Title

CoolWave: Towards the generalizable system identification with feature separation for HVAC Systems

Author

Kashyap, Naveen

Publication Date

2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

CoolWave: Towards the generalizable system identification with feature separation for HVAC Systems

A thesis submitted in partial satisfaction of the
requirements for the degree
Master of Science

in

Computer Science

by

Naveen Kashyap

Committee in charge:

Professor Rajesh Gupta, Chair
Professor Jan Kleissl
Professor Julian McAuley

2020

The thesis of Naveen Kashyap is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This thesis is coauthored with Koh, Jason and Hong, Dezhi and Gupta, Rajesh K. The thesis author was the primary author of this chapter.

ABSTRACT OF THE THESIS

CoolWave: Towards the generalizable system identification with feature separation for HVAC Systems

by

Naveen Kashyap

Master of Science in Computer Science

University of California San Diego, 2020

Professor Rajesh Gupta, Chair

Heating, ventilating, and air conditioning (HVAC) systems are widely used and can constitute up to 48% of a commercial building's energy consumption. Computational methods of optimizing the energy consumption of these systems require a model of the system to be identified. Modeling the complexities of HVAC systems is well-suited for machine learning algorithms and has been the focus of an increasing amount of research on this topic. Existing works have approached modeling HVAC systems using data driven methods, physics based methods, or a mixture of the two. In this thesis, we propose a data driven modeling method that generalizes well to data outside the bounds of the training data by modeling the ambient temperature of

the building and the influence of the HVAC system separately. Such a model can simulate an HVAC system and can be used as part of larger generalization frameworks, such as reinforcement learning, to optimize the HVAC system. We evaluate our model by its ability to accurately predict the temperature of a zone given a window of environmental data and find that CoolWave produces a strong improvement in prediction quality and generalizability when compared the performance to an implementation of Wavenet, an Auto-Regressor, and a Gaussian Process Regressor.

1 Introduction

Heating, ventilating, and air conditioning (HVAC) systems are widely used and can constitute up to 48% of a commercial building's energy consumption [1]. Optimizing for energy consumption is an involved process that requires adjusting an HVAC system's parameters to drive down energy consumption while maintaining the building's environmental requirements. Given the complexity and scale of HVAC systems, optimizing a system by hand is tedious and ineffective. Some techniques have been developed to optimize HVAC systems by observing and modeling the energy consumption of the individual components of an HVAC system and then algorithmically deriving the appropriate HVAC system parameters using these models [12] [13]. However, these techniques require extensive and rich energy consumption data for every component of the HVAC system which may not be widely available, so effort has been made to develop a technique to model the entire HVAC system as it relates to the building's internal environment. Such a model serves as an HVAC system simulator and allows optimization techniques that require the use and exploration of the HVAC system, such as reinforcement learning [17], to be used without the expensive and uncomfortable process of experimenting with the actual HVAC system [6].

However, creating this model is costly and non-trivial. HVAC systems have many complex subsystems that work together to regulate a building's environment, so the non-linear environmental effect of adjusting one parameter may be obscured by the reaction of the other subsystems, each with their own set of changing inputs. Furthermore, the sheer number of system parameters increases the complexity of modeling its parametric behavior.

The problem of modeling the complexities of HVAC systems is well-suited for machine learning algorithms and has been the focus of an increasing amount of research on this topic [2]. However, the main struggle with this approach is the inability for a machine learning model to generalize beyond the bounds of the training data [20]. HVAC system data tends to have low variance by nature; the purpose of the system is to maintain a steady environment which yields repetitive, statistically similar data [3]. Thus, a machine learning algorithm applied to HVAC system data will accurately model the behavior of the HVAC system under typical conditions but will struggle to accurately model the behavior otherwise. Developing a method to improve the generalizability of such a model is the focus of this thesis.

Existing works have approached modeling HVAC systems using data-driven methods, physics-based methods, or a mixture of the two. Data driven methods aim to model HVAC system parameters using only historical sensor data and therefore can be used without detailed knowledge on the building's physical layout, but struggle to capture the system operating outside of the bounds of the training data. Physics based methods are derived from the laws of physics and use complex physics engines to model HVAC system parameters but require detailed building schematics and do not account for varying heat loads, such as room occupancy or sunlight. Hybrid modeling methods use physics-based methods to build a rough model of the HVAC system parameters and then use data-driven methods to fine tune the model. However, this approach still relies upon detailed knowledge of the building's use and physical layout.

In this thesis, we propose a data-driven modeling method, named CoolWave, that shows a strong improvement in prediction quality and generalizability when compared to competing modeling methods. We do this by modeling the ambient temperature of the building and the influence of the HVAC system separately. Intuitively, the temperature inside a building is influenced by two main factors: the ambient temperatures of surrounding masses and the HVAC system. For example, if a room is left unattended and the HVAC system is off for an extended period of time, the temperature in the room is influenced only by the temperatures of the surrounding areas, for example the outside air temperature or the sun. However, if the HVAC system is on, the temperature inside the room is influenced by both the ambient temperatures and the HVAC system. CoolWave learns the building's indoor temperature absent of any HVAC system influence by observing sensor data during times of HVAC inactivity. Based on this, CoolWave predicts the temperature of a room during times of HVAC activity and learns the influence of the HVAC system by observing the difference between its prediction and the true temperature.

Our model is trained and tested using building and system data gathered from the Computer Science and Engineering building on the campus of University of California, San Diego. In Section 2 we will discuss the background of HVAC system identification and other works aimed at doing this. In Section 3 we will outline our approach to a solution and then detail our methodologies. In Section 4 we evaluate our model in two ways: a quantitative measure of prediction accuracy and a qualitative measure of prediction accuracy on data both within and outside of the bounds of the training data. Finally, we speculate potential future works in Section 5.

2 Background and Related Works

2.1 Building HVAC Systems

Heating, ventilating, and air conditioning (HVAC) systems are widely used, accounting for up to 48% of all energy consumed by commercial buildings [1]. These systems are responsible for maintaining a specific environment inside the building regardless of the environment outside the building. They are complex but work on one basic principle: when appropriate, collect outside air, heat or cool it, and then force it inside. At a high level, HVAC systems draw outside air into the building via ductwork. This air is then cooled in an Air Handling Unit (AHU) and then supplied to a Variable Air Volume (VAV) unit that can heat the incoming supply air if needed. The VAV adjusts the Supply Air Flow Set Point (SAFSP) which is the measure of how much cooled air is supplied from the VAV and into the area being serviced, called a "zone". By modulating the SAFSP, the VAV is able to adjust the zone's temperature (ZNT) [11]. Figure 2.1 shows a simplified diagram of how a VAV services a zone. The Supply Air is provided to the VAV by the AHU. Within the VAV, a Reheating Coil carrying hot water can reheat the Supply Air if needed before injecting the air into the zone.

