The Impact of Design Space on the Accuracy of Predictive Models in Predicting Chiller Demand Using Short-Term Data

Rodwan Elhashmi*, Kevin P. Hallinan, and Salahaldin Alshatshati
Department of Mechanical and Aerospace Engineering University of Dayton, Dayton, OH
*Corresponding Author's Email: elashmir1@udayton.edu

Abstract— Predicting cooling load is essential for many applications such as diagnosing the health of existing chillers, providing better control functionality, and minimizing peak loads. In this study, short-term chiller and total building demand are acquired for five different commercial buildings in the Midwest USA. Four different machine learning models are then used to predict the chiller demand using the total building demand, outdoor weather data, and day/time information. Two data collection scenarios are considered. The first relies upon use of multiple weeks of data collection that includes very warm periods and season transitional periods where the outdoor temperature ranged from very warm to cool conditions in order to envelope all cooling season weather conditions. The second scenario employs use of contiguous data for a several weeks during only the warmest period of the year. The results show that using two or more separate time periods to envelope most of the weather data yields a much more accurate model in comparison to use of data for only one time period. These research findings have importance to energy service companies which often do short term audits (measurements) in order to estimate potential savings from chiller system upgrades (controls or otherwise).

Keywords— Cooling load, Machine Learning, Building energy consumption

I. INTRODUCTION

The cooling load is associated with the amount of removed heat energy from a space in order to maintain the inside air temperature within an acceptable range for human comfort. Predicting the cooling load for a building is essential for: developing building appropriate heating, ventilating, and air conditioning (HVAC) system design; optimally governing the existing cooling system functionality; evaluating the performance of existing cooling systems, e.g., continuously commissioning; and minimizing peak loads, all while maintaining the desired comfort in the space. The cooling load is influenced by weather conditions, building geometrical and envelope energy characteristics, building usage, building location, occupant schedule and behavior, lighting systems and controls, heating, ventilation, and air conditioning (HVAC) system, controls, and scheduling, and also all sources of heat gains (solar fenestration, people, appliances, infiltration, and heat gain through the envelope).

In practice, many buildings employ building automation systems to manage HVAC and other power systems operations and scheduling; ideally in a way that optimizes the energy performance of the buildings. in general, conventional building controls do not achieve optimal building energy use due to the increase in nonlinearity and complexity of modern HVAC systems [1] thus often the building energy systems manager will circumvent the controller in order to just keep things operational. Optimal performance is generally not realized. A new market for energy service providers (ESCOs) has emerged in the U.S. to leverage data available from Building Automation Systems (BAS) in order to estimate potential savings from their service of improving the HVAC controls to reduce energy. These service providers typically do not have access to the data contained in the BAS system, which may or may not have been archived, or may in fact not exist. So, these ESCOs will put data loggers in buildings for short periods of time to monitor things like chiller power. From this data they attempt to estimate the realizable savings for a year. Ultimately, their value estimate will guide a proposal to the building owner / manager for a performance contract, whereby they guarantee annual savings sufficient to justify their continuous commissioning service. Savings are inferred by comparing actual consumption to the predicted consumption in the absence of ESCO services. Thus, it is essential to be able to accurately forecast annual consumption.

The question this research asks is this: Can short-term chiller and total building demand be used to accurately predict annual chiller demand in order for an ESCO estimate potential energy savings from improvements.

II. BACKGROUND

Many methods have been developed to predict the cooling load in both the design and operational phases of a building. These methods can be generalized into two categories; namely physical based models and data based models. These are described in more detail in the next sub-section.

A. Physical Based Models

Physical based modeling relies upon use of detailed specification of the parameters affecting the building cooling load described previously. It also relies upon mathematical models to predict the thermal behavior of an entire building or specific zone within a building (which mainly includes calculations of heat gain to a building). Presently, there are many commercial softwares available to do such modeling, such as TRNSYS, ESP-r, Trane Trace 7000, and the Department of Energy's EnergyPlus. Even with this software, the calculation of building consumption based on physical models is still challenging. The building level data is often hard to obtain and fought with significant user error. Moreover, this type of modeling requires substantial expertise on the part of the user as well as substantial time to collect the needed data and process. Thus, this type of modeling is often cost prohibitive [2] [3].

Such modeling has been documented to be both inaccurate and inconsistent. For example, Beausoleil-Morrison documented a study where eight expert users were tasked with predicting annual space heating and cooling energy for a building using different simulation tools. The variation in prediction was roughly 20% [4].

Beausoleil-Morrison followed this study with another where users were considered non-expert. This type of user was deemed more realistic than the expert users employed in the prior study. Specifically, he relied upon twenty-one of his students to predict annual cooling and heating load using only two different simulation tools. A wider variation in the predictions was observed (+-30%) [4].

Berkeley et al. conducted such a study premised on this question: "What would happen if expert users employed an identical simulation tool and asked to predict the same loads?" Twelve professionals were given the same energy systems specifications and building drawings. These professionals were asked to predict annual gas and electrical consumption. For electricity, there was as much as 100% difference between predictions. The spread was much bigger for gas energy. Additionally, they found that the relative predictions from a specific expert were consistent, meaning, that when one expert predicted more consumption than others, they did so for all months [5].

