Data-driven Approaches for Predictive Thermal Modelling in Small and Medium Commercial Buildings

Brent Huchuk[1], Farid Bahiraei[1]
[1]National Research Council Canada, Ottawa, Canada

Abstract

Small and medium buildings comprise 75% of Canada's commercial and industrial building stock. However, small and medium buildings are currently underserved by energy efficiency tools because of their dispersion and relatively low payback potential. In this work, we leveraged the connected thermostat data from 16 small and medium commercial buildings to develop multi-zone data-driven models to be used for predictive control and planning applications. Specifically, we sought to understand which of the candidate data-driven models (i.e., grey and black box models) could most accurately predict the indoor temperatures across a diverse sample of buildings using only typical data collected from a connected thermostat. Our results show that data-driven models, specifically black box models with appropriately selected parameters, had the best predictive performance.

Introduction

Small and medium commercial buildings (SMBs) make up a significant portion of Canada's building stock and generally have higher energy use intensities than other building sizes (Natural Resources Canada (2013, 2018)). Over half of SMB's energy usage is related to the space heating and cooling of the building (Natural Resources Canada (2016)). The heating, ventilation, and air conditioning (HVAC) systems within SMBs are often composed of multiple independently controlled components that each serve their own thermal zone. Despite the energy ramifications and complicated equipment configurations, operation of SMB's HVAC is generally left to reactive and uncoordinated control systems. To better operate these buildings, control systems and operators need to consider predictive actions based on the predicted thermal response. Practically, this requires the ability to quickly and easily train customized thermal models for each building that accounts for coupled effects between the zones with their systems. It has been estimated that 70% of the complication and effort of implementing predictive controls is developing accurate models (Henze (2013)).

Generally, thermal models can be classified as white box, black box, or grey box. Practically, white box models are detailed representations built from first-principle relationships. They are not easy to implement outside simple building configurations and they require detailed knowledge of the equipment and building construction (Prívara et al. (2013)). In comparison, data-driven methods, referred to as grey or black box (ASHRAE (2017)), are based on simplified or no physical relationships and are more often used in predictive control applications (Prívara et al. (2013)).

Previous investigations have compared and contrasted grey and black box models using connected thermostat (CT) data (Huchuk et al. (2020); Baasch et al. (2021)). These studies found black box models to outperform the alternative grey box models, but did not consider the multi-zone/equipment configurations seen in SMBs. Alternatively, Arroyo et al. (2020) did consider the multi-zone case of SMBs but only tested various grey box model configurations and did not consider exclusively the data available with standard CTs. As such, questions remain on how different model types perform in the SMB domain.

With the advent of low-cost sensing and control through the installation of CTs in SMBs, there is now an accessible source of operational data to train and evaluate data-driven multi-zone thermal models. Moreover, installation of CTs potentially enable SMB operators to engage in more energy efficiency programs and management (Rovito et al. (2014)). Unfortunately, given the relative recency in the CT hardware upgrades, and the overall lack of investigation on SMBs, little research has been conducted on which model types and which features provide the best combinations for multi-hour temperature predictions in these buildings. To this end, this paper uses a unique dataset from a small portfo-

lio of SMBs to evaluate data-driven thermal models for this purpose.

Methodology

In this section, we discuss the critical elements of our investigation. First, we discuss our unique dataset from a collection of SMBs. Then we discuss our candidate thermal models and the required model training and evaluation.

Data

We conducted our analysis with historical data logged by installed CTs within a sample of 16 SMBs located in Ontario, Canada. The data was logged from January 2019 to August 2021. Each CT is connected to its own rooftop unit (RTU) or a series of electric baseboard heaters. It is not possible to differentiate between the two equipment configurations from the data collected by each CT. The CTs and their equipment are not aware of other devices in the building and operate independently based on typical dead-band control.

Each CT records operational data at a five-minute interval. This data includes: indoor air temperature, fan runtime, setpoint temperatures, and the amount of runtime for each stage of equipment attached (e.g., first and second stages of heating or cooling equipment). We downsampled the data to a 30-minute resolution to help smooth the data and better align with potential planning applications based on the intervals as the CTs' programmed setpoint schedules. For temperature values, we calculated the mean measurement over the 30-minute interval. Meanwhile, for equipment runtime, we converted the runtime to a ratio over the interval. For example, if 10 minutes of stage 1 heating runtime was measured over the 30-minute window, we would record that as 1/3.

