SAND2016-7252J

Next Day Building Load Predictions based on Limited Input Features Using an On-Line Laterally Primed Adaptive Resonance Theory Artificial Neural Network

C. Birk Jones[a,1], Matt Robinson[b], Yasser Yasaei[c], Thomas Caudell[c], Manel Martínez-Ramón[c], Andrea Mammoli[b]

[a] *Photovoltaic & Grid Integration Group, Sandia National Laboratory, Albuquerque, New Mexico*
[b] *Department of Mechanical Engineering, University of New Mexico, Albuquerque, New Mexico*
[c] *Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, New Mexico*

Abstract

Optimal integration of thermal energy storage within commercial building applications requires accurate load predictions. Several methods exist that provide an estimate of a buildings future needs. Methods include component-based models and data-driven algorithms. The current work implemented a previously untested algorithm for this application that is called a Laterally Primed Adaptive Resonance Theory (LAPART) artificial neural network (ANN). The LAPART algorithm provided accurate results over a two month period where minimal historical data and a small amount of input types were available. These results are significant, because common practice has often overlooked the implementation of an ANN. ANN have often been perceived to be too complex and require large amounts of data to provide accurate results.

The LAPART neural network was implemented in an on-line learning manner. On-line learning refers to the continuous updating of training data as time occurs. For this experiment, training began with a singe day and grew to two months of data. This approach provides a platform for immediate implementation that requires minimal time and effort. The results from the LAPART algorithm were compared with statistical regression and a component-based model. The comparison was based on the predictions linear relationship with the measured data, mean squared error, mean bias error, and cost savings achieved by the respective prediction techniques. The results show that the LAPART algorithm provided a reliable and cost effective means to predict the building load for the next day.

Keywords: Energy storage, load predictions, optimization, Artificial Neural Network, Gaussian Process, Laterally Primed Adaptive Resonance Theory

---

1. Introduction

Commercial buildings waste a considerable amount of energy in the United States [1]. A major contributor are the heating, ventilation, and air conditioning (HVAC) systems [2]. Therefore, in order to mitigate consumption and increase the feasibility of renewable energy integration, demand side energy optimization of HVAC systems has been gaining attention [3, 4]. Tools such as the Distributed Energy Resources Customer Adoption Model (DER-CAM) have been implemented into actual building control systems [5]. The optimization, provided by DER-CAM, define next day operations that are based on local utility tariffs, storage capacities, and building energy load requirements. Estimating the energy load requirements can be a difficult, but very important task for the optimization [6]. Three approaches for completing this task have been highlighted in the literature and include statistical regression, component-based models, and machine learning techniques [7].

The implementation of a particular method for estimating the next day's load is based on two essential considerations. First, can the predictive model produce accurate results. This is critical, because the model provides an estimate, and costs will increase as the accuracy of the estimate decreases [3]. The second consideration is the initial implementation cost for the prediction model. These two factors are combined to determine the feasibility for implementation by calculating the return on investment (ROI). To insure a high ROI the predictive model must be developed and applied in a timely and accurate manner.

Existing literature that used data driven prediction tools, such as machine learning techniques, acquired at least one year of data to train the particular models [3]. The acquisition of one year of data is necessary to train the learning algorithm so that it understands different working conditions throughout the year. Yet, in many cases facilities do not have significant amounts of actual data that can be used to train these models [8]. This has convinced some to use a component-based model created in a well known software such as TRNSYS, eQuest, or EnergyPlus [9]. However, this approach is often not cost effective. The development of energy models is expensive because a large number of input parameters are required and the process of collecting and applying physical elements is time consuming [8]. Moreover, the accuracy of a given physical model can vary by 50% [10]. For instance, Zhao and Magoules found that component-based models had large variations in accuracy and required considerable upfront investments [11]. The learning-based approaches, on the other hand, have a considerably low upfront cost because they can learn system behavior without prior knowledge.

1.1. Laterally Primed Adaptive Resonance Theory

Artificial neural network (ANN) algorithms are a form of machine learning that emulate a simplified version of an animal's nervous system. The algorithms

[1]P.O. Box 5800, Mail Stop 0951, Albuquerque, New Mexico 87185 (cbjones@sandia.gov)

are used for acquiring and storing knowledge by learning system behaviors during a training process. Then, the algorithm can be subjected to new previously unseen data and provide system performance predictions. The predictions can be generalized, which means that the ANN can provide reasonable outputs for inputs not encountered during training. The ANN are also very advanced and can solve complex problems that have either linear or nonlinear behavior.

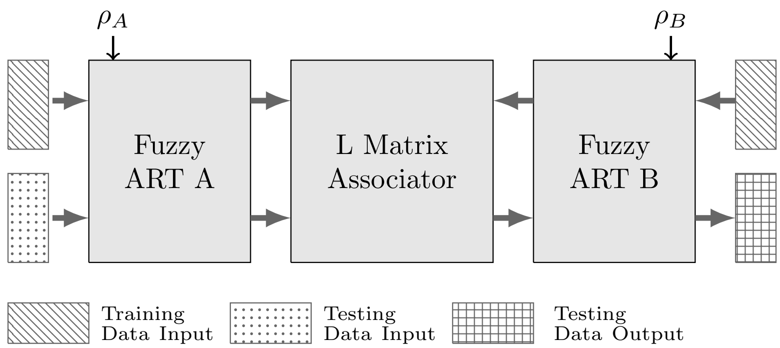

Figure 1: LAPART algorithm training uses two Fuzzy ART (A&B) algorithms connected by an associator matrix (L). During training inputs are applied to both the A and B sides. The algorithm produces A and B templates. It also produces an L matrix that link the templates in the A and B side to one another. During testing the B side learning is turned off and only A side inputs are applied. The inputs resonate with the stored weights in the A, and through an association in the L propagate to the B side template that provides the prediction output.

