Predicted Error) next day forecast error of 3.20%.[3] The best publicly available electricity demand forecast is the Australian Energy Market Operator (AEMO) next day forecast with a MAPE error of 2.04%. The AEMO forecasting model is not publicly available and it is reasonable to infer the AEMO forecasting model uses multivariate prediction variables with advanced but opaque forecasting techniques such as Deep Neural Networks or Gradient Boosting. The parsimonious GAM model allows for a transparent understanding of the high frequency interaction of temperature and demand and provides a solid foundation for the development of more accurate and complex forecasting models.

The remainder of the paper is organised as follows. Section 2 describes data used in our analysis. GAM model specification and its variations tested in the paper are introduced in Section 3. An extensive empirical analysis demonstrating the quality of fit of the proposed models to the entire data set as well as seasonal models is presented in Section 4. Section 5 deals with the prediction results for the electricity demand, and Section 6 concludes the paper.

2. DATA

This section discusses and provides some preliminary analysis on the data used in this study, which will enable us to formulate appropriate model specifications in Section 3.

2.1 Data Description

We use instantaneous intra-day electricity demand in Megawatts (MW), available at 5 minute frequency for the Australian state of New South Wales (NSW) and the Australian Capital Territory (ACT) for the year 3-February-2014 to 2-February-2015. Demand is aggregate data (i.e. including households, companies, industrial and public sectors) that has been downloaded from the Australian Energy Market Operator (AEMO) website.[4] The electricity demand observations are merged with instantaneous temperature data over the same period and frequency. The temperature data was obtained from the Australian Government Bureau of Meteorology.[5]

We have restricted our consideration to business days only,[6] which results in 250 days of data and each day of data has 12*24 = 288 five minute demand observations, from 00:00–00:05 until 23:55–24:00. Thus, a total of 72,000 five-minute demand and temperature data points will be used for the empirical analysis.

The 5-minute temperature data is recorded in the Sydney suburb of Homebush, which is a suburb located close to the population centre of the Greater Sydney urban area. If we assume that the Homebush temperature represents the instantaneous temperature in Greater Sydney, then this temperature observation is valid for 61% of population of the NSW/ACT electricity demand area.

3. The yearly model 5 forecast MAPE error, for further details refer to the section on prediction (Section 5).

4. Current and archived data are available from http://www.nemweb.com.au/REPORTS/CURRENT/ and http://www.nemweb.com.au/REPORTS/ARCHIVE/, respectively. Python software has been used to download these forecasts.

5. Available from http://www.bom.gov.au.

6. Time dependent intra-day variation of electricity is significantly different for weekends and public holidays compared to business days. Since we are primarily interested in modelling the high frequency relationship between temperature and demand, we have restricted our consideration to business days only.

However, it should be noted that the assumption that we can represent temperature related electricity demand in the NSW/ACT demand area with a single temperature is a deliberate simplification to preserve the parsimonious property of our modelling. An obvious improvement to the accuracy of modelling electricity demand as a function of temperature would be to use multiple temperature (and potentially humidity) measurements from different suburban, urban and rural areas and set up an adequate average temperature index weighted by population.

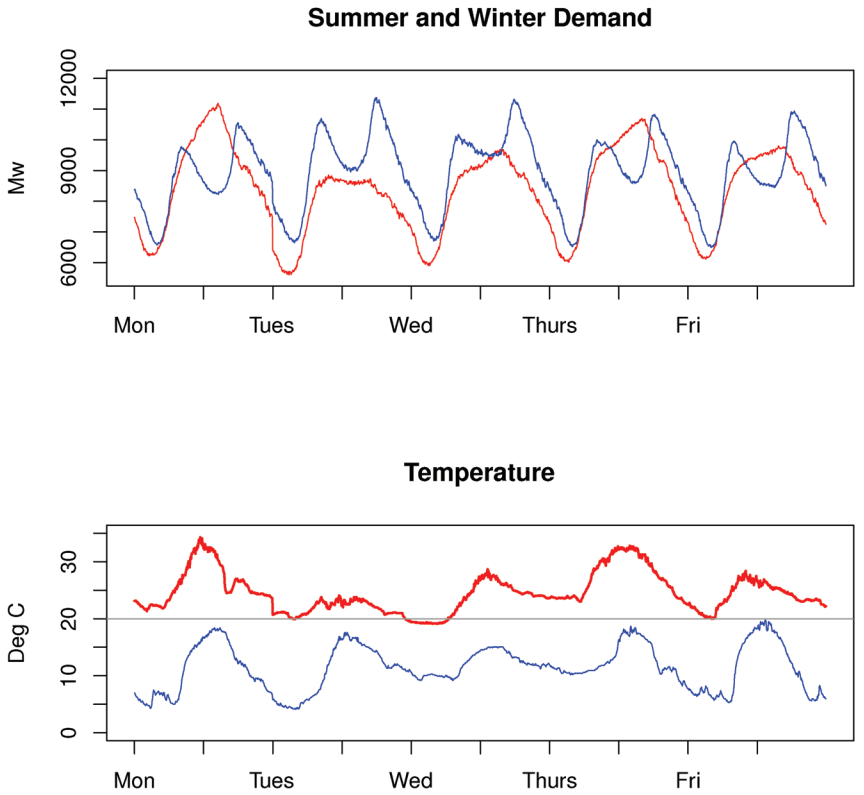
To provide an idea of the relationship between demand and temperature data, we show in Figure 1 winter and summer patterns of demand (top panel) and temperature (bottom panel) over the five business days of a typical week. In both, summer (red line, 12-January-15 to 16-January-15) and winter (blue line, 14-July-14 to 18-July-14) graphs we observe a cyclical pattern in evolution of temperature and demand throughout the week. The winter demand graph experiences two daily peaks, which correspond to an increasing usage of heating during the cold morning and evening hours, while the demand during the day (when the outside temperatures are relatively high) drops to a lower level. The summer demand graph shows peak demand during the hot afternoon hours caused by the use of coolers and air-conditioners. The winter and summer demand patterns suggest presence of a so-called minimum demand 'comfort' temperature, such that the magnitude of the difference between the current temperature and the 'comfort' temperature is strongly correlated with demand. We will demonstrate in Section 3 that the value of the minimum demand 'comfort' temperature can be set equal $20°C$. The magnitude of the difference between the current temperature and the 'comfort' temperature (abs[$Temp-20$]) is strongly correlated with demand. For example, Monday, January 12, 2015 was a hot day and the corresponding demand graph shows a pronounced spike mid-afternoon. The double morning and evening peaks of the winter pattern are pronounced because temperatures are well below $20°C$, with a dip in demand in the afternoon as the temperature rises. However, it is interesting to note that these peaks do not correspond to the minimum temperature early in the morning, but correspond to *lower temperatures when people are active*. This is clearly seen by examining the winter demand pattern in the early morning. The minimum demand for Tuesday, July 15, 2014 at 4am is essentially the same as the minimum demand for Thursday, July 17, 2014 at 4am, even though Tuesday 4am is considerably colder than Thursday 4am. The reason for this is intuitive: There are low levels of activity at 4am and the effect of the stronger temperature demand signal on Tuesday is attenuated by low personal and economic activity. Therefore, the *temperature demand signal is time dependent*. The relationship between temperature and demand shown in Figure 1 is our primary motivation for modelling intra-day electricity demand as a function of temperature.

2.2 Data Exploration

To get an idea about characteristics of the data, and in particular, obtain some insights on sensitivity of the demand with respect to temperature, we perform some preliminary analysis that will assist us in model development in Section 3. One advantage of using high frequency temperature and demand data is that it gives us a large amount of data: We have 250 days, and 288 5-minute observations on each day, which allows us to perform cross sectional regressions using 250 observations at each individual 5 minute data point, thus, leading to a total of 288 regressions. We emphasise that these cross sectional regressions are not considered to be a model (unlike models that will be specified in Section 3, which will be developed in order to predict demand out-of-sample), but are used merely as a tool to examine the data.

