

PSYCHROMETRIC BIN ANALYSIS FOR DATA CENTER COOLING MODES

USING ARTIFICIAL NEURAL NETWORKS

by

ABHISHEK UDAY WALEKAR

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements for the Degree of

MASTER OF SCIENCE IN MECHANICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2018

Copyright © by Abhishek Uday Walekar 2018

All Rights Reserved



Acknowledgements

I would like to thank Dr. Dereje Agonafer for his sustained encouragement and support over the last few years of my thesis work in University of Texas Arlington. I would also like to thank him for providing opportunities to work on industry-funded projects and advancing various networking possibilities through attending conferences. I would like to thank Dr. Haji-Sheikh and Dr. Veerendra Mulay of Facebook for serving on my thesis committee.

I owe a great deal of gratitude to Ashwin Siddarth for being my mentor with regards to my thesis and various other projects. He was a constant source of knowledge and an inspiration to solve research problems. I thank Abhishek Guhe for his guidance and genuine interest to help. I am also thankful to Nikita Sukthankar, Raghvendra Parinam and other EMNSPC team members for their help and support.

Most of all, I would like to thank my family for their emotional and financial support at all times. My mother (Aruna Walekar) and my father (Uday Walekar) who have always been there in my journey, providing enormous strength with their affection and love.

May 20, 2018

Abstract

PSYCHROMETRIC BIN ANALYSIS FOR DATA CENTER COOLING MODES USING ARTIFICIAL NEURAL NETWORKS

Abhishek Uday Walekar, M.S.

The University of Texas at Arlington, 2018

With an increase in the need for energy efficient data centers, a lot of research is being done to increase the use of Air Side Economizers (ASEs), Direct Evaporative Cooling (DEC), Indirect Evaporative Cooling (IEC) and multistage I/DEC cooling. The cooling strategies used to control these systems is based on typical meteorological year (TMY) weather data and thermodynamic principles. But the main drawback of these control strategies is that they do not account for the nonlinearities developed by the conditions inside the data center. So, the primary objective of this study is to use Artificial Neural Networks (ANN) for predicting the cold aisle humidity and temperatures for different modes of cooling. These results can then be studied and then utilized to come up with new bins for each cooling mode. These results will account for the nonlinearities in the data center which are difficult to model using traditional methods.

Table of Contents

Acknowledgements.....	i
Abstract.....	ii
List of Illustrations.....	iv
Chapter 1 Introduction.....	1
1.1 Data Center (DC).....	1
1.2 Air Cooling Unit (ACU).....	2
1.3 Air Side Economization (ASE).....	2
1.4 Direct Evaporative Cooling (DEC).....	3
1.5 Indirect Evaporative Cooling.....	3
1.6 Multi-Stage I/DEC.....	4
1.7 Power Usage Efficiency (PUE).....	5
1.8 Artificial Neural Network (ANN).....	5
Chapter 2 Literature Review.....	7
Chapter 3 Psychrometric Bin Analysis.....	8
3.1 Introduction.....	8
3.2 Psychrometric Bins.....	8
3.3 ANN Based Psychrometric Bin Analysis.....	11
Chapter 4 ANN Modeling and Applications.....	12
4.1 Data Collection and Preprocessing.....	13
4.2 Input and Output Parameters.....	15
4.3 ANN Configuration.....	16
4.4 ANN Training Results.....	17
Chapter 5 Results and Discussions.....	19
5.1 Cold Aisle (CA) Bins.....	19
5.2 Outside Air (OA) Bins.....	23
5.3 Cooling Mode - Hours of Operation.....	27
5.4 ANN Predicted PUE Vs Actual PUE.....	31
Chapter 6 Summary.....	33
References.....	34
Appendix A – MATLAB Code for ANN Modeling.....	35
Appendix B – MATLAB Code for ANN Based Prediction App.....	39
Biographical Information.....	41

List of Illustrations

Figure 1.1 MDC at Mestex	1
Figure 1.2 Air Cooling Unit (ACU)	2
Figure 1.3 Thermodynamic Process for DEC	3
Figure 1.4 Thermodynamic Process for IEC	4
Figure 1.5 Thermodynamic Process for IDEC	5
Figure 1.6 ANN Working - Block Diagram	6
Table 1 ASHRAE Recommended and Allowable Envelopes for IT Equipment.....	9
Figure 3.1 Psychrometric chart showing ASHRAE recommended and allowable environmental envelopes for IT equipment	9
Figure 3.2 Psychrometric chart displaying bin zones for the recommended envelope.....	10
Figure 3.3 Psychrometric chart displaying bin zones for the Allowable A1 envelope	10
Table 2 Cooling strategies for each Bin/Region	11
Figure 4.1 ANN Modelling Flowchart.....	13
Figure 4.2 Exported Data File from Mestex MDC.....	14
Figure 4.3 Anomalies in Rearranged Data Set.....	14
Figure 4.4 Cleaned Data Set.....	15
Figure 4.5 ANN Model.....	16
Figure 4.6 Trained ANN Model (MATLAB Toolbox).....	17
Figure 4.7 Performance Plot.....	17
Figure 4.8 Error Plot.....	18
Figure 4.9 Regression Plot.....	18
Figure 5.1 Predicted CA conditions for ASE ON.....	19
Figure 5.2 Box and Whisker Plot - Predicted CA Conditions ASE ON.....	20
Figure 5.3 Predicted CA conditions for DEC ON.....	20
Figure 5.4 Box and Whisker Plot - Predicted CA Conditions DEC ON.....	21
Figure 5.5 Predicted CA conditions for IEC ON.....	21

Figure 5.6 Box and Whisker Plot - Predicted CA Conditions IEC ON.....	22
Figure 5.7 Predicted CA conditions for IDEC ON.....	22
Figure 5.8 Box and Whisker Plot - Predicted CA Conditions IDEC ON.....	23
Figure 5.9 OA Bin for Recommended Region Using ASE.....	24
Figure 5.10 OA Bin for Recommended Region Using DEC.....	24
Figure 5.11 OA Bin for Recommended Region Using IEC.....	25
Figure 5.12 OA Bin for Recommended Region Using IDEC.....	25
Figure 5.13 OA Bin for A1 Allowable Envelope Using ASE.....	26
Figure 5.14 OA Bin for A1 Allowable Envelope Using DEC.....	26
Figure 5.15 OA Bin for A1 Allowable Envelope Using IEC.....	27
Figure 5.16 OA Bin for A1 Allowable Envelope Using IDEC.....	28
Figure 5.17 Percentage Hourly ASE usage for Recommended and A1 Envelopes.....	28
Figure 5.18 Percentage Hourly DEC usage for Recommended and A1 Envelopes.....	28
Figure 5.19 Percentage Hourly IEC usage for Recommended and A1 Envelopes.....	29
Figure 5.20 Percentage Hourly IDEC usage for Recommended and A1 Envelopes.....	29
Figure 5.23 Percentage Hours of Operation in Recommended Envelope.....	30
Figure 5.24 Percentage Hours of Operation in Allowable Envelope.....	31
Table 3 Average PUE.....	31
Figure 5.25 ANN Based Cold Aisle Prediction App.....	32

Chapter 1

Introduction

1.1 Data Center (DC)

Data center is a facility that houses the IT equipment which includes the servers, racks, power supply and backup units, switches, and the connecting cables. Along with these data centers also include other important monitoring and security devices like sensors, fans, fire protection systems and environmental controls. IT operations are crucial part of most organizations all over the world. Most of them rely on their IT systems to run their operations. Thus, maintaining the data centers at the most optimum conditions and running them without any down times is very important for the organization's growth. Cooling systems are used to maintain the favorable conditions inside the data centers. Thus, a lot of energy is spent for running and operating the data centers. A large data center is an industrial-scale operation using as much electricity as a small town.

Modular Data Center systems are a portable form of data centers that can be added or moved as per the requirement. They are modelled to fit the data center equipment into a standard shipping container or similar cabinets. MDC are designed for rapid deployment, energy efficiency and better computing density and require power supply, internet access and chilled water supply for deployment.

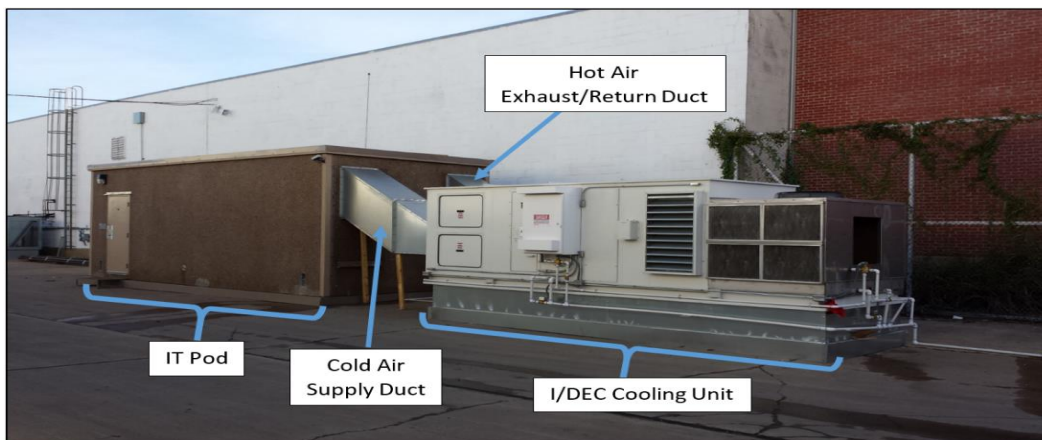


Figure 1.1 MDC at Mestex

1.2 Air Cooling Unit (ACU)

Cooling units are used to maintain the favorable conditions inside the data center. These units include different type of cooling modes, ducts, fans, and controls. The type of cooling is mainly controlled using the outside air condition and the cold isle condition. The MDC at Mestex comprises of three modes of cooling: Air side Economizer (ASE or Free Cooling), Direct Evaporative Cooling (DEC) and Indirect evaporative cooling (IEC). Precise control and switching between the modes of cooling is important for running the data center efficiently with minimum cost and down times.

