Forecasting the electric demand of an HVAC system with deep learning-based techniques

Jason Runge and Radu Zmeureanu
Centre for Net-Zero Energy Buildings Studies, Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Quebec, Canada.

Abstract

This paper presents the development of deep learning-based (DL) models for the forecasting of electric demand of a heating, ventilation and air conditioning (HVAC) system during the summer operation, which includes the fans, pumps, chillers, and cooling towers. The paper reviews the recent applications of DL models for the performance forecasting of HVAC systems. Models are built and applied to a case study of an institutional building with synthetic data obtained from a calibrated eQuest simulation. The models use hourly time step data and provide a forecast horizon of six-hours ahead. Multiple long short-term memory (LSTM) models are coupled together in parallel into a homogenous ensemble to provide a combined output forecast. The proposed DL models, with several hidden layers, are compared with a simple forecasting (SFA) approach, and with a feed-forward neural network (FFNN) with a single hidden layer.

Introduction

With the growth of our cities and populations, our energy consumption usages have been increasing both domestically and worldwide. The ramifications of the increasing trends have raised many concerns related to environmental costs and degradations, geopolitics, and energy security. Within Canada, the residential and commercial sectors consumes approximately 19% of our total primary usage (Natural Resources Canada, 2018). Furthermore, buildings are responsible for approximately 40% of the total direct and indirect CO_2 emissions in most countries (International Energy Agency , 2020). Therefore, increasing the energy efficiency for building and building operations is critical for achieving global sustainability.

Forecasting energy consumption in buildings has received a lot of attention as it underpins many strategies for improving and/or managing energy use. Such applications for forecasting include fault detection and diagnosis, energy optimization, demand response, model predictive control, and demand side management. It also has further applications for energy security and smart grid integration. So far, the majority of the applications for forecasting has been applied to short-term horizons, as this has strong applications for the daily operations. Review papers (Runge & Zmeureanu, 2019; Amasyali & El-Gohary, 2018; Wang & Srinivasan, 2017) which catalogued a significant portion of the published research work, have found that the majority of such models focused on forecasting overall energy, heating, cooling, and total electricity loads for buildings, using hourly data with a horizon up to 24 hours ahead. Furthermore, such reviews found that artificial neural network (ANN) models are among the highest performing machine learning algorithms, and recommended that future work should focus on the application of deep learning models.

Today's buildings stock can be quite energy-intensive, while simultaneously being monitored and recorded in increasing amounts. Thus, big data problems have become a growing issue over the past couple of decades. Recent advances in the field of artificial intelligence (AI), deep learning (DL), have begun addressing such big data issues with a variety of new approaches and techniques. As a result, DL techniques have begun to be applied over a variety of fields including forecasting energy in buildings.

Amarasignhe et al. (2017) compared deep learning and conventional neural network techniques (i.e., ANN, SVM) to forecast the electric demand of a building using measured data as inputs. The forecast horizon was 60 hours ahead and used 4 years of training data (Amarasinghe, Marino, & Manic, 2017). The results demonstrated that the deep learning based forecasting models obtained less forecasting error when compared with standard machine learning based techniques (e.g., reduction of RMSE of 0.028 to 0.189 kW).

Cai et al. (2019) compared deep learning with time series techniques in order to forecast the overall electric demand of an institutional building 24 hours ahead. The results from this work showed that the deep learning based techniques had a reduced error when compared with the time series techniques (e.g., reduction of CV(RMSE) of 0.4% to 11.15%) (Cai, Pipattanasompom, & Rahman, 2019).

Fan et al. (2019) applied deep learning techniques for feature extraction which were then coupled to forecasting models such as multiple linear regression (MLR), a neural network (ANN), support vector regression (SVR), and extreme gradient boosting algorithm, in order to forecast future cooling loads for an institutional building. The forecast horizon was 24 hours ahead using data at 30 minute intervals. The results of this work demonstrated that the neural network was among the best-performing algorithms (i.e., CV(RMSE) of 20.1% to 25.1%). However, it was surpassed by extreme gradient boosting algorithm (i.e., CV(RMSE) of 17.7% to 19.3%) (Fan, Sun, Zhao, Song, &

Wang, 2019). Furthermore, it was observed that deep learning based techniques may provide an improved performance for feature extraction (i.e., reduction of CV(RMSE) of up to 4.6%).