We fit the following linear model 288 times to each 5 minute period during the day:

Figure 1: Top panel: Summer demand (red line 12-Jan-15 to 16-Jan-15) and winter demand (blue line: 14-Jul-14 to 18-Jul-14). Bottom panel: Temperature in summer (red line) and winter (blue line).



Notes: The 'comfort' (black line) in the temperature graph is 20°*C*. The graph shows that the temperature demand signal is time dependent.

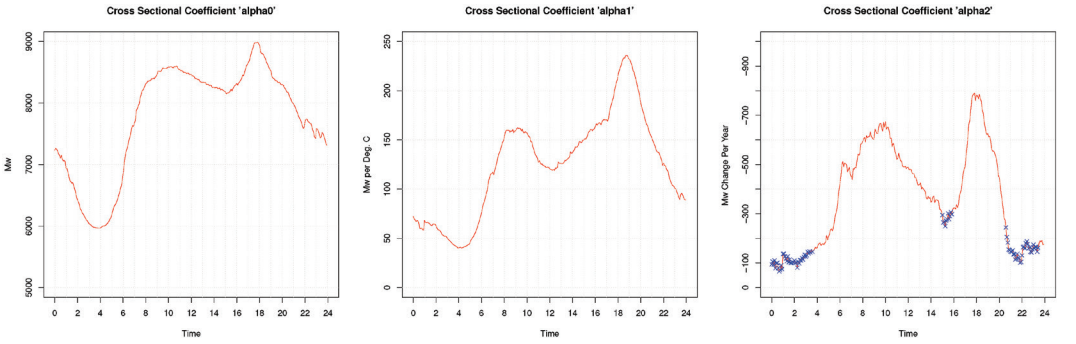
$$D = \alpha^0 + \alpha^1 |Temp - 20.0| + \alpha^2 Year + \varepsilon, \tag{2.1}$$

where the dependent variable D is electricity demand in Megawatts (MW).

The independent variable $|Temp - 20.0|$ is the absolute value of the difference of the recorded temperature and 20°*C*. We refer to the constant temperature of 20°*C* as the 'comfort' or minimum demand temperature (see Section 3 for comfort temperature analysis). This result is well known in the literature and is basis for "degree-days" commonly used in modelling of the temperature dependence of different economic variables, such as electricity demand, as well as energy derivatives. Further details on the approaches studying the impact of temperature on electricity demand using heating degree days (HDD) and cooling degree days (CDD) can be found, e.g., in Al-Zayer and Al-Ibrahim (1996), Sailor and Munoz (1997), Valor et al. (2001), Sailor (2001), Pardo et al. (2002); Amato et al. (2005) and Xiao et al. (2007).

The independent variable *Year* is the scaled time of the year with values in the interval [0,1), where 0 corresponds to the first data record of electricity demand for the 5 minute standard time period 3-February-2014 00:00–00:05 (DST 3-February-2014 01:00–01:05) and ((365*288)–1)/ (365*288)=0.999990487 corresponds to the final electricity demand record for the standard time period 2-February-2015 23:55–24:00 (DST 3-February-2015 00:55–01:00).

Figure 2: This graph displays 288 time indexed coefficients for the exogenous time dependent demand $\alpha_t^0, \in [1,...,288]$ (left panel), time demand sensitivity to temperature $\alpha_t^1, \in [1,...,288]$ (middle panel) and cross sectional change in yearly demand $\alpha_t^2, \in [1,...,288]$ (right panel), obtained using cross sectional regression in Equation (2.1).



Notes: All coefficients α_t^0 and α_t^1 are highly significant. In the right panel the coefficients marked with a blue cross are not significant at the 99% confidence level. The y-axis is (inverted) decline in demand in MW for the year.

Fitting cross sectional regressions from Equation (2.1) to the data gives us a 288 point vector of time indexed coefficients α^0, α^1 and α^2. Here, $\alpha_t^0, t \in [1,\ldots,288]$ are the exogenous daily demand cycle coefficients. If temperature corresponds to 'comfort' or minimum demand temperature of $20°C$, and $Year = 0$, which corresponds to the first data record of electricity demand, then the dependent demand D will correspond to exactly α_t^0. The empirical time indexed demand sensitivity to temperature is given by coefficients $\alpha_t^1, t \in [1,\ldots,288]$, and $\alpha_t^2, t \in [1,\ldots,288]$ are coefficients for the cross sectional change in yearly demand. Figure 2 summarises coefficient estimates α_t^0, α_t^1 and α_t^2 in the left, middle and right panel, respectively. Each of the 288 fitted coefficients α_t^0 (left panel) is highly significant with a median t-statistic of 130 and a median standard error of 59. The cross sectional empirical time demand sensitivity to temperature coefficients $\alpha_t^1, t \in [1,\ldots,288]$ (middle panel) represent the time variation in demand due to temperature in MW per degree. They are highly significant with a median t-statistic of 15.1 and a median standard error of 8.0. The y-axis represents additional demand in MW generated by each degree variation from the 'comfort' temperature ($|Temp - 20.0|$). The pattern of the cross-sectional empirical time demand sensitivity to temperature is intuitive and relates to the daily activity cycle: The minimum morning sensitivity is at 4:00am, the morning maximum is reached at 9:00am and the night decline begins at 18:30pm. We observe that electricity demand due to temperature variation away from the maximum comfort temperature ($20°C$) is time sensitive. For example, if the temperature at 4am is $12°C$ which is $8°C$ colder than the minimum demand comfort temperature, the increase in demand (the coefficient α^1 in MW per degree) will be relatively small because of low human activity at 4am. Conversely, if it is $12°C$ at 9am (again $8 = |20 - 12|$), temperature related increase in demand (the coefficient α^1 in MW per degree) will be much greater because of the higher levels of human activity at 9am. Thus, electricity consumption is less sensitive to temperature variation away from the 'comfort' temperature late at night and early in the morning. The result on time sensitivity of demand due to temperature is a major result in this paper, which will be confirmed using models in Section 3 and empirical results in Section 4. To our knowledge, this result does not appear in the literature. Finally, a 288 point vector of year coefficients from the cross sectional regressions $\alpha_t^2, t \in [1,\ldots,288]$ (right panel) shows the decrease in electricity demand across the year as a function of time (the coefficient α^2 in MW

per year). One observes that the yearly decline in demand has been concentrated during the daylight hours (demand replacement with solar power), and in particular, the greatest demand falls have been in the morning and afternoon peak periods. This may also indicate that peak demand pricing is causing time insensitive electricity consumers to shift demand to non-peak periods. This is supported by the explanation for the decline in electricity demand, proposed by the Australian Energy Market Operator (AEMO, 2014), which is related to "energy efficiency savings in response to high electricity prices over recent years". The fact that the decline in demand is concentrated in the peak demand periods and not spread more uniformly across the day has very important implications for models that we will introduce in Section 3.

3. MODEL SPECIFICATION

The objective of our analysis is to develop a comprehensive high frequency modelling framework to link electricity demand to the outside temperature that can be used for out-of-sample prediction of demand. For these purposes we introduce the Generalised Additive Model (GAM), see e.g. Hastie and Tibshirani (1990), Wood (2006). Specifically, at each point in time t, $t = 1,...,N$ with N being the total number of observations, we link the demand to the temperature in the following way:

$$\textbf{Model 1}: \quad D_t = \beta_0 + s(Time_t) + \beta_1 \,|\, Temp_t - 20.0 \,| + \beta_2 Year_t + \varepsilon_t, \tag{3.1}$$

where D_t is demand and $Temp_t$ is temperature. $Time_t$ is a number in the interval [0,1) where 0 is the time recorded for the electricity demand in the 5 minute period 00:00–00:05, $1/(12*24) = 0.0034722$ is the time recorded in the period 00:05–00:10 and $((12*24)-1)/(12*24) = 0.9965278$ is the time recorded for the 5 minute period 23:55–00:00. There are two time fields for each demand record in the data; standard (astronomical) time and the DST.[7] During the period when DST is active, the DST field is advanced by 1 hour or $0.041667 = 1/24$. For example, for the 5 minute time period on the 3rd of February 2014 (DST is active) where the standard time is recorded as 00:00–00:05 ($Time_t = 0.0$), the DST time is recorded as 01:00–01:05 ($Time_t = 0.041667$). The instantaneous electricity demand (D_t) recorded for this 5 minute period was 7135.67 megawatt/hour. Outside of the period where DST is active, the DST and standard time fields are equivalent.