Last of all, a 2000 study by Pigg and Nevius compared actual heating energy consumption for 147 houses to predicted space heating using Home Energy Rating System (HERS) software. The results indicated that the simulation tool failed to correctly predict space heating for leaky houses, houses with higher space heating bills, poorly insulated houses, and likely older houses [6]. Moreover, Delzendeh et al. (2017) documented several studies illustrating that the building actual consumption is up to three times higher than simulated building energy consumption [7].

Fundamentally, this review points to the shortcomings associated with physical-based energy modeling. Consumption predictions are seemingly dependent on both individual users and software tools.

B. Data Based Modeling

The burgeoning amount of building energy data is overwhelming researchers who are trying to utilize this data to drive energy reduction. Data based models employ statistical techniques to find the relationship between predictors (any characteristic which influences energy consumption) and the response (energy consumption or demand). This approach is now widely implemented to predict energy consumption with a high degree of accuracy. Some advantages of data based modeling are as follows: 1) it is valid for online application; 2) it has the ability to run different models at the same time and compare the results to get the best model; 3) it has fewer predictors compared to physics-based modeling methods; and 4) it does not require building and energy system details [2] [3].

The problem of estimating cooling and heating loads based on data driven models was first addressed by Kashiwagi and Tobi (1993). The authors used three months (June-August) of measured summer data to develop a Neural Network along with Kohonen's Feature Map and Vector Quantization (LVQ) to predict heating and cooling loads. The validation of the model involved model application to September and October meter periods. The results were very promising [8]. Since that time, estimating cooling load based on data driven models have been thoroughly described in the literature in various ways. Table I provides a summary of the researches conducted to predict chiller power. Shown in the table are the data type, the features used to predict chiller power, the target, the duration of training and testing date used to develop a model.

Table I demonstrates that some researches rely on physical based modeling to generate the training and testing set. The bias that physical based modeling cause was discussed in section II.A. In addition, some studies focused on estimating annual cooling load which can be useful for early design of the HVAC systems for new buildings but does not help with the control strategies or estimating savings from applying energy efficiency measurements. Moreover, Predicting next day load also might not be sufficient for long future plan for upgrading or implementing new retrofit. Finally, the prediction of seasonally hourly load was discussed in two studies but neither study investigate the effect of utilizing multiple time period on predicting cooling load.

In summary, prior studies have sought to develop a model to predict chiller demand strictly from knowable time related data and overall building demand. Additionally, prior research has often been casual in terms of documenting the extent of the training data. Prior research has not considered use of multiple data collection periods in terms of developing a model generalizable to all seasons. Lastly, Prior research has not predicted 15 minutes interval cooling load for entire season.

III. Goals

In this context, we are effectively asking "Can short-term seasonal chiller data taken at one period of time or at multiple periods of time be used to construct a model to predict chiller demand based upon total demand that is generally applicable to long-term (annual) future prediction?" This type of data could be collected by an ESCO in an energy audit over a two- or three-week period or over several periods at different points in time. Doing so will permit chiller health to be assessed in the future simply from whole building demand. Ultimately, this option could represent a low-cost solution for chiller health monitoring in particularly small-sized buildings.

More specifically, this research thus seeks to:

- Predict chiller demand relying solely upon short-term data, taken over two or three periods of time, with machine learning from total load.
- Evaluate and compare four different data mining prediction methods for estimating chiller demand (Boosting, Random Forest (RF), support vector machine (SVM), and artificial neural network (ANN).

Table I: Literature review summary

Article	Data Type	Primarily Engineering Features	Target	Duration of Training Data	Duration of Testing Data
Cooling load prediction in a district heating and cooling system through simplified robust filter and multi-layered neural network [9]	Measured	Temperature, and day of the week	Next day hourly load	Four fifth of five-week data was used for training, while the remainder was used for testing	
Hourly cooling load prediction by a combined forecasting model based on analytic hierarchy process [10]	Calculated based on actual measurement	Load at t − 24 hours, ambient temperatures, and solar insolation data	Next day hourly load	One day	One day
Predicting cooling loads for the next 24 hours based on general regression neural network: methods and results [11]	Measured	Ambient temperature, relative humidity, and solar radiation intensity	Next day hourly load	Building 1: June 2011 to March 2012 Building 2: July 2011 to june 2012	Building 1: April 2012 Building 2: July and august 2012
Cooling-load prediction by the combination of rough set theory and an artificial neural-network Based on data-fusion technique [12]	Measured	Return air temperature, return air relative humidity supply, air temperature supply air relative, humidity supply air pressure, room air temperature, room air relative humidity room, and air pressure,	Next day hourly load	240 hours	24 h for testing 24 h for validation
Development and validation of a simplified online cooling load prediction strategy for a super high-rise building in Hong Kong [13]	Measured	Temperature Relative humidity Solar radiation Cloudiness	Next day hourly load	From mid-june to early august	
Heating and cooling load prediction using a neural network system [8]	Measured	Maximum temperature, and minimum temperature	Average daily cooling load	One month (aug,1992)	Two months (sep. And oct,1992)
An improved office building cooling load prediction model based on multivariable linear regression [15]	Measured	Daily maximum dry-bulb temperature, daily maximum wet-bulb temperature, daily mean dry-bulb temperature, and daily mean wet-bulb temperature.	Average daily cooling load	July 25 to august 31, 2012	July 25 to august 31, 2013 (updated after correcting prediction error)
Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and NARX recurrent neural network: a comparative study on district scale [16]	Measured	Dew point temperature, mean wind direction, mean wind velocity, outdoor temperature, precipitation intensity, precipitation quantity, relative humidity, school holiday time, working time schedule	Monthly loads	Two -years data	
Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks [17]	Simulated	Outdoor dry-bulb temperature Previous hour outdoor dry-bulb temperature Previous 2 hours outdoor dry-bulb temperature Relative humidity Solar radiation intensity Previous hour solar radiation intensity	Seasonally hourly load	July	May, June, august, and October.
Building cooling load forecasting model based on LS-SVM [18]	Simulated	Dry-bulb temperature Relative humidity Solar radiation intensity	Hourly load for 3 months	May and June	July, august, and September
Re-evaluation of building cooling load prediction models for use in humid subtropical area [19]	Simulated	Dry-bulb temperature, horizontal solar radiation, room temperature set point, and cooling loads of previous 4 hours	Predict cooling load for different thermostat set points 22°, 23°, 24°, 25° and 26°	12 cases	48 cases