The floor area and number of CTs vary in our sample of buildings. We treated each CT as part of its own independent thermal zone within the building. This thermal zone assumption was validated based on discussions with the building operators. Figure 1 shows the distribution in the floor area of the sample of buildings. The average building was 5897 ft^2 (548 m^2). Over 60% of the buildings had four CT devices, with the remaining few having either three or five CTs.

In addition to the thermostat data, we collected weather data for each building from the Government of Canada (2021). Specifically, we collected hourly outdoor air temperature from the nearest weather station to each building based on provided latitude and longitude.

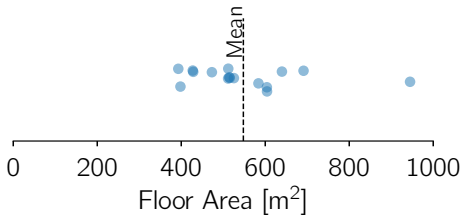


Figure 1: Distribution in floor areas of the buildings.

Models

We trained a collection of models to predict the indoor air temperature at each CT at future timesteps (t). The models we selected included a baseline, a physics-based grey box model and two black box models (i.e., linear and random forest regressions). We restricted the candidate grey box and black box models based on previous investigations (Huchuk et al. (2020); Baasch et al. (2021)). All of the models predicted an output vector ($\hat{Y}$) that was of dimension M (i.e., the number of CTs/zones in the building). We provide details for each model in the following subsections.

Baseline model

We trained a simple model to provide a baseline to the other more complex models. Specifically, the baseline model took the last measured temperature at that thermostat at hour 00:00, 06:00, 12:00 and 18:00 and predicted that single temperature until the next held value was reached. For example, all predictions from 00:30 to 06:00 for a single CT used the same value which was measured by that CT at 00:00. The baseline model would not be useful in most practical applications, but rather is intended to allow us to quantify the performance of the other models to naive assumptions.

Grey box model

Given the multiple (and variable) number of zones within the buildings, our grey box models were required to be customized for each building. We based our models on a single resistance (R) and capacitance (C) for each zone. More complicated networks have been used in previous studies (Baasch et al. (2021)). However, we decided to avoid including unobserved states as it may hinder the estimation of the parameters. The finite difference equation representing a single thermal

zone is:

$$T_{tstat,m,t+1} = T_{tstat,m,t} + \frac{\Delta t}{C_m}\left[\sum_j \frac{(T_{j,t} - T_{tstat,m,t})}{R_{m\leftrightarrow j}} + Q_{m,t}\right] \quad (1)$$

where t is the timestep, m is the index of a thermostat, Q is the energy load and j is the index of all connecting nodes to m.

For our application, we reformulated Equation 1 to Equation 2 to separate out our heating and cooling runtime fractions (q_{heat} and q_{cool}). The runtime fractions are multiplied by equipment capacity terms (β) which we left as free parameters to learn. We needed the model to learn the capacity terms because standard CT thermostats are unaware of the equipment sizing.

$$T_{tstat,m,t+1} = T_{tstat,m,t} + \frac{\Delta t}{C_m}\left[\sum_j \frac{(T_{j,t} - T_{tstat,m,t})}{R_{m\leftrightarrow j}} + \beta_c q_{cool,m,t} + \beta_h q_{heat,m,t}\right] \quad (2)$$

For buildings with multiple zones and CTs the indoor temperature nodes (T_{tstat}) were connected with an additional resistance similar to the method of Arroyo et al. (2020). Figure 2a shows the RC network for a building with two thermostats. Figure 2b expands the network for a building with M thermostats. In both figures, T_{out} is the outdoor temperature.

