# Unlocking the Potential of Demand Response Schemes

## Using Machine Learning Techniques


**SUMMARY**

- Demand response (DR) programs can benefit electricity consumers, distribution network service providers, and system operators. However, the complexity and characteristics associated with the loads connected to customer premises and the baseline calculation process still require considerable research to exploit maximum benefits from DR programs.
- An essential part of demand response programs for commercial and industrial (C&I) customers is the baseline used to calculate the monetary benefits to customers. This article studies the potential use of interpretable machine learning (ML) tools for enhanced and more accurate baselines.
- The study investigated how temperature combined with consumption history and simple and interpretable ML methods can be used to provide more precise demand forecasts and thus baselines closer to actual load profiles.
- The analysis shows that for most C&I portfolios, using a simple ML model, such as a polynomial regression model, results in a more accurate baseline than the current baselines used in the market. The proposed ML models are not black box and thus can be fully explained and interpreted.


### WHAT IS DEMAND RESPONSE?

Demand response (DR) is an essential tool in power grid operation. It encourages customers to reduce non-essential consumption of electricity and switch to onsite generation when a relevant aggregator gives a DR signal. It is an agreement between the customers and the aggregator based on certain conditions such as time intervals, price, and load. DR can help manage the supply and demand of electricity in a balanced way during peak hours, reducing operating costs and installation costs and mitigating potential power grid failures. It helps to optimise investments, thus keeping affordability under control while maintaining the reliability and security of the power supply.

### DEMAND RESPONSE FOR COMMERCIAL AND INDUSTRIAL CUSTOMERS

Demand response for C&I customers provides a range of curtailment solutions for businesses with at least 250 kW of curtailable load and/or 250 kW of backup generation capacity. These DR programs require customer action, which depends on program performance and measurement. An essential part of DR programs for C&I customers is the baseline used to calculate the monetary benefits to

customers. It is necessary because performance is calculated by comparing the actual load measurement to what it would have been if an event was not called. This can be challenging when the load profile is not flat and/or dependent on the conditions of the event, such as temperature. DR events are more likely to occur on extreme heat days, meaning more energy consumption than on regular days. So, the calculation needs to account for the temperature conditions on the day. The current baseline calculations, such as the averaging or maximum load of historical interval meter data, have a limitation in accurately considering weather information, such as temperature.

### HOW CAN MACHINE LEARNING TECHNIQUES HELP?

Machine learning (ML) is a rapidly growing field with its roots in Artificial Intelligence. It combines the fields of mathematics, engineering, and computer science and may be thought of as creating computer-generated inferences based on past observations and future parameters. Over time, as the model encounters more examples of behaviour, its predictions become more accurate. The Baseline model [1] uses the same historical time stamps to calculate the value at the same future time (average value), e.g., using the 10:00 historical data to predict 10:00 in the future with a time window (the number of historical days used as inputs). However, this does not consider the continuity and dependency of the time series.

Moreover, meteorological information such as temperature and humidity may influence the demand. Thus, building the prediction model by considering these factors may help obtain a more accurate demand prediction. ML models provide more possibilities and flexibilities for the input data and learn the relationship between the inputs and outputs (predicted values). This paper proposes ML-based baseline calculation methods to improve the baseline calculation process of two different real-world example case studies.

## CORRELATION OF DEMAND WITH METEOROLOGICAL DATA

The relationship between electricity demand and climatic conditions is often location-dependent because it depends on people's geographical location, style of living, type of industries and commerce, type of urbanism, and type of implemented building codes. We use temperature data from the Australian Government Bureau of Meteorology to study the effect of ambient temperature on electricity demand and DR.

We investigate the correlation between the total daily energy consumption and the Heating Degree Days (HDD) or Cooling Degree Days (CDD) calculated in accordance with [2]:

"CDDs & HDDs, which indicate the level of comfort, are based on the average daily temperature. The average daily temperature is calculated as follows: [maximum daily temperature + minimum daily temperature] / 2. If the average daily temperature falls below comfort levels, heating is required and if it is above comfort levels, cooling is required. The HDDs or CDDs are determined by the difference between the average daily temperature and the BASE (comfort level) temperature. The BASE values used are 12 and 18 degrees Celsius for heating and 18 and 24 degrees Celsius for cooling."

We use the conventional Pearson correlation coefficient [3] for 12 C&I customer portfolios included in the study, as listed in Table 1. The table shows that the portfolios that may potentially be considered to use temperature as part of its ML model are the Telecom and Shopping Centre portfolios. However, in the following section, we still use temperature as an exogenous variable for ML models to demonstrate the effect of temperature on demand forecasting. On the other hand, the Shopping Centre- and University portfolios use relatively high energy-consuming cooling- and heating systems to maintain the inside temperature close to acceptable comfort levels.