It is assumed that the *temperature independent electricity demand* $s(Time_t)$ is a daily periodic cyclic empirical function of $Time_t$ over the sample period. We use GAM regression to determine the periodic function $s(\cdot)$ of $Time_t$. This periodic function is a *cyclic cubic spline*. A cyclic cubic spline function is a piecewise cubic function continuous up to second derivatives at the knots. At the endpoints of each daily cycle, the function values and derivatives up to the second order are equal, which creates a smoothed periodic function.[8] We notice that the function $s(\cdot)$ can be specified with a smoothing parameter (number of spline knots), or the number of degrees of freedom (df, the number of spline knots-2), which is assumed to be larger than one (with $df = 1$ corresponding to a linear fit). The df parameter is chosen in such a way that it leads to the best goodness-of-fit measured by the Akaike Information Criterion (AIC), and we find that $df = 10$ is optimal. The GAM in Equation (3.1) has a linear predictor that is specified in terms of a sum of smooth functions of

7. It is intuitive (and will be shown below) that personal and economic activity is linked to daylight saving time (DST), i.e. clock time rather than the actual (standard) time. The standard time is astronomical time. Daylight saving time (+1 hour) in Sydney commences at 2am on the first Sunday in October and the change from daylight saving (–1 hour) to standard time is 3am on the first Sunday in April. In our sample period (from February 3, 2014 to February 2, 2015) daylight saving ends on April 6, 2014 (clocks turned back from 3am to 2am) and starts on October 5, 2014 (clocks turned forward from 2 am to 3 am).

8. R function *s(Time, bs = "cc")* is used.

predictor variables (Wood, 2006). This technique is particularly suited to modelling intra-day actual electricity demand as a function of the time dependent electricity demand due to daily personal and economic activity.

The *temperature dependent electricity demand* in Equation (3.1), $|Temp - 20.0|$, is the absolute value of the difference of the recorded temperature and the 'comfort' or minimum demand temperature $20°C$. The value of $20°C$ is determined empirically as as follows: We fit Model 1 in Equation (3.1) using different temperatures ranging from $17°C$ to $23°C$. These temperatures are used as minimum demand temperatures $Temp_{min}$ in the temperature dependent term of the demand regression, $\beta_1 |Temp_t - Temp_{min}|$, i.e., $Temp_{min} = \{17,18,19,20,21,22,23\}$. The objective is to analyse which minimum demand temperature results in a better fit. The regression results (DST is used as the time index) for the temperature dependent demand term $\beta_1 |Temp_t - Temp_{min}|$ are reported in Table 1. The results show that $20°C$ is the *marginally* optimal constant for the minimum demand temperature since the regression temperature dependent coefficient $\beta_1 |Temp_t - 20|$ has a slightly higher t-statistic and the regression has a slightly higher R^2_{adj}.[9] This result will be confirmed graphically in Section 4 when fitting the demand as a function of temperature. We note, however, that the difference between $20°C$ and $19°C$ or $21°C$ is very small, which implies that temperature dependent demand is a non-linear function of the difference between the minimum demand temperature and the measured temperature, with small differences producing little or no change in temperature dependent demand. This possibility has been examined in empirical Section 4.3 where we fitted Model 3 (Equation (3.3)) that captures the relationship between the temperature and demand via a non-linear function modelled by non-periodic splines.

Table 1: Estimation results for regression in Equation (3.1) (Model 1) with different minimum demand temperatures.

	β_1 Estimate	β_1 Std.error	β_1 t-value	R^2_{adj}
$\beta_1 \|Temp - 17\|$	95.2***	0.58	163.9	0.795
$\beta_1 \|Temp - 18\|$	108.4***	0.54	202.4	0.820
$\beta_1 \|Temp - 19\|$	112.0***	0.49	226.2	0.835
$\beta_1 \|Temp - 20\|$	108.1***	0.47	229.3	0.837
$\beta_1 \|Temp - 21\|$	99.3***	0.46	215.5	0.829
$\beta_1 \|Temp - 22\|$	89.4***	0.46	194.9	0.816
$\beta_1 \|Temp - 23\|$	79.4***	0.46	172.7	0.801

Notes: ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

The third term $Year_t$ in Equation (3.1) is a linear long term drift in average electricity demand, as introduced in Section 2.2.

The GAM regression models assume the residual term ε_t to be Gaussian with zero mean.[10] We also introduce extensions of Model 1, which will be shown to enhance the quality of fit provided by Model 1 (refer to Section 4). Model 2, that incorporates a weighted temperature demand signal, is given by the following equation:

$$\textbf{Model 2:} \quad D_t = \beta_0 + s(DST_t) + \beta_1 (w(DST_t) * |Temp_t - 20.0|) + \beta_2 Year_t + \varepsilon_t \tag{3.2}$$

9. Note that this minimum demand temperature is slightly higher than the $18°C$ used as the reference minimum demand temperature in energy derivatives. For a description of the over-the-counter (OTC) weather derivatives traded on the Chicago Board of Trade (CBOT), refer to Alaton et al. (2002).

10. For the residual term, using the R GAM regression software, one can select any distribution from the exponential family of distributions.

This model is a time weighted temperature demand model. The difference between Model 1 and Model 2 is the term $w(DST_t)^*|Temp_t - 20.0|$, where the function $w(\cdot)$ is a piecewise continuous sinusoidal function of DST that returns values between 0 and 1. As can be readily seen from the formulation above, this function weights the temperature demand signal where 1 represents the 'full' temperature signal ($|Temp_t - 20.0|$) and 0 completely attenuates the temperature signal. It is intuitive and reasonable that the demand sensitivity to the temperature signal $|Temp - 20.0|$ depends on the human and economic activity. This daily activity cycle can be readily determined by examining the sample daily electricity demand cycle with the temperature demand component removed. In other words, $w(DST_t)$ has approximately the same shape as $s(DST_t)$ and, therefore, like $s(DST_t)$, is driven by the daily activity cycle. The cross sectional regressions performed in Section 2.2 clearly show the sensitivity to the daily exogenous demand cycle.

The third model uses non-periodic splines to model the non-linear relationship between temperature and demand:

$$\textbf{Model 3:} \quad D_t = \beta_0 + s(DST_t) + h(Temp_t) + \beta_2 Year_t + \varepsilon_t. \tag{3.3}$$

This model can be referred to as the nonlinear temperature dependent demand model. Here, in addition to the cyclic spline function of time $s(DST_t)$ (that will be present in all models), we incorporate a non-periodic (non-cyclic) spline function of temperature $h(Temp_t)$ instead of using the function from Model 1 (second term on the right hand side of Equation (3.2)).

The fourth model uses non-periodic splines to model the non-linear relationship (interaction) between time weighted temperature and demand:

$$\textbf{Model 4:} \quad D_t = \beta_0 + s(DST_t) + h(Temp_t * w(DST_t)) + \beta_2 Year_t + \varepsilon_t. \tag{3.4}$$

The fifth model applies non-periodic splines to model the non-linear relationship of long-term change in demand as a function of the *Year* using the fitted spline term $k(Year_t)$:

$$\textbf{Model 5:} \quad D_t = \beta_0 + s(DST_t) + h(Temp_t * w(DST_t)) + k(Year_t) + \varepsilon_t. \tag{3.5}$$

Models 1 through 5 use the entire data (one year) for the estimation, thus, they all have the $Year_t$ term included in the model. An implicit assumption of these models is that exogenous time dependent demand $s(DST)$ is stationary across the year. However, when examining the cross sectional regression in Section 2.2, it was shown that the term $s(DST_t)$ is not stationary across the year. Thus, the evolution of exogenous time dependent demand $s(DST)$ may be better captured by the seasonal demand model, which caters for non-stationarity of the exogenous time dependent demand by fitting regressions over shorter periods. Thus, we fit the following regressions for each calendar month:

$$\textbf{Model 6:} \quad D_t = \beta_0 + s(DST_t) + \beta_1 (w(DST_t)^* | Temp_t - 20.0|) + \varepsilon_t. \tag{3.6}$$

The seasonal demand Model 6 in Equation (3.6) is a simple two-term version of Model 2 in which the yearly regression term ($\beta_2\ Year_t$) has been removed. This seasonal model, which will be fitted for each calendar month, is expected to better capture demand fluctuations compared to the yearly demand models (Models 1 to 5).

