Detection and interpretation of anomalies in building energy use through inverse modeling

Burak Gunay, H.; Shen, Weiming; Newsham, Guy; Ashouri, Araz

Publisher's version / Version de l'éditeur:
https://doi.org/10.1080/23744731.2019.1565550
Science and Technology for the Built Environment, 25, 4, pp. 488-503, 2019-04-09

Detection and interpretation of anomalies in building energy use through inverse modelling

Abstract: This paper presents a study in which three inverse modelling techniques were applied to hourly heating and cooling load data extracted from 35 office buildings in Ottawa, Canada. These modelling techniques were: three-parameter change point models, regression trees, and artificial neural networks. The change point models were trained with outdoor temperature data; whereas the other two models were trained with four regressors: outdoor temperature, wind speed, horizontal solar irradiance, and a binary workhours indicator. The correlations among the change point model parameters of individual buildings were analyzed. The sensitivity of heating and cooling load intensities to the four regressors was examined. The models were used to identify several types of energy use anomalies. The anomalies detected by different modelling techniques were generally in agreement. The results indicate that nearly half of the buildings did not have effective afterhours schedules to save energy. In all but three buildings, the cooling change point temperature was lower than the heating change point temperature – indicating a simultaneous heating and cooling problem. Moreover, a few buildings with anomalies potentially related with high air-infiltration or over-ventilation, high thermal conductance, and high solar heat gains during summer were identified.

Keywords: Inverse modelling; Anomaly detection; Change point models; Regression trees; Artificial neural networks

1. Introduction

Efficient delivery of heating and cooling services in commercial and institutional buildings requires many interconnected systems and devices to function harmoniously. Consequently, maintaining, controlling, and retrofitting these systems and devices becomes a complex operations research problem. It is estimated that poorly maintained, degraded, and improperly controlled building equipment wastes 15 to 30% of the energy used in commercial buildings (Katipamula and Brambley 2005a; Katipamula and Brambley 2005b).

Out of sight devices fail at random instances during the service life of a building (e.g., faulty dampers, valves, sensors) (Kim and Katipamula 2017). Aside from physical faults, the systems are sometimes set to operate at inappropriate setpoints and schedules. For example, Abdelalim et al. (2017) demonstrated that a simple controls programming mistake in a heat recovery wheel wasted about 20% of the heating energy used in an academic office building. Perhaps more alarming than the fault itself is the fact that the programming mistake remained undetected for over five years, despite an ongoing commissioning contract. In another study, Gunay et al. (2017b) identified that inappropriate economizer programming increased the cooling loads by 25%. Beyond issues related with the maintenance and control of the existing building systems, retrofits and equipment upgrades can become cost-effective in time as technologies emerge (Ferreira et al. 2016). However, a tangible research challenge is to identify the top energy saving opportunities in a reliable, low-cost, and non-invasive manner.

Energy auditing is a process in which an energy consumption baseline is established and recommendations to improve energy efficiency are developed (Kelsey and Pearson 2011; NRCan 2011). The process often involves walk-through surveys, and collection and analysis of building level energy use data. In some cases, the auditors can collect short-term equipment specific data, and develop calibrated system or building-level energy models. ASHRAE categorizes the energy auditing process into three groups

based on the level of complexity (ASHRAE 2011). Energy audits rely on an expert's knowledge in building systems; thus, they can be labour-intensive and cost-prohibitive, particularly for small-to-medium-sized commercial buildings. In addition, energy audits can be invasive, as walk-through surveys and data collection from occupied spaces are needed. Thus, energy audits are typically carried out much less frequently than once per year (Friedman et al. 2003; Mills 2011). Over time, energy savings through the implementation of the findings of an energy audit often reduce due to changes in operational characteristics and new faults in building systems (Bourassa et al. 2004; Cho 2008). For example, temperature setpoints and operating schedules can be overridden following occupant complaints, or the sensors and actuators involved in the economizer, heat recovery, and demand control ventilation programs can malfunction. This situation highlights the necessity for analytical methods that continuously monitor a building's energy performance.

Inverse modelling can be used as an analytical tool to continuously monitor the energy performance of a building and to identify and interpret its energy use anomalies. This process, relying on automatically-recorded building systems data rather than on-site inspection, is sometimes generically referred to as "remote auditing". Considering the challenges in acquiring as-built geometry and thermophysical information from existing buildings, greybox and blackbox model formalisms have been more commonly used in inverse modelling than the physics-based models in the reviewed literature (Pérez-Lombard et al. 2009).

The greybox models have been typically built as thermal network models (Zhou et al. 2008; Fux et al. 2014) in which heat transfer and storage processes are mimicked through a number of thermal resistances and capacitances. For example, the model proposed by Zhou et al. (2008) consisted of fourteen parameters – eight resistances and six capacitances. The parameters were estimated by using the measured data and assumptions for indoor and outdoor environmental variables such as the outdoor solar irradiance, indoor and outdoor temperatures, casual heat gains, infiltration, and ventilation schedules. The greybox models can represent the space heating and cooling loads of a building, when they are designed at appropriate complexity and with sensor inputs that can characterize the thermal disturbances. Consequently, they can be built with shorter data records (or shorter training periods for online algorithms) than blackbox models (Braun and Chaturvedi 2002). However, greybox modelling often requires high resolution sensor data from the indoor environment (e.g., indoor temperature, setpoints and schedules, humidity, occupancy). Thus, greybox modelling may not be a suitable method for buildings in which a permanent sensor data archival system is not available. It is worth noting that in most commercial buildings, the sensor data from building automation systems are stored temporarily (typically for less than a week) within the buffer memory of each controller as trendlogs (Gunay and Shen 2017).

Non-physics-based inverse model formalisms (i.e., blackbox or regression models) used in characterizing the energy performance of a building include the change-point regression models (Zhang et al. 2015), the bin method (Thamilseran and Haberl 1995; Golden et al. 2017), multiple linear regression models (Thomas Ng et al. 2008), artificial neural network models (Krarti et al. 1998; Karatasou et al. 2006; Gunay et al. 2017a), support vector machines (Dong et al. 2005; Li et al. 2009), Gaussian mixture regression models (Srivastav et al. 2013), clustering (Jalori and Reddy 2015a), and decision tree models (Yu et al. 2010). As an alternative to using a single model formalism, a family of models can be trained with the same dataset to remedy the deficiencies of each other and to improve the overall performance. This modelling

technique is called the ensemble learning method; and in the context of this paper, it was employed by Wang and Srinivasan (2017) to predict the energy use patterns in buildings.

Non-physics-based inverse models input several weather variables (e.g., outdoor temperature, humidity ratio, solar irradiance, wind speed, precipitation, sky clearness) and categorical variables (e.g., time of day, day of week, workhours / afterhours, data from plug load and lighting submeters). Then, they map these input variables onto monthly, daily, hourly, or subhourly heating or cooling energy use data – often normalized against the floor area. It is worth noting that these data types are typically available in most commercial buildings, regardless of the availability of a permanent data archiver for the building automation system (Gunay and Shen 2017). Therefore, non-physics-based inverse modelling can be a viable solution in most commercial buildings.

In the reviewed literature, inverse modelling with building energy use data has been employed for several reasons. Establishing an energy use baseline was one of the primary reasons behind inverse modelling (Abushakra 1997; Singh et al. 2014; Jalori and Reddy 2015b). Abushakra (2001) demonstrated that only two weeks' worth of data collected at hourly intervals can be sufficient for baseline model development; and introduced an algorithm to systematically search for the best two-weeks to formulate a multivariate linear regression baseline model. Subsequently, based on the findings of the ASHRAE Research Project 1404, Singh et al. (2014) investigated the minimum length of the energy use data record needed to develop a change point model which inputs only the outdoor temperature as a regressor. By using daily energy use data from three buildings, they identified that two to three months' worth of data can be sufficient to develop a change point model. Singh et al. (2013) presents a hybrid inverse modelling approach which uses monthly utility bill data in tandem with a short dataset of monitored daily energy use, internal loads, and weather variables. They identified that with this hybrid approach only one month's worth of data is often sufficient to establish an accurate baseline model. Abushakra and Paulus (2016a, 2016b, 2016c) also reported from the findings of the ASHRAE Research Project 1404. They studied multivariate, three- to five-parameter change point models which input weather variables such as outdoor temperature, humidity, occupancy, and submetered lighting and equipment loads. Some of these variables are multicollinear. For example, occupancy and lighting / equipment energy use, and outdoor temperature and humidity ratio are expected to be closely related. To avoid overfitting due to this multicollinearity, Abushakra and Paulus (2016a, 2016b, 2016c) recommended using the occupancy variable or the lighting and equipment loads variable individually – not together. Furthermore, they identified that despite the collinearity between outdoor humidity and temperature it can be beneficial to include outdoor humidity in baseline models which characterize the cooling energy use. Abushakra and Claridge (2001) examined the viability of filtering out the impact of occupant-driven loads from the energy use data, so that the models defining the energy use baseline can account for the occupancy variable. Using a synthetic dataset generated via the building performance simulation tool EnergyPlus, Paulus et al. (2015) developed an algorithm to automate the process of change point model development. The algorithm systematically searches for the best change point temperatures and the number of change points.

Quantifying energy savings after a retrofit was another common reason behind inverse modelling in the reviewed literature (Thamilseran and Haberl 1995; Haberl and Thamilseran 1996; Liu et al. 2011b; Jalori and Reddy 2015b; Carpenter et al. 2018). This use case requires an inverse model to establish an energy use baseline before and after a retrofit. Jalori and Reddy (2015b) list other application fields for inverse

modelling as the short-term load forecasting for utility demand response management (Chae et al. 2016), condition monitoring (Liu et al. 2011a), and fault detection (Jacob et al. 2010). They proposed a unified inverse modelling methodology that combines various application-specific model formalisms (Jalori and Reddy 2015b; Jalori and Reddy 2015a).

