

Challenge from Transpower: Determining the effect of the aggregated behaviour of solar photovoltaic power generation and battery energy storage systems on grid exit point load in order to maintain an accurate load forecast

C. Z. W. Hassell Sweatman¹ N. Wichitaksorn²
A. Jiang³ T. Farrell⁴ N. Bootland⁵ G. Miskell⁶
G. Pritchard⁷ C. Chrystall⁸ G. Robinson⁹

(Received 2 September 2019; revised 8 June 2020)

Abstract

With limited data beyond the grid exit point (GXP) or substation level, how can Transpower determine the effect of the aggregated behaviour of solar photovoltaic power generation and battery energy storage systems on GXP load in order to maintain an accurate load forecast? In this initial study it is assumed that the GXP services a

residential region. An algorithm based on non-linear programming, which minimises the financial cost to the consumer, is developed to model consumer behaviour. Input data comprises forecast energy requirements (load), solar irradiance, and pricing. Output includes both the load drawn from the grid and power returned to the grid. The algorithm presented is at the household level. The next step would be to combine the load drawn from the grid and the power returned to the grid from all the households serviced by a GXP, enabling Transpower to make load predictions. Various means of load forecasting are considered including the Holt–Winters methods which perform well for out-of-sample forecasts. Linear regression, which takes into account comparable days, solar radiation, and air temperature, yields even better performance.

Contents

1	Introduction	M3
2	Data	M5
2.1	Weather, pricing and load	M5
2.2	GXP load forecasting: Holt–Winters methods and ARIMAX	M6
2.3	GXP load forecasting: focus on solar radiation	M7
3	Load forecasting: Holt–Winters methods and ARIMAX	M8
4	Load forecasting: focus on solar radiation	M15
5	Modelling power drawn from and solar power returned to the grid	M21
5.1	The non-linear programming model: household level	M21
5.2	Results and discussion	M24
6	Summary	M36

1 Introduction

Transpower is the operator of New Zealand's national electricity transmission grid. At present, about 0.24% of the power generated in New Zealand is photovoltaic (PV).¹ Solar panels for PV power generation have traditionally been expensive, but are becoming cheaper and more commonly fitted in both new and old houses. Battery storage systems are also becoming cheaper and are likely to alter consumer load requirements in the future. With limited data beyond the grid exit point (GXP) (or substation) level, how can Transpower determine the effect of the aggregated behaviour of solar PV power generation and battery energy storage systems on GXP load in order to maintain an accurate load forecast? This was the challenge brought to MINZ by Transpower in 2018.

Currently, the GXP load comprises residential, commercial and small industrial components taking power, plus small embedded generation delivering power. Transpower anticipate that solar PV power generation and battery storage systems will have a significant uptake and are expected to cause a noticeable effect on the grid. Based on this observation, the residential, commercial and small industrial components may be partitioned into those without solar PV panels, those with solar PV panels and no battery storage, and those with both solar PV panels and battery storage. Consumers will have different behaviours, depending on pricing structure, on the availability of solar PV panels, and the capacity for storage. As the proportions of both those with solar PV generation and no battery storage, and those with both solar PV generation and battery storage rise, how will this affect the GXP load?

The injection of solar PV power into the grid during sunny days from many households simultaneously may be difficult for the operator of the grid to manage. One solution is to limit the PV power fed into the grid (Weniger, Bergner, and Quaschnig 2014). This approach requires a means of wasting

¹<https://www.eeca.govt.nz/energy-use-in-new-zealand/renewable-energy-resources/solar>

excess energy or limiting production. Another solution is to charge consumers to feed power into the grid.

It was decided at MINZ to focus on a GXP servicing a residential region. This meant industrial and commercial power use could be ignored. Reasonable assumptions for a residential home concerning roof area for solar panels and battery capacity were made. An algorithm based on non-linear programming was developed to model the behaviour of a typical consumer (Section 5). This algorithm could be incorporated into a home energy management system (HEMS) and minimises the financial cost to the consumer. Input data comprises forecast energy requirements, solar irradiance and pricing. Output includes both the load drawn from the grid and power returned to the grid. Combining the output from all the residential homes serviced by a GXP output would enable Transpower to make load predictions.

The electricity load forecasting was investigated since Transpower expected that more accurate energy requirement forecasts would provide better prediction with respect to power drawn from and returned to the grid for households with solar power generation and batteries. In the first approach, load data from an Auckland node were used in an experiment to assess various forecasting methods (Section 3). The forecasting results from a preliminary assessment revealed that the multiple seasonalities that exist in the time series create a stream of fluctuations. This makes the forecasting task more challenging and led to further investigation. Holt–Winters methods, time series linear regression, an autoregressive integrated moving average (ARIMA) model and an ARIMA model with regressors/explanatory variables (ARIMAX) were investigated.

A second approach to load forecasting was investigated, based on linear regression (Section 4). Given a stable residential location, household loads (energy requirements) may be predicted by considering historic use at the same time of year. This takes into account public holidays, school holidays, the working week, and weekends. A more dynamic approach for a changing population uses historic data from, for example, the previous week or the

previous day, but this means care must be taken to avoid comparing working days to holidays, et cetera. More accurate forecasting takes into account the forecast weather and its likely effect on energy consumption. Household loads may be satisfied by various means, including the grid, solar PV generation, and batteries. Solar radiation has considerable effects on electricity demand even *without* taking photovoltaic generation into account.

There has been a considerable interest in HEMS which incorporate domestic solar PV generation and related challenges such as load forecasting over the past few years. Reviews of this literature have been published (Beaudin and Zareipour 2015; Zhou et al. 2016). To give a few examples: Niimura et al. (2012) described a simple residential PV output profiling based on a weekly weather forecast for a HEMS in Tokyo. A simulation-based solution was presented at the Australian Mathematics in Industry Study Group 2016 to address the problem of estimating the value of various combinations of PV generation, storage and tariffs (Boland et al. 2016). Shakeri et al. (2017) described an intelligent system architecture in a HEMS for efficient demand response in smart grid, incorporating real-time information. Artificial neural networks have been applied to load forecasting (Collotta and Pau 2017). Kikusato et al. (2019) described the problem of charging an electric vehicle efficiently using solar PV generation.

2 Data

2.1 Weather, pricing and load

Transpower provided an hourly updated weather forecast obtained from the National Institute of Water and Atmospheric Research (NIWA) for the Auckland residential ROS0221 GXP region, for 2016 and 2017. This includes solar irradiance, temperature, wind speed and direction, rainfall, relative humidity, apparent temperature, and cloud cover forecasts. Grid exit point

data is provided by the New Zealand Electricity Authority under the product code ROS0221.²

Solar irradiance is measured in kW/m². Assuming 25 m² of PV panels, a yield of 15%, a production rate of 75% and transmission losses of 16%, a solar power generation time series for one household with a time step of one hour was calculated.

Transpower provided June 2016 and December 2017 consumer buying prices, per half hour trading period. It was assumed that December prices in 2016 and 2017 were similar. Initially, the selling price was fixed negative (at −100 cents/kWh) to discourage feeding power back into the grid. As there is no way of wasting excess solar generated power in our model, it is fed back into the grid, at a small cost to the producer.

Transpower provided an hourly updated forecast of the ROS0221 GXP load in megawatts (MW), for 2016 and 2017.

The data provided by Transpower for 1–2 June 2016 were combined to produce winter input, for a single typical household, including load, PV generated power, and buying price, at hourly intervals, from midnight to midnight. Similarly, the data for 1–2 December 2016 were combined to produce summer input.

Hydro-electric storage and supply and demand are the most critical factors in New Zealand power pricing.

2.2 GXP load forecasting: Holt–Winters methods and ARIMAX

As the subject of interest, the out-of-sample forecasts were made with the hourly electricity load in March 2016 as the training set and March 2017 as the test set. By doing this, the seasonality caused by the monthly fluctuations can be controlled in this experiment.

²https://www.emi.ea.govt.nz/Wholesale/Datasets/Metered_data/Grid_export

Table 1: Summary of data used in forecasting

Variable	Unit
Electricity Load	megawatt (MW)
Temperature	degrees Celsius
Wind Direction	degree
Wind Speed	knot
Rainfall	millimetre
Relative Humidity	percent
Apparent Temperature	degrees Celsius
Solar Radiation	kJ/m^2

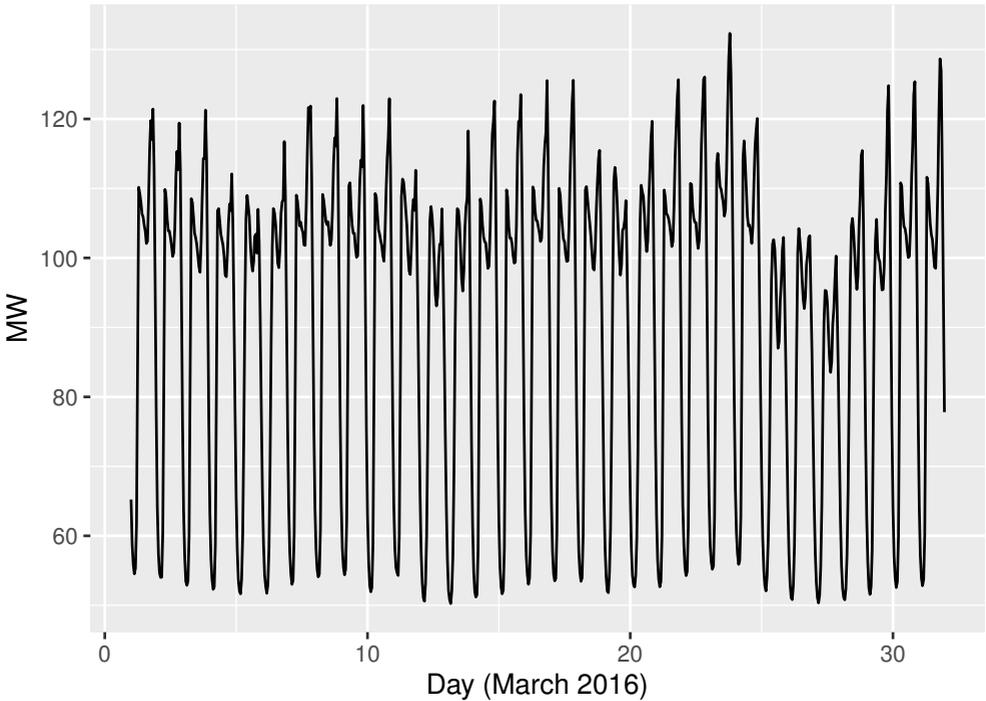
The other predictors included in the analysis are wind direction, wind speed, rainfall, relative humidity, apparent temperature, and solar radiation. [Table 1](#) summarises the data used in the analysis.

Dummy variables were included to control the effects of the weekdays and the time of the day: 6–11am, 11am–5pm, 5–9pm, and 9pm–12am, with 12–6am as the base dummy variable (due to its lowest electricity consumption period in a day).