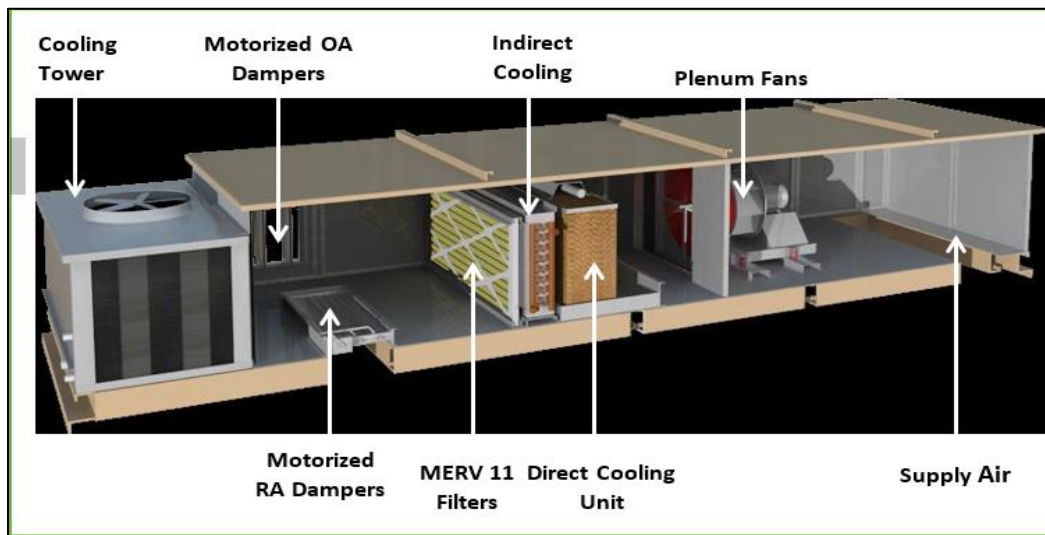


Figure 1.2 Air Cooling Unit

1.3 Air Side Economization (ASE)

Air side economization is used when the ambient air conditions are favorable for cooling the data center. This minimizes the need of external cooling, thus significantly reducing the energy consumption of a data center. ASHRAE defines ASE as “a duct and damper arrangement and automatic control system that together allow a cooling system to supply outdoor air to reduce or eliminate the need for mechanical cooling during mild or cold weather” [1]. In this mode of cooling, outside air is drawn using fans and then filtered for contaminants before being supplied to the cold isle.

ASE can be used fully (free cooling) or partially depending on the ambient air conditions. The major limitation for ASE is that it can be used only if the outside temperature, humidity, and the air contaminants are

within a specified range. This limits the use of ASE to only a few hours a year. To increase the use of ambient air for compressor less cooling, ASE is used along with Direct Evaporative Cooling (DEC), Indirect Evaporative Cooling (IEC) or a two-stage Indirect/Direct Evaporative Cooling (I/DEC).

1.4 Direct Evaporative Cooling (DEC)

Direct Evaporative Cooling (DEC) is a method of cooling air through direct contact with water. In this method, the warm air is directly blown over a water damped media. As the air comes in contact with the water, it loses its energy to evaporate the water in the form of latent heat of vaporization thus reducing the temperature of air. As the total energy content of the system remains same, this mode of cooling is also called as adiabatic cooling. Thus, the wet bulb temperatures of the ambient air and the conditioned air remains the same. The DEC process can be represented on the psychrometric chart as shown below in figure 1.3.

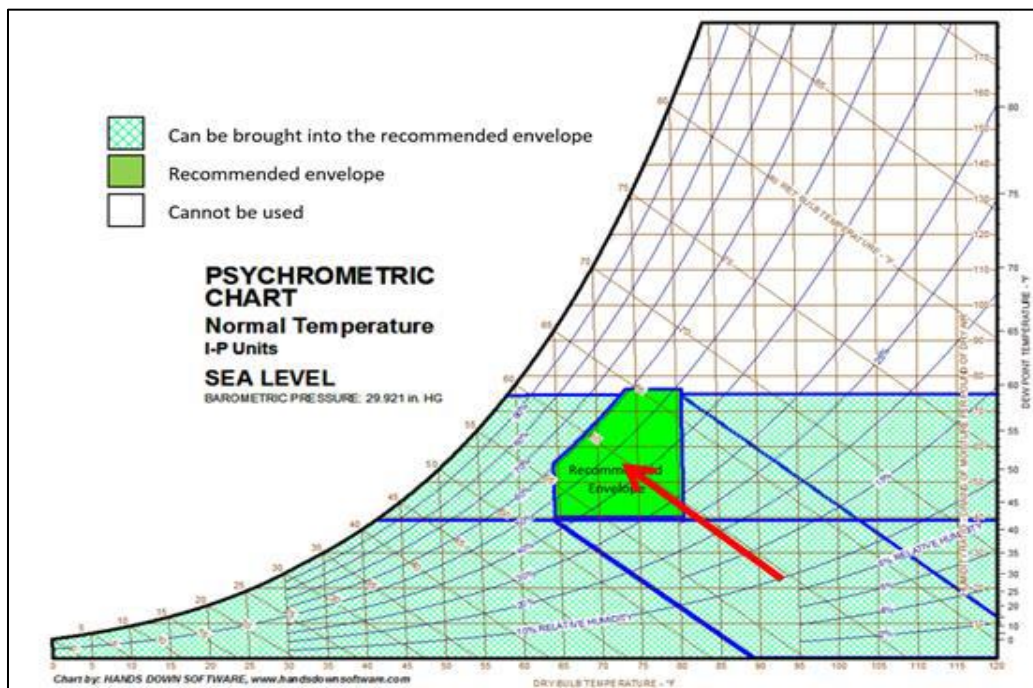


Figure 1.3 Thermodynamic Process for DEC [2]

1.5 Indirect Evaporative Cooling

Indirect Evaporative Cooling (IEC) is a method of cooling air through indirect contact with water. In this method, warm ambient air is passed over a cooling coil that has cold water flowing through it. The air loses

energy to the cold water flowing through the coil thus reducing its temperature. As there is no direct contact of air with the water, the moisture content or specific humidity of the air remains constant while the dry bulb decreases and the relative humidity increases.

As the energy lost by the air to the water flowing through the coils, the outlet temperature of the water is higher than its inlet temperature. This water is then ducted to the cooling tower and distributed over a DEC media. The cooling tower fans pull the ambient air over the media pads thus cooling the water flowing down the pads. This water is then collected in the sump at the bottom of cooling tower and then pumped back into the coil. This process is shown on the psychrometric chart below.

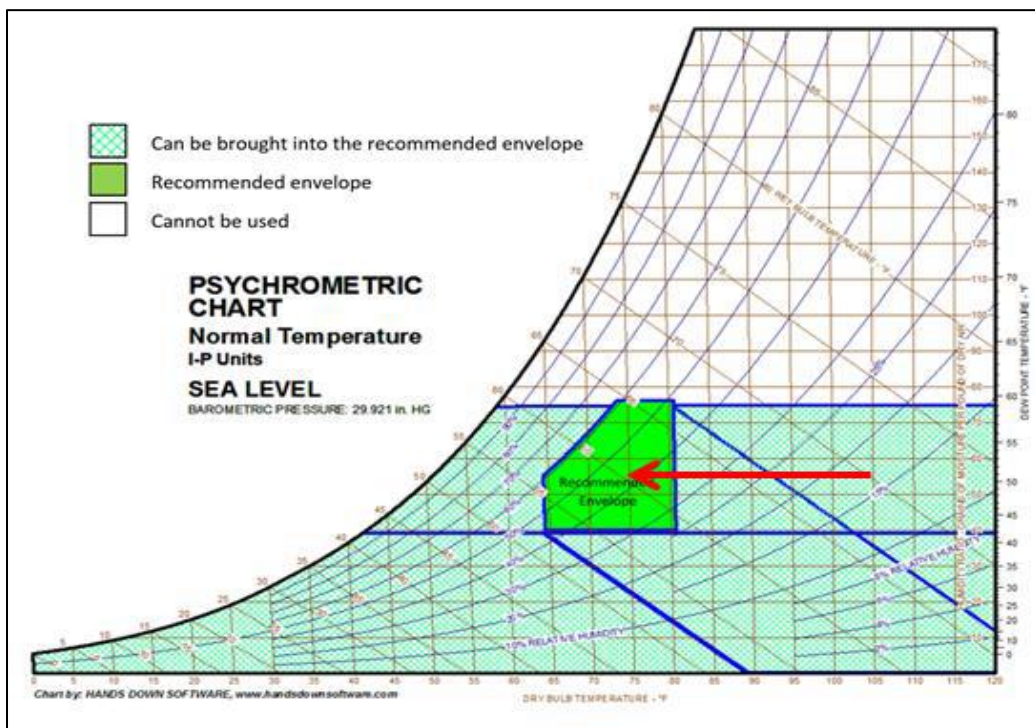


Figure 1.4 Thermodynamic Process for IEC

1.6 Multi-Stage I/DEC

Multi-Stage I/DEC allows further wider range of ambient air conditions to be used for compressor-less cooling of data centers. In multi-stage cooling, the air is first cooled indirectly by passing it over the coils. In this first stage, the dry bulb temperature is decreased thus cooling the air sensibly. In the second stage DEC is used to cool the air further. In this stage, there is an increase in the humidity content and the air is cooled

adiabatically. Multi-stage is generally used when the temperature is hot and dry. The multi-stage I/DEC is shown in figure 1.5 below.

