Data-Driven Modelling of Pressurized Corridor Ventilation System Performance in a Multi-Unit Residential Building

Helen Stopps[1], Cara H. Lozinsky[2], Marianne F. Touchie[3]
[1]Toronto Metropolitan University, Toronto, Canada
[2]Carleton University, Ottawa, Canada
[3]University of Toronto, Toronto, Canada

Abstract

The performance of pressurized corridor (PC) ventilation systems in multi-unit residential buildings (MURBs) can be highly influenced by occupant behaviours (e.g., window/exhaust fan operation), wind; or indoor-outdoor temperature differentials. We developed an XGBoost regression model to assess the impacts of these factors on system operation; identify when depressurization events may occur (which may influence ventilation and/or contaminant transfer); and understand how to improve system operation. The model was built using data collected from a 17-storey MURB, including differential pressures, window/door operation, corridor supply air flow rates, and interior/exterior temperature and relative humidity data. This paper describes the model development and evaluates the model as a potential diagnostic tool and testing environment to improve ventilation system performance.

Introduction

Pressurized corridor (PC) ventilation systems have been used extensively to supply make-up air in existing multi-unit residential buildings (MURBs). In these systems, ventilation air is supplied directly to the common corridor. In theory, the supply air flow rate is sufficient to positively pressurize the corridor, relative to the dwelling unit, which drives the make-up air into the dwelling units via intentional undercuts at the entry doors. The positive corridor pressurization also helps to control inter-dwelling unit odour/contaminant transfer, which can occur via the corridor. In practice, these systems are sensitive to interference from occupant behaviours, such as window and exhaust fan operation (Maxwell, Berger, & Zuluaga, 2014; Proskiw & Phillips, 2008; Ricketts, 2014; Sheltair Scientific Ltd., 1998); wind and stack effect (Air Solutions Inc., 2005; Fine & Touchie, 2021; Proskiw & Phillips, 2008; Ricketts, 2014); and interior leakage pathways (Enermodal Engineering Limited, 2002, 2003; Maxwell et al., 2014). This can cause uneven distribution of make-up air to the dwelling units and/or corridor depressurization, resulting in poor indoor air quality outcomes.

Previous studies assessing PC system performance have typically relied on simultaneous measurements of corridor supply air flow rates, inter-zonal differential pressures, occupant behaviours (e.g., window and/or exhaust fan operation), and/or weather variables over short or long periods (Air Solutions Inc., 2005; Lozinsky, Stopps, & Touchie, 2023; Maxwell et al., 2014; Ricketts, 2014). In these studies, typical corridor-to-dwelling unit pressures ranged from -5Pa to +30Pa, with lower pressure differentials found on floors near the bottom of the building (under this sign convention, a negative corridor-to-dwelling unit differential pressure indicates that the corridor is depressurized relative to the dwelling unit, i.e., the PC system is not acting as intended). In-unit exhaust fan operation depressurizes dwelling units relative to the corridor, resulting in increased corridor-to-dwelling unit differential pressures; however, the magnitude of the effect is highly variable, with previous studies reporting changes in corridor-to-dwelling unit differential pressures ranging from +1Pa to +75Pa, depending on the number of fans and their respective flow rates (Air Solutions Inc., 2005; Enermodal Engineering Limited, 2003; Ricketts, 2014; Sheltair Scientific Ltd., 1998). Estimating effects on ventilation system performance from simultaneous measurements of differential pressure, occupant behaviours, and/or weather has its limitations as an analysis technique, as it is often difficult to fully isolate the different factors. To our knowledge, no one has implemented a model-based approach to evaluate the effects of external factors – including weather and occupant behaviours – on PC system performance.

This paper describes the development of an extreme gradient boosting (XGBoost) regression model for evaluating the performance of a PC system in a contemporary MURB. The model was built using data collected from a prior field study (Lozinsky et al., 2023).

Methodology

The following section describes (1) the case study building; (2) data collection procedures for the dwelling unit data used to build the model; and (3) model development and assessment, including performance metrics, feature selection, hyperparameter tuning, and feature importance assessment.

Case Study Building Description and Data Collection

The case study building is a 17-storey, reinforced concrete MURB located in Toronto, ON, constructed in 2017. The building is a double-loaded corridor configuration running north-south, with most dwelling units facing either due east or due west. There are 343 dwelling units in the building, ranging from 35 m^2 to 85 m^2 in floor area (377 ft^2 to 915 ft^2).

Ventilation air is delivered to the corridors via two supply grilles located at each end of the corridor. The make-up air units (MAUs) are operated on a set-back schedule, with reduced air flow rates between 9:00am and 11:00am; and 1:00pm and 4:00pm; and 9:00pm and 7:00am. Average supply air flow rates for the normal and setback periods are 218 L/s and 146 L/s, respectively (462 cfm and 309 cfm), measured as part of the previous field study (Lozinsky et al., 2023). Each dwelling unit is equipped with exhaust fans in the bathrooms (47 L/s/fan (100 cfm/fan)); kitchen (71 L/s/rangehood (150 cfm/rangehood)); and dryer (71 L/s/dryer (150 cfm/dryer)). The bathroom fan operates concurrently with the bathroom light, while the kitchen rangehood is occupant-controlled. Dwelling unit heating and cooling is provided via recirculating fan-coil units.