Shan et al. (2019) compared deep learning techniques (LSTM, GRU, ensemble GRU) with other forecasting algorithms including: a variety of regression-based models, support vector machine, decision trees, ARIMA, and standard ANNs. The forecasting models were applied in order to forecast the electricity demand of a hotel and offices using hourly data. The results of this work demonstrated that the deep learning based forecasting models were among the top performing algorithms. For example, the forecasting error of LSTM model was 3.72%, while the ordinary least square regression had the mean absolute percentage error (MAPE) of 13.55%, for the office case study (Shan, Cao, & Wu, 2019).

To summarize the results found, the application of deep learning models for forecasting whole energy loads within buildings may provide a lower forecasting error when compared to other forecasting approaches. While the results look promising for the applications of deep learning techniques to forecast the overall energy loads in buildings, the application of such models to forecast some of the target variables for the heating, ventilation, and air conditioning systems (HVAC) are still in their infancy.

Lui et al. (2019) compared conventional and deep learning based techniques in order to forecast the energy demand of a ground source heat pump. Data was collected at five-minute time steps and the forecast horizon was one step (i.e., 5 min) ahead. The results demonstrated improved performance for the application of deep learning techniques. The lowest RMSE of 15.32 kW was found with a coupled auto encoder and a deep determine policy gradient approach compared to the largest RMSE of 20.81 kW obtained with a support vector machine approach (Liu, Xu, Guo, & Chen, 2019).

In spite of the papers which have applied DL techniques to date, the forecasting of the HVAC energy demand remains an area which has not received as much attention. Yet, it is of significant importance as the HVAC can contribute a large portion to a building's overall electric demand e.g., 19-67% for a commercial building (U.S. Energy Information Agency, 2012). To date, the overwhelming majority of HVAC forecasting publications have applied "shallow" machine learning based techniques. Such "shallow" models typically consists of one to three non-linear transformations. With regards to the ANN, the "conventional" ANN which has been applied to date, has been the feed forward neural network (FFNN); typically, this model has been applied with one or two hidden layers and without any recurrent connections.

In spite of the good work by previous researchers, there remain numerous gaps related to the applications of deep learning based techniques for the forecasting of HVAC systems. Therefore, this paper proposes to address the gaps with the following objectives:

1. Design and testing of a univariate multistep ahead DL model to forecast the HVAC's overall electric demand
2. Development of a fast to deploy deep learning model
3. Comparison of different forecasting techniques in terms of performance. Such techniques include: (i) a homogeneous ensemble of LSTM networks, (ii) a single LSTM model, (iii) a feed forward neural network, and (iv) a simple forecasting approach

Synthetic data is used for this work as a proxy to real measurement data from a building. This approach was selected as the synthetic data does not contain noise, missing data and faults commonly recorded in measurement data. Thus, the models built on this data set and their performances can be used as benchmarks to compare with future models developed on other case studies with measurement data.

The approaches presented in this paper may be beneficial for improved decision making strategies pertaining to a building HVAC system for building owners and building energy managers. Furthermore, the approaches and techniques applied within this work may help demand side management, control, and fault detection and diagnosis as forecasting is a fundamental aspect within these applications.

Methods

Forecasting methods

The forecasting methodology followed within this work is based on the methodology proposed by Makridakis et al (2003) and is presented in Table 1.

Table 1: Forecasting methodology

Definition of Tasks	
Step 1: Problem definition	- Problem definition - Identification of objectives
Step 2: Data gathering	- Data extraction from eQuest simulation
Step 3: Preliminary analysis	- Visualization and pre-processing of data
Step 4: Model Construction	- Data partitioning - Hyperparameter selection
Step 5: Application	- Application of forecasting models to testing data
Step 6: Comparison	- Comparison of forecasting methods

Long short term memory networks

Time series data is often characterized with high variations with rapid transients. Modelling such systems with linear models can therefore, be quite challenging. In such a context, non-linear techniques are therefore desired. Artificial neural networks (ANN) are a sub-category of machine learning (ML) inspired by biological systems. A neural network consist of interconnected neurons, joined together with the connections typically being weighted.

Many different types of ANN models exist; however, for this work, a long short term memory neural network (LSTM) is applied along with the standard feed-forward neural network.

Similarly, DL are a sub-category of ML; DL has been defined as "representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level" (LeCun, Bengio, & Hinton, 2015). Deep learning techniques have become popular in recent years due to their ability to handle large amounts of data, improved feature extraction abilities, and improved model performances.