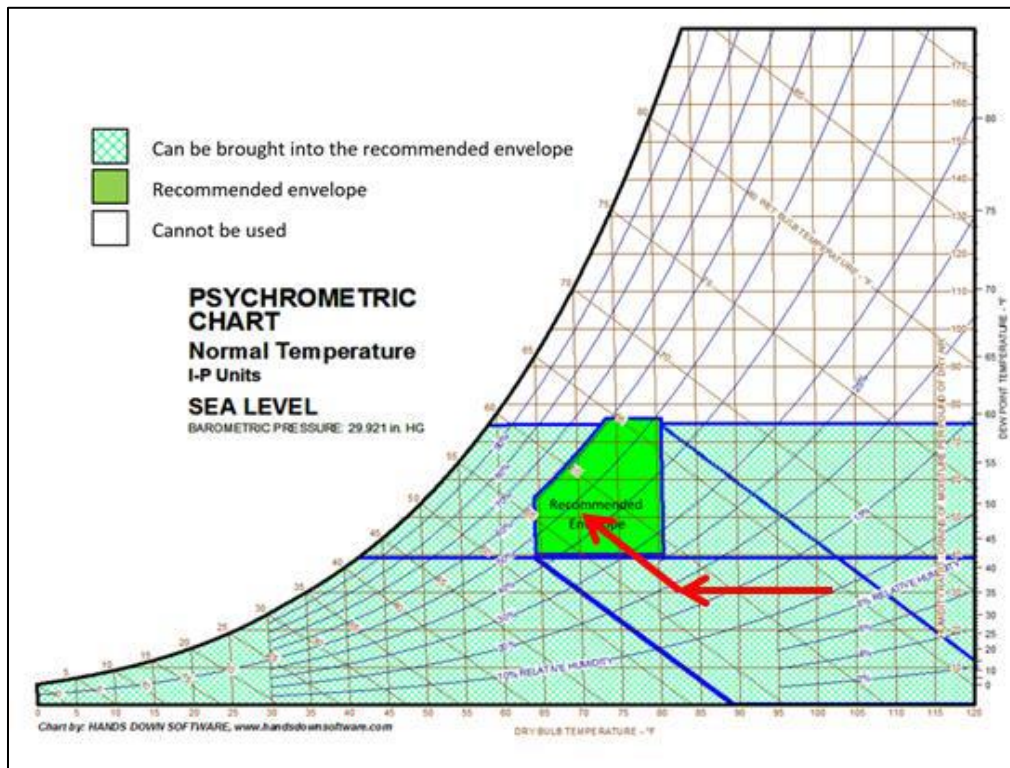


Figure 1.5 Thermodynamic Process for IDEC

1.7 Power Usage Efficiency (PUE)

Power usage effectiveness (PUE) is a ratio that describes how efficiently a computer data center uses energy; specifically, how much energy is used by the computing equipment (in contrast to cooling and other overhead). An ideal PUE is 1.0. Anything that isn't considered a computing device in a data center (i.e. lighting, cooling, etc.) falls into the category of facility energy consumption [3].

$$PUE = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}} = 1 + \frac{\text{Non - IT Facility Energy}}{\text{IT Equipment Energy}}$$

1.8 Artificial Neural Network (ANN)

ANN is a machine learning tool which works similar to the human brain. It is a black box model which learns based on the data that we feed to the network while training and validation. The ANN comprises of a large number of neurons similar to the neurons in the brain. The ANN model works on the machine learning algorithm

which uses the find out precise relations between the input parameters and the outputs through these hidden layers and neurons.

Figure 1.6 shows the basic working mechanism of the ANN algorithm. First the input is normalized and then feed to the neuron in the input layer. Every neuron in the network has an activation function defined. It can be either be linear or non-linear as per the type of data and its variation with the output. Generally, a non-linear activation function is used for input and hidden layers whereas, a linear activation function is used for the output layer. The output value from the neuron is calculated based on the input value and its activation function. This output from the first neuron is then multiplied with the weights and is then given as an input to the neuron in the next layer.

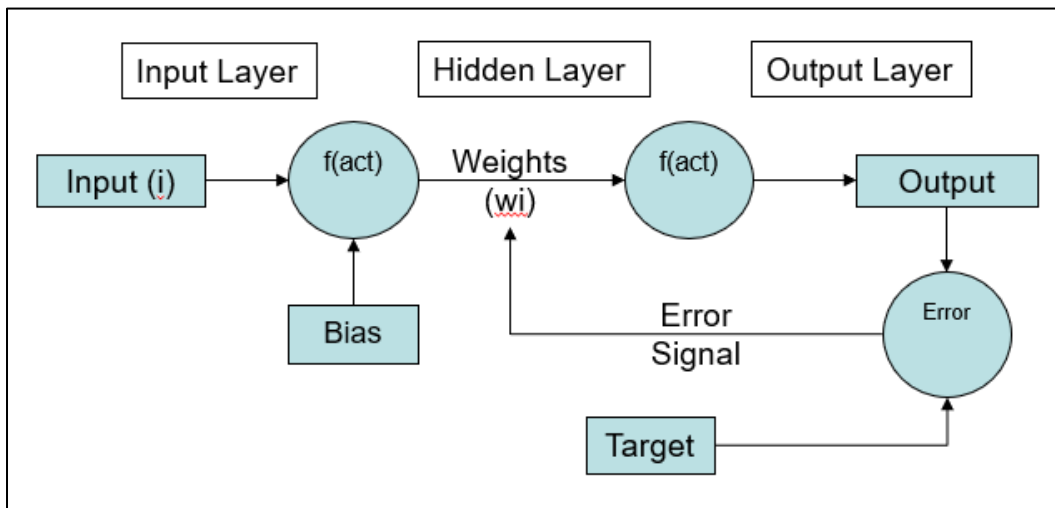


Figure 1.6 ANN Working - Block Diagram

Finally, the output is calculated from the neuron in the output layer with a linear activation function. This output is then compared with the actual target value to calculate the error value. This error is then backpropagated to adjust the weights between the neurons reducing the error value in the next iteration. This loop is carried on till the error is small enough to be neglected or is within the required accuracy.

Chapter 2

Literature Review

In the last couple of decades, extensive research has been done to reduce the cooling cost of data centers. Some of the focus of this research has been on mineral oil immersion cooling [4], segregated cooling of IT servers [5] and reducing the use of Direct Expansion (Compression). Direct expansion (DX) cooling accounts for almost 50% of the total cooling power consumption. The use of the DX can be reduced by using other compression less modes of cooling like air side economization and evaporative cooling.

As we know, data center is a combination of complex mechanical and electrical system with non-linear interdependencies and thus making it impossible to model and optimize it analytically. The control strategies being currently used are based on ASHRAE recommended regions [2] [6] and it does not account for these non-linearities present in the data center. Machine learning tools like Artificial Neural Networks (ANN) can easily account for these non-linearities by using actual operation sensor data to train the network [7]. This study focuses on increasing the usage of compression less cooling by optimizing the control strategies of the air cooling unit using a well-trained ANN framework.

Previous studies have used ANN for applications like airflow and temperature optimization [8], fault diagnosis for temperature and flow rate sensors [7] [9], plant configuration optimization [7], predicting energy demands [10], task scheduling [11] and cooling predictions [12]. The objective of this study is to use the well-trained ANN framework to predict the cold aisle conditions and PUE and then use these results for defining new bins for optimized control of the cooling system.

Chapter 3

Psychrometric Bin Analysis

3.1 Introduction

Data centers are high energy consumption devices and require continuous cooling to maintain the favorable conditions inside the IT pod. Majority of data centers use direct-expansion cooling which accounts for almost 50% of the energy usage of the data center. Alternatively, to reduce the power consumption, a combination of cooling modes is implemented which include air side economization, direct evaporative cooling, indirect evaporative cooling, and supplemental direct expansion. Using these alternative modes of cooling can reduce the data center cooling energy consumption by up to 80% [6].

The different cooling modes are examined individually and in combination with others to categorize seven distinct alternative cooling strategies.¹ These strategies are represented on the psychrometric chart by bins or zones. The bounds of the bins are defined by the capability of each cooling mode's thermodynamic processes to achieve supply air in the ASHRAE recommended or allowable environmental envelopes for the IT equipment [6].

3.2 Psychrometric Bins

A psychrometric chart represents the thermodynamic properties of moist air, i.e. its graphical equation of state. It can be used to graph the thermodynamic properties of air, thus the condition of the air. In this study, we have used the ASHRAE recommended and A1 allowable envelopes to define the favorable conditions for the IT equipment. The bins for the different control strategies have been defined based on these envelopes. Table 1 shows the temperature and moisture boundaries that define the ASHRAE recommended and A1 allowable environmental envelopes. Figure 3.1 shows these boundaries on the psychrometric chart.

