

Article

Machine Learning for Benchmarking Models of Heating Energy Demand of Houses in Northern Canada

Behrad Bezyan and Radu Zmeureanu * 

Centre for Zero Energy Buildings Studies, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec H3G 1M8, Canada; bbezian@yahoo.com

* Correspondence: radu.zmeureanu@concordia.ca

Received: 7 February 2020; Accepted: 2 March 2020; Published: 4 March 2020



Abstract: In most cases, the benchmarking models of energy use in houses are developed based on current and past data, and they continue to be used without any update. This paper proposes the method of retraining of benchmarking models by applying machine learning techniques when new measurements are made available. The method uses as a case study the measurements of heating energy demand from two semi-detached houses of Northern Canada. The results of the prediction of heating energy demand using static or augmented window techniques are compared with measurements. The daily energy signature is used as a benchmarking model due to its simplicity and performance. However, the proposed retraining method can be applied to any form of benchmarking model. The method should be applied in all possible situations, and be an integral part of intelligent building automation and control systems (BACS) for the ongoing commissioning for building energy-related applications.

Keywords: benchmarking models; linear regression; machine learning; building automation and control system; ongoing commissioning; measurements; Northern houses

1. Introduction

One of the first steps of the ongoing commissioning of heating, ventilation, and air conditioning (HVAC) of buildings consists in the comparison of the recorded measurements with the predictions of benchmarking models to detect unusual operation conditions, faults in sensors, or degradation of equipment performance. This comparison should be applied in all possible situations and be an integral part of building automation and control systems (BACS).

The benchmarking models can be classified in public benchmarking and internal benchmarking [1]. The public benchmarking models are developed through the statistical analysis of energy performance of a sample of buildings by using independent variables or regressors such as conditioned floor area, number of floors, location, or building main activity type. The internal benchmarking models for assessing the building energy performance can be classified mainly in (i) physics-based or white-box models (e.g., detailed energy simulation models such as EnergyPlus and TRNSYS) and (ii) data-driven (inverse models or black-box models), which are developed from the measurements of normal operating conditions of the building under analysis, or from synthetic data obtained by the simulation with calibrated programs such as EnergyPlus and eQuest. The experience proved that the data-driven models are easy to develop, train, and retrain, and give good predictions if sufficient data of good quality are available. A hybrid class called gray-box models can also be developed by coupling simple physics-based models with data-driven models.

So far in most cases, the benchmarking models of energy use in houses are developed based on current and past data, and they continue to be used without any update. This paper proposes

the method of retraining benchmarking models by applying machine learning techniques when new measurements are made available. The method uses as a case study the measurements of heating energy demand from two semi-detached houses of Northern Canada. The results of the prediction of heating energy demand using static or augmented window techniques are compared with measurements.

The paper presents, in Section 2, the use of benchmarking data-driven models for the assessment of building energy performance. Section 3 discusses the application of machine learning models that use statistical techniques to process and learn from data recorded from current and past measurements, and improve the performance of data-driven models using previous knowledge or data without being explicitly programmed. Section 4 presents the retraining of proposed benchmarking model using machine-learning techniques, when new data become available. Section 5 presents the case study of two semi-detached houses of Inuvik, NWT, Canada. Finally, Section 6 presents the conclusions of this study.

2. Benchmarking Data-Driven Models

The use of energy signatures of the heating energy use of houses is a well-established procedure. The energy signature is based on a physical model written as a regression model, which is obtained by system identification also known as inverse modeling [2]. The most common method used for the identification of model parameters is the least squares method. Some features of the benchmarking data-driven models, presented in this section, are listed in Table 1.

Fels [3] developed the Princeton Scorekeeping Method (PRISM) that evaluates the energy savings after the retrofit, by calculating the difference between the normalized annual energy consumption before and after retrofit. Each normalized annual consumption (NAC) is obtained by using the energy signature as the linear relationship between the daily average energy use, calculated from the utility bills, and heating or cooling degree-days of the corresponding period from the closest airport or local weather station. The NAC estimate is a reliable and stable index of performance. The three parameters a , b , and T_{REF} are physical meaningful indicators, whose changes over time might not be statistically significant. Decicco et al. [4] presented a case study of gas consumption for space and water heating in a multifamily building, and discussed the physical parameters extracted from the data. They compared the PRISM results using monthly utility bills versus daily measurements.

Zmeureanu [5] showed by using synthetic data that the linear energy signature that was developed for one year, using the daily energy consumption and the daily outdoor temperature, does not change for subsequent years, provided that no modifications of the building envelope or HVAC systems are performed. Although it is easy to obtain the daily average energy consumption from the monthly utility bills, the energy signature is less accurate compared with the case of using the daily measured values.

Other types of energy signatures are displayed in ASHRAE Guideline 14-2014 [6]. It is beyond the scope of this paper to present all different types of energy signature developed over the past three decades. Only a few extensions of the original energy signatures are presented herein.

Yu and Chan [7] developed the energy signature of chillers, in which the hourly electric demand is linearly correlated with a climatic index that is defined as the product of outdoor air temperature and humidity ratio. They analyzed 16 combinations of 4 design options and 4 operating strategies for chillers, using synthetic data from the simulation of a hypothetical hotel.

Monfet and Zmeureanu [8] developed multivariate linear and nonlinear models and artificial neural networks (ANNs) to benchmark the energy performance of water-cooled electric chillers.

Catalina et al. [9] proposed a multivariate non-linear regression model for prediction of heating energy demand by using three regressors: The overall building heat loss coefficient, the south equivalent window surface area, and the difference between the indoor set point temperature and the sol-air temperature. They used both synthetic data and measurements from 17 apartment buildings.

Korolija et al. [10] developed non-linear, multivariate, and linear regression models for the prediction of annual heating, cooling, and auxiliary energy requirements as a function of office building heating and cooling demands, respectively. They used synthetic data from the simulation of a large number of office building models each one served by five different HVAC systems.

Zhang et al. [11] compared four benchmarking models for prediction of hot water energy consumption of an existing office building, by using the three-parameter change-point regression model, the Gaussian process regression model, the Gaussian mixture regression model, and the artificial neural network model. They concluded that the three-parameter model was the most appropriate for this case study in terms of accuracy and efforts spent for the modeling. The daily models have a higher R^2 than the hourly models.

Hitchin and Knight [12] discussed the use of energy signature of air conditioning system of an office building using measurements of energy consumption at 15 min intervals. The model parameters have limited diagnostic power when considered individually, but combinations of values can suggest causes of unusual consumption levels. They concluded that daily energy signatures can generate more robust energy consumption benchmarks and provide additional insight compared to monthly or weekly signatures.

Abushakra and Paulus [13] concluded that hourly data of two weeks in swinging season (spring or autumn) are sufficient for developing the benchmarking model of annual building energy use. In order to predict the long-term energy use based on the short-term measurements, they recommended the optimum selection of (i) the length of the observation period, (ii) the time or season of the observation, (iii) the required variables, and (iv) the technique for developing a benchmarking model that is effective and acceptable by the user.

Table 1. Features of quoted benchmarking data-driven models.