2.3 GXP load forecasting: focus on solar radiation

The dataset used is again the one for the Roskill 22 kV GXP (ROS0221), a largely residential load. Five years of data were used, covering June 2013 through to May 2018, a period during which there was little penetration of embedded photovoltaics in the area. There was no meteorological data available specific to this part of Auckland, so the temperature data used are an average of measurements at four meteorological stations in the Auckland region (Airport, Mangere, Henderson North, and Albany). The radiation data are averaged across three stations (Airport, Mangere, and Albany).

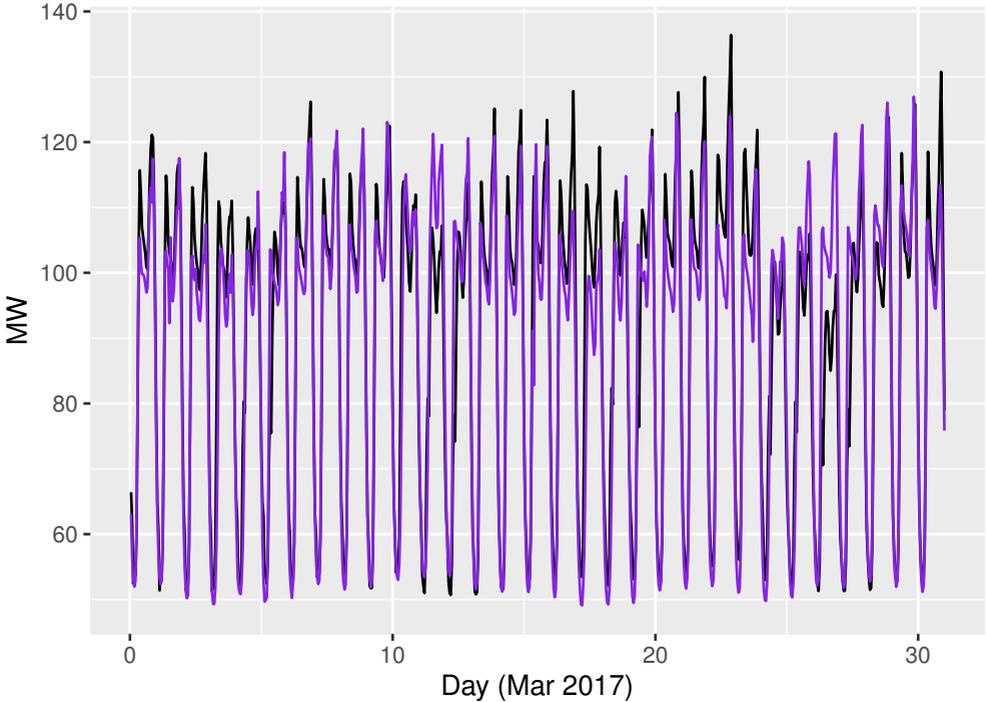
Figure 1: Hourly electricity load: an Auckland node, March 2016



3 Load forecasting: Holt–Winters methods and ARIMAX

Figure 1 plots time series for the hourly electricity load in March 2016. The plot exhibits two levels of seasonality. One completes in a day and the other completes in a week. In a day, the highest consumption usually occurs at 8 pm while the lowest consumption occurs around 3 am or 4 am. Mondays to Thursdays usually have the higher consumption levels while people consume less during the weekends and public holidays. Friday is the only workday that has lower consumption than the others, particularly in the evening.

Figure 2: Forecasts from the Holt–Winters additive method: test data plotted in black, forecasts plotted in purple

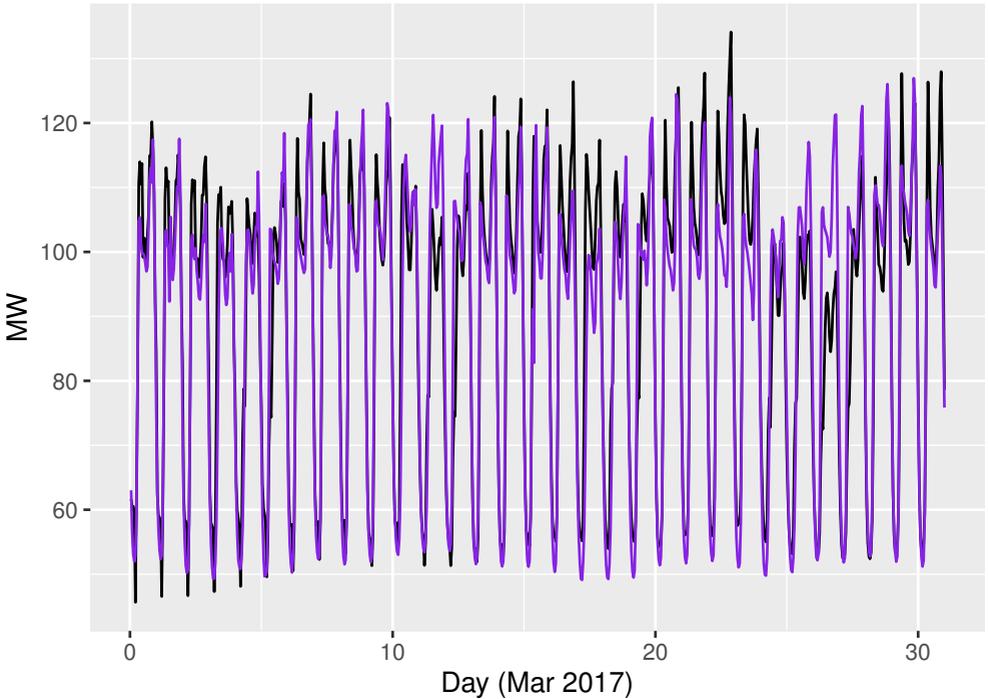


Figures 2 to 4 show the results from various Holt–Winters methods including additive, multiplicative, and the multiplicative with exponential trend. The forecasts from these three methods seem to capture the weekly and daily seasonalities where the multiplicative with exponential trend appears to reach the peaks and troughs better than the other two methods.

The forecasts were evaluated using the root mean square error (RMSE) and the mean absolute percentage error (MAPE). The three methods return similar results in terms of both RMSE and MAPE (Table 2).

These metrics are internationally recognised standards for forecasting models.

Figure 3: Forecasts from the Holt–Winters multiplicative method: test data plotted in black, forecasts plotted in purple



They treat every hour or half-hour trading period as equally important. Pricing in New Zealand depends on many factors. Peak periods within a day may have higher prices, but this is not always the case. It is important for Transpower as the system operator to forecast the load for any time of day, depending on system conditions at the time.

A linear regression model was fitted to the data. This approach allowed a set of explanatory variables to be included in the analysis. Forecasts were made. [Figure 5](#) shows the forecasts from the time series regression model where the extremes and seasonalities are not well captured. This might be caused by the inability to analyse the complex characters of the load by the regression

Figure 4: Forecasts from the Holt–Winters multiplicative with exponential trend method: test data plotted in black, forecasts plotted in purple

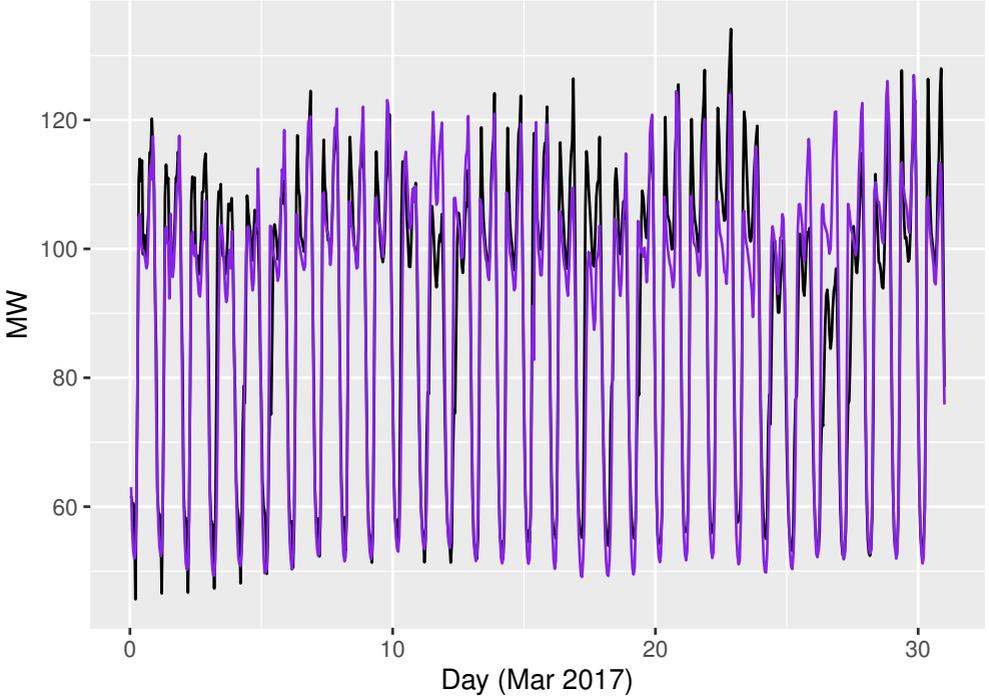
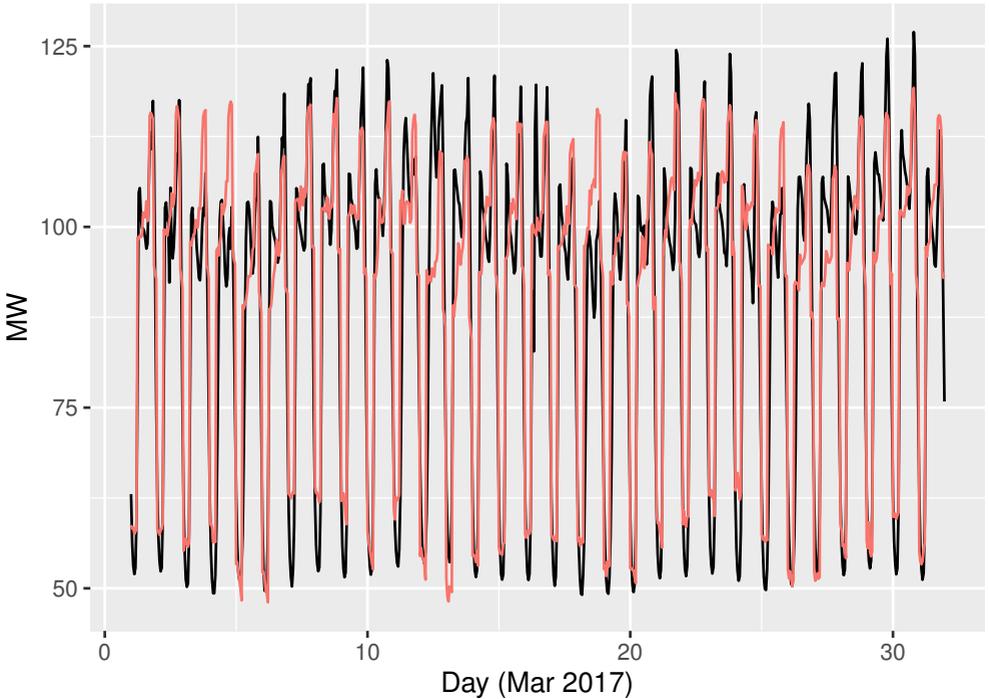