Deep neural networks are a sub category of DL, such models incorporate deep learning neural networks. DNN's have typically been applied in recent years as the main approach with DL based techniques applied for forecasting energy in buildings. To the best of the author's knowledge, there have been three main ways to date in which DNN have been applied as DL based techniques:

1. Increasing the number of feed forward neural networks with multiple hidden layers
2. The application of recurrent neural networks. Such models may have a single or multiple hidden layers, however, even single layer recurrent neural networks may be considered deep learning due to training approaches. Unfolded, which occurs in training of the network, such models are consider networks with very deep structures as information from previous states is passed to current states. (Goodfellow, Bengio, & Courville, 2016).
3. Through the sequential coupling of different types of neural networks into overall structures (example, an autoencoder coupled with a ANN forecasting model)

The benefits and drawbacks of each approach are beyond the scope of this work; the primary focus of this work is on the implementation of recurrent neural networks.

Recurrent neural networks (RNN) are a family of AI and deep learning models. Such networks are specialized for handling sequences of values y(t), y(t-1),…, y(t-n) (Goodfellow, Bengio, & Courville, 2016). Such models recursively pass information through time from previous states; thus, current outputs are based on previous outputs.

The LSTM is a type of recurrent neural network first introduced in 1997 by Hochreiter and Schmidhuber (Hochreiter & Schmidhuber, 1997). Unlike "normal" FFNNs, RNN contain recurrent connections between neurons, even in training. The LSTM unit, is a unique RNN which contains a memory cell. The purpose of this is to store information over time. Within the LSTM unit there are three main gates which control the flow of information: the forget gate, the input gate, and the output gate. All three gates are provided the same input vector of information and then passed through weights and activation functions. The forget gate decides whether the cell state forgets the previous state, the input gate decides whether the input state enters the cell state, and the output gate decides whether the cell state passes values to output and hidden state of the next time step.

The LSTM neural network has the advantage of constant error backpropagation, thus allowing the neural network to learn long term dependencies compared with previous types of recurrent neural networks. Furthermore, such models may have reduced hyperparameter tuning and are able to generalize well. Equations 1-6 provides the governing equations for the LSTM models applied within this work.

$$f_t = \sigma\,(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$
$$i_t = \sigma\,(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o\,) \tag{3}$$
$$\tilde{C}_t = \mathrm{ReLu}(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{5}$$
$$h_t = o_t * \mathrm{ReLu}(C_t) \tag{6}$$

Where f_t, i_t, and o_t are the forget gate, input gate, and output gate. Furthermore, x_t is the input and h_t is the cell output. The W's and b's refer to the weights and biases of the LSTM unit. C_t is the current cell state and $\tilde{C}_t$ is the new cell state. The activation function for the models applied in this work are the rectified linear unit, ReLu, while the recurrent activations are the sigmoid function (σ). Figure 1 presents the architecture of an LSTM unit. Such units are applied in parallel as a hidden layer within the overall forecasting model. In addition to the layer(s) of LSTM units, a fully connected neuron layer is attached at the outer most layer of the network; this is shown in Figure 2.

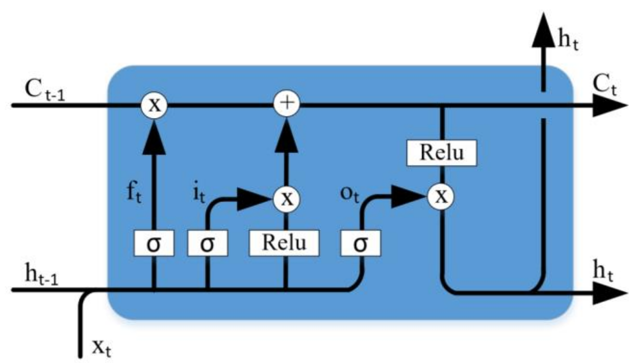

Figure 1: Schematic for a LSTM unit

Ensemble forecasting

Within the field of machine learning, ensemble models consist of multiple single-point forecasting models (e.g. single ANN, or SVR, etc.) trained individually, whose output forecasts are combined into a single overall output (Opitz & Maclin, 1999). The resulting output forecasts are generally more accurate and consistent than either of the individual single-point forecasting models within the ensemble. Such models can be broken into two main classifications (i) heterogeneous models consisting of different single-point forecasting models (e.g., LSTM and SVR), and (ii) homogenous models consisting of similar single-point forecasting models with different hyperparameters (e.g., LSTM and LSTM). A key feature of the ensemble forecasting model is the combination of the single-point forecasts into an overall output forecast. While, different approaches exist, for this work the equal weights approach was applied as this is shown to have robust results as demonstrated in references (Jovanović, Sretenović, & Živković, 2015) and (Runge, Zmeureanu, & Le Cam, 2020).