ANN have been applied to evaluate energy systems. For instance, ANN have been used to predict building cooling loads. Kalogirou *et al.* used a back propagation neural network to estimate loads [12]. In addition, Ben-Nakhi and Mahmoud used a general regression neural network [13] for the same task. One such algorithm that has not been applied to predict building cooling loads is the Laterally Primed Adaptive Resonance Theory (LAPART) neural network. The LAPART neural network steps beyond the current state-of-the-art and provides accurate results with minimal historical data and a small amount of input types.

The LAPART algorithm was introduced by Healy and Caudell for logical inference and supervised learning [14]. The algorithm can be used as a prediction tool and has been show to provide accurate weather forecasts [15]. It has also been applied successfully to solar micro-forecasting to predict solar irradiance at small time steps [16]. LAPART was applied to building system analysis and was used to detect faults within an air handling unit [1].The architecture couples two Fuzzy adaptive resonance theory algorithms to create a mechanism for making predictions that is based on learned associations. This approach has a unique stability that allows the algorithm to converge rapidly unlike the multi-layer perceptron. It differs from the popular multi-layer perceptron because it does not rely on the gradient descent method to converge to the optimal solution. Additionally, the gradient descent approach is susceptible to issues that could result in slow or incorrect convergence [17].

The LAPART architecture couples two Fuzzy Adaptive Resonance Theory (ART) [18] algorithms to create a mechanism for making predictions. The coupling of the two Fuzzy ARTs to create the LAPART algorithm is described graphically in Figure. 1. The A and B Fuzzy ARTs are connected through the L matrix that associates the A and B memory templates. Each Fuzzy ART has its respective vigilance parameters ρA and ρB, and during the learning process inputs are presented to the A and B side simultaneously. The A and B side create and update their memory while at the same time producing links between one another. The updating of weights on the A and B side occurs by first finding the template, $\mathbf{T}_j$, that satisfies Equation 1.

$$\frac{|\mathbf{X} \wedge \mathbf{T}_j|}{|\mathbf{T}_j|} \geq \rho \tag{1}$$

Equation 1 finds the templates where the minimum value of the normalized and complement coded input matrix ($\mathbf{X}_j$) and $\mathbf{T}_j$ divided by $\mathbf{T}_j$ is greater than or equal to the respective free parameter ρ. If multiple templates satisfy Equation 1, the best template is identified as the one with the max value $\mathbf{c}$ which is computed by dividing the minimum of $\mathbf{X}$ and $\mathbf{T}_j$ divided by a choice parameter (α), which was set to be 0.0000001, plus $\mathbf{T}_j$ as shown in Equation 2.

$$\mathbf{c} = \frac{|\mathbf{X} \wedge \mathbf{T}_j|}{\alpha + |\mathbf{T}_j|} \tag{2}$$

The template that resonated with the given input $\mathbf{X}_j$ was updated based on Equation 3.

$$\mathbf{T}_j^{update} = \beta(\mathbf{X} \wedge \mathbf{T}_j^{old}) + (1-\beta)\mathbf{T}_j^{old} \tag{3}$$

The update was the minimal value of $\mathbf{X}$ and the old $\mathbf{T}_j$ multiplied by the learning rate, β. In this case the learning rate was set to 1. However, if the input value did not resonate with any template, then a new template was created that was equal to the input as shown in Equation 4.

$$\mathbf{T}_j^{new} = \mathbf{I} \tag{4}$$

After training was complete, testing inputs were only applied to the A side and were allowed to resonate with the previously stored memory. Then the associations in the L matrix were used to connect with the B side memory and provide the prediction outputs.

1.2. Comparison Methods

Next day load predictions can be performed by many different types of process history and component-based models. Therefore, this paper includes a comparison of the LAPART results with other common techniques. The other methods include a component-based and two empirical models. The component-based model used in this experiment was a commercially available building energy simulation software. The two empirical models were least squares regression and a gaussian process regression learning algorithm. The building energy simulation software was called *TRNSYS* which stands for Transient System Simulation.

1.2.1. Component-Based Model

TRNSYS is a flexible graphically based software environment used to simulate the behavior of transient systems using algebraic and differential equations. In this case, the tool was used to estimate actual air handling unit loads by considering required air flows and temperatures. The simulations considered building geometries, external and internal loads, weather, controls, and ventilation rates. The set-up of the model began with the creation of a three dimensional building in *Trimble Sketch Up* as shown in Figure 2. Building geometries and fenestration locations were replicated to scale and were oriented in the correct cardinal direction. This file was then imported into the *TRNSYS* simulation software where internal loads, and material types and thickness for walls, roofs, ceilings, windows, and other elements were assigned. Then the model was arranged to accept certain inputs such as time, weather, ventilation rates and occupancy.

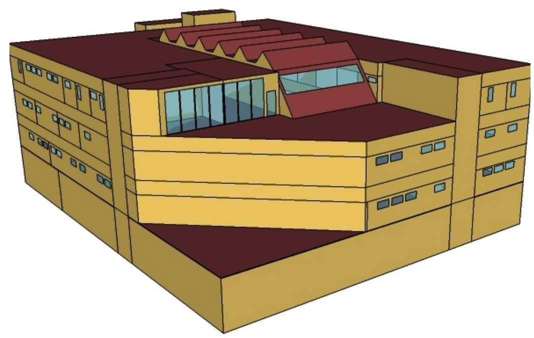

Figure 2: Geometric physical model of the UNM Mechanical Engineering building.

The simulation effort required actual weather, occupancy estimates, and HVAC operations schedules and control parameters as inputs. More specifically, the outside air temperature, relative humidity, and solar irradiance for the simulation time period where provided to the model and defined the external environmental conditions. The model output a single load prediction for each hour of the next day.