Table 2: Forecast error measures from various models and methods

Method/Model	RMSE	MAPE
Holt–Winters Additive	8.0	6.4
Holt–Winters Multiplicative	8.2	6.7
Holt–Winters Multiplicative with exponential trend	7.7	6.7
Time-series Linear Regression	9.7	8.9
ARIMAX	8.2	6.6

Figure 5: Forecasts from the time series linear regression model: test data plotted in black, forecasts plotted in red

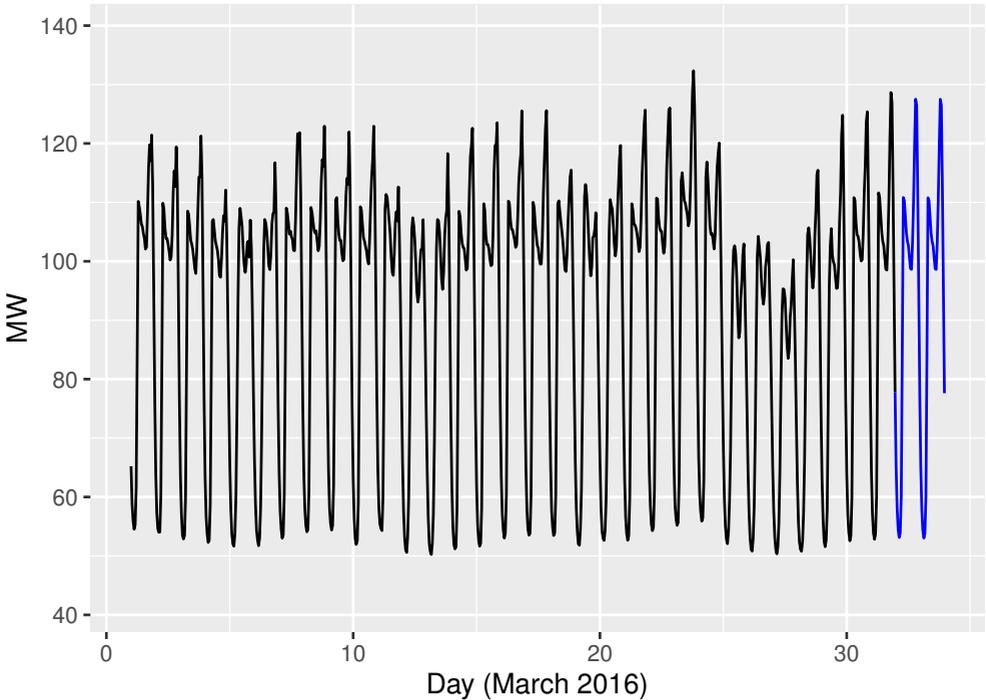


model. The RMSE and MAPE in [Table 2](#) confirm the inferior performance of the forecasts from the time series regression model.

The load was analysed using an ARIMA model using only the training set to preliminarily assess the performance of the model ([Figure 6](#)). Similar to the Holt–Winters method and due to its specification, the ARIMA model seems to capture the seasonalities well with promising error measures (not shown here).

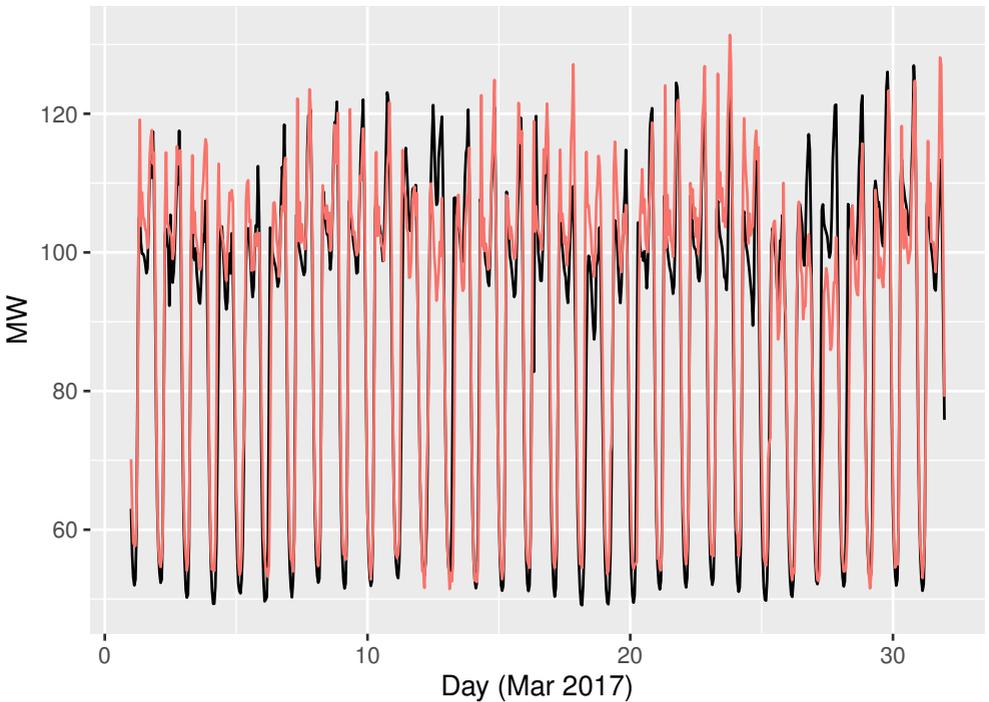
Analysis and forecasting of the load data continued using ARIMA with the set of explanatory variables or regressors through the ARIMAX model. Using the lag orders from the ARIMA but with a set of regressors, more reasonable

Figure 6: Forecasts from the $ARIMA(5, 0, 1)(0, 1, 1)_{24}$ model; autoregressor order 5, moving average order 1, seasonal differencing order 1, seasonal moving average order 1, seasonal lag 24: test data plotted in black, forecasts plotted in blue



results were obtained from the ARIMAX than those of the time series regression model (Figure 7). Although the forecast errors from the ARIMAX model are slightly worse than those of the Holt–Winters methods (Table 2), ARIMAX results seem to provide more room for future research.

Figure 7: Forecasts from the $ARIMAX(3, 0, 2)(0, 0, 2)_{24}$ model; this refers to the ARIMA model with regressors and autoregressive order 3, moving average order 2, seasonal moving average order 2, seasonal lag 24: test data plotted in black, forecasts plotted in red



4 Load forecasting: focus on solar radiation

In this section, a one day-ahead load forecasting model is constructed in which solar radiation has a starring role. The intent here is not to model embedded solar photovoltaic generation, but rather to establish a baseline load model to which other forecasts of solar generation could be added.

The model has three main predictors:

- previously-measured load at the same time on a recent comparable day,
- air temperature (daily maximum or minimum only, this choice is discussed later in [Section 4](#)), and
- solar radiation (daily total only).

We assumed that the temperature and radiation values are known on a one day-ahead basis, from sufficiently accurate weather forecasts. Previous-load and temperature predictors are known to be adequate for a basic load forecasting model. Adding solar radiation improves the model's performance a little further ([Figure 9](#)). But the question of real interest here is to characterise the way in which solar radiation affects the model's predictions. That is, what is the average difference in load between a cloudy day and a sunny day, given that the two days have the same air temperature and the same recent history (as represented by the previous day to which both are being compared)?

'Comparable' days are defined in the following way. First, all days are classified into work days and holidays (Saturdays, Sundays, and other public holidays). Then, a time t on a date d_1 is comparable to the same time t on another date d_2 if, and only if,

- d_1 and d_2 have the same work-day/holiday classification;
- if t is before noon, then $d_1 - 1$ and $d_2 - 1$ have the same work-day/holiday classification; and

- if t is after noon, then $d_1 + 1$ and $d_2 + 1$ have the same work-day/holiday classification.

This means, for example, that for times on a typical Monday morning, the most recent comparable day is the previous Monday (allowing for a beginning-of-the-week effect), while on the same Monday, for times in the afternoon, the most recent comparable day is the previous Thursday (allowing for an end-of-week effect that gives Friday evenings a different behaviour).

The model is a linear regression. In this context, the recent comparable day's load is included as a single regressor: its contribution to the forecast is simply itself, multiplied by a regression coefficient (which is presumably close to one). But when it comes to the temperature and radiation effects, some thought should be given to feature engineering to extract the best predictive value. A forecast should consider not only the temperature for the forecast day, but also the temperature for the previous comparable day, as this affects the load on that day. A reasonable way to incorporate this is to consider the *difference* in temperature between the two days. Similarly for solar radiation: the difference in radiation total between the forecast and comparable days is likely to be a better predictor than the forecast day's radiation alone.

However, even this is not sufficient. The effect of temperature on load is non-linear, being greatest (most positive) for the lowest and highest temperatures, and least (most negative) for clement intermediate conditions. We therefore allow for a temperature effect of the form

$$(T_f - T_c)g(T_f), \quad (1)$$

where T_f and T_c are the temperatures on the forecast and comparable days, and g is a fitted non-linear function. To model a temperature effect that is positive for the lowest forecast temperatures (when $T_f - T_c < 0$) and highest forecast temperatures (when $T_f - T_c > 0$), g was chosen so that $g(T) < 0$ for the lowest values of T , and $g(T) > 0$ for the highest values. In order to remain fairly flexible as to the form of g , a cubic spline was chosen, and the

function expressed as the following linear combination

$$g(T) = \sum_{i=1}^n \gamma_i b_i(T)$$

where the b_i are a set of spline basis functions. The temperature-related regressors are thus of the form $(T_f - T_c)b_i(T_f)$ for $i = 1, \dots, n$.