Dwelling unit-level data was collected between February and September 2020 and included: indoor temperature and relative humidity; corridor-to-dwelling unit differential pressure; exterior-to-dwelling unit differential pressure; and sliding door, window, and dwelling unit entry door operation. Six dwelling units were originally recruited to participate in the study: two each on the 3rd, 10th, and 14th floors (one on each side of the building). This was done to evaluate differences in PC system performance based on elevation, i.e., wind exposure, as well as location relative to the neutral pressure plane. Unfortunately, due to on-site conditions, we were unable to install the corridor-to-dwelling unit differential pressure sensors in the east-facing 10th floor dwelling unit and the 14th floor dwelling units. All other sensors were installed in these locations. Exterior weather variables including wind speed, wind direction, temperature and relative humidity were collected at 15-min intervals by a weather station installed on the roof of the building across the street. Air density was calculated using the MetPy package in Python (May et al., 2022), which employs Equation 1 for the calculation.

$$\rho = \frac{\epsilon p(1+\omega)}{R_d T(\omega+\epsilon)} \qquad (1)$$

where ρ is the density of the air (Pa), ϵ is the molecular weight ratio (assumed to be 0.622), ω is the mixing ratio (dimensionless), R_d is the gas constant for one kilogram of dry air, and T is the outdoor air temperature (K). The mixing ratio is a function of ϵ, the partial pressure of water vapour (calculated using outdoor temperature and relative humidity measurements), and the total air pressure.

Figure 1 shows a typical dwelling unit layout with equipment locations and Table 1 lists the equipment specifications. Dwelling units were recruited based on their location and residents' willingness to participate in the study. Participating dwelling units were compensated with a smart thermostat. This research was approved by the University of Toronto's Research Ethics Board (protocol number: 38815).

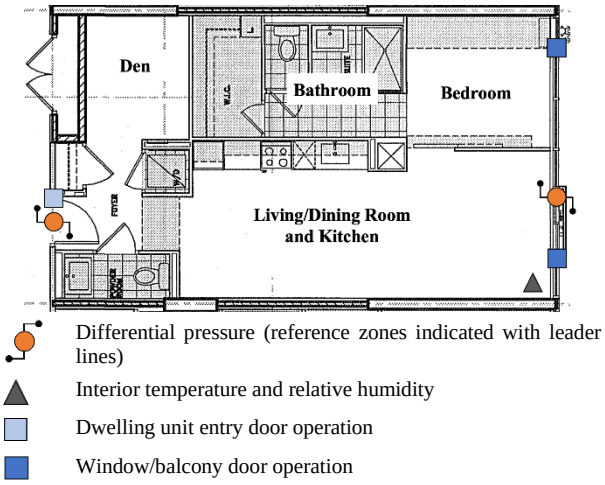

Fig. 1: Typical dwelling unit layout, with equipment locations

While we did not measure dwelling unit-level exhaust fan operation explicitly, exhaust fan operation was easy to visually detect in the differential pressure data. Exhaust fan operation periods were first identified using heuristic classification models, which used the first and second derivatives of the differential pressure measurements (magnitude and direction of change) as well as the magnitude of the differential pressure measurement. First order, lowpass Butterworth signal filters were applied as necessary to reduce signal noise. These models were able to capture the majority of exhaust fan operations in the dwelling units. Exhaust fan operation data was then adjusted through visual inspection to correct both false positives and false negatives produced by the models.

Data Preparation, Model Development and Testing

Extreme gradient boosting (XGBoost) regression models (Chen & Guestrin, 2016) to predict corridor-to-dwelling unit differential pressure for each individual dwelling unit were implemented in Python using Scikit-Learn (Pedregosa et al., 2011). The XGBoost algorithm is a supervised learning ensemble model. It was chosen as it has demonstrated strong performance in building energy use prediction (M. W. Ahmad, Mourshed, & Rezgui, 2017; T. Ahmad & Chen, 2018; Candanedo, Feldheim, & Deramaix, 2017; Divina, Torres, Vela, Luis, & Noguera, 2019; Lu, Li, Bai, & Wang, 2019; Robinson et al., 2017) and has

