Evaluation of Data-driven Models for Short-term Electrical Load Forecasting in Office Buildings

Farid Bahiraei*[1], Hadia Awad[1], Araz Ashouri[1]
[1] National Research Council, Ottawa, Canada

Abstract

Load forecasting algorithms play a key role in the successful management of building operation schemes such as peak management and time-of-use optimization. Such implementations require reliable historical energy use data to train the models, which might be unavailable for newer buildings, or in cases of major interruptions in data collection. In this paper, we present data-driven approaches to predict electrical loads for commercial buildings with limited training data. The models are based on temporal, autoregressive, and exogenous variables and are designed to reduce the impacts of the COVID-19 pandemic on their prediction performance. Model validation is performed with data from an existing commercial building connected to a smart metering system. The results indicate that, compared to simple and linear models, non-linear models provide enhanced and acceptable forecast accuracy even when trained over relatively short periods of time, and despite changes in the energy use pattern during the pandemic.

Introduction

As per the International Energy Agency (Cozzi and Gould 2021), in 2021 Canada had the largest per capita energy consumption, three times the global average. In 2019, the building sector had become the second highest energy consumer, after transportation, in OECD countries Worldwide, final energy use in buildings grew from 118 EJ in 2010 to 128 EJ in 2019, and reached an all-time high of related CO_2 emissions. Effective and reliable approaches need to be implemented in order to mitigate energy demand and to pro-actively control power peak events. Energy load forecasting plays a major role in power utility operation and design to model consumption and generation (Collotta and Pau 2017, Ahmad *et al.* 2018).

Accurate building-level load forecasting can support many energy management tasks, such as peak reduction, predictive control, demand-response programs and optimization of distributed energy resources. Depending on the time horizon, there are multiple categories for energy load forecasting (Koukaras *et al.* 2021). For example, Very Short-Term Load Forecasting (VSTLF) predicts in a range of one minute to one or several hours, typically using historical data from the same day (Guan *et al.* 2013). Short-Term Load Forecasting (STLF), on the other hand, predicts energy loads for the next 30 minutes up to two weeks (Jacob *et al.* 2020). Mid-Term Load Forecasting (MTLF) (Amjady and Daraeepour 2011) and Long-Term Load Forecasting (LTLF) (Agrawal *et al.* 2018) deal with energy use over the next few months or years,. In the current study, we focus on STLF at an hourly interval, where the forecast horizon is one day ahead. The methods used for VSTLF and STLF are developed mainly from classical statistical techniques or artificial intelligence techniques (Raza and Khosravi 2015, Yildiz *et al.* 2017). As buildings are more commonly equipped with smart meters that gather related energy use data with high granularities, artificial intelligence techniques have attracted more attention (Bourdeau *et al.* 2019). Artificial neural networks are one of the most common methods for load forecasting in buildings (Wang and Srinivasan 2017). (Kim *et al.* 2019) presented a hybrid multi short-term electrical load forecasting model. They used 21 inputs to forecast the load of the next day. Their model was a recurrent inception convolution neural network (RICNN) that combined RNN and 1-dimensional CNN (1-D CNN). They used the 1-D convolution inception module to calibrate the prediction time and the hidden state vector values calculated from nearby time steps. The RICNN model outperformed the benchmarked multi-layer perceptron, RNN, and 1-D CNN in day-ahead electric load forecasting. (Muzaffar and Afshari 2019) proposed a recurrent neural network to do the long-term and short-term load forecasting in forecast horizons of 24 and 48 hours, 7 and 30 days. They compared their results to a time-series approach and concluded that ANN results in a better performance. A study by (Jones et al. 2012) investigated Modified Learning From Experience (MLFE) with a recursive least squares (RLS) method and compared it against the LNBL's Baseline Profile Model to predict electricity loads in commercial buildings. The MLFE RLS method has indicated an acceptable level of accuracy during mornings and early afternoons, however, in late afternoon hours the BLP3 performed better. It was concluded that taking occupancy into account can potentially improve load

predictions, especially in a California climate that requires higher cooling loads during late afternoon hours.