For solar radiation, a similar approach is taken. However, the difference in radiation $R_f - R_c$ between the forecast and comparable days should not be modulated by a function of R_f (which can take most of its possible values at any time of year, even though the effects in summer and winter are expected to be rather different). Rather, we posited a seasonal modulation through an effect of the form

$$(R_f - R_c)h(t), \quad (2)$$

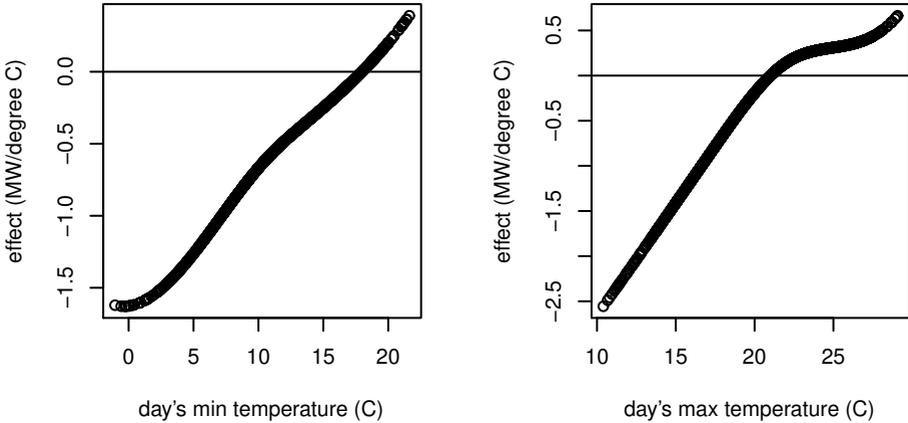
where t is time, and h a fitted non-linear function. The function h should be periodic with a period of one year, and is modelled by a trigonometric polynomial:

$$h(t) = \eta_0 + \sum_{j=1}^m [\eta_j^s \sin(j\omega t) + \eta_j^c \cos(j\omega t)],$$

where $\omega = 2\pi \text{ year}^{-1}$, and η_0 , η_j^s , and η_j^c are fitted coefficients. The radiation-related regressors are thus of the form $(R_f - R_c) \sin(j\omega t)$ and $(R_f - R_c) \cos(j\omega t)$ for $j = 0, \dots, m$.

A separate model was fitted for each of the 48 half-hours of the day. After some experimentation, the authors decided that there was little advantage to be gained by including two sets of temperature-related regressors, using both the minimum and maximum temperature of the day. A fit almost as good can be obtained by including only one set of temperature-related regressors, if they are based on the day's minimum temperature for predictions of half-hours before 10 am, and on the day's maximum temperature thereafter.

Figure 8: The temperature-modulation function g in two trading periods, TP 18 (left) and TP 31 (right)



The dataset used is again the one for the Roskill 22 kV GXP, a largely residential load ([Section 2.3](#)).

To verify the forecasting performance of the model, a cross-validation was performed in which the data were partitioned into five time-contiguous subsets of one year each, and an out-of-sample forecast made for each subset by a model fitted to the data contained in the other four subsets. The results are quite satisfactory for a model of this complexity: the RMSE is 3.18 MW and the MAPE is 4.1%.

The roles played by temperature and radiation in the model are discovered by inspecting the fitted functions g and h described in (1) and (2). [Figure 8](#) shows the function g , that is, the coefficient of $(T_f - T_c)$, for two typical half-hours: TP 18 (8:30–9:00 am) and TP 31 (3:00–3:30 pm). As might have been expected, this coefficient takes mostly negative values (that is, reducing the temperature on the forecast day will increase load), with the strength of

the effect increasing at low temperature. At high temperature, the sign of the coefficient changes (increasing the temperature on the forecast day will increase load), although this is a much weaker effect. This is presumably reflective of electric heating and (to a lesser extent) cooling loads, which are both common in Auckland.

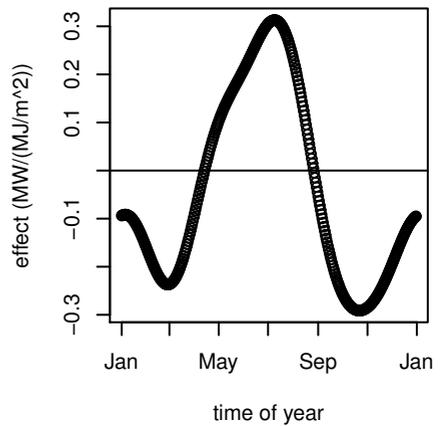
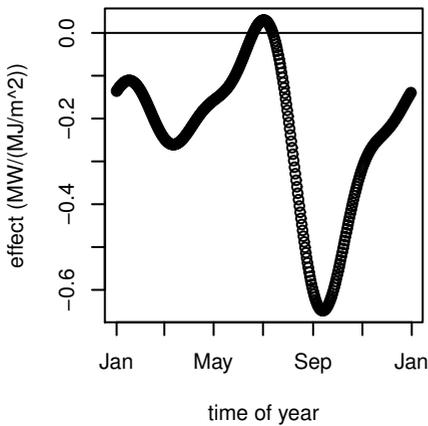
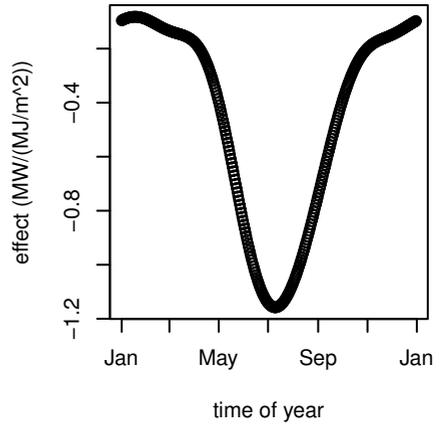
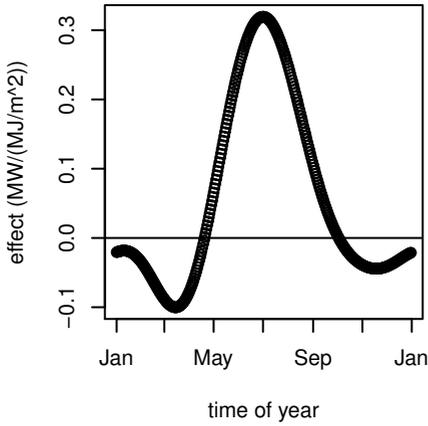
Of more interest here are the radiation effects. [Figure 9](#) shows the function h (i.e., the coefficient of $R_f - R_c$) for four half-hours. Considering first TP 16 (7:30–8:00 am), the strongest apparent effect is at first surprising: sunshine in winter *increases* load. This becomes explicable when it is considered that at this time of day, the sun has barely risen. But the prospect of a sunny day is associated with clear-sky conditions, and this is apparently enough to drive the morning-peak load upwards—even though the overnight-low air temperature is separately accounted for in this model.

Turning to TP 32 (3:30–4:00 pm), a more expected result is encountered: late-afternoon sunshine reduces load, though the effect is largely confined to winter. It is a large effect: the predicted difference between a sunny and a cloudy day can be as much as 25 MW (at a GXP where the overall average load is 52 MW), not including the effect of sunshine in raising air temperature, which is separately accounted for in this model.

In the evening peak (TP 37, 6:00–6:30 pm), there is other behaviour to observe. In general, sunshine reduces load, with the strongest effect occurring in spring (when conditions may still be cold, but the sun has not yet set at this time of day). There is almost no effect in mid-winter, when this time period falls after sunset.

Finally, the later evening (TP 40, 7:30–8:00 pm) shows how the lingering after-effects of a sunny day may either increase or decrease load, depending on the time of year.

Figure 9: The radiation-modulation function h in four trading periods, TP 16 (left upper), TP 32 (right upper), TP 37 (left lower) and TP 40 (right lower)



5 Modelling power drawn from and solar power returned to the grid

5.1 The non-linear programming model: household level

A typical household load in 2016 was forecast by dividing the total forecast load at the residential GXP (provided by Transpower) by the number of residences serviced. This assumes there was negligible solar PV power generation in 2016. How would the amount of power drawn from the grid change for a household with solar PV power generation and battery storage?

Let $P(t)$, $L(t)$, $G(t)$, $A_1(t)$, $A_2(t)$, $A_3(t)$ and $B(t)$ denote elements of time series in PV generated power, load, power drawn from the grid, PV generated power supplied to the load, PV generated power supplied to the battery, PV generated power supplied to the grid and battery discharge, respectively, as shown in [Figure 10](#). All quantities are powers, measured in kW. The time step is $1/n$ hour, where $n = 1$ corresponds to no interpolation and integer $n > 1$ corresponds to linear interpolation between data points. Let $SOC(t)$ denote battery state of charge, a fraction between zero and one. Let $C(t)$ denote the battery capacity and C_{\max} denote the maximum battery capacity (in kWh). Then, at any time step, $C_{\max} SOC(t) = C(t)$.

Let d be the number of days considered. For each $t \in \{1, 2, \dots, 24nd + 1\}$, the following financial cost function (grid expense minus PV income) is minimised

$$F(t) = (G(t) \cdot G_{\text{cost}}(t) - A_3(t) \cdot A_{3,\text{cost}}(t)) \Delta t,$$

where $G_{\text{cost}}(t)$ is the buying price (the cost to the consumer of power supplied by the grid in cents/kWh), $A_{3,\text{cost}}(t)$ is the selling price and $\Delta t = 1/n$ hours, subject to the following constraints:

$$A_1(t) + A_2(t) + A_3(t) = P(t),$$

$$\begin{aligned}
A_1(t) + B(t) + G(t) &= L(t), \\
0 &\leq A_1(t) \leq L(t), \\
0 &\leq G(t) \leq L(t) - P(t) \quad \text{if } L(t) \geq P(t), \\
0 &\leq A_3(t) \leq P(t) - L(t) \quad \text{if } P(t) \geq L(t), \\
0 &\leq B(t) \leq nC(t-1), \\
0 &\leq A_2(t) \leq n[C_{\max} - C(t-1)], \\
\text{where } C(t) &= C_{\max} \left(\text{SOC}_0 + \frac{1}{C_{\max}} \sum_{s=1}^{t-1} (A_2(s) - B(s)) \Delta t \right),
\end{aligned}$$

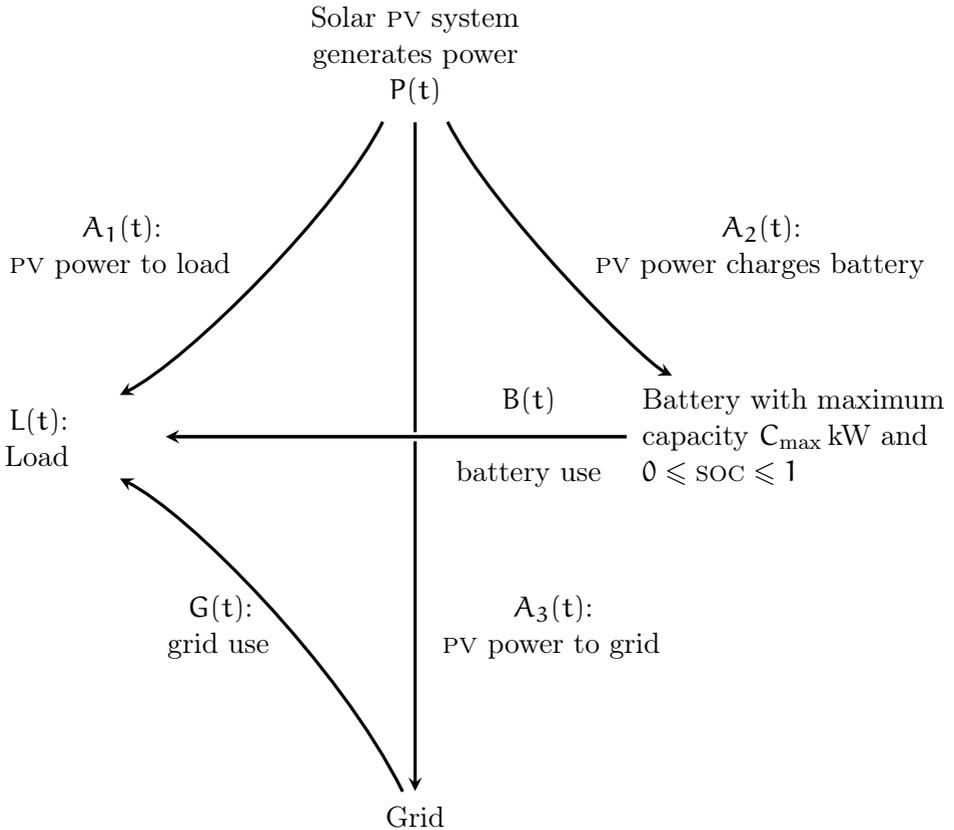
and battery charge and discharge cannot happen simultaneously, modelled by $0 = A_2(t)B(t)$.