Model	Time Interval of Data	Type of Data	Time Interval of Prediction	Training R^2	Ref.
$WBEU = a + b.T_o$	5-mn	Real	Hour	0.423	[2]
$WBEU = a + b.T_o + c.GHR$			Day	0.711	
			Hour	0.579	
			Day	0.740	
$NAC = 365.a + b.H_o(T_{REF})$	Month	Utility bills	Year	0.985	[3]
$NAC = 365.a + b.H_o(T_{REF})$	Month	Utility bills		0.82–0.97	
	Month	Synthetic	Year	0.997	[4]
	Hour	Real		0.76–0.86	
$WBEU = a + b.(T_o - T_{REF})$	Hour	Synthetic	Day	0.80–0.86	[5]
	Month		Day	0.93–0.99	
$BLR = a + b.(T_o.W)$ $E_{CH} = a + b.(T_o.W)$	Hour	Synthetic	Hour	0.80–0.857	[7]
$E_{CH} = \alpha \cdot \frac{Q_E}{Q_{Edes}} + \beta.T_{CNDS} + \gamma.T_o$	15-mn	Real	15-mn	0.994–0.999	[8]
$WBH = \alpha + f(G, SES, \Delta\theta)$	Hour	Synthetic	Year	0.974	[9]
$HVC = \alpha.WBC$ $HVC = \alpha.WBC^\beta$	Hour	Synthetic	Year	0.654–0.994 0.951–0.998	[10]
$HVHW = a + b.(T_o - T_{REF})^+$	5-mn	Real	Hour	0.85	[11]
			Day	0.96	
$WBEL = a.T_o + b$ $WBG = a.T_o + b$	Day	Real	Day	0.863 0.86	[14]
$E_{AC} = a + b.(T_o - T_{REF})^+$	Hour	Measured	Month	0.57–0.94	[15]
$WBEU = \alpha.T_o + \beta.LPD + \gamma.PLPD + \delta.HSP + \varepsilon.CSP + \eta.SAT$	Hour	Synthetic		0.944	[16]
		Real	Day	0.796	

Ko et al. [14] used the linear energy signature to develop the benchmark of whole building daily electricity use and gas energy use, respectively, versus the outdoor air temperature. They used the daily clusters to improve the predictions from $R^2 = 0.1$ to 0.863 for electricity use, and from $R^2 = 0.468$ to 0.86 for gas use.

Perez et al. [15] applied the daily house energy signature, defined as the three-parameter change-point model, to disaggregate air-conditioning smart meter data from 45 single-family houses. They performed the statistical analysis of the energy signature slope to determine significant energy related variables of those houses, with the final scope to identify and target those houses for energy efficiency improvements.

Chen et al. [16] developed a multivariate linear regression model of the whole building daily energy use by using as regressors some variables such as the outdoor dry-bulb temperature, interior lighting power density, thermostat set point temperature, and supply air temperature. They used synthetic data from the simulation an office building. The approach was also tested with data from an existing office building.

The literature review by Fumo and Biswas [2] supported the feasibility of using the simple and multiple linear regression models for the prediction of whole building energy use of single-family houses. Their results showed that the simple daily linear energy signature has a better-quality parameters, as proved by the coefficient of determination $R^2 = 0.711$ compared with $R^2 = 0.423$ for multiple hourly energy signature. The use of multiple linear models does not bring a significant increase of R^2 : 0.740 versus 0.711 of the single linear model.

In summary, the benchmarking models presented in the literature review are developed using; (1) hourly or monthly synthetic data obtained from computer simulation, (2) monthly or bimonthly utility bills, and (3) real measurements at 5-mn, 15-mn, 1-hr, or 1-day time interval. The highest values of the coefficient of determination R^2 for training of the benchmarking model were obtained by using synthetic data. The use of synthetic data for the development and testing of a reference data-driven benchmarking model has some advantages: Data is noise free; there are no measuring errors, operation errors, faults, or degradation of energy performance of equipment in time. However, the simulated data cannot be used for the ongoing commissioning, but only as reference prediction performance. The daily average energy signature derived from the utility bills has a high R^2 value, but the application must wait for one or two months for having access to the utility bills. This is not practical for the ongoing commissioning. The simple linear two-parameter or three-parameter regression models, known as energy signature, have higher values of training R^2 when they are developed by using real data, compared with multivariable models. The application of such models for the ongoing commissioning is a practical solution with good results.

3. Machine Learning (ML) for Benchmarking Models

The term of artificial intelligence (AI) is often used in relation with current applications for the building operation and maintenance of HVAC systems. However, the AI systems that truly imitate or even exceed human cognitive abilities such as processing, learning, and reasoning are still unavailable for the optimization of operation and maintenance of buildings and HVAC systems. There are high expectations that the future generations of building automation and control systems will use such AI methods.

In our opinion, an intelligent BACS installed in a changing environment must be able to: (1) continuously have access to all available measurements, (2) analyze and select those measurements that help in the decision-making, (3) select the most appropriate benchmarking model among a set of generic available models, (4) continuously/periodically update the parameters of that model by learning from changes, or even select another model and improve it with new data, (5) control the HVAC system to achieve the desired performance, and (6) verify the HVAC system performance. All these tasks must be performed without being explicitly programmed or without any interference

from the user. Under this concept, such intelligent BACS do not exist yet. The development of such systems is still a goal for research.

In the past two to three decades most publications related to HVAC systems have presented methods such as neural networks, genetic algorithms, patterns recognition, expert systems, or case-based reasoning. However, such publications have not claimed connection with AI. Other authors indicated that the proposed methods are based on AI, without claiming that those methods have full computational intelligence. Other papers attached the terms of AI technologies and computational intelligence to different new systems. The term of soft computing was also used for building and HVAC applications that gives approximate and uncertain solutions by using methods such as evolutionary algorithms (e.g., genetic algorithms, ant colony optimization, particle swarm optimization) and machine learning (e.g., neural networks, and support vector machines).

Kim and Katipamula [17] expanded the review of 2005 [18,19] with additional 118 new studies of automated fault detection and diagnostics (AFDD). They classified the AFDD models in three categories: Processed-based (e.g., data-driven models), qualitative model-based (e.g., rule-based), and quantitative model-based (e.g., detailed physical models). They also discussed the combination of some AFDD models. Among the data-driven models, the black box models were the most commonly used (62% of studies), followed by the rule-based models (26%). The review revealed that most AFDD methods are applied to small or large commercial buildings associated mostly with variable air volume-air handling units (42% of papers), chillers and cooling towers (17%), rooftop units (17%), whole building energy use (12%), and the remaining for commercial refrigerators, lighting, and other HVAC units. Apparently, the number of studies that applied AFDD to residential buildings was negligible. Among the conclusions, the authors recommended future work for the improvement of methods that eliminate the need for manual model identification or algorithm training; this is a topic related to the scope of this paper. A few papers used in the review [17] are discussed shortly below.

Yu et al. [20] developed a steady-state gray-box based virtual meter of the supply airflow rate in packaged rooftop units from the measurements in the laboratory settings. The nonlinear correlation models use as regressors the supply and outdoor air temperatures, and outside air damper status. The total uncertainty under laboratory settings of the virtual air flow meter is $\pm 13.8\%$ under cooling mode and $\pm 6.9\%$ under heating mode, while in field applications the uncertainty of the virtual meter might be higher due to a variety of factors such as uneven air distribution, gradual drifting, and faulty installation.

Sun et al. [21] developed a gray-box model that uses the statistical process control (SPC) and Kalman filter-based method for fault detection in chillers and cooling towers. The model coefficients are obtained by applying the least-squares method to historical measured data. If there is a device fault, the model parameters may deviate from its normal range and the fault can be detected by a SPC rule.

Bynum et al. [22] presented the Automated Building Commissioning Analysis Tool (ABCAT) that is a prototype fault detection and diagnostic tool. It is a first principle-based whole building level top down tool that compares the measurements of whole building electricity, whole building heating, and whole building cooling, with the predicted energy consumption. The faults are detected if the differences between the whole building predictions and measurements are statistically significant. The prediction model uses some assumptions, in the absence of detailed information, such as constant COP of the heat pump and constant boiler efficiency.

Capozzoli et al. [23] presented a simplified approach for the detection of anomalies in the active electrical power for lighting and the total active electrical power of each building. The method uses statistical pattern recognition techniques and artificial neural ensembling networks coupled with outliers' detection methods for fault detection.

Seem [24] presented a data analysis method for the automated detection of abnormal energy consumption in buildings. The method accounts for weekly variation in energy consumption by grouping days of the week with similar power consumption. A robust outlier detection method is used to determine if the energy consumption is significantly different than previous energy consumption.

The generalized extreme studentized deviate (ESD) many-outlier procedure is used to identify the outliers in the data set.