Random forest regression has also attracted attention in load forecast applications. (Moon *et al.* 2018) proposed a hybrid short-term load forecast model for a large building complex by using random forest and multilayer perceptron. They collected six years of electrical load data from a university campus. They used decision trees to classify the data with input parameters including date, day of the week, holiday, and academic year. The authors considered various configurations for random forest and multilayer perceptron and evaluated their prediction performance using the validation set to determine the optimal configuration. They reported that the proposed hybrid forecast model performs better than other popular single forecast models. (Chen *et al.* 2019) developed a prediction model with a pattern recognition scheme to directly predict the desired levels using simpler classifier models without undergoing regression. The proposed pattern recognition classifier was compared to its regression method using a similar algorithm applied to a real-world energy dataset. A random forest algorithm outperformed other widely used machine learning techniques in previous research. However, when the number of desired levels increased, the performance of their proposed model approached the accuracy of the conventional regression methods.

Studies on load forecasting in buildings commonly rely on relatively long periods of data collection (Bourdeau *et al.* 2019). The COVID-19 pandemic and the preventive measures taken to contain the spread of the disease have drastically changed the patterns of energy use in buildings. As economies recover from the pandemic and the initial operational measures in reaction to it, it is expected that some inconstancies in electrical load patterns will persist during the post-pandemic era. Therefore, it is desired to minimize these impacts on the performance of the forecasting models. In this study, we developed models with relatively short training data requirements so the models can quickly adapt to changes in electricity use patterns.

Case Study

We collected the building electricity use data from a commercial building located in Maritime Canada. The building recently went through a major retrofit, which included upgraded heat pumps and a PV and battery storage system. The building is also connected to fossil fuel boilers. In the current HVAC sequence of operation, the heat pumps are used for cooling and the boilers are used when the heat pumps cannot meet the building's heating demand. We pre-processed the measured electricity use data to remove the noise and outlier datapoints. If the electricity use measurements were missing for less than 4 consecutive hours, a linear interpolation was used to fill the missing values. If data were missing for more than 4 consecutive hours, that period of time was flagged as unusable for the model development. The resulting dataset, shown in Figure 1, contains electricity use data recorded between June 2019 and the end of December 2020 at hourly time intervals.

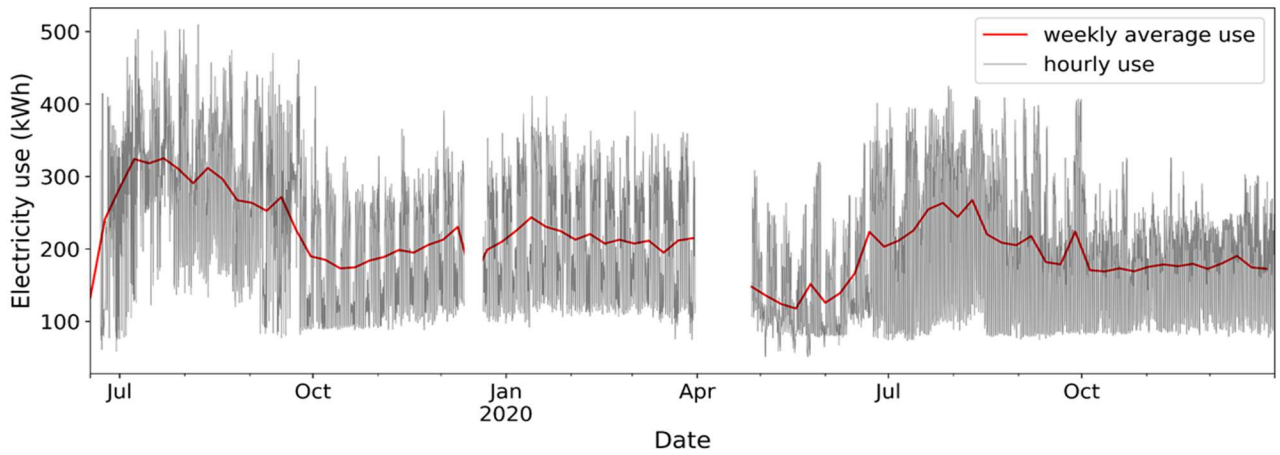

Figure 1: Electricity load of the building

Figure 1 shows three gaps in the data. We excluded these periods of time from our analysis. This figure also shows around 24% reduction in the average electricity use after the COVID-19 pandemic that was declared in March 2020.