The constraints on $B(t)$ and $A_2(t)$ mean that the battery cannot discharge or charge in less than one time step. This ensures $0 \leq \text{SOC}(t) \leq 1$ at every time step. These constraints also model batteries which discharge slowly when nearly empty, and charge slowly when nearly full. Since $\Delta t = 1/n$ is measured in hours, n has units per hour.

For one household with solar PV panels and one battery with a capacity of C_{\max} kWh; energy consumption, energy production, and the battery state of charge are modelled as time series. Power for the load could be drawn from the grid, from PV panels, or from the battery (if charged). The battery may be charged by PV generation or discharged to the load. PV power generated may feed into the load, the battery, or the grid. See [Figure 10](#), which is based on a similar figure published by Weniger, Bergner, and Quaschnig (2014).

The model is implemented in MATLAB, using the non-linear optimisation function `fmincon`, minimising a financial cost function (grid expense minus PV income). The input data comprises the hourly forecast load, estimated PV power generation and buying prices. The simulation results plotted in [Section 5.2](#) were obtained using data corresponding to the first 48 hours in June 2016 and the first 48 hours in December 2016, taking into account winter and summer patterns in Auckland, New Zealand.

Figure 10: Path of energy in a solar PV system with battery storage but no means of dumping excess PV generated power (Weniger, Bergner, and Quaschnig 2014). All time dependent quantities $P(t)$, $L(t)$, $G(t)$, $A_1(t)$, $A_2(t)$, $A_3(t)$ and $B(t)$ are powers. The battery state of charge (SOC) is a fraction of the maximum capacity.



5.2 Results and discussion

Time series are plotted for variables, namely $P(t)$, $L(t)$, $G(t)$, $A_1(t)$, $A_2(t)$, $A_3(t)$, $B(t)$ and $SOC(t)$. Winter and summer patterns are considered. In the simulations a time step of 20 minutes was used, that is $n = 3$, over the first 48 hours in both June and December 2016, from midnight to midnight. The selling price $A_{3,\text{cost}}(t)$ was set to -100 cents/kWh to discourage selling PV generated power to the grid, $C_{\text{max}} = 13.5$ kWh, and the area of PV panels per household roof was set to 25 m^2 .

Winter and summer weather patterns differ markedly. Hence, winter and summer power use is predicted to differ markedly too. Winter and summer examples are compared below.

- The predicted peak winter solar PV power generated is lower, about half that predicted for summer (Figure 11(a) and Figure 15(a)).
- The model predicts that, in winter, less solar energy would be generated than in summer. Compare the area under the curves in Figure 11(a) and Figure 15(a).
- The model predicts that the peak winter load is higher than the peak summer load (Figure 11(b) and Figure 15(b)).
- Higher energy requirements are expected in winter. Compare the area under the curves in Figure 11(b) and Figure 15(b).
- Grid use is predicted to be high in winter, accounting for most of the load (Figure 12(a)).
- Grid use is predicted to be negligible on sunny days in summer (Figure 16(a)).
- The model predicts that, in winter, almost all solar PV power is used by the load. Very little is left to charge the battery and the amount fed to the grid would be negligible (Figure 12(b), Figure 13(a) and Figure 13(b)).

Figure 11: Winter time series over 48 h, 1–2 June 2016, midnight to midnight, ROS0221 GXP, per household

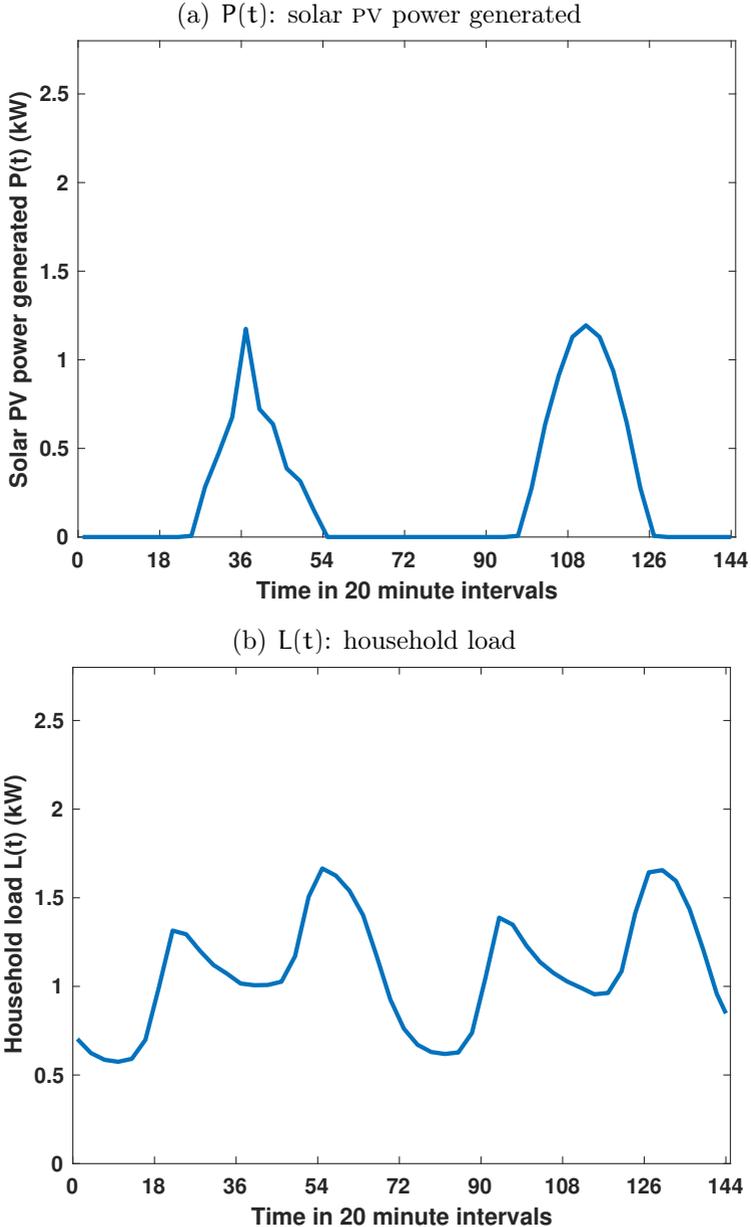


Figure 12: Winter time series over 48 h, 1–2 June 2016, midnight to midnight, ROS0221 GXP, per household

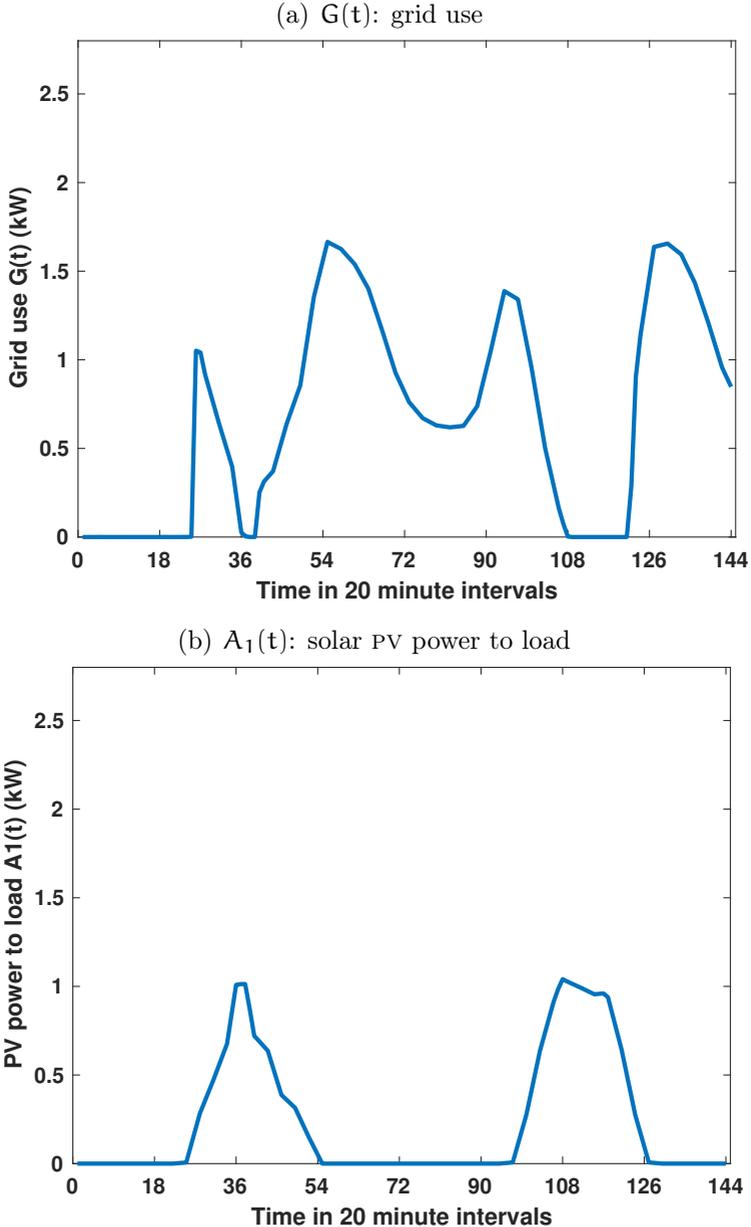


Figure 13: Winter time series over 48 h, 1–2 June 2016, midnight to midnight, ROS0221 GXP, per household