4. EMPIRICAL ANALYSIS

In this section we perform empirical analysis using the demand and temperature data described in Section 2 and models introduced in Section 3. Each of the following sub-sections 4.1

through 4.6 will discuss the results obtained from fitting data using Models 1 through 6, respectively, comparing the fit of each model to the previously obtained results.

4.1 Model 1

Model 1 is a time dependent demand model (Equation (3.1)), which is indexed by DST or standard (astronomical) time. The results from the model fit obtained using DST and standard time are tabulated in panels A and B of Table 2, respectively.

Table 2: Model 1 using DST and standard time: Estimation results for regression in Eq. (2) (Model 1) where *Time* variable is given by the DST (panel A) and standard time (panel B).

	Estimate	Std.error	t-value	F test (p-value)	R^2_{adj}
Panel A: Model 1 with daylight saving time					
Intercept	7795.2***	4.08	1908.55		0.837
s(*DSTime*)				35513(0.000)	
|*Temp*−20|	108.146***	0.5134	211.29		
Year	–319.01***	6.00	–53.15		
Panel B: Model 1 with standard time					
Intercept	7804.9***	4.39	1776.69		0.811
s(*Time*)				29693(0.000)	
|*Temp*−20|	106.05***	0.51	209.36		
Year	–318.8***	6.46	–49.37		

Notes: ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

From both panels we observe that all terms are highly statistically significant.[11] In particular, both regressions show a long-term decline in electricity demand (the term *Year* in the regression), with daily demand falling 319 Megawatts over a one year period (–3.5%). This decline in electricity demand is in line with the media release from the Australian Energy Market Operator (AEMO) (2014).[12] The temperature regression term ($|Temp_t - 20|$) shows the expected positive relationship to electricity demand. In Section 3 we have discussed the choice of the 'comfort' temperature using regressions with different 'comfort' temperatures, and have demonstrated that $20°C$ is the optimal constant for the minimum demand temperature. We notice that using DST (panel A) rather than standard time (panel B) as the independent variable gives an improvement of fit compared to using standard time, which is reflected in the higher R^2_{adj} and larger value of the F-statistic. This result is intuitive as the daily personal and economic demand cycle depends on clock (DST) time rather than standard time.

4.2 Model 2

In this section we report the results from fitting Model 2 (Equation (3.2)), which is a time weighted temperature demand model. Here, temperature is time weighted using a constant piecewise continuous smooth function of DST, $w(DST_t)$. The results are reported in Table 3. We observe that Model 2 produces a superior fit compared to Model 1 (Table 2), which is evident

11. In this and the following tables we use the autocorrelation-robust standard errors.

12. *"The 2014 NEFR [National Electricity Forecasting Report] shows reduced residential and commercial consumption in most NEM regions due to strong growth in rooftop photovoltaic (PV) system installations and ongoing energy efficiency savings in response to high electricity prices over recent years."*

from the higher value for R^2_{adj} and more significant temperature term (higher value for the t-stat for $w(DST_t)^*|Temp-20|$.

Table 3: Model 2 using time weighted temperature: Estimation results for regression in Eq. (3) (Model 2).

	Estimate	Std.error	t-value	F test (p-value)	R^2_{adj}
Intercept	7772.4***	3.7	2100.91		0.862
s(*DSTime*)				28250(0.000)	
$w(DST_t)^*\|Temp-20\|$	216.5***	0.7884	274.57		
Year	–360.55***	5.52	–65.28		

Notes: ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

4.3 Model 3

To examine the possibility that the relationship between temperature difference and demand is non-linear, we fit Model 3 (Equation (3.3)) where the relationship between the temperature and demand is modelled by non-periodic splines. The results of this regression are reported in Table 4. We observe that the term *h*(*Temp*) is highly significant, so is the *s*(*DSTime*) term. However, R^2_{adj} from Model 3 is marginally lower compared to Model 2, which is due to the fact that we do not time weight the temperature signal in Model 3 using function *w*(·).

Table 4: Model 3 using nonlinear temperature: Estimation results for regression in Eq. (3.3) (Model 3).

	Estimate	Std.error	t-value	F test (p-value)	R^2_{adj}
Intercept	8323.4***	3.4	2435.38		0.841
s(*DSTime*)				30722 (0.000)	
h(*Temp*)				7987 (0.000)	
Year	–338.4***	6.0	–56.17		

Notes: ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

4.4 Model 4

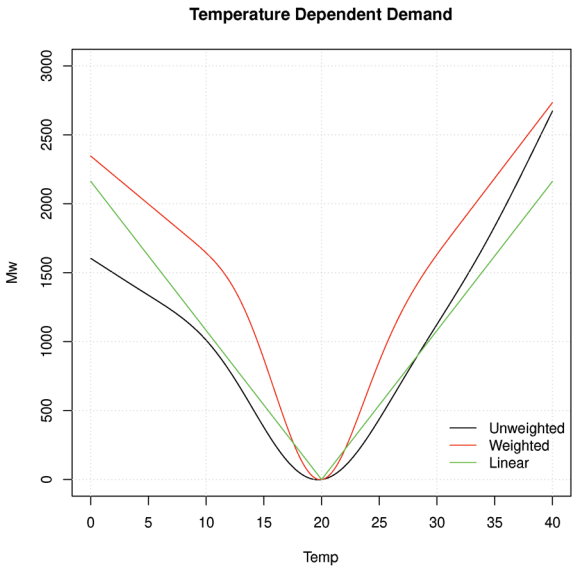
In this subsection we present the results for the time weighted temperature model (Model 4) in Equation (3.4). Table 5 reports estimation results. The results from the t-tests indicate that both variables, temperature and year are highly statistically significant, and the value of the F-statistic for the weighted term *h*(*Temp***w*(*DST*)) is larger compared to the F-statistic for the unweighted term *h*(*Temp*) from Model 3 (Table 4). We also observe the largest (compared to Models 1 through 3) value for the R^2_{adj} corresponding to 0.869, which again points towards a superior fit for the weighted temperature model.

Table 5: Model 4 using *time weighted* nonlinear temperature: Estimation results for regression in Eq. (3.4) (Model 4).

	Estimate	Std.error	t-value	F test (p-value)	R^2_{adj}
Intercept	8328.6***	3.1	2687.42		0.869
s(*DSTime*)				22115(0.000)	
h(*Temp***w*(*DST*))				11754(0.000)	
Year	–349.0***	5.5	–63.92		

Notes: ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

Figure 3: Nonlinear temperature dependent electricity demand using non-periodic splines.



Notes: Black line shows the resulted fitted demand when using Model 3 (Equation (3.3)) with the unweighted temperature spline function $h(Temp_t)$. Red line shows the resulted fitted demand when using Model 4 (Equation (3.4)) with the time weighted spline temperature function $h(Temp_t{*}w(DST_t))$. For comparison, the green line is the linear relationship of Model 1 (Equation (3.1)).

We compare Models 3 and 4 in Figure 3: Black line shows fitted demand when using Model 3 with the unweighted temperature spline function $h(Temp_t)$; red line corresponds to the non-linear temperature dependent demand function resulting from fitting Model 4 that uses a spline function $h(Temp_t{*}w(DST_t))$. For comparison, the green line corresponds to the linear relationship of Model 1. It is clear from this graph that the relationship is 'U' shaped at the minimum demand 'comfort' temperature and this explains the optimal $20°C$ minimum demand temperature, as discussed above. We notice that for both weighted (red) and unweighted (black) functions, fitted demand takes nearly identical values for large temperatures, and the curves deviate from each other for small temperatures. The unweighted model is less sensitive to low temperatures than it is to high temperatures. This is intuitive and supports the time-weighted model of temperature sensitivity. A cold morning at 04:00am generates a much lower increase in demand compared to a cold morning at the same temperature at 09:00am due to the difference in human activity between these two times.