1.2.2. Least-Squares Regression

The regression considered is multiple linear regression. It is well known that building thermal loads are nonlinear but linear models are so common that exclusion in the present study would yield an incomplete work. The independent variables in the model are outside air temperature (°C), time of day (hour), and occupancy (percentage of total occupants). It has been shown in other works that these are the primary drivers of the thermal load (particularly in dry climates and thick concrete buildings and few windows). The relationship for the linear model is therefore

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{5}$$

where $\boldsymbol{y}$ is a column vector of the cooling load, $\boldsymbol{X}$ is a matrix of the input variables, $\boldsymbol{\beta}$ is a column vector of coefficients, and $\boldsymbol{\epsilon}$ is a random error term. The coefficients can be estimated by using ordiinary least squares and assigning the hat notation to the estimates.

$$\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}} \tag{6}$$

$\hat{\beta}$, the estimate of the column vector of coefficients, can be obtained by the least squares method.

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X'}\boldsymbol{X})^{-1}\boldsymbol{X}\boldsymbol{y} \tag{7}$$

Eqn. 7 is calculated each day in the experimental window of the year. Each time there are 24 more data pointys added to the set.

1.2.3. Gaussian Process

The GP establishes a probabilistic model over the estimator as shown in Eqn. 8.

$$y_n = \mathbf{w}^\top \phi(\mathbf{x}_n) + \epsilon_n \tag{8}$$

The nonlinear transformation, $\phi(\cdot)$, was used to map the input into a Kernel Reproducing Hilbert Space. In this case, the Mercer theorem [**?**] was used by applying a dot product

$$k(\mathbf{x}, \mathbf{x'}) = \phi(\mathbf{x})^\top \boldsymbol{\Sigma} \phi(\mathbf{x'}) \tag{9}$$

between the transformed observations known as the covariance properties. $\boldsymbol{\Sigma}$ is a positive semidefinite matrix if and only if $k(\cdot, \cdot)$ is a positive semidefinite function.

The error or noise term ε is modelled through a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ and is independent and identically distributed. A priori over $\mathbf{w}$ is established with the form of a Gaussian random variable of zero mean and covariance matrix $\boldsymbol{\Sigma}_w$. The expression of the covariance $\mathbb{E}(y_n y_m|\mathbf{X})$ is

$$\begin{aligned}
\mathbb{E}(y_n y_m|\mathbf{X}) &= \mathbb{E}\left((\mathbf{w}^\top \phi(\mathbf{x}_n) + \epsilon_n)(\mathbf{w}^\top \phi(\mathbf{x}_m) + \epsilon_m)\right) \\
&= \phi^\top(\mathbf{x}_n)\mathbb{E}(\mathbf{w}\mathbf{w}^\top)\phi(\mathbf{x}_m) + \mathbb{E}(\epsilon_m \epsilon_m) \\
&= \phi^\top(\mathbf{x}_n)\boldsymbol{\Sigma}_w \phi(\mathbf{x}_m) + \mathbb{E}(\epsilon_m \epsilon_m) \\
&= k(\mathbf{x}_n, \mathbf{x}_m) + \sigma^2 \delta(m - n)
\end{aligned} \tag{10}$$

where $\mathbf{X}$ is a column matrix containing all input observations $\mathbf{x}_n$. Since $\boldsymbol{\Sigma}_w$ is positive definite, then $k(\cdot, \cdot)$ is a dot product and the covariance matrix can be written as

$$\mathbf{K}_{y,y} = \mathbf{K} + \sigma \mathbf{I} \tag{11}$$

where $\mathbf{K}$ is the matrix of kernel dot products between observations.

A predictive distribution can be constructed for a test sample $\mathbf{x}_*$ whose predicted output is $\mathbf{f}_* = \mathbf{w}^\top \phi(\mathbf{x}_*)$. This can be solved by computing the joint probability distribution of the training and test samples, which is a Gaussian with zero mean and covariance

$$\begin{pmatrix} \mathbf{K}_{*,*} & \mathbf{K}_{y,*} \\ \mathbf{K}_{*,y} & \mathbf{K}_{y,y} \end{pmatrix} \tag{12}$$

where $\mathbf{K}_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$ and $\mathbf{K}_{*,y} = \mathbf{K}_{y,*}^\top$ is the row vector of all dot products $k(\mathbf{x}_*, \mathbf{x}_n)$. The *predictive* posterior over the new sample $\mathbf{x}_*$ given the training data can be computed using the Baye's rule, resulting in a Gaussian with mean and variance given by

$$\begin{aligned} \boldsymbol{\mu}_* &= \mathbf{K}_{y,*}\mathbf{K}_{y,y}^{-1}\mathbf{y} \\ \boldsymbol{\sigma}_*^2 &= \mathbf{K}_{*,*} - \mathbf{K}_{y,*}^\top \mathbf{K}_{y,y}^{-1}\mathbf{K}_{y,*} \end{aligned} \tag{13}$$

The GP then offers a posterior distribution over the prediction rather than a prediction alone. Its variance is reduced with respect to the variance of its prior in a quantity $\mathbf{K}_{y,*}^\top \mathbf{K}_{y,y}^{-1}\mathbf{K}_{y,*}$. The noise variance σ^2 and the kernel parameters are adjusted by optimization of the log-likelihood of the regressors y_n with respect to this parameter by taking derivatives over the parameters and applying a standard gradient ascent.

2. Methodology

The present work considered the implementation of the LAPART algorithm in comparison with regression, gaussian process, and a component-based model. Each of the methods were used to predict the next day's cooling load of the Mechanical Engineering Building (MechBldg) that is located on the campus of the University of New Mexico (UNM). A robust data communication system was established to monitor performance and provide inputs into the existing Energy Management Control System (EMCS). Each of the prediction methods considered the same input features from the National Oceanic and Atmospheric Administration (NOAA) forecast and estimated the next day's load. The experiment was conducted in an on-line learning manner and results from each method were compared based on the linear relationship between actual and predicted, mean square error, mean bias error, and overall energy costs.