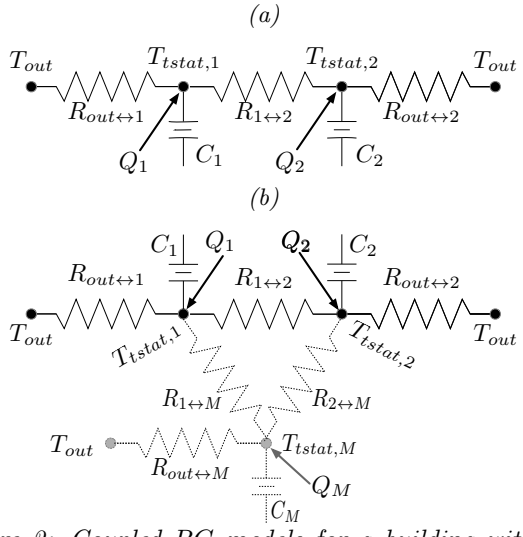


Figure 2: Coupled RC models for a building with (a) two and (b) M thermostats.

The model for each building was trained all at once (i.e., all R's, C's, and β's simultaneously). The models were trained using the `lmfit` (Newville et al. (2021)) package in python.

Linear regression

The first black box model we considered was a linear regression trained with an elastic net (Zou and Hastie (2005)). An elastic net regression is a typical linear regression with an adjusted minimization loss function ($\mathcal{L}$). The minimization function, shown in Equation 3, has both ℓ_1 and ℓ_2 regularization components balanced by the term α. The regularization of the weights (w) prevents the model from becoming overfit.

$$\mathcal{L} = \frac{1}{2\mathcal{N}}\left\|Y - Xw\right\|_2^2 + \alpha\ell_{1ratio}\left\|w\right\|_1 + 0.5\alpha(1-\ell_{1ratio})\left\|w\right\|_2^2 \quad (3)$$

where X is an array of the input features and $\mathcal{N}$ is the number of training examples. The input features of X were the same as those used in the training of the grey box model. Specifically, we did not consider the inclusion of additional recent-history terms.

We implemented the model in `Scikit-Learn` (Pedregosa et al. (2011)). During model training, we used five-fold cross validation to tune α and ℓ_{1ratio} values and selected the combination of values which resulted in the lowest mean squared error score.

Random forest regression

The other black box model we considered was the random forest regression (RF). As a tree-based method, the RF provides a non-linear regression. As an ensemble method, constructed of multiple decision trees, RFs are less susceptible to overfitting than a single decision tree (Breiman (2001)). Just as with the linear regression, we used five-fold cross validation to tune the hyper-parameters for each model trained and the best combination was selected based on mean squared error. Specifically for RF, we tuned: the number of estimators, max number of features, max depth, minimum number of samples for a split, minimum number of samples at each leaf, and whether to bootstrap or not. Similar to the linear regression, the model was implemented in `Scikit-Learn` (Pedregosa et al. (2011)).

Model training and evaluation

Our model training and evaluation method were designed to maximize the number of high-quality days of prediction for each building and evaluate across a wide range of day types and conditions. For each building, we identified *available testing days* based on the data

quality. We required 90% of the expected data (logged by the CT and weather stations) to be present for the previous 14 days. Similarly, we required 90% data existence on the testing day. Figure 3 shows the number of *available testing days* for each building. The average number of *available testing days* was 119 with the fewest number for a building being 58.

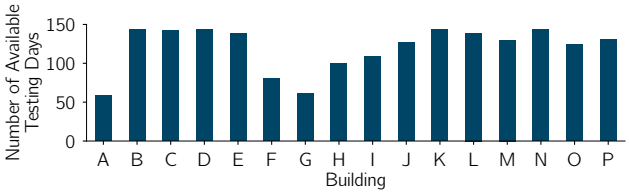


Figure 3: Number of available testing days for each building.

During our evaluation, any missing temperature measurements for an identified *available testing days* was filled using linear interpolation. No interpolation was done for potentially missing equipment runtime. For each *available testing day*, each of the four models (previously introduced) were trained using the 14 previous days of data for the building. To test the performance of the models on the testing day, we made four six-hour prediction sequences. The four sequences of 12 predictions (i.e., two predictions for each of the six hours) were conducted starting at 00:00, 06:00, 12:00, and 18:00. In each sequence of predictions, estimated temperatures ($\hat{T}_{tstat,m,t+1}$) were fed back into the model to be used by the next prediction (i.e., at t+1). This strategy leads to the accumulation of modelling errors over the six-hour prediction sequence. A similar method was used in the evaluation of single zone thermal models based on CT data in Huchuk et al. (2020).