Amasyali and El-Gohary [25] presented an extensive review of papers about the data-driven building energy consumption predictions that utilized machine learning algorithms. They compiled information about studies that use support vector machines (SVM), artificial neural networks, decision trees, and statistical algorithms such as general linear regression, multiple linear regression, autoregressive integrated moving average, and case-based reasoning. They compiled information from 63 papers in terms of scope of predictions, learning algorithm, building type, data temporal granularity, type of data set and regressors, and performance metrics. They found out that: (i) only 19% of the reviewed studies focused on residential buildings, (ii) only 20% of studies focused on the building heating energy use, and (iii) the coefficient of variance was the most-commonly used evaluation metric. They concluded that all the models have their own strengths and weaknesses; there is no one single model which can be used for all conditions. They recommend the development of application-specific models.

Most of the surveyed artificial neural networks models, a subset of ML models, related to the building energy predictions are static in nature. The prediction model is set up in advance using historical data and does not change when new measurements become available.

The class of adaptive models, e.g., ANN models, can help for eliminating this limitation. The adaptive models are capable of adapting themselves by retraining the models to unexpected pattern changes observed when new measurements are made available [26–28]. Most such models were found in the literature related to the electric load forecasting for power system. This class of adapting models can be used for the ongoing commissioning of building energy systems. A few examples of adapting benchmarking models are presented in this section, with the corresponding coefficient of variance of root mean squared error (CV(RMSE)) as the performance metric listed in Table 2.

Table 2. Performance metric, coefficient of variance of root mean squared error (CV(RMSE)) in percent (%), of predictions over the testing data set using adapting benchmarking models.

Benchmarking Model	Type of Data	Time Interval	Static Training	Augmented Window Training	Sliding Window Training	Ref.
ANN	Synthetic	Hour	7–16	15–17	15–40	[29]
Multi-linear	Real	15-mn	5.1–18.6	3.1–12.3	3.8–27.1	[8]
ANN	Real	15-mn	8.4–13.8	8.0–9.4	8.7–9.2	[30]

Yang et al. [29] developed adaptive ANN models for the forecasting of electric demand of chillers, for the cluster of working hours of weekdays, by using synthetic data from the simulation of a large office building. They compared two adaptive training techniques, i.e., the accumulative retraining and sliding window retraining techniques, against the static training. The initial data set used for the static training was composed of the first 20 days of June. For the accumulative retraining the data set was increased by the measurements of one day, when they become available. This updated data set was used to retrain the ANN model for carrying out subsequent predictions. For the sliding window retraining technique, the size of the training data set was kept constant and new measurements of one day were added, while some of the oldest data were dropped from the training set.

Monfet and Zmeureanu [8] presented a new approach for the development and use of benchmarking models of the electric demand of chillers of a university campus, in the context of ongoing commissioning. The model was the energy signature presented in Table 1. Different techniques were explored with different sizes of training and testing data sets. For instance, the static approach was trained with a data set of 10 weekdays in June, and tested with two days in July. The augmented window technique used the initial training data of 10 days in June, followed by the increase of training data set by 7 days, and tested with data of 4 days. The sliding window technique

used a fixed window of 14 days, where the first 10 days of data were to establish the model, and the last 4 days for testing.

Chae et al. [30] proposed a short-term forecasting model of the whole building electricity usage based on an ANN model with Bayesian regularization algorithm. The model was applied to measurements from a commercial office building complex. For the static training, the model was trained using four weeks of data in July, and it forecasts for August and September without retraining the model. For the accumulative retraining, the initial data set of July was augmented every day with new available; the ANN model retrained on a daily basis. The sliding windows retraining used a fixed data set size of four weeks (equal to the data set of July). The window was shifted by a day by removing the first day of the old training data set and adding the new measurements into the data set. The ANN model was retrained daily with the new training data set. Table 2 shows the CV(RMSE) of the forecasting results for the weekdays of August 2012.

Since the results depend on the application conditions such as the variation with time of target value, the type and quality of forecasting models, and the size of data set used for the adaptive retraining, the generalization of results is a challenging undertake. A case-base adopted retraining technique should be the solution.

The authors prefer to focus on the application of machine learning algorithms to the development of systems for the operation and maintenance of HVAC systems. ML is a subset of AI in the field of computer science that uses statistical techniques to give computer systems the ability to process data and learn, and hence continuously improve performance by using previous knowledge or data without being explicitly programmed.

A different approach is proposed in this paper. The paper contributes to the body of knowledge regarding the use of machine learning techniques along with measurements recorded by the BACS from the HVAC systems for the training, testing, and retraining, when new data become available, of benchmarking models of building heating energy demand. The proposed method can be implemented for the purpose of ongoing commissioning of energy performance of houses by using only the available sensors, without requiring a dedicated monitoring system.

The main contribution of this paper is the proposed ML technique for the training and testing of benchmarking models of heating energy demand of houses by using measurements of the reference period, and for the periodic update of models by learning from new measurements. The proposed method is applied to measurements from two semi-detached houses in Northern Canada. The paper presents the comparison between the predicted and measured daily heating energy demand.

4. Method

4.1. Benchmarking Gray-Box Model

Any type of benchmarking model can be used with the proposed method. The use of large numbers of regressors might give the impression that the prediction model is more accurate. However, when a short list of regressors is used, the uncertainty due to the error propagation from the measurements up to the dependent variable is smaller than when many regressors are used. Moreover, the use of a large number of regressors requires a large number of sensors, which might not be available in most HVAC systems because of budget reasons.

A benchmarking gray-box model was used that combines some knowledge about the physical phenomena with the measurements, as “data alone is not enough” for the generalization of the model [31]. The physical-based model is the daily signature of space heating energy demand, a two-parameter model (Equation (1)), which is developed from easily available measurements of two variables: The outdoor air temperature and the heating energy demand.

$$E = a.T_o + b \quad (1)$$

where, E is the daily heating energy demand in MJ/(m² day), a is the slope of the weather-dependent energy demand in MJ/(m² °C day), b is the intersect or reference energy demand at 0°C in MJ/(m² day), and T_O is the daily average outdoor air temperature in °C. The coefficient a (Equation (1)) reveals the sensitivity of daily heating energy demand of the house to changes in the daily average outdoor air temperature. Once the coefficients a and b are identified, for instance, from the measurements of one year, they do not change when the measurements from to another year are used, provided that the building thermal characteristics, the efficiency of HVAC system, or the occupancy pattern do not change, such as after a major renovation of the building envelope, or change of heating equipment.

This model is one of the sample models presented in [6] that gives good predictions of heating energy use in buildings [32–34]. The linear regression models belong to the family of ML algorithms along with SVM and ANN [30]. Multivariate models can also be used, but the main obstacle is the availability of measurements of all regressors.

Machine-learning techniques are used for the training and retraining of the benchmarking model by using recorded measurements of a small number of variables. The model was trained to identify the initial relationship between the heating energy demand and regressors, and then it was retrained periodically with new measurements.

If the data set contains measurements over a long time period such as two heating seasons, then the training could be more effective. However, if the available data set is smaller, the analyst is constraint to use what is available for training, to be able to deploy as soon as possible the benchmarking model for the ongoing commissioning. This second situation is used as the case study in the paper.

Once the models are trained and tested, they can be used for the prediction of daily heating energy demand over the following operation periods, and for the comparison with actual measurements. Significant differences should alert of possible faults or deterioration of system equipment and sensors, or even changes in peoples' occupation patterns. When new data becomes available, the benchmarking models should be retrained to benefit for the most recent information. Two retraining techniques are compared in the paper: The static window versus augmented window.

4.2. Training, Testing and Application Data Sets

The data set of measurements of December 2014 was selected as the reference, and used for the initial development of benchmarking model. The data set was composed of; (1) the training data set, which is used to identify the model coefficients and (2) the testing data set, which deploys the balance of data set to verify the model accuracy. For the purpose of presenting the method, the model was initially trained with a data set of the first three weeks of December 2014 (i.e., December 1–21), and tested with a data set of the last week of December (i.e., December 22 to 31).

The tested model was then used along with the application (prediction) data set to estimate the daily space heating energy demand of the following periods. The tested model can be used unchanged over the application time interval by applying the so-called static window technique, or the model can be periodically retrained with new measurements (i.e., the update of the model coefficients a and b) by using the augmented window technique.

In the case of static window technique, the daily heating energy demand was predicted by the tested benchmarking model over the remaining part of the heating season, for instance from January 1 to March 31, 2015.