4.5 Model 5

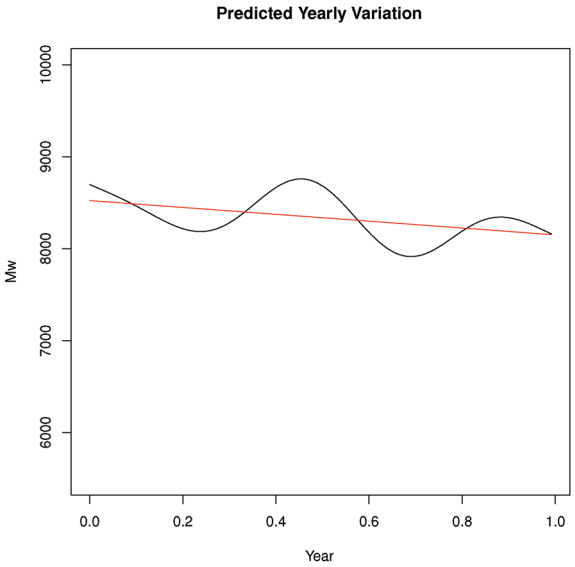
Now we assume that changes in demand as a function of the *Year* variable are non-linear. This is achieved, as suggested above, by fitting a non-periodic spline function. We fit Model 5 (Equation (3.5)), which uses the spline term $k(Year)$. The results are reported in Table 6. We observe that the addition of the non-linear term $k(Year_t)$ produces an R^2_{adj} of 0.898, which is higher than R^2_{adj} obtained for Model 4 ($R^2_{adj}=0.869$). Figure 4 shows the fitted functional form of the term $k(Year_t)$ (black line). This was fitted with DST set to noon ($DST=0.5$) and temperature set to the optimal 'comfort' temperature (no temp signal; $Temp=20°C$). The red line is a linear approximation of $k(Year_t)$; it shows a decline in demand across the year. This is consistent with the linear fitted decline in Model 4.

Table 6: Model 5 using nonlinear yearly variation in demand.

	Estimate	Std.error	t-value	F test (p-value)	R^2_{adj}
Intercept	8156.8***	1.36	6000		0.898
s(*DSTime*)				34731(0.000)	
h(*Temp***w*(*DST*))				8971(0.000)	
k(*Year*)				4190(0.000)	

Notes: Estimation results for regression in Eq. (6) (Model 5); ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

Figure 4: Fitted spline function of yearly changes in demand.



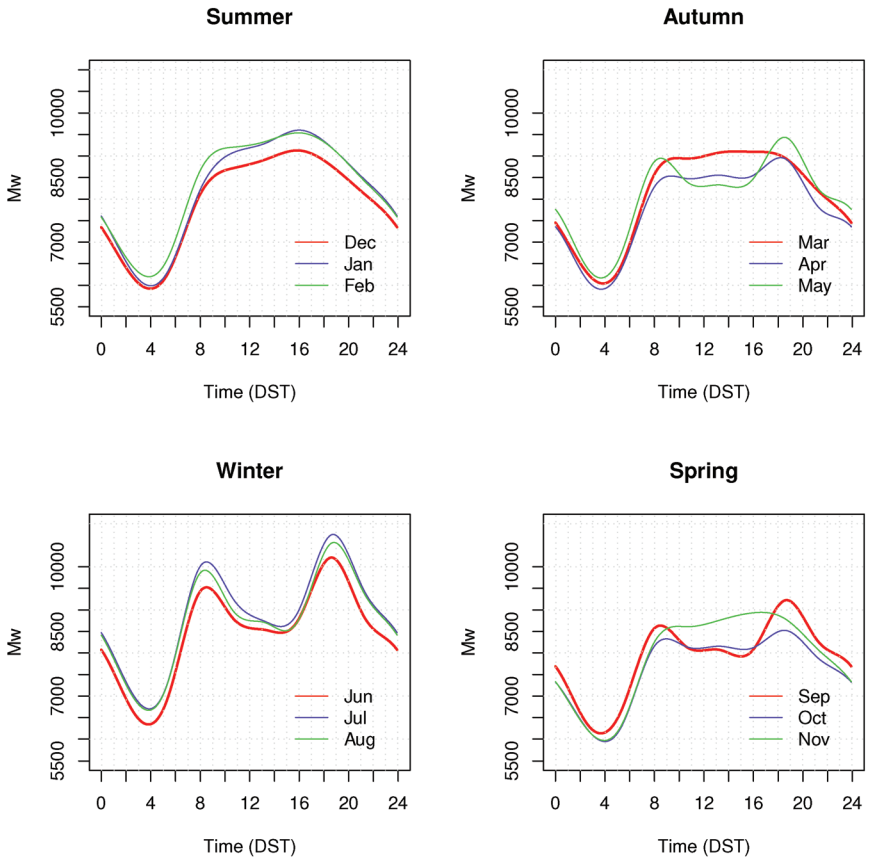
Notes: The fitted functional form of the $k(Year_t)$ term (black line) in Model 5 (Equation (6)), fitted using DST set to noon ($DST = 0.5$) and temperature set to the optimal 'comfort' temperature (no temp signal; $Temp = 20°C$). The red line is a linear approximation of $k(Year_t)$.

4.6 Model 6

We now examine the performance of the seasonal demand model (Model 6) that is fitted for each calendar month.

The daily smoothed (periodic splines) electricity demand curves, $s(DST_t)$, for different months of the year that result from fitting Model 6 are shown in Figure 5. It is important to notice that even if the DST (activity) demand cycle is homogeneous across months, the $s(DST_t)$ term will change each month (as it is evident from the figure) because of the change in the daily temperature cycle between months. From the figure we can observe the formation of two peaks from autumn (April and May, not March) to winter (June–August). These peaks disappear from spring (September and October still have two peaks, but not November) to summer (December–February) with an increase of the maximum demand. Furthermore, the peak demands are larger in winter than in summer. It is not the purpose of our paper to disentangle that part of the daily $s(DST_t)$ term that is due to intrinsic DST (activity time) and that part that is due to the daily temperature cycle. However, performing the regression each month has two implications:

Figure 5: **Daily smoothed (periodic splines) electricity demand curves for different months resulted from Model 6, for different months of the year (indexed by DST).**



- Any seasonal change in the daily DST (activity time) related demand will be captured in the $s(DST_t)$ term, resulting in a better fit to the data.
- The seasonal daily temperature variation will also be captured in the $s(DST_t)$ term. This effect is obvious when we examine the empirical $s(DST_t)$ functions in Figure 5. The winter $s(DST_t)$ functions (in June, July, August) show twin peaks of demand in the morning and evening when the temperature is significantly colder than the 'comfort' temperature ($20°C$). The daytime demand is lower as the temperature rises towards the 'comfort' temperature. Conversely, in summer (December, January, February) the peak demand reaches its maximum in the afternoon when temperatures are above the 'comfort' temperature.