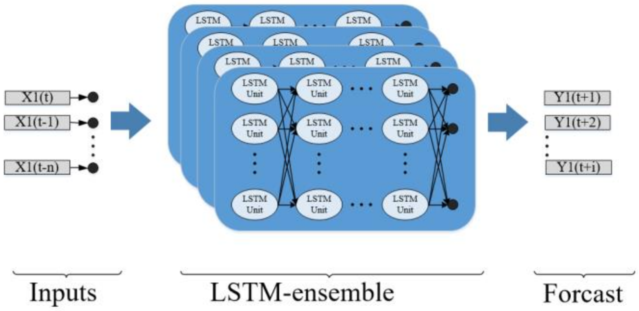

Figure 2: LSTM ensemble forecasting model

Figure 2 depicts the homogenous LSTM ensemble forecasting model. For this work, an ensemble consisting of four LSTMs is applied with various levels of hidden layers. This work limits the single point forecasting models to four in order to try and keep a reduced computational time. The number of hidden layer neurons is randomized at training and retraining iteration. However, boundaries are applied to the randomized values at each layer based on heuristics found during the hyperparameter optimization.

This overall ensemble approach is based on an ASHRAE competition in which a top performing method within the competition foregone hyperparameter optimization, leveraging an ensemble of many single point forecasting models, FFNNs, to produce an overall high performing model (Feuston & Thurtell, 1994). The approach presented in this work differs in that simple heuristics are applied in order to reduce the overall number of single point forecasting models within the ensemble.

Hyperparameter Optimization

In addition to the quantity and quality of the data obtained, the performance of ML models are determined by the hyperparameters of the model. Hyperparameters are such parameters which may be adjusted in the creation of the model (Goodfellow, Bengio, & Courville, 2016). With neural networks a few of such parameters include: input variables, activation functions, length of training, hidden layers, amount of hidden layer neurons, epochs etc. The adjustment of the hidden layer neurons is of great importance for the successful selection of hyperparameters. A neural network with too few connections may fail to learn the underlying representations in the data and as a consequence have large forecasting errors; in contrast, a network with too many neurons may overlearn noise within the data and similarly result in a model with large forecasting errors (Yao, 1999). One of the biggest challenges with hyperparameter selection, is that there is no standard way in order to select the optimal hyperparameters and most approaches rely on manual adjustments. In order to find the amount of hidden layer neurons, a grid search was conducted.

Performance indices

The performance indices presented in this paper are the root means square error (RMSE) and the coefficient of variation of the root mean square error CV(RMSE) shown in equations 7 and 8 (ASHRAE, 2014).

$$\text{RMSE [kW]} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{7}$$

$$\text{CV (RMSE) [\%]} = \frac{\text{RMSE}}{\bar{y}} \times 100 \tag{8}$$

Where $\hat{y}_i$ is the forecasted value of the electric demand, y_i is the recorded value, and $\bar{y}$ is the average of the recorded values over the selected time period.

Case Study

System overview

The case study for this work is the Research Centre for Structural Genomics (GE) of Concordia University located in Montreal, Canada. The building is a multi-level research centre with a total floor area of 5,400 m^2. The research building contains laboratories with fume hoods which make up approximately 30% of the total floor area. Furthermore, approximately 53% of the total floor area is occupied by offices, corridors and conference rooms. Figure 3 provides an illustration for the building.

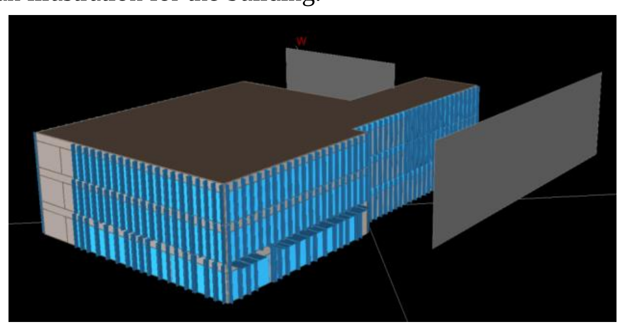

Figure 3: eQuest Case Study building picture take from source (Mihai, 2014)

A calibrated eQuest simulation of the GE building was used as a data source for this study (Mihai, 2014). The building is supplied with cooled water from an electric chiller rated at 1,758 kW (500 ton). A constant speed pump circulates chilled water to the air handling unit of the building at about 7°C. The air handling unit uses a variable speed drive to supply conditioned air at 12^{0}C to 13^{0}C. The chiller is connected to an open cooling tower which rejects the heat from the condenser to the environment. Synthetic data was extracted from an eQuest simulation for the electric demand for major HVAC components at hourly time steps. The total electric demand of the HVAC system was calculated with equation 9 for all hours over the summer period. This calculated data set was then used in order to train and apply the proposed forecasting models. Thus for this work, the target variable to be forecasted is the future electric demand of the HVAC system using the current and past history of the system as inputs to the forecasting model.