Such an HVAC system typically works on a simple control policy: given a range of acceptable temperatures, set SAFSP to zero when the zone's temperature is within the range, and set the SAFSP to a non-zero value when the zone's temperature is outside the range. For example, if the temperature in the zone rises above the upper threshold of the acceptable temperature range,

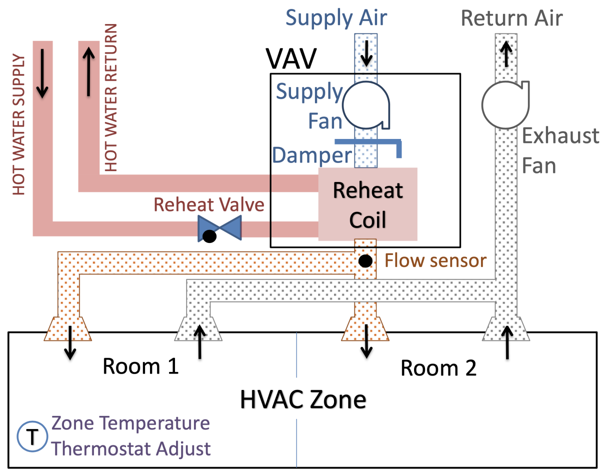

Figure 2.1: A simplified diagram of an VAV adapted from Quiver [11]. Given supply air from the AHU, the VAV is able to use coils of hot water to heat the supply air if needed. By adjusting the SAFSP, the VAV controls the ZNT of the zone being serviced.

the VAV will increase the SAFSP to inject cool air into the zone until the temperature in that zone drops below the upper threshold. Additionally, if the temperature continues to rise even after setting the SAFSP to a non-zero value, continue to increase the SAFSP until the temperature in that zone begins to drop. However, a policy of employing the HVAC system only when the temperature has already left the range of acceptable temperatures may be energy-inefficient and thus motivates policy optimization.

2.2 Existing HVAC Models

To optimize a building's HVAC system, the system itself must be modeled such that its parameters can be appropriately tuned to fit the building's requirements while reducing energy consumption. A variety of computational methods have been developed to model the complexities of HVAC systems. There are three main types of modeling methods for HVAC systems: the first type is data-driven models usually requiring some form of artificial intelligence tasked with learning the system parameters, the second is physics-based models usually requiring a detailed physical model of the building and an understanding of the underlying properties of heat transfer to calculate the system parameters, and the third is a mixture of the two wherein a physics-based model provides a loose structure of the model and the data-driven model learns to fine-tune the parameters [2].

2.2.1 Data-Driven Models

Current data-driven models offer the advantage of modeling the HVAC system without extensive knowledge on the building and system specifications. One such method of HVAC system identification consists of the formulation of a mathematical model of the system dynamics from measured data [22]. Given real-time sensor data, an appropriately chosen recursive estimation algorithm is used to identify system parameters online. This recursive approach allows the parameters of the approach to be continuously monitored and updated, but is unable to accurately generalize to unseen data. Some methods propose using a differentiable Model Predictive Control (MPC) to encode domain knowledge on system dynamics [6]. Typical MPC methods are planning-based and solve an optimal control problem iteratively over a receding time horizon, but requires a model of the building. To remove the dependence upon a model of the building, research has been conducted to differentiate through the MPC by using first derivative tests of the convex approximation at a fixed point of the controller of the system [4]. Using this strategy, the model learns the cost and dynamics of a controller via end-to-end learning. However, this method presents a set of problems that can lead to the inability to find such a fixed point which prevents the MPC from being differentiated, regressing the method back to having to use a model of the building. By contrast, another method takes a more straight-forward approach of using neural networks to model the non-linearities typically characteristic of dynamic systems [9]. It shows that neural networks can be used for predicting the indoor temperature of building based only on a small numbers of input variables. However, this technique does not address the wide range of training cases or the inability to generalize beyond the training data. CoolWave uses the advantages that come with neural networks but generalizes better by using a combination of two neural networks that independently model the factors that influence the temperature in a room. We will detail our approach in Section 3.

2.2.2 Physics-Based Models

Physics based models are also known as analytical first-principal models, forward models, or white box models and are more generalizable than data-driven models, but require detailed understanding of the system. This prohibits modeling of any changes to the system. Most famously is EnergyPlus, a modular physics-based model that uses physics engines and building system simulation modules to predict the temperature at the next time step [8]. This requires a detailed model of the physical properties of the building, environment, and HVAC system itself and cannot adapt to variations in the heat loads in the zone, such as an open window or typical human traffic. Furthermore, an EnergyPlus input file, while readable, is cryptic and definitely not user-friendly – it is not intended to be the main interface for typical end-users making it difficult for the system to adapt to changes in the building, environment, or HVAC system. Another such method for modeling HVAC systems is a numerical method that generates equations inspired by Newton's law of cooling without any consideration for the effects of outside heating loads [5]. The paper uses semiparametric regression [18] to identify a linear relationship between each AC unit and the temperature in the room, leaving the non-parametric heat load to be estimated as a function of time. However, this is estimation makes the major assumption that the heating loads themselves are not parametric and can be estimated simply as a function of time. Another class of modeling methods use physical equations to mathematically describe the heating and cooling loads of each component of the HVAC system, such as the reheating and cooling coil, the fans and pumps, and the damper [2]. These complex and non-trivial time-domain differential equations accurately describe each component of the HVAC system but are specific to a building's HVAC system and require extensive understanding of the physics of each HVAC component, rendering them impractical at a large scale [2].

2.2.3 Hybrid Models

Hybrid models couple the generalizability of physics-based models and the accuracy of data-driven models. Also called Grey Box models, hybrid models provide physical meaning of the HVAC system and can capture the effects of any unmodeled dynamics of the system that were left out of the physics equations. Still, knowledge of both the underlying physical principals of the HVAC system is required [2]. One such example of a hybrid model employs a numerical model to estimate indoor temperature, and then passes the output to an ANN to calibrate the model [16]. The benefit of this method is that the calibration requires only the measurement room temperatures and does not need to know the details of the building systems, thus significantly reducing the effort needed to model the building. Another method takes an interesting approach and represents the model of a building as a resistance-capacitance (RC) model, based off of the RC models used in electrical circuit calculations, and then uses a genetic algorithm to find the model's parameters, such as the thermal resistances and capacitances [15]. Other works enhance data-driven models with physics-based models, using differentiable physical equations to express the steady-state of the energy balance of a room and then uses the well-known ARMAX method to identify system parameters using architectural parameters, such as room volume, wall and window thickness, window-to-wall ratio, and heat conduction and convection constants [24]. However, the nature of hybrid models requires an underlying understanding of the physical phenomenon of the inputs and outputs of the HVAC system, which themselves may be extremely complex. Furthermore, if a hybrid model is found, the parameters often need retuning as operating conditions drift from the initial conditions making hybrid models the hardest to implement [2].

2.2.4 Problem Formulation

Overall, an appropriate model of an HVAC system will take as input a history of each zone's temperature as well as a history of all factors which might influence the temperature of a zone, including the HVAC system itself. As output, the model will predict the temperature of the zone at the next timestep and the model will use SAFSP as the only control variable while the input history of outside air temperature and intensity of the sun will follow the natural behavior. Most importantly, the model will generalize beyond the bounds of the training data, meaning that it should be accurate when modeling the behavior of the system when given inputs outside of the range seen in the training data.