Although the inverse modelling field has been rapidly progressing, there are several gaps in the reviewed literature. In most cases, the model development relied on limited datasets from only a few buildings or synthetically generated data from simulation tools (Jalori and Reddy 2015a; Paulus et al. 2015). Application of the inverse modelling techniques with datasets from large building clusters would enable comparative studies among models trained for different buildings (e.g., the relationship between the heating and cooling change point temperatures in different buildings). The focus of the model development process was to improve the predictive accuracy. However, each modelling method has different strengths. For example, a decision trees model is akin to deriving rules which are suitable for human interpretation (Yu et al. 2010). In contrast, a neural network model, despite its superior predictive accuracy, is not directly suitable for human interpretation – as the inputs are mapped onto the outputs through non-linear activation functions. Beyond accuracy and ease of interpretation, an inverse model can be structured not only to detect energy use anomalies but also to help diagnose them (Wang and Srinivasan 2017). For example, if the heating and cooling load intensities do not change between workhours and afterhours, this can be interpreted as a result of an inappropriate HVAC operation schedule. In summary, the focus on inverse modelling has been on establishing an energy use baseline, predicting the energy use intensity, and estimating energy savings after a retrofit – rather than characterizing the operation for the identification of the energy use anomalies.

This paper presents a study conducted on the annual hourly heating and cooling load data gathered from 35 office buildings in Ottawa, Canada. An overview of the buildings and dataset is provided in Section 2. Three different inverse model formalisms are examined: change point models, regression trees, and artificial neural network (ANN) models. These models are assessed in terms of their ability to extract operations-related information for the identification of energy intensive anomalies. An overview of these model forms is presented in Section 3. In Section 4, a transverse study is carried out to analyze the models developed for the 35 buildings comparatively. A systematic inverse modelling methodology is demonstrated upon a limited number of data types. The limitations of inverse modelling in the operational decision-making process are identified, and future work recommendations are developed.

2. Overview of the buildings and dataset

The analysis presented in this paper was conducted on the heating and cooling load data gathered from 35 federal office buildings in Ottawa, Canada. The heating and cooling to these buildings were provided through three different district heating and cooling plants. For heating, two of these plants contained natural gas boilers generating steam; the third plant had natural gas boilers producing hot water. For cooling, the three plants had centrifugal water-cooled chillers supplying chilled water. The hot water / steam and chilled water were circulated through underground loops; and individual buildings extracted thermal energy through heat exchangers. The heating and cooling energy consumed by each building was metered separately through flow meters and temperature sensors. The calibration of these metering devices was maintained annually through a third-party commissioning contract. The thermal energy supplied to each building was logged at hourly intervals to a database. For the analysis presented in this paper, the authors extracted the data gathered in 2015 only.

Table 1 provides an overview of the characteristics of these buildings. The floor area, vintage, and the window-to-wall ratio (WWR) were extracted from recent energy audit documents. The heating and cooling load intensity values were computed from the metered data. The buildings were of vastly different age, size, and WWR. Their annual heating and cooling energy use intensities per square meter ranged from 78 kWh/m^2 to 1090 kWh/m^2. The space heating and cooling energy use intensities did not exhibit any significant correlation with WWR or building age (p-value >> 0.05). Along with the heating and cooling load data from individual buildings, the outdoor temperature, horizontal solar irradiance, and wind speed data were collected from a local weather station at 15-min intervals during 2015. Figure 1 presents the distribution of these data records. The weather data exhibited a wide range of outdoor conditions, typical for this location – ensuring that the relationships between these weather data records and the concurrent heating / cooling data can be identified via inverse modelling.

Table 1: An overview of the characteristics of the buildings. The averages and the standard deviations of the values for each column are highlighted in bold font.

Building	Floor area (1000 m^2)	Heating energy use intensity (kWh/m^2-yr)	Cooling energy use intensity (kWh/m^2-yr)	Vintage	WWR
B1	11	244	172	1847	0.2
B2	11	200	73	1911	0.6
B3	6	154	91	1924	0.2
B4	8	156	49	1930	0.2
B5	6	101	90	1919	0.4
B6	149	146	72	1977	1.0
B7	8	212	141	1911	0.6
B8	14	220	91	1913	0.2
B9	75	227	114	1974	0.6
B10	12	220	183	1866	0.2
B11	26	193	125	1949	0.4
B12	7	242	101	1889	0.4
B13	60	55	71	2015	0.8
B14	72	18	81	1990	0.8
B15	32	281	165	1967	0.2
B16	62	159	190	1973	0.4
B17	13	327	253	1940	0.4
B18	31	89	56	1960	0.4
B19	12	86	80	1961	0.6
B20	63	83	55	1970	0.4
B21	8	334	149	1990	0.2
B22	21	160	127	1990	0.2
B23	17	142	50	1962	0.4
B24	8	721	369	1965	0.2
B25	34	458	101	1978	0.4
B26	7	133	71	1952	0.4
B27	17	73	17	1957	0.2

B28	12	170	34	1960	0.4
B29	33	79	82	1970	0.4
B30	61	33	56	1979	0.4
B31	12	815	181	1954	0.2
B32	19	54	24	1965	0.4
B33	41	70	70	1974	0.4
B34	4	271	55	1955	0.2
B35	39	121	67	1952	0.4
Average	**29**	**201**	**106**	**1951**	**0.4**
Standard deviation	**29**	**168**	**69**	**35**	**0.2**

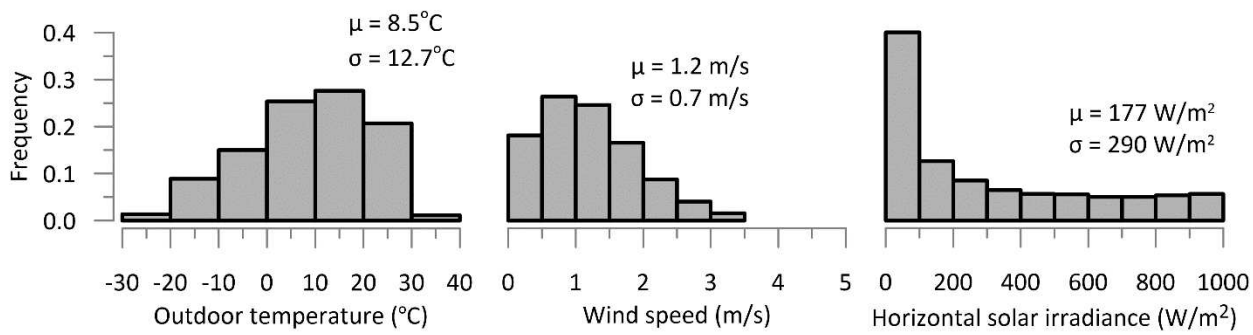

Figure 1: The distribution of the outdoor temperature, wind speed, and horizontal solar irradiance data records. The symbols μ and σ stand for the mean and the standard deviation, respectively. The data in these distributions are for all-hours (not for workhours only).

3. Models

At increasing model complexity, three different inverse model formalisms were examined: univariate change point models, regression trees, and ANN models. The choice of the change point models and regression trees formalisms was based on their ease of interpretation and the ability to derive rules from their outputs, whereas the selection of ANNs was based on their ability to capture the non-linearities in heating and cooling energy use patterns accurately.

3.1. Change point models

Three-parameter and four-parameter univariate change point regression models were developed for heating and cooling load intensities (Q_{htg} or Q_{clg}) in each of the 35 buildings. Figure 2 presents a schematic representation of these models. The only regressor used with these models was the outdoor temperature (T_{out}). The main difference between these two regression models is that the four-parameter model defines a secondary slope instead of a constant above the heating change point temperature ($T_{b,htg}$) and below the cooling change point temperature ($T_{b,clg}$). The unknown parameters of the models were estimated by using the least-squares method (using Matlab's *fitlm* function) and an iterative search of the best change point temperature. The change point temperatures for heating and cooling were separately shifted at one-degree Celsius increments between 0°C and 25°C, and the models with change point temperatures which yield the smallest coefficient of variation of the root-mean-square error (RMSE) was selected. Note that the one-degree Celsius increment size was deemed appropriate after a preliminary sensitivity analysis

revealing that the change point temperatures in individual buildings varied over a wide temperature range (>15°C). Subsequently, the coefficient of variation of the RMSE of the selected three- and four-parameter change point models for each building were compared. It was determined that the three- and four-parameter change point models performed similarly; and the secondary slopes of the four-parameter models for heating load intensity were nearly zero in the sample studied. A potential interpretation for the absence of a significant secondary slope is that the data used in model development were the heating and cooling energy extracted from a district heating and cooling energy system rather than the energy used by the plant equipment. Consequently, the data used in this study were not affected by the non-linearities in plant equipment efficiencies at part load conditions. Therefore, the results and discussion presented subsequently do not cover the four-parameter change point models.

With the three-parameter change point model, each building was defined with six parameters – three for heating and three for cooling load intensities. Two of these parameters were the slope parameters defining the rate of change in the heating ($x_{2,htg}$) and cooling ($x_{2,clg}$) load intensities in response to a one-degree Celsius increase in outdoor temperature. Two of them were the change point temperatures for heating ($T_{b,htg}$) and cooling ($T_{b,clg}$). The last two parameters were the y-intercept values at the change point temperatures for the heating and cooling load intensities ($x_{1,htg}$, $x_{1,clg}$). In functional form, this simple regression model can be represented as follows:

$$\begin{aligned} Q_{htg} &= x_{1,htg} + x_{2,htg}\left(T_{out} - T_{b,htg}\right)^{-} \\ Q_{clg} &= x_{1,clg} + x_{2,clg}\left(T_{out} - T_{b,clg}\right)^{+} \end{aligned} \quad (1)$$

where the superscript plus ($^{+}$) and minus ($^{-}$) notations indicate that the values in the parenthesis shall be set to zero when they are negative and positive, respectively. Note that this notation style follows ASHRAE Guideline 14 (2014)'s definition of change point models. The relationships among the six parameters ($x_{1,clg}$, $x_{2,clg}$, $T_{b,clg}$, $x_{1,htg}$, $x_{2,htg}$, $T_{b,htg}$) defining the heating and cooling energy performance of the 35 buildings were examined. The outlier buildings were identified, and the underlying reasons behind their anomalous behaviour were interpreted. The results of this cross-sectional analysis are presented in Section 4.