2.1. Building

The MechBldg has a total of four floors which add up to about 6,503m^2 of office, classroom, laboratory, and common area space. The building was originally constructed in 1980 as a living laboratory for building energy system research. In the period following commissioning the building performed extremely well, with an energy consumption less than one third of comparable buildings [19]. In the years following its construction, energy prices started to fall, and so did interest in improving energy efficiency [20]. The building has now been revitalized and includes a solar thermal array used for heating and cooling, cold and hot storage tanks, and an upgraded mechanical room that includes multiple air handling units and an absorption chiller.

Multiple experiments have been conducted within the MechBldg testbed that include the integration of DER-CAM [4]. In the present work, DER-CAM has acted as a software-as-a-service provider and was used to optimize chilled water storage and absorption chiller operations [21]. The optmizatino was based on a load prediction for the next day. The prediction was produced by the TRNSYS simulation software described in Section 1.2.1. The present work compared the LAPART capabilities with the utilized TRNSYS results. An accurate prediction of the building load was important because the DER-CAM tool controlled the operations of actual building components.

The DER-CAM tool was integrated with the existing building controls. This meant that it updated the equipment control schedules, based on the optimization results, on a nightly basis. The controlled equipment and component in-

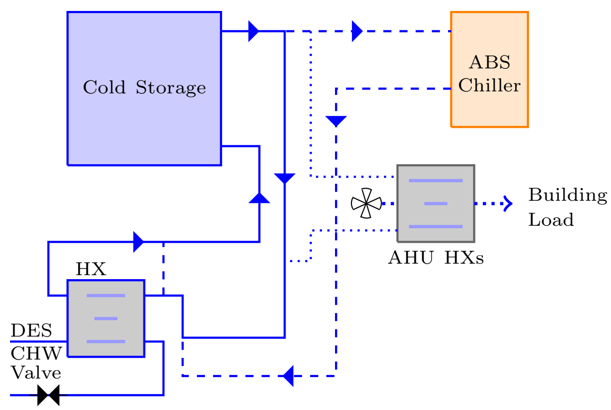

Figure 3: UNM MechBldg chilled water (CHW) storage system that includes storage tanks, absorption chiller (ABS), district energy system (DES) CHW heat exchanger (HX), and air handling units (AHU).The solid lines and arrows connecting the equipment define the operations for charging the cold storage tank. The dashed lines describe the routing of chilled water during the tank discharge conditions.

cluded an absorption (ABS) chiller and chilled water (CHW) flow valve from the district energy system (DES) heat exchanger (HX). These elements were used to charge the cold storage tank as shown in Figure 3. During charging operations the DES provided CHW to the HX that was connected to the cold storage tank. The ABS chiller system, which included a solar thermal array and hot water storage [20], could also provided CHW to the storage tanks as well. When the building required cooling the tank was discharged and CHW was routed to the air handling units (AHUs) HXs. The AHUs then supplied cold air to the building zones to maintain desirable room temperatures.

2.2. Data communications

The data communication process that supported the load prediction, optimization, and building controls platform is shown in Figure 4. It included the interaction of three main components: (A) EMCS for the Mechanical Engineering Building, (B) optimization tool, and the (C) forecast and database system. The EMCS used in this experiment was a proprietary control system called Delta Controls. The optimization system could be any third party or embedded

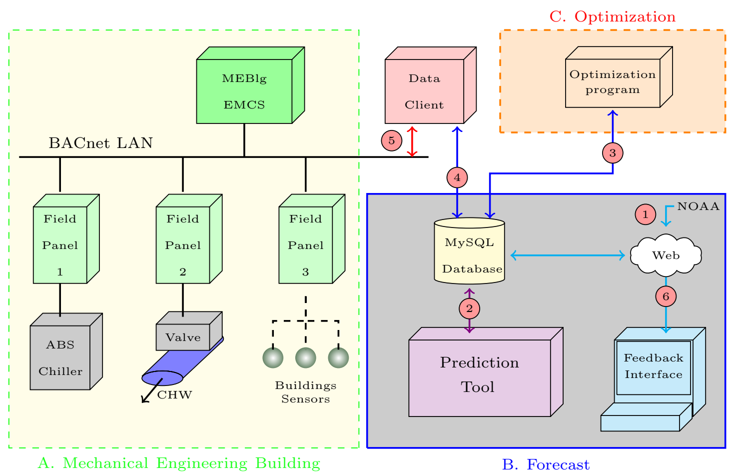

Figure 4: The IT infrastructure for the transfer and storage of forecast, optimization, and actual operations data.

algorithm that has the capability to consider existing tariffs, weather, building load predictions, and equipment specifications to provide an operations schedule for the next day. For example, the Mechanical Engineering Building used DER-CAM as the optimization tool. The third component, which is the focus of this paper, is the forecast system that can predict the next day's thermal load for the building.

The process began at point one with the extraction of weather forecast data from the NOAA webpage. The data was then inserted and stored in a MySQL database. Second, the three prediction tools accessed the data and used it as inputs for each of their simulations. The results from each of the predictions tools were then stored in the MySQL databases. From there, the optimization tool accessed the predictions stored in the database and produced an optimized schedule. In step four, the schedule was sent through a data client connected to the MechBldg EMCS to specific control points in the Delta control system. The optimization platform controlled the DES CHW valve and ABS chiller during both charging and discharging of the cold storage tank operations. Finally, the results from the predictions, optimization, and actual operations were viewed through a web-based interface that plotted time-series operations data.