*Table 1: Pearson correlation (%) between the daily total energy consumption, HDD and CDD.*

| Customer Portfolio | Correlation to HDD (18°C) % | Correlation to CDD (24°C) % |
|---|---|---|
| Chemical Plant | - | 0.8 |
| Medium Manufacturing | - | 7.2 |
| Metal Recycling | - | -4.8 |
| Sandstone Quarry | - | -18.4 |
| Shopping Centre | - | 81.1 |
| Telecom | 45.6 | -9.9 |
| Telecom VIC | -16.5 | 17.0 |
| University | - | 15.9 |
| Water Utility 1 | - | -6.7 |
| Water Utility 2 | - | 11.7 |
| Water Utility 5 | - | -0.3 |
| Water Utility 6 VIC | -16.0 | - |

### Case Study 1: Twelve different C&I customers

The dataset in this case study contains daily power consumption in kWh for 12 C&I customer portfolios, listed in Table 1. The data is collected every 30 minutes for different periods. It also includes the power consumption of an event day (actual DR event), which has been excluded during the training and testing of the ML prediction models. We use different ML techniques for DR prediction and compare them with the baseline model. ML models, including polynomial linear regression (a nonlinear predictor, NLR), support vector regression (SVR), and Bayesian linear regression (BLR), are used to compare against the Baseline model. The accuracy (ACC) metric has been used to evaluate the performance of all models.

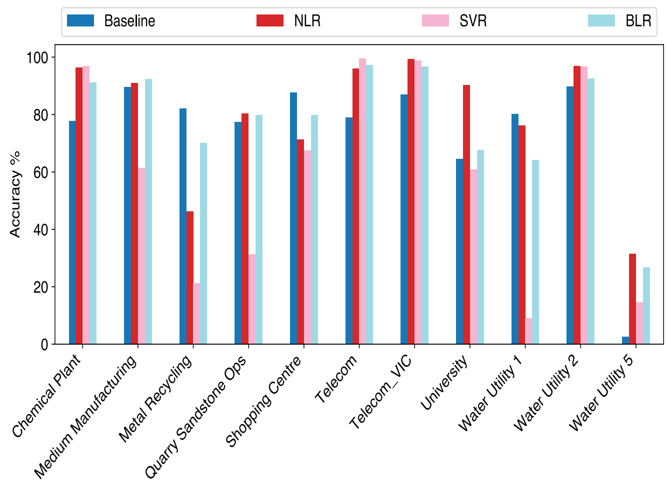


*Figure 1: Testing accuracy performance compared to Baseline NLR, SVR and BLR models using only energy consumption.*

For each model, the following settings are explored to verify the effectiveness of weather-related data on the demand prediction: 1) energy consumption, 2) energy consumption and temperature.

Figure 1 shows the percentage accuracy results on the test set for 11 C&I customer portfolios compared to the baseline model for setting one, using only energy consumption. The NLR model achieves better accuracy for 8 out of 11 C&I customers than the Baseline model. The BLR model shows better accuracy for 6 out of 11 C&I customers compared to the average baseline. Lastly, the SVR model shows the least accuracy, with 4 out of 11 better accuracies than the Baseline model. The overall accuracy of the case study 1 dataset for setting one is summarised in Table 2.

*Table 2: Overall accuracy of comparison of Baseline- and Machine Learning predictor models for both settings (Case Study 1).*

| Model | Overall Accuracy % | |
|---|---|---|
| | *Energy Consumption* | *Energy Consumption and maximum temperature* |
| Baseline | 74 | 74 |
| NLR | **79** | **79.4** |
| SVR | 59 | 60 |
| BLR | 78 | 77.8 |

The NLR model achieves the highest accuracy, with an accuracy of 79%. Figure 2 shows the results of the ML models using daily maximum temperature as the exogenous variable for the predictive model. In general, we do not observe much improvement in the model performance compared to setting one. The only difference in the performance is observed for the Shopping Centre portfolio. Both the NLR and SVR accuracies have improved by including the maximum temperature, as shown in Figure 2. In the next paragraph, we investigate the C&I monetary benefits