In the case of the augmented window technique, the measurements of two weeks were added to the initial data set. The first week of new data was added to the initial training period. The model was retrained with a data set of five weeks (December 1, 2014 to January 4, 2015), the model coefficients a and b were updated, and the model was tested with a data set from the following week, i.e., January 5 to 11, 2015. The new retrained model was then used for prediction of the daily heating energy demand from January 12 to March 31, 2015.

The trained and tested model predicts the daily heating energy demand under the assumption that the physical phenomena through the building envelope and of the operation of the heating system

remain within the range of values experienced during the training period. When this model is applied to other time intervals, the model should predict the daily heating energy demand that should be obtained under the reference conditions.

Any large difference between the predictions and measurements would detect abnormal results due to faults, degradation in operation, or changes in the house operation. This method can be used as the first step (i.e., detection) in fault detection and diagnosis (FDD) methods. The identification of the source of abnormal results is the second step of FDD methods. The method presented in this paper can be used only for the detection phase.

4.3. Quality of Predictions of Benchmarking Model

Three statistical indices are used: The coefficient of determination R^2 (Equation (2)) for the model training, and the root mean squared error (RMSE) (Equation (3)) and the CV(RMSE) (Equation (4)) for the model testing and application by comparing the predictions and measurements of space heating energy demand.

$$R^2 = \left[1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \right] \cdot 100 \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n - 1}} \quad (3)$$

$$CV(RMSE) = \frac{\sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n - 1}}}{\bar{y}} \cdot 100 \quad (4)$$

where y_i is the measured value, \hat{y}_i is the predicted value, \bar{y}_i is the average measured value, and n is the number of observations.

According to ASHRAE guideline 14 [6] the model predictions of the whole building energy consumption, when using hourly data, are acceptable if CV(RMSE) is less than 30%; and the model predictions using monthly data are acceptable if CV(RMSE) is less than 15%. Since this paper uses daily data of heating energy demand, and this case is not identified in [6], the authors followed the recommendations by Kaplan et al. [35]. This last reference proposed the use of different levels of acceptable differences in terms of the end-use type and the interval of time for comparison. For instance, in the case of HVAC energy use, the maximum acceptable differences should be CV(RMSE) = 25–35% for daily values, and 15–25% for the monthly values. Therefore, in this paper, the daily predictions were acceptable if CV(RMSE) is less than 30%. The readers should also be aware that the maximum CV(RMSE) listed in [6] were obtained by consensus rather than by a scientific method.

4.4. Estimation of Total Heating Energy Demand of Application Period

The benchmarking model can be used along with the outdoor air temperature bins to estimate the total heating energy demand [5] over the application period (Equation (5)).

$$E_p = \sum_{i=1}^n (a \cdot T_{O,j} + b) \cdot \text{BIN}(T_{O,j}) \quad [\text{MJ}/\text{m}^2] \quad (5)$$

where, E_p is the predicted total energy demand by the benchmarking model; a is the slope and b is the intercept of the non-weather dependent energy demand, both identified during the training phase (Equation (1)); T_O is the daily average outdoor temperature ($^{\circ}\text{C}$); and $\text{BIN}(T_O)$ is the number of days of occurrence of the daily average outdoor air temperature bin having T_O as center.

5. Case Study

5.1. Houses

The authors had access to the measurements from October 1, 2014 to September 30, 2015 in two semi-detached houses A and B of Inuvik, NWT, Canada, which were offered for the analysis by Arctic Energy Alliance, and Northwest Territories Housing Corporation, NWT, Canada.

Inuvik is situated at 68.36° N and 133.72° W, about 200 km north of the Arctic Circle. For the comparison of climatic conditions, Inuvik has about 9560 heating degree days below 18 °C, compared with about 4500 in Montreal.

The design thermal resistance of the house envelope exceeds the minimum requirements used at the design stage in compliance with the Model National Energy Code of Canada for Houses [36]. For instance, the thermal resistance of exterior walls is 8.1 m²K/W compared with 4.75 m²K/W in [36], and 14.1 m²K/W for roofs compared with 10.6 m²K/W in [36]. The houses are supported by space frame foundations which are proved to work well in permafrost conditions. Therefore, the floors are exposed to the outdoor air. The thermal resistance of floors was 9.3 m²K/W compared with 8.1 m²K/W in [36]. The air infiltration rate at 50 Pa pressure difference was about 50% of the maximum value of 4.55 ACH required by [36].

Measurements of the heating water flow rate, and the supply and return hot water temperatures for each house were recorded at one minute time step, from which the daily values of the space heating energy demand were calculated. The values of daily heating energy demand were almost normally distributed. The outliers, which have values outside the range $\bar{y} \pm 2 \cdot \sigma$ [37], were removed, where, \bar{y} is the average value and σ is the standard deviation. Thus, 95.5% of the available data remained in the analysis data set. The outliers were removed for each training and testing periods.

The annual measured heating energy demand is 98.1 kWh/(m²·year) for house A, and 101.7 kWh/(m²·year) for house B, with the average of 99.9 kWh/(m²·year). The total energy demand for space heating and domestic hot water was 122.4 kWh/(m²·year), which was covered by a gas-fired high efficiency condensing boiler of 41.3 kW that serves both semi-detached houses A and B. The annual measured natural gas energy use was 178.2 kWh/(m²·year). The hot water is used for the space heating through radiators, the pre-heating of outdoor air before entering the heat recovery ventilators, and for the domestic hot water.

Detailed information about the Inuvik northern sustainable house (NSH) project is presented in [38].

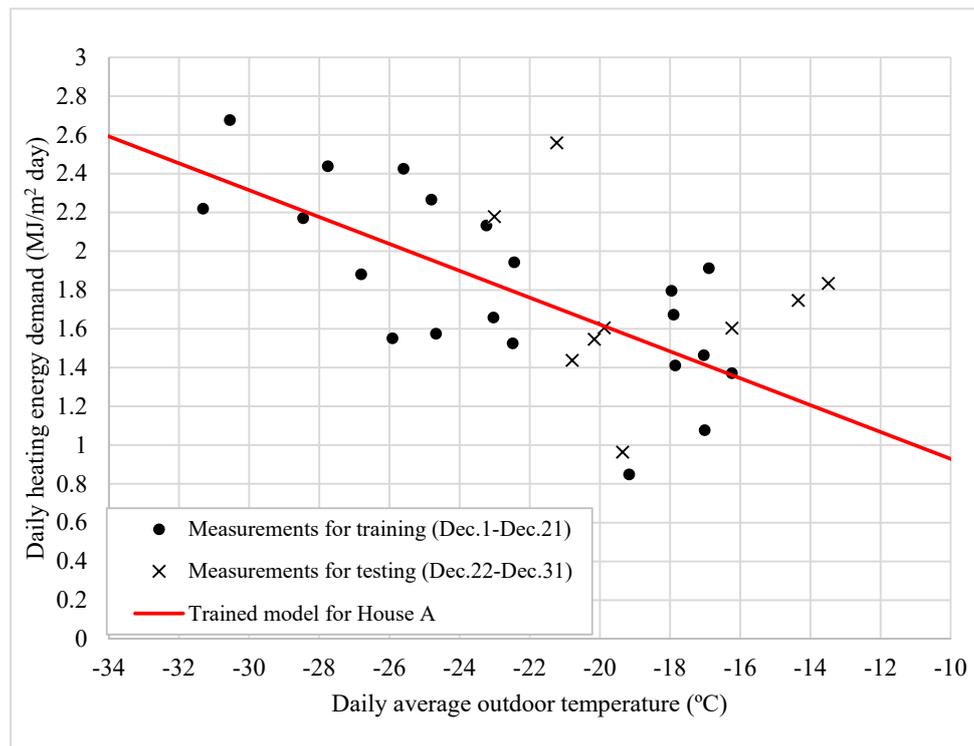
5.2. Training and Testing of Benchmarking Models

The benchmarking models of daily space heating energy demand of two semi-detached houses of Inuvik, NWT, Canada, which are used as a case study, are trained by using the measurements of December 2014, as the reference or initial training data set.