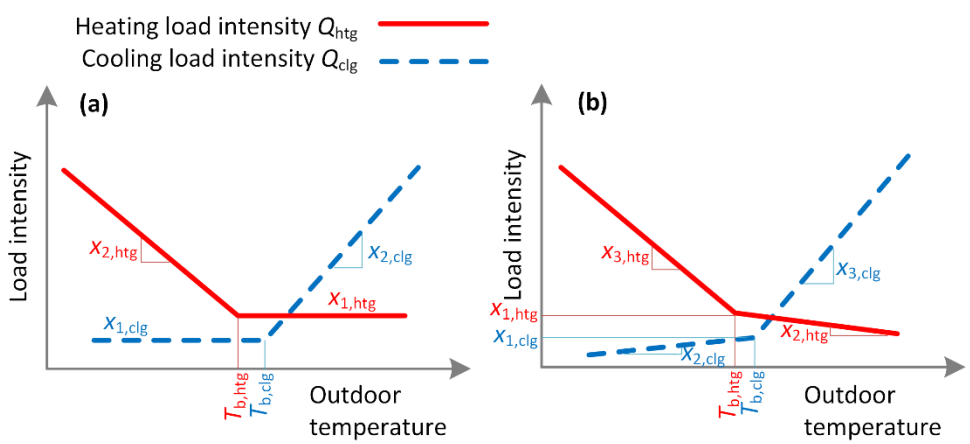

Figure 2: A schematic representation of univariate (a) three-parameter and (b) four-parameter change point models.

3.2. Regression tree models

The second model formalism examined in this study is regression trees. The regression tree models of this study were developed with four potential regressors: the outdoor temperature (T_{out}), the wind speed

(S_{wd}), the solar irradiance (Q_{sol}), and the binary workhours indicator (B_{wh}). The binary workhours indicator takes the value "1" between 8 am and 5 pm on weekdays, and otherwise it takes the value "0". The purpose of this variable was to detect whether or not there is a noticeable change in the heating and cooling energy use patterns between nominal workhours and afterhours. Considering that these buildings are office buildings which are mainly used by government employees with standard occupancy schedules, operational strategies such as equipment on-off and temperature setback are expected to introduce changes in the heating and cooling energy use patterns, which can be filtered out through appropriate inverse models. It is worth noting that this simplification was needed to represent the effect of occupant-driven thermal loads on the heating and cooling loads in the absence of sensors to estimate the occupancy counts in these buildings. Arguably, inverse models of higher accuracy can be trained with the use of occupancy count estimation technologies.

The regression tree models for each of the 35 buildings were developed by employing a binary recursive partitioning procedure – the CART algorithm (Breiman et al. 1984) as implemented in Matlab's *fitrtree* function. Figure 3 presents an illustrative example of this model form for building B30. In this example, there are nine branch nodes (i.e., representing decision splits) and ten leaf nodes (i.e., ten potential predicted outcomes for the heating load intensities). The algorithm automatically makes decision splits in recursion and keeps branching out until a stopping criterion is satisfied. The decision splits at each branch node are made with respect to the regressors which yield the greatest information gain, whereby the information gain is defined as the difference between the entropy (a metric to quantify level of uncertainty) before and after a decision split. Thus, the regressors of the greatest importance are expected to appear at the top of the tree – and likely to appear multiple times. Insignificant regressors may be used infrequently at lower tier decision splits – or they may not be used at all. As in the example shown in Figure 3, six of the decision splits are at different outdoor temperature levels, and the remaining three splits are at different solar irradiance levels. There are not any splits across the wind speed and the binary workhours indicator. Thus, the model form provides insights into the hierarchical importance of predictors.

The trees are expected to grow to a maximal size without the use of stopping rules. To create parsimonious trees (i.e., to avoid overfitting), different statistical measures were taken in this study. Firstly, the splitting was terminated when there were less than 24 hours' worth of data remaining at a given node. Secondly, the sections of a tree that provide no additional information were removed – an action also known as pruning. This is the reason behind the asymmetry of the tree shown in Figure 3. Note that in Figure 3 some of the branch nodes split four times before reaching to a leaf node, whereas some of the decision paths ended after three splits. Lastly, for each building, a repeated random subsampling cross-validation approach was employed. The maximum number of decision splits were shifted from 1 to 50, and the ones providing the best predictive performance (in terms of RMSE) for the validation sets were selected. The regression tree model of a building's heating / cooling load intensity stopped splitting further once its maximum permissible number of decision splits was reached.

Further information on classification and regression trees can be found elsewhere (Reddy 2011; Witten et al. 2016). In Section 4, the sensitivity of heating / cooling load intensities to the studied regressors was examined by using the regression tree models developed for each building. The outlier buildings in terms of their sensitivity to different regressors were identified, and the physical meaning of these sensitivity results was discussed.

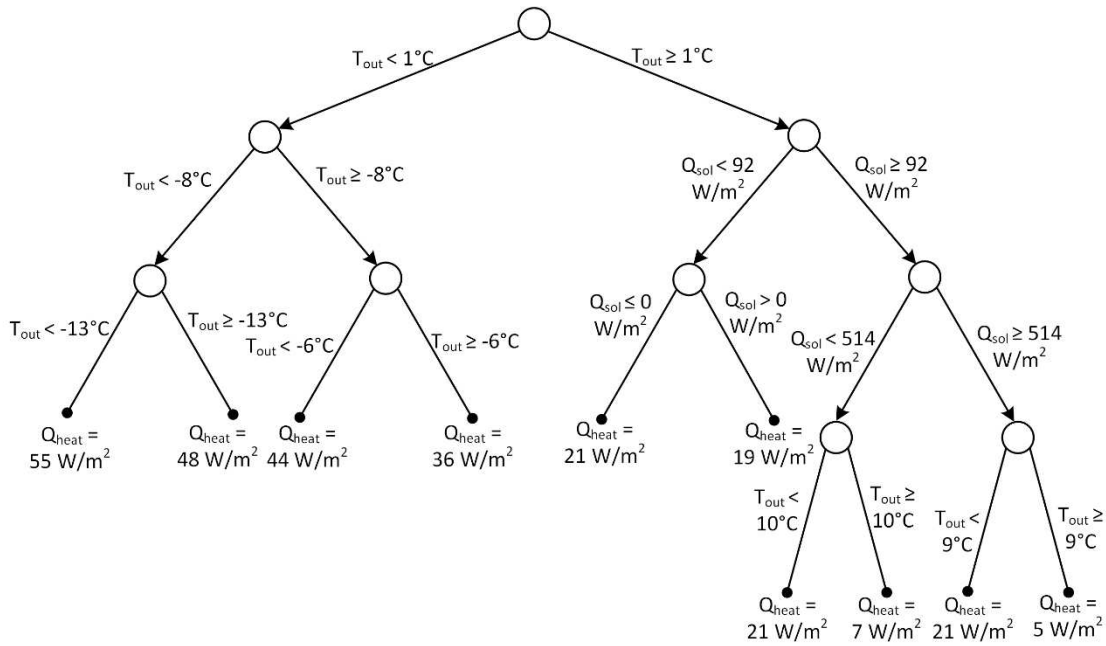

Figure 3: An illustrative example for the regression tree models. The illustrative example shown in the figure is developed to characterize the heating load intensities of building B30.

3.3. Artificial neural network models

The ANN models for each of the 35 buildings input the same four regressors as the regression trees: the outdoor temperature (T_{out}), wind speed (S_{wd}), horizontal solar irradiance (Q_{sol}), and the binary workhours indicator (B_{wh}). However, unlike the regression trees, the ANN models map these inputs onto the heating and cooling load intensities in a non-linear fashion. The ANN models were two-layer feedforward neural networks (see Figure 4). The hidden layer of the ANN models was designed with sigmoid activation functions, whereas the output layers were designed with linear activation functions. The parameters of the ANN models were estimated by using the Levenberg-Marquardt backpropagation method. Training of a neutral network was stopped when a minimum performance gradient is reached such that the parameter estimates no longer change or when repetitions exceed a certain number. After investigating the sensitivity of the models to these stopping criteria, the minimum performance gradient was selected as 10^{-6} and the maximum number of iterations was selected as 1000. For each input scenario, the number of nodes in the hidden layer of the ANN models were varied from one to twenty; and the number of nodes achieving the smallest cross-entropy were selected. The ANN models were developed and assessed by using Matlab functions *fitnet*, *perform*, and *train*. The ANN models had the following functional form:

$$Q_{\text{htg}} \text{ or } Q_{\text{clg}} = f\left(\phi_o + \sum_{j=1}^{n} \phi_j \cdot g\left(\phi_{oj} + T_{out}\phi_1 + S_{wd}\phi_2 + Q_{sol}\phi_3 + B_{wh}\phi_4\right)\right) \quad (2)$$

where φ's are the parameters to be estimated, g is a sigmoid activation function, f is a linear activation function, and n is the number of nodes in the hidden layer.

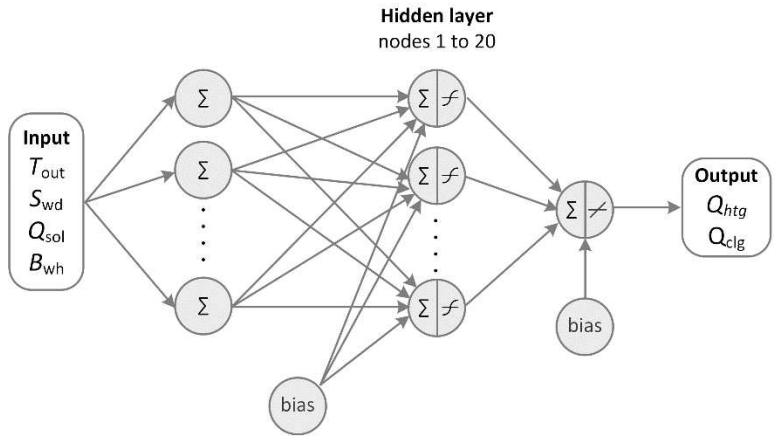

Figure 4: The structure of the ANN models of this study.

4. Results and discussion

Figure 5 presents the distribution of the change point models developed for the 35 buildings. The results indicate that there is a large variation in the change point models developed for individual buildings. It was noted that the same three buildings were outliers for both heating and cooling energy use intensities. Consequently, the mean heating and cooling energy use intensities of these buildings were considerably higher than their median. Recall that these three buildings (B31, B24, and B21) were small (under 12,000 m^2) low-rise (under 3 storeys) buildings with WWRs of about 20%. They were constructed prior to 1990s.