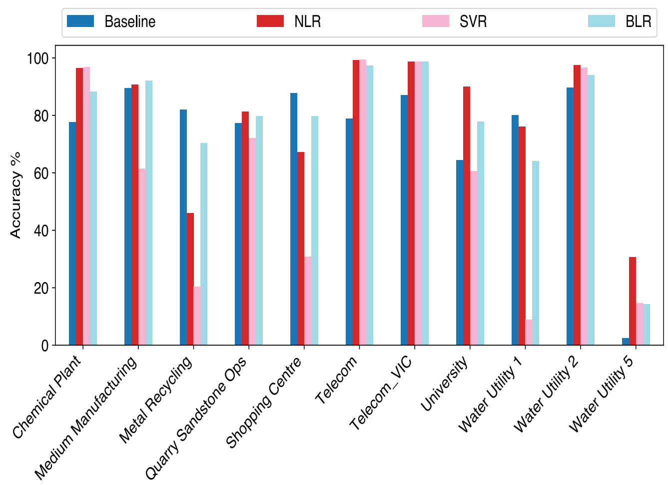


*Figure 2: Testing accuracy performance compared to Baseline, NLR, SVR and BLR models using energy consumption- and daily maximum temperature data..*

of using ML models compared to the baseline model.

We chose the best performing model, i.e., the NLR model - based on the accuracy presented for the DR analysis and compared it with the Baseline model. Figure 3 shows the DR prediction of the NLR model and Baseline model compared to the actual event consumption. Based on the DR event, we calculate the percentage kWh difference between the ML model and the Baseline model compared to the actual DR event, as shown in Table 3. The results show that the NLR model will increase monetary benefits for 9 out of 11 C&I customer portfolios compared to the baseline model for these DR events. The only C&I customer portfolios where the baseline model shows improved monetary benefits are the Water Utility 2- and Water Utility 5 portfolios.

*Table 3: The demand difference (i.e., the predicted values of Baseline/NLR - the actual values) percentage (i.e., demand differences divided by the actual values) for Case Study 1.*

| Portfolio | Event Date | Mean % Difference | |
|---|---|---|---|
| | | *Baseline* | *NLR* |
| Chemical Plant | 1/14/2021 | 10.37% | 13.36% |
| Medium Manufacturing | 1/31/2020 | 0.04% | 0.32% |
| Metal Recycling | 12/17/2020 | -0.19% | 1.59% |
| Sandstone Quarry | 1/31/2020 | 2.82% | 4.22% |
| Shopping Centre | 1/31/2020 | -0.48% | -0.18% |
| Telecom | 6/10/2021 | 0.98% | 1.41% |
| Telecom VIC | 5/20/2021 | 0.31% | 0.52% |
| University | 1/31/2020 | -0.35% | -0.14% |
| Water Utility 1 | 1/31/2020 | -0.03% | -0.04% |
| Water Utility 2 | 1/31/2020 | 0.71% | 0.83% |
| Water Utility 5 | 1/31/2020 | 0.13% | -0.18% |