2.3. Input Features

The input features for each of the prediction tools were forecasted weather, and an estimate of occupancy levels inside the building. The forecast weather data included outside air temperature and solar irradiance. These values were stored in a database during the experiment, and a comparison over an eight day period of the forecast versus the actual observed weather is shown in the top two graphs of Figure 5. The outside air temperature was a direct extraction from NOAA. However, the solar irradiance forecast was based on an equation that used the predicted cloud cover from NOAA, the predicted irradiance value for that time of the year, and parameters derived from the minimization of the

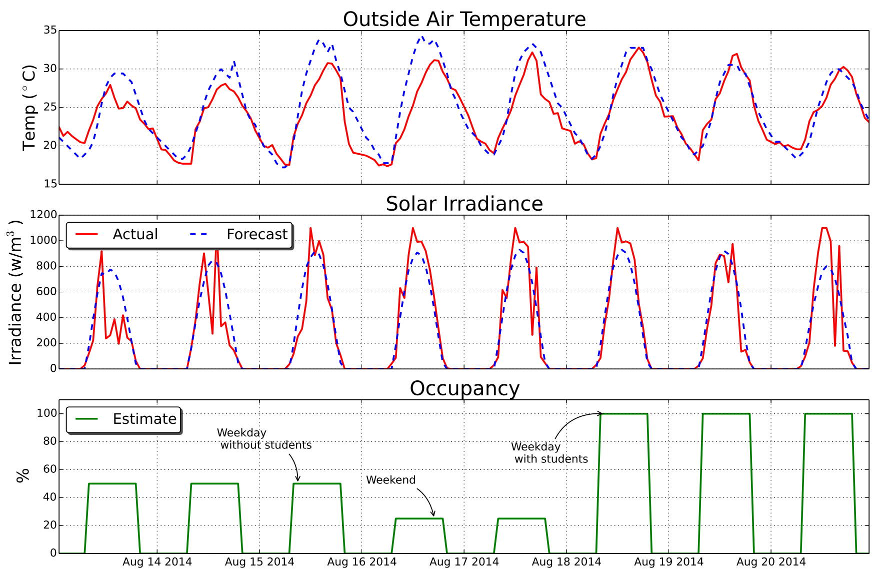

Figure 5: The input features presented to each of the prediction tools were outside air temperature, solar irradiance, and occupancy levels. The weather values were compared with actual observations and display significant correlations. The occupancy levels were simple assumptions based on observations of weekend and weekday activities.

squared error as described by Mammoli *et al.* [4]. In addition, Mammoli *et al.* described a significant correlation between the forecast and actual observations for the temperature and solar irradiance.

The occupancy values were estimated based on a simple approximation and are shown in the bottom graph of Figure 5. The occupancy level was assumed to be at 50% during the weekdays when school was not in season. The weekends the occupancy level was dropped to 25% because many students and faculty were either still working or had left their computers on. Finally, when school was in session the occupancy levels rose to 100%.

2.4. Experiment

The experiment was conducted over a two month period that begin on August 6, 2014 and ended on October 4, 2014. Predictions where made by each of the three methods in a different manner. The TRNSYS model was initial calibrated prior to the August 6th start data and was not updated at any point. The calibration of the TRNSYS model was based on ASHRAE 2001 guidelines whereby the Normalised Mean Bias Error (NMBE) and Coefficient of Variation of the Root Mean Squared Error (CVRMSE) was required to be below 10% and 30%, respectively [22]. The model met NMBE and CVRMSE requirements as described by Jones *et al.* [21]. The data-drive models, on the other hand, were updated using an on-line learning approach. This approach, described in Figure 6 began after the first day of data collection. The data from day 0 was collected and used to train the respective data-driven methods. After training was complete the algorithm predicted day 1, Then, after each subsequent day the new data was added to the bank of training and each algorithm was trained again to predict the next day.

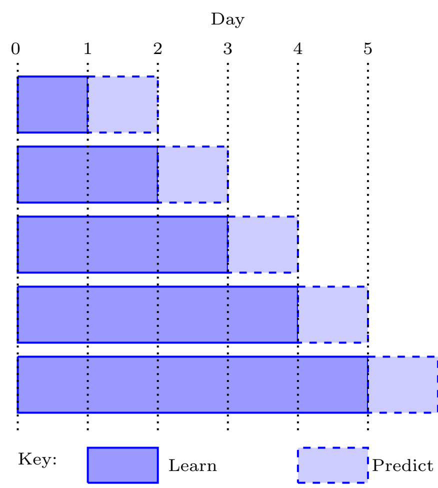

Figure 6: On-line learning began after the first day of data collection. The data from day 0 was collected and used to train the algorithm that then predicted day 1 load. The data used for learning was accumulated and added to the training as time went on to predict the next day's load.

The training process included a two part process for the LAPART algorithm. First, a cross validation was conducted to set the optimal free parameters for the given data set. This was accomplished through a K-Folds cross-validation procedure that defined the free parameters with the lowest error [23]. Second, the complete training data set was presented to the algorithm with the optimal free parameters. This process allowed LAPART to learn the data and develop "memory" that was stored in a database. The stored memory was then used in conjunction with the input features for the next day to produce an estimate of the next day's load.

2.5. Prediction Tool Comparison

The comparison of the prediction tools evaluated in the present work was based on the mean square error, mean bias error, overall energy and associated cost savings, and finally the linear relationship between actual and predicted daily energy values. The mean squared error (MSE) and the mean bias error (MBE) equations were used to evaluate the overall error and predictors tendency to over or under predict. For example, a prediction method that estimates values with a bias of zero would be considered unbiased. The total energy was a sum of the total energy calculated for each day over the life of the experiment. The cost savings considered the efficiency of the chilled water system used to charge the cold storage tank and the time of use rates for the UNM. The linear fit relationship tests to see if the actual and predicted data sets are homogeneous. The analysis considered the R^2 and p-values for a best fit regression line to determine how well the prediction method performed for the given data set..