Table 1: Equipment specifications for dwelling unit monitoring

Measurement	Equipment	Measurement Range	Listed Accuracy
Dwelling Unit Differential Pressure[1]	Veris PX Series Pressure Sensors, connected to Onset HOBO UX120-006M Data Logger	+/- 100Pa	+/- 1% F.S. of selected range
Dwelling Unit Temperature/ Relative Humidity[2]	Onset HOBO U12-013 Temperature/Relative Humidity Data Logger	Temperature: -20ºC to 70ºC (-4ºF to 158ºF) RH: 5% to 95%	Temperature: +/- 0.35ºC from 0ºC to 50ºC (+/- 0.64ºF from 32ºF to 122ºF) RH: +/- 2.5% from 10% to 90%, to a maximum of 3.5%
Window and Door Operation[3]	Reed switch (4 – 20mA), connected to Onset HOBO Data Loggers	4 to 20.1mA	N/A
-	Onset HOBO U12-013 Data Logger (windows)	4 to 20.1mA	+/- 2.5% of absolute reading
-	Onset HOBO UX120-006M Data Logger (balcony and dwelling unit entry doors)	4 to 20.1mA	+/- 0.001mA, +/-0.2% of reading

[1] 10s sampling frequency

[2] 10min sampling frequency

[3] Awning windows = 5min sampling frequency; balcony dwelling unit entry doors = 10s sampling frequency

consistently strong performance in a wide range of regression applications in data science competitions (Bekkerman, 2015). Each dwelling unit was modelled individually, as the small sample size of dwelling units limited the information available for the model to learn the influence of dwelling unit characteristic data (e.g., building floor, orientation) on corridor-to-dwelling unit differential pressures. Table 2 shows the features available for model training. Model features were linearly normalized between zero and one and categorical variables were one-hot encoded.

Table 2: Features (i.e., independent variables) available for model training.

Feature	Units, Description
Air Density	kg/m^3, calculated based on measured temperature, barometric pressure, and relative humidity values.
Wind Speed	m/s, 15-min average wind speed measured by weather station. Down-sampled to five-minute intervals.
Wind Direction	degrees, down-sampled to five-minute intervals. Converted from letter values (e.g., NE, NNW) to degrees with 22.5° resolution, following the DOE convention (North was defined as 0°, clockwise increase in value).
Fenestration Status	Unitless, 1 if all fenestration is closed (balcony door and windows), 0 if any fenestration is open.
Corridor Door Status	Unitless, 1 if corridor entrance door is closed, 0 open.
(MAU) Status	Unitless, 1 if MAU is operating in set-back mode, 0 otherwise.
Cooling/ Heating Season	Unitless, 1 if the month is between June and September, 0 otherwise.
Weekday/ Weekend	Unitless, 1 if it is a weekday, 0 otherwise.

Prior to model training, available data were split into two subsets: 75% of the data were assigned to a training set for training and validation during model development and 25% of the data were assigned to a testing set for performance assessment. Additionally, 25% of the training dataset (i.e., 18.75% of the total data) was reserved for model validation during the training process, including to enable the use of the early stopping parameter in XGBoost. Early stopping was implemented for all models to enable the model to develop the necessary number of estimators to learn data trends, while limiting overfitting.

Model Performance Assessment

Model performance was evaluated on the testing data set using mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE) (shown in Equations 2-4).

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|\left|A_t - F_t\right|\right| \quad (2)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\left|\frac{A_t - F_t}{A_t}\right|\right| \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}} \quad (4)$$

where n is the number of test values, A_t is the actual corridor-to-dwelling unit differential pressure and F_t is the predicted corridor-to-dwelling unit differential pressure.

Bland-Altman analysis (Bland & Altman, 1986), which is a standard approach to assess differences between clinical measurements methods, was used for visualization of model performance. This included plotting the residuals against the measured data, and calculating the upper and lower limits of agreement (LoA) using Equation 5.

$$LoA = \bar{d} \pm 2s \quad (5)$$

where $\bar{d}$ is the mean difference between measurements and s is the standard deviation of the measurement differences. Calculated LoA's represent the 95% confidence intervals of the combined systemic and random differences between measurements for normally distributed residuals.

Feature Selection

Due to the low computational intensity of the models, feature selection was completed using an exhaustive grid search. The hyperparameters used during model training are outlined in Table 3. Values are listed for the three dwelling units with corridor-to-dwelling unit differential pressure measurements. The dwelling unit labels indicate the floor of the dwelling unit, and its orientation (e.g., 3W = the west-facing dwelling unit on the 3rd floor). An XGBoost model was developed for each possible combination of features and the RMSE was calculated. The combination of feature sets that resulted in the lowest RMSE were selected for use in the final model. For all of the dwelling units, using all of the features outlined in Table 2 yielded the lowest RMSE.

Hyperparameter Tuning

Hyperparameters were selected using six stages of sequential tuning, using RMSE to assess model performance for each hyperparameter set. Parameters were tuned in the following order: (1) maximum tree depth and minimum child weight, (2) gamma (regularization), (3) proportion of total features and samples included in each decision tree, (4) alpha (regularization), (5) lambda (regularization), and (6) the learning rate. Finally, the effect of the tree-building method (exact solution or GPU-enabled histogram approximation – gpu_hist) was tested. The final parameter values selected are shown in Table 3. Overall, hyperparameter tuning only produced a slight improvement in model accuracy. However, for one of the three dwelling units, using the "gpu_hist" tree-building method resulted in significantly lower accuracy than the "exact" method. This is worth noting, as gpu_hist is a commonly used tree-building method due to its improved computation efficiency. If increased model accuracy is needed, it may be worthwhile to conduct a sensitivity analysis to determine whether the exact method can improve model accuracy.