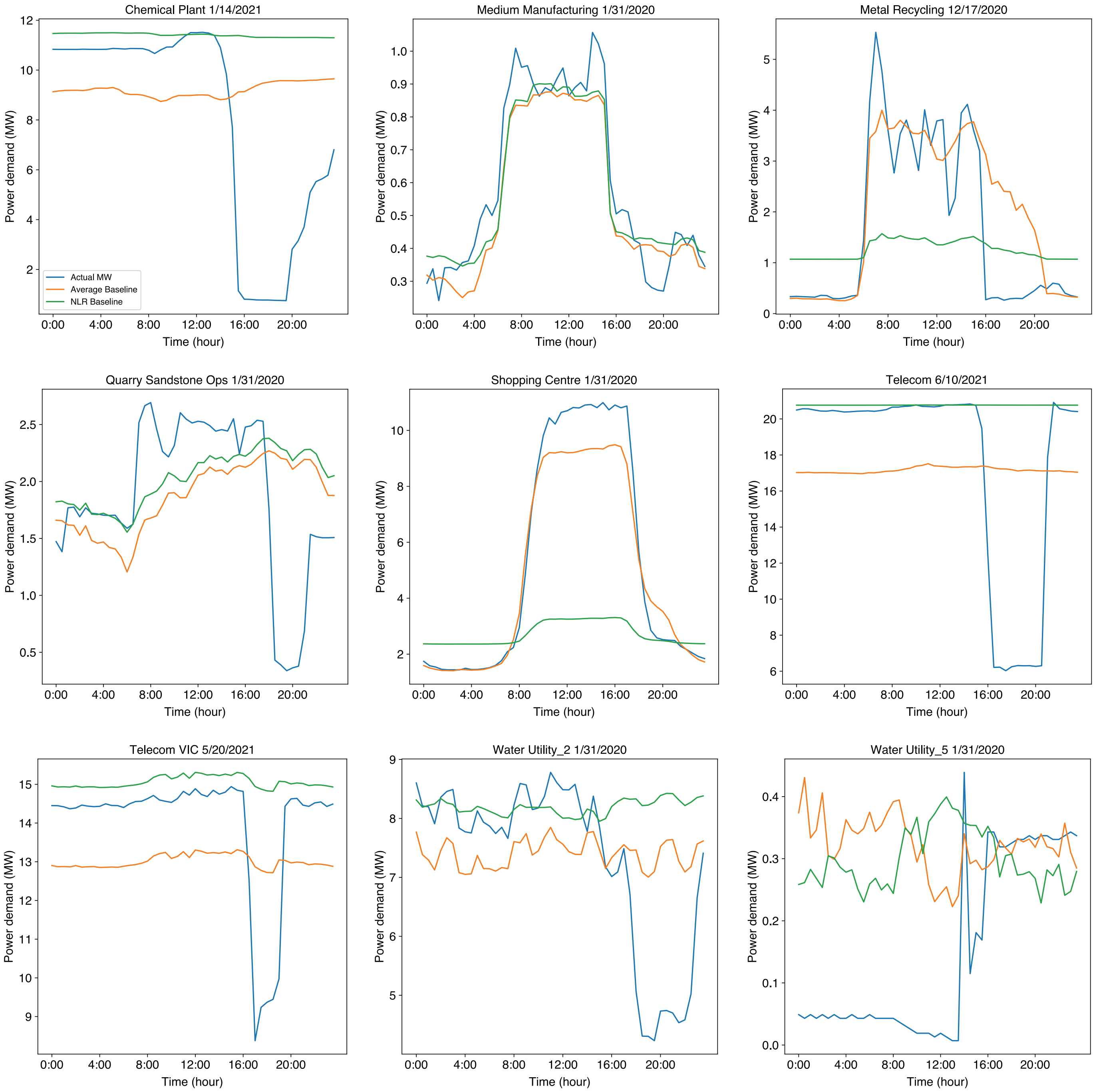


*Figure 3: Demand response predictions for the NLR model compared to the baseline for the Case Study 1 dataset containing 11 C&I customer portfolios (results shown for 9 portfolios).*

## Case Study 2: Water Utility 1

In this case study, the data from a water utility company is collected every 30 minutes from 01/04/2018 to 31/12/2020. We use the data before 28/12/2020 (training set) to train the prediction model and the data from 28/12/2020 to 30/12/2020 (testing set) to test the trained model. We choose linear regression (LR) and NLR as the predictors. We design two different models: **Model 1** uses the same historical time stamps as the baseline, but can add more exogenous factors, to predict the future event; **Model 2** uses the time stamps before the predicted time stamps as inputs with a time window (the number of previous time stamps denoted as TW), e.g., using 7:00, 7:30, 8:00, 8:30, 9:00, and 9:30 data to predict 10:00 for the same day. Since this dataset has more weather-related information, we explore the performance of the LR- and NLR predictors for the following scenarios: energy consumption (Demand), energy consumption and temperature (DemandTemp), and energy consumption, temperature, daily precipitation rate and solar radiation information (DemandTempOthers).

Figure 4 shows the root mean square error (RMSE - smaller is better) testing performance over the Baseline, LR-, and NLR predictors across different scenarios for **Model 1** and **Model 2** with different time window (TW) settings, where TW=10 in **Model 1** and **Model 2** means that the same time stamps of the previous 10 days and the previous 10 time stamps are regarded as the inputs, respectively. From the results, **Model 2** leads to better performance than **Model 1**, which means considering the continuity and dependen-