Estimation results for Model 6 are summarised in Table 7. We observe that all variables are highly statistically significant, with R^2_{adj} ranging between 0.803 to 0.963. It is worth noticing that the smallest R^2_{adj} of 0.803 is observed for the month of December. If however, we exclude the last three days of the year (29, 30, 31 December), the R^2_{adj} increases to 0.906. This result can be explained referring to Figure 6 that shows the actual demand (black line) and predicted demand (red line) for December in the top panel, and associated temperatures in the bottom panel. The period between Christmas and New Year (December 29, 30, 31) is a 'defacto' holiday period with many businesses closed (even though they are officially business days). Although this period is not formally a holi-

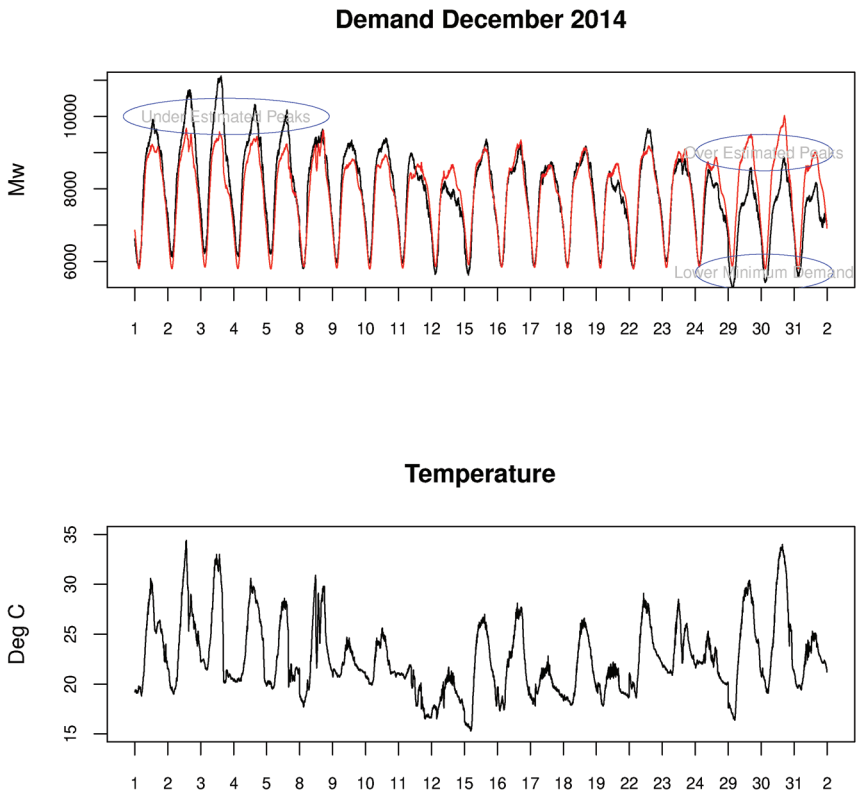
Table 7: Model 6: Estimation results for monthly seasonal regression model in Eq. (3.6) (Model 6).

Month	Coefficent	Estimate	Std.error	t-value	F test (p-value)	R^2_{adj}
February 2014	Intercept	7961.7***	7.3	1095.7		0.928
	$s(DST_t)$				3174(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	197.3***	3.5	56.9		
March 2014	Intercept	8018.1***	5.5	1446.0		0.963
	$s(DST_t)$				8855(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	46.2***	2.7	17.0		
April 2014	Intercept	7775.9***	6.7	1167.2		0.919
	$s(DST_t)$				5872(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	27.5***	3.2	8.5		
May 2014	Intercept	7792.0***	6.9	1134.5		0.952
	$s(DST_t)$				12085(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	109.0***	2.6	41.6		
June 2014	Intercept	8031.1***	12.6	637.3		0.939
	$s(DST_t)$				6574(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	129.3***	3.5	37.3		
July 2014	Intercept	8223.1***	11.7	700.7		0.941
	$s(DST_t)$				6683(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	147.9***	2.5	58.3		
August 2014	Intercept	7956.1***	13.1	607.5		0.932
	$s(DST_t)$				3964(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	184.2***	2.9	62.9		
September 2014	Intercept	7496.7***	8.2	914.7		0.870
	$s(DST_t)$				3179(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	139.6***	2.6	53.4		
October 2014	Intercept	7363.4***	4.8	1529.4		0.934
	$s(DST_t)$				6765(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	126.9***	1.7	76.4		
November 2014	Intercept	7509.0***	6.3	1192.6		0.900
	$s(DST_t)$				2848(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	233.0***	2.4	95.2		
December 2014	Intercept	7615.2***	10.8	703.7		0.803
	$s(DST_t)$				1108(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	139.3***	4.1	34.1		
December 2014	Intercept	7585.7***	8.0	948.0		0.906
December 29,30,31	$s(DST_t)$				1839(0.000)	
excluded	$w(DST_t)^*\|Temp_t-20\|$	231.0***	3.3	70.3		
January 2015	Intercept	7498.6***	8.3	898.1		0.930
	$s(DST_t)$				1862(0.000)	
	$w(DST_t)^*\|Temp_t-20\|$	279.8***	2.9	97.9		

Notes: ***, ** and * indicate significance at 0.001, 0.01 and 0.05 significance level, respectively.

day, it is characterized by a very low electricity demand. An implicit assumption in the regression is that the DST (activity) based demand cycle is homogenous across days and that demand innovations are driven by temperature. However, by including the 'defacto' holiday period, the underlying demand in the data is not homogeneous and the regression fitted using this low demand data is 'biased' downwards causing temperature driven electricity peaks earlier in the month (e.g., December 2: Max 34°*C*; December 3: Max 33°*C*) to be underestimated. Conversely, the regression overestimates temperature driven electricity peaks during the low demand 'defacto' holiday period (e.g., December 29: Max 30°*C*; Decembers 30: Max 34°*C*).

Figure 6: **Top panel: actual (black line) and predicted (red line) demand for December 2014; bottom panel: the associated temperatures.**



5. PREDICTING ELECTRICITY DEMAND

This section deals with electricity demand prediction using Model 5 (Equation (3.5)), which has been identified as the best performing model for modelling yearly data (i.e. when modelling the entire sample), as well as aggregated monthly prediction using Model 6. We will fit the following three models:

1. 'Oracle' model, which specifies the input parameter temperature $Temp_t$ as the actual temperature, i.e., it assumes that the actual temperature is known. This is the benchmark model as it contains all the information about the temperature variable.
2. 'Zero' model, which uses calculated temperatures based on long-term seasonal temperature variations and a physics based model of intra-day temperatures. The model of intra-day temperatures of the 'zero' model is based on the physics of daytime solar radiation (sinusoidual) and night-time cooling (exponential decline) using a model developed by Göttsche et al. (2001), with time calculations developed in Due and Beckman (2013). In this model the intra-day (diurnal) temperature $T_{day}(t)$ between the time of the minimum daily temperature t_{min} (before sunrise) and the time of sunset t_{sunset} is given by:

$$T_{day}(t) = T_{min} + (T_{max} - T_{min})\sin\left(\frac{\pi}{2}\left(\frac{t - t_{min}}{t_{max} - t_{min}}\right)\right), \quad t_{min} \le t \le t_{sunset}. \tag{5.1}$$

We also define

$$T_{sunset} = T_{day}(t_{sunset}). \tag{5.2}$$

The intra-day (diurnal) temperature $T_{night}(t)$ between the time of sunset t_{sunset} and the time of the minimum temperature $t_{min(nextday)}$ on the next day is given by

$$T_{night}(t) = T_{sunset} \exp\left(\log(T_{min(nextday)} / T_{sunset}) \frac{t - t_{sunset}}{t_{min(nextday)} - t_{sunset}} \right), \tag{5.3}$$
$$t_{sunset} < t < t_{min(nextday)}.$$

This is a 'zero knowledge' model and, thus, we refer to this model as the 'Zero' model.

3. 'Forecast' model uses the publicly available next day forecast maximum and minimum temperatures. Specifically, 16:20 DST ('the 6 o'clock news forecast') Australian Bureau of Meteorology forecast of next day maximum and minimum temperatures at the Sydney suburb of Paramatta[13] and then applies Göttsche et al. (2001)''s physics to interpolate intra-day temperatures for the next day.

We split the data into in-sample data (used for estimation of the model) and out-of sample (used for prediction) in the following way: We sample one out of ten data points randomly and use it as the out-of-sample dataset, the remaining (non-sampled data) is used as the in-sample dataset. This results in 90% of all data points used for the estimation and the remaining 10% used for the prediction and validation of the results. We notice that the 'Oracle' model assumes that all the information about the temperature variable is available, which appears unrealistic for any predictive model of electricity consumption, as actual temperatures are unknown at the time of prediction. On the other hand, the 'Zero' model assumes that no information about next day temperature is available, and is also unrealistic since any forecaster of next day electricity demand will have access to meteorological temperature forecasts. The 'Forecast' model is, thus, the most realistic electricity forecasting model.