In this thesis, we propose a data-driven model that generalizes beyond the training data while maintaining comparable accuracy within the training data range.

3 CoolWave

CoolWave is a data-driven model that uses the history of the Supply AirFlow Set Point, outside air temperature, position of the sun, and the zone temperature ending at time t to predict the zone temperature at time $t+1$.

The intuition behind CoolWave's design is derived from the domain knowledge that a zone's temperature is a composition of the heat produced by the outside air temperature and the intensity of the sun, and the cool air injected by the HVAC system. Modeling the influences of the outside air temperature and the intensity of the sun separately from the influence of the HVAC system allows a comprehensive model of the zone under all conditions to be composed. Modeling each influence requires detecting the non-linear complexities of the HVAC system with little statistical training, which is one of the advantages offered by neural networks [21]. Thus, CoolWave involves two separate neural networks, one learning from the other.

Both neural networks within CoolWave are a standard implementation of Wavenet [23]. Wavenet is a popular convolutional neural network (CNN) especially effective at modeling high-frequency sequential data. A typical CNN takes advantage of the hierarchical nature in data to construct a model of complex patterns from simpler, smaller patterns. Each convolutional layer trains an $m \times m$ matrix of weights, called a "filter", which is then convolved across a $d \times d$ input matrix, where $d \geq m$. The result is an abstraction of the input matrix containing the patterns identified by the filter. The power of Wavenet lies in the stacks of dilated convolutions that allow the model to access a wide range of history with high efficiency. A dilated convolution is a

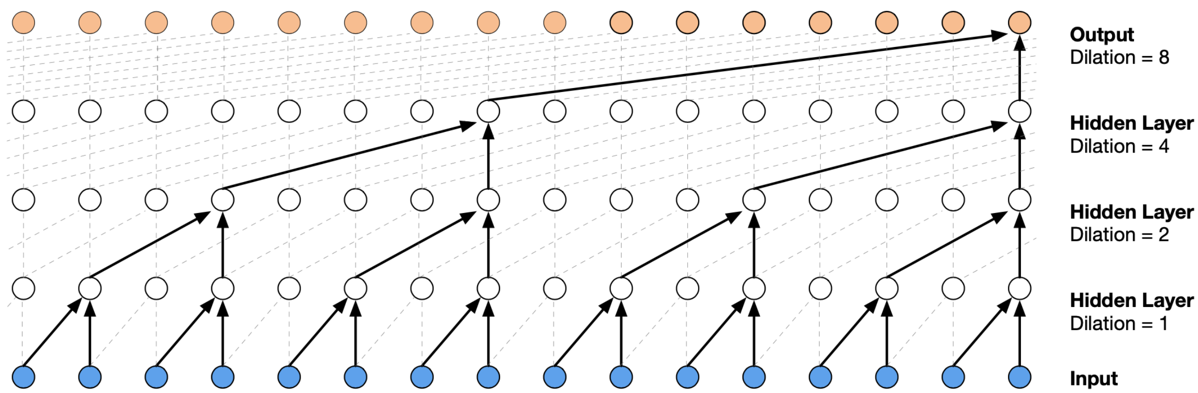

Figure 3.1: A diagram of a stack of 1-dimensional dilated convolutions adapted from the original Wavenet paper [23]. The first hidden layer has a dilation rate of 1, skipping every other node. The hidden layers after that all have a dilation rate twice that of the layer before it.

convolution where the filter is applied over an area larger than its length by skipping input values with a certain step, allowing the filter to be applied on a courser scale. As a result, Wavenet is faster to train than a traditional convolutional neural network and yields competitive results compared to a recursive neural network, a natural alternative for modeling timeseries data. Figure 3.1 shows how wavenet is constructed, using a stack of dilated convolutional layers to abstract a complex pattern at the output from the simpler patterns at the input.

In CoolWave, the first network models the temperature of a room absent of any influence by the HVAC system. This model is appropriately named the Heating Model. The second network models the temperature of the room given the influence of the HVAC system. This model is appropriately named the Cooling Model. Figure 3.2 shows a basic diagram of how the Heating and Cooling models interact. When the HVAC system is off, the Heating Model can observe and learn the influence of the outside air temperature and the sun on the temperature of a zone. Using the Heating Model to predict the temperature in a zone with the HVAC system off, the Cooling Model can observe and learn the influence of the HVAC system on a zone's temperature by observing the relationship between the SAFSP and the difference between the Heating Model's prediction and the actual temperature in the zone. Then, combining the Heating Model's prediction with the Cooling Model's prediction yields the composed temperature at the

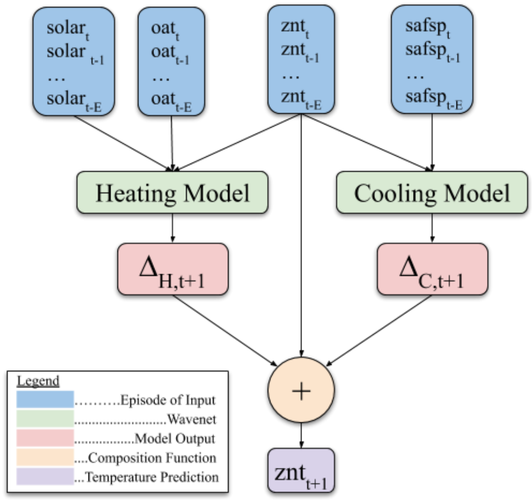

Figure 3.2: A diagram of CoolWave's two internal models. The Heating Model uses a window of *solar*, *oat*, and *znt* to predict $\Delta_{H,t+1}$. The Cooling Model uses a window of *safsp* and *znt* to predict $\Delta_{C,t+1}$. The final prediction is composed as: $znt_{t+1} = znt_t + \Delta_{H,t+1} + \Delta_{C,t+1}$.

next timestep.

We define a *window size*, w, to be the amount of history of data, in seconds, that the Heating and Cooling Models are given as input. For example, an window size of 6 hours ($w = 21600$) means that the models are given data starting at time $t - w$ and ending at time t. Larger values of w mean that the models have more historical context to make their predictions whereas smaller values of w mean that the models have less historical context to make their predictions.

3.1 Heating Model

The temperate climate in San Diego means that the only factors that heat a zone's temperature are the outside air temperature and the intensity of the sun, so the Heating Model observes and learns a zone's temperature as it relates to these two features only when the HVAC system is off at all times of the window. In our building, this mainly occurs outside of typical business hours and on weekends.

The Heating Model takes as input an window of the outside air temperature, the intensity of the sun, and the zone temperature ending at time t to predict the change in temperature between time t and $t+1$, $\Delta_{H,t+1}$. A brief description of the features are provided in Table 3.1.