Modeling heating and cooling loads by artificial intelligence for energy-efficient building design [1]	Simulated	Relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing distribution	Annual cooling load	768 cases
Prediction of heating load and cooling load of buildings using neural network [2]	Measured	Relative compactness, surface area, wall area, roof area, total height, orientation, glazing area, and glazing distribution	Annual cooling load	768 buildings
Early predicting cooling loads for energy-efficient design in office buildings by machine learning [3]	Simulated	No. Of floors, building plan aspect ratio, window-to-wall ratio, outdoor air rate, floor area, window glass u-factor, exterior opaque wall's u-factor, occupant density, lighting density, equipment density, indoor thermostat set point, cation, floor height, and overhang depth	Annual cooling load	243 buildings
Predicting heating and cooling loads in energy-efficient buildings using two hybrid intelligent models [4]	Online source	Relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, and glazing area distribution of a residential building	Annual cooling load	768 sample 80% (614) training and 20% (154) testing

- Demonstrate the value in using multiple short-term measurement periods in developing a model capable of forecasting over a much longer period of time.

The remainder of this paper provides an overview of these techniques, previous research applying these approaching to building and chiller demand prediction, and methodology with results for this research.

IV. Case Studies

With the intent to determine if short-term training data including both building chiller and total demand can be used to develop a forecasting models for chiller demand applicable to long periods of time, different types of buildings are considered. Data from five buildings total comprise this study. The buildings include:

- two university buildings (an academic building and a dormitory)
- one K-12 school buildings
- a place of worship, gathering space, and some educational space; and
- a health care / lab building

Three weeks of measured data for the university buildings was collected; one week during the spring semester, one week between the spring and summer semesters, and one week during the fall. For the rest of the buildings, only two weeks of data during transition months when there is large outside temperature variation as well as significant cooling was used to build the models. For the academic buildings, two training data date periods minimally to account for seasonal occupancy. The measured data for all buildings are summarized in the Table II. Later the value of using multiple training periods relative to a singular one will be investigated. For all of these buildings, chiller and total demand data (15 minutes interval) was collected.

Table II: Available measured data for all buildings.

Building	Training Data	New Data (Testing)
Acad. 1and Acad. 2	Week 1 from 03/02/2017 to 09/02/2017 Week 2 from 08/05/2017 to 08/11/2017 Week 3 from 09/21/2017 to 09/27/2017	03/02/2017 - 10/31/2017
Church	Week 1 from 08/20/2017 to 08/26/2017 Week 2 from 10/30/2017 to11/05/2017	08/18/2017 - 11/15/2017
Health care / lab	Week 1 from 08/17/2017 to 08/23/2017 Week 2 from 10/30/2017 to 11/05/2017	08/02/2017 - 11/06/2017
K-12	Week 1 from 08/17/2017 to 08/23/2017 Week 2 from 10/25/2017 to 10/31/2017	08/02/2017 - 10/31/2017

V. Methodology

The methodology is organized as follows. First, possible predictors for predicting the cooling load are hypothesized. Second, the factor or feature space is defined to establish the predictor 'space' for which a developed model will be valid. This step is one of the most critical elements of the research reported here. The third step is to pre-process the data, by handling missing values, detecting collinearity and multicollinearity, and dealing with categorical variables. Next, various data mining approaches are used to predict model chiller demand using random subsets of data. Finally, using only the identified critical predictors, the data models are tested in estimating chiller performance in a variety of operating conditions. Figure 1 outlines the methodology used in this research.

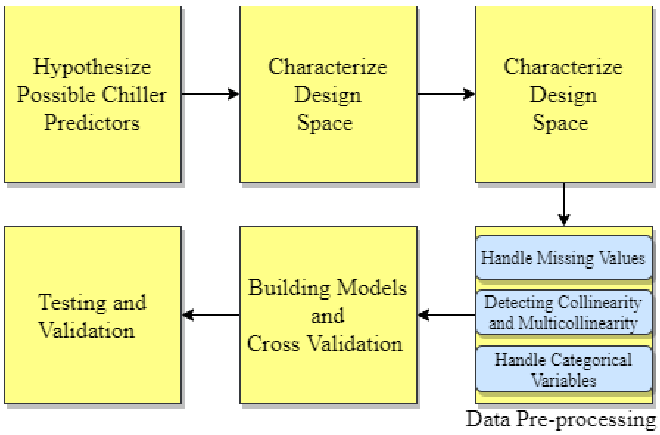


Figure 1: Methodology

A. *Hypothesize Possible Chiller Predictors*

The goal here is to identify and quantify all variables that likely influence chiller demand. Time parameters and weather parameters are both considered, as demand in every one of the buildings depends on the hour of day, day of week, and various weather conditions. Here the potential predictors include: working hours, weekends, holidays, hour of the day index, day of week index, occupancy fraction (for educational buildings only), outside temperature, actual demand, previous 24 hours demand, and previous 24 hours outside air temperature.