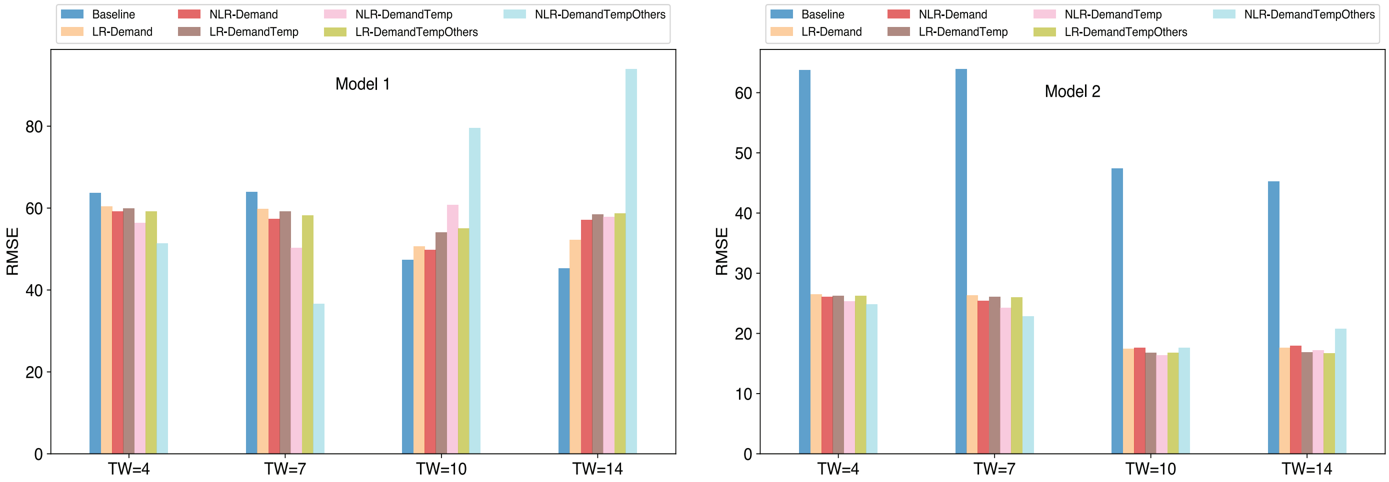


*Figure 4: The testing RMSE performance of the baseline, LR and NLR predictors for the Demand-, DemandTemp- and DemandTempOthers scenarios for Model1 and Model2.*

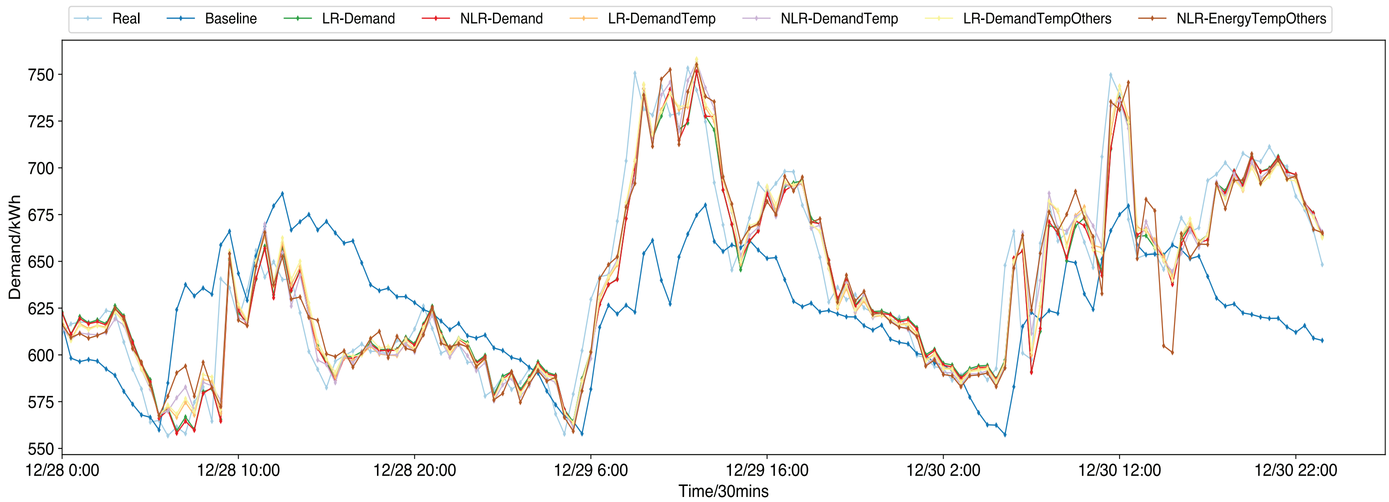


*Figure 5: The comparison of the actual- and predicted demand with the Baseline-, LR- and NLR predictors.*

cy of the time series is more accurate and reasonable for demand prediction. Also, NLR-DemandTemp with TW=10 (i.e., the best performance of NLR across all cases and TW settings in **Model 2**) outperforms LR-DemandTemp with TW=10 (i.e., the best LR performance across all cases and TW settings in **Model 2**). The following analysis is based on **Model 2**.

Figure 5 illustrates the actual and predicted demand of the Baseline-, LR-, and NLR predictors using different combinations of demand and weather-related information. It shows that the LR- and NLR predictors with only demand data (and weather-related data) lead to more accurate predicted values than the Baseline model. To further explore the influence of demand prediction on DR, we choose the optimal performance of the LR and NLR predictors among the investigated combinations to be compared with the Baseline model.