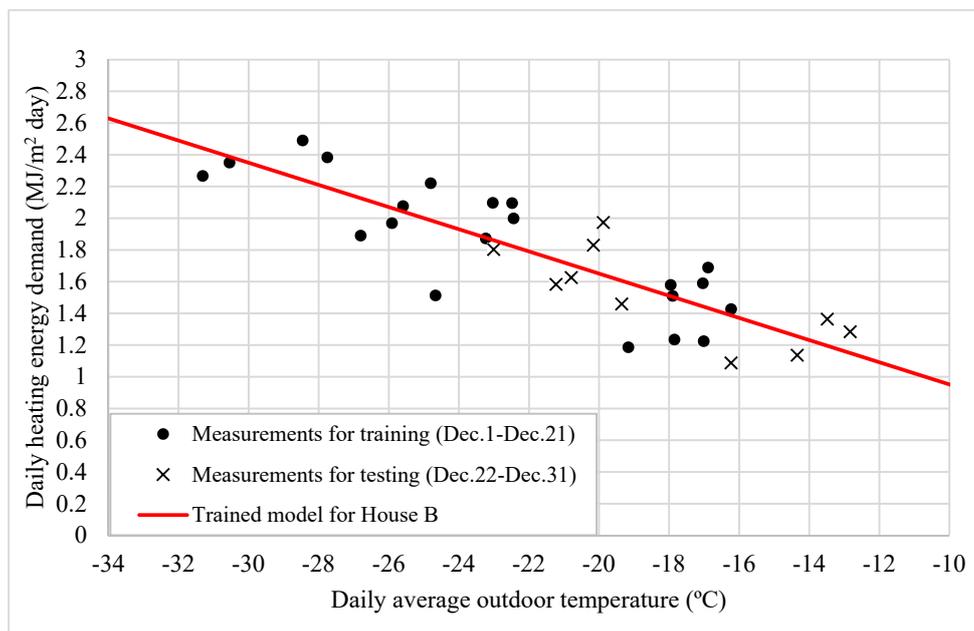
First, the daily signature of space heating energy demand (Equation (1)) of houses A and B were developed from the training data set of December 1–21, 2014 (Equations (6) and (7), and Figure 1). The two parameters, a and b , of linear model (Equation (1)) were identified by the least squares error method, which is a statistical method used to determine a line of best fit by minimizing the sum of squares created by a mathematical function.

$$E = -0.069 \cdot T_o + 0.24 \left[\text{MJ}/\text{m}^2 \text{ day} \right] \text{ for house A} \quad (6)$$

$$E = -0.070 \cdot T_o + 0.25 \left[\text{MJ}/\text{m}^2 \text{ day} \right] \text{ for house B} \quad (7)$$



(a)



(b)

Figure 1. Daily signature of space total heating energy demand used as the benchmarking model from the training data set of December 1–21, 2014: (a) house A and (b) house B.

In the case of the static window technique, these two signatures are not retrained when new data become available, for instance in January 2015. When the augmented window technique was applied, the coefficients a and b of the benchmarking models were updated every time with a new training data set, in which the previous training data set is augmented with new data of two weeks. Table 2 presents the coefficients of the daily signature that were identified from the training data set, and the statistical

indices of the difference between predictions and measurements over the testing period. For instance, a second signature was developed for house A by using the augmented training data set of December 1 to January 4.

Some relevant information was extracted directly or interpreted from the daily signature of these two houses (Table 3):

- The values of the parameter a of daily signature of both houses, obtained from the first training period of Dec 1–21, 2014, were almost equal, i.e., -0.069 and -0.070 MJ/(m² °C day). Hence, the two houses had the same sensitivity to changes in outdoor air temperature. The two houses were identical in terms of the thermal insulation of envelope, the air leakage, and the efficiency of heat recovery ventilators.
- Figure 1 shows that the daily space heating energy demand was always positive, hence one can conclude that the heating system was on for the whole heating season in both houses.
- The parameters b of daily signature of both houses, were compared for instance at $T_O = -14$ °C; they were almost equal: 0.23 MJ/(m² day) for house A, and 0.22 MJ/(m² day) for house B. Hence, on average over the initial training period the internal loads are identical in both houses. However, there was a larger dispersion of daily space heating energy demand for house A that give $CV(RMSE) = 29\%$, while for house B is only 13% . Since the measurements related to the internal gains were not available, the authors assumed, based on these results, that the daily variation of internal load was higher in house A than in house B.
- When the daily signature of house A was retrained using the augmented window technique, the parameter b increases for all training periods compared with the initial period. This could be the result of the reduction of internal heat gains, due to the reduction of occupancy and activities, from January to March 2015 compared with December 2014.
- The benchmarking models with $CV(RMSE)$ values less than 30% , calculated over the testing data set (Table 3) for both houses A and B, had an acceptable accuracy, and thus can be used for the prediction purposes. There were only three exceptions with $CV(RMSE)$ of 33% , 45% , and 46% .

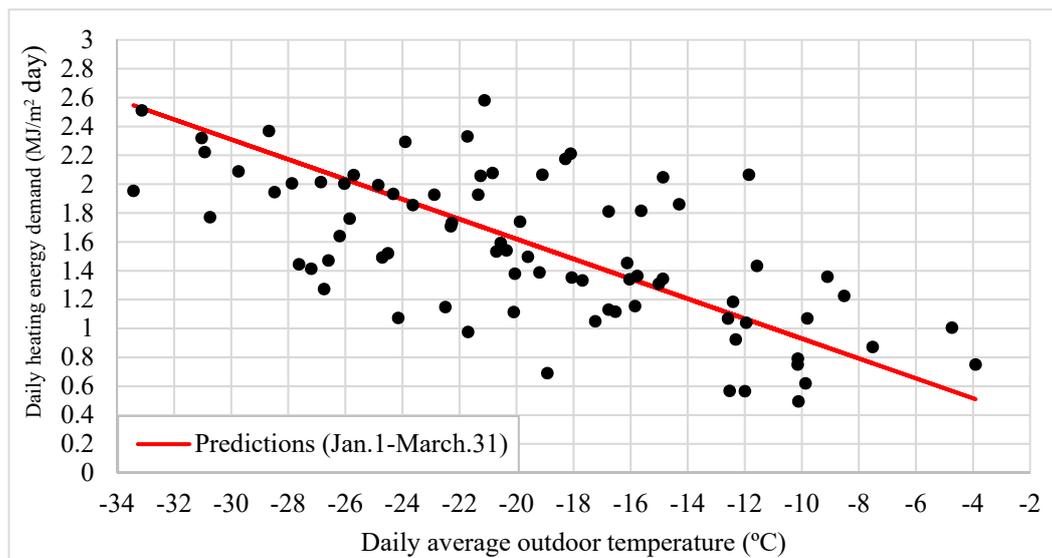
Table 3. Coefficients of the trained benchmarking models, and statistical indices of the difference between measurements and predictions over the testing period.