The working hours factor is considered binary; it is assigned 1 during typical working hours and 0 otherwise. So too is the weekends variable, which is assigned to 1 if a weekend and 0 otherwise. For the university educational buildings, the occupancy fraction represents the normalized enrollment of students at any term during the year (1 for Fall – Spring semesters, 0.16 for intercessions, and 0.4 for Summer). These values correspond to the fraction of students on campus relative to the normal academic year.

B. *Characterize Design Space*

With potential predictors identified, it is important to characterize the factor space with respect to these variables, as any prediction model of the chiller will not be able to predict chiller performance if the model design space does not envelope the minimum and maximum chiller consumption. All new data must minimally fall between these minimum and maximum bounds.

In order to ensure that all data points are within the design space, the chiller data for the coldest and the hottest weeks are chosen to build the model. From a practical perspective, this would be associated data logging during hot summer months and cold winter months when the outdoor temperature range varies from a most extreme value relative to heating or cooling to a value where heating or cooling is note present. Also, the reults section provides an idea of how much training data is required as well as the requirements for the training data.

C. *Pre-processing Steps*

Before applying the data mining techniques to develop a predictive model, typical pre-processing steps are applied to prepare the datasets. In order to avoid bias errors, the first step is to remove observations from the raw data with missing values in any of the columns.

Collinearity occurs when two of the predictor variables in the dataset are highly correlated. Multicollinearity occurs in a dataset when one variable can be predicted with high accuracy from a linear combination of other variables. This causes the coefficients in the multiple regression to vary widely with small changes to the data, making it difficult to perform calculations with individual predictor variables. If the effects of individual predictors need to be evaluated, small degrees of multicollinearity can significantly affect the analysis. Most of the time multicollinearity has a negative effect in multiple regression models. Further, small sample sizes are more effected by multicollinearity [24].

To correct for multicollinearity, the variable that depends linearly on a set of other variables is removed from the dataset, retaining the variable or variables with highest variable importance. Several tests in R, such as the variance inflation factor (VIF) and the condition number (which is the ratio of max to min eigenvalues), are available to determine whether or not collinearity is present in the dataset. The VIF is calculated by:

$$VIF = \frac{1}{1 - R_i^2} \tag{1}$$

where R_i^2 is the R^2-value obtained by regressing the i^{th} predictor on the remaining predictors.

Lastly, predictors such as working hours, minutes since occupied index, minutes since occupied index, weekends, holidays, hour of the day index, and day of the week index are categorical variables even they have been written as numbers. In order to use them accurately in the models, these variables must be coded as *dummy* variables. For each level of categorical variables, a new dummy variable will be created. This means, if a categorical variable has five levels, then five dummies variables are created to replace this variable.

D. Building Models and Cross Validation

Neural network, Support Vector Machine, Random Forest, and Boosting Random Forest techniques were used to build the model. Cross validation was used to enhance each model. Random Forest does not really need cross validation since the out-of-bag performance for a Random Forest is similar to cross validation. However, since we here compare RF against other models that do not use bagging in the same way, it is wise to use cross validation with all models including RF.

The general procedure of developing a predictive model is to build a model with training data set, and then test the model on one or several validation data sets which weren't used for training. Model overfitting can be a real issue for any regression model. It can lead to an unreal high R-squared value. Model overfitting occurs when the number of predictors are too many compared to the number of observations. An overfit model would often not fit new data, which almost always is associated with error as compared to the original data or may have points that fall in between the training data feature values.

A k-fold cross-validation approach is a good way to avoid model overfitting. In k-fold cross-validation, the training data set is split into k equally exclusive subsets (folds) that have almost the same size [25]. For example, in 10-fold cross validation, the training data is divided into 10 subsets. In the process, the first of the ten folds of the sample is used for testing and the rest for training. The error in testing is calculated. The same procedure is applied to the remaining folds. Lastly, the average error rate from all of the folds is calculated. The goal of using cross validation is to avoid model overfitting.

E. Model Evaluation Metrics

The performance of each algorithm was evaluated based on a new data that does not include the training nor the testing data sets. The following sections show the R-squared value and the mean square error (MSE) for both models and new data for all algorithms and all buildings.. R-squared value and the MSE, defined as equations 2 and 3 respectively:

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} \tag{2}$$

where $SS_{Regression}$ is the sum squared regression error and SS_{Total} is sum squared total error

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y_\iota})^2 \tag{3}$$

Here, n is the number of data points, $\widehat{Y_\iota}$is the observed value, and $\widehat{Y_\iota}$ is the predicted value.

VI. RESULTS

This section is organized to discuss the results of the pre-processing data (collinearity and multicollinearity), model performance for different measurement time periods using different algorithms, and model performance using single one-period training data using different algorithms.

- *Collinearity and Multicollinearity*

After conducting a variance inflation factor (VIF) test to detect multicollinearity, variables that have multicollinearity were excluded from the training set before building the model. The results of VIF test for each building are summarized in Table III. Only two of the buildings exhibited multi-collinearitybetween predictors.