The empirical results obtained from forecasting electricity demand using the 'Oracle' model (complete temperature information), the 'Zero' model (no temperature information) and the 'Forecast' model (one day ahead forecast for the maximum and minimum temperatures) are presented in Table 8. Panel A corresponds to the monthly seasonal model (Model 6, Equation (3.6)) while panel B corresponds to the best performing yearly model (Model 5, Equation (3.5)). The following errors are reported in Table 8: (i) the median error, which is the median computed across all errors, whereby the error is defined as the difference between predicted and actual demand, scaled by the actual demand and expressed in percentages; (ii) the standard deviation error, which is the standard deviation computed across all errors defined in (i); (iii) the MAPE forecasting error, defined as the average of the absolute value of the errors defined in (i).

The Australian Energy Market Operator (AEMO) publishes electricity demand forecasts every 2 hours. We use the shortest AEMO forecast period (16.5 hours), published at 12:00pm DST the previous day for 48 half hour periods (24 hours) beginning at 04:30am the following day. Each demand forecast is made at the 10%, 50% (median) and 90% deciles. We compare the median forecasts for all business days from June 1, 2014 to May 31, 2015.[14] These are tabulated below for

13. Paramatta is the closest suburb (approx 10 kilometres) to Homebush where Australian Bureau of Meteorology forecasts are available. The median difference between the forecast minimum at Paramatta and the actual minimum at Homebush was –0.2°*C* (std dev. 1.4°*C*) and the difference between the forecast maximum at Parramatta and the actual maximum at Homebush was –0.1°*C* (std dev. 1.4°*C*).

14. Unfortunately, we were unable to obtain AEMO forecasting data to exactly match our February 3, 2014 to February 2, 2015 data period.

comparison with the models developed above.[15] The AEMO forecast data can be regarded as the best publicly available commercial electricity demand forecasts. However, the AEMO forecasting *model and methodology* is not publicly available.

Table 8: Out-Of-Sample Electricity demand prediction statistics.

Panel A: Model 6 (Monthly)			
	Median Error	Std Dev. Error	MAPE
'Oracle'	0.34%	3.58%	2.72%
'Forecast'	0.89%	4.25%	3.12%
'Zero'	0.53%	4.39%	3.23%
Panel B: Model 5 (Yearly)			
	Median Error	Std Dev. Error	MAPE
'Oracle'	0.09%	3.87%	2.94%
'Forecast'	0.53%	4.15%	3.20%
'Zero'	0.26%	4.15%	3.28%
AEMO	0.07%	2.67%	2.04%

Notes: Electricity demand prediction statistics for the 'Oracle' model (complete temperature information), the 'Zero' model (no temperature information) and the 'Forecast' model (one day ahead forecast for the maximum and minimum temperatures). These are compared to Australian Energy Market Operator (AEMO) demand forecasts.

As expected, using 'Oracle' temperature data results in the best prediction statistics, leading to the smallest standard deviation and the smallest MAPE (Mean Average Predicted Error) forecasting error. The 'Forecast' (realistic) temperature data leads to the intermediate performance, and the 'Zero' (no knowledge) temperature data has the least predictive power (refer to Table 8). We note that the yearly Model 5 (Equation (3.5)) has good predictive power and is very similar in performance to Model 6 (Equation (3.6)).

The MAPE 'Forecast' yearly prediction error is 3.20%, 1.16% higher than the AEMO MAPE forecast error of 2.04%. It is reasonable to infer that in order to minimise forecast error, the AEMO model uses multiple weighted temperature forecasts, unlike our parsimonious single temperature model. We discuss the direction of further forecast model development in the Conclusion (Section 6).

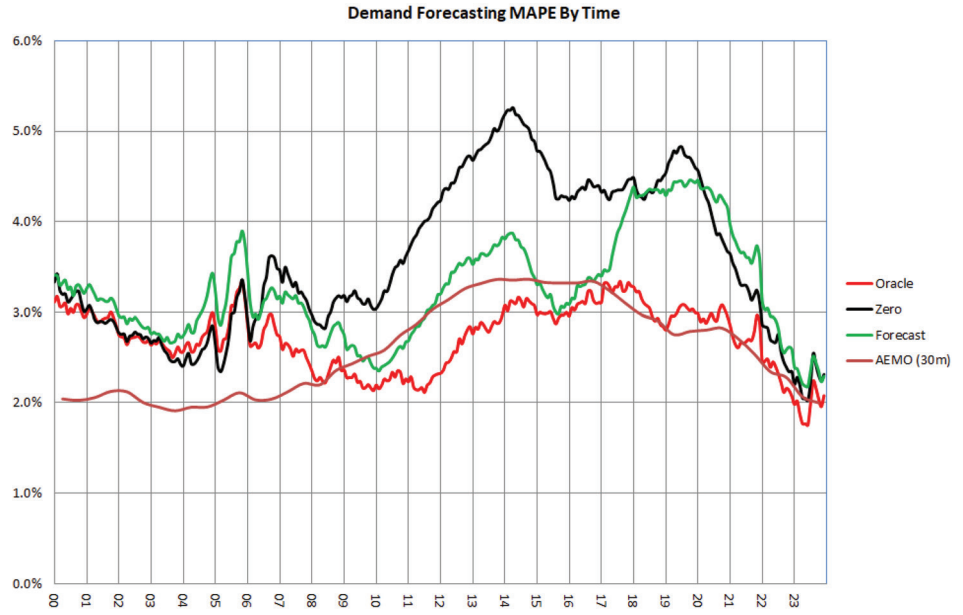
Finally, in Figure 7 we plot prediction errors (MAPE) as a function of time in 5 minute intervals, which allows us to distinguish the quality of prediction during the peak and off peak hours. We observe that all three forecasting models behave similarly during the times of low human activity (in the early morning hours and late evening hours), resulting in smaller percentage errors (typically, below 3%). The discrepancy between errors becomes more pronounced between 10:00am and 21:00pm (increased human activity), where the largest errors are observed at 14:00pm and 21:00pm. In particular, the more realistic 'Forecast' model results in the largest MAPE of above 4% during the evening hours (between 18:00pm and 21:00pm).

6. CONCLUSION

This paper introduces a parsimonious Generalised Additive Model (GAM) to relate high frequency (5-minute) electricity demand in Australia to the time of the day, time of the year and

15. These forecasts can be found on the AEMO web server at the URLs: http://www.nemweb.com.au/REPORTS/CURRENT/Short_Term_PASA_Reports/ and http://www.nemweb.com.au/REPORTS/ARCHIVE/Short_Term_PASA_Reports/ for current and archived reports, respectively.

Figure 7: Monthly prediction errors (MAPE) as a function of time (5 minute intervals) for the 'Oracle' model (complete temperature information), the 'Zero' model (no temperature information) and the 'Forecast' model (one day ahead forecast for the maximum and minimum temperatures). The AEMO 30 minute forecasting model is also shown for comparison.



outside temperature. Using yearly and seasonal demand models, we document a strong relationship between high frequency electricity demand and intra-day temperature. We establish a link between electricity demand and human activity cycle (modelled through the time of the day), and show that using Daylight Saving Time (DST) as an independent variable for the time indexed daily periodic demand consumption function provides a small but highly significant improvement of fit compared to using standard (astronomical) time.

The major novel result of this paper is that the temperature demand signal is time weighted. This relates the magnitude of the temperature demand signal to the daily activity cycle based on DST. This result is intuitive: A cold morning at 04:00am generates a much lower increase in demand compared to a cold morning with the same temperature at 09:00am due to the difference in personal and economic activity between the two times. Regression models performed using time weighted temperature demand outperform models that were not time weighted. The result is also confirmed when using cross-sectional regressions of the change in demand as a function of the temperature for all 5 minute periods across the day.

Our parsimonious GAM model is accurate with a MAPE forecasting error of 3.2% (see section 5 for details). This is an excellent result, given the limitation of temperature data to only one temperature observation at each point in time. The parsimonious GAM model, thus, provides a solid foundation for the development of more elaborate and accurate models for forecasting high frequency electricity demand.