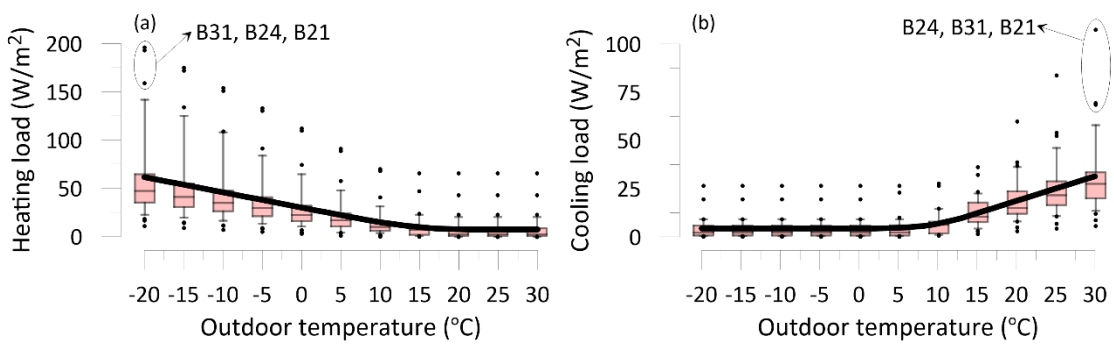

Figure 5: The distribution of the change point models developed for the 35 buildings. The solid line indicates a change point model for the mean of the model parameters for individual buildings. The whiskers enclose 5th and 95th percentiles.

Figure 6 presents the distribution of the change point model parameters for heating and cooling load intensities of the 35 buildings – i.e., the y-intercepts at the change point temperatures, the slopes, and the change point temperatures. Note that the intercept values indicate the minimum expected heating and cooling energy use intensities independent of the outdoor temperature. As shown in Figure 6.a, the change point models for the cooling energy use in buildings B7, B17, and B24 had intercept values greater than 9 W/m^2. This observation can be interpreted as these buildings tend to require significantly more cooling energy per unit floor area than the other buildings during the heating season. A potential reason behind this operational anomaly is inappropriate thermal zoning. If zones with significantly different thermal loads are served by the same AHU, the return air temperature at the AHU may not be high enough to trigger the economizer cycle – even though some of the zones need cooling. Another plausible explanation is the absence of a functional air-side economizer program. While it is logical that the core

spaces in a large building require cooling even during the heating season, the need for cooling should be mostly offset by the cool outdoor air – in lieu of chillers.

The intercepts of the change point models developed for heating load intensities indicate that B31, B24, and B15 are the outliers in terms of their high cooling season heating energy demands. A potential explanation for this operational anomaly is leaking or stuck open heating / reheat coil valves. Another explanation is inappropriate automation control logic (e.g., high temperature setpoints for heating), or a manual operator override.

The slopes of the change point models for heating and cooling load intensities both indicate that B24, B31, and B21 are the outliers in terms of the impact of outdoor temperatures on their heating and cooling energy performance. In these three buildings, a one-degree Celsius increase in the outdoor temperatures tends to decrease the heating energy use intensity by 4 to 5 W/m^2 and increase the cooling energy use intensity by 3 to 5 W/m^2. Note that the median value for the slope parameters was about 1 W/m^2 for both heating and cooling. Plausible explanations behind these anomalies are poor envelope performance (e.g., high air infiltration or high thermal conductance) or over-ventilation (i.e., excessive use of outdoor air in ventilation).

It is also worth noting that the best-fits for the hourly heating and cooling load data were achieved at very different change point temperatures in individual buildings. The median change point temperatures for cooling and heating were 8°C and 16°C, respectively. This can be interpreted as the median building in our dataset requires space heating at outdoor temperatures below 16°C and space cooling at outdoor temperatures above 8°C. B13, B17, and B10 had the lowest cooling change point temperature of 1°C; and B3, B10, and B23 had the highest heating change point temperature of 21°C. Underlying reasons behind these anomalous change point temperature values can be extremely low or high internal and solar heat gains, or inappropriate controls programming leading to simultaneous heating and cooling in some of the zones. The distribution of the differences between the cooling and heating change point temperatures in individual buildings is shown in Figure 7. The results indicate that the cooling change point temperatures were lower than the heating change point temperatures in 32 of the 35 buildings, meaning that most of the studied buildings tend to require both space heating and cooling between these temperatures. This problem was most evident with B3, B10, and B17. In these three buildings, the heating change point temperatures were at least 15°C higher than the cooling change point temperatures.

Figure 8 presents the Pearson's correlation coefficients among the six parameters of the heating and cooling change point models of the 35 buildings. The results indicate that the slope parameters for cooling and heating ($x_{2,\mathrm{clg}}$ and $x_{2,\mathrm{htg}}$) were significantly correlated with the intercept parameters for cooling and heating ($x_{1,\mathrm{clg}}$ and $x_{1,\mathrm{htg}}$). This can be interpreted that the heating and cooling energy performance of a building are closely related. However, the change point temperatures for heating and cooling did not exhibit a significant correlation with the slope or the intercept parameters.

Note that the analysis presented with these change point models was suitable to detect anomalous energy use patterns, and broadly interpret potential causes of these anomalies. However, a multivariate inverse model with variables such as the solar irradiance, wind speed, and binary workhours indicator could help in interpreting and isolating the causes of these anomalies. For example, the sensitivity of heating load intensities to the wind speed can be associated with a building's airtightness. Similarly, the sensitivity of the heating and cooling load intensities to the binary workhours indicator variable can reveal the

effectiveness of afterhours AHU on-off and temperature setback schedules to reduce the energy use. As an illustrative example, Figure 9 overlays separate change point models for B30 derived from hourly heating and cooling energy use data grouped into categories for three binary variables – i.e., afterhours / workhours, windy / still, and sunny / overcast. The threshold for windy class was defined as the instances at which the wind speed was greater than 1.2 m/s – recall that 1.2 m/s was the mean wind speed. The threshold for the sunny class was defined as the instances at which the horizontal solar irradiance was greater than 177 W/m^2 – recall that 177 W/m^2 was the mean horizontal solar irradiance. A visual inspection of Figure 9 underlines that the variations among the eight data categories are subtle, and that other model forms such as regression trees and ANN models may be more suitable than change point models to reveal the impact of these secondary predictors on the heating and cooling load intensities.

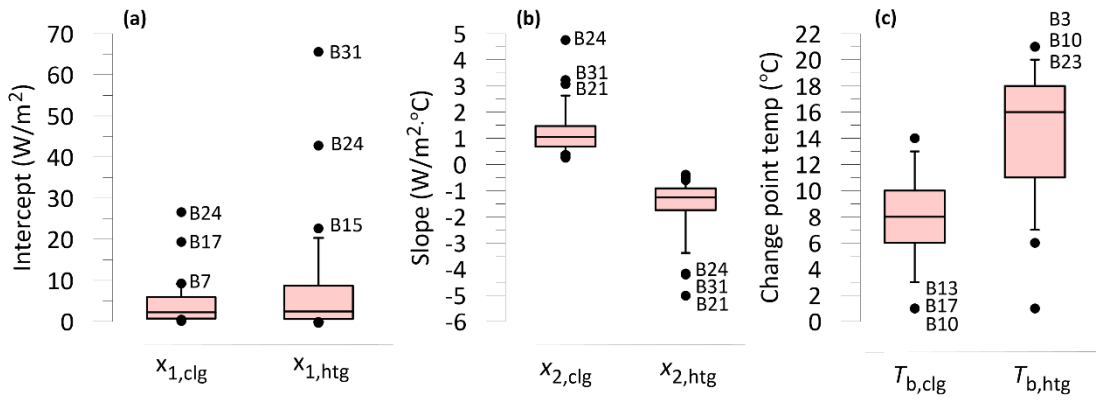

Figure 6: The distribution of the change point model (a) y-intercepts at the change point temperatures, (b) slopes, and (c) change point temperatures for individual buildings. The outlier buildings were annotated. The whiskers enclose 5th and 95th percentiles.

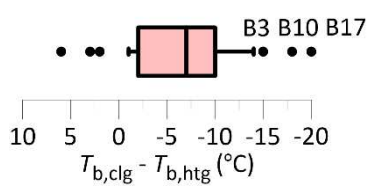

Figure 7: The difference between the cooling and heating change point temperatures in individual buildings. The whiskers enclose 5th and 95th percentiles.

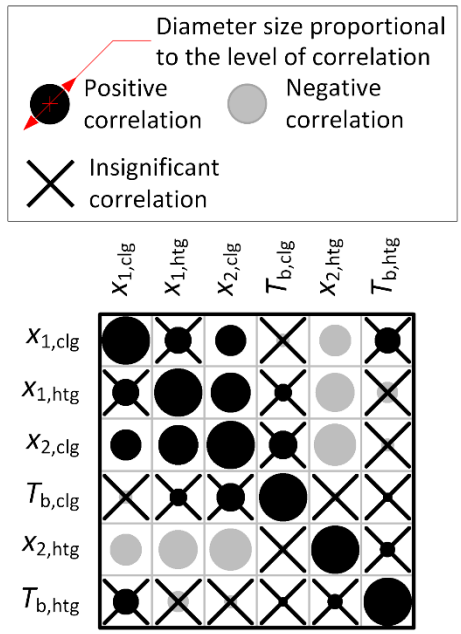

Figure 8: Pearson's correlation coefficients for the six heating and cooling change point model parameters from the 35 buildings. A correlation is deemed insignificant when the p-value is greater than 0.05.