Table 3.1: Features used by the Heating Model

Feature Name	Description
Outside Air Temperature (OAT)	Temperature of the air flowing in from outside
Solar	Intensity of the sun on the roof of the building
Zone Temperature (ZNT)	Temperature of the zone

3.2 Cooling Model

The only factor that cools a zone's temperature is the HVAC system. To model this, the Heating Model must first provide $\Delta_{H,t+1}$ for all times regardless of whether or not the HVAC system is on. Then the Cooling Model must observe and learn the relationship between SAFSP and the difference between $\Delta_{H,t+1}$ and the actual change in temperature, $\Delta_{actual,t+1}$, only at times when SAFSP is non-zero. This difference in change of temperature is denoted as $\Delta_{C,t+1}$ and is calculated as such:

$$\Delta_{C,t+1} = \Delta_{actual,t+1} - \Delta_{H,t+1}$$

where

$$\Delta_{actual,t+1} = znt_{t+1} - znt_t$$

In our building, this mainly occurs during typical business hours on weekdays.

The Cooling Model takes as input an window of the SAFSP and zone temperature ending at time t to predict the change in temperature due to the HVAC system. A brief description of the features are provided in Table 3.2.

Table 3.2: Features used by the Cooling Model

Feature Name	Description
Supply Air Flow Set Point (SAFSP)	A measure of how much air is coming out of the vents
Zone Temperature (ZNT)	Temperature of the room

3.3 Combining the Heating and Cooling Models

Following the intuition that the temperature in a zone is the composition of the influence of the sun and the outside air temperature, which both serve as the sole factors in heating the zone, and the influence of the HVAC system, which serves as the sole factor in cooling the zone, the temperature in a zone at time $t+1$ can be calculated simply as such:

$$znt_{t+1} = znt_t + \Delta_{H,t+1} + \Delta_{C,t+1}$$

3.4 Training Method

The methods of training the Heating Model and the Cooling Model are similar. Summarily, each model takes an window of data as input and is trained on the appropriate change in temperature.

3.4.1 Heating Model

Training the Heating Model is straight-forward. As input, the model takes in a $w \times f$ window of data, where w is the number of rows within the window and f is the number input features. Specifically, $f = 3$ because the model takes solar, oat, and znt as input. The target for each training data point is the change in temperature from time t to *t+1*. This training signal is calculated as such:

$$\Delta_{actual,t+1} = znt_{t+1} - znt_t$$

The Heating Model is only trained on data samples for which the HVAC system was off (SAFSP = 0) for the entirety of the window. Thus, $\Delta_{actual,t+1}$ represents the influence of the sun and the outside air temperature on the temperature of the zone.

3.4.2 Cooling Model

Training the Cooling Model requires the Heating Model's predictions, $\Delta_{H,t+1}$ for all time t. These predictions represent the change in temperature in a room with the HVAC system off. As input, the Cooling Model takes in a $w \times f$ window of data, in this case $f = 2$ because the model only takes SAFSP and znt as input. The target of the model for each training data point is the difference between the change in temperature predicted by the Heating Model (representing the change in temperature of the room if the HVAC system is off) and the actual change in temperature of the zone. It is calculated as such:

$$\Delta_{C,t+1} = \Delta_{actual,t+1} - \Delta_{H,t+1}$$

Because this model is only trained on data samples for which the HVAC system was on (SAFSP $\neq 0$) at any time during the window, $\Delta_{C,t+1}$ represents the change in temperature of the zone due to the HVAC system. For example, if the Heating Model predicts that the temperature in a zone will increase by 1 °F, and the actual temperature decreases by 2 °F, then the HVAC system is effectively responsible for a 3 °F decrease in the zone's temperature.

3.5 Prediction Method

The nature of modeling any dynamical system is to use online predictions; the prediction at time *t* should be included in the input data for the prediction at time *t+1*. First, an window of data ending at time t is given to the Heating Model which then predicts $\Delta_{H,t+1}$. Similarly, an window of data also ending at time t is given to the Cooling Model which then predicts $\Delta_{C,t+1}$. The zone temperature at time t+1 is then calculated as such:

$$znt_{t+1} = znt_t + \Delta_{H,t+1} + \Delta_{C,t+1}$$

Importantly, znt_{t+1} is then included in the window of input data ending at time t+1. By doing so, CoolWave uses its previous predictions to inform its next prediction allowing it to simulate behavior unseen in the training data.

4 Evaluation

4.1 Experimental Setup

4.1.1 Data Collection

The data for all input features were collected in the Computer Science and Engineering (CSE) building at the University of California, San Diego. Over the course of the 2018 calendar year, temperature and HVAC system component values were regularly recorded at 5 minute intervals. All data was recorded using sensor data in the zone and from within the HVAC system itself. The outside air temperature and the intensity of the sun data were collected using sensors placed on the roof of the building. The training set was constructed using data starting at midnight on April 2nd, 2018 and ending at 11:55pm on May 31st, 2018. The validation set was constructed using data starting at midnight on June 1st, 2018 and ending at 11:55pm on June 30th, 2018. The training set start and end dates were chosen because weather patterns are most consistent during April and May. By contrast, the validation set start and end dates were chosen because the weather patterns in June are distinctly different than in April and May, meaning that the validation set includes data outside the bounds of the data in the training set.

We used three separate rooms as our test bed, training and testing separately for each. Figure 4.1 shows the organization of the rooms we used. All three rooms are in the southwestern wing. The first room, CSE 2108, is a northwestward facing room with a large window facing in the northwestern direction. The second room, CSE 2136, is a westward facing room with a small

Figure 4.1: A floor plan of the second floor of the CSE building which contains the rooms: 2108, 2136, 2150 in the southwestern wing of the building.

window facing in the western direction. The third room, CSE 2150, is a large southeastward facing room with multiple windows facing in the southeastern direction. We discuss the results for each room in Section 4.2.2.

4.1.2 Metrics

The results of this model can be quantitatively assessed by looking at its accuracy, measured as the root mean squared error (RMSE), in predicting the temperature of a zone within the bounds of the training data as well as outside the bounds of the training data. Additionally, the RMSE over weekends, when the HVAC system is mainly off, is a measure of how well the model has learned that the influence of the HVAC system is zero when the SAFSP is zero. However, a model which only predicts the average of all temperature might yield a low RMSE so other qualitative metrics are also discussed to measure model performance: znt responsiveness to changes in SAFSP indicates that the model has learned an appropriate relationship between the znt and the SAFSP, and the general ability for znt predictions to follow the *trends* of the actual znt indicates an understanding of the overall physics of heat transfer within the zone.

4.1.3 Model Hyperparameters

CoolWave has three hyperparameters: the window size, which is the measure of how much historical context the model receives as input, the kernel regularization, which is used to help prevent the model from overfitting to the training data, and the learning rate, which controls the speed of gradient descent on the error function. For our building, we found that an window size of 6 hours ($w = 21600$) and a kernel regularization value of *0.001* yields the best results. Kernel regularization is a technique used to keep the internal model weights as small as possible, motivated by keeping the model as general as possible [7]. We used the RMSE between CoolWave's prediction and the ground truth as the error function and optimize the learning rate using Adam, a computationally efficient optimization algorithm that that has proven to lead to better results [10]. During training, we used an early stopping technique to stop training the Heating Model when the RMSE on the validation set converged within a margin of 10^{-7} and stop training the Cooling Model when the RMSE on the validation set converged within a margin of 10^{-3}.