House	Training period	Benchmarking Model $E=a.T_o+b$		Statistical Indices over the Testing Period		
		a (MJ/m ² °C day)	b (MJ/m ² -day)	Testing period	RMSE (MJ/m ² day)	CV(RMSE) (%)
Static window technique						
A	Dec 1–21, 2014	-0.069	0.24	Dec 22–Dec 31	0.51	29
B		-0.070	0.25	Dec 22–Dec 31	0.20	13
Augmented window technique						
A	Dec 1–21	-0.069	0.24	Dec 22–Dec 31	0.51	29
	Dec 1–Jan 4	-0.054	0.64	Jan 5–Jan 11	0.46	23
	Dec 1–Jan 18	-0.044	0.89	Jan 19–Jan 25	0.38	21
	Dec 1–Feb 1	-0.042	0.92	Feb 2–Feb 8	0.16	8
	Dec 1–Feb 15	-0.042	0.91	Feb 16–Feb 22	0.44	33
	Dec 1–March 1	-0.053	0.62	Mar 2–Mar 8	0.56	45
B	Dec 1–21	-0.070	0.25	Dec 22–Dec 31	0.2	13
	Dec 1–Jan 4	-0.070	0.25	Jan 5–Jan 11	0.29	16
	Dec 1–Jan 18	-0.068	0.32	Jan 19–Jan 25	0.25	14
	Dec 1–Feb 1	-0.063	0.42	Feb 2–Feb 8	0.16	8
	Dec 1–Feb 15	-0.067	0.36	Feb 16–Feb 22	0.82	46
	Dec 1–March 1	-0.050	0.78	Mar 2–Mar 8	0.25	17

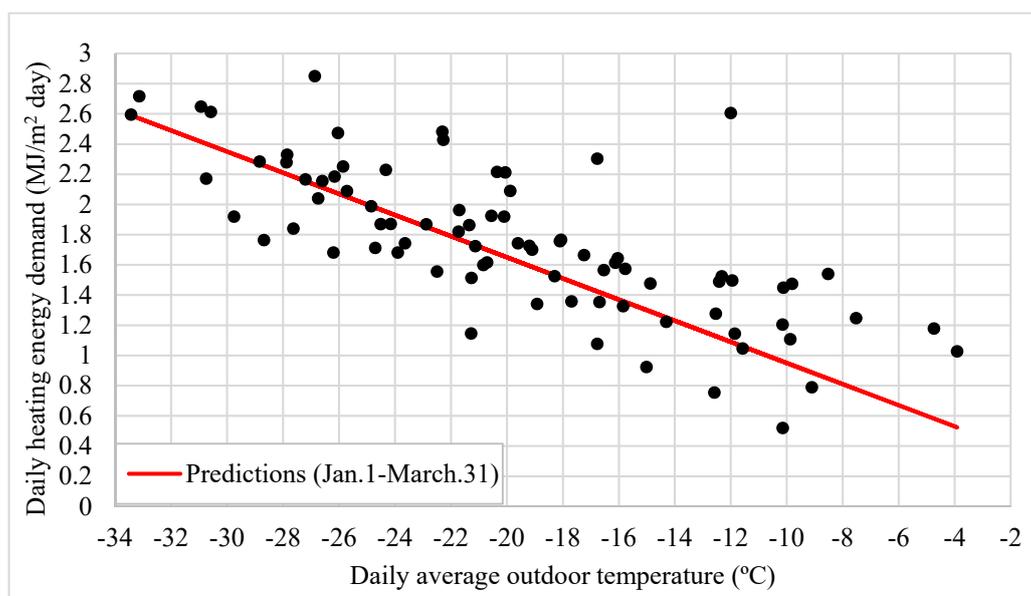
5.3. Comparison of Predicted Daily Heating Energy Demand with Measurements over the Application Period

The benchmarking models were used to detect differences between the predictions and measurements of daily space heating energy demand. Large differences might indicate changes in the operation of heating system, changes in the number of occupants and activities, and faults in sensors. This was the first step in the ongoing commissioning, which is normally followed up by the identification of causes of such changes.

Figure 2 shows an example of such a comparison, when the models trained and tested with data of December 2014 were used for the predictions from January 1 to March 31, 2015, without retraining (i.e., static window technique). The larger dispersion of measurements might have been triggered by the larger variation of internal gains in house A, compared with house B.



(a)



(b)

Figure 2. Measurements of daily total heating energy demand from January 1 to March 31, 2015 compared with the predictions of benchmarking models using static window technique: (a) house A and (b) house B.

The CV(RMSE) values of the difference between the measurements of the application period and predictions from various benchmarking models are summarized in Table 4 (static window technique) and Table 5 (augmented window technique). In most cases, the CV(RMSE) values were almost equal or greater than the corresponding values over the testing period.

Table 4. CV(RMSE) values over the testing and application periods with the static window technique.

House	Application Period	CV(RMSE) (%)	
		Testing	Application
A	Jan 1–March 31	29	27
	Jan 12–March 31		27
	Jan 26–March 31		28
	Feb 9–March 31		34
	Feb 23–March 31		36
	March 9–March 31		37
B	Jan 1–March 31	13	21
	Jan 12–March 31		21
	Jan 26–March 31		22
	Feb 9–March 31		24
	Feb 23–March 31		21
	March 9–March 31		17

Table 5. CV(RMSE) values over the testing and application periods with the augmented window technique.

House	Testing Period	Application Period	CV(RMSE) (%)	
			Testing	Application
A	Dec 22–Dec 28	Jan 1–March 31	29	27
	Jan 5–Jan 11	Jan 12–March 31	23	28
	Jan 19–Jan 25	Jan 26–March 31	21	32
	Feb 2–Feb 8	Feb 9–March 31	8	38
	Feb 16–Feb 22	Feb 23–March 31	33	41
	Mar 2–Mar 8	March 9–March 31	45	37
B		Jan 1–March 31	13	21
	Jan 5–Jan 11	Jan 12–March 31	16	21
	Jan 19–Jan 25	Jan 26–March 31	14	21
	Feb 2–Feb 8	Feb 9–March 31	8	23
	Feb 16–Feb 22	Feb 23–March 31	46	20
	Mar 2–Mar 8	March 9–March 31	17	21

In the case of house B, the CV(RMSE) values over the application period were lower than 30% (i.e., between 17% and 24%), hence, the measurements over the application period were considered to be normal compared with the training period. In the case of house A, the CV(RMSE) values were greater than 30% starting with February 9, for both static and augmented window techniques. In the absence of measurements of other physical variables, it is difficult to write an evidence-based conclusion about the reason of difference between the CV(RMSE) values obtained from the analysis of these two houses. One can speculate that the difference might be attributed to changes in the daily internal loads due to people and cooking, which were greater in house A compared with house B.

5.4. Comparison of Predicted Total Heating Energy Demand with Measurements over the Application Period

The outdoor air temperature bins, each of 1 °C width, over three months of heating season that were used with Equation (5) are presented in Table 6. For each application period, a similar table is used.

Table 6. Outdoor air temperature bins in Inuvik, NWT, Canada, from January 1 to March 31, 2015.

Temperature BIN (C)	T_o (°C)	Number of Days
−33.5, −32.5	−33	2
−32.5, −31.5	−32	-
−31.5, −30.5	−31	3
−30.5, −29.5	−30	1
−29.5, −28.5	−29	1
−28.5, −27.5	−28	4
−27.5, −26.5	−27	4
−26.5, −25.5	−26	4
−25.5, −24.5	−25	3
−24.5, −23.5	−24	4
−23.5, −22.5	−23	1
−22.5, −21.5	−22	5
−21.5, −20.5	−21	7
−20.5, −19.5	−20	5
−19.5, −18.5	−19	3
−18.5, −17.5	−18	4
−17.5, −16.5	−17	4
−16.5, −15.5	−16	5
−15.5, −14.5	−15	3
−14.5, −13.5	−14	1
−13.5, −12.5	−13	2
−12.5, −11.5	−12	6
−11.5, −10.5	−11	-
−10.5, −9.5	−10	5
−9.5, −8.5	−9	2
−8.5, −7.5	−8	1
−7.5, −6.5	−7	-
−6.5, −5.5	−6	-
−5.5, −4.5	−5	1
−4.5, −3.5	4	1
−3.5, −2.5	−3	-
Total	-	82

Tables 7 and 8 summarize the comparison of the measurements with the predicted total heating energy demand for houses A and B over different application intervals. When the models were trained with data of December 2014 and used for the prediction of rest of heating season from January to March (for static and augmented window techniques), the difference was 7.8% (house A) and 3.6% (house B). This result indicates that over a longer prediction time interval, the measurements were within an acceptable difference from the predictions of total heating energy demand.

Table 7. Measured versus predicted total heating energy demand for houses A and B with the static window technique.