Methods

Based on the review on short-term data-driven load forecasting, we decided to use three models: a linear (elastic-net) model, a non-linear (random forest) model and a baseline model. We used the baseline as a benchmark to provide context to the elastic-net and random forest models' accuracy.

Linear model

We used elastic-net as the candidate linear model to lower the risk of over-fitting that may happen during the typical training of linear regression models. Elastic-net is an ordinary least squares regression that linearly combines the l_1 and l_2 penalties of the lasso and ridge methods. The loss function of an elastic-net model can be defined as:

$$L = \frac{\|y - X\beta\|}{2n} + \lambda\left(\frac{1-\alpha}{2}\|\beta\|_2^2 + \alpha\|\beta\|_1\right) \quad (1)$$

The mixing parameter (α) can vary between 0 (ridge) and 1 (lasso). Note that α and λ should be tuned. The values of $[10^{-4}, 10^{-3}, 10^{-2}]$ were used for both α and λ. The tuning process is discussed in the model development section.

Non-linear model

As discussed in the Introduction section, neural networks have been widely used for load forecasting in buildings. In this study, we used random forest regression as the candidate non-linear model. Random forest is able to provide a non-linear regression. Also, the random forest algorithm can reduce the multicollinearity problem by reducing the possibility of the highly correlated features being selected (Liaw and Wiener 2002). Although random forests cannot completely resolve the multicollinearity issue, the contributions of the highly correlated features are retained, and the rank of the most important features will not be much affected (Jin *et al.* 2020). The hyper-parameters tuned were maximum depth limit ϵ[10,15], minimum samples split ϵ[2,4] and minimum samples leaf ϵ[1,2,4]. The number of trees was fixed at 200.

Baseline

We used a baseline method as the benchmark to evaluate the performance of the proposed models. In our baseline method, we assume that the load at any hour is equal to the load at the same hour a week before. This method is applicable to both weekdays and weekends, which is desirable for a simple and intuitive baseline model.

Features

Based on the literature review, we considered three types of features. Considering the nature of electricity use in commercial buildings, day of week (DoW) and hour of day (HoD) are used to reflect the cyclic patterns of electricity use. In order to capture the proximity and ordering among DoW and HoD, we converted them into cyclic variables using a linear combination of the two sinusoidal terms (Ashouri *et al.* 2019). The second type of feature is autoregressive parameters. We used electricity use of the previous time step and electricity use of the same hour previous week as features. The last type of feature is weather data. We collected historical outdoor temperature data from the nearest weather station (Government of Canada 2021). In order to offer a more practical approach, instead of building-specific solar data (e.g. accounting for orientation of surfaces) we obtained the solar irradiance data from a publicly available source (NASA research). The weather data used in this study are true historical values, which adds more uncertainty to the linear and non-linear models in this study.

Model development

As mentioned in the Introduction section, we designed a model development approach with a short training requirement in order to achieve a more flexible framework for varying operation conditions during the post-pandemic return to work scenarios. The framework is shown in Figure 2.