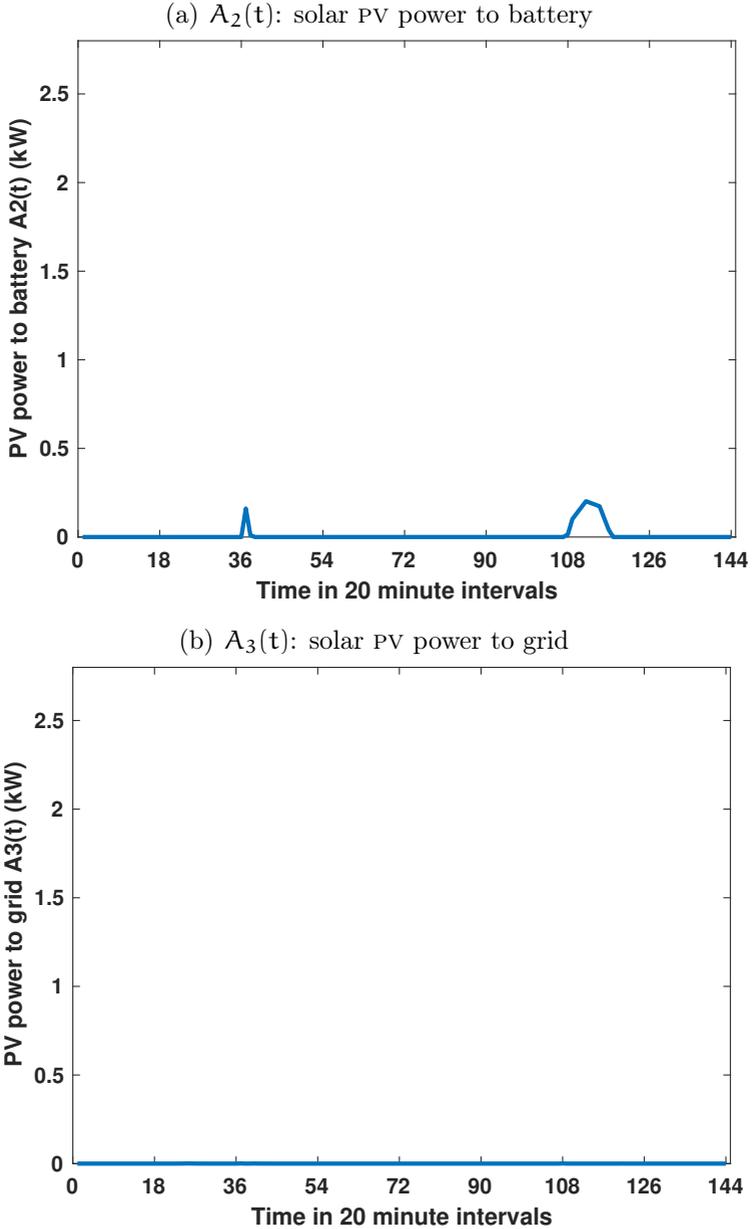


Figure 14: Winter time series over 48 h, 1–2 June 2016, midnight to midnight, ROS0221 GXP, per household

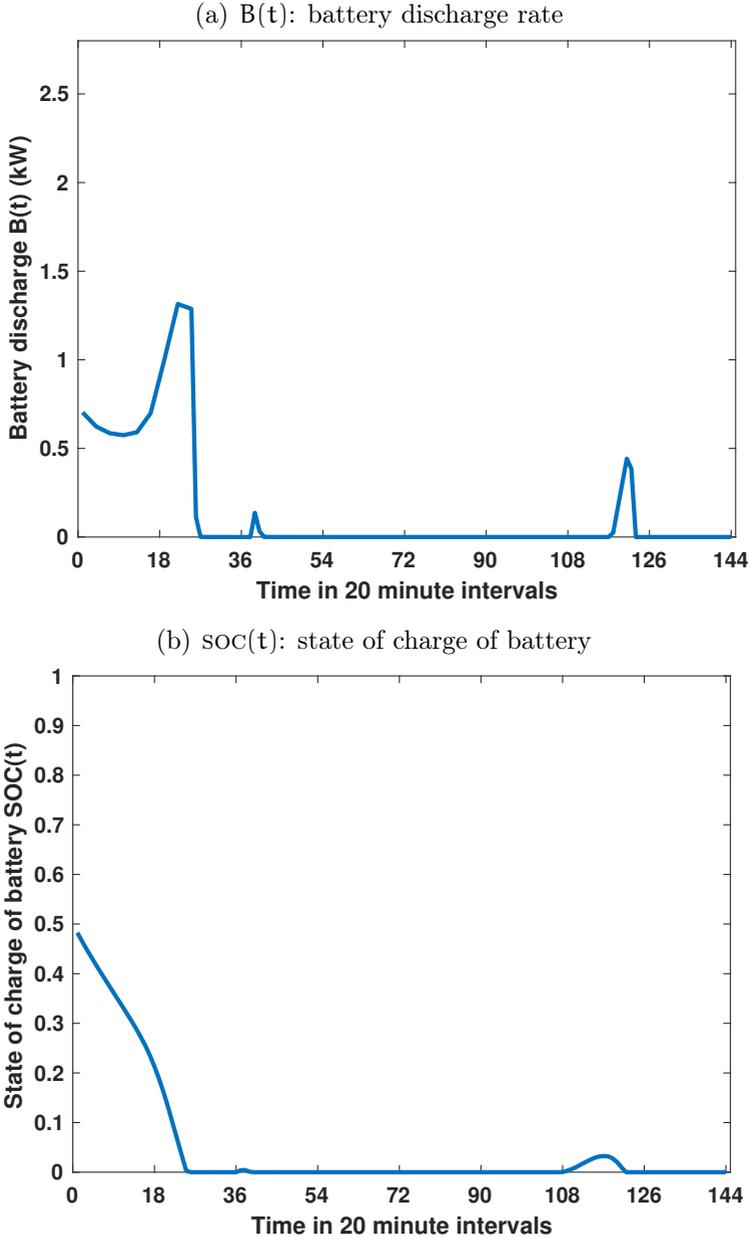


Figure 15: Summer time series over 48 h, 1–2 December 2016, midnight to midnight, ROS0221 GXP, per household

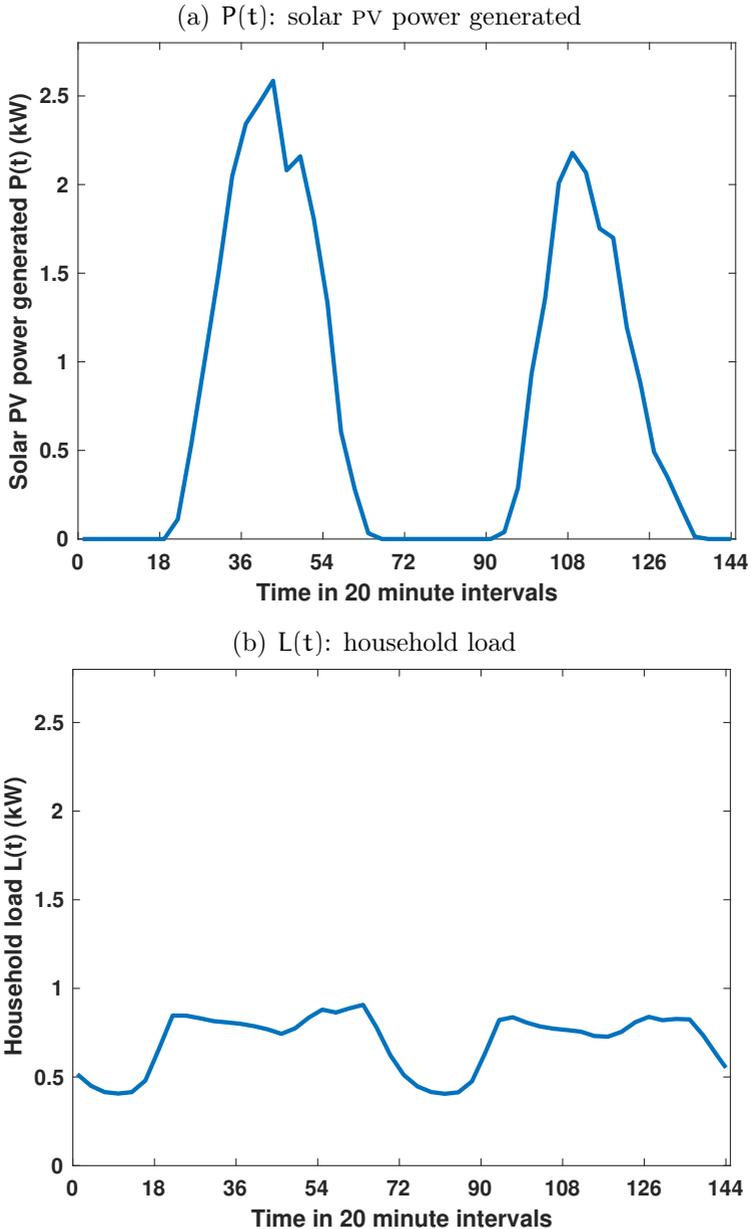


Figure 16: Summer time series over 48 h, 1–2 December 2016, midnight to midnight, ROS0221 GXP, per household

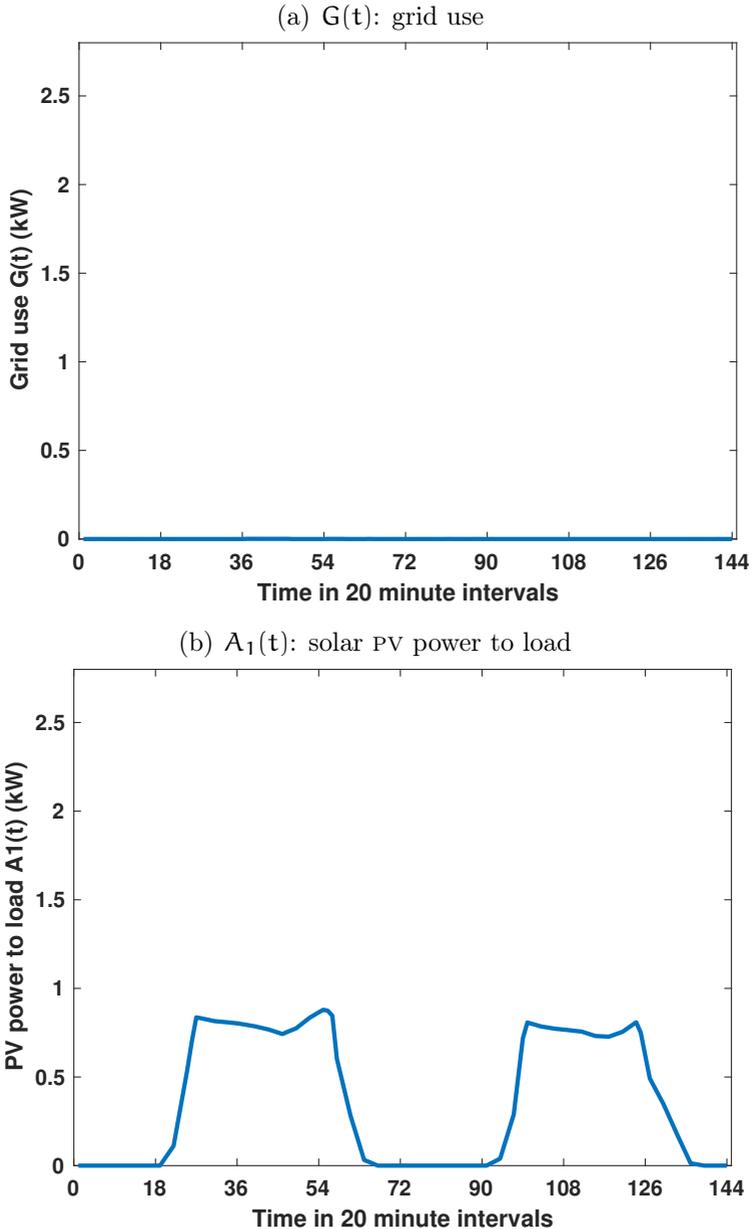


Figure 17: Summer time series over 48 h, 1–2 December 2016, midnight to midnight, ROS0221 GXP, per household