As directions for future research, further development of the parsimonious GAM model to increase forecasting accuracy will depend on improved temperature and climate data. Temperature variations of $10°C$ or more between coastal and inland suburbs of Sydney are common. In addition, there are several substantial cities several hundred kilometres from Sydney. A more accurate model

would include demand weighted temperatures from these locations, in addition to other environmental variables such as wind and humidity. If the goal is to provide the most accurate demand forecast, then advanced but opaque forecasting techniques such as Deep Neural Networks or Gradient Boosting could be utilised.

REFERENCES

Al-Zayer, J. and A.A. Al-Ibrahim (1996). "Modelling the impact of the temperature on electricity consumption in eastern province of Saudi Arabia." *Journal of Forecasting* 15: 97–106. https://doi.org/10.1002/(SICI)1099-131X(199603)15:2<97::AID-FOR608>3.0.CO;2-L.

Alaton, P., B. Djehiche, and D. Stillberger (2002). "On modelling and pricing weather derivatives." *Applied Mathematical Finance* 9(1): 1–20. https://doi.org/10.1080/13504860210132897.

Amato, A.D., M. Ruth, P. Kirshen, and J. Horwitz (2005). "Regional energy demand response to climate change: methodology and application to the commonwealth of Massachusetts." *Climate Change* 71: 175–201. https://doi.org/10.1007/s10584-005-5931-2.

Australian Energy Market Operator (AEMO) (2014). "NEM electricity demand continues downward trend." *Media Release*. http://www.aemo.com.au/News-and-Events/News/2014-Media-Releases/NEM-Electricity-Demand-Continues-Downward-Trend.

Bašta, M. and K. Helman (2013). "Scale-specific importance of weather variables for explanation of variations of electricity consumption: The case of Prague, Czech Republic." *Energy Economics* 40: 503–514. https://doi.org/10.1016/j.eneco.2013.07.023.

Bessec, M. and J. Fouquau (2008). "The non-linear link between electricity consumption and temperature in Europe: a threshold panel approach." *Energy Economics* 30: 2705–2721. https://doi.org/10.1016/j.eneco.2008.02.003.

Cancelo, J.R., A. Espasa, and R. Grafe (2008). "Forecasting the electricity load from one day to one week ahead for the Spanish system operator." *International Journal of Forecasting* 24: 588–602. https://doi.org/10.1016/j.ijforecast.2008.07.005.

Cottet, R. and M. Smith (2003). "Bayesian modeling and forecasting of intraday electricity load." *Journal of the American Statistical Association* 98: 839–849. https://doi.org/10.1198/016214503000000774.

Darbellay, G. and M. Slam (2000). "Forecasting the short-term demand for electricity: do neural networks stand a better chance?" *International Journal of Forecasting* 16: 71–83. https://doi.org/10.1016/S0169-2070(99)00045-X.

Duffie, J.A. and W.A. Beckman (2013). *Solar engineering of thermal processes*. Fourth Edition, New York, USA: Wiley. https://doi.org/10.1002/9781118671603.

Göttsche, F.M., F.S. Olesen, and Schädlich (2001). "Influence of land surface parameters and atmosphere on meteosat brightness temperatures and generation of land surface temperature maps by temporally and spatially interpolating atmospheric correction." *Remote Sensing of Environment* 75: 39–46. https://doi.org/10.1016/S0034-4257(00)00154-1.

Hahn, H., S. Meyer-Nieberg, and S. Pickl (2009). "Electricity load forecasting methods: tools for decision making." *European Journal of Operational Research* 199: 902–907. https://doi.org/10.1016/j.ejor.2009.01.062.

Hastie, T.J. and R.J. Tibshirani (1990). *Generalized Additive Models*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Hekkenberg, M., R.M.J. Benders, H.C. Moll, and A.J.M. Schoot Uiterkamp (2009). "Indications for a changing electricity demand pattern: the temperature dependence of electricity demand in the Netherlands." *Energy Policy* 37(4): 1542–1551. https://doi.org/10.1016/j.enpol.2008.12.030.

Kim, M.S. (2013). "Modelling special-day effects for forecasting intraday electricity demand." *European Journal of Operational Research* 230(1): 170–180. https://doi.org/10.1016/j.ejor.2013.03.039.

Lam, J.C., K.K.W. Wan, and K.L. Cheung (2009). "An analysis of climatic influences on chiller plant electricity consumption." *Applied Energy* 86: 933–940. https://doi.org/10.1016/j.apenergy.2008.05.016.

Mirasgedis, S., Y. Sarafidis, E. Georgopoulou, D.P. Lalas, M. Moschovits, F. Karagiannis, and D. Papakonstantinou (2006). "Models for mid-term electricity demand forecasting incorporating weather influences." *Energy* 31: 208–227. https://doi.org/10.1016/j.energy.2005.02.016.

Moral-Carcedo, J. and J. Pérez-García (2015). "Temperature effects on firms' electricity demand: An analysis of sectorial differences in Spain." *Applied Energy* 142: 407–425. https://doi.org/10.1016/j.apenergy.2014.12.064.

Moral-Carcedo, J. and J. Vicéns-Otero (2005). "Modelling the non-linear response of Spanish electricity demand to temperature variations." *Energy Economics* 27(3): 477–494. https://doi.org/10.1016/j.eneco.2005.01.003.

Pardo, A., V. Meneu, and E. Valor (2002). "Temperature and seasonality influences on Spanish electricity load." *Energy Economics* 24: 55–70. https://doi.org/10.1016/S0140-9883(01)00082-2.

Sailor, D.J. (2001). "Relating residential and commercial sector electricity loads to climate—Evaluating state level sensitivities and vulnerabilities." *Energy* 26: 645–657. https://doi.org/10.1016/S0360-5442(01)00023-8.

Sailor, D.J. and J.R. Munoz (1997). "Sensitivity of electricity and natural gas consumption to climate in the USA—Methodology and results for eight states." *Energy* 22: 987–998. https://doi.org/10.1016/S0360-5442(97)00034-0.

Smith M. (2000). "Modeling and short-term forecasting of New South Wales electricity system load." *Journal of Business and Economic Statistics* 18: 465–478. https://doi.org/10.1080/07350015.2000.10524885.

Soares, L.J. and M.C. Medeiros (2008). "Modeling and forecasting short-term electricity demand: a comparison of methods with an application to brazilian data." *International Journal of Forecasting* 24: 630–644. https://doi.org/10.1016/j.ijforecast.2008.08.003.

Taylor, J.W. (2003). "Short-term electricity demand forecasting using double seasonal exponential smoothing." *Journal of the Operational Research Society* 54: 799–805. https://doi.org/10.1057/palgrave.jors.2601589.

Taylor, J.W. (2010). "Triple seasonal methods for short-term electricity demand forecasting." *European Journal of Operational Research* 204: 139–152. https://doi.org/10.1016/j.ejor.2009.10.003.

Taylor, J.W. (2012). "Short-term load forecasting with exponentially weighted method." *IEEE Transactions on Power Systems* 27: 458–464. https://doi.org/10.1109/TPWRS.2011.2161780.

Tung, C., T. Tseng, A. Huang, T. Liu, and M. Hu (2013). "Impact of climate change on Taiwanese power market determined using linear complementarity model." *Applied Energy* 102: 432–439. https://doi.org/10.1016/j.apenergy.2012.07.043.

Valor, E., V. Meneu, and V. Caselles (2001). "Daily air temperature and electricity load in Spain." *Journal of Applied Meteorology* 408: 1413–1421. https://doi.org/10.1175/1520-0450(2001)040<1413:DATAEL>2.0.CO;2.

Wood, S. (2006). *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC Monographs on Statistics and Applied Probability. https://doi.org/10.1201/9781420010404.

Xiao, N., J. Zarnikau, and P. Damien (2007). "Testing functional forms in energy modeling: An application of the Bayesian approach to the US electricity demand." *Energy Economics* 29: 158–166. https://doi.org/10.1016/j.eneco.2006.02.005.