Heating demand and indoor air temperature prediction in a residential building using physical and statistical models: a comparative study

Y Sun1, M M Joybari1, K Panchabikesan1, A Moreau2, M Robichaud3 and F Haghighat1,*

1 Department of Building, Civil & Environmental Engineering, Concordia University, H3G 1M8, Montreal, QC, Canada;
2 Laboratoire des technologies de l’énergie d’Hydro-Québec, G9N 7N5, Shawinigan, Canada;
3 Ouellet Canada, 180 3e Ave, G0R 2B0, L’Islet, QC, Canada;
* Fariborz.Haghighat@concordia.ca

Abstract. In Canada, space heating accounts for the largest proportion of energy consumption in residential buildings. Therefore, accurately predicting the heating demand and interior temperature of a residential building plays a vital role in estimating the building’s total energy consumption with the consideration of thermal comfort. The prediction results obtained through different models could be used to develop predictive controllers to achieve peak shifting as well as to provide utility providers with valuable information for electric power distribution. Common methods to predict heating demand mainly include physical models and statistical methods. This study used two physical models (i.e. TRNSYS model and TRNSYS-CONTAM model) and one statistical model using supervised machine learning algorithm to predict the heating demand as well as the indoor temperature of a residential building, located in Quebec, Canada. Results show that TRNSYS-CONTAM model has higher accuracy than TRNSYS model no matter in terms of interior air temperature or heating demand prediction, while the statistical model shows better interior air temperature prediction result than physical models.

1. Introduction

Space heating accounts for a huge proportion of energy consumption in the buildings located in cold climate regions. For instance, space heating accounted for 63% of annual energy consumption for residential buildings located in Canada in 2013 [1]. Due to the increased heating demand during cold season, the energy consumption in residential buildings gets intense further to maintain the indoor thermal comfort. In addition, occupant living habits would cause daily peaks of energy consumption during the morning and/or late afternoon [2]. As a result, utility companies in several countries adopted time-of-use tariffs [3]. This means that changing the energy consumption pattern or taking methods to reduce peak demand will not only release the stress of the electricity grid but also enable the consumers to save their electricity bill. Based on this concern, shifting peak power using thermal energy storage (TES) systems have shown to be beneficial [4]. Although many buildings use potential TES system, proper control strategies are needed to achieve and ensure peak shifting. In the recent years, predictive controller which optimizes peak shifting strategy based on energy consumption/heating demand as well as interior temperature prediction are being used widely [5]. On the other hand, if the utility company can predict the energy consumption of individual consumers and proper electricity distribution, then the electricity storage strategies can be designed.

Common methods to predict heating demand and indoor temperature mainly include physical models and statistical methods. Physical models predict the thermal behaviour by considering the heat transfer in the building, while for statistical methods, machine learning algorithms are normally utilized based
on the available dataset. Sun et al. [6] presented a TRNSYS-CONTAM model to predict heating demand and indoor air temperature. However, the developed model cannot predict one-day ahead heating demand. Besides, developing the physical models are extremely time consuming [7]. Rahman et al. [8] proposed an algorithm to predict electricity consumption for commercial and residential buildings by using recurrent neural networks. Their model considered the time series features, and thus, would be advantage for system/building with energy storage capacity. However, neural network would require huge amount of training data and high computational power. By contrast, relatively simple machine learning algorithms can be applied, but they usually do not consider time series features [9] and most of them are applied for entire building instead of individual rooms [10].

In order to give a guidance for model selection to predict heating demand and indoor air temperature, in this study a comparison is made between two physical models (TRNSYS model and TRNSYS-CAONTAM model) and one statistical model (linear regression) in terms of prediction accuracy.

2. Experimental house and data collection
The experimental house considered in this study is in Quebec, Canada. It has one semi-underground floor (basement), and two floors above the ground. Floor plans for the experimental house are shown in Figure 1. This building is equipped with three types of heating devices: buried electrically heated floor, surface electrically heated floor and baseboards. Detailed heating system distribution of the considered house can be inferred from the referred paper [6]. Red blocks in Figure 1 represents the location of thermostats. Thermostats can be utilized to collect data utilized in this study, such as time, interior temperature, and instant power consumption with 5 minutes interval. Additionally, information about construction material and infiltration characteristics are considered for physical model development.

Figure 1. Floor plan of (a) Basement (b) Ground floor (3) Second floor

3. Methodology
3.1. Physical models
3.1.1 TRNSYS model. TRNSYS model in the present study can be split into four blocks and the description of each block is described as below.
(1) **Climatic conditions and others.** This mainly include weather data obtained from the closest weather station (Jean Lesage International Airport), soil temperature, as well as infiltration data which was calculated by distributing the whole air leakage rate through separating the global flow coefficient to each zone based on the component area and retaining the flow exponent for each zone the same as the global value.

(2) **Multi-zone building model.** Type 56 (the multi-zone building model) was utilized to predict the zonal temperature or zonal energy demand. To develop this model, several data such as room volume, envelopes and their constituting materials and properties, inter-zonal airflow (assumed to be zero in this TRNSYS model) and parameters from ‘Climatic conditions’ and ‘Thermostat controllers’ are fed as the input.