Table 1 ASHRAE Recommended and Allowable Envelopes for IT Equipment

	Recommended Envelope	A1 Allowable Envelope
Low End Temperature	64.4°F (18°C)	59°F (15°C)
High End Temperature	80.6°F (27°C)	89.6°F (32°C)
Low End Moisture	41.9°F DP (5.5°C)	20% RH
High End Moisture	60% RH & 59°F DP (15°C)	80% RH & 63°F DP (16.8°C)

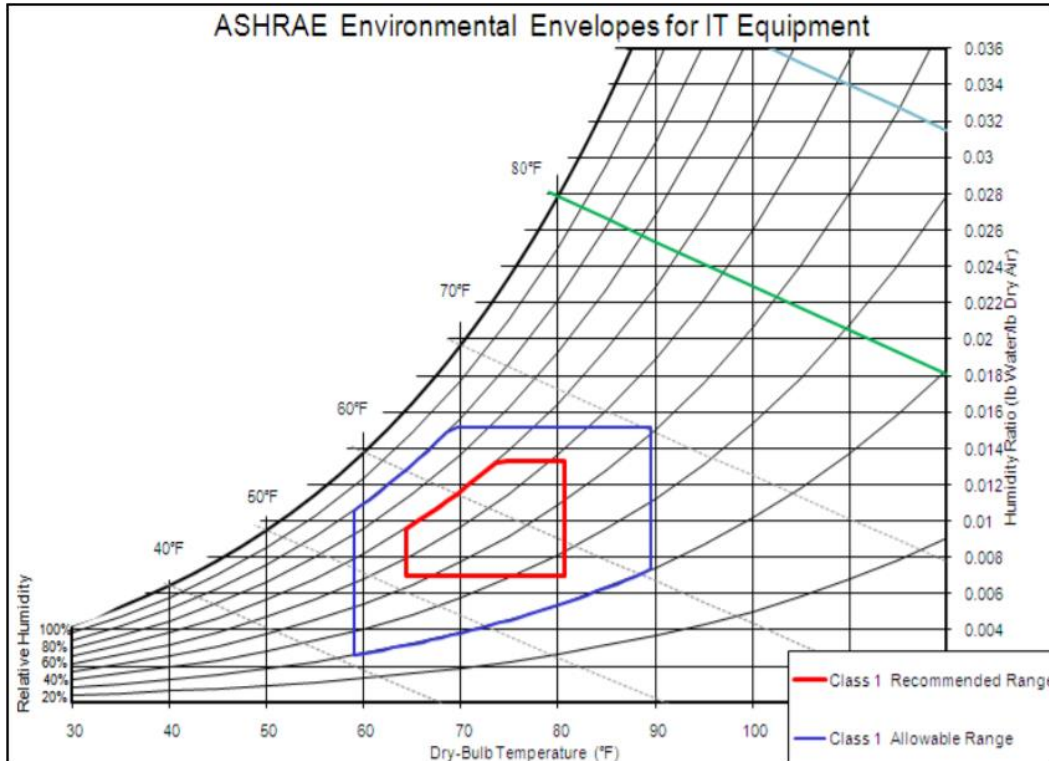


Figure 3.1 Psychrometric chart showing ASHRAE recommended and allowable environmental envelopes for IT equipment [13].(Source: Ian Metzger, NREL)

Figure 3.2 shows the defined bins or regions base on the different modes of cooling that can be used to condition the outside air to the recommended envelope. Similarly, Figure 3.3 shows the bins for the allowable A1 envelope. Table 2 represents the cooling mode associated with each region.

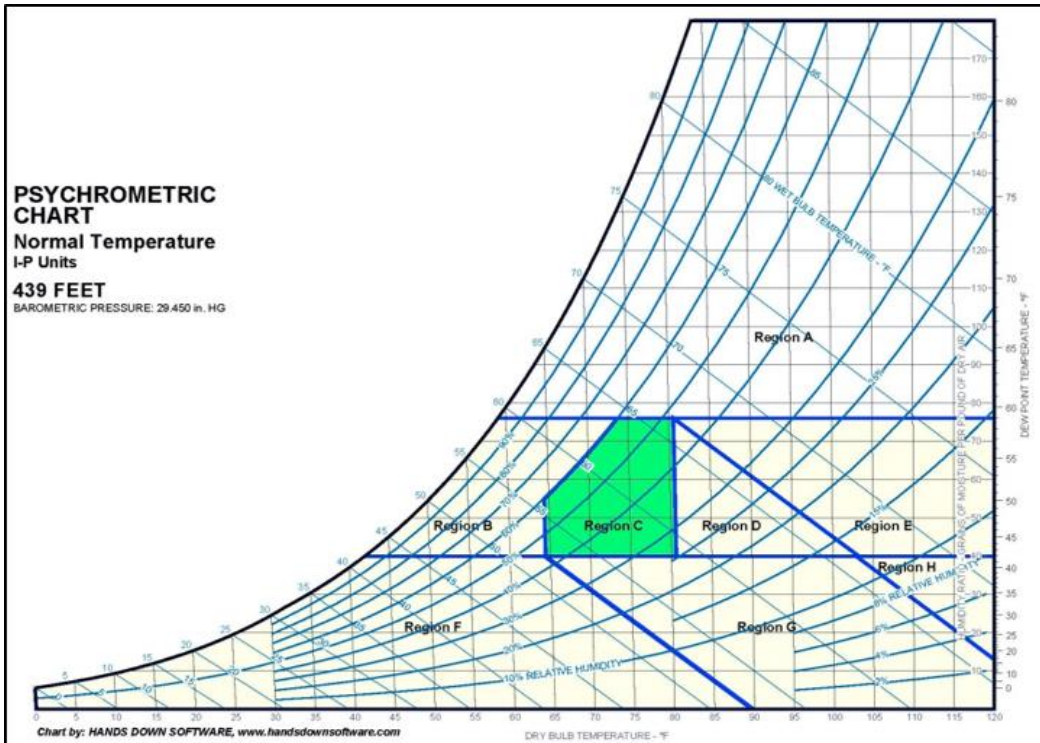


Figure 3.2 Psychrometric chart displaying bin zones for the recommended envelope

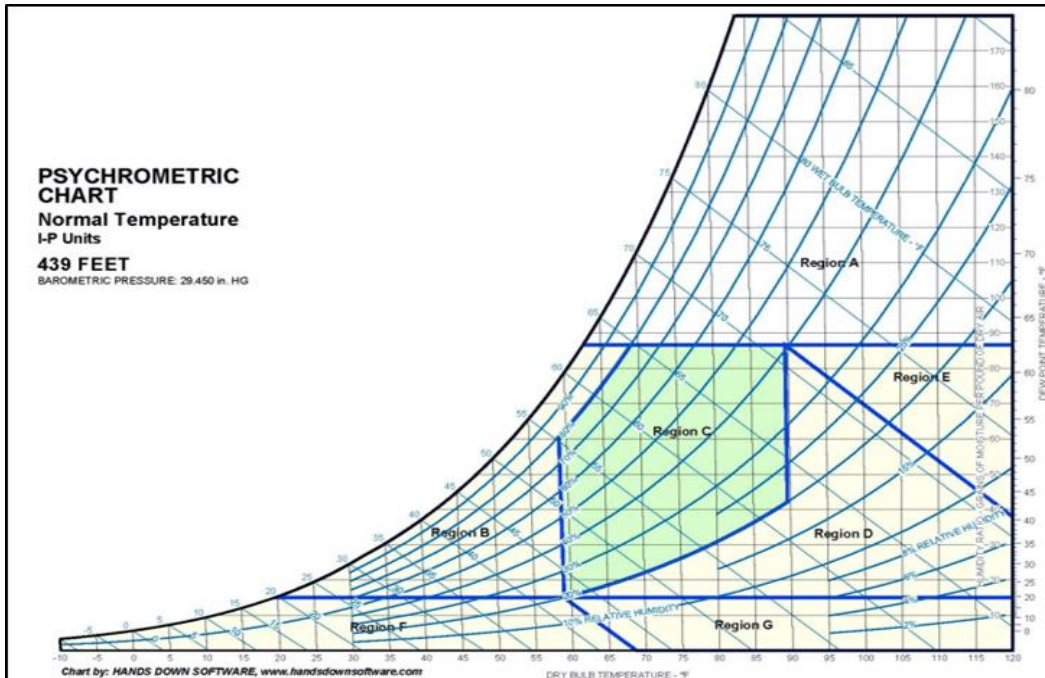


Figure 3.3 Psychrometric chart displaying bin zones for the Allowable A1 envelope

Table 2 Cooling strategies for each Bin/Region

<i>Bin/Region</i>	<i>Cooling Mode</i>
Region A	Direct Expansion (DX)
Region B	Air Side Economizer (ASE)
Region C	Free Cooling
Region D	Indirect Evaporative Cooling (IEC) or Direct Evaporative Cooling (DEC)
Region E	Indirect Evaporative Cooling (IEC)
Region F	Air Side Economizer (ASE) + Direct Evaporative Cooling (DEC)
Region G	Direct Evaporative Cooling (DEC)
Region H	Indirect Evaporative Cooling (IEC) + Direct Evaporative Cooling (DEC)

The bound for the regions in figure 3.2 are mainly based on the thermodynamic process of each cooling mode and their effectiveness. If the outside air conditions fall in region A, only direct expansion cooling must be used as the air needs to be dehumidified and then cooled which is not possible using compression less cooling. In region B, as the air is cool and humid, air side economization must be used to mix the outside air with the return air to reduce the relative humidity of the outside air. Region C represents the recommended envelope itself and thus here no conditioning is required. Hence the outside air can be directly supplied to the IT pod. For the air conditions under region D, either of the direct or indirect evaporative cooling can be used. Here air can be cooled sensibly or adiabatically to condition it to the recommended envelope. Air in region E must be cooled sensibly using indirect evaporative cooling as the moisture content of air in this region is high. For region F, as the air is cool and humid, a combination of air side economization and direct cooling must be used. First the air will be dehumidified and heated by mixing it with the return air and then cooled adiabatically to achieve the required condition. Direct evaporative cooling must be used for region G as air is hot and dry. In region H, as the air is hot and dry, multistage direct indirect cooling can be used, where air will be first cooled sensibly using indirect cooler to region G and then adiabatically by the direct cooler to finally get to the recommended condition.

3.3 ANN Based Psychrometric Bin Analysis

Data center facilities comprise of many sensors and control algorithms to optimize it's working. This introduces many non-linearities inside the data center which are impossible to capture mathematically or analytically. Also, the psychrometric bin associated with the cooling mode and hence, it cannot account for these

non-linearities inside the data center. This may result in overcooling of the data center, increasing its energy consumption.

Using ANN, it is possible to account for these non-linearities thus giving more realistic results. In this study, we have used an ANN model which utilizes the sensor data from a real data center at Mestex, Dallas for training and validation. As the readings from the sensor data are used as inputs and outputs, the model captures the variation in the data center caused due to the non-linearities. Further, we have then used the TMY3 data from Dallas Love Field as input to the trained ANN network and used these results to define new psychrometric bins.