We determined the performance of the predictions for a building's period of predictions (p) using the root mean squared error (RMSE) and the mean absolute error (MAE) as:

$$RMSE_p = \sqrt{\frac{\sum_{t=1}^{N}\sum_{m=1}^{M}\left(T_{tstat,m,t+1} - \hat{T}_{tstat,m,t+1}\right)^2}{N \cdot M}} \quad (4)$$

$$MAE_p = \frac{\sum_{t=1}^{N}\sum_{m=1}^{M}\left|T_{tstat,m,t+1} - \hat{T}_{tstat,m,t+1}\right|}{N \cdot M} \quad (5)$$

The error metrics were aggregated for all CTs (M) within the building at each timestep t and averaged over all points in that period N. For our prediction lengths, $N = 12$.

Results

In this section we present the outcomes of the training and testing of our models.

Model training

For each building's *available testing days*, each of the four models were trained. Figure 4a shows the distribution in RMSE for an *available testing day*'s prediction sequence separated by model and building. Figure 4b shows the same distributions but for MAE for the same breakdowns by buildings and models as in Figure 4a. We observe that the baseline is, unsurprisingly, the worst performing model based on both metrics. Generally, we see the two black box models (i.e., the linear and random forest regression) have better performance than the grey box model.

Model testing

Once trained, each model was used for a whole day of predictions as described in the "Model training and evaluation" section. Figure 5a and Figure 5b show the distributions in the RMSE and MAE for the testing days. Similar to the training results, we observe the linear and random forest models being the best performers across the portfolio of buildings with the median performance around 0.4°C. The linear regression model ultimately has the lowest median error across over two-thirds of the test buildings.

Demonstration of predictions

We expected the predictions from the various models to differ given the observed overall performance differences in testing (Figure 5a and Figure 5b). Figure 6 shows the prediction for a single building (where M=4) for a single *available testing day*. We note that at the start of a new prediction sequence (i.e., at 00:00, 06:00, 12:00, and 18:00) there is an intentional break in the prediction sequence. We observe in this example case that the grey box model tends to accumulate errors more drastically and the coupled relationship of the temperatures drives oscillations – visible in the top two thermostat plots. The linear model, through the regularization of its parameters, and the random forest through its tuning and inherent bounds on values, does not suffer the same parameter explosion resulting in more robust predictors.

Temporal effect to model errors

Figure 7 shows the median performance for each model and building based on the prediction window (i.e., periods starting at 00:00, 06:00, 12:00 and 18:00). The baseline and grey box models appear to have the most varied performance based on when in the day the predictions were being made. The highest errors appear in the 2^{nd} and 4^{th} prediction window (i.e., periods starting at 06:00 and 18:00). The linear and random forest mod-

els are less sensitive to the prediction window. However, the most variation in the performance of these models happens during the 2^{nd} and 3^{rd} windows. Considering the business hours of these buildings, we expect these periods would be the most transient for the building and could lead to higher median errors.

Discussion

The results above are an informative exercise for future researchers investigating the training of multi-zone thermal models from CT data. Under the same informative desire, we highlight a couple of key points of discussion.

Challenges with the grey box model

Similar to the results observed in Huchuk et al. (2020), tuning the parameters of the grey box model was challenging when applied to any generic building and the parameters were not easily validated. Given that the number of feasible solutions can be vast, grey box models should be verified against reasonable value ranges to ensure reasonable physical meaning (Brastein et al. (2018); Dimitriou et al. (2015)) but that is difficult to accomplish on a large sample of buildings.

While the grey box models may be outperformed when predicting indoor air temperatures by the black box models, the grey box model does still allow for physically-meaningful parameters to be learned and compared if the model can be properly validated. This may provide an alternative use case for the model class if potentially linked to high-quality data related to building audits, building envelope characteristics, or energy usage.