(3) **Thermostat controllers.** Either simulated power consumption from Type 108 (5-stage room thermostats) or measured power consumption from Type 9 (data reader) to heat zones in the developed Type 56 module.

(4) **Output export.** All required simulation results were exported to text files by Type 65 (online plotter with the file). The exported results were further processed and analyzed in MATLAB.

3.1.2 **TRNSYS-CONTAM model.** Consideration of a constant value for inter-zonal airflow might end up in predicting the temperature and energy consumption of a multi-zone building with more deviations compared to the measured data. To overcome this problem, TRNSYS-CONTAM model calculates dynamic infiltration rate and inter-zonal airflow by an integrated CONTAM module.

3.2. **Statistical model**
Linear regression assumes that the output can be predicted by given input data with linear function, as shown in Equation 1. Note that the input data/features selected in this model are weather information (e.g. exterior temperature, wind direction, wind speed and solar radiation), time of day, day of the week, and last 24 hours data (e.g. indoor air temperature, instant heating power and adjacent room air temperature). The outputs of the statistical model are next day’s heating demand and indoor air temperature. The goal of the statistical model is to find optimum parameters, ‘w’, which ensures minimum square error (Equation 2).

\[ F_w(X) = Xw \]  
\[ \text{Err}(w) = (y - Xw)^T(y - Xw) \]  

where ‘X’ is the matrix of input data, ‘w’ is the matrix of parameters, \( F_w(X) \) is the predicted output. \( \text{Err}(w) \) is the square error.

To avoid overfitting, this study added L2 regularization to the error function.

3.3. **Validation**
Prediction accuracy of all models developed in this study is evaluated in terms of two interpretable criteria: Normalized Mean Bias Error (NMBE) and Coefficient of Variance of the Root Mean Square Error (CV(RMSE)). Physical models are validated based on two months data, while the statistical model utilized one-month data as training set and the other month data as validation set.

4. **Results**
4.1. **Physical models**
Temperature validation results of TRNSYS and TRNSYS-CONTAM models are compared in Table 1, while heating demand validation results are shown in Table 2. Since the status of exterior doors (open/closed) on the ground floor cannot be monitored, validating this floor is complicated. To simplify the research work at current stage, during the validation process, air temperatures and floor surface temperatures on the ground floor were kept the same as the measured data.
Table 1. Temperature validation results of TRNSYS and TRNSYS-CONTAM models

| Room/floor | T model NMBE (%) | T-C model NMBE (%) | T model CV(RMSE) (%) | T-C model CV(RMSE) (%) |
|------------|------------------|--------------------|----------------------|-----------------------|
| B1         | 0.57             | 1.42               | 5.06                 | 3.11                  |
| B2         | 12.5             | 2.33               | 14.23                | 3.33                  |
| B3         | 4.94             | 0.17               | 12.34                | 3.93                  |
| B4         | -0.52            | 0.35               | 5.66                 | 3.08                  |
| B5         | -1.08            | -0.72              | 4.87                 | 4.43                  |
| SF1        | 8.57             | 2.76               | 10.47                | 4.54                  |
| SF2        | 2.25             | -2.34              | 12.03                | 6.00                  |
| SF3        | -8.07            | -1.88              | 15.65                | 6.29                  |
| SF4        | 11.49            | 2.50               | 15.84                | 4.82                  |
| SF5        | 3.32             | -2.89              | 9.76                 | 6.30                  |
| SF6        | 2.24             | -2.36              | 11.46                | 7.11                  |
| SF7        | 6.23             | -1.05              | 10.56                | 4.25                  |

Note: T model means TRNSYS model. T-C model means TRNSYS-CONTAM model.

Table 2. Heating demand validation result of TRNSYS and TRNSYS-CONTAM model

| Floor      | Half-daily | Daily | Monthly |
|------------|------------|-------|---------|
|            | T model NMBE (%) | T-C model NMBE (%) | T model CV(RMSE) (%) | T-C model CV(RMSE) (%) |
| Basement   | -4.50      | -1.68  | 29.01   | 29.03       |
| Second floor | -24.02    | 3.77   | 39.45   | 33.71       |
| Total      | -9.07      | 0.40   | 26.46   | 25.52       |

The inference from Table 1 is that for the basement temperature prediction, TRNSYS model provided a greater absolute value of NMBE (abs(NMBE)) for most of the rooms except room B1, and higher CV(RMSE) for each room when compared to TRNSYS-CONTAM model. As for the second floor, TRNSYS model has high abs(NMBE) (more than 5%) for the temperature prediction for four rooms (SF1, SF3, SF4, and SF7). Particularly, in SF4, the NMBE is 11.49%, which means the TRNSYS model underestimated the room air temperature. Meanwhile, in TRNSYS model, CV(RMSE) values for most rooms in the second floor were higher than 10%. By contrast, in the TRNSYS-CONTAM model, the abs(NMBE) values for rooms in the second floor was always lower than 5%, and the CV(RMSE) values for these rooms were always lower than 10%. Based on the above analysis, the TRNSYS-CONTAM model had better temperature prediction for both the basement and the second floor.