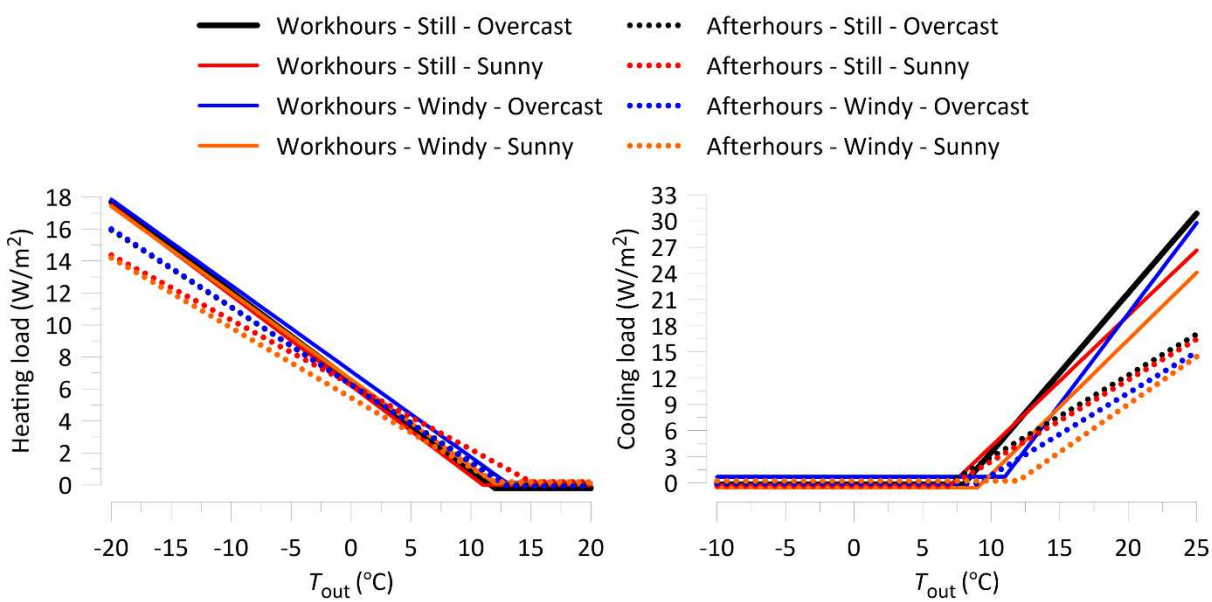

Figure 9: An illustrative example to demonstrate the difficulty in visually inspecting the influence of S_{wd}, Q_{sol}, and B_{wh} with change point models.

Figure 10 presents the complexity of the selected regression tree models that predict the heating and cooling load intensities in each of the 35 buildings. The median model for heating and cooling both had 13 branch nodes (i.e., 13 decision splits), and 14 leaf nodes (i.e., 14 possible outcomes). The number of branch and leaf nodes of selected regression tree models for individual buildings ranged from 6 to 35. Recall that the model selection was based on a repeated random subsampling cross-validation approach. Alternatively, the viability of fixing a regression tree form with the same number of decision splits and

predictors for all buildings could be studied by testing the universality of various regression tree forms with data from different buildings.

Figure 11 presents the distribution of the fraction of decision splits across each of the four predictors. Of the selected regression tree models for cooling, 30% to 55% of the decision splits were across different outdoor temperature levels, 0% to 34% of the decision splits were across different wind speed levels, 11% to 34% of the decision splits were across horizontal solar irradiance, and 3% to 25% of the splits were across the binary workhours indicator. Of the selected regression tree models for heating, 31% to 69% of the decision splits were at different outdoor temperature levels, 9% to 35% of the decision splits were across different wind speed levels, 0% to 30% of the decision splits were across different horizontal solar irradiance levels, and 0% to 23% of the decision splits were made across the binary workhours indicator variable. Based on these observations, for heating the hierarchical rank of the continuous predictors in descending order is outdoor temperature, wind speed, and solar irradiance, respectively. For cooling, the importance rank is outdoor temperature, solar irradiance, and wind speed, respectively. 6 out of the 35 regression tree models for the heating load intensities did not have any decision splits across the binary workhours indicator. This can be interpreted as the afterhours AHU on-off and temperature setback strategies did not seem to affect the heating energy use intensity in these six buildings – B25, B24, B19, B18, B14, B1. A potential reason behind this anomaly is the absence of a functional weekly operating schedule.

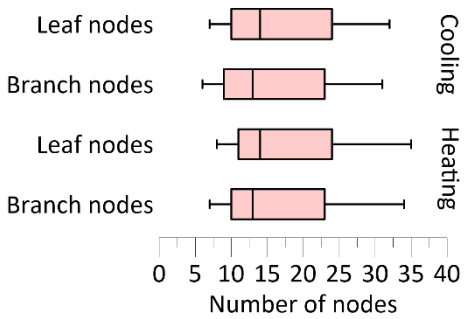

Figure 10: The number of leaf and branch nodes of the selected regression trees for the 35 buildings. The whiskers enclose all 35 values.

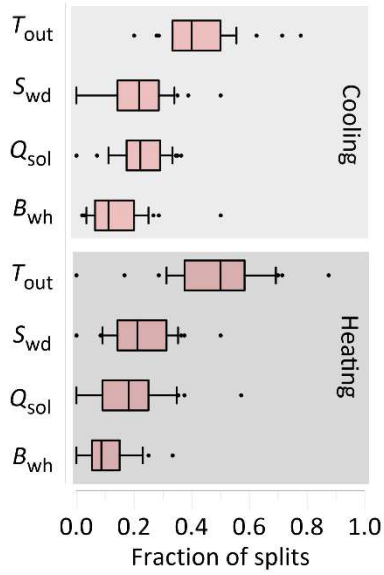

Figure 11: The distribution of the fraction of decision splits across each of the four predictors. The whiskers enclose 5th and 95th percentiles.

By using the regression tree models, we assessed the influence of an incremental change in each of the four regressors on the predicted heating and cooling load intensities. Figure 12 presents the influence of an incremental change in each of the four regressors in the predicted heating and cooling load intensities of individual buildings. Firstly, we assessed the average change in the predicted heating and cooling load intensities for each building, if the outdoor temperature was increased by 1°C for the whole data record. Similarly, the predicted influence of a 1 m/s increase in the wind speed and a 50 W/m^2 increase in the horizontal solar irradiance on the heating and cooling load intensities was estimated independently. Lastly, the influence of setting the binary workhours indicator value to "1" was estimated. Note that the purpose of this exercise was to isolate the outlier buildings based on their relationships with the four regressors. Therefore, we were merely interested in the significance of these regressors on a building's heating and cooling energy performance relative to other buildings. It is worth noting that the incremental change values were all in the expected direction – e.g., for all buildings, the cooling load increases and the heating load decreases with a one-degree Celsius increase in outdoor temperatures. However, despite the use of a random sampling-based cross-validation approach during model development, the readers should recognize the risk of exaggerating the influence of a regressor due to its collinearity with other regressors.

The results shown in Figure 12 indicate that a one-degree Celsius change in outdoor temperatures throughout the year tends to create the greatest changes with B21, B31, and B24. Note that this observation is in line with the change point model slope parameters reported in Figure 6.b. However, recall that the insights from the change point models were not adequate to isolate whether this anomaly was due to poor envelope performance, high air infiltration or over-ventilation. A 1 m/s change in the wind speed was estimated to cause the greatest changes in the heating and cooling load intensities of B24. This information can be used to prioritize an audit to inspect the AHU airflow and outdoor air fraction rates, and the condition of the weather-stripping in B24. In addition, the overall thermal performance of the envelope should be inspected in B21 and B31. The sensitivity analysis conducted upon the regression tree models also pointed out that the cooling load intensity in B17 appears to be highly dependent on the

horizontal solar irradiance. Based on this observation, a detailed energy model of B17 can be built, and retrofit options to reduce solar heat gains during the summer (e.g., automated solar shades) can be examined. It was also identified that the binary workhours indicator (B_{wh}) did not make any impact on the heating or the cooling load intensities of 14 of the 35 buildings. In other words, there were not any discernable differences between the workhours and the afterhours heating and cooling energy use patterns in two-fifths of the studied buildings. Based on this observation, the availability of weekly AHU on-off and temperature setback schedules can be inspected in these fourteen buildings.

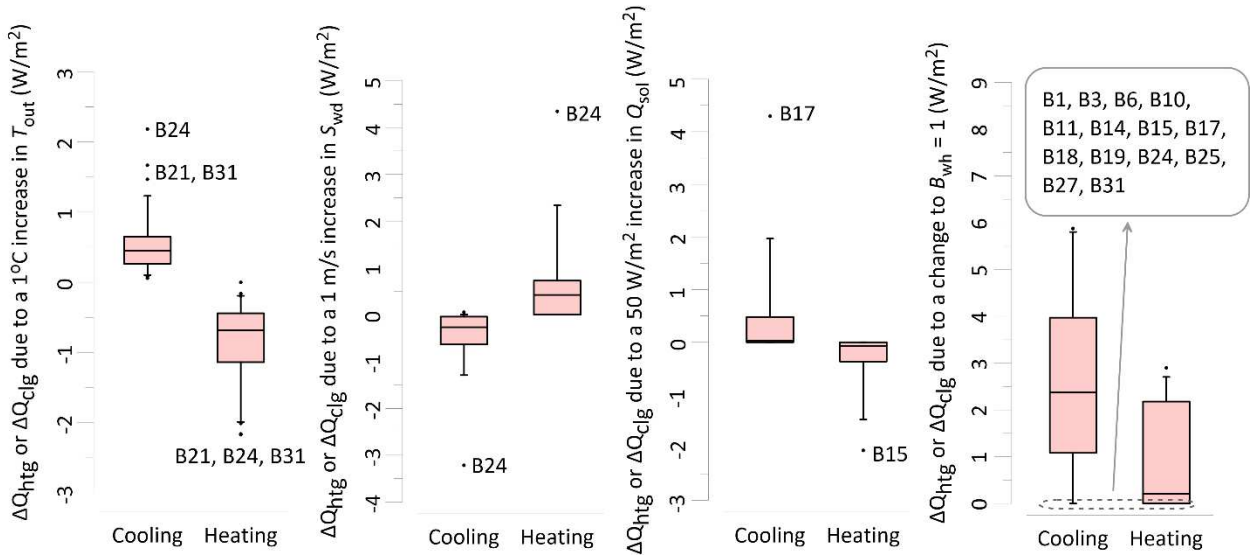

Figure 12: The influence of an incremental change in each of the four predictors in the predicted heating and cooling load intensities of individual buildings. These changes are predicted by the regression tree models. The whiskers enclose 5th and 95th percentiles.

After developing the change point and regression tree models, ANN models for the heating and cooling load intensities of individual buildings were developed as described in Section 3.3. The purpose of this exercise was to compare the insights gathered from the regression tree models to another data-mining algorithm. Figure 13 presents the number of hidden layer nodes of the selected ANN models. The median of these distributions for heating and cooling was three and five, respectively. For individual buildings, the selected ANN models had 1 to 11 hidden layer nodes for heating and 2 to 16 hidden layer nodes for cooling load intensities.