Application of the models

From a practical implication perspective, the relative error achieved by our best performing model (i.e., linear model) may not be suitable for the active and real-time control of the CT for three main reasons. First, we only evaluated the model at a 30-minute resolution while the CT needs to operate at near real time. Second, the median accuracy of the thermal models, is nearly the same as a size of a control dead-band on these CTs (0.5°C). This uncertainty could often lead to incorrect control decisions being made. Finally, the conversion of the runtime equipment ratios (q) need to be translated into actionable control sequences which needs to be further explored. More appropriately, the black box models could be used by facility managers to understand the temperature ramifications of planned

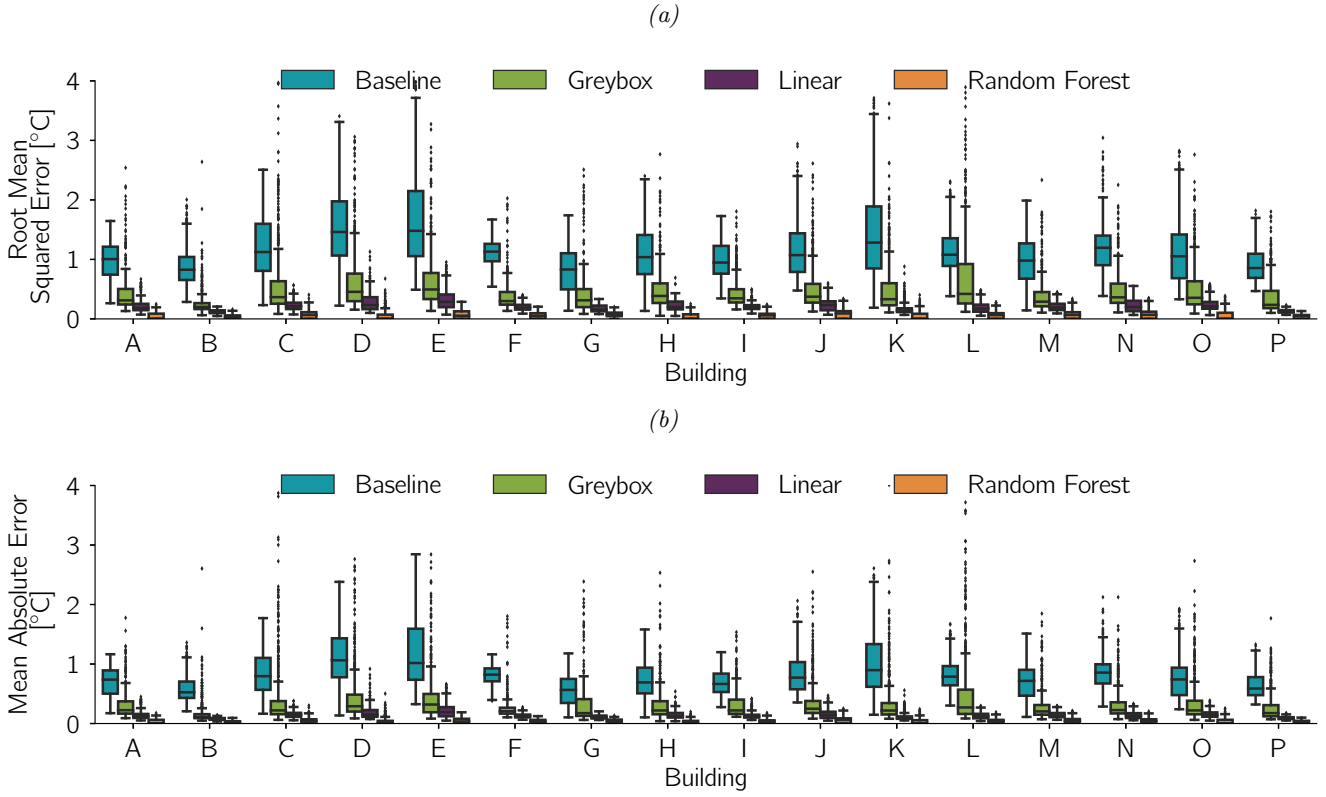
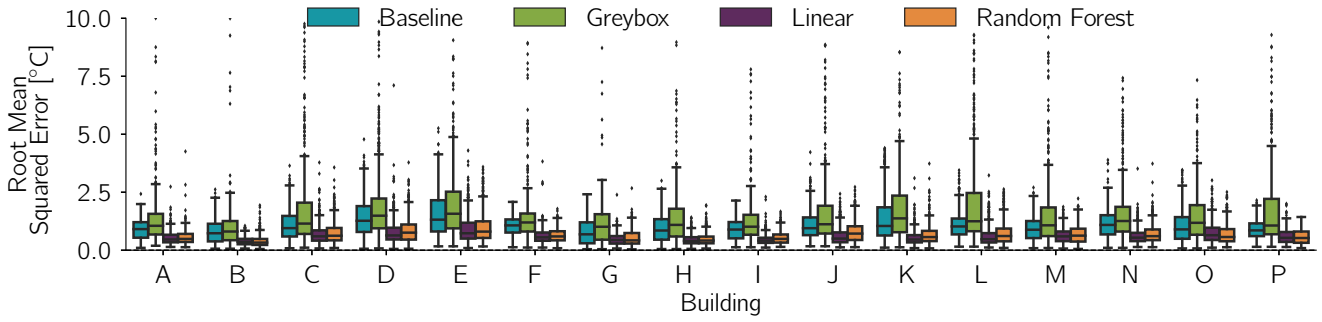