4.1.4 Heating and Cooling Model Architecture

The Heating and Cooling Models both had identical Wavenet architectures: each using 6 layers dilated as shown in Table 4.1. Each layer consists of a 1-dimensional filter using the *tanh* activation function which is then multiplied with another 1-dimensional gate using the *sigmoid* function to squash the output of the filter.

Table 4.1: The first layer has a dilation rate of 2, every layer after that has a dilation rate twice that of layer before it.

Layer	Dilation Rate
1	2
2	4
3	8
4	16
5	32
6	64

4.2 Results

4.2.1 Baselines

To evaluate our results, we compare the performance of CoolWave with the performance of a few other known modeling techniques: a standard Wavenet implementation of the same architecture as the Heating and Cooling Models, an Auto-Regressive timeseries model implementation shown to model HVAC systems [14], and a Gaussian Process Regressor implemented by a popular Python library, Sci-Kit Learn [19].

4.2.2 Comparison

To evaluate our results, we compare both the quantitative results and the qualitative results of CoolWave against all baseline models.

Quantitative Results

We compare the quantitative results of CoolWave with all other baseline models using the RMSE over different time ranges. Firstly, we calculate the RMSE of each model over the range of the training data to evaluate each model's ability to predict znt using data from within the bounds of the training data. Secondly, we calculate the RMSE of each model over the range of the testing data to evaluate each model's ability to predict znt using data outside the bounds of the training data. Finally, we calculate the RMSE of each model only on weekends over the range of the testing data to further evaluate each model's ability to predict znt specifically when the HVAC system is off.

Table 4.2 shows the comparison of these results for CSE 2108, CSE 2136, and CSE 2150, respectively. In most cases, CoolWave and Wavenet outperformed the Auto-Regressor and the Gaussian Process Regressor. In CSE 2108, over the range of the training data, CoolWave performed better than Wavenet by a large margin however, over the range of the testing data and on weekends over the range of the testing data, Wavenet performed better than CoolWave by a small margin. In CSE 2136, Wavenet performed better than CoolWave over the range of the training data, the range of the testing data, and on weekends over the range of the testing data. In CSE 2150, over the range of the training data, CoolWave performs worse than both Wavenet and Auto-Regressor, and over the range of the testing data and on weekends over the range of the testing data, CoolWave performed worse than all baseline models.

Qualitative Results

As discussed in Section 4.1.2, while the RMSE provides valuable insight into the overall capabilities of each model, it does not fully describe the behavior of each model. We compare qualitative results by looking at the responsiveness of the znt predictions and the SAFSP. We define "responsiveness" as having two major qualities: a clear change in the trend of a prediction local to the time the HVAC system changes in value, and a clear adherence to the ground truth znt

Table 4.2: RMSE values of each model calculated over various data ranges.

Room	Data Range	CoolWave	Wavenet	Auto-Regressor	Gaussian Process Regressor
2108	Training Data	**1.12**	1.41	1.42	1.49
	Testing Data	0.95	**0.80**	1.64	2.30
	Weekends, Testing Data	1.01	**0.92**	2.50	3.17
2136	Training Data	0.58	**0.39**	0.88	19.7
	Testing Data	0.62	**0.36**	0.73	4.59
	Weekends, Testing Data	0.67	**0.30**	0.99	7.19
2150	Training Data	0.93	**0.67**	0.83	0.97
	Testing Data	1.00	**0.56**	0.71	0.60
	Weekends, Testing Data	0.90	**0.48**	0.80	0.55

patterns, specifically the predictions should not simply resemble the average znt over time. This definition encapsulates the requirements of a model that demonstrates a learned understanding of the physics of heat transfer within a zone.

Figures 4.2, 4.3, and 4.4 show the qualitative comparison of CoolWave against the baseline models when predicting the znt over the range of data outside the bounds of the training data for CSE 2108, CSE 2136, and CSE 2150, respectively. *2018-06-02 - 2018-06-03* is a weekend, *2018-06-04 - 2018-06-09* are weekdays, and *2018-06-10 - 2018-06-11* is another weekend.

In CSE 2108 (Figure 4.2), as the RMSE results suggest, the Auto-Regressor and Gaussian Process Regressor perform notably worse than CoolWave and Wavenet for all rooms. Using our definition of responsiveness, neither the Auto-Regressor nor the Gaussian Process Regressor show a clear change in the trend of prediction with respect to changes in SAFSP values. Likewise, neither model clearly adheres to the general ground truth znt pattern. However, unlike the RMSE results suggested, Wavenet performs worse than CoolWave. While both CoolWave and Wavenet show a clear change in the trend of prediction with respect to changes in SAFSP values, Wavenet does not clearly follow the general pattern of the ground truth znt. Specifically, looking at the predictions between *2018-06-07* and *2018-06-08*, it is clear that CoolWave predictions plateau when the ground truth znt plateaus but Wavenet predictions do not. However, most notably in the hours following, CoolWave predictions drop sharply as does the ground truth znt but Wavenet predictions do the opposite. Looking at the overall prediction trends on weekends, it is clear that both CoolWave and Wavenet closely adhere to the trends of the ground truth znt. However looking at the overall prediction trends on weekdays, Wavenet seems to follow the average of the ground truth znt while CoolWave better models the ground truth znt.