Chapter 4

ANN Modeling and Applications

The main object of this study was to come up with a well-trained ANN model that could predict the conditions inside the data center and then use this model for various applications. Some of the important steps in modelling the ANN were data preprocessing, selection of input and output parameters, understanding and selecting the right ANN architecture and configuring it to get the best accuracy and finally validate it with a different data set. Figure 4.1 shows the steps we followed for modelling the neural network.

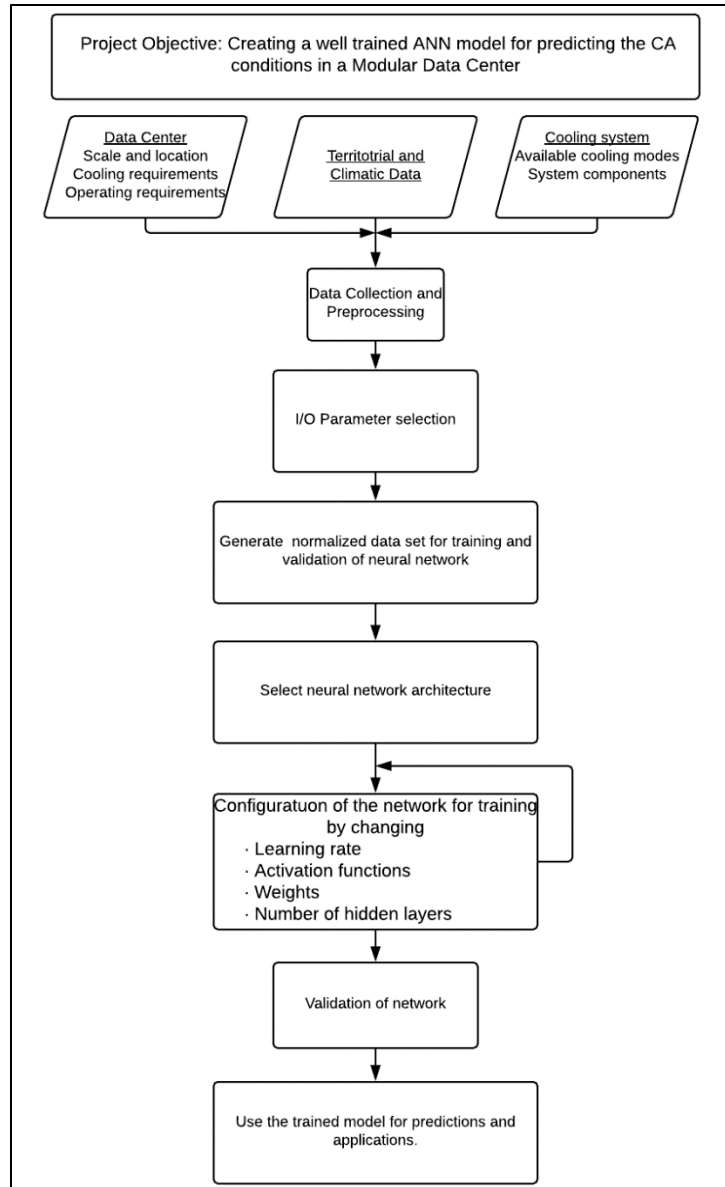


Figure 4.1 ANN Modelling Flowchart

4.1 Data Collection and Preprocessing

The data collected from the sensors at a fully functional modular data center was used for training and validation of the ANN model. This data was exported in the form of access files with each sensor denoted by a specific TID and each file containing a minute-wise data for a month. Also, the exported data was arranged as per the TID and the date and time stamp. A snapshot of the data file is shown in figure 4.2. The main challenge with the data was to rearrange it as each file contained around 800000 columns. We used python scripts and

tableau to rearrange this data. Also, each sensor had different recording intervals and thus the date stamp for each reading had to be matched to get a complete data set.

TID	SEQUENCENU	DATE_STAMP	RECORD_TYP	DATA_VALUE	STATUS_FLAGS
0	1	4/30/2015 5:09:00 AM		2 13.1008	0
0	2	4/30/2015 5:10:00 AM		2 13.084698	0
0	3	4/30/2015 5:11:00 AM		2 13.096198	0
0	4	4/30/2015 5:12:00 AM		2 13.126101	0
0	5	4/30/2015 5:13:00 AM		2 13.1652	0
0	6	4/30/2015 5:14:00 AM		2 13.149097	0
0	7	4/30/2015 5:15:00 AM		2 13.130702	0
0	8	4/30/2015 5:16:00 AM		2 13.1652	0
0	9	4/30/2015 5:17:00 AM		2 13.137602	0

Figure 4.2 Exported Data File from Mestex MDC

Another issue with the data was the spikes and anomalies caused due to the faulty sensor readings. The data from each sensor had to be plotted to find out these spikes and then removed to get a clean data set. Figure 4.3 shows the sensor errors or spikes while figure 4.4 represents the cleaned data. A new set of hourly TMY3 data from Dallas Love Field was used to run the tests and the predicted output data was then utilized for cooling mode bin and weather bin analysis.

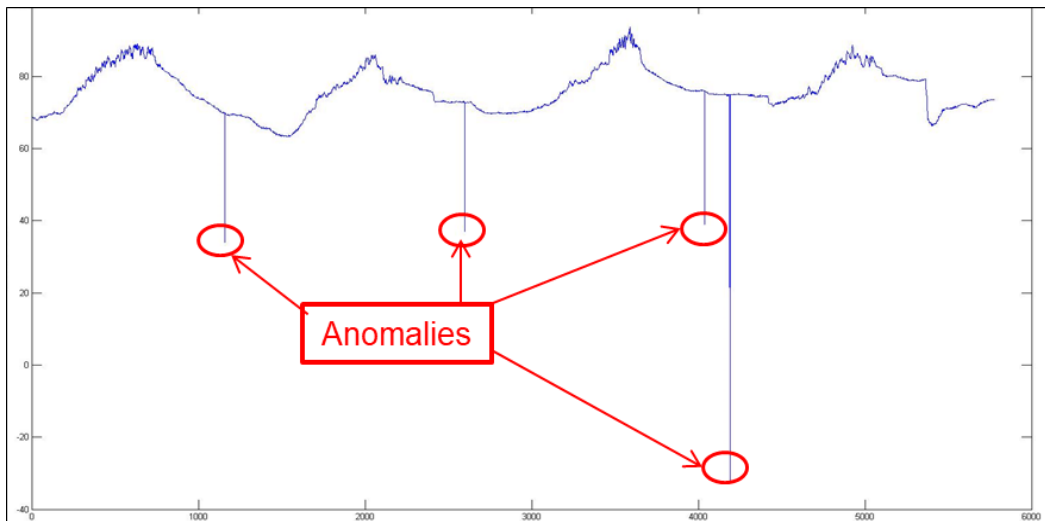


Figure 4.3 Anomalies in Rearranged Data Set

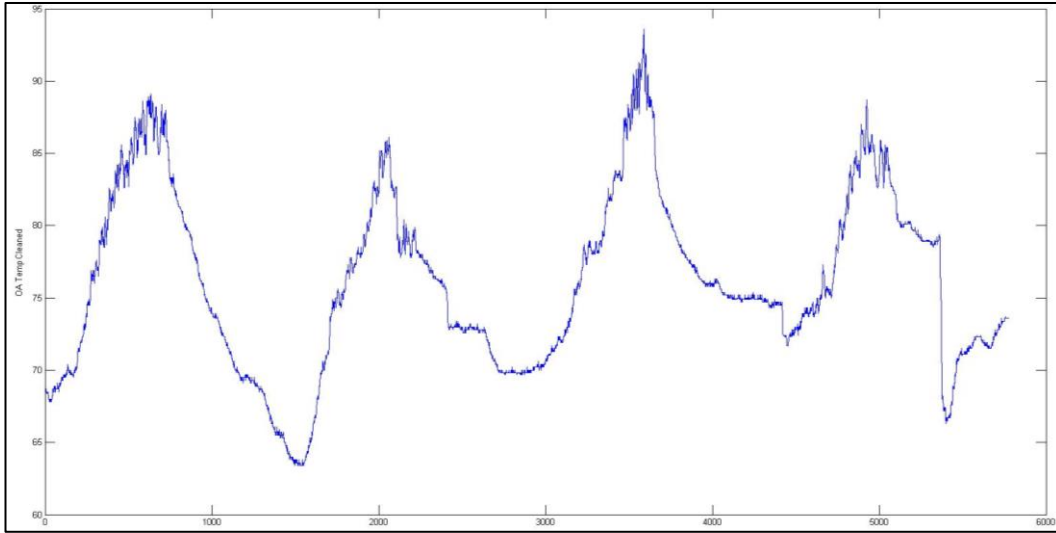


Figure 4.4 Cleaned Data Set

4.2 Input and Output Parameters

The selection of the correct input variables is very crucial for training and accuracy of the model. Only the parameters that have a major impact on the output must be selected. Also, the inputs must not be correlated with each other as it creates redundancy. In this study, we have used seven independent parameters as inputs, which are not affected by the conditions inside the data center. The model has been trained to predict the PUE, cold aisle temperature and the cold aisle humidity. The list of input and output parameters is as given below.

Input Parameters

Outside Air Temperature

Outside Air Humidity

IT Load

Temperature difference Across the Racks
(Delta T between the CA and the HA)

Cooling Types (ASE, DEC, IEC)

Output Parameters

PUE

Cold Aisle Temperature

Cold Aisle Humidity

4.3 ANN Configuration

The ANN model learns based on the training data that is given to it. Thus, an ANN architecture with back propagation algorithm must be used for training the model. In this study, Levenberg-Marquardt algorithm with feedback was used to train the model. Levenberg-Marquardt optimization algorithm is a supervised learning algorithm that updates the weights and biases based on the error value that is backpropagated. Thus, the accuracy of the model is increased with each iteration. The learning is stopped once the model reaches a certain minimum value or if it becomes constant and doesn't change with increasing iterations.