Figure 4: Distribution in (a) root mean squared error and (b) mean absolute error values for each building and each model type for all training data.

(a)



(b)

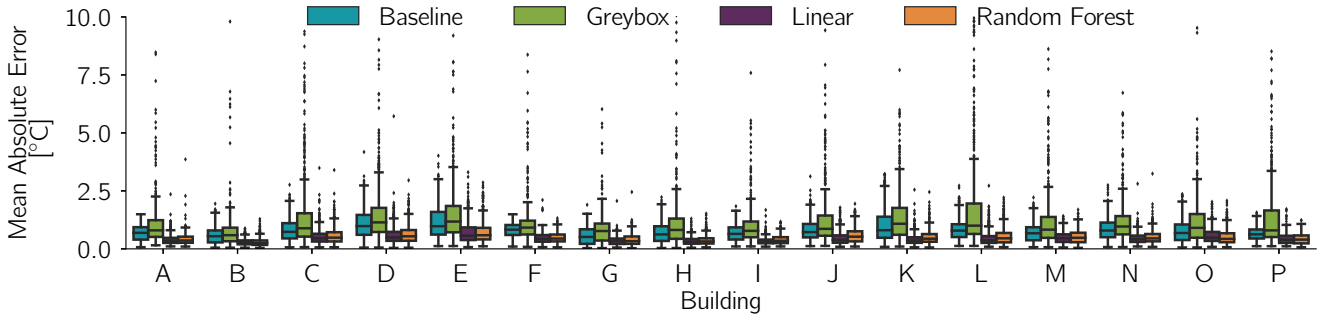


Figure 5: Distribution in (a) root mean squared error and (b) mean absolute error values for each building and each model type for all testing data.