Similar to Figure 12, Figure 14 presents the influence of a change in the four regressors on the average predicted heating and cooling load intensities, albeit with the ANN models in lieu of the regression trees. The outliers identified based on the predicted sensitivity of the heating and cooling load intensities to outdoor temperature were identical to those detected with the regression trees (B21, B24, B31). The ANN and regression tree model formalisms were also in agreement when detecting the anomalous heating and cooling load intensity behaviour in response to a change in the wind speed. Although they were minor, there were also some discrepancies between outliers identified by regression trees and ANNs. For example, based on the regression tree models, B15 appears to be the most sensitive building to an increase in the horizontal solar irradiance. However, based on the ANN models, B9 was the most sensitive building to an increase in the horizontal solar irradiance. Based on the ANN model formalism, it was identified that 16 of the 35 buildings exhibit only minimal differences between their workhours and their

afterhours heating or cooling energy use patterns (less than 0.5 W/m^2). Recall that this number was 14 with the regression trees; 12 of these buildings were common in both ANN and regression tree model detections.

Table 2 presents a summary of the energy intensive anomalies, their potential underlying reasons, and the model formalisms that can be used in detecting them. An important research question which arises from these observations is: how can we formally blend the information extracted by individual inverse modelling approaches? Ensemble methods such as bagging, stacking, or boosting can act as meta-algorithms to combine the outcomes of several models developed for the same buildings. The research on the application of ensemble techniques in building energy inverse modelling remains anecdotal. For example, Touzani et al. (2018) employed the gradient boosting machine algorithm on a large building energy use dataset and achieved a substantial improvement in predictive accuracy in contrast to a time-of-week and outdoor temperature piecewise regression model. Ahmad et al. (2017) employed another ensemble modelling technique *random forest* on an HVAC energy use data record from a hotel in Madrid, Spain. Araya et al. (2017) put forward an ensemble learning framework for anomaly detection in building energy consumption and demonstrated that their ensemble learning technique can accurately detect anomalous energy consumption patterns. Future research is needed to further study ensemble modelling techniques in building energy use anomaly detection.

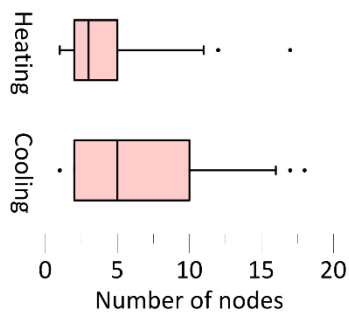

Figure 13: The number of hidden layer nodes of the selected ANN models predicting the heating and cooling load intensities of the 35 buildings. The whiskers enclose 5th and 95th percentiles.

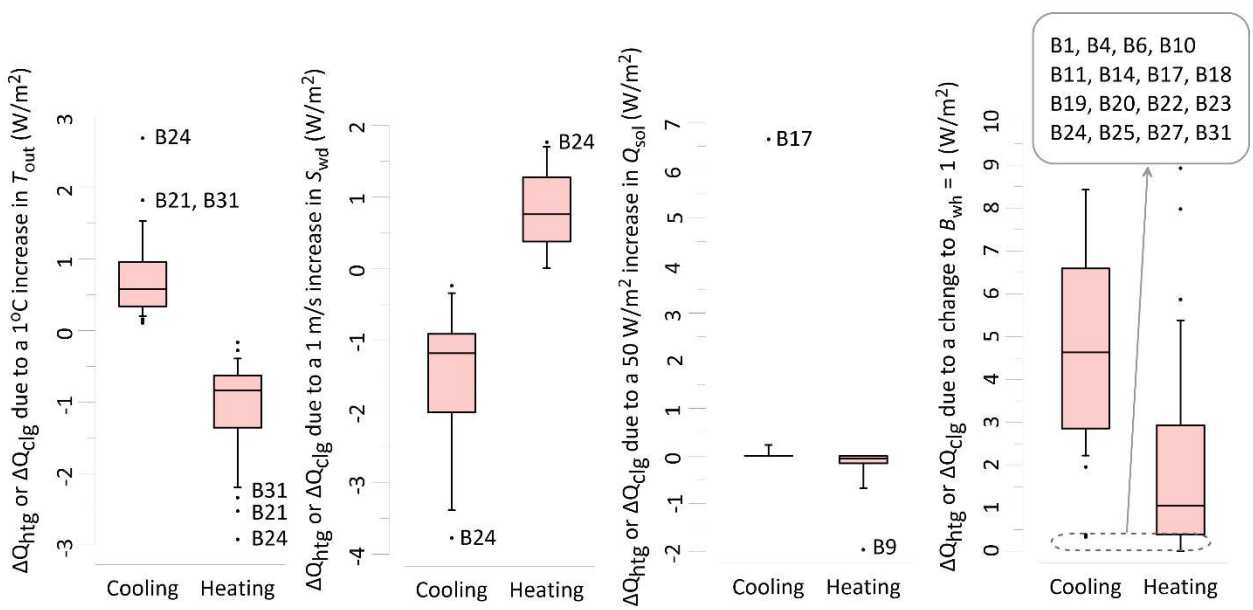

Figure 14: The influence of an incremental change in each of the four predictors in the predicted heating and cooling load intensities of individual buildings. These changes are predicted by the ANN models. The whiskers enclose 5th and 95th percentiles.

Table 2: A summary of the energy intensive anomalies, their potential underlying reasons, and the model formalisms that can be used in detecting them. CPM, RT, and ANN stand for the change point models, regression trees, and artificial neural networks, respectively.

Anomaly	**Potential reasons**	**CPM**	**RT**	**ANN**
High heating load intensity when the outdoor temperature was above the heating change point temperature	• Leaking or stuck open heating / reheat coil valves • High temperature setpoints for heating • Manual operator overrides to heating / reheat coil valves	✓		
High cooling load intensity when the outdoor temperature was below the cooling change point temperatures	• Absence of an economizer • Leaking or stuck open cooling coil valves • Low temperature setpoints for cooling • Manual operator overrides to cooling coil valves	✓		
Large rate of change in the heating and cooling load intensities in response to a change in the outdoor temperature	• High air infiltration • High envelope thermal conductance • Over-ventilation	✓	✓	✓
High heating change point temperature	• Low internal and solar heat gains • Inappropriate controls programming leading to simultaneous heating and cooling	✓		
Low cooling change point temperature	• High internal and solar heat gains • Inappropriate controls programming leading to simultaneous heating and cooling	✓		
Large rate of change in the heating and cooling load intensities in response to a change in the wind speed	• High air infiltration • Wind speeds exacerbated by building location and geometry		✓	✓
Large rate of change in the heating and cooling load intensities in response to a change in the horizontal solar irradiance	• High solar heat gains		✓	✓
Negligible difference between afterhours and workhours heating and cooling load intensities	• Ineffective afterhours AHU on-off and temperature setback schedules		✓	✓

Figure 15 presents a comparison of the predictive accuracy of the three-model formalisms. The results are presented as RMSE normalized with the mean heating / cooling load intensities – $CV(RMSE) = \left(\frac{RMSE}{mean(Q_{htg}\ or\ Q_{clg})}\right)$ – and mean of the prediction residuals normalized with the mean heating / cooling load intensities – $NMBE = \left(\frac{mean(residuals)}{mean(Q_{htg}\ or\ Q_{clg})}\right)$. The results indicate that CV(RMSE) metric for heating load intensity models tends to decrease slightly from a range of 4-20% with change point models to a range of 4-18% with the use of regression trees, and to a range of 3-16% with the use of ANNs. Similarly, the CV(RMSE) of the cooling load intensity models tends to decrease slightly with the use of regression trees and ANNs in lieu of change points. In most buildings, the normalized mean bias errors (NMBE) were less than 5%; meaning that studied model forms were able to generate unbiased heating and cooling load

predictions. Among the studied model forms, the lowest NMBE was achieved with regression trees. On average, the three inverse model formalisms were able to achieve higher predictive accuracies for the heating load intensities than for the cooling load intensities. This situation can be interpreted as the heating loads in these cold climate buildings (ASHRAE Climate Zone 6) being driven primarily by environment-driven loads, while steady-periodic thermal disturbances such as plug-in equipment or lighting use play a greater role on the cooling loads in these buildings.

However, it is important to note that different inverse model formalisms, despite not offering significant improvements in the predictive accuracy of heating and cooling load intensities, provide opportunities to isolate different energy use anomalies at varying resolutions. Therefore, advanced inverse modelling methods such as ANNs and regression trees can be beneficial in detecting and isolating energy use anomalies but added complexity may not be necessary for establishing an energy use baseline. However, note that univariate change point models may not be suitable to establish an energy use baseline for buildings in which the heating and cooling load intensities are significantly affected by casual and solar heat gains, as well as the daily and seasonal variations in the AHU and temperature setpoint scheduling. In these cases, univariate change point models were observed to be sufficient in characterizing load profiles at temporal resolutions coarser than weekly only. To illustrate this, Figure 16 presents examples comparing the predictions of the three model forms for heating and cooling load intensities of B30 over a one-month period. The use of regression trees and ANN models in lieu of univariate change point models appears to improve the ability of inverse models to characterize the hourly heating and cooling load intensity patterns. As an alternative to the multivariate regression trees and ANNs demonstrated in this paper, the hybrid multivariate change point model method which combines short term energy use data, weather variables, and internal loads with utility bill data could be employed (Abushakra 2016a, 2016b, 2016c).