The overall errors for each of the methods were analyzed using MSE and MBE equations. The MSE measured the average of the squares of the errors as shown in Equation 14.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mathrm{y}}_i - \mathrm{y}_i)^2 \tag{14}$$

The MBE is a measure of overall bias error or systematic error, and is written as a percentage error as shown in Equation 15.

$$\mathrm{MBE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{\mathrm{y}}_i - \mathrm{y}_i) \tag{15}$$

The total energy for each of the prediction methods was computed for each day and over the entire experiment. This energy number for each method was used to evaluate the energy cost impact. The impact considered the cost to charge at night at UNM, which was about \$0.04/kWh, and during the day at a cost of \$0.08/kWh. Finally, the present work computed the coefficient of determination which is often referred to as R^2. This statistical measure has been used to define the dependent variable variance that can be estimated given a set of independent variables.

3. Results

The results for each of the prediction tools was an accumulation of two months of one hour actual and predicted data. The data covered multiple occupancy ranges and weather conditions for the MechBldg. The outside temperatures ranged from 25°C to 30°C and occupancy levels changed from 15% to 100%. Furthermore, the measured load during operating hours ranged from 20kW thermal ($\mathrm{kW}_{thermal}$) to a maximum value of $220\mathrm{kW}_{thermal}$. The four prediction tools were presented with a difficult task of predicting load patterns based on a very limited set of input features. The tested approaches could not replicate load patterns at hourly intervals with a high degree of accuracy (Figure 7) because many of the building dynamics, such as occupancy levels, were not accurately represented in the input features. However, the present work did identified that the LAPART algorithm could provide the most accurate results of the four methods tested. The LAPART algorithm had the smallest MSE,

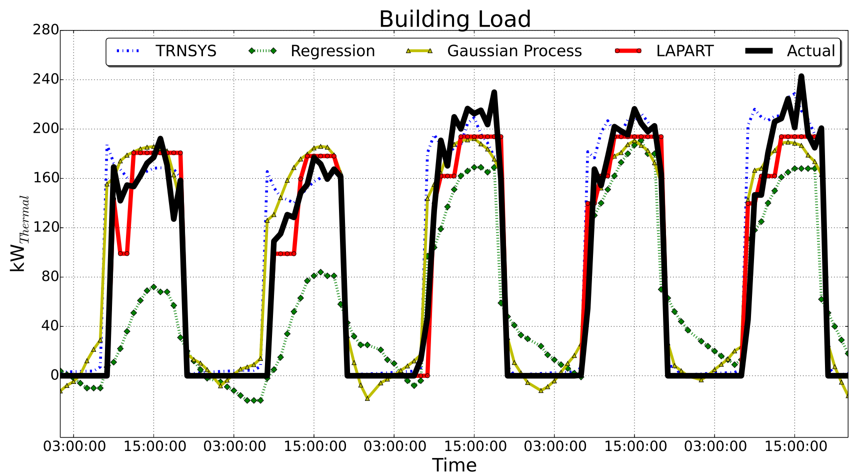

Figure 7: The predicted thermal load produced by the four estimators and the actual load is plotted with respect to time from September 06 to September 10, 2014.

and MBE. It also had an overall energy that matched well with the measured load and provided the highest cost savings for optimizing cold storage charging and discharging schedules.

3.1. Thermal Energy Error

The prediction methods attempted to represent the actual thermal load that was measured in the MechBldg over the two month period. The total

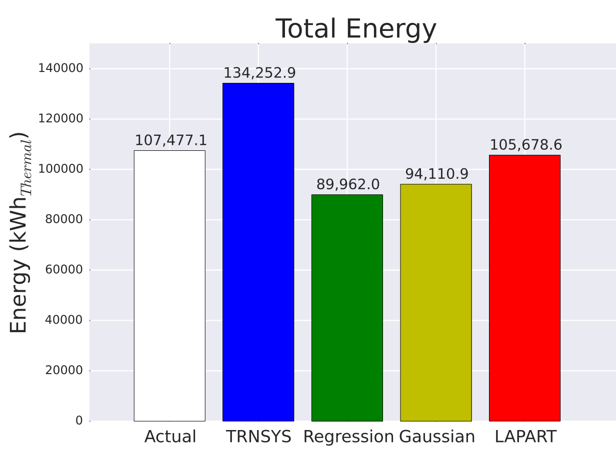

Figure 8: The total thermal energy for the MechBldg was about 107,477kWh$_{thermal}$. The LAPART algorithm was able to provide daily estimates that equated to a 2% difference over the entire two month period of the experiment.

energy over the entire two months of the experiment was measured in the actual building to be 107,477kWh$_{thermal}$ as shown in Figure 8. The TRNSYS model over estimated the actual load by about 24%, which was a difference of around 26,776kWh$_{thermal}$. The regression approach under estimated the overall load by about 117,785kWh$_{thermal}$ or 16% and the Gaussian Process underestimated by about 12%. Whereas, the LAPART algorithm results under predicted by about 1,798 which was only a 2% difference. The error evaluation for the present work also considered the MSE and MBE for each method.

The overall MSE and MBE for daily energy are described in Figures 9 and 10 respectively. The LAPART algorithm produced results that had the lowest MSE that was 1,033.1. The LAPART MSE was about half of the regression MSE

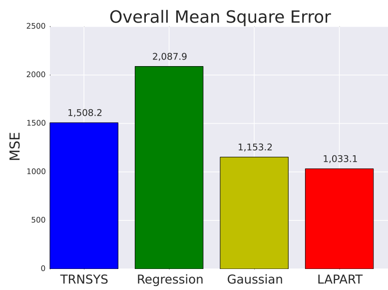

Figure 9: Mean square error for each of the estimation methods showed that the LAPART algorithm produced the smallest overall error.