Figure 6 shows the demand difference using the Baseline-, LR-, and NLR predictors for 28/12/2020, 29/12/2020, and 30/12/2020. The green line represents the demand difference of the actual values (all zeros). For most time stamps, the LR- and NLR predictors have smaller demand differences than the Baseline model, which means that their prediction is more accurate and that the Baseline model has the most negative impact on the demand response over each day. Since the tariff structure is not used to calculate the DR monetary benefit, we summarise the average demand difference percentage (demand difference divided by the actual value) for the Baseline-, LR-, and NLR predictors over the testing days in Table 4. Since an average value close to (or equal to) 0% is maximally beneficial, the NLR- and LR predictors perform alternately the best on the testing days, and both of them lead to better demand response than the Baseline model.

*Table 4: The demand difference (i.e., the predicted values of Baseline/LR/NLR - the actual values) percentage (i.e., demand differences divided by the actual values).*

| **Date** | **Mean % for each Model** | | |
|---|---|---|---|
| | *Baseline* | *LR* | *NLR* |
| 28/12/2020 | 3.91% | 0.36% | 0.19% |
| 29/12/2020 | -4.26% | -0.16% | -0.17% |
| 30/12/2020 | -5.02% | -0.20% | -0.21% |

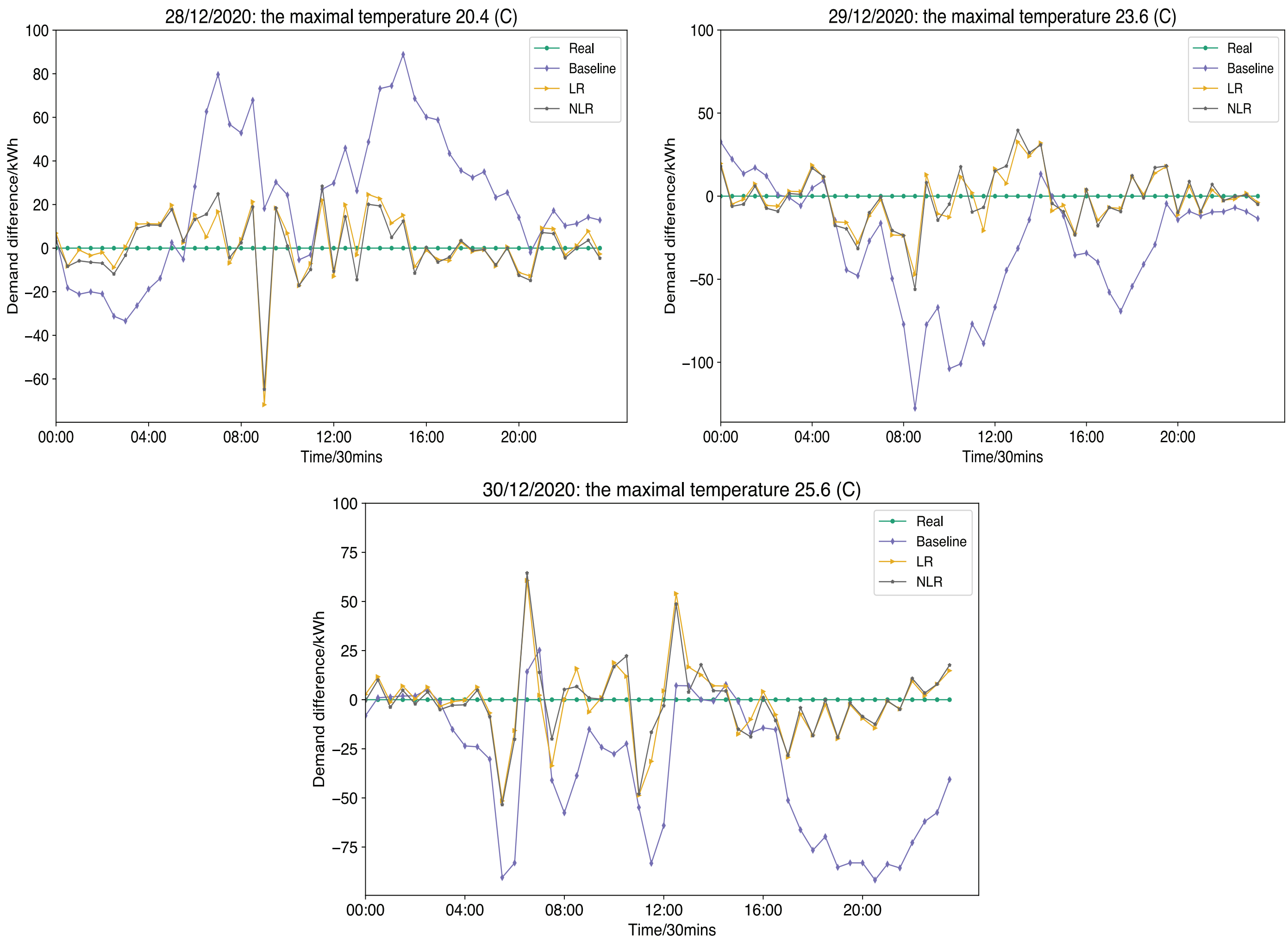


*Figure 6: Demand difference (predicted values of Baseline/LR/NLR - the actual values) for three test days.*