$$\dot{E}_{GE,HVAC}^{t+i} = \dot{E}_{fan,sup}^{t+i} + \dot{E}_{CHWS,pump}^{t+i} + \dot{E}_{Chiller}^{t+i} + \dot{E}_{CDS,pump}^{t+i} + \dot{E}_{CTower}^{t+i} \tag{9}$$

Where $\dot{E}_{fan,sup}$ is the electric demand of the supply fans, $\dot{E}_{CHWS,pump}$ is the electric demand of the chilled water supply pump, $\dot{E}_{chiller}$ is the electric demand of the chiller, $\dot{E}_{CDS,pump}$ is the electric demand of the condenser pump, and $\dot{E}_{CTower}$ is the electric demand of the cooling tower.

Exploratory analysis of the GE HVAC systems electric demand

Figure 4 presents the exploratory analysis of the HVAC system over its daily operation. The time interval for the data is from June 1st to August 1st, 2014. The colour of each cell represents the total HVAC electric demand ($\dot{E}_{GE,cooling}$). The red colour corresponds to the high electric demand in kW, while the blue colour corresponds to low demand. The figure demonstrates a range of electric demand for the HVAC system of 60 kW to 90 kW during unoccupied periods and 100 kW to 170 kW during occupied periods.

Hyperparameter analysis

A grid search technique was conducted for both the LSTM and FFNN models on an Intel Core 2.8 GHz CPU, with 8 GB ram, Windows 10, and a 64 bit operating system. Table 2 provides a sample of the search results for the LSTM model. Within Table 2, H1 to H4 refer to the number of hidden neurons in the first hidden layer to fourth hidden layer respectively.

Features of the model were restricted to only the historical electric demand of the HVAC system. Inputs for the model were the previous time lags of that usage. It was observed that the model performed well using the past 24 hours as inputs; that is values at time (t) to (t-23). Therefore, such values will be applied as inputs to the forecasting models in this work.

For this work, the hidden layer neurons were varied at steps sizes 1 to 10, 15, 25, 50, 75, 100, 125, 150, 175, and 200 for single layered neural networks. Furthermore, an initial search demonstrated that larger models with a greater number of LSTM units were observed to have non-convergence issues. Therefore, the boundaries were set in order to avoid such issues.

For two layered LSTM models, the H1 neurons were varied from 5, 10, 25, 50, 75, 100, 125, 150, and 175 for each H2 value in the range of 5, 10, 15, 20, and 25. The reduction in search range for H2 comparable to H1, arose as a result of the computational time which increased significantly with the larger models without a reduction in performance.

For more than two hidden layered LSTM models, the steps sizes for the hidden layer units were reduced further ranging from 5, 10, and 15 each. This was done as the length of training time began to increase drastically. The step sizes were selected based on observations and the desire to keep model development time to a minimum. Smaller steps sizes could have been accomplished, however, such sizes would come with an increase in the development time. Based on the general trends observed in the search with large step sizes, a local grid search could then be conducted with narrower boundary conditions and a smaller step size to identify the top performing single point architectures.