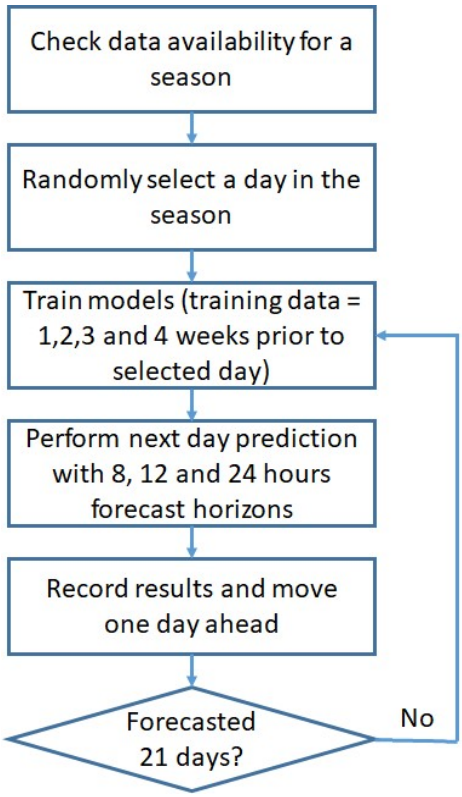

Figure 2: Model development steps for each season

In this approach, we divided the historical data into seasons. The summer season was defined from the first of June to the end of August. The winter season was defined from the first of December until the end of February. The periods of time between summer and winter seasons are labelled as shoulder seasons. In each season, a testing day is randomly selected based on the data availability in that season. The models are trained using the data from the hourly historical data prior to the testing day. To examine the sensitivity of prediction to the training data length, we tested four distinct training data lengths, i.e. one week (168 hours) to four weeks (672 hours). For each test time, we trained the models and forecasted the next-day electricity use and repeated this process for the next 21 days (i.e. 21 testing days). Therefore, 8 new models (4 linear and 4 non-linear models) are trained for each day. All test days are within the same season. During each testing day, we considered three forecast

horizons of 8, 12 and 24 hours. In the case of an 8-hour forecast horizon, the tests started at 01:00, 9:00 and 17:00. The second forecast horizon is 12 hours (starting at 01:00 and 13:00). During each prediction horizon, we used a recursive prediction method in which the forecasted electricity load is fed forward to the next timestep. At the end of each forecast horizon, the autoregressive term from the previous hour was reset to the real measured value and used for the first prediction of the next forecast horizon. This approach limits the error propagation to the length of the forecast horizon.

We use 4-fold cross validation with grid-search to tune the hyper-parameters of these models. The root mean square error was used to find the best performing combination of hyper-parameters for each model. We calculated the root mean square error and accuracy of forecasts, as follows:

Forecast horizon

The results of load forecasting models for the last season (winter of 2020) were used to represent the effects of forecast horizon on the model performance. Figure 3 shows the distribution of model accuracy for different forecast horizons.

Season and length of training data

The performance of all models for each season and training data length is shown in Figure 4. In all seasons, the accuracy of elastic-net and random forest models does not improve with training data lengths of more than 2 weeks. Note that the performance of the baseline method is not a function of training data length. However, we have shown it in all cases for easier comparison with other models. In all cases, the random forest model outperforms the elastic-net and baseline models. One potential explanation is that, unlike