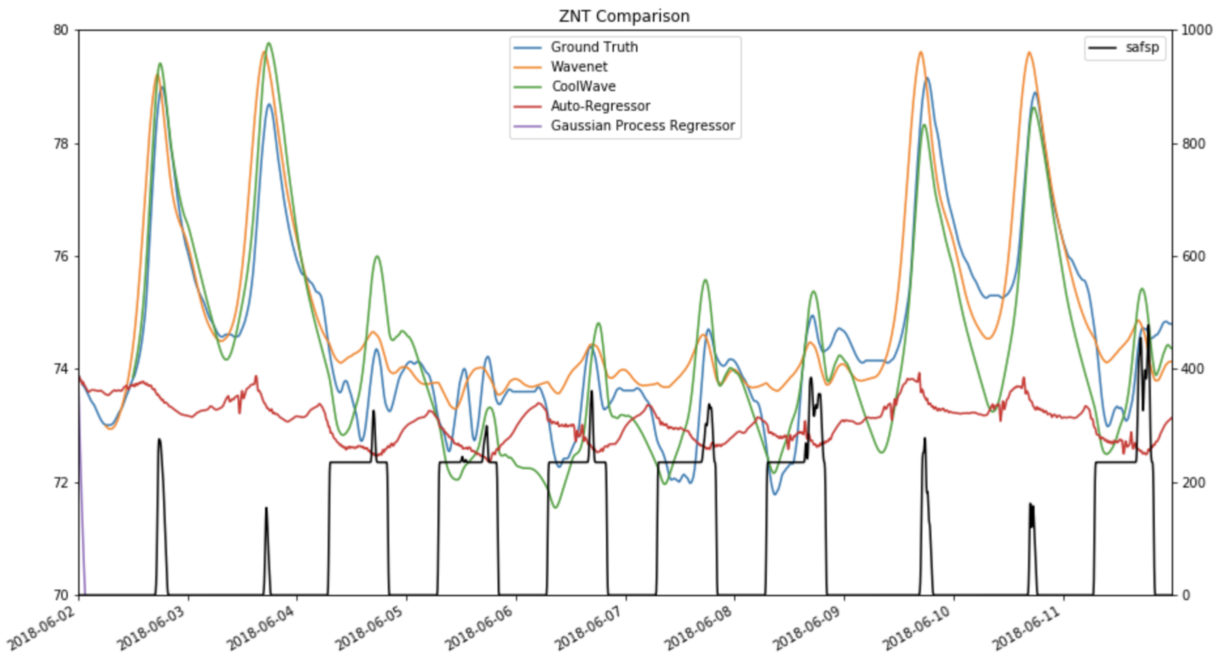

Figure 4.2: A comparison of ZNT predictions over a range of data outside the bounds of the training data in CSE 2108.

In CSE 2136 (Figure 4.3), all but the Auto-Regressor show a strong performance in prediction quality. While the Auto-Regressor does adhere to the general trend of the ground truth znt pattern during the week, it does not do so during the weekend nor does it show a clear change in the trend of prediction local to the time when the SAFSP changes value. Comparing CoolWave, Wavenet, and Gaussian Process Regressor, all follow the general trend of the ground truth znt pattern during the week and weekend. However, the Gaussian Process Regressor does not show a clear change in the trend of prediction local to the time when SAFSP changes value while CoolWave and Wavenet do. In this room, Wavenet appears to qualitatively outperform CoolWave. It is important to note that the SAFSP pattern in this room is more consistent and regular than the SAFSP pattern seen in CSE 2108 which can likely be attributed to the fact that CSE 2136 has a smaller window that receives far less sunlight than CSE 2108. Less direct sunlight means that the room maintains a cool temperature throughout the day which would lead to a regular and consistent SAFSP pattern.

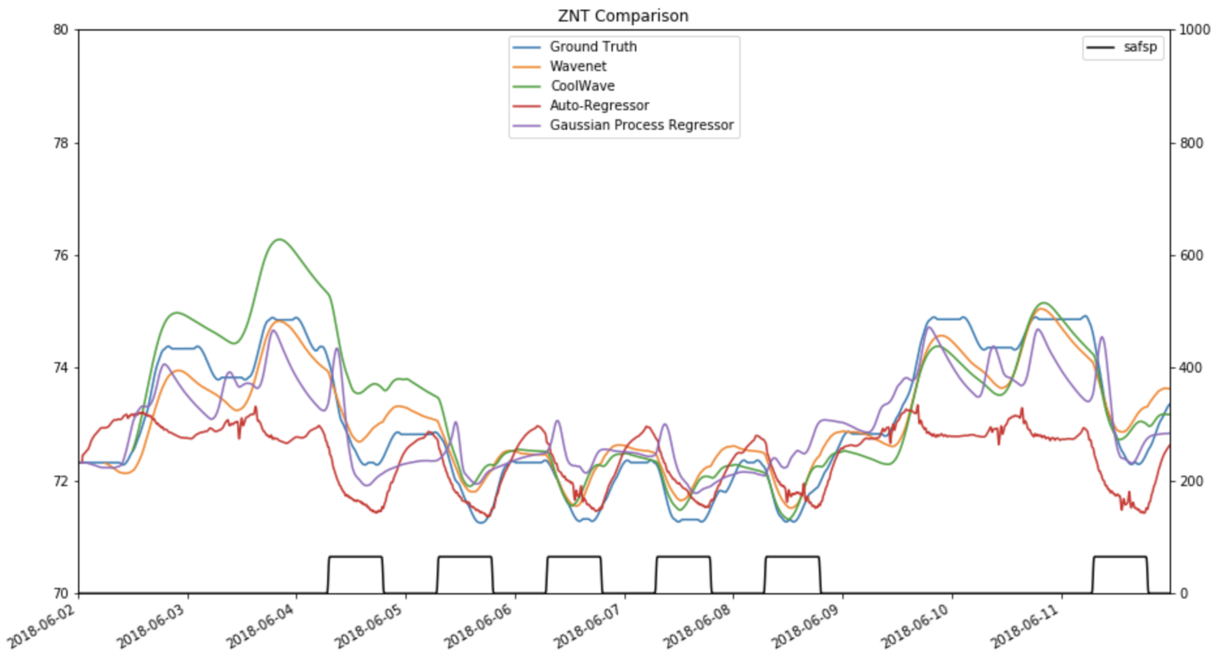

Figure 4.3: A comparison of ZNT predictions over a range of data outside the bounds of the training data in CSE 2136.

In CSE 2150 (Figure 4.4), CoolWave and all three baseline models generally follow the trend of the ground truth znt pattern both during the week and weekends. However, inspecting a closer, CoolWave follows the ground truth znt pattern more closely than the baseline models. Specifically, during the day of *2018-06-07*, CoolWave follows the peaks and valleys of the ground truth znt while the baselines do not. Wavenet in particular does not follow the smaller peaks and valleys of the ground truth znt and instead follows more of an average at the fine-grained level. Most notably, however, is the SAFSP pattern in this room. Specifically, the SAFSP does not reach a value of zero, which would mean that the HVAC system is off, but instead goes no lower than 330, which means that the HVAC system is consistently on. The spikes in the SAFSP value indicate the HVAC system increasing its output as a reaction to increasing znt despite the HVAC system already being on. Also notice that there does not appear to be a clear distinction between weekdays and weekends like there is in CSE 2108 and CSE 2136. This can be attributed to the fact that this room is larger and more open than the other rooms. The effects of the HVAC system are slower to realize and the temperature of the room is harder to change, even as a result of the outside air temperature, the intensity of the sun, and the HVAC system itself. Smaller changes in znt may be the result of human traffic or open doors yet CoolWave is able to capture this behavior as the human traffic must also be expressed in the training data.