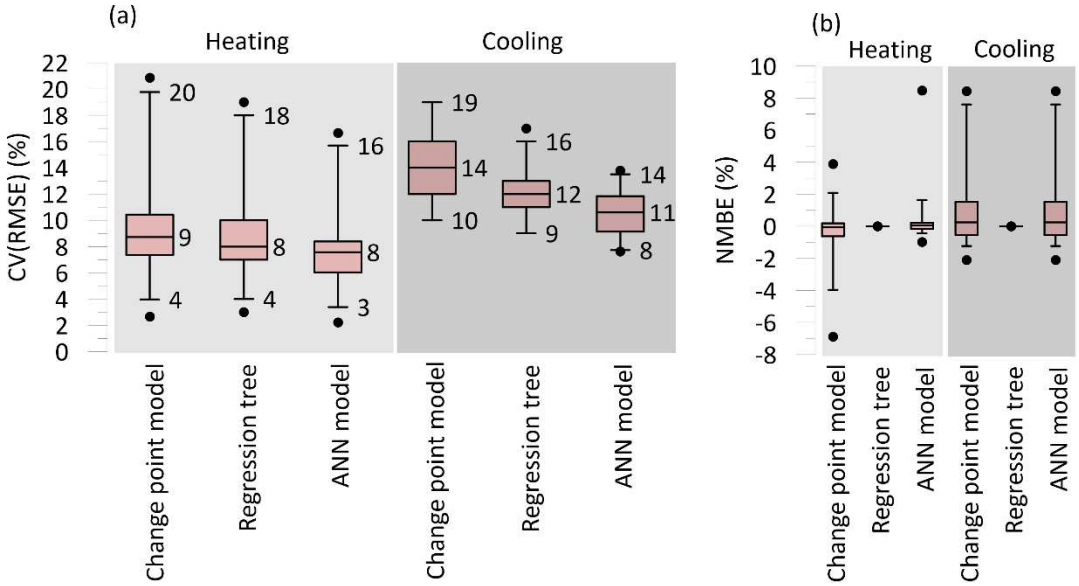

Figure 15: The cross-validated predictive accuracy of the three modelling formalisms for individual buildings in terms of (a) CV(RMSE) and (b) NMBE.

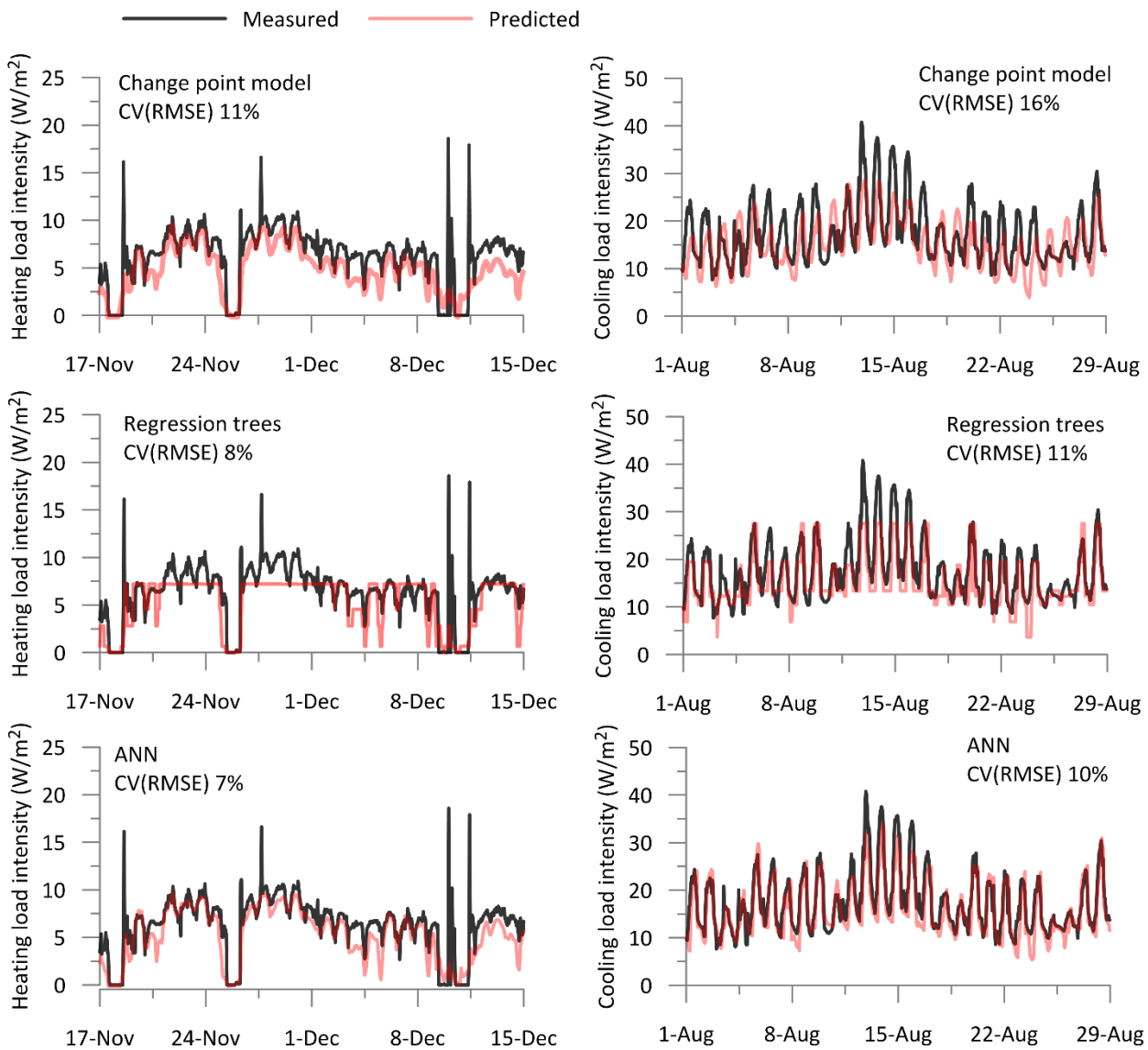

Figure 16: Illustrative examples comparing the predictions of the three models for heating and cooling load intensities of B30.

4.1. Unresolved issues

In this study, three inverse model formalisms were applied to characterize the heating and cooling load patterns of 35 office buildings. We demonstrated that the three model forms can be used to detect energy use anomalies consistently. It is also argued that they can be used to interpret potential reasons behind these anomalies. Although the results demonstrated in this study are promising, there are several methodological limitations.

In the absence of electricity submeters for lighting and plug-in equipment loads and sensors for occupancy count estimation, we treated occupant-driven thermal loads in a simplistic way. The regression trees and ANN models were designed to input a binary occupancy indicator as a regressor during nominal workhours to represent the impact of occupant-driven thermal loads on the buildings' energy use. As the

sensing and metering infrastructure in commercial and institutional buildings improve, it is likely that inverse modelling will lend itself to the detection of many other energy use anomalies.

Using inverse models in anomaly (i.e., outlier) detection is an unsupervised learning technique. Thus, the physical significance of the anomalies needs to be assessed by an expert. However, the anomalies detected, and their potential explanations, have not been verified by on-site inspections in this paper – which is an important next research step.

To illustrate the process of inverse model-based anomaly detection, the fifth percentile and the ninety-fifth percentile values were selected as anomaly thresholds in Figure 6, 7, 12, and 14. As needed, these threshold values can be modified. However, because this analysis is intended to be applied continuously over time, as the issues causing anomalies are addressed, new outliers are anticipated to emerge and to be prioritized.

Note that the methodology presented in this paper cannot be used to detect non-energy intensive anomalies – e.g., issues specific to some thermal zones. Despite being capable of helping an expert interpret potential reasons behind anomalies (e.g., high air-infiltration, high thermal conductance, absence of AHU on-off schedule), the method cannot diagnose the root-cause of a problem. Hence, it is not a replacement for conventional energy auditing and fault detection and diagnostics (FDD) methods – which build upon distributed indoor sensing inherent in modern building automation systems. Complementary to energy audits and FDD, inverse modelling can be used to establish an energy use benchmark in a building cluster, to monitor its energy performance continuously, and to detect energy intensive anomalies and interpret potential reasons behind them. This may prompt a physical inspection by an expert, and, in helping to target the resources for such inspections, make the process of successful building maintenance more efficient.

In this study, the inverse modelling methods were applied to a dataset gathered from 35 office buildings of varying age, size, and characteristics. However, these buildings were in the same city; thus, they were subject to the same climatic conditions (Climate Zone 6). The predictive performance of inverse models is anticipated to be different in other climates. Although the methodology is transferable, the readers should be cautious in extending the results of this paper to other buildings.

Although the insights gathered from different inverse modelling techniques (e.g., regression trees and ANNs) were generally in agreement, there were some minor discrepancies in their anomaly detections. This situation highlights the need for ensemble learning techniques, which can systematically blend the predictions from different model formalisms. Future work is planned to examine the use of ensemble methods in inverse modelling of the heating and cooling load intensities in buildings.

5. Conclusions

Three different inverse modelling methods were applied to hourly heating and cooling data extracted from 35 office buildings in Ottawa, Canada. In this modelling exercise, concurrent weather data from a local weather station were used as regressors. The purpose of this inverse modelling exercise was to establish an energy use baseline and continuously monitor energy performance; and detect energy intensive anomalies and interpret the potential reasons behind them.

Firstly, for each building, two univariate three-parameter change point models were developed to characterize the heating and cooling load intensities. The relationship among these six parameters from each building was explored. Using these simple change point models, five types of anomalies were identified, and the potential reasons behind them were discussed. These five anomaly types were: (1) a high heating load intensity when the outdoor temperature was above the heating change point temperature, (2) a high cooling load intensity when the outdoor temperature was below the cooling change point temperatures, (3) a large rate of change in the heating and cooling load intensities in response to a change in the outdoor temperature, (4) a high heating change point temperature, and (5) a low cooling change point temperature.

The other two inverse modelling methods studied in this paper were regression trees and the artificial neural networks. These models were structured with four regressors: the outdoor temperature, the wind speed, the horizontal solar irradiance, and a binary workhours indicator. The sensitivity of heating and cooling load intensities to the studied regressors was examined by using the regression tree models developed for each building. The outlier buildings in terms of their sensitivity to different regressors were identified, and the physical meaning of these sensitivity results was discussed. Aside from the previous five, three additional anomaly types were identified: (1) a large rate of change in the heating and cooling load intensities in response to a change in the wind speed, (2) a large rate of change in the heating and cooling load intensities in response to a change in the horizontal solar irradiance, (3) a negligible difference between afterhours and workhours heating and cooling load intensities. The results from this analysis indicate that 14 to 16 of the 35 buildings did not appear to have an effective afterhours AHU on-off and/or temperature setback strategy. All except three had cooling change point temperatures lower than their heating change point temperatures – a potential indication of a simultaneous heating and cooling problem. The analysis also identified a few buildings with anomalies potentially related to high air-infiltration or over-ventilation, high thermal conductance, and high solar heat gains during the summer. However, it is worth noting that these three additional anomalies can also be identified through a multivariate change point modelling approach – which was not studied in this paper.

Moreover, the insights gathered by the studied inverse modelling methods were compared. It was found that although the anomalies detected by different inverse modelling techniques were generally in agreement, there were some minor discrepancies. This situation highlighted the need for future research to develop meta-modelling methods that can blend the insights gathered from different inverse modelling techniques.