Table 3: Model hyperparameters used during feature selection and final values after hyperparameter tuning.

Hyper-parameter	Value Used During Feature Selection	Final Value After Hyperparameter Tuning		
		3W	3E	10W
Maximum Depth	9	22	21	20
Minimum Child Weight	1	0	6	6
Gamma	0	0	0	0
Proportion of Features per Tree	1	0.9	0.8	0.8
Proportion of Samples per Tree	1	0.7	0.8	0.8
Alpha	0	0.0001	0.00015	0.00015
Lambda	1	7	9	7
Learning Rate	0.3	0.4	0.4	0.5

Feature Importance Assessment

After feature selection and hyperparameter tuning were completed, the relative importance of each parameter on model performance was assessed. Three approaches were used: (1) mean decrease in impurity (MDI), (2) permutation importance, and (3) Shapley Additive explanations (SHAP) values.

MDI is commonly used to assess feature importance for tree-based models and is a function of the decrease in node impurity when a tree splits on an individual feature, averaged across all splits. While MDI is computationally efficient, this calculation of feature importance has several notable biases (Agarwal, Kenney, Tan, Tang, & Yu, 2023). These include bias towards variables with high cardinality, misleading importance of correlated features, and skewing based on data distribution.

Permutation importance is calculated by first assessing a model's performance on a dataset, then randomly shuffling each feature in the dataset one at a time and re-assessing performance. Each feature's importance is determined by how much the performance metric decreases when that feature's values are shuffled. This process does not have the same biases related to feature cardinality and dataset distribution; however, it can result in misleading importance values for correlated variables.

SHAP values adapt methods from cooperative game theory and accounts for interactions with other features (Lundberg & Lee, 2017). SHAP values offer a robust method for assessing feature importance; however, their calculation is computationally intensive.

Results and Discussion

First, overall model performance is described, followed by a discussion of feature importance/model interpretation.

Model Performance

Table 4 lists the feature importance metrics for the three modelled dwelling units for (1) the full data set, (2) time steps when the window and balcony door were closed, and (3) time steps when the exhaust fans were turned off. While the values of the mode errors are quite small when considered physically, they do represent substantial percentage errors for many periods of system operation. Based on the field measurements, the *average* winter-time corridor-to-dwelling unit differential pressure was 4.7 Pa, -0.1 Pa, and 13.2 Pa, for 3W, 3E, and 10W, respectively (Lozinsky et al., 2023).

Figure 2 shows the Bland Altman residual plots for each dwelling unit model. Plots of the model residuals as compared to measured data showed a bias of 0.00 Pa for both third-floor dwelling units and a bias +0.01 for dwelling unit 10W (i.e., the model predicted a differential pressure that was 0.01 Pa higher than the actual measured data on average). Based on the calculated LoAs, 95% of the residuals for the two third-floor dwelling units are expected to fall between the bounds of +2.36 Pa and -2.35 Pa and 95% of the residuals of 10W are expected to fall between +5.67 Pa and -5.65 Pa. Visually, it is clear that there is little proportional bias (e.g., biases where the residuals/error increase or decrease with the magnitude of measurement),

Table 4: Feature importance metrics

Dwelling Unit	Evaluation Subset	MAE (Pa)	RMSE (Pa)	MAPE (%)	Testing Set Average (Pa)	Percentage of Dataset
3W	Full	0.60	1.20	30%	10.28	100%
	Fenestration Closed (Open)	0.56 (1.31)	1.14 (2.02)	29% (39%)	10.51 (5.44)	95% (5%)
	Exhaust Fan Off (On)	0.60 (0.59)	1.15 (1.28)	44% (5%)	4.01 (21.23)	64% (36%)
3E	Full	0.65	1.18	89%	2.94	100%
	Fenestration Closed (Open)	0.51 (1.65)	0.85 (2.54)	92% (69%)	2.18 (8.56)	88% (12%)
	Exhaust Fan Off (On)	0.65 (1.12)	1.18 (1.84)	90% (12%)	2.89 (10.11)	99% (1%)
10W	Full	1.36	2.83	17%	13.85	100%
	Fenestration Closed (Open)	1.23 (2.06)	2.65 (3.70)	14% (36%)	13.73 (14.49)	85% (15%)
	Exhaust Fan Off (On)	1.27 (2.28)	2.46 (5.47)	18% (6%)	10.88 (46.68)	92% (8%)

as the residuals for all plots are roughly evenly distributed around the bias line. Several graphs of model predictions versus measured data for 24-hour periods in each of the dwelling units were also considered, to understand how model predictions trended with actual data. Two examples of these plots are shown in Figure 3. Visually, these graphs show that the models generally perform well at identifying trends, however they have challenges in accurately estimating the actual magnitude, particularly around spikes due to sudden changes in wind speed and/or direction. In part, this can be attributed to mismatches in data collection frequencies: differential pressures were measured at 10s intervals, while weather variables were measured at 15-min intervals. Wind can exhibit high spatial and temporal variability, which was not fully captured by the 15-min weather measurements. During initial data post-processing, we aggregated differential pressure data to 1-min intervals, to remove some of the noise caused by variations in wind speed and direction. Aggregating the data to a larger interval may have improved model accuracy, however we would have lost significant temporal resolution.