Figure 4.5 represents the ANN model used in this study. It consists of 7 input parameters and 3 output parameters with 20 hidden layers. The value for the hidden layers has been selected by trial and error. The model was trained with increasing number of hidden layers starting at 5. Most accurate results were recorded when the model was trained with 20 hidden layers. Further increasing the number of hidden layers increased the computation time but didn't significantly decrease the errors.

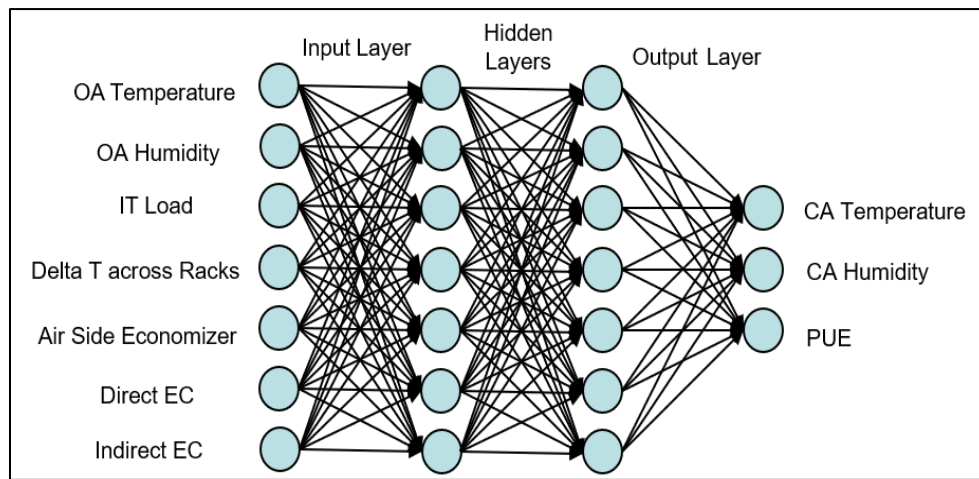


Figure 4.5 ANN Model

The learning rates and the initial weights and biases are pre-defined by the Levenberg-Marquardt algorithm. Each node has an activation function defined to it which calculates the output from each neuron. A sigmoidal non-linear activation function is used for the input and hidden layers whereas a linear activation function is used for the output layer. These activation functions are as shown below.

$$f(x) = e^x / (e^x + 1)$$

Sigmoidal Function

$$f(x) = a * x + b$$

Linear Function

4.4 ANN Training Results

The baseline model configured as above was tested for accuracy and training state. Figure 4.6 represents the trained network model as seen in the MATLAB toolbox. As defined before, it consists of the 7 input and 3 output parameters with 20 hidden layers and a single output layer.

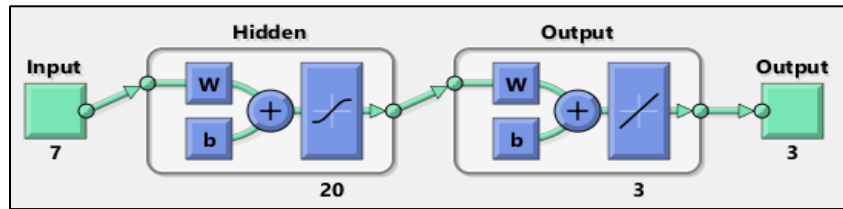


Figure 4.6 Trained ANN Model (MATLAB Toolbox)

The result in figure 4.7 shows the maximum error present during validation of the network. As seen from the figure 4.7, the mean square error starts above 10000 but decreases with an increase in the number of iterations or epochs. The error value remains constant after 50 epochs finally stops at 10.2059 after 334 iterations

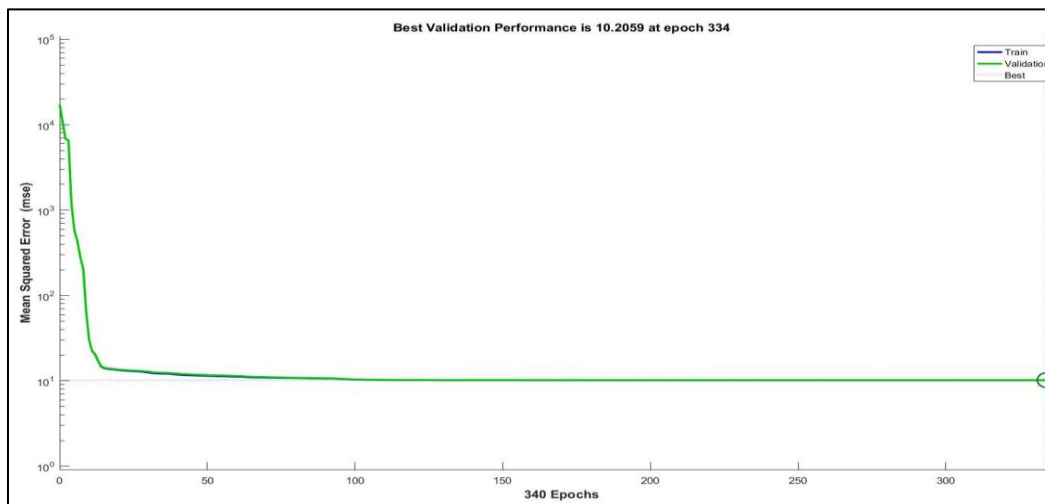


Figure 4.7 Performance Plot

Figure 4.8 and figure 4.9 shows the error and regression plots respectively. In figure 4.8, the x-axis represents the error value where as the bars represent the instances. It can be seen from the error plot that for maximum instances the error is recorded to be between ± 3.5 with maximum validation error of 10.206.

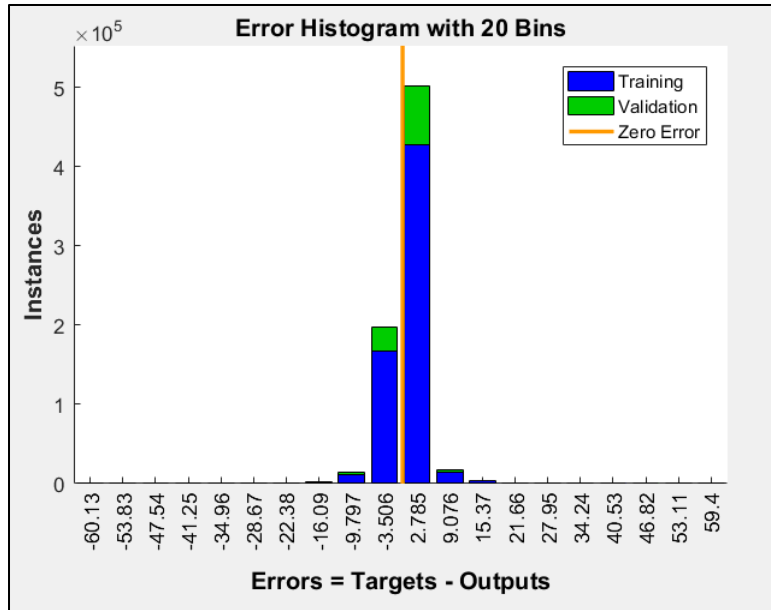


Figure 4.8 Error Plot

The regression plot in figure 4.9 shows the actual fit represented by the red line and the output data points represented by the black circles. The distance between the line and the circle represents the error. This plot gives an overview of how well the model is trained and its accuracy to predict the outputs.

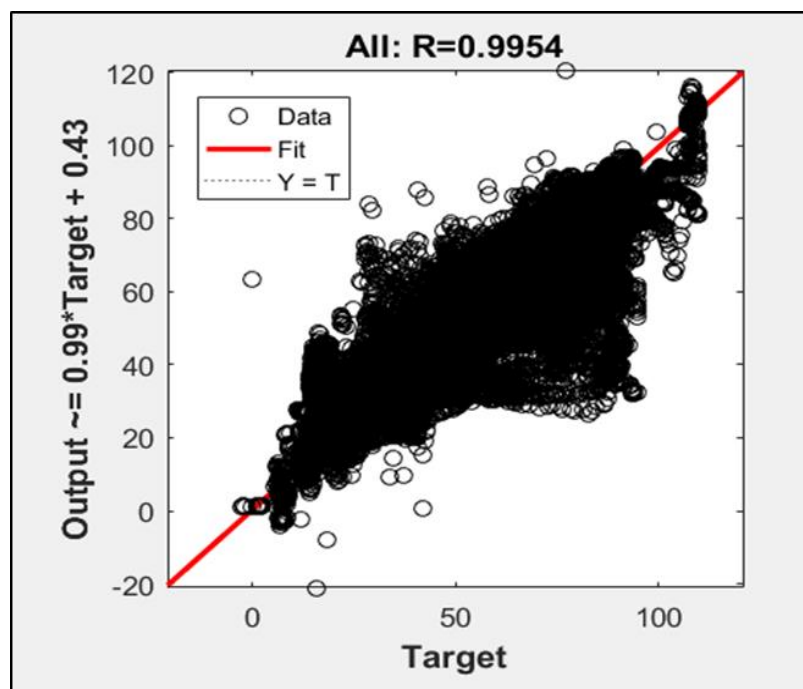


Figure 4.9 Regression Plot

Chapter 5

Results and Discussions

The cold aisle conditions were predicted using the trained ANN model. The OA temperature and humidity from the TMY3 data for Dallas Love Field was used as input along with the IT load, delta T across racks and the different cooling modes as On or OFF. Four different sets of predicted results were obtained by fixing the mode of cooling in the input parameters for the whole year. These predictions show how the data center would react if only one type of cooling was being used for the whole year. This was repeated for all four cooling modes; i.e. ASE on, DEC on, IEC on, and multistage I/DEC on.