Table III: VIF test results.

Building	Collinear variables that should be excluded	Correlation Coefficients	
		Minimum correlation	Maximum correlation
Acad. 1	No Multicollinearity	Hour of the Day ~ Occupancy Fraction: 0	Bldg. Total kW ~ Previous 24hrs kW: 0.90
Acad. 2	No Multicollinearity	Working Hours ~ Day of The Week: -0.001	Last 24hrs OAT ~ OAT: 0.85
Church	No Multicollinearity	Day of the Week ~ Hour of The Day: -0.001	OAT ~ Previous 24hrs OAT: 0.90
Health care / lab	No Multicollinearity	Previous 24hrs Bldg. kW ~ Hour of The Day:0.012	Previous 24hrs OAT ~ Total Bldg. kW: 0.85
K-12. 1	No Multicollinearity	Day of the Week ~ Hour of The Day: 0	Is Weekends ~ Day of the Week: 0.79

- *Model Performance for Different Measurement Time Periods*

In this section, results are presented to show the potential for developing models to predict chiller using only total demand and demonstrating the value of using multiple training periods to develop a more generalizable model.

The performance of the four machine learning algorithms was tested and evaluated for each of the buildings employing all of the training data available. Table 3 shows the validation metrics for each building. It also shows the error metrics for model application to longer-range new data. The duration and extent of this new data was defined in Table 1 for each of the buildings.

Above all, Table IV shows that there is not one algorithm that provides the absolute best predictive model based upon short duration training data capable of forecasting extended periods. The best models for the other buildings generally performed well, although the statistical metrics for application of the models to new data was in all cases below the similar metrics for model training validation. In some of the cases (see Academic Building 2), the testing metrics were well below the model training validation metrics. Additionally, as shown in Figure 2 which shows plots of the predicted versus actual chiller demand as applied to the extended new data, none of the models were effective in predicting 0 actual demand. The plot for Academic Building 2 particularly reveals the inability of the developed model to accurately predict the chiller demand.

Table IV: R^2 and MSE for NN, SVM, RF, and B for training data and new data

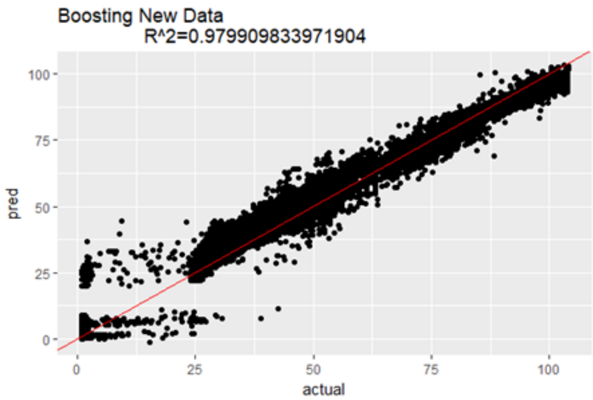
Building	Algorithms	Training Metrics		Testing Metrics for Model Application to New Data	
		R2	MSE	R2	MSE
Acad. 1	NN	99.8 %	1.70	94.8 %	45.92
	SVM	99.6 %	3.56	94.8 %	45.99
	RF	99.8 %	2.08	96.6 %	29.90
	B	99.8 %	1.74	97.9 %	18.04
Acad. 2	NN	95.1%	1.63	43.6 %	11.72
	SVM	98.9 %	0.37	55.9 %	9.16
	RF	98.5 %	0.48	74.4 %	5.31
	B	95.2 %	1.60	55.7 %	9.20
Church	NN	99.9 %	0.46	99.4 %	14.96
	SVM	99.2 %	17.31	91.9 %	219.02
	RF	99.6 %	8.40	97.7 %	60.40
	B	99.7%	5.71	99.2 %	21.10
Health care / lab	NN	92.4 %	669.26	81.0 %	932.13
	SVM	99.1 %	72.08	73.0 %	1324.71
	RF	99.3 %	59.40	88.8 %	549.18
	B	99.1 %	77.47	88.2 %	577.69
K-12. 1	NN	83.0 %	108.43	86.1 %	47.06
	SVM	84.0 %	101.76	76.9 %	78.12
	RF	84.2 %	100.34	88.1 %	40.12
	B	86.4 %	86.60	88.2 %	39.88

Moreover, almost all models poorly predicted zeros values. This might due to the fact that the data used here is 15 minutes interval period where we see a lot more zeros compared to hourly data. Possibly, this issue could be solved by either applying moving average filter or use or by using two step machine learning where first we use classifier algorithm to predict zeros and no-zeros values. Then use regression to predict all non-zero value.

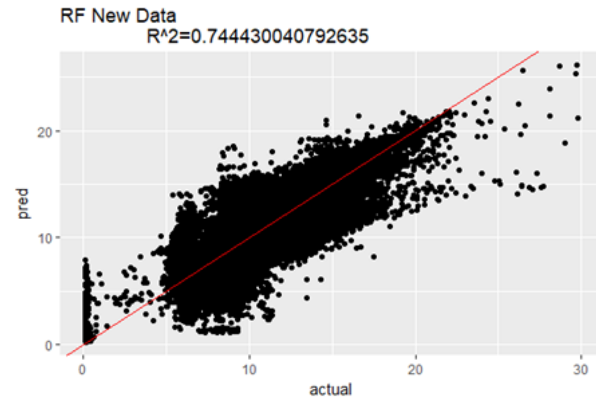
The question is "Why are some of the models developed from short duration measurements of chiller and total demand more effective in predicting future chiller demand? Figures 3 and 4 help to illustrate why. Figure 3 shows the training data distributions and the new data distributions for the most important features for Academic Building 2, for which the developed model performed the worst on the new data. Figure 4 show similar data for the building with the best performing model on the new data.