Limitations of available data also contribute to accuracy limitations. For example, as shown in Figure 3, the predicted effect of exhaust fan operation shows poor agreement with the measured data. This is likely a result of different types and combinations of exhaust fans operating in the dwelling unit (e.g., kitchen rangehood, bathroom, dryer). As noted in the Methodology section, exhaust fan operation was not measured directly, but was inferred from the differential pressure data. This method was able to detect when an exhaust fan was operating, but could not identify which fan, or if multiple fans were operating. Depending on which exhaust fans were used, there would be differing impacts on corridor-to-dwelling unit pressure differentials. This information is not captured in the explanatory variables (which simply indicate if dwelling unit exhaust fans are on or off) and as such, it is not possible for the models to learn these subtleties.

Feature Importance

Figure 4 shows the MDI, Permutation Importance, and SHAP values for each feature in the models. Note: MDI is shown on a separate x-axis from the Permutation Importance and SHAP values as it is a unitless value between 0 and 1, whereas the other two metrics are represented in the same units as the model output (Pa). While MDI is commonly used for assessing feature importance in tree-based models, it does not show agreement with the overall feature importance suggested by Permutation Importance or SHAP methods (more rigorous methods of evaluation). MDI and SHAP values have relatively high agreement for exhaust fan status in suites 3W and 10W and, to a lesser extent, weekday/weekend in suites 3W and 3E. This may be due to the low cardinality of these features. As discussed in the Methodology, MDI is biased towards features with high cardinality; however, this does not explain why other low cardinality features (MAU modulation, heating/cooling season, fenestration status, corridor door status) show poor agreement. Further analysis is required to determine the cause of this trend. The calculated feature MDIs suggest that many features that we know to significantly impact corridor-to-dwelling unit pressures based on physical models are not important in producing accurate predictions (e.g. outdoor air density, which together with indoor air density drives stack effect). SHAP values and, to a lesser extent, Permutation Importance, on the other hand, highlighted that these physical parameters do, in fact, impact prediction accuracy.

Dwelling unit 3E had low MDI and SHAP values for exhaust fan operation, compared to 3W and 10W. This could be explained by 3E's infrequent use of exhaust fans. When exhaust fans were used, the induced differential pressure changes were always very small in magnitude and range, suggesting a single fan used infrequently, and in isolation. In contrast, 3W and 10W used their exhaust fans more frequently, with larger ranges and magnitudes for the corresponding induced differential pressures, indicating that