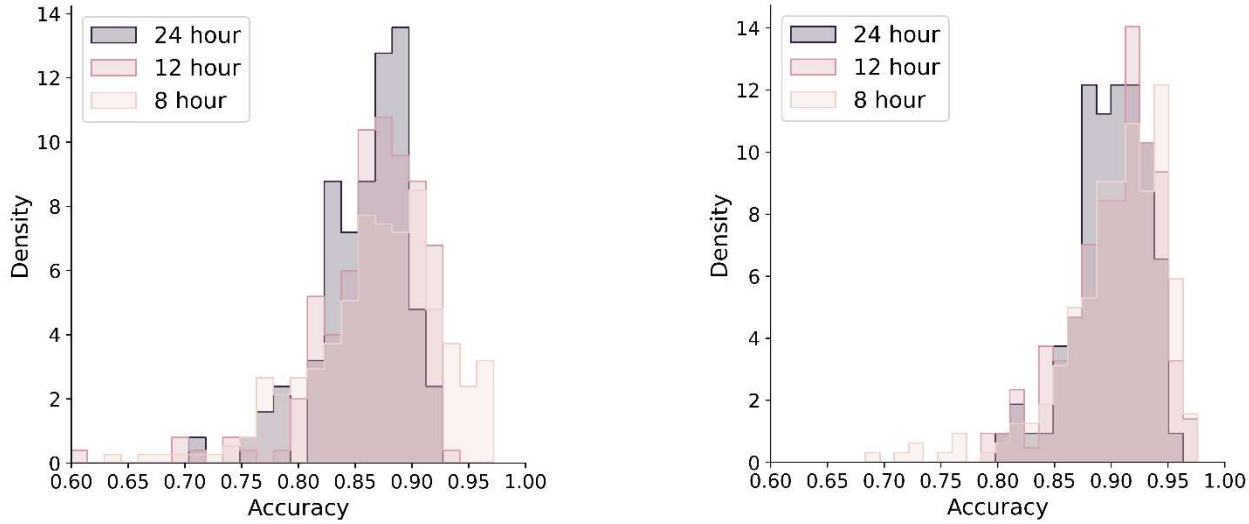

Figure 3: Distributions of accuracy for various forecast horizons. Elastic-net (left) and random forest (right)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(L_i - \hat{L}_i)^2}{N}} \quad (2)$$

$$CVRMSE = \frac{RMSE}{L_{mean}} \quad (3)$$

$$Accuracy = 1 - CVRMSE \quad (4)$$

We implemented both models in *scikit-learn* (Pedregosa *et al.* 2011).

Results

The results of the model performance evaluations, including the effects of training data length and forecast horizon, are presented in this section.

the elastic-net model, a random forest model is able to capture the nonlinear relation between the electrical load and influencing factors.

From Figure 4, it is evident that the performance of models has not been affected by the effects of pandemic restrictions. Note that we did not test the models' performance during the spring of 2020 due to the data gap in April (Figure 1). This figure shows a noticeable decrease in models' performance during the shoulder seasons (September to November). Although the median accuracy of 24 hour-ahead prediction of random forest model with 2 weeks of training data remains higher than 87% for both shoulder seasons; there are outliers in the accuracy with values as low as 60%.