Figure 3 shows why the developed model from short-term measured data failed to forecast chiller demand for an extended period. Whereas the training and new data probability density distributions for outdoor temperature and total demand were comparable, the occupancy fraction distributions for the training and new data were quite different. The new data distribution for occupancy fraction included data during the summer semester for which there was no data in the training set. This period of time was associated with an occupancy fraction of 40% relative to the regular academic year.

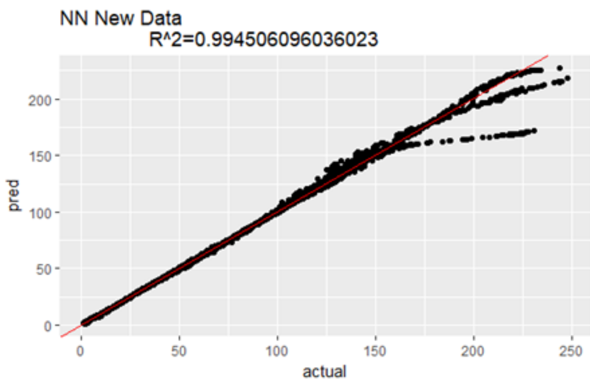
In contrast, the church building has consistent occupancy throughout the year. the training and new data probability density distributions for outdoor temperature and total demand are comparatively similar. This means that the training data well represent the entire season as shown in Figure 4.



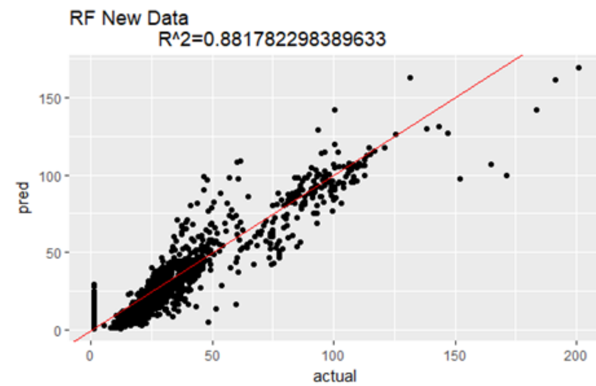
Acad. 1



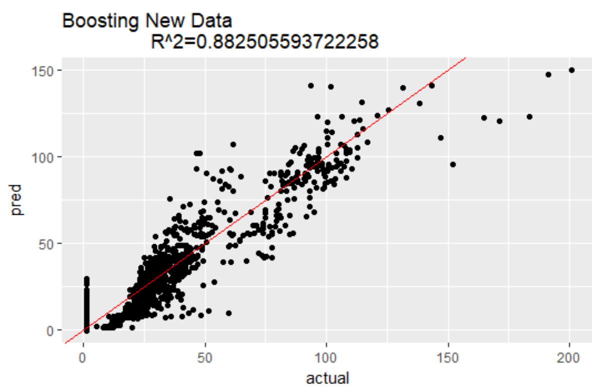
Acad. 2



Church



Health care / lab



K-12

Figure 2: Best models results for actual chiller demand vs. predicted chiller demand for the new data

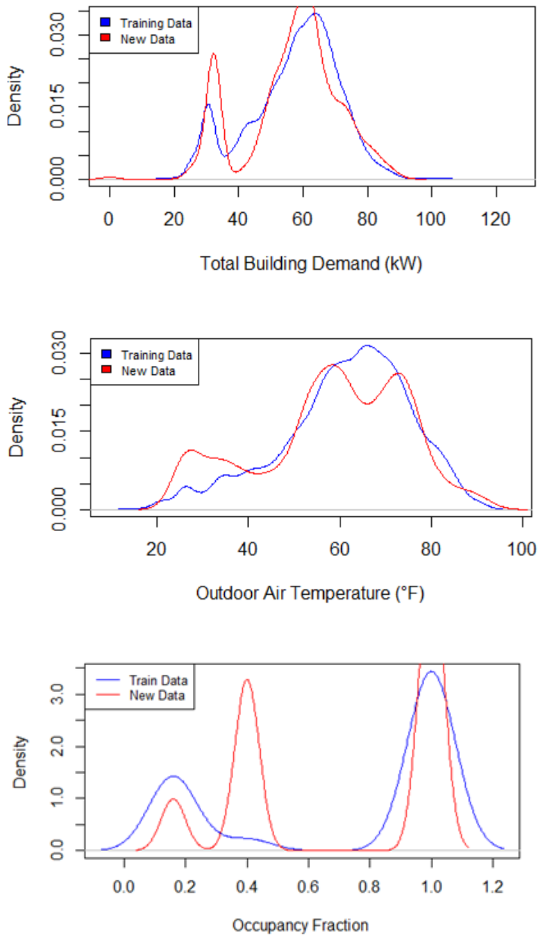


Figure 3: Probability density of most important predictors for training data and new data for Academic Building 2

- *Model Performance Using Single Time Training Data Period (continuous period for three weeks for Acad. 1and Acad. 2 and two weeks for the remaining buildings)*

The selection of the data used to build the final model requires full understanding of building scheduling and behavior in order to help the algorithm find the underlying patterns of the training data. In this study for example, the training data used to develop the chiller demand models described in the previous section included data taken from multiple weeks that included very warm periods and season transitional periods where the outdoor temperature ranged from very warm to cool conditions. In general, the training data seemed to well represent the long-term weather data and thus the developed model generally performed well except for Academic Building 2. However, for the university buildings, three weeks of measured data was used because the scheduling for university buildings is different and two weeks only were not enough to represent the whole season.