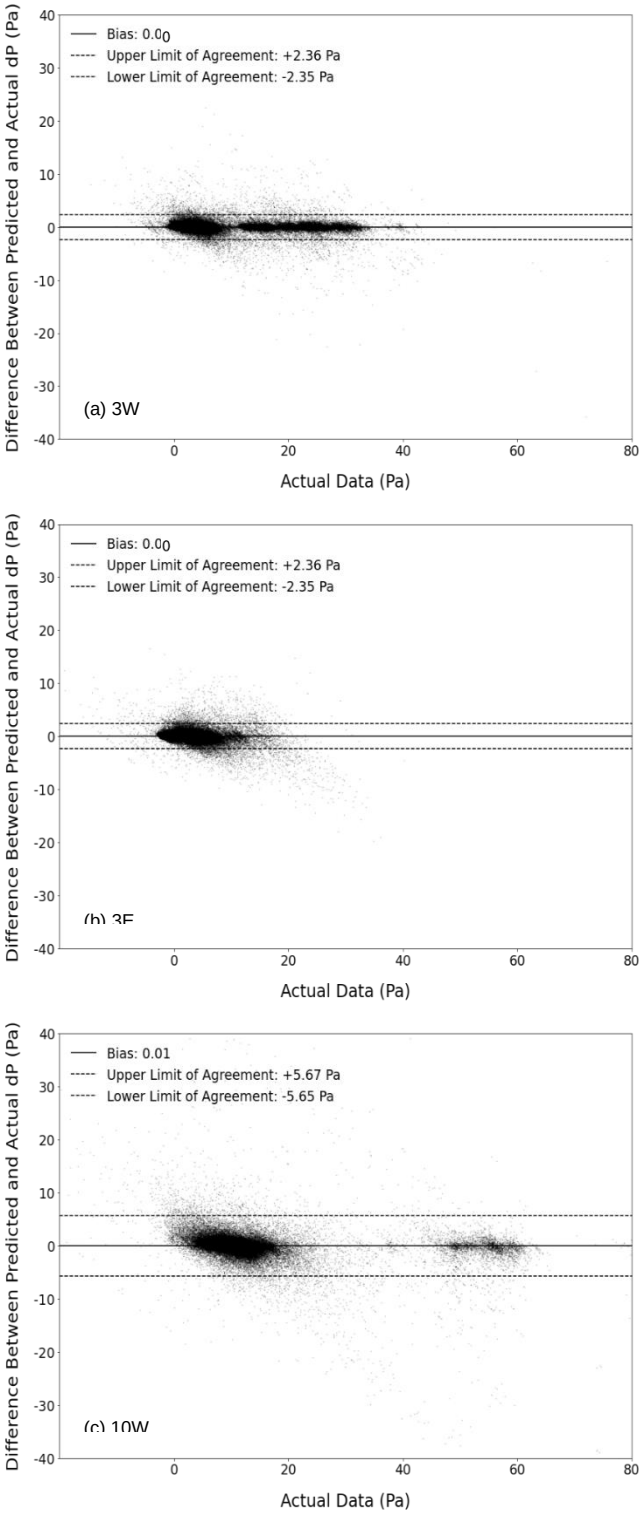

Figure 2: Plots of the differences between measured and predicted corridor-to-dwelling unit differential pressures for the test data sets for (a) 3W, (b) 3E, and (c) 10W.

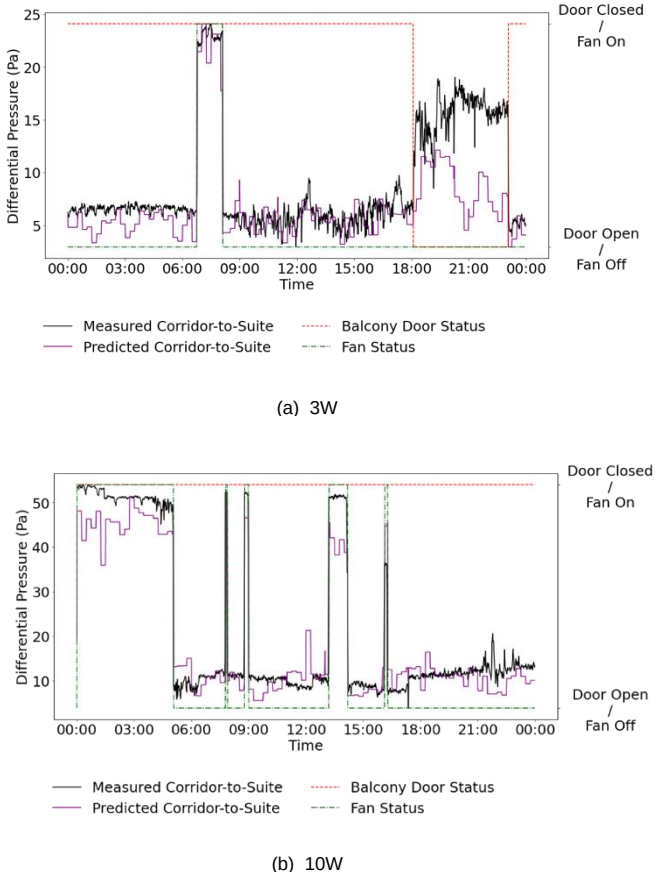

Figure 3: Plots of the measured and predicted corridor-to-dwelling unit differential pressures for (a) 3W, and (b) 10W, for a single day in June. The data for both dates were exlcuded from model training data.

the occupants used multiple fans in different configurations. Unlike MDI and SHAP values, which are biased against rare events, Permutation Importance is not affected by event frequency, which could explain the difference between the three metrics. Wind direction was more important than wind speed for all models, with relatively similar Permutation Importance and SHAP values. This is consistent with previous studies, which found that wind has a largely local effect on inter-zonal differential pressures near the windward elevation (Ricketts, 2014).

Interestingly, the effect of window and/or balcony door operation ("Fenestration Status") is inconsistent, both between the three metrics and the three models. Similar to the exhaust fan operation, this variability may be due to differences in fenestration operation patterns and the frequency of fenestration opening events in the datasets.