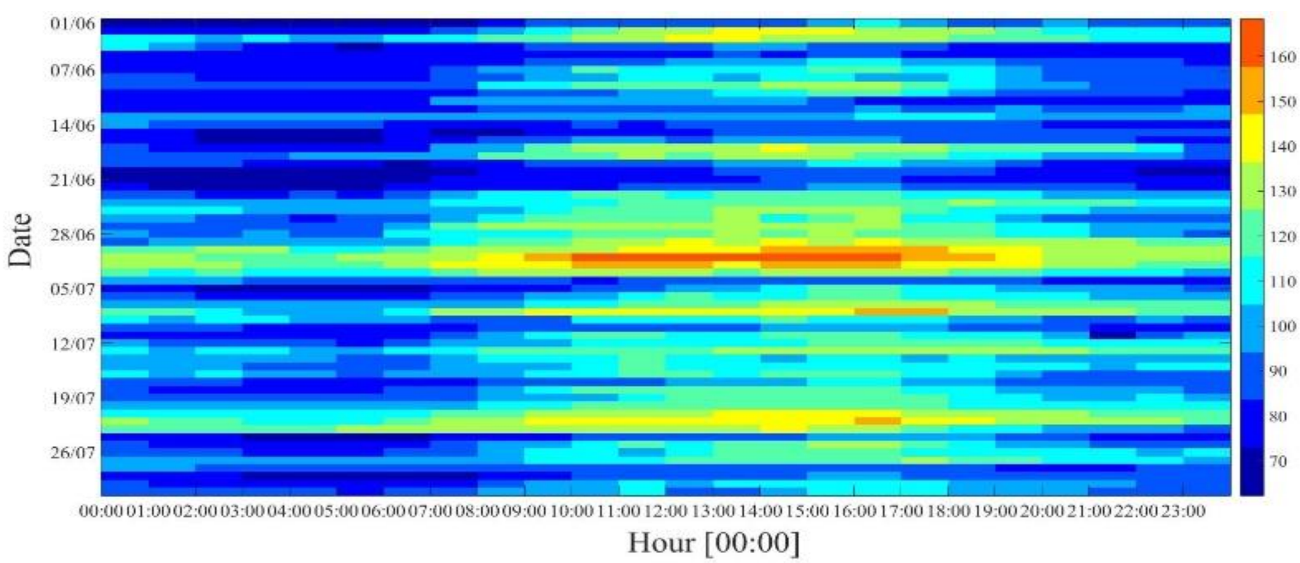

Figure 4: Carpet plot for the electric demand of the HVAC system

Table 2 provides a sample showing the error obtained for architectures during training and the computational time required for each. From Table 2 it can be noticed that the performance did not drastically change with the addition of another hidden layer. It was also observed that increasing the neurons in a hidden layer did not provide a significant reduction in performance. The results found appear consistent with papers exploring LSTM hyperparameter searches in other fields (time series, natural language processing, etc.). Such results were explored over other fields as there are few papers which have explored LSTM hyperparameters searches applied to forecasting electrical loads in buildings. For instance, it was found for time series forecasting of the M4 competition, that increasing the number of hidden layer neurons did not result in lower mean error, rather it led to non-convergence (Peter & Matskevichus, 2019). Furthermore, it was shown that LSTM models are more sensitive to hyperparameters such as epochs, activation function, and random weight initialization while the number of hidden units may be of less importance (Reimers & Gurevych, 2017).

Table 2: Sample of hyperparameter search results for LSTM for the electrical demand of the HVAC system

H1	H2	H3	H4	RMSE (kW)	CV(RMSE) (%)	Training Time (sec)
15	0	0	0	5.75	5.66	198
15	15	0	0	5.68	5.58	309
15	15	15	0	13.72	13.39	481
15	15	15	15	8.57	8.39	860

In this case study, it was observed that the performance for a single layer LSTM model obtained the lowest forecasting error within the training data set with a range of 75 to 125 hidden layer neurons. However, due to space limitation such results are not provided. It should be noted that the single layered LSTM models require less computational time to train for this case study and provided among the best performances. Due to both those reasons, such models were selected to be applied.

The heuristics found through the grid search for the LSTM model, 75 to 125 hidden layer neurons, were then applied as boundary conditions for selecting the number of hidden layer neurons in each of the single point forecasting models within the overall homogenous ensemble.

Furthermore, a local search was conducted finding the best optimal number of hidden layer neurons for a single point forecasting model. This local search consisted of reduced step sizes, 1 hidden layer neuron at a time, around the minima found. The single point LSTM forecasting model will consist of 24 inputs, 100 hidden layer neurons, and six outputs (one to six hour ahead).

A similar local grid search was conducted for the FFNN in order to select the appropriate number of hidden layer neurons. However, the FFNN search was restricted to a single hidden layer in order to compare the results of a conventional and deep learning approach. The single point FFNN forecasting model consists of 24 inputs, with 50 hidden layer neurons, and six outputs.

Results and discussion

This section presents the results of the forecasting models applied. The target variable included the future electric demand for the HVAC system and the input values consisted of the current and previous usages (t to t-23). The outputs for the models consists of the forecast for the next six-hours (t+1 to t+6).

The data was broken into three sections: training, validation, and testing. All forecasting models were trained with data starting from 6/01/14 00:00 until 12-hours prior to the testing data set. The full testing data set was from 30/07/14 09:00 to 01/08/14 00:00. At each hour in the testing data set, the forecasts of the next six-hours were generated by the forecasting models and the performance was recorded comparing the estimated values to eQuest data. The machine learning models were periodically retrained every 12 hours. As new data became available, it was added to the original training data set (i.e., accumulative window approach) and the ML models were retrained on this new, larger accumulated data set.

Therefore, for the first data set the training data consisted of 6/01/14 00:00 to 7/29/2014 20:00, the validation set was from 7/29/2014 21:00 to 7/30/2014 08:00 with forecasts generated hourly from 7/30/2014 09:00 to 7/30/2014 14:00. Forecasts were generated at each hour until 7/30/2014 21:00 when a new batch of data (09:00 to 21:00) was added to the original training set and retraining occurred.