5.1 Cold Aisle (CA) Bins

The predicted CA conditions for each type of cooling were separately plotted on the psychrometric chart to come up with new bins. Each plot shows very distinguished patterns for the predicted CA conditions. Figure 5.1 shows the predicted CA conditions when the data center is run using ASE alone. It can be observed that the CA temperature ranges between 63F and 90F whereas the relative humidity ranges from 0% to 95%.

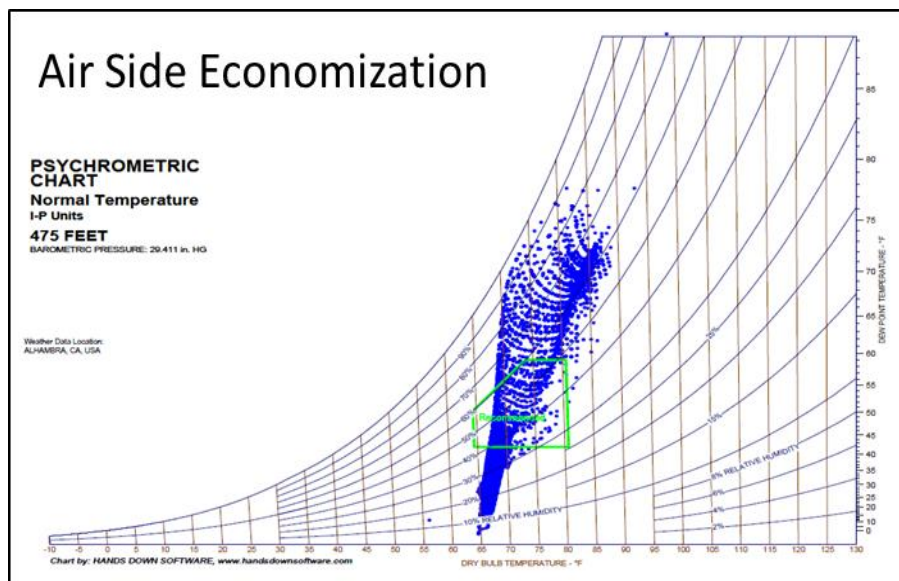


Figure 5.1 Predicted CA conditions for ASE ON

Figure 5.2 represents the same results distributed month-wise on a box and whisker chart. The average CA condition are within the recommended region from November through mid-April and thus ASE can be used

during this period. But the CA temperature and humidity increases during summers from May to October and this shows that ASE is not recommended during these months.

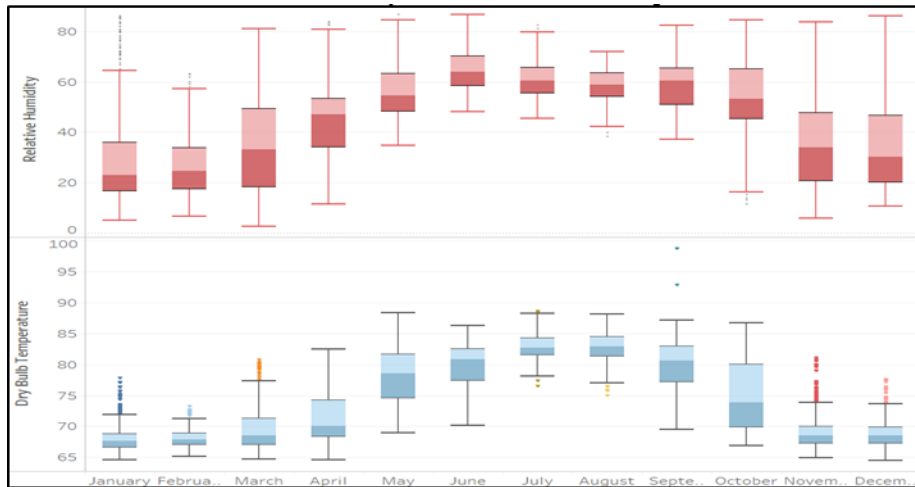


Figure 5.2 Box and Whisker Plot - Predicted CA Conditions ASE ON

Figures 5.3-5.8 represent similar results for DEC ON, IEC ON and IDEC ON respectively. Figure 5.3 shows the plot for DEC ON and it can be observed that the humidity is increased compared to the plot in figure 5.1. This is because adiabatic cooling takes place during DEC and the evaporated water is carried by the air into the CA thus increasing the relative humidity. Similar changes can be seen in the box and whisker for the DEC ON as shown in figure 5.4

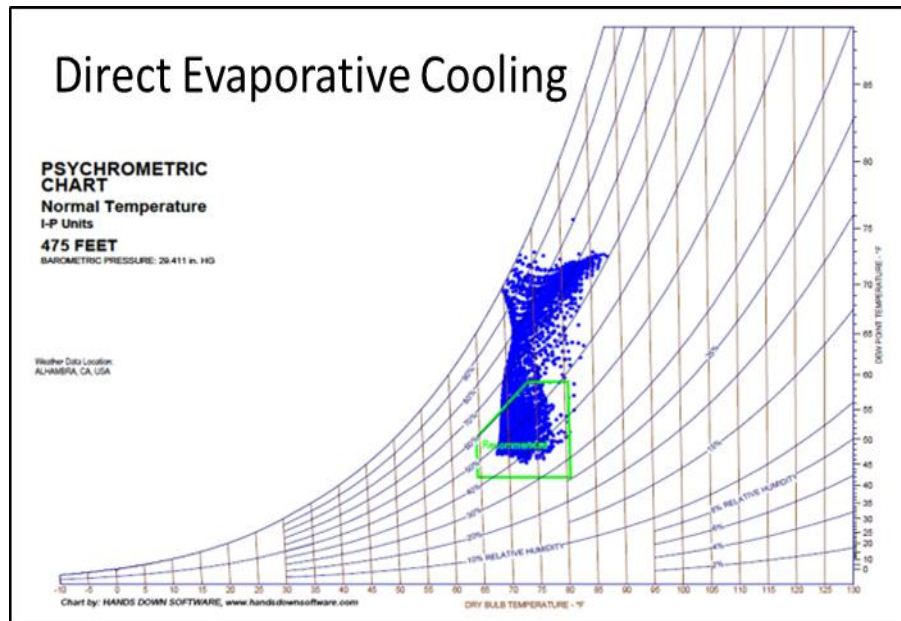


Figure 5.3 Predicted CA conditions for DEC ON

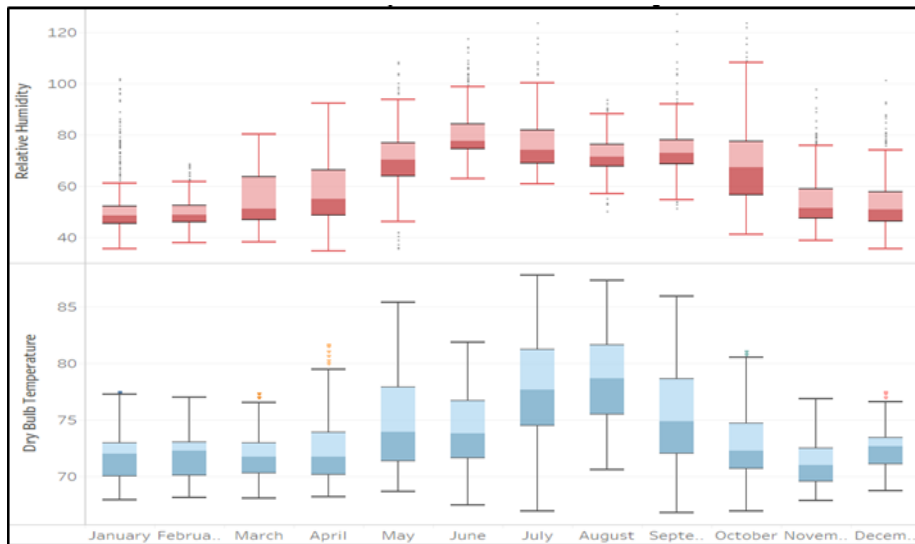


Figure 5.4 Box and Whisker Plot - Predicted CA Conditions DEC ON

Figure 5.5 shows the CA condition plots for IEC ON. During IEC, sensible cooling takes place and hence no major rise in humidity is seen. Thus, we can see that there is a drop in the humidity compared to the DEC results in figure 5.3. Similar changes can be seen in figure 5.6.

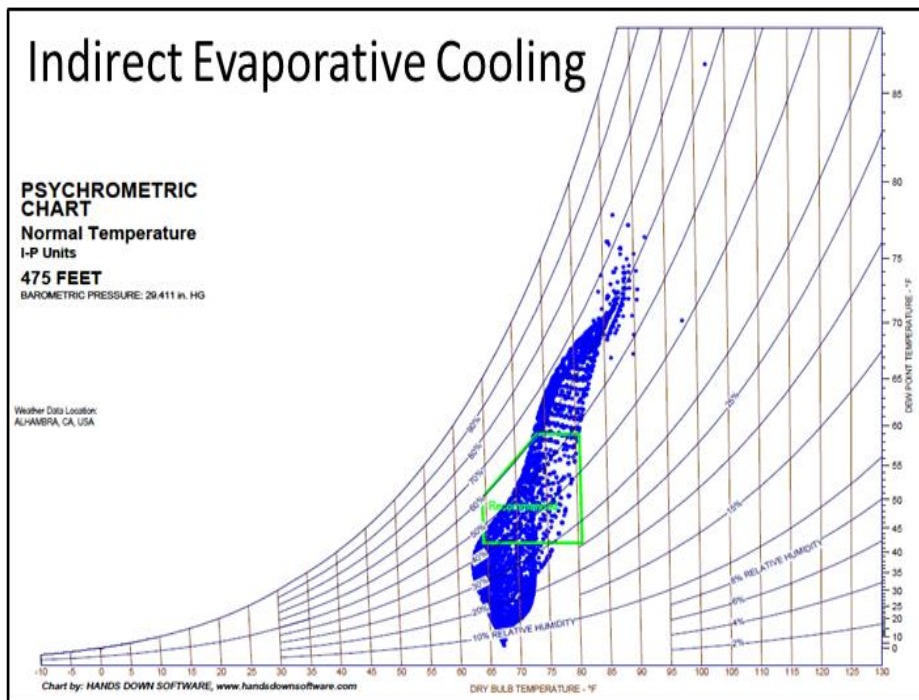


Figure 5.5 Predicted CA conditions for IEC ON



Figure 5.6 Box and Whisker Plot - Predicted CA Conditions IEC ON

Final couple of figures show the results for multistage IDEC ON. In IDEC, air is cooled sensibly by IEC in stage 1 followed by adiabatic cooling using DEC in stage 2. As both the evaporative cooling modes are used in series, a rise in the humidity is observed as shown in figure 5.7. Also, as the cooling effect is doubled, we can see a drop in the higher temperature limit. The temperature for CA can always be maintain between 60F and 81F, but the relative humidity is above 60% for almost half of the plots.