Data Sets		Measured Space Heating Demand (MJ/m ²)		Predicted Space Heating Demand (MJ/m ²)		(Measured–Predicted)/Measured (%)	
Training	Prediction	House A	House B	House A	House B	House A	House B
Dec 1– Dec 21, 2014	Jan 1–March 31, 2015	136.26	156.08	147.84	150.56	−7.8	3.6
	Jan 12–March 31, 2015	113.56	133.78	126.88	129.22	−10.5	3.5
	Jan 26–March 31, 2015	96.17	116.66	101.02	102.9	−4.8	13.4
	Feb 9–March 31, 2015	67.06	84.21	73.38	74.76	−8.6	12.6
	Feb 23–March 31, 2015	43.87	55.09	49.45	50.4	−11.2	9.3
	March 2–March 31, 2015	27.7	33.87	31.15	34.8	−11	−2.7

Table 8. Measured versus predicted total heating energy demand for houses A and B with the augmented window technique.

Data Sets		Measured Space Heating Demand (MJ/m ²)		Predicted Space Heating Demand (MJ/m ²)		(Measured)–Predicted/Measured (%)	
		House A	House B	House A	House B	House A	House B
Training	Prediction						
Dec 1, 2014–Dec 21, 2014	Jan 1–March 31, 2015	136.26	156.08	147.84	150.56	−7.8	3.6
Dec 1, 2014–Jan 4, 2015	Jan 12–March 31, 2015	113.56	133.78	134.11	129.22	−15.3	3.5
Dec 1, 2014–Jan 18, 2015	Jan 26–March 31, 2015	96.17	116.66	110.85	104.82	−13.2	11.3
Dec 1, 2014–Feb 1, 2015	Feb 9–March 31, 2015	67.06	84.21	82.58	76.84	−18.8	9.6
Dec 1, 2014–Feb 15, 2015	Feb 23–March 31, 2015	43.87	55.09	56.84	52.46	−22.8	5
Dec 1, 2014–March 1, 2015	March 2–March 31, 2015	27.7	33.87	36.25	38.69	−23.6	−12.5

For both model training techniques and all application time intervals, the measurements of total heating energy demand of house A were lower than the predictions by 13.2% to 23.6%; while in the case of house B the measurements were higher than the predictions by 3.5% to 11.3%, except the last application period when it underestimated by 12.5%.

5.5. Discussion of Results

The paper presented a method for retraining a benchmarking model, the two-parameter energy signature of heating energy demand, by using machine learning techniques. The proposed method was tested by using measurements from two houses in Northern Canada, which experienced heating energy demand for the entire season from December to March. As a consequence, the two-parameter energy signature model was developed, which used the parameters a and b . The three-parameter model should be used when the base load and heating slope intersects at the outdoor air temperature, called the reference temperature T_{REF} . These are not the conditions of the case study.

The results from the initial training phase of the energy signature revealed that:

- (i) the two houses were identical in terms of the thermal insulation of envelope, the air leakage, and the efficiency of heat recovery ventilators;
- (ii) on average over the first training period the internal loads were identical in both houses; however, there was a larger dispersion of daily space heating energy demand for house A; and
- (iii) the benchmarking models with CV(RMSE) values less than 30%, can be used for the prediction purposes.

The use of trained energy signature models over the application period showed that:

- (i) In most cases of retraining with static and augmented window techniques the CV(RMSE) values were almost equal or greater than the corresponding values over the testing period;
- (ii) the analysis of CV(RMSE) values between the predictions and measurements from house B, over the application period, indicated that the measurements are considered to be normal compared with the training period; and
- (iii) there was a larger variation of daily heating energy demand of house A compared with house B.

The energy signature was used along with the outdoor air temperature bins to predict the total heating energy demand for houses A and B over different application intervals, and to compare with measurements. The analysis of results revealed that:

- (i) when the models were trained with data of December 2014 and used for the prediction of rest of heating season from January to March (for static and augmented window techniques), the measurements over a longer prediction time interval were within the predictions of total heating energy demand; and
- (ii) the predictions by the benchmarking models, which were retrained with the augmented window technique, were useful for the comparison with measurements over shorter time intervals.

6. Conclusions

The paper presented the use of machine learning for the training, retraining, and use of benchmarking models of the space heating energy demand of houses. A benchmarking gray-box model was used that combines some knowledge about the physical phenomena with the measurements. The physical-based model was the daily signature of space heating energy demand, a two-parameter model, which was developed from easily available measurements of two variables: The outdoor air temperature, and the heating energy demand. The daily energy signature was used as a benchmarking model due to its simplicity and performance.

However, any type of benchmarking model can be used with the proposed method. The use of multivariable regression models might be problematic in an ongoing commissioning system due to: (i) the unavailability of all sensors and (ii) the uncertainty due to the higher error propagation from the measurements of several sensors up to the dependent variable, compared with the case of only two physical variables as proposed in this paper.

The paper uses the static and augmented window retraining techniques, by using the measurements of space heating energy demand of two Inuvik houses from December 1, 2014 to March 31, 2015, and the corresponding outdoor air temperature.

The statistical indices over the testing period indicated that the trained benchmarking models have an acceptable accuracy, and thus can be used for the prediction purposes. The results indicated a larger variation of daily heating energy demand of house A throughout the heating season. In the case of house B, the measurements over the application period are considered to be normal compared with the training period.

The method should be applied in all possible situations, and be an integral part of intelligent building automation and control systems for the ongoing commissioning for building energy-related applications. The trained models should be used for the prediction of daily heating energy demand, and for the comparison with actual measurements. Significant differences should alert of possible faults or deterioration of system equipment and sensors, or even changes in peoples' occupation patterns. An immediate application for the proposed benchmarking method is the assessment of energy performance and potential improvements by using the measurements from the smart meters, installed in numerous houses in North America. In this case, the installation cost of dedicated monitoring system is avoided.

The storage in a database of different models, which were trained throughout the time, and corroboration with other information about the occupancy, habits, and other factors not considered in this study, might lead to better understanding of the pattern of energy usage in houses. The generalization of such energy usage patterns can be generated only through the compilation of many cases of monitoring and analysis.

Author Contributions: Conceptualization, R.Z.; methodology, R.Z.; formal analysis, B.B.; investigation, B.B.; resources, R.Z.; data curation, B.B.; writing—original draft preparation, B.B.; writing—B.B. and R.Z.; supervision, R.Z.; project administration, R.Z.; funding acquisition, R.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The financial support received by the first author from the Discovery Grant Program of Natural Sciences and Engineering Research Council of Canada (NSERC) (N00271) is acknowledged. The authors also acknowledge the financial support from the Faculty of Engineering and Computer Science of Concordia University (VE0017).

Acknowledgments: The financial support received by the first author from the Discovery Grant Program of Natural Sciences and Engineering Research Council of Canada (NSERC) is acknowledged. The authors also acknowledge the financial support from the Faculty of Engineering and Computer Science of Concordia University. This study was made possible by the collaboration of L. Robinson, S. Outlet, N. Walker, L. Azzolini from Arctic Energy Alliance, J. M. Robson from Northwest Territories Housing Corporation, and Aurora Research Institute.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

a	Base level energy consumption
b	Heating/cooling slope of energy use
BLR	Whole building cooling load ratio
CSP	Cooling set point
CV(RMSE)	Coefficient of variation of root mean squares error
E_{AC}	Whole house electricity use for air-conditioning
E_{CH}	Electric demand of chiller
G	Overall building heat loss coefficient
GHR	Global horizontal radiation
$H_O(T_{REF})$	Heating degree-days (base T_{REF})
HVC	HVAC system cooling energy use
HSP	Heating set point
HVHW	HVAC hot water energy use
LPD	Interior/exterior lighting power density
PLPD	Plug load power density
NAC	Normalized annual consumption
Q_E	Chiller cooling load
Q_{Edes}	Chiller cooling load at design conditions
R^2	Coefficient of determination
SAT	Supply air temperature
SES	South equivalent window surface area
T_{CNDS}	Condenser-water leaving temperature
T_O	Dry-bulb outdoor air temperature
T_{REF}	Reference temperature
W	Humidity ratio
WBC	Whole building cooling demand
WBEL	Whole building electricity use
WBEU	Whole building energy use
WBG	Whole building gas use
$\Delta\theta$	Temperature difference between the heating set point and sol-air temperature