Within this paper, a forecasting set refers to a set of forecasts generated by a specific forecasting model over the next six hours. The paper will use the notation F+x in reference to the forecasting sets, where x refers to the number of the forecasting set. For example, at the first time step of the testing data set 08:00, the LSTM model will generate its forecast set over the next six hours (09:00 to 14:00). The notation for the first forecasting set will then be F+1. As new data becomes available, the ML models are then applied using the new set of inputs in order to generate the next forecast set (10:00 to 15:00), or F+2.

For comparison purposes, a single layer LSTM model is used in addition to a single layer feed forward neural network and a simple forecasting approach (SFA). The SFA uses the recorded value at the time t-24, which is 24 hours before, as a forecast value at time t. For example, the SFA forecast value of the electric demand (kW) on 30/07/14 at 09:00 is equal to the electric demand (kW) one day before that on 29/07/14 at 09:00.

As the forecasting models are applied to the testing data set, the performance is calculated. At each hour, the forecasting models generate a forecast over the horizon (the next six-hours) and the performance is calculated comparing the generated forecasts to the synthetic data. For example, the

CV(RMSE) calculated at F+1. The process then continues over the entire testing data set calculating the forecasting error at each time step (F+2, F+3,…F+n).

After all the forecasting models were applied to the full testing data set and the performance was recorded at each forecasting step, the average performance was then calculated. Table 3 presented the results for the average RMSE (kW) and the average CV(RMSE) obtained for each forecasting model over the testing data set. In addition, the standard deviation of the CV(RMSE) over the testing data set is provided to present the deviation of the performance.

Table 3: Performance indices for the comparison of forecasting approaches over full testing data set

Forecasting Model	RMSE (kW)	CV(RMSE) (%)	Stand deviation of the CV(RMSE)
LSTM (24-100-6)	6.55	6.72	4.45
LSTM ensemble	6.18	6.33	3.31
FFNN (24-50-6)	7.26	7.53	3.70
SFA	9.15	9.39	4.32

The results of this work and case study can be summarized in the following points:

- A slight improved forecasting performance can be seen with the application of deep learning based forecasting models (6.33% compared with 7.53% CV(RMSE)) over the testing data set. Similar findings were found by others with the comparison of DNN models to other forecasting approaches when applied to forecast overall building electrical loads.

  For instance, Amarasignhe et al. (2017) found a reduction of RMSE of 0.028 to 0.189 kW of DL models compared with conventional ML models forecasting the overal electric load for a building (Amarasinghe, Marino, & Manic, 2017).In addition, Cai et al. (2019) which obtained a reduction of CV(RMSE) of 0.4% to 11.15% comparing DL models and time series forecasting models targeting the electric demand of a building (Cai, Pipattanasompom, & Rahman, 2019).

  Furthmore, Bouktif et al. (2020) found a reduction of 0.1625 to 0.668% CV(RMSE) for LSTM forecasting models compared with ML based models (Bouktif, Fiaz, Ouni, & Serhani, 2020).
- The development time needed for the forecasting model (often used mostly in hyperparameter optimization) was significantly reduced with the application of simple heuristics and an ensemble forecasting approach. This was a result of not needing to extend out the search range or conducting a localized search. A trade off occured with the LSTM ensemble approach in which development time was reduced, however, this increased the computational time needed in training when compared to a single point forecasting model (a training time of approximately 16 minutes for the ensemble approach compared with 4 minutes for a single LSTM).
- The application of the ensemble model helped reduced the forecasting error and the standard deviation of the forecasting error over the testing data set.
- All models presented showed acceptable performance with a CV(RMSE of less than 10%) while the ML models had an average CV(RMSE) of less than 8%.

Therefore, depending on accetpable performance criteria for the implementation of each forecasting model, the following can be noted. The SFA was the simplest to implement and achieved adequate performance, however, the error was among the largest and the standard deviation of the CV(RMSE) was also among the largest. The machine learning models achieved improved performance results and reduced fluctuations in error. Specifically, the ensemble model was among the highest performing algorithms, achieved the lowest variations of performance, and had the lowest developmental time. However, it had the drawback of requiring longer computational time for training and at retraining periods.

Conclusion

This paper presents the development of univariate deep-learning based models in order to forecast the electric demand of an HVAC system of an institutional building located in Montreal, Canada. Synthetic data was obtained from a calibrated eQuest model for the building and used to train and test the forecasting models presented herein. The data obtained was on an hourly time step and the forecasting models provided a forecast horizon of up to six hours ahead. The hyperparameters for the forecasting models presented herein were found through a heuristic grid search approach. The results demonstrated that the LSTM ensemble achieved the lowest forecasting error of 6.33%, followed by the single point LSTM model 6.72%, FFNN with 7.53%, and the SFA with 9.39% CV(RMSE). Therefore, it can be concluded that the deep learning models can provide acceptable forecasts in comparison to the target variable.