References

Abdelalim, A., W. O'Brien and Z. Shi (2017). "Development of Sankey diagrams to visualize real HVAC performance." *Energy and Buildings* 149: 282-297.

Abushakra, B. (1997). *An inverse model to predict and evaluate the energy performance of large commercial and institutional buildings*. Building Simulation, Prague, Czech Republic, http://www.ibpsa.org/proceedings/BS1997/BS97_P183.pdf IBPSA.

Abushakra, B. and D. Claridge (2001). "Accounting for the occupancy variable in inverse building energy baselining models." International Conference for Enhanced Building Operations. https://oaktrust.library.tamu.edu/handle/1969.1/5161, Texas A&M University.

Abushakra, B. (2001). Accounting for the occupancy variable in inverse building energy baselining models, https://oaktrust.library.tamu.edu/handle/1969.1/154781, Texas A&M University.

Abushakra, B. and M. T. Paulus (2016). "An hourly hybrid multi-variate change-point inverse model using short-term monitored data for annual prediction of building energy performance, part I: Background (1404-RP)." *Science and Technology for the Built* Environment 22(7): 976-983.

Abushakra, B. and M. T. Paulus (2016). "An hourly hybrid multi-variate change-point inverse model using short-term monitored data for annual prediction of building energy performance, part II: Methodology (1404-RP)." *Science and Technology for the Built Environment* 22(7): 984-995.

Abushakra, B. and M. T. Paulus (2016). "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)." *Science and Technology for the Built Environment* 22(7): 996-1009.

Ahmad, M. W., M. Mourshed and Y. Rezgui (2017). "Trees vs Neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption." *Energy and Buildings* 147: 77-89.

Araya, D. B., K. Grolinger, H. F. ElYamany, M. A. Capretz and G. Bitsuamlak (2017). "An ensemble learning framework for anomaly detection in building energy consumption." *Energy and Buildings* 144: 191-206.

ASHRAE (2011). Procedures for commercial building energy audits second edition. Atlanta, American Society of Heating, Refrigerating, and Air-Conditioning Engineers.

ASHRAE (2014). *Guideline 14: Measurement of energy, demand, and water savings*. Atlanta, American Society of Heating, Refrigerating and Airconditioning Engineers.

Bourassa, N. J., M. A. Piette and N. Motegi (2004). *Evaluation of retrocommissioning persistence in large commercial buildings*. National Conference on Building Commissioning, https://cloudfront.escholarship.org/dist/prd/content/qt9tj6h68r/qt9tj6h68r.pdf.

Braun, J. E. and N. Chaturvedi (2002). "An inverse gray-box model for transient building load prediction." *HVAC&R Research* 8(1): 73-99.

Breiman, L., J. Friedman, C. J. Stone and R. A. Olshen (1984). *Classification and regression trees*, CRC press.

Carpenter, J., K. A. Woodbury and Z. O'Neill (2018). "Using change-point and Gaussian process models to create baseline energy models in industrial facilities: A comparison." *Applied Energy* 213: 415-425.

Chae, Y. T., R. Horesh, Y. Hwang and Y. M. Lee (2016). "Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings." *Energy and Buildings* 111: 184-194.

Cho, S. Y. (2008). *The Persistence of Savings Obtained from Commissioning of Existing Buildings*. Master's Thesis, http://oaktrust.library.tamu.edu/handle/1969.1/55022, Texas A&M University.

Dong, B., C. Cao and S. E. Lee (2005). "Applying support vector machines to predict building energy consumption in tropical region." *Energy and Buildings* 37(5): 545-553.

Ferreira, M., M. Almeida, A. Rodrigues and S. M. Silva (2016). "Comparing cost-optimal and net-zero energy targets in building retrofit." *Building Research & Information* 44(2): 188-201.

Friedman, H., A. Potter, T. Haasl, D. Claridge and S. Cho (2003). *Persistence of benefits from new building commissioning*. National Conference on Building Commissioning, https://aceee.org/files/proceedings/2002/data/papers/SS02_Panel3_Paper11.pdf.

Fux, S. F., A. Ashouri, M. J. Benz and L. Guzzella (2014). "EKF based self-adaptive thermal model for a passive house." *Energy and Buildings* 68: 811-817.

Golden, A., K. Woodbury, J. Carpenter and Z. O'Neill (2017). "Change point and degree day baseline regression models in industrial facilities." *Energy and Buildings* 144: 30-41.

Gunay, B. and W. Shen (2017). "Connected and Distributed Sensing in Buildings: Improving Operation and Maintenance." *IEEE Systems, Man, and Cybernetics Magazine* 3(4): 27-34.

Gunay, B., W. Shen and G. Newsham (2017a). "Inverse blackbox modeling of the heating and cooling load in office buildings." *Energy and Buildings* 142: 200-210.

Gunay, B., W. Shen and C. Yang (2017b). "Characterization of a building's operation using automation data: A review and case study." *Building and Environment* 118: 196-210.

Haberl, J. S. and S. Thamilseran (1996). Great energy predictor shootout II: Measuring retrofit savings--overview and discussion of results, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, GA (United States).

Jacob, D., S. Dietz, S. Komhard, C. Neumann and S. Herkel (2010). "Black-box models for fault detection and performance monitoring of buildings." *Journal of Building Performance Simulation* 3(1): 53-62.

Jalori, S. and T. A. Reddy (2015a). "A new clustering method to identify outliers and diurnal schedules from building energy interval data." *ASHRAE Transactions* 121: 33.

Jalori, S. and T. A. Reddy (2015b). "A unified inverse modeling framework for whole-building energy interval data: Daily and hourly baseline modeling and short-term load forecasting." *ASHRAE Transactions* 121: 156.

Karatasou, S., M. Santamouris and V. Geros (2006). "Modeling and predicting building's energy use with artificial neural networks: Methods and results." *Energy and Buildings* 38(8): 949-958.

Katipamula, S. and M. R. Brambley (2005a). "Review article: methods for fault detection, diagnostics, and prognostics for building systems—a review, Part I." *HVAC&R Research* 11(1): 3-25.

Katipamula, S. and M. R. Brambley (2005b). "Review article: methods for fault detection, diagnostics, and prognostics for building systems—a review, Part II." *HVAC&R Research* 11(2): 169-187.

Kelsey, J. and D. Pearson (2011). "Updated Procedures for Commercial Building Energy Audits." *ASHRAE Transactions* 117(2).

Kim, W. and S. Katipamula (2017). "A Review of Fault Detection and Diagnostics Methods for Building Systems." *Science and Technology for the Built Environment* 24: 3-21.

Krarti, M., J. F. Kreider, D. Cohen and P. Curtiss (1998). "Estimation of Energy Savings for Building Retrofits Using Neural Networks." *Journal of Solar Energy Engineering* 120(3): 211-216.

Li, Q., Q. Meng, J. Cai, H. Yoshino and A. Mochida (2009). "Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks." *Energy Conversion and Management* 50(1): 90-96.

Liu, F., H. Jiang, Y. M. Lee, J. Snowdon and M. Bobker (2011a). *Statistical modeling for anomaly detection, forecasting and root cause analysis of energy consumption for a portfolio of buildings*. Building Simulation, Sydney, Australia, http://www.ibpsa.org/proceedings/BS2011/P_1789.pdf, IBPSA.

Liu, J., J.-C. Baltazar and D. E. Claridge (2011b). "Improvements to a Methodology for Estimating Potential Energy Savings from Existing Building-Commissioning/Retrofit Measures." *ASHRAE Transactions* 117(1).

Mills, E. (2011). "Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the United States." *Energy Efficiency* 4(2): 145-173.

NRCan (2011). Energy savings toolbox: an energy audit manual and tool. Ottawa, Natural Resources Canada.

Paulus, M. T., D. E. Claridge and C. Culp (2015). "Algorithm for automating the selection of a temperature dependent change point model." *Energy and Buildings* 87: 95-104.

Pérez-Lombard, L., J. Ortiz, R. González and I. R. Maestre (2009). "A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes." *Energy and Buildings* 41(3): 272-278.

Reddy, T. A. (2011). *Applied data analysis and modeling for energy engineers and scientists*, Springer Science & Business Media.

Singh, V., T. A. Reddy and B. Abushakra (2013). "Predicting Annual Energy Use in Buildings Using Short-Term Monitoring and Utility Bills: The Hybrid Inverse Model Using Daily Data (HIM-D)." *ASHRAE transactions* 119(2).

Singh, V., T. A. Reddy and B. Abushakra (2014). "Predicting annual energy use in buildings using short-term monitoring: The dry-bulb temperature analysis (DBTA) method." *ASHRAE Transactions* 120: 397-405.

Srivastav, A., A. Tewari and B. Dong (2013). "Baseline building energy modeling and localized uncertainty quantification using Gaussian mixture models." *Energy and Buildings* 65: 438-447.

Thamilseran, S. and J. Haberl (1995). *A bin method for calculating energy conservation retrofit savings in commercial buildings*. Solar Energy Conference, Maui, Hawaii, https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/6640/ESL-HH-94-05-21.pdf, ASME-JSES-JSME.

Thomas Ng, S., M. Skitmore and K. F. Wong (2008). "Using genetic algorithms and linear regression analysis for private housing demand forecast." *Building and Environment* 43(6): 1171-1184.

Touzani, S., J. Granderson and S. Fernandes (2018). "Gradient boosting machine for modeling the energy consumption of commercial buildings." *Energy and Buildings* 158: 1533-1543.

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

Witten, I. H., E. Frank, M. Hall and C. Pal (2016). *Data Mining: Practical machine learning tools and techniques*, Morgan Kaufmann.

Yu, Z., F. Haghighat, B. C. M. Fung and H. Yoshino (2010). "A decision tree method for building energy demand modeling." *Energy and Buildings* 42(10): 1637-1646.

Zhang, Y., Z. O'Neill, B. Dong and G. Augenbroe (2015). "Comparisons of inverse modeling approaches for predicting building energy performance." *Building and Environment* 86: 177-190.

Zhou, Q., S. Wang, X. Xu and F. Xiao (2008). "A grey-box model of next-day building thermal load prediction for energy-efficient control." *International Journal of Energy Research* 32(15): 1418-1431.