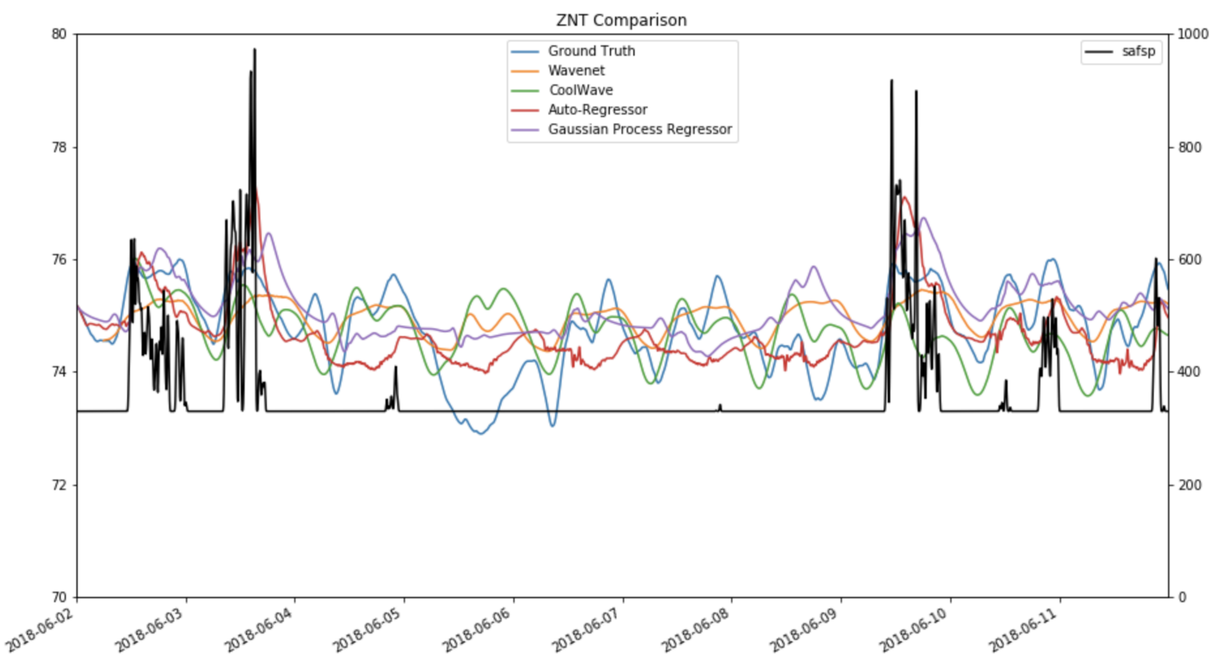

Figure 4.4: A comparison of ZNT predictions over a range of data outside the bounds of the training data in 2150.

Artificial SAFSP Pattern

We conducted an experiment on CoolWave and the baseline models in CSE 2150 to test their ability to generalize to an SAFSP pattern that does not exist in the training or testing data set. In the experiment, the outside air temperature, intensity of the sun, and the initial znt of a zone was taken from the testing data set and the SAFSP pattern was set artificially. Figure

4.5 shows the SAFSP pattern over the course of a single day, June 4th, 2018. The SAFSP is zero from midnight until 7am, at which point the SAFSP alternates every 30 minutes between a value of zero and a value of 450 until 6pm. After 6pm, the SAFSP is zero again. The SAFSP pattern is set artificially so we can not compare the models' predictions to a ground truth znt for quantitative or qualitative comparisons. Instead, we look to the overall effect the HVAC system has on the predictions and consider the intuitive likelihood of the predictions. Firstly, it is clear that the Gaussian Process Regressor's znt predictions are unusual because the predictions contain erratic spikes and drops in znt which is unlikely to occur given our domain knowledge of the physics of heat transfer. Secondly, Auto-Regressor has a distinct pattern from both CoolWave and Wavenet and, given the repeated demonstrated accuracy of both CoolWave and Wavenet, most likely is not an accurate representation of the znt. Furthermore, small unexplained patterns in the predictions such as the sudden increase in znt at the time when the SAFSP first jumps from zero to 450. Now, comparing CoolWave and Wavenet requires more attention to the details of each prediction because the general trend of predictions of both models are similar. The most outstanding difference between the two predictions is the presence of small perturbations that exist in the Wavenet predictions during and after the SAFSP patterns. Again, drawing on our domain knowledge of the physics of heat transfer, it is unlikely that the pattern of znt predictions rapidly alternate as they do starting after the time the SAFSP changes in value from 450 to zero. This supports our conclusion that CoolWave is the most likely predictor of znt given the SAFSP pattern.

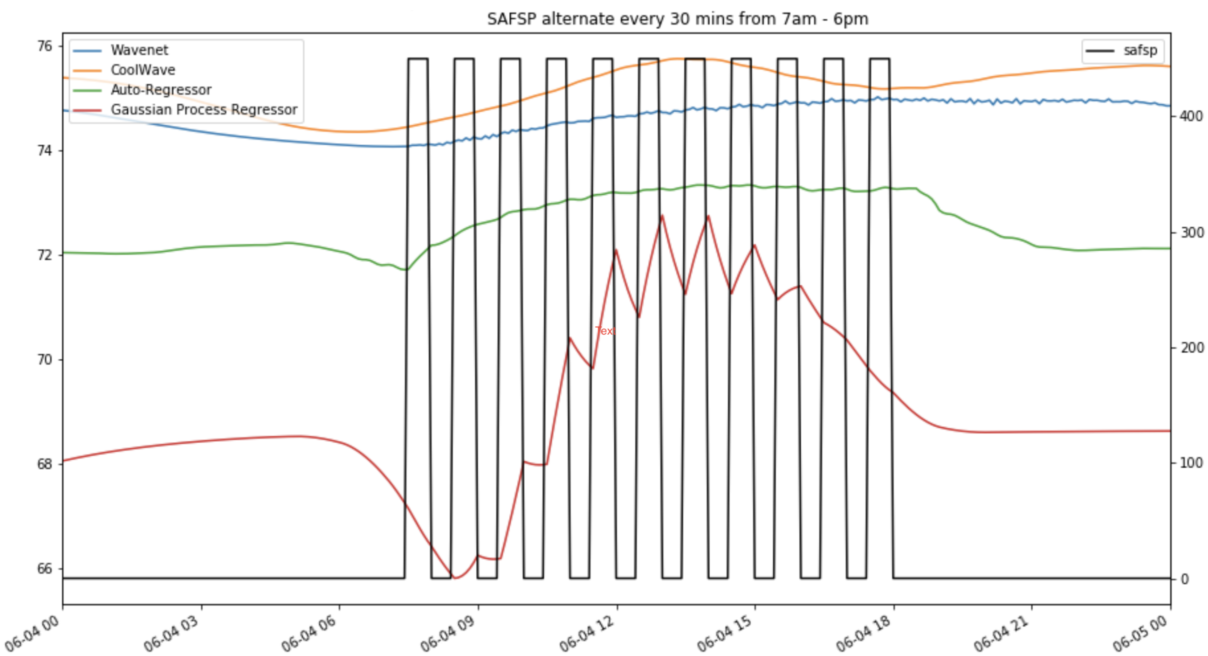

Figure 4.5: An experimental SAFSP pattern in CSE 2150 using the outside air temperature, the intensity of the sun, and the initial znt from the testing data set. CoolWave and the baseline models predict a znt pattern given this SAFSP pattern.