From Table 2, the heating demand prediction results of TRNSYS-CONTAM model was found to be better than TRNSYS model, in terms of both abs(NMBE) and CV(RMSE). Table 2 shows that the energy consumption pattern predicted by TRNSYS model for the second floor is not in good agreement with the measured one. The CV(RMSE) is always higher than 30% regardless of the period, i.e. half-daily, daily, or monthly. The NMBE of the second floor even reached -24.02%, which means the TRNSYS model overestimated the energy consumed by the heating devices on the second floor. The reason is that TRNSYS model does not consider the inter-zonal airflow between the first floor and second floor.

4.2. Statistical model

Three rooms (i.e. B1, GF1 and SF6 which is heated by buried electrically heated floor, baseboards and surface electrically heated floor, respectively) are selected to test the applicability of statistical model. Table 3 shows the temperature and heating demand validation results. The temperature prediction of the statistical model shows a high accuracy, i.e. both NMBE and CV(RMSE) are lower than 10%. However, the heating demand prediction for all rooms shows extremely high CV(RMSE). The possible reasons for high heating demand prediction error of the statistical model could be: (1) poor data quality caused by improper location of adjacent thermostat; (2) offline heating devices in the same room, such as that an extra off-line baseboard is installed in SF6; (3) training data is not enough. On the other hand, poor performance of the statistical model lead to a question, whether the accurate performance of physical models are reliable.
Table 3. Temperature and heating demand validation result of the considered statistical model

| Room | NMBE_Temp (%) | CV(RMSE)_Temp (%) | NMBE_Heat (%) | CV(RMSE)_Heat (%) |
|------|---------------|-------------------|---------------|-------------------|
| B1   | 1.53          | 2.53              | 11.51         | 68.04             |
| GF1  | 2.76          | 5.64              | 25.74         | 97.08             |
| SF6  | 0.22          | 2.36              | -4.62         | 64.24             |

Figure 2 shows the detailed indoor air temperature and heating demand prediction results for the three rooms with respect to time. The inference from the figure is that the predicted temperature or heating demand by the statistical model follows the trend of measurement data most of the time.

5. Conclusion
In this study, two physical models (TRNSYS and TRNSYS-CONTAM) and a statistical model (linear regression) were used to predict both heating demand and indoor temperature for multi-rooms in a
residential building. Between the considered physical models, it is found that the simulation accuracy of temperature and energy prediction got improved when the interzonal airflow was considered (TRNSYS-CONTAM model) as compared to when it was ignored (TRNSYS) to the pure building energy (TRNSYS) model. Besides, statistical model shows higher accuracy compared to physical models in terms of indoor temperature prediction. However, the accuracy of statistical model for heating demand prediction is worse. This may be caused due to poor data quality, lack of training data, false selection of features. Further investigations are required to justify the above said reasons. On the other hand, poor performance of the statistical model lead to a question, whether the accurate performance of physical models are reliable.

Future work will focus on proper feature selection for statistical models, model selection methods, integration of separate room predictor for entire building.

References

[1] Natural Resources Canada 2016 Energy Efficiency Trends in Canada 1990 to 2013
[2] Thieblemont H, Moreau A, Lacroix G and Haghighat F 2017 Simplified anticipatory control for load management: application to electrically heated floor. Presented at the International Conference & Workshop REMOO 2017
[3] Yang L, Dong C, Wan CJ and Ng CT 2018 Electricity time-of-use tariff with consumer behavior consideration. International Journal of Production Economics. 146(2):402-10
[4] Romaní J, Belusko M, Alemu A, Cabeza LF, de Gracia A and Bruno F 2018 Control concepts of a radiant wall working as thermal energy storage for peak load shifting of a heat pump coupled to a PV array. Renewable energy. 118:489-501
[5] Killian M and Kozek M 2016 Ten questions concerning model predictive control for energy efficient buildings. Building and Environment. 105:403-12
[6] Sun Y, Panchabikesan K, Joybari MM, Olsthoorn D, Moreau A, Robichaud M and Haghighat F 2018 Enhancement in peak shifting and shaving potential of electrically heated floor residential buildings using heat extraction system. Journal of Energy Storage. 18:435-46
[7] Thieblemont H, Haghighat F, Ooka R and Moreau A 2017 Predictive control strategies based on weather forecast in buildings with energy storage system: A review of the state-of-the art. Energy and Buildings. 153:485-500
[8] Thieblemont H, Haghighat F, Moreau A and Lacroix G 2018 Control of electrically heated floor for building load management: A simplified self-learning predictive control approach. Energy and Buildings. 172:442-58
[9] Gray FM and Schmidt M 2016 Thermal building modelling using Gaussian processes. Energy and Buildings. 119:119-28
[10] Luo N, Hong T, Li H, Jia R and Weng W 2017 Data analytics and optimization of an ice-based energy storage system for commercial buildings. Applied Energy. 204:459-75