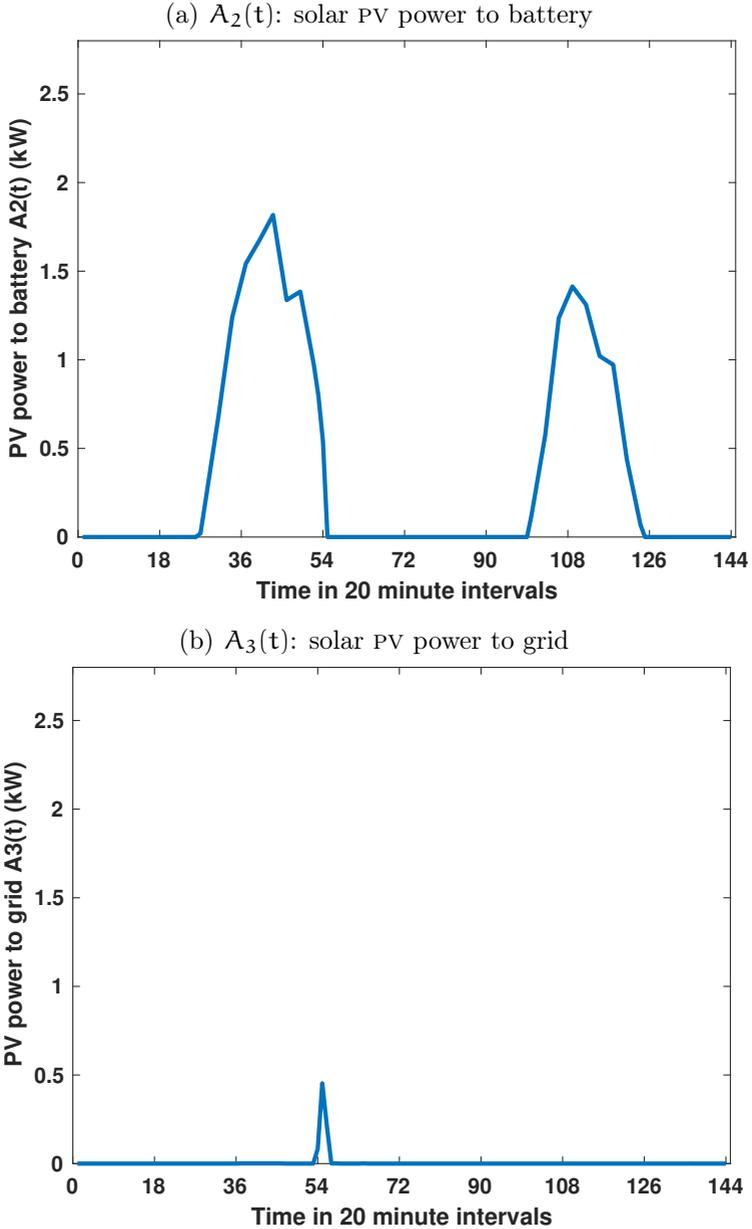
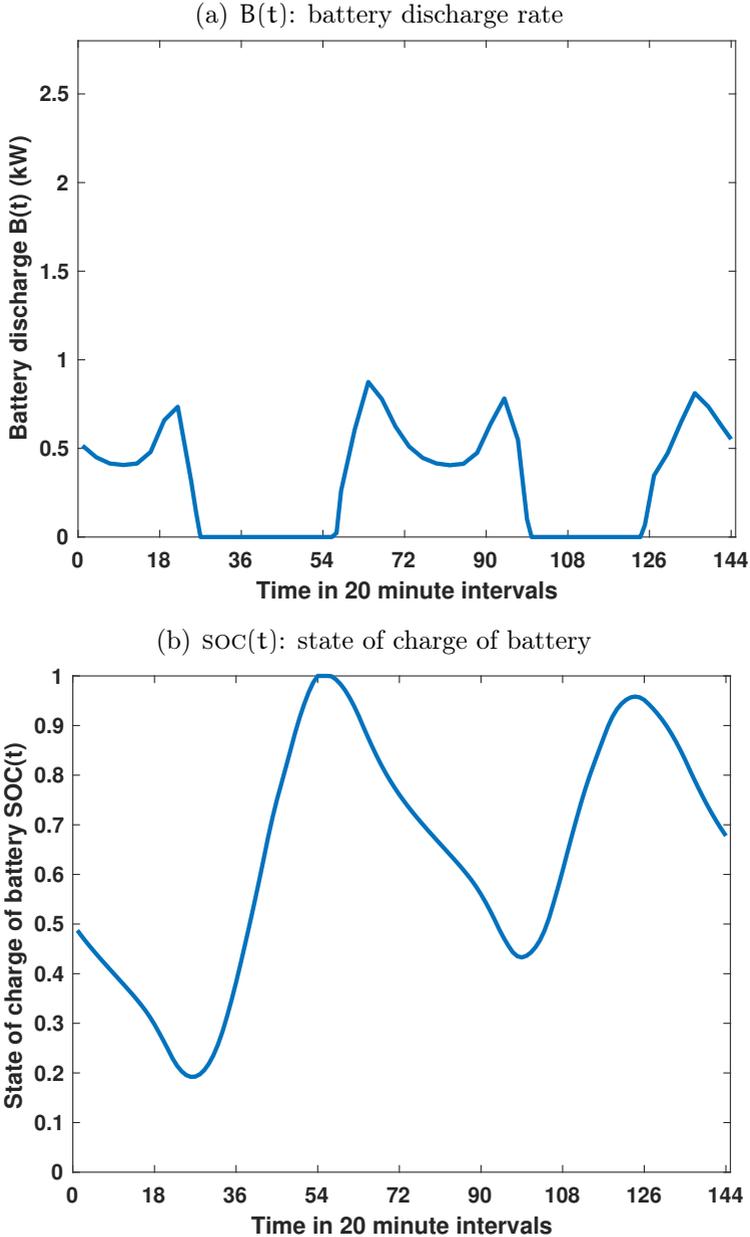


Figure 18: Summer time series over 48 h, 1–2 December 2016, midnight to midnight, ROS0221 GXP, per household



- The model predicts that, in summer, solar PV would charge the battery in the middle of the day, attaining maximum capacity (more or less). The slightly higher PV generation predicted on the first day, compared to the second day (Figure 15(a)), corresponds to slightly higher solar PV supplied to the battery (Figure 17(a)), full battery charge by evening (Figure 18(b)) and a slight excess of solar PV power fed into the grid (Figure 17(b)). This would happen at a cost to the consumer, as there is no way to waste power in the model. In the simulations, the battery is at half maximum capacity at the start (midnight) of the 48 hours (Figure 18(b)).
- The battery SOC is predicted to be very low in winter and hence so is the battery discharge rate (Figure 14(a) and Figure 14(b)).
- The battery SOC is predicted to reach capacity about 6 pm in summer. By about 9 pm, the battery is predicted to have taken over from solar PV power, satisfying the load. Battery capacity is predicted to fall to a minimum of 20%–40% maximum capacity by about 8 am (Figure 18(a) and Figure 18(b)).
- Scaling the input $P(t)$ on a day in winter by a factor of 2.23 results in recovering the initial 50% SOC after 48 hours. Hence, on a very sunny day in winter (during which the heating load would likely reduce), battery use could be expected (Figure 19). In this example, the battery is emptied by 10 pm.
- Figure 20 plots contributions to the load for winter and summer. The model predicts that on a day in winter, most of the load is satisfied by the grid; whereas on a day in summer, solar PV generated power during the day and stored solar PV generated power from the battery during the night satisfy the load.

Figure 19: Time series for a very sunny day in winter over 48 h, 1–2 June 2016, midnight to midnight, ROS0221 GXP. The original winter PV data was multiplied by 2.23.