Conclusion

We developed an XGBoost model to assess the relative impact of various factors, including weather and occupant behaviours, on PC system performance. The model was trained and validated using measurements from a 17-storey

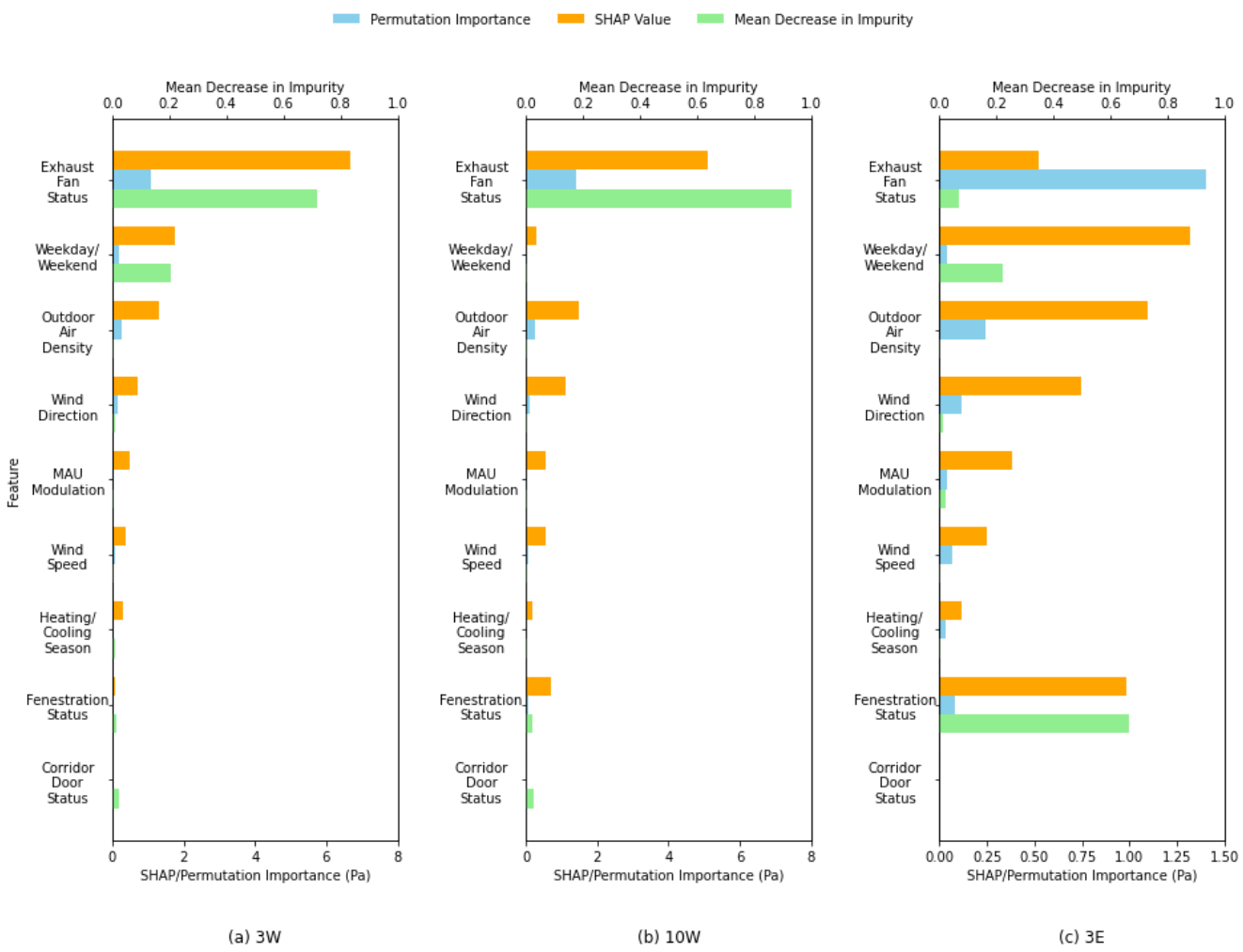

Figure 4: Feature importance metrics (MDI, SHAP,and Permutation Importance) for (a) 3W, (b) 10W, and (c) 3E. Note the difference in scale for the x-axis in (c), in order to better visualize the relative importance of different SHAP & Permutation Importance values for 3E.

contemporary MURB located in Toronto, ON. This paper summarizes the model development including feature selection, hyperparameter tuning, and accuracy assessment, with a specific emphasis on evaluating the potential of using the model as a diagnostic tool and testing environment to improve ventilation system performance.

Dwelling unit-specific feature importance was influenced by the relative frequency of specific events, as the model did not have sufficient data on which to train. This effect could be mitigated by either increasing the data collection period or introducing synthetic data, to ensure that feature events are more evenly represented. SHAP and/or Permutation Importance metrics were more useful than MDI for evaluating feature importance: the use of MDI as the sole metric for assessing feature importance would have under-represented the impact of many features (e.g., wind direction) on differential pressures. Likely, this is due to the lower bias of SHAP values and Permutation Importance based on dataset distribution and feature cardinality as compared to MDI. Differences in data resolutions between the differential pressure and weather data (wind speed and direction) also introduced model inaccuracies. We originally aggregated the differential pressure data to 1-min intervals. A larger bin size, closer to the data collection frequency of the weather data would likely reduce model inaccuracies caused by data resolution mismatch.