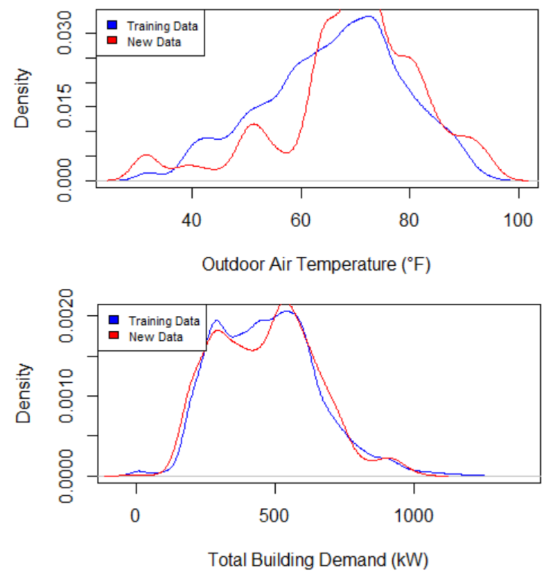


Figure 4:Probability density of most important predictors for training data and new data for Church building

In this section, the authors did not consider the design space. Instead of using two separate time periods to envelope most of the data points, we used continuous weeks the represent the warmest time of the year. In this context, for Acad. 1 and Acad. 2 buildings three continuous weeks were used to build the model while only two continuous weeks were used to build the models for the remaining buildings. Table V summarizes the training and testing data periods for each of the buildings.

Table VII shows the training validation metrics and the error metrics associated with application of the model to extended new data period. Clear from this table is that the models developed, excluding that developed for the church building, do not accurately forecast chiller demand. The MSE compared to that reported in Table IV using multiple training periods has increased by 295%, 184.55 %, 533.30%, and 25 % for Academic Building 1, Academic Building 2, the Church Building, the Health care / lab, and the K-12 Building respectively.

Table V: Training and testing data periods for each of the buildings (continuous period)

Building	Training Data	New Data
Acad. 1and Acad. 2	8/8/2017 - 08/29/2017	03/02/2017 - 10/31/2017
Church	8/19/2017 - 9/2/2017	08/18/2017 - 11/15/2017
Health care / lab	08/17/2017 – 08/31/2017	08/02/2017 - 11/06/2017
K-12. 1	08/17/2017 – 08/31/2017	08/02/2017 - 10/31/2017

TABLE VII: R2 and MSE for NN, SVM, RF, and B for training data and new data

Building	Algorithms	Training Validation Error Metrics		Testing Error Metrics for New Data	
		R2	MSE	R2	MSE
Acad. 1	NN	99.10 %	2.58	88.16 %	113.68
	SVM	99.40 %	1.89	59.80 %	386.05
	RF	98.40 %	4.72	67.60 %	311.09
	B	99.00 %	2.86	92.60 %	71.27
Acad. 2	NN	91.70 %	1.19	20.50 %	16.75
	SVM	97.00 %	0.43	- 4.30 %	21.98
	RF	95.20%	0.68	28.30 %	15.11
	B	94.30 %	0.81	-5.30 %	22.18
Church	NN	99.99 %	0.53	97.50 %	94.74
	SVM	99.50 %	13.95	68.30 %	1095.82
	RF	99.90 %	3.92	95.00 %	173.14
	B	99.85 %	4.34	96.40 %	124.14
Health care / lab	NN	96.10 %	103.27	69.00 %	1522.58
	SVM	97.80 %	58.05	37.20 %	3087.37
	RF	97.40 %	69.53	67.10 %	1618.16
	B	97.00%	80.11	67.70 %	1589.78

VII. CONCLUSIONS AND DISCUSSION

Different modeling approaches (Physical Based Model and Data Based Model) for predicting cooling load were reviewed and compared based on literature reports. Physical based model seems to be the only feasible approach for new buildings since it does not required measured data. In addition, physical based modeling requires detailed information about building envelope and HVAC system. This information is not required in data based models. Moreover, data based model is applicable for online application while physical based model is not. In term of accuracy, physical based model appears to have a substantially lower accuracy compared to data based model.

In this study, a data based model approach was implemented to predict chiller kWh and chillier kW for 5 different buildings using two different serious for the training set. The first scenario is data taken from multiple weeks that included very warm periods and season transitional periods where the outdoor temperature ranged from very warm to cool conditions. While the second scenario is to utilize continuous data that represent the warmest period of the season only. In both scenarios, the length of the training data is the same. For university buildings, three weeks measured chiller data was used train the model whereas of two weeks for measured chiller data for each building was used to train the model. The results show that short period data (2-3 weeks) can possibly present long time period (one year for example). However, these short-term data have to be taken for the warmest and coldest weeks to envelope all data points. The ability of the training data to forecast the future depends upon how it reflects all future data.