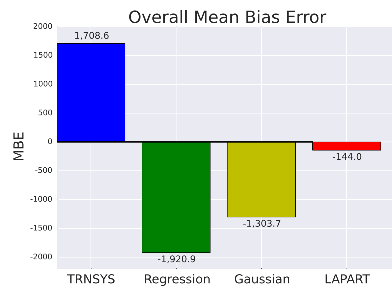

Figure 10: Mean bias error for each of the estimation methods indicates that the LAPART algorithm had a much smaller overall bias.

and about two thirds of the TRNSYS model results. The Gaussian Process algorithm was very close to the LAPART with a MSE error of 1,153. The MBE results described in Figure 10 define the difference between the prediction method results and the desired value. An estimator with a bias equal to zero is considered unbiased. The LAPART results had a value of -144 which was much closer to zero than the other two methods. The regression and Gaussian Process methods had a MBE of -1,920 and -1,303 respectively, which indicated that their predictions tended to be significantly lower than the actual. The TRNSYS model results had a MBE of 1,708 which showed that it tended to over predict.

The four methods abilities to estimate the daily energy required by the MechBldg was evaluated by defining the linear relationship between the prediction and the actual value as shown in Figures 11, 12, 13, and 14. None of the prediction methods had a desired R^2 value very close to one. They also did not have an intercept and slope of the best fit line close to zero and one respectively. Instead, the TRNSYS results produced the best R^2 value equal to 0.77 followed by regression, LAPART, and then the GP. The results produced by the LAPART algorithm had a R^2 value equal to 0.65 and the GP had a very low R^2 value equal to 0.19.

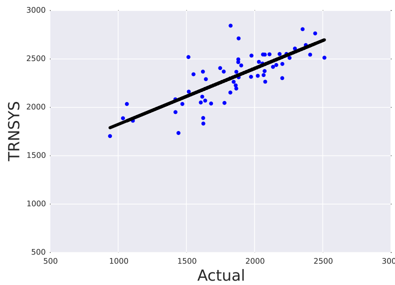

Figure 11: The TRNSYS predictions had an R^2 value equal to 0.77. The best fit line had an intercept and slope equal to 1246 and 0.57.

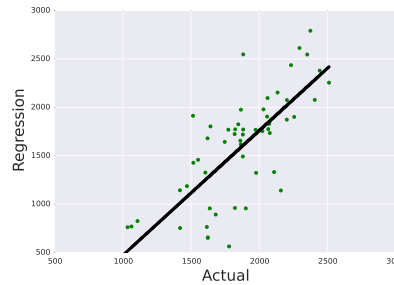

Figure 12: The regression predictions had an R^2 value equal to 0.75. The best fit line had an intercept and slope equal to -818 and 1.2.

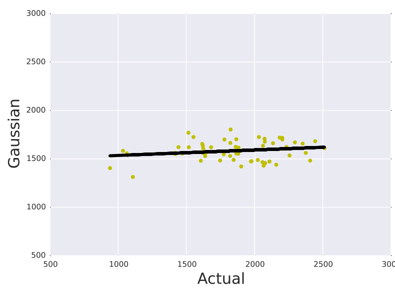

Figure 13: The GP predictions had an R^2 value equal to 0.19. The best fit line had an intercept and slope equal to 1479 and 0.05.

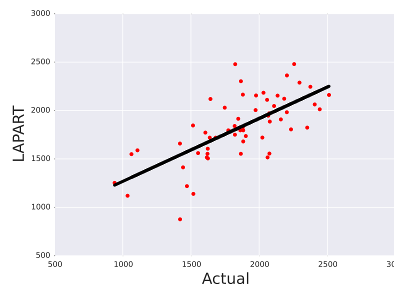

Figure 14: The LAPART predictions had an R^2 value equal to 0.65. The best fit line had an intercept and slope equal to 617and 0.65.

3.2. Cost Savings

The intent of the thermal storage optimization, discussed in the present work, was to reduce the overall energy consumption and cost used to run a chiller. The chiller was used to charge the storage tanks in the MechBldg at night and supply CHW during the day if necessary. A complete charge of the storage tanks each night required about 3,000kWh$_{thermal}$ of energy. The actual measured energy required by the building fluctuated between 1,000kWh$_{thermal}$ and 2,500kWh$_{thermal}$. The regression, TRNSYS, GP, and LAPART predictions attempted to predict the overall energy required to charge the thermal tanks. However, because the estimates of thermal energy where not perfect there were missed opportunity cost associated with each prediction method as well as the baseline scenario where the tank was charged completely.

The baseline cost for the entire experiment was calculated to be \$1,302. This overall cost was based on the scenario where the tank was charged to its capacity each night. If the tank was charged only to the amount required for the next day then the total cost would be \$807 for a savings of \$495 over the two month period observed in the present work. However, the prediction methods were not perfect and TRNSYS provided a savings of $291, regression was $308, GP was $339, and the LAPART performed the best with a savings of $354 as shown in Figure 15.

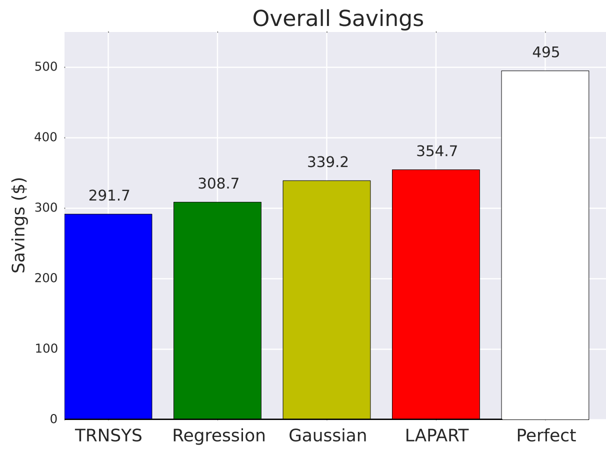

Figure 15: The LAPART method had the best overall savings for the experiment followed by Gaussian Process, regression and then TRNSYS.