References

1. Chung, W. Review of building energy-use performance benchmarking methodologies. *Appl. Energy* **2011**, *88*, 1470–1479. [[CrossRef](#)]
2. Fumo, N.; Biswas, M.R. Regression analysis for prediction of residential energy consumption. *Renew. Sustain. Energy Rev.* **2015**, *47*, 332–343. [[CrossRef](#)]
3. Fels, M.F. PRISM: An introduction. *Energy Build.* **1986**, *9*, 5–18. [[CrossRef](#)]
4. Decicco, J.M.; Dutt, G.S.; Harrje, D.T.; Socolow, R.H. PRISM Applied to Multifamily Building: The Lumley Homes Case Study. *Energy Build.* **1986**, *9*, 77–88. [[CrossRef](#)]
5. Zmeureanu, R. Assessment of the energy savings due to the building retrofit. *Build. Environ.* **1990**, *25*, 95–103. [[CrossRef](#)]
6. Guideline 14, A.S.H.R.A.E. *Measurement of Energy, Demand, and Water Savings*; American Society of Heating Refrigerating and Air-Conditioning Engineers Inc.: Atlanta, GA, USA, 2014.
7. Yu, F.W.; Chan, K.T. Energy signatures for assessing the energy performance of chillers. *Energy Build.* **2005**, *37*, 739–746. [[CrossRef](#)]
8. Monfet, D.; Zmeureanu, R. Ongoing commissioning of water-cooled electric chillers using benchmarking models. *Appl. Energy* **2012**, *92*, 99–108. [[CrossRef](#)]
9. Catalina, T.; Iordache, V.; Caracaleanu, B. Multiple regression model for fast prediction of the heating energy demand. *Energy Build.* **2013**, *57*, 302–312. [[CrossRef](#)]
10. Korolija, I.; Zhang, Y.; Majanovic-Halburd, L.; Hanby, V. Regression models for predicting UK office building energy consumption from heating and cooling demands. *Energy Build.* **2013**, *59*, 214–227. [[CrossRef](#)]

11. Zhang, Y.; O'Neill, Y.Z.; Dong, B.; Augenbroe, G. Comparisons of inverse modeling approaches for predicting building energy performance. *Build. Environ.* **2015**, *86*, 177–190. [[CrossRef](#)]
12. Hitchin, R.; Knight, I. Daily energy consumption signatures and control charts for air-conditioned buildings. *Energy Build.* **2016**, *112*, 101–109. [[CrossRef](#)]
13. Abushakra, B.; Paulus, M.T. An hourly hybrid multi-variate change-point inverse model using short-term monitored data for annual prediction of building energy performance, part III: Results and analysis (1404-RP). *Sci. Technol. Built Environ.* **2016**, *22*, 996–1009. [[CrossRef](#)]
14. Ko, J.-H.; Kong, D.-S.; Huh, J.-H. Baseline building energy modeling of cluster inverse model by using daily energy consumption in office buildings. *Energy Build.* **2017**, *140*, 317–323. [[CrossRef](#)]
15. Perez, K.X.; Cetin, K.C.; Baldea, M.; Edgar, T.F. Development and analysis of residential change-point models from smart meter data. *Energy Build.* **2017**, *139*, 351–359. [[CrossRef](#)]
16. Chen, Z.; Freihaut, J.; Lin, B.; Wang, C.D. Inverse energy model development via high-dimensional data analysis and sub-metering priority in building data monitoring. *Energy Build.* **2018**, *172*, 116–124. [[CrossRef](#)]
17. Kim, W.; Katipamula, S. A review of fault detection and diagnostics methods for building systems. *Sci. Technol. Built Environ.* **2018**, *24*, 3–21. [[CrossRef](#)]
18. Katipamula, S.; Brambley, M. Review article: Methods for fault detection, diagnostics, and prognostics for building systems—A review, part I. *Hvacr Res.* **2005**, *11*, 3–25. [[CrossRef](#)]
19. Katipamula, S.; Brambley, M. Review article: Methods for fault detection, diagnostics, and prognostics for building systems—A review, part II. *Hvacr Res.* **2005**, *11*, 169–187. [[CrossRef](#)]
20. Yu, D.; Li, H.; Yang, M. A virtual supply airflow meter in rooftop air conditioning units. *Build. Environ.* **2011**, *46*, 1292–1302. [[CrossRef](#)]
21. Sun, B.; Luh, P.; Jia, Q.; Neil, Z.O.; Song, F. Building energy doctors: An SPC and Kalman filter-based method for system-level fault detection in HVAC systems. *IEEE Trans. Autom. Sci. Eng.* **2015**, *11*, 215–229. [[CrossRef](#)]
22. Bynum, J.D.; Claridge, D.E.; Curtin, J.M. Development and testing of an automated building commissioning analysis tool (ABCAT). *Energy Build.* **2012**, *55*, 607–617. [[CrossRef](#)]
23. Capozzoli, A.; Lauro, F.; Khan, I. Fault detection analysis using data mining techniques for a cluster of smart office buildings. *Expert Syst. Appl.* **2015**, *42*, 4324–4338. [[CrossRef](#)]
24. Seem, J.E. Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy Build.* **2007**, *39*, 52–58. [[CrossRef](#)]
25. Amasyali, K.; El-Gohary, N.M. A review of data-driven building energy consumption prediction studies. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1192–1205. [[CrossRef](#)]
26. Djukanovic, M.R.S.; Babic, B.; Sobajic, D.J.; Pao, Y.H. A neural-based short term load forecasting using moving window procedure. *Electr. Power Energy Syst.* **1995**, *17*, 391–397. [[CrossRef](#)]
27. Mohammed, O.; Park, D.; Merchant, R.; Dinh, T.; Tong, C.; Azeem, A.; Farah, J.; Drake, C. Practice experience with an adaptive neural networks short-term load forecasting system. *IEEE Trans. Power Syst.* **1995**, *10*, 254–265. [[CrossRef](#)]
28. Charytoniuk, W.; Chen, M.S. Very short-term load forecasting using artificial neural networks. *IEEE Trans. Power Syst.* **2000**, *15*, 263–268. [[CrossRef](#)]
29. Yang, J.; Rivard, H.; Zmeureanu, R. On-Line Building Energy Prediction Using Adaptive Artificial Neural Networks. *Energy Build.* **2005**, *37*, 1250–1259. [[CrossRef](#)]
30. Chae, Y.T.; Horesh, R.; Hwang, Y.; Lee, Y.M. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. *Energy Build.* **2016**, *111*, 184–194. [[CrossRef](#)]
31. Domingos, P. A Few Useful Things to Know about Machine Learning. *Commun. ACM* **2012**, *55*, 78–87. [[CrossRef](#)]
32. Jacobsen, F.R. Energy signature and energy monitoring in building energy management systems. In Proceedings of the Clima 2000 World Congress on Heating, Ventilating and Air-Conditioning, Copenhagen, Denmark, 25–30 August 1985; Volume 3, pp. 25–31.
33. Deeble, V.C.; Probert, S.D. Straight-line correlations for annual energy-consumption predictions? *Appl. Energy* **1986**, *25*, 23–39. [[CrossRef](#)]
34. International Energy Agency. *Source Book for Energy Auditors*; Lyberg, M.D., Ed.; Energy Conservation in Buildings and Community Systems Programme, Annex XI; Energy Auditing; Swedish Council for Building Research: Stockholm, Sweden, 1987; Volume 1, pp. 38–42.

35. Kaplan, M.B.; McFerran, J.; Jansen, J.; Pratt, R. Reconciliation of a DOE-2.1C model with monitored end-use for a small office building. *Ashrae Trans.* **1990**, *96*, 981–993.
36. National Resources of Canada. *Model National Energy Code of Canada for Buildings*; NRCC: Ottawa, ON, Canada, 1997.
37. ASHRAE Guideline 2-2010 (RA 2014). *Engineering Analysis of Experimental Data*; American Society of Heating Refrigerating and Air-Conditioning Engineers Inc.: Atlanta, GA, USA, 2014.
38. CMHC. *Full Report—Inuvik Northern Sustainable House Energy Consumption Monitoring: Year 1 & 2 Results*; Canada Mortgage and Housing Corporation: Ottawa, ON, Canada, 2017.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).