This model was trained and validated using data from a single building and likely has poor external validity. Additional data would be required to improve the external validity. That said, we found that this was a useful way to characterize system performance, especially for hard-to-isolate variables like wind speed and direction, which do not lend themselves well to manual estimation techniques. This paper provides other researchers with guidance on model development and analysis techniques for similar applications. Future work will focus on using the model to predict corridor depressurization events and understand how we can better improve PC system performance.

Acknowledgements

The authors would like to thank the condominium board, residents, and operations and property management personnel for their co-operation during the study and their willingness to accommodate this research initiative. This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant RGPIN-2016-06234 and NSERC Post-Graduate Scholarship PGSD3-547141-2020), and ASHRAE's New Investigator Award. The authors would also like to thank ecobee for their financial and in-kind support.

References

Agarwal, A., Kenney, A. M., Tan, Y. S., Tang, T. M., & Yu, B. (2023). *MDI + : A Flexible Random Forest-Based Feature Importance Framework*. 1–83.

Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs Neurons : Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy & Buildings*, *147*, 77–89. https://doi.org/10.1016/j.enbuild.2017.04.038

Ahmad, T., & Chen, H. (2018). Energy & Buildings Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. *Energy & Buildings*, *166*, 460–476. https://doi.org/10.1016/j.enbuild.2018.01.066

Air Solutions Inc. (2005). *Assessment of Suite Compartmentalization and Depressurization in New High-Rise Residential Buildings: Research Highlight*. Ottawa, ON.

Bekkerman, R. (2015). The present and the future of the kdd cup competition: an outsider's perspective.

Bland, J. M., & Altman, D. G. (1986). Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement. *Lancet*, 307–310. https://doi.org/10.1016/j.ajo.2008.09.032

Candanedo, L. M., Feldheim, V., & Deramaix, D. (2017). Data driven prediction models of energy use of appliances in a low-energy house. *Energy & Buildings*, *140*, 81–97. https://doi.org/10.1016/j.enbuild.2017.01.083

Chen, T., & Guestrin, C. (2016). XGBoost : A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Divina, F., Torres, M. G., Vela, F. A. G., Luis, J., & Noguera, V. (2019). *A Comparative Study of Time Series Forecasting Methods for Short Term Electric Energy Consumption Prediction in Smart Buildings*. https://doi.org/10.3390/en12101934

Enermodal Engineering Limited. (2002). *Monitored Performance of an Innovative Multi-Unit Residential Building: Final Report*. Ottawa, ON.

Enermodal Engineering Limited. (2003). *Field Testing of an Integrated Ventilation Space-Conditioning System for Apartments: Final Report*. Ottawa, ON.

Fine, J. P., & Touchie, M. F. (2021). Evaluating ventilation system retrofits for high-rise residential buildings using a CONTAM model. *Building and Environment*, *205*, 108292. https://doi.org/10.1016/j.buildenv.2021.108292

Lozinsky, C. H., Stopps, H., & Touchie, M. F. (2023). In-situ performance of a pressurized corridor ventilation system: A long-term study in a high-rise multi-unit residential building. *Journal of Building Engineering*, *70*(March), 106320. https://doi.org/10.1016/j.jobe.2023.106320

Lu, S., Li, Q., Bai, L., & Wang, R. (2019). Performance predictions of ground source heat pump system based on random forest and back propagation neural network models. *Energy Conversion and Management*, *197*(July), 111864. https://doi.org/10.1016/j.enconman.2019.111864

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768–4777. Red Hook, NY, USA: Curran Associates Inc.

Maxwell, S., Berger, D., & Zuluaga, M. (2014). *Evaluation of Ventilation Strategies in New Construction Multifamily Buildings*.

May, R. M., Goebbert, K. H., Thielen, J. E., Leeman, J. R., Camron, M. D., Bruick, Z., … Marsh, P. T. (2022). MetPy: A Meteorological Python Library for Data Analysis and Visualization. *Bulletin of the American Meteorological Society*, *103*(10), E2273–E2284. https://doi.org/10.1175/BAMS-D-21-0125.1

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn : Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Proskiw, G., & Phillips, B. (2008). An Examination of Air Pressure and Air Movement Patterns in Multi-Unit Residential Buildings. *Building Enclosure Science and Technology (BEST) Conference*. Retrieved from http://c.ymcdn.com/sites/www.nibs.org/resource/resmgr/BEST/BEST1_043.pdf

Ricketts, L. (2014). *A field study of airflow in a high-rise multi-unit residential building*. University of Waterloo.

Robinson, C., Dilkina, B., Zhang, W., Guhathakurta, S., Brown, M. A., & Pendyala, R. M. (2017). Machine learning approaches for estimating commercial building energy consumption. *Applied Energy*, *208*(September), 889–904. https://doi.org/10.1016/j.apenergy.2017.09.060

Sheltair Scientific Ltd. (1998). *Field Testing to Characterize Suite Ventilation in Recently Constructed Mid- and High-Rise Residential Buildings*. Ottawa, ON.