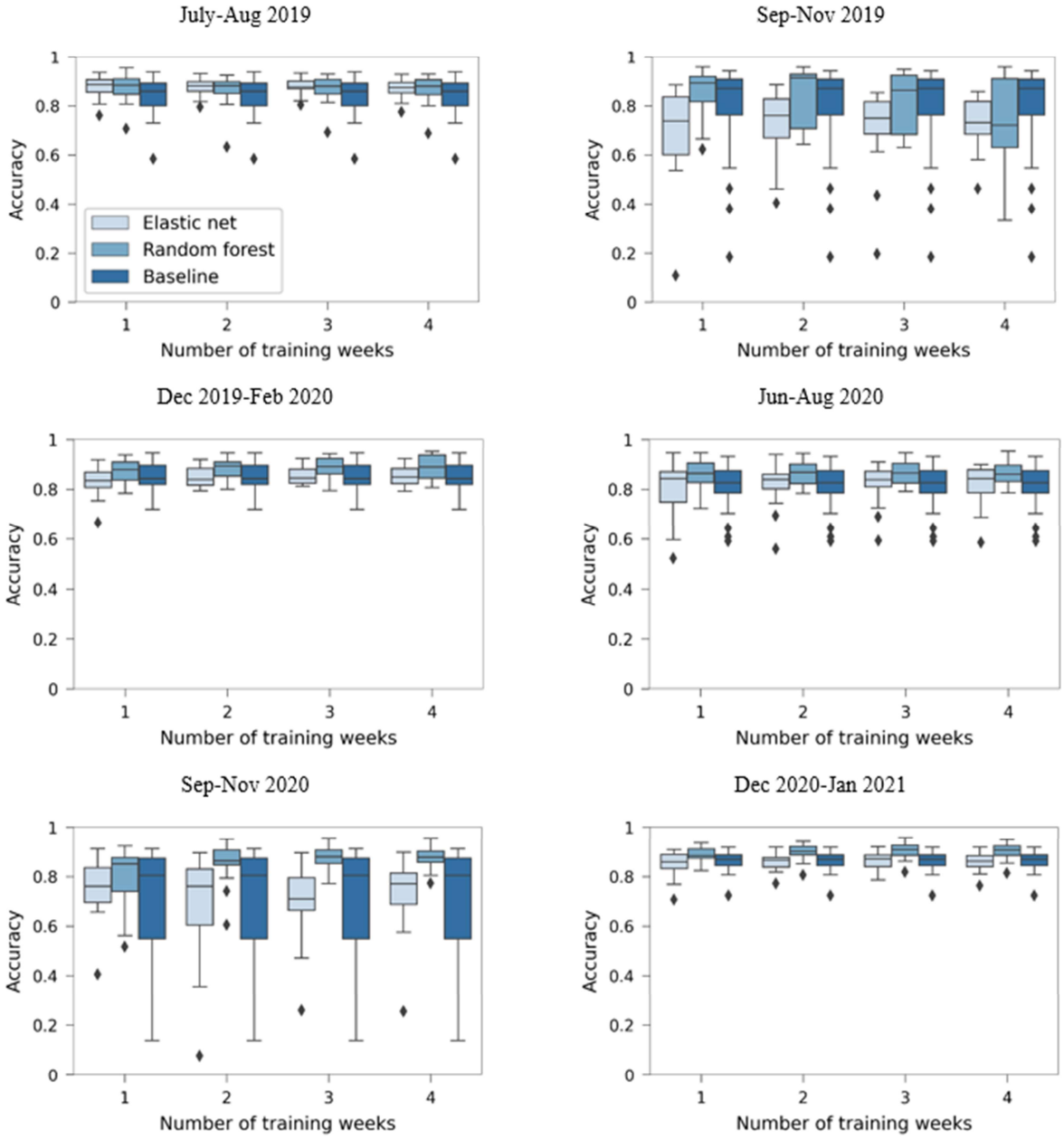

Figure 4: Distributions of accuracy for various forecast horizons. Elastic-net (left) and random forest (right)

Figure 5 shows an example of 24 hour-ahead forecasting of electricity load during the worst performing season (Sep-Nov 2019). During this period of time, it is evident that both linear and baseline models fail to provide an accurate forecast while the random forest model tends to be more accurate. However, this model significantly under-predicts the load on Sep 30 (the first day shown in the figure).

In order to better understand the behaviour of the random forest model, Figure 6 shows two weeks of real load data, and predicted load by the random forest model for the second week using the first week as training data. The outdoor temperature is also shown on the right vertical axis. This figure suggests that higher electricity use during Sep 28th, 29th and 30th is mainly associated with elevated outdoor temperature. However, the maximum outdoor temperature on Sep 30 is similar to that on Sep 24th, while the electricity load is about 35% higher on Sep 30. One possible explanation is the relatively higher outdoor temperatures during the prior night on Sep 30 which leads to a high morning cooling demand compared to Sep 24th.

These observations suggest that using history terms as additional features may improve the performance of models. We intend to test this idea in our future work.