Finally, based on the results from the privous section, the best and the worst models in predicting new data varies from building to another. This leads us to conclude that there is not an absolute best predictive model that work for all data.

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

VIII. REFERENCES

[1] C. Jui-Sheng and B. Dac-Khuong, "Modeling heating and cooling loads by artificial intelligence forenergy-efficient building design," *Energy and Buildings*, vol. 82, pp. 437-446, 2014.

[2] V. V. Mokeev, "Prediction of Heating Load and Cooling Load of Buildings Using Neural Network," in *2019 International Ural Conference on Electrical Power Engineering (UralCon)*, Chelyabinsk, Russia, 2019.

[3] N.-T. Ngo, "Early predicting cooling loads for energy-efficient design in office buildings by machine learning," *Energy and Buildings*, vol. 182, pp. 264-273, 2019.

[4] D. T. Bui, H. Moayedi, A. I. Dounis and L. K. Foong, "Predicting heating and cooling loads in energy-efficient buildings using two hybrid intelligent models," *Applied Sciences*, vol. 17, p. 3543, 2019.

[5] F. Behrooz, A. R. Ramli and K. Samsudin, "A survey on applying different control methods approach in building automation systems to obtain more energy efficiency," *International Journal of the Physical Sciences*, vol. 6, pp. 2308-2314, 2011.

[6] X. Li and J. Wen, "Review of building energy modeling for control and operation," *Renewable and Sustainable Energy Reviews*, vol. 37, p. 517–537, 2014.

[7] F. M. Hai-xiang Zhao, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586-3592, 2012.

[8] B.-M. Ian, Composer, *The world needs building simulation : Are we ready?.* [Sound Recording]. 2016.

[9] P. Berkeley, P. Haves and E. Kolderup, "IMPACT OF MODELER DECISIONS ON SIMULATION RESULTS," in *ASHRAE/IBPSA-USA Building Simulation Conference*, Atlanta, GA, 2014.

[10] B. Polly, N. Kruis and D. Roberts, "Assessing and Improving the Accuracy of Energy Analysis for Residential Buildings," U.S. Department of Energy (DOE), 2011.

[11] E. Delzendeh, S. Wu, A. Lee and Y. Zhou, "The impact of occupants' behaviours on building energy analysis: A research review," *Renewable and Sustainable Energy Reviews*, vol. 80, pp. 1061-1071, 2017.

[12] K. Norihito and T. Toshikazu, "Heating and Cooling Load Prediction Using a Neural Network System," in *International Joint Conference on Neural Networks*, 1993.

[13] S. M., U. S., K. K. and O. K., "Cooling load prediction in a district heating and cooling system through simplified robust filter and multi-layered neural network," in *IEEE Xplore*, Tokyo, Japan, 1999.

[14] Y. Ye, L. Zhiwei, L. Shiqing and H. Zhijian, "Hourly cooling load prediction by a combined forecasting model based on Analytic Hierarchy Process," *International Journal of Thermal Sciences*, vol. 43, no. 11, pp. 1107-1118, 2004.

[15] S. Yuying, W. Wei, Z. Yaohua and S. Pan, "Predicting Cooling Loads for the Next 24 Hours Based on General Regression Neural Network: Methods and Results," *Advances in Mechanical Engineering*, vol. 2013, 2013.

[16] H. Zhijian, L. Zhiwei, Y. Ye and Y. Xinjian, "Cooling-load prediction by the combination of rough set theory and an artificial neural-network based on data-fusion technique," *Applied Energy*, vol. 83, no. 9, pp. 1033-1046, 2006.

[17] S. Yongjun, W. Shengwei and X. Fu, "Development and validation of a simplified online cooling load prediction strategy for a super high-rise building in Hong Kong," *Energy Conversion and Management*, vol. 68, pp. 20-27, 2013.

[18] C. Fan, F. Xiao and Y. Zhao, "A short-term building cooling load prediction method using deep learning algorithms," *Applied Energy*, vol. 195, pp. 222-233, 2017.

[19] G. Qiang, T. Zhe, D. Yan and Z. Neng, "An improved office building cooling load prediction model based on multivariable linear regression," *Energy and Buildings*, vol. 107, pp. 445-455, 2015.

[20] D. Koschwitz, J. Frisch and C. v. Treeck, "Data-driven heating and cooling load predictions for non-residential buildings based on support vector machine regression and NARX recurrent neural network: a comparative study on district scale," *Energy*, vol. 165, pp. 134-142, 2018.

[21] L. Qiong, M. Qinglin, C. Jiejin, Y. Hiroshi and M. Akashi, "Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks," *Energy Conversion and Management*, vol. 50, no. 1, pp. 90-96, 2009.

[22] X. Li, J.-h. Lu, L. Ding, G. Xu and J. Li, "Building Cooling Load Forecasting Model Based on LS-SVM," in *Asia-Pacific Conference on Information Processing*, 2009.

[23] L. Zhengwei and H. Gongsheng, "Re-evaluation of building cooling load prediction models for use in humid subtropical area," *Energy and Buildings*, vol. 62, pp. 442-449, 2013.

[24] T. Baguley, Serious Stats: A guide to advanced statistics for the behavioral sciences, Macmillan Education UK, 2012.

[25] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 1995.