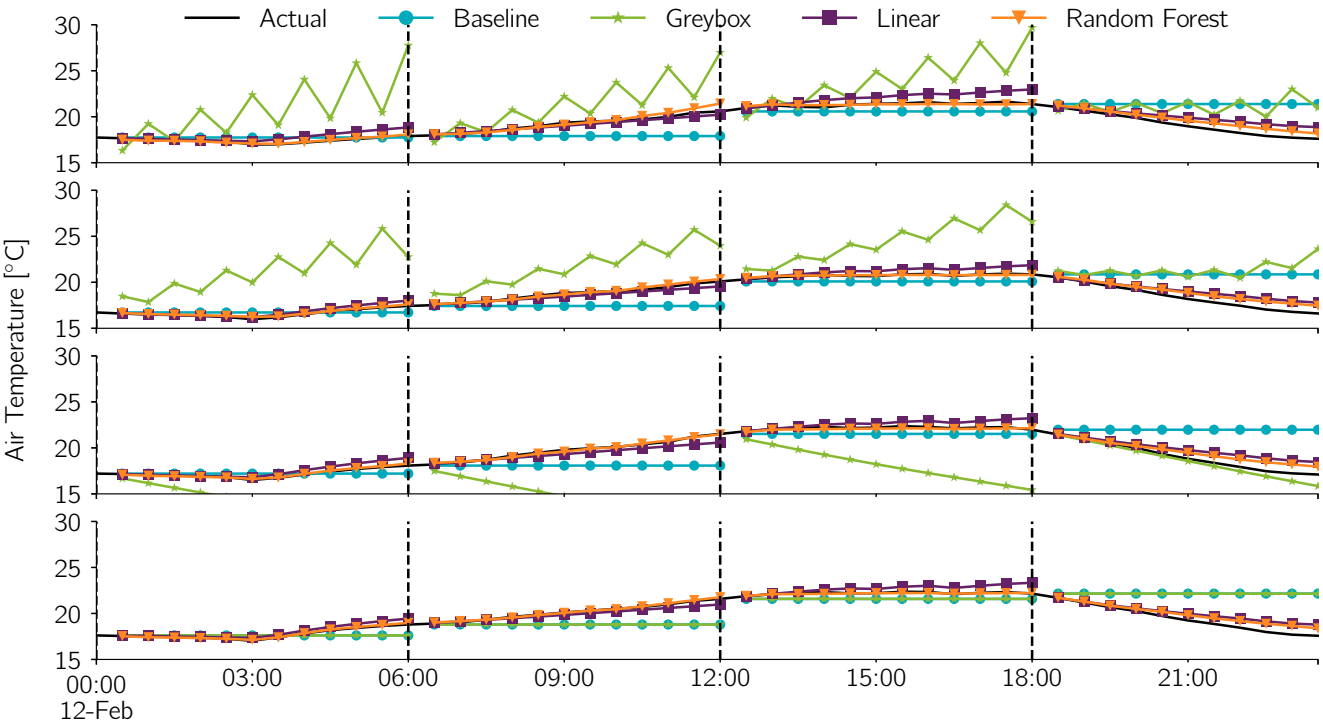


Figure 6: Demonstration of predictions for an available testing day for a random building with four thermostats.

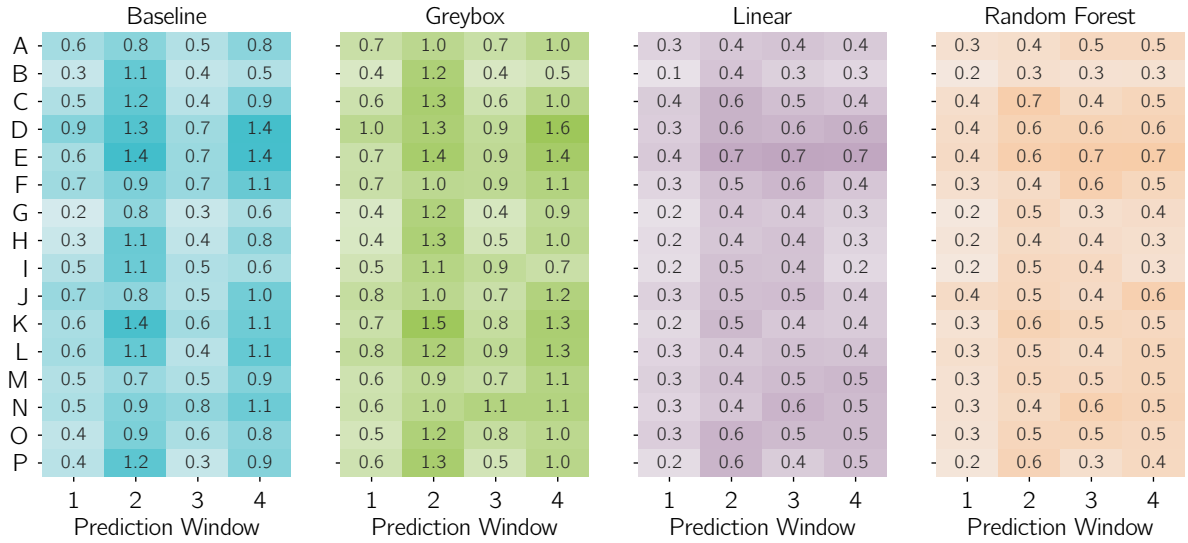


Figure 7: Median root mean squared error scores for each model and building based on the prediction window.

events such as a pre-cooling/heating events or optimal starting of equipment.