5 Discussion

5.1 Implications of The Results

The results of CoolWave clearly demonstrate the ability to model individual classes of contributing influences, which themselves can be considered simpler, smaller systems, and then combine the results of each to compose a comprehensive model of a more complex system. Underlying this method is the observation that projecting data onto a smaller subspace may expose a higher variance between sets of features and a target feature. In our case, the data was projected onto two separate subspaces: first onto a subspace containing only the outside air temperature, intensity of the sun, and the znt, and second onto a subspace containing only the SAFSP and the znt. In each case, we observed a higher variance in the znt with respect to each group of input features separately and, as a result, we were able to combine them to learn a more general model.

5.2 Limitations of CoolWave

Such a model as CoolWave requires the domain knowledge of the individual factors that influence the temperature of a zone. It was critical to the success of CoolWave that we were able to split the features into two independent classes. If it was not possible to split the features into two different classes, this technique could not work. This imposes a limitation on the applications of such techniques to data that is known to be separable in this manner. Furthermore, CoolWave is limited by the heating and cooling factors that are used to compose the overall predictions. Specifically, if a room is not directly affected by the intensity of the sun or the outside air because it is located in a central location in the building, then there will need to be other considerations for the factors that heat the room, such as hallway temperatures, room occupancy, and other nearby heat loads. If it is the case that these factors are not discernible through the sensor data, then other measures will need to be used to define the factors that influence the heating of a zone.

5.3 Future Work

Here, we identified the outside air temperature and the intensity of the sun as the two sole factors that heated a zone, and the HVAC system as the sole factor that cooled a zone. However, perhaps there are other factors that could contribute to each class. For example, the temperature of the air being injected would influence how quickly the temperature in the zone changes and thus might be a good feature to include in the Cooling Model. It should be noted that we did try to include such a feature, named the Discharge Air Temperature (DAT), however we found that it contributed more noise than signal to our model so future works will need to either find another metric of the temperature of the air being injected into the zone, or find a way to distinguish the noise from the signal.

We have identified only two classes of influences on the temperature of a zone in this thesis. However, future work might find that there are in fact more influencing classes of features that contribute to a zone's temperature. Given more expressive data, there may be other influences within the HVAC system that contribute to a zone's temperature. Finding this internal influences might provide insight into the complex, non-linear relationships between the subsystems of the larger HVAC system.

We also required that the SAFSP be the only control variable available to the model. Yet, using control variables, such as the DAT, may be useful in modeling a zone's temperature more accurately. While this would add complexity to the model, it would expose a relationship to the model giving it data with higher variance, helping with generalizability.

This thesis is coauthored with Koh, Jason and Hong, Dezhi and Gupta, Rajesh K. The thesis author was the primary author of this chapter.

Bibliography

[1] U.s. energy information administration - eia - independent statistics and analysis.

[2] A. Afram and F. Janabi-Sharifi. Review of modeling methods for hvac systems. *Applied Thermal Engineering*, 67(1):507–519, 2014.

[3] C. Agbi, Z. Song, and B. Krogh. Parameter identifiability for multi-zone building models. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, page 6951–6956, Dec 2012.

[4] B. Amos, I. D. J. Rodriguez, J. Sacks, B. Boots, and J. Z. Kolter. Differentiable mpc for end-to-end planning and control. *arXiv:1810.13400 [cs, math, stat]*, Oct 2019. arXiv: 1810.13400.

[5] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin. Identifying models of hvac systems using semiparametric regression. In *2012 American Control Conference (ACC)*, page 3675–3680, Jun 2012.

[6] B. Chen, Z. Cai, and M. Bergés. Gnu-rl: A precocial reinforcement learning solution for building hvac control using a differentiable mpc policy. In *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation - BuildSys '19*, page 316–325. ACM Press, 2019.

[7] C. Cortes, M. Mohri, and A. Rostamizadeh. L2 regularization for learning kernels. *arXiv:1205.2653 [cs, stat]*, May 2012. arXiv: 1205.2653.

[8] D. B. Crawley, L. K. Lawrie, F. C. Winkelmann, W. F. Buhl, Y. J. Huang, C. O. Pedersen, R. K. Strand, R. J. Liesen, D. E. Fisher, M. J. Witte, and J. Glazer. Energyplus: creating a new-generation building energy simulation program. *Energy and Buildings*, 33(4):319–331, Apr 2001.

[9] M. Gouda, S. Danaher, and C. Underwood. Application of an artificial neural network for modelling the thermal dynamics of a building?s space and its heating system. *Mathematical and Computer Modelling of Dynamical Systems*, 8(3):333–344, Sep 2002.

[10] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980 [cs]*, Jan 2017. arXiv: 1412.6980.

[11] J. Koh, B. Balaji, V. Akhlaghi, Y. Agarwal, and R. Gupta. Quiver: Using control perturbations to increase the observability of sensor data in smart buildings. *arXiv:1601.07260 [cs]*, Jan 2016. arXiv: 1601.07260.

[12] A. Kusiak, M. Li, and F. Tang. Modeling and optimization of hvac energy consumption. *Applied Energy*, 87(10):3092–3102, Oct 2010.

[13] A. Kusiak and G. Xu. Modeling and optimization of hvac systems using a dynamic neural network. *Energy*, 42(1):241–250, Jun 2012.

[14] N. Lazic, C. Boutilier, T. Lu, E. Wong, B. Roy, M. Ryu, and G. Imwalle. *Data center cooling using model-predictive control*, page 3814–3823. Curran Associates, Inc., 2018.

[15] J. Li and Z. Pan. Annihilator-preserving maps, multipliers, and derivations. *Linear Algebra and its Applications*, 432(1):5–13, Jan 2010.

[16] T. Lu, X. Lü, and C. Kibert. A hybrid numerical-neural-network model for building simulation: A case study for the simulation of unheated and uncooled indoor temperature. *Energy and Buildings*, 86:723–734, Jan 2015.

[17] M. Ponsen, M. E. Taylor, and K. Tuyls. *Abstraction and Generalization in Reinforcement Learning: A Summary and Framework*, volume 5924, page 1–32. Springer Berlin Heidelberg, 2010.

[18] J. L. Powell. *Chapter 41 Estimation of semiparametric models*, volume 4 of *Handbook of Econometrics*, page 2443–2521. Elsevier, 1994.

[19] C. E. Rasmussen and C. K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, 2006.

[20] A. Trask, F. Hill, S. E. Reed, J. Rae, C. Dyer, and P. Blunsom. *Neural Arithmetic Logic Units*, page 8035–8044. Curran Associates, Inc., 2018.

[21] J. V. Tu. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, 49(11):1225–1231, Nov 1996.

[22] W. J. N. Turner, A. Staino, and B. Basu. Residential hvac fault detection using a system identification approach. *Energy and Buildings*, 151:1–17, Sep 2017.

[23] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W. Senior, and K. Kavukcuoglu. Wavenet: A generative model for raw audio. *CoRR*, abs/1609.03499, 2016.

[24] S. Wu and J.-Q. Sun. A physics-based linear parametric model of room temperature in office buildings. *Building and Environment*, 50:1–9, Apr 2012.