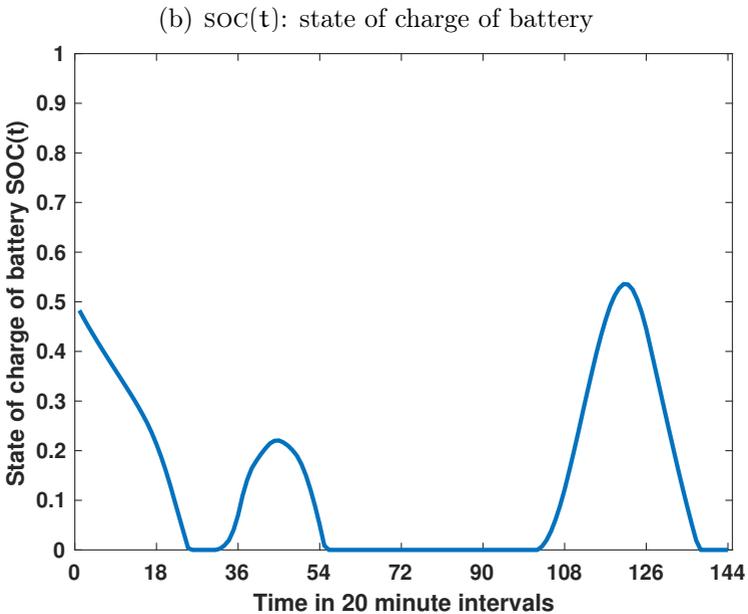
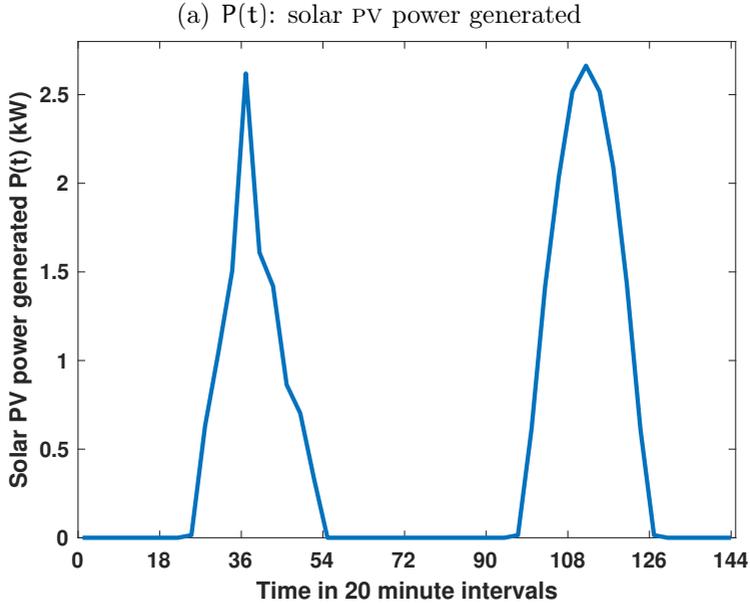
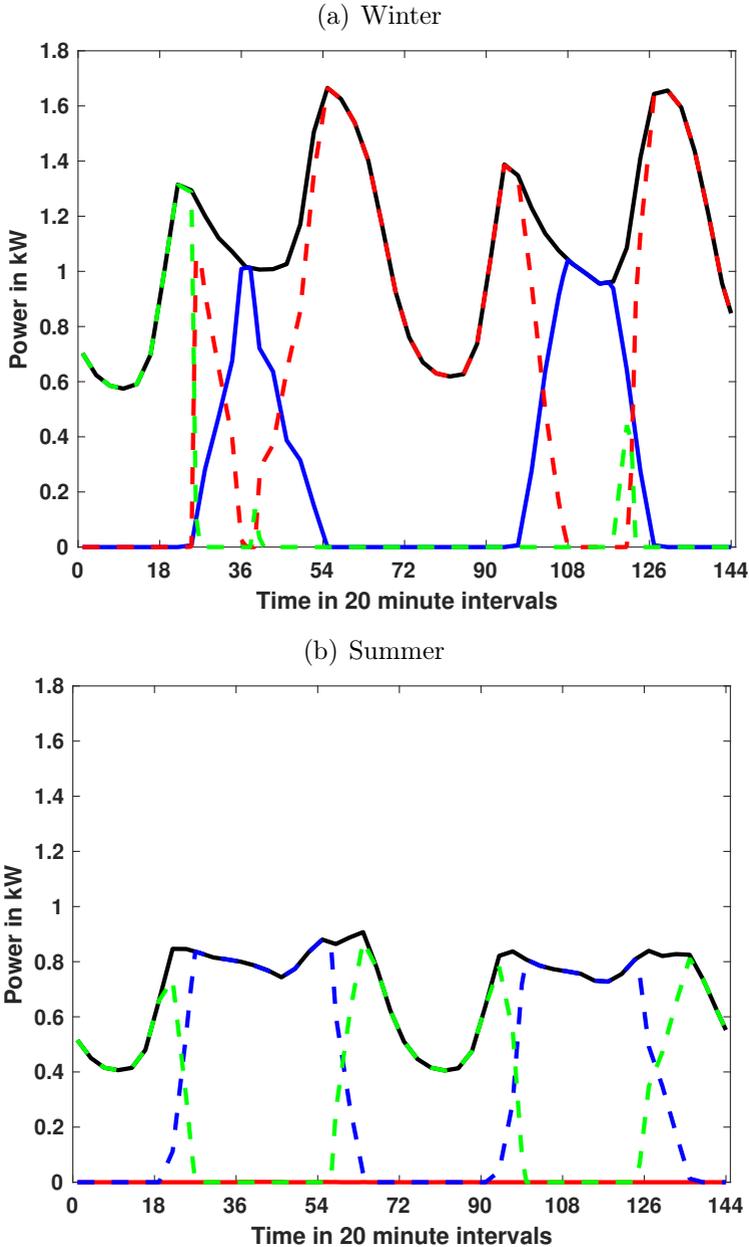


Figure 20: Contributions $A_1(t)$ (blue, solar PV power to load), $G(t)$ (red, grid use), and $B(t)$ (green, battery discharge rate) add to the load $L(t)$ (solid black line), winter and summer 2016



6 Summary

The non-linear programming model of [Section 5](#) predicts that grid use varies greatly from season to season, even in one location. Sunny summer days could result in negligible grid use by those with average needs, average solar power generation capacity, and one battery of average capacity. However, this is not the case on wet summer days, which are part of a New Zealand summer.

Grid use is predicted to satisfy most of the load in winter. The non-linear programming model predicts that a battery could soon be discharged in mid-winter, with little opportunity for charging, given that PV generated power is used locally to satisfy the household load in preference to battery charging or feeding into the grid. Only on very sunny days in winter could a battery be expected to be charged and then discharged. However, even on winter days, the availability of solar PV generated power decreases grid use during the middle of the day.

Households in one residential GXP region could be partitioned into groups, depending on characteristics such as area of solar panels and battery capacity. For each group, the non-linear programming model could be used to predict grid use and the solar PV generated power injected into the grid, as time series. The results for these groups of households could be scaled and combined to predict the power drawn from the grid and fed back to the grid, enabling Transpower to predict grid use in a particular location at a particular time of the year. This is work for the future.

Two approaches to electricity load forecasting were investigated, assuming no solar PV generation. In the first approach ([Section 3](#)), load data from an Auckland node were used in an experiment to assess and compare various forecasting methods, namely Holt–Winters methods, time series linear regression, an autoregressive integrated moving average (ARIMA) model and an ARIMA model with regressors/explanatory variables (ARIMAX). Preliminary forecasting results revealed that the multiple seasonalities that exist in the time series create a stream of fluctuations. This makes the forecasting task

more challenging and led to further investigation. The advantage of using Holt–Winters methods is that only the time series (the electricity load in our case) is needed to make the forecasts. However, that limits the inclusion of all other predictors into the analysis. With the time series linear regression models, other predictors may be incorporated into the model but the linear regression models can only capture the trend and well-specified seasonalities in the time series, not any other erratic fluctuations. The best candidate for the electricity load data might be the ARIMA with regressors or explanatory variables (ARIMAX) as all the important features of the time series can be identified and analysed, and all other predictors can also be included. The Holt–Winters additive method and the ARIMAX model return the best forecasts for the electricity load consumption in March 2017.

The second approach to load forecasting ([Section 4](#)) is an alternative form of linear regression, taking into account historic use on comparable days, air temperature, and solar radiation. Five years of Transpower data was combined with five years of data from four meteorological stations in Auckland. This performs even better than the first approach, yielding the lowest RMSE and MAPE.

Acknowledgements The authors are grateful to Transpower for bringing this problem to MINZ2018 and for the valuable help of their representatives during the week. H. M. Nasir and R. Choudhary also worked on this problem during MINZ2018 and we thank them for their contributions.

References

- Beaudin, M. and H. Zareipour (2015). “Home energy management systems: A review of modelling and complexity.” In: *Renewable and Sustainable Energy Reviews* 45, pp. 318–335. ISSN: 13640321. DOI: [10.1016/j.rser.2015.01.046](https://doi.org/10.1016/j.rser.2015.01.046) (cit. on p. [M5](#)).

- Boland, J. et al. (2016). “Optimisation of photovoltaic system and storage”. In: *ANZIAM Journal* 58, pp. 1–32. DOI: [10.21914/anziamj.v58i0.11471](https://doi.org/10.21914/anziamj.v58i0.11471) (cit. on p. M5).
- Collotta, M. and G. Pau (2017). “An Innovative Approach for Forecasting of Energy Requirements to Improve a Smart Home Management System Based on BLE”. In: *IEEE Transactions on Green Communications and Networking* 1.1, pp. 112–120. ISSN: 2473-2400. DOI: [10.1109/TGCN.2017.2671407](https://doi.org/10.1109/TGCN.2017.2671407) (cit. on p. M5).
- Kikusato, H. et al. (2019). “Electric Vehicle Charge–Discharge Management for Utilization of Photovoltaic by Coordination Between Home and Grid Energy Management Systems”. In: *IEEE Transactions on Smart Grid* 10.3, pp. 3186–3197. ISSN: 1949-3061. DOI: [10.1109/TSG.2018.2820026](https://doi.org/10.1109/TSG.2018.2820026) (cit. on p. M5).
- Niimura, T. et al. (2012). “Profiling residential PV output based on weekly weather forecast for home energy management system”. In: *2012 IEEE Power and Energy Society General Meeting*, pp. 1–5. DOI: [10.1109/PESGM.2012.6345020](https://doi.org/10.1109/PESGM.2012.6345020) (cit. on p. M5).
- Shakeri, M. et al. (2017). “An intelligent system architecture in home energy management systems (HEMS) for efficient demand response in smart grid”. In: *Energy and Buildings* 138, pp. 154–164. ISSN: 0378-7788. DOI: [10.1016/j.enbuild.2016.12.026](https://doi.org/10.1016/j.enbuild.2016.12.026) (cit. on p. M5).
- Weniger, J., J. Bergner, and V. Quaschnig (2014). “Integration of PV power and load forecasts into the operation of residential PV battery systems”. In: *Proceedings of the 4th Solar Integration Workshop, Berlin 10–11 November 2014*. DOI: [10.13140/2.1.3048.9283](https://doi.org/10.13140/2.1.3048.9283) (cit. on pp. M3, M22, M23).
- Zhou, B. et al. (2016). “Smart home energy management systems: Concept, configurations, and scheduling strategies.” In: *Renewable and Sustainable Energy Reviews* 61, pp. 30–40. ISSN: 13640321. DOI: [10.1016/j.rser.2016.03.047](https://doi.org/10.1016/j.rser.2016.03.047) (cit. on p. M5).

Author addresses

1. **C. Z. W. Hassell Sweatman**, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, NEW ZEALAND.
<mailto:catherine.sweatman@aut.ac.nz>
orcid:0000-0002-1616-8209
2. **N. Wichitaksorn**, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, NEW ZEALAND.
<mailto:nuttanan.wichitaksorn@aut.ac.nz>
3. **A. Jiang**, School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, NEW ZEALAND.
<mailto:Aaron.Jiang@genesisenergy.co.nz>
4. **T. Farrell**, School of Mathematical Sciences, Queensland University of Technology, Queensland 4000, AUSTRALIA.
<mailto:t.farrell@qut.edu.au>
5. **N. Bootland**, Mathematical Institute, University of Oxford, UNITED KINGDOM.
<mailto:bootland@maths.ox.ac.uk>
6. **G. Miskell**, School of Chemical Sciences, University of Auckland, NEW ZEALAND.
<mailto:georgia.miskell@gmail.com>
7. **G. Pritchard**, Department of Statistics, University of Auckland, NEW ZEALAND.
<mailto:g.pritchard@auckland.ac.nz>
8. **C. Chrystall**, Transpower, NEW ZEALAND.
<mailto:charles.chrystall@transpower.co.nz>
9. **G. Robinson**, Transpower, NEW ZEALAND.

<mailto:gareth.robinson@transpower.co.nz>