**Case Study 3: Water Utility 2**

In this case study, the data from another water utility company is collected every 15 minutes from 28/05/2019 to 28/05/2021. Given the ML predictors, the inputs used to predict the future demand will include ten historical demand data points. Before training, data cleansing is performed on the consumption days, days with undefined values, weekends, and days with power outages. ML models, including polynomial linear regression (a nonlinear predictor, NLR), support vector regression (SVR), and Bayesian linear regression (BLR), are used to compare against the Baseline model. The accuracy (ACC) metric was used to evaluate the performance of all models. Similar to the other case studies, the following settings were explored to verify the impact of weather-related data on demand prediction: 1) energy consumption and 2) energy consumption and temperature. Figure 7 shows prediction accuracies of NLR, SVR, BLR, and the Avg. Baseline models when only energy consumption data is used. Figure 8 shows the prediction accuracies of NLR, SVR, BLR, and the Avg. baseline models using energy consumption and maximum temperature data. There is an observed difference in the performance between the ML predictive models and the Baseline model. However, the ML predictive models' performance decreases when the maximum temperature is included in the model, as shown in Figure 8. Figure 9 shows the DR predictions of the NLR, SVR, and BLRs models and the Avg. Baseline model for 27/05/2021 compared to the actual event consumption data.

**Future Outlook**

- Machine learning methods can provide considerably more accurate predictions than the baselines, especially when sufficient data is available (e.g., Case Study 2).
- From the prediction accuracy results for the Case Study 1 dataset, we observe that the NLR model performs better than the average Baseline model for several C&I customers.
- Accordingly, we highly recommend using the NLR model for the DR baseline calculation for the following C&I customer portfolios: Chemical Plant, Telecom, Telecom VIC, Water utility 2, and Water utility 5.
- However, for other C&I customer portfolios, we recommend using the Baseline model since the difference in performance is negligible. The earlier recommendation may also change based on the availability of more data.

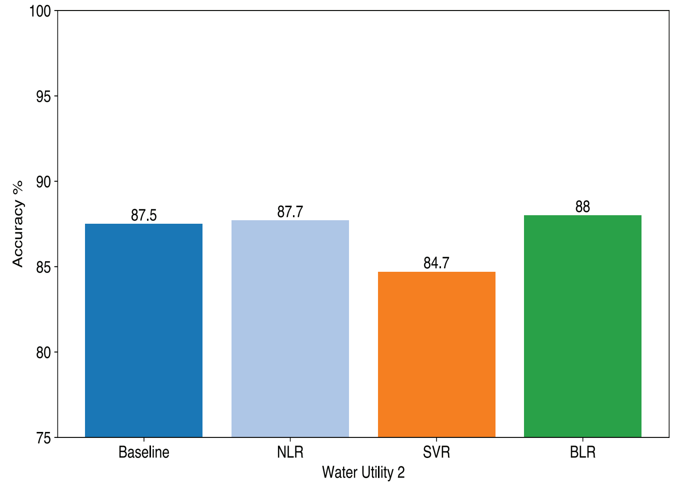


*Figure 7: Testing accuracy performance over Baseline, NLR, SVR and BLR models using only energy consumption data.*

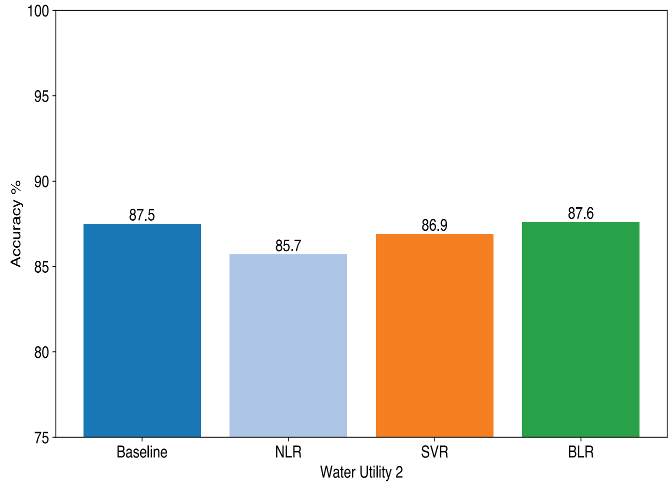


*Figure 8: Testing accuracy performance over Baseline, NLR, SVR and BLR models using energy consumption and daily maximum temperature data.*