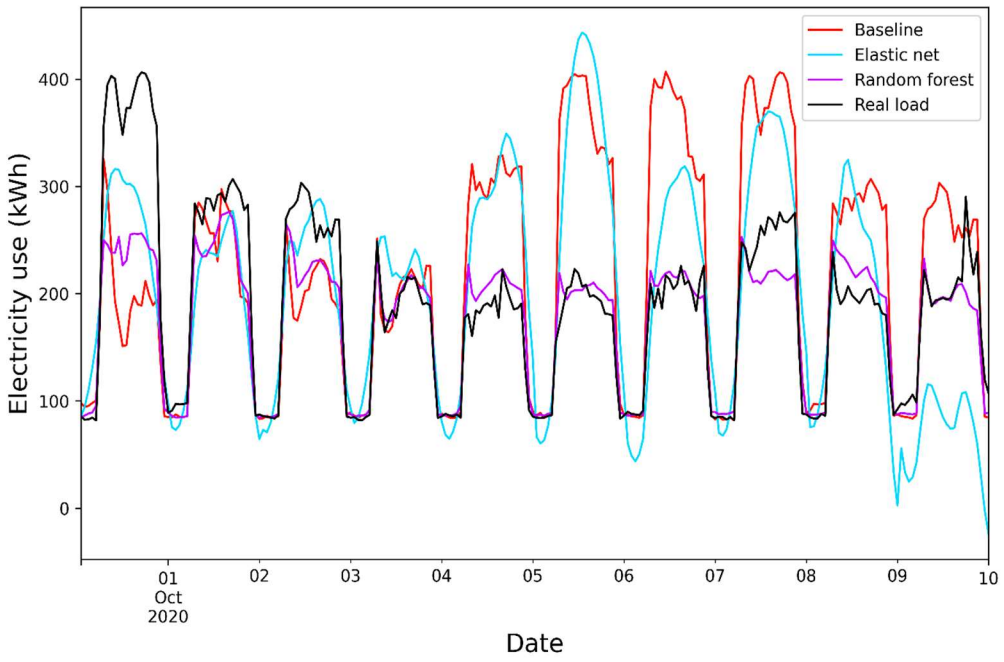

Figure 6: Variation of real and forecasted load

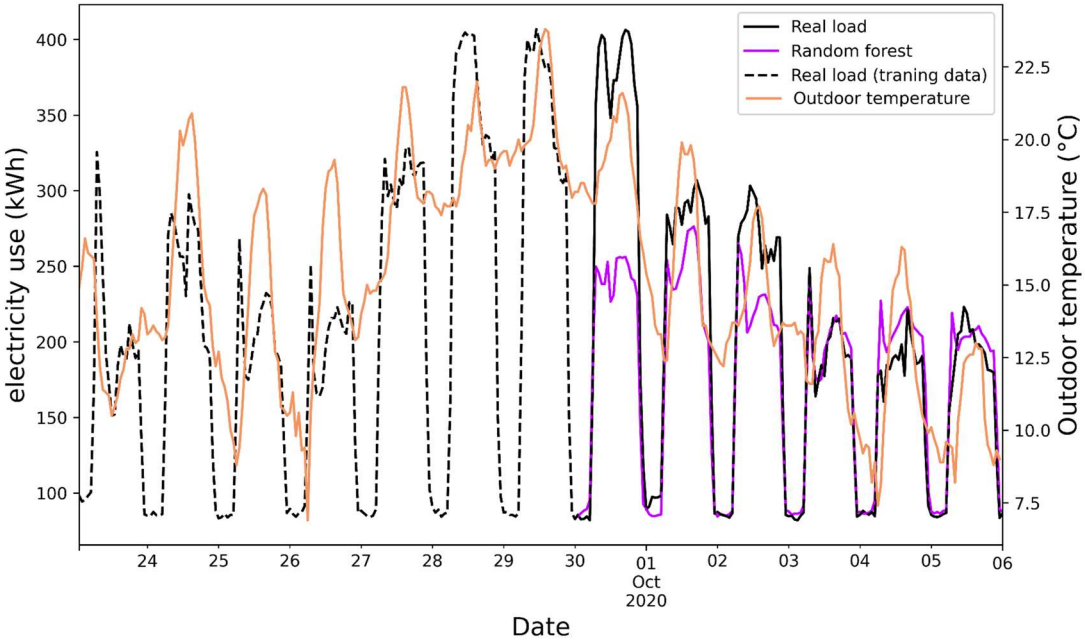

Figure 5: Variation of real and forecasted load and outdoor temperature

Conclusions

In this work, we tested two models, a linear (elastic-net) and a non-linear (random forest), for short-term forecasting of the electricity load in an office building. We also used a baseline method as a benchmark. We adopted a sliding-window prediction method with various forecasting horizons. For each day of the forecast, eight models were developed using relatively short periods of training data varying from 1 to 4 weeks. We used three different forecast horizons, namely 8, 12 and 24 hours for day-ahead predictions. Each model was tuned using the grid-search method with 4-fold cross validation. The results indicate that the accuracy of the linear and non-linear models does not improve with training data lengths of more than 2 weeks. Although the highest accuracies are achieved when an 8-hour forecast horizon is used, in the case of the 24-hour horizon the median accuracy is higher than 87%, which might be considered acceptable for most applications. Moreover, the non-linear model outperforms the linear and baseline models. All models show a lower performance during the shoulder seasons. We discussed the potential mechanisms behind this observation.

There are a number of possible ways to improve the performance of both models. In particular, adding recent historical information may improve the models' performance during shoulder seasons. We intend to test a variety of historical features in our future work.