The implementation of each of these methods does have an associated cost. The component-based model is the most time consuming of the three and would require considerable engineering effort to learn system behaviors. Costs may vary but could easily reach $12,000 to build a detailed energy model of a building comparable to the MechBldg. The learning based tools used in the present work, LAPART, GP, and regression, require considerably less time to develop. The development time would be the integration of actual building data with the algorithms. A simple estimation for programming time would be about $1,800 for regression and $2,800 for LAPART. The increase in cost for LAPART would be due to the set-up of a database for more complex outputs. The Return on Investment (ROI) for the component-based model compared to the learning based approach was considerably different. The present work showed that the TRNSYS model would have a simple payback of about 13.5 years if it performed at the given rate for the rest of the cooling season. The regression and LAPART approaches had a simple payback of about 1.9 and 2.6 years respectively.

4. Conclusions

The effective integration of thermal storage into commercial building applications requires accurate load predictions. Component-based and data-driven models have been used to estimate future loads. This paper introduced the LAPART ANN for predicting the building thermal loads and tested on the MechBldg that is 6,503m^2 of office, classroom, laboratory, and common area building that has CHW storage. The LAPART algorithm was compared to a component-based model, Gaussian Process, and least squares regression. LAPART results outperformed the other approaches in terms of overall error, MSE, and MBE. However, the LAPART did not have the best linear relationship with the actual data. Yet, the LAPART algorithm does provide the best option for cost savings and overall ROI.

References

[1] C. B. Jones, Fault detection and diagnostics of an HVAC sub-system using adaptive resonance theory neural networks, Ph.D. thesis, University of New Mexico, Albuquerque, NM (Jun. 2015).
URL `http://repository.unm.edu/handle/1928/27794`

[2] U. S. Department Energy, Commercial Buildings Energy Consumption Survey (CBECS), EIA, 2012.
URL `http://www.eia.gov/consumption/commercial/`

[3] G. P. Henze, R. H. Dodier, M. Krarti, Development of a Predictive Optimal Controller for Thermal Energy Storage Systems, HVAC&R Research 3 (3) (1997) 233–264.

[4] A. Mammoli, M. Stadler, N. DeFrost, H. Barsun, R. Burnett, C. Marnay, Software-as-a-Service Optimised Scheduling of a Solar-Assisted HVAC System with Thermal Storage, International Conference on Microgeneration and Related Technologies 19 (112) (2013) 888–896.

[5] K. M. Maribu, R. M. Firestone, C. Marnay, A. S. Siddiqui, Distributed energy resources market diffusion model, Energy Policy 35 (9) (2007) 4471 – 4484.

[6] K. Roth, R. Zogg, J. Brodrick, Cool Thermal Energy Storage, ASHRAE Journal (2006) 94–96.

[7] L. Pedersen, J. Stang, R. Ulseth, Load prediction method for heat and electricty demand in buildings for the purpose of planning for mixed energy distribution systems, Energy and Buildings 40 (2008) 1124–1134.

[8] C. Scheuer, G. A. Keoleian, P. Reppe, Life cycle energy and environmental performance of a new university building: modeling challenges and design implications, Energy and Buildings 35 (10) (2003) 1049 – 1064.

[9] C. Marnay, G. Venkataramanan, M. Stadler, A. Siddiqui, R. Firestone, B. Chandran, Optimal Technology Selection and Operation of Commercial-Building Microgrids, Vol. 23, IEEE Transactions on Power Systems, 2008.

[10] Z. Li, G. Huang, Re-evaluation of building cooling load prediction models for use in humid subtropical area, Energy and Buildings 62 (2013) 442–449.

[11] H.-x. Zhao, F. Magouls, A review on the prediction of building energy consumption, Renewable and Sustainable Energy Reviews 16 (6) (2012) 3586–3592.

[12] S. A. Kalogirou, Applications of artificial neural-networks for energy systems, Applied Energy 67 (12) (2000) 17–35.

[13] A. E. Ben-Nakhi, M. A. Mahmoud, Cooling load prediction for buildings using general regression neural networks, Energy Conversion and Management 45 (1314) (2004) 2127–2141.

[14] M. Healy, T. Caudell, S. Smith, A neural architecture for pattern sequence verification through inferencing, IEEE Transactions on Neural Networks 4 (1) (1993) 9–20.

[15] M. Healy, T. Caudell, Acquiring rule sets as a product of learning in a logical neural architecture, IEEE Transactions on Neural Networks 8 (3) (1997) 461–474.

[16] A. Mammoli, A. Menicucci, T. Caudell, A. Ellis, S. Willard, J. Simmins, Low-cost solar micro-forecasts for PV smoothing, in: 2013 1st IEEE Conference on Technologies for Sustainability (SusTech), 2013, pp. 238–243.

[17] S. Haykin, Neural Networks A Comprehensive Foundation, 2nd Edition, Prentic-Hall, Inc, 1999.

[18] G. A. Carpenter, S. Grossberg, D. B. Rosen, Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system, Neural Networks 4 (6) (1991) 759–771.

[19] M. Wildin, Results from use of thermally stratified water tanks to heat and cool the Mechanical Engineering Building at the University of New Mexico, DOE Contract W-7405-ENG-26, 1983.

[20] A. Mammoli, P. Vorobieff, H. Barsun, R. Burnett, D. Fisher, Energetic, economic and environmental performance of a solar-thermal-assisted HVAC system, Energy and Buildings 42 (9) (2010) 1524–1535.

[21] C. B. Jones, C. Marnay, M. Robinson, H. Barsun, L. Ghanbari, S. Mashayekh, W. Feng, A. Mammoli, Software-as-a-service optimal scheduling of New Mexico buildings, in: ECEE, Presquile de Giens, France, 2015.

[22] ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings (2002).

[23] K. P. Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.

[24] PNM, PNM Electricty Rates and Riders (2015).
URL `https://www.pnm.com/rates`