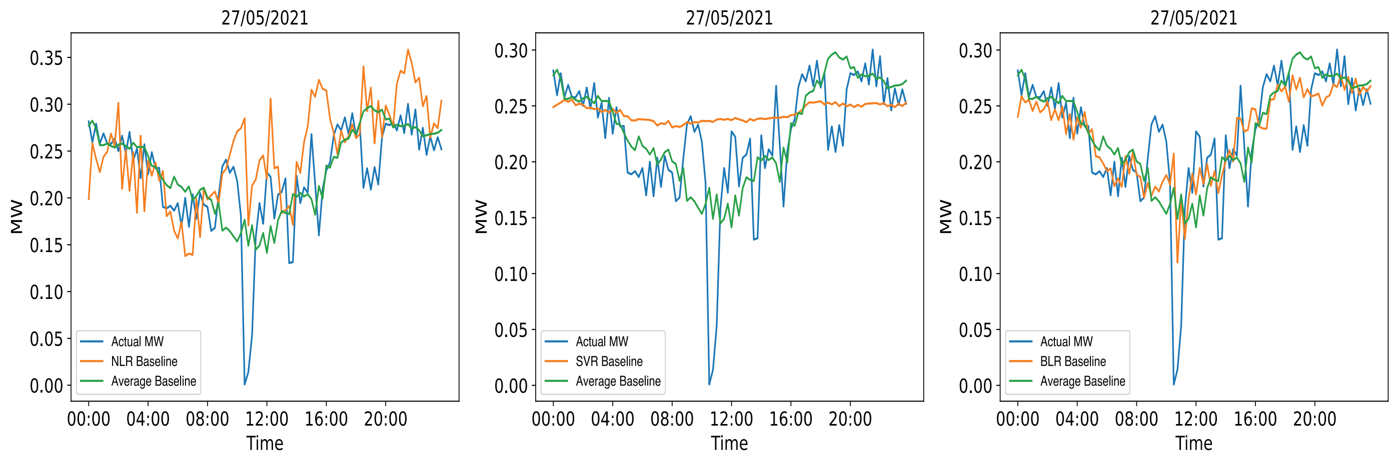


*Figure 9: Demand response predictions for NLR, SVR and BLR compared with the Average Baseline using the Case Study 3 dataset.*

- Furthermore, the addition of hourly temperature data in the ML model should be included in future work. Case Study 2 shows that adding temperature data may result in more accurate predictions, demonstrating that temperature influences demand.
- Even without weather-related information, the NLR- and LR models outperform the Baseline model.
- Since the LR- and NLR models have almost similar performance and the LR model is simpler and easier to understand, we recommend using the LR model as the Baseline model. The DR analysis shows that the predicted energy demand values of the NLR- and LR models are much closer to the actual values. The average demand difference percentage of the NLR- or LR models is much less than the Baseline model.
- In future, we recommend using the tariff structure to calculate the monetary benefits to validate this finding further.


## Acknowledgement

The project is supported by the Centre for New Energy Technologies (C4NET), AGL, Gippsland Water Factory, Greater Western Water and the Victorian Department of Environment, Land, Water and Planning.



**For further information contact:**

**Assoc./Prof. Lasantha Meegahapola**

**Email:** lasantha.meegahapola@rmit.edu.au


## REFERENCES


[1] Y. Zhang, W. Chen, R. Xu, and J. Black, "A cluster-based method for calculating baselines for residential loads," IEEE Transactions on Smart Grids, vol. 7, no. 5, pp. 2368-2377, 2015.

[2] Australian Government Bureau of Meteorology. "Annual and monthly heating and cooling degree days." Available: http://www.bom.gov.au/climate/map/heating-cooling-degree-days/documentation.shtml.

[3] R. J. Mislevy, "Recent developments in the factor analysis of categorical variables," Journal of Educational Statistics, vol. 11, no. 1, pp. 3-31, 1986.



Written by **Hui Song, Ammar Kamoona, Mahdi Jalili, Richardt Wilkinson, Kedem Levy, Xinghuo Yu, Brendan McGrath, Kazi Hasan and Lasantha Meegahapola**