Conclusion

To properly manage the HVAC systems of small and medium commercial buildings, customized predictive models need to be developed for each building. Realistically these models need to be based only on the measured data from low-cost and easily deployed connected thermostats. Using a unique sample of 16 small and medium commercial buildings which had undergone a control retrofit, we compared the performance of candidate data-driven thermal models on multi-zone and multi-hour predictions.

In the training of the multiple thermal zones simultaneously, we found a linear model which had been regularized using an elastic net model outperformed other black box and grey box alternatives. While the linear model does not easily allow for additional insights to be generated (i.e., in comparison to the physically meaningful values of a grey box model), the linear form makes it ideal for easy utilization across a wide variety of buildings and as part of larger optimizations.

Acknowledgment

This research was supported by the Office of Energy Research and Development (OERD) of Natural Resources Canada. We also thank Jared Goodman, Rajendran Avadaippan, Andre Roman, Ruth Weiner, and John Sobrinho for their support.

References

Arroyo, J., F. Spiessens, and L. Helsen (2020). Identification of multi-zone grey-box building models for use in model predictive control. *1493*(May).

ASHRAE (2017). *Handbook of Fundamentals: American Society of Heating, Refrigerating Air-Conditioning Engineers*.

Baasch, G., P. Westermann, and R. Evins (2021, June). Identifying whole-building heat loss coefficient from heterogeneous sensor data: An empirical survey of gray and black box approaches. *Energy and Buildings 241*, 110889.

Brastein, O. M., D. W. Perera, C. Pfeifer, and N. O. Skeie (2018, Jun). Parameter estimation for grey-box models of building thermal behaviour. *Energy and Buildings 169*, 58–68.

Breiman, L. (2001). Random Forests. *Machine Learning 45*(1), 5–32.

Dimitriou, V., S. K. Firth, T. M. Hassan, T. Kane, and M. Coleman (2015). Data-driven simple thermal models: The importance of the parameter estimates. *Energy Procedia 78*, 2614–2619.

Government of Canada (2021). Historical Climate Data.

Henze, G. P. (2013). Model predictive control for buildings: a quantum leap? *Journal of Building Performance Simulation 6*(3), 157–158.

Huchuk, B., S. Sanner, and W. O'Brien (2020). Evaluation of data-driven thermal models for multi-hour

predictions using residential smart thermostat data. *Journal of Building Performance Simulation*.

Natural Resources Canada (2013). *Survey of commercial and institutional energy use: buildings 2009 : summary report*.

Natural Resources Canada (2016). Energy Efficiency Trends in Canada 1990 to 2013.

Natural Resources Canada (2018). National Energy Use Database.

Newville, M., R. Otten, A. Nelson, A. Ingargiola, T. Stensitzki, D. Allan, A. Fox, F. Carter, Michał, R. Osborn, D. Pustakhod, Lneuhaus, S. Weigand, Glenn, C. Deil, Mark, A. L. R. Hansen, G. Pasquevich, L. Foks, N. Zobrist, O. Frost, A. Beelen, Stuermer, Azelcer, A. Hannum, A. Polloreno, J. H. Nielsen, S. Caldwell, A. Almarza, and A. Persaud (2021, Oct). lmfit/lmfit-py: 1.0.3.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research 12*, 2825–2830.

Prívara, S., J. Cigler, Z. Váňa, F. Oldewurtel, C. Sagerschnig, and E. Žáčeková (2013). Building modeling as a crucial part for building predictive control. *Energy and Buildings 56*, 8–22.

Rovito, M., G. Subramony, L. Duffy, and P. Savio (2014). Advanced Thermostats for Small- to Medium-Sized Commercial Buildings. In *2014 ACEEE Summer Study on Energy Efficiency in Buildings*.

Zou, H. and T. Hastie (2005, Apr). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67*(2), 301–320.