Future work

Future work will focus on the increase of database in terms of variables, and long-term measurements to assess the performance of the deep learning techniques. This may be achieved through two different approaches (i) to apply multiple input variables for each model and/or (ii) the inclusion of previous years profiles. In addition, exploring the effects of coupling multiple neural network models (autoencoder and LSTM) may be beneficial for exploring the effect of deep learning based techniques. Furthermore, the application of the applied models to measurement data from real HVAC systems would be beneficial to explore the performance of such models, and then compare with results from synthetic (simulation) data.

Acknowledgements

The authors would like to acknowledge the financial support of Natural Sciences and Engineering Research Council of Canada and the Faculty of Engineering and Computer Science of Concordia University.

References

Amarasinghe, K., Marino, D., & Manic, M. (2017). Deep neural networks for energy load forecasting. 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE) (pp. 1483-1488). Edinburgh: IEEE.

Amasyali, K., & El-Gohary, N. (2018). A review of data-driven building energy consumption prediction studies. Renewable and Sustainable Energy Reviews, 81, 1192-1205.

Bouktif, S., Fiaz, A., Ouni, A., & Serhani, M. (2020). Multi-Sequence LSTM-RNN Deep Learning and Metaheuristics for Electric Load Forecasting. Energies.

Cai, M., Pipattanasompom, M., & Rahman, S. (2019). Day-ahead building-level load forecasts using deep learning vs. traditional time-series techinques. Applied Energy, 1078-1088.

Fan, C., Sun, Y., Zhao, Y., Song, M., & Wang, J. (2019). Deep learning-based feature engineering methods for improved building energy prediction. Applied Energy, 240, 35-45.

Feuston, B., & Thurtell, J. (1994). Generalized nonlinear regression with ensemble of Neural Nets: The Great Energy Predictor Shootout. ASHRAE Transactions, 100(2).

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. Cambridge: MIT Press.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

International Energy Agency . (2020, January 24). Buildings A source of enormous untapped efficiency potential . Retrieved January 2020, 2020, from https://www.iea.org/topics/buildings

Jovanović, R., Sretenović, A., & Živković, B. (2015). Ensemble of various neural networks for prediction of heating energy consumption. Energy and Buildings, 94, 189-199.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521, 436-444.

Liu, T., Xu, C., Guo, Y., & Chen, H. (2019). A novel deep reinforcement learning based methodology for short-term HVAC system energy consumption prediction. International Journal of Refrigeration, 107, 39-51.

Makridakis, S., Wheelwright, S., & Hyndman, R. (2003). Forecasting Methods and Applications. New Jersey: Wiley & Sons.

Mihai, A. (2014). Calibration of a Building Energy Model Using Measured Data for a Research. Montreal, QC: Concordia University.

Natural Resources Canada. (2018). Energy Fact Book. Ottawa: Minister of Natural Resources.

Opitz, D., & Maclin, R. (1999). Popular ensemble methods: an empirical study. Journal of Artifical Intelligence Research, 169-198.

Peter, G., & Matskevichus, M. (2019). Hyperparameters Tuning for Machine Learning Models for Time Series Forecasting. International Conference on Social Networks Analysis, Management and Security (pp. 328-323). Granada: IEEE.

Reimers, N., & Gurevych. (2017). Optimal Hyperparameters for Deep LSTM-Networks for Sequence. Retrieved from Cornell University: https://arxiv.org/abs/1707.06799

Runge, J., & Zmeureanu, R. (2019). Forecasting Energy Use in Buildings Using Artificial Neural Networks: A Review. Energies , 12(3254).

Runge, J., Zmeureanu, R., & Le Cam, M. (2020). Hybrid short-term forecasting of the electric demand of supply fans using machine learning. Journal of Building Engineering, 29.

Shan, S., Cao, B., & Wu, Z. (2019). Forecasting the Short-Term Electricity Consumption of Building Using a Novel Ensemble Model. IEEE Access, 7, 88093-88106.

U.S. Energy Information Agency. (2012). Commercial Building Energy Consumption Survery. Retrieved from https://www.eia.gov/consumption/commercial/reports/2012/energyusage/

Wang, Z., & Srinivasan, R. (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.

Yao, X. (1999). Evolving Artificial Neural Networks. Proceedings of the IEEE, 87(9), 1423-1477.