Acknowledgements

This research was supported by the Office of Energy Research and Development (OERD) of Natural Resources Canada. The authors are grateful to Troy Shanks and Emera Inc. for their support.

References

Agrawal, R.K., Muchahary, F., and Tripathi, M.M., 2018. Long term load forecasting with hourly predictions based on long-short-term-memory networks. *2018 IEEE Texas Power and Energy Conference, TPEC 2018*, 2018-February, 1–6.

Ahmad, T., Chen, H., Guo, Y., and Wang, J., 2018. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. *Energy and Buildings*, 165, 301–320.

Amjady, N. and Daraeepour, A., 2011. Midterm demand prediction of electrical power systems using a new hybrid forecast technique. *IEEE Transactions on Power Systems*, 26 (2), 755–765.

Ashouri, A., Shi, Z., and Gunay, H.B., 2019. Data-driven short-term load forecasting for heating and cooling demand in office buildings. *Journal of Physics: Conference Series*, 1343 (1).

Bourdeau, M., Zhai, X. qiang, Nefzaoui, E., Guo, X., and Chatellier, P., 2019. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustainable Cities and Society*, 48 (April), 101533.

Chen, Y.T., Piedad, E., and Kuo, C.C., 2019. Energy consumption load forecasting using a level-based random forest classifier. *Symmetry*, 11 (8), 1–9.

Collotta, M. and Pau, G., 2017. An Innovative Approach for Forecasting of Energy Requirements to Improve a Smart Home Management System Based on BLE. *IEEE Transactions on Green Communications and Networking*, 1 (1), 112–120.

Cozzi, L. and Gould, T., 2021. World Energy Outlook 2021, 1–386.

Government of Canada, 2021. Historical Climate Data [online]. Available from: https://climate.weather.gc.ca [Accessed 11 Mar 2021].

Guan, C., Luh, P.B., Michel, L.D., and Chi, Z., 2013. Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation. *IEEE Transactions on Power Systems*, 28 (4), 3806–3817.

Jacob, M., Neves, C., and Vukadinović Greetham, D., 2020. Short Term Load Forecasting. *In*: *Forecasting and Assessing Risk of Individual Electricity Peaks*. Cham: Springer International Publishing, 15–37.

Jin, Z., Shang, J., Zhu, Q., Ling, C., Xie, W., and Qiang, B., 2020. RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12343 LNCS, 503–515.

Jones, T.C., Auslander, D.M., Taneja, J., Trager, J., and Sankur, M., 2012. Improved Methods to Load Prediction in Commercial Buildings Introduction to Current Methods. *Proceedings of the …*, 173–184.

Kim, J., Moon, J., Hwang, E., and Kang, P., 2019. Recurrent inception convolution neural network for multi short-term load forecasting. *Energy and Buildings*, 194 (2019), 328–341.

Koukaras, P., Bezas, N., Gkaidatzis, P., Ioannidis, D., Tzovaras, D., and Tjortjis, C., 2021. Introducing a novel approach in one-step ahead energy load forecasting. *Sustainable Computing: Informatics and Systems*, 32 (November), 100616.

Liaw, A. and Wiener, M., 2002. Classification and Regression by randomForest. *R News*, 2 (3), 18–22.

Moon, J., Kim, Y., Son, M., and Hwang, E., 2018. Hybrid short-term load forecasting scheme using random forest and multilayer perceptron. *Energies*, 11 (12), 1–20.

Muzaffar, S. and Afshari, A., 2019. Short-term load forecasts using LSTM networks. *Energy Procedia*, 158, 2922–2927.

NASA research, n.d. The POWER Project [online]. Available from: https://power.larc.nasa.gov/.

Pedregosa, F., Varoquaux, G., and Gramfort, A., 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12 (9), 2825–2830.

Raza, M.Q. and Khosravi, A., 2015. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews*, 50, 1352–1372.

Wang, Z. and Srinivasan, R.S., 2017. A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews*, 75 (October 2016), 796–808.

Yildiz, B., Bilbao, J.I., and Sproul, A.B., 2017. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 73 (December 2016), 1104–1122.