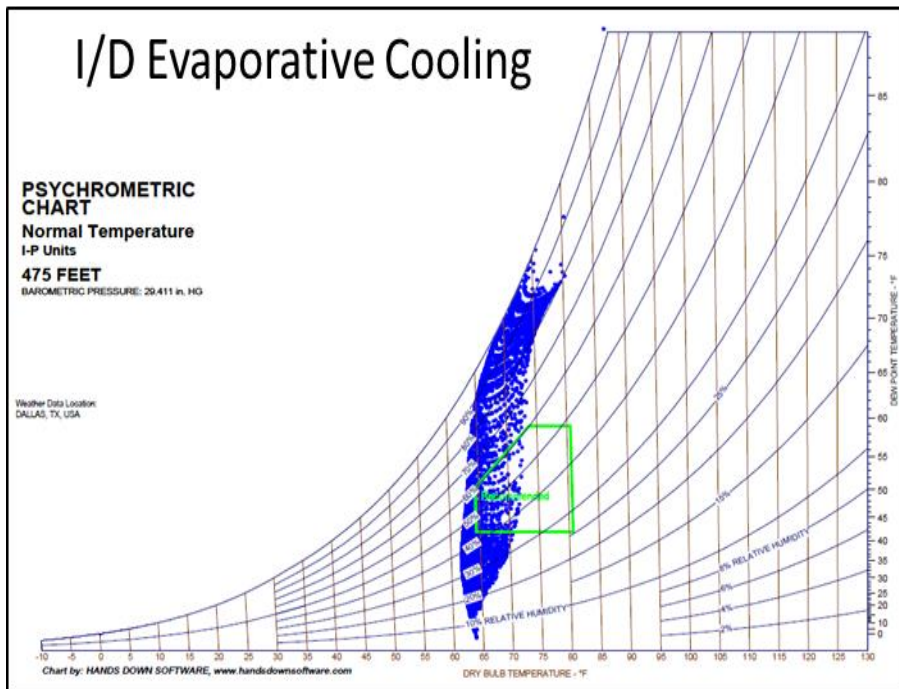


Figure 5.7 Predicted CA conditions for IDEC ON

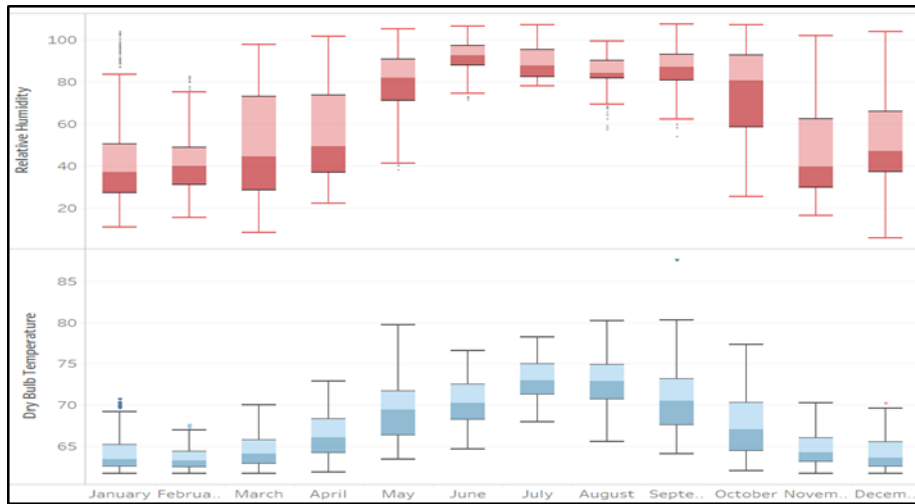


Figure 5.8 Box and Whisker Plot - Predicted CA Conditions IDEC ON

5.2 Outside Air (OA) Bins

The next set of results represent the psychrometric bins for the OA conditions which can be processed to fall into the recommended bin using the different cooling modes separately. Figures 5.9-5.12 represent the OA conditions that can be processed to the recommended zone using ASE, DEC, IEC and IDEC respectively.

Most of the points in figure 5.9 fall below the humidity limit of the recommended envelope. This is because only mixing of the OA air and the return air from hot aisle takes place during ASE operation. Thus, ASE cannot be used for hot and humid condition. Also, no cooling takes place when ASE is being used and thus most points in this bin fall below 85F. ASE can also be used for cold and humid conditions, where supply air is heated on mixing with the hot and dry air from the hot aisle and its relative humidity is decreased.

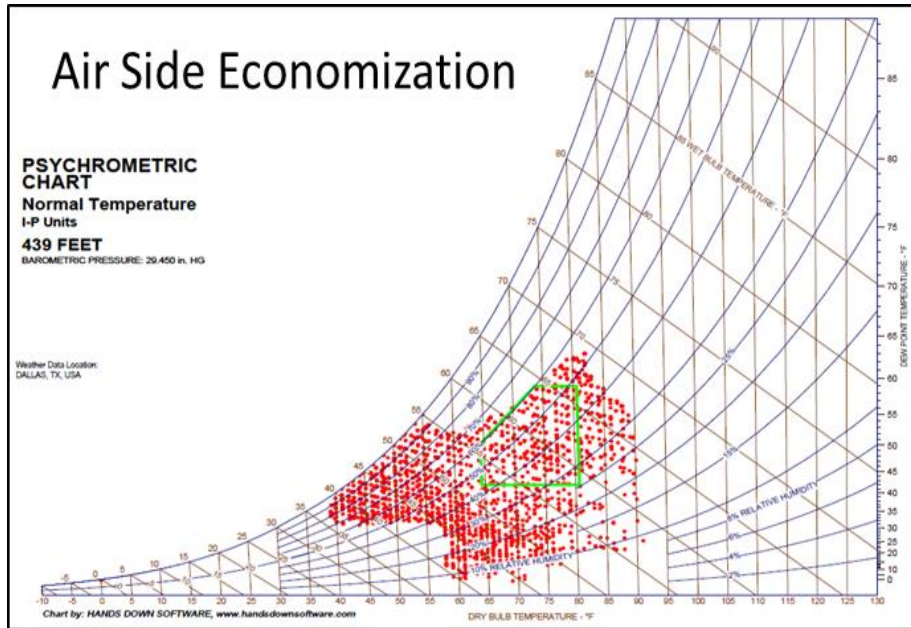


Figure 5.9 OA Bin for Recommended Region Using ASE

Figure 5.10 represents the similar results for DEC ON. Again, as humidity increases during DEC, most of the plots in this bin fall below the 40% relative humidity line. For cold and humid air conditions, the air is first heated and dehumidified using ASE and then DEC is used to cool it to the recommended region.

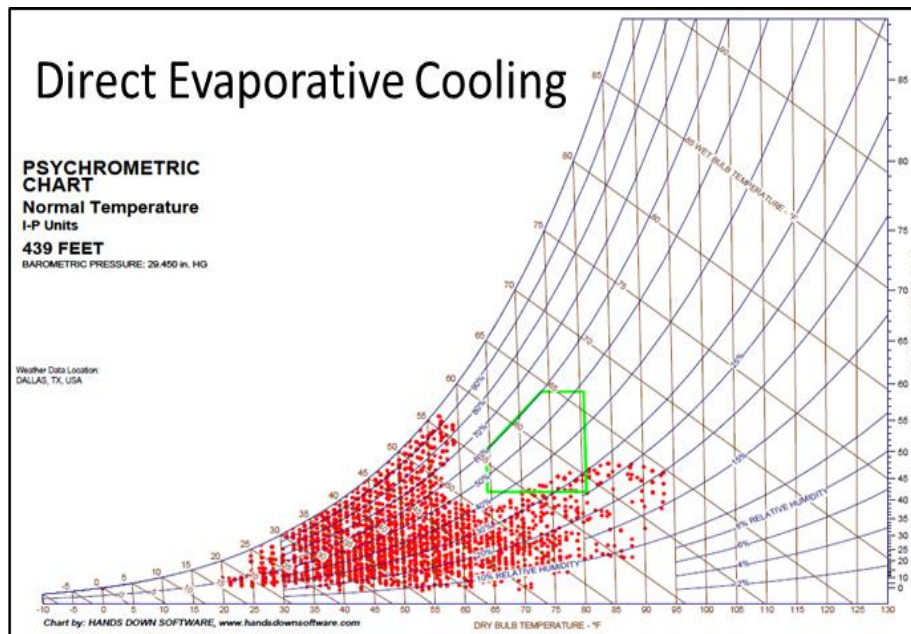


Figure 5.10 OA Bin for Recommended Region Using DEC

The OA plot for IEC ON is shown in figure 5.11. Here as there is no rise in humidity, we can see a wider range of operating bin. The plots below ranges from 30F to 95F and 30% RH to almost 100% RH. Figure 5.12 shows the plots for IDEC ON. IDEC is mostly used when the temperature is very hot and dry. As the recorded OA temperature rarely goes above 100F in Dallas, we can see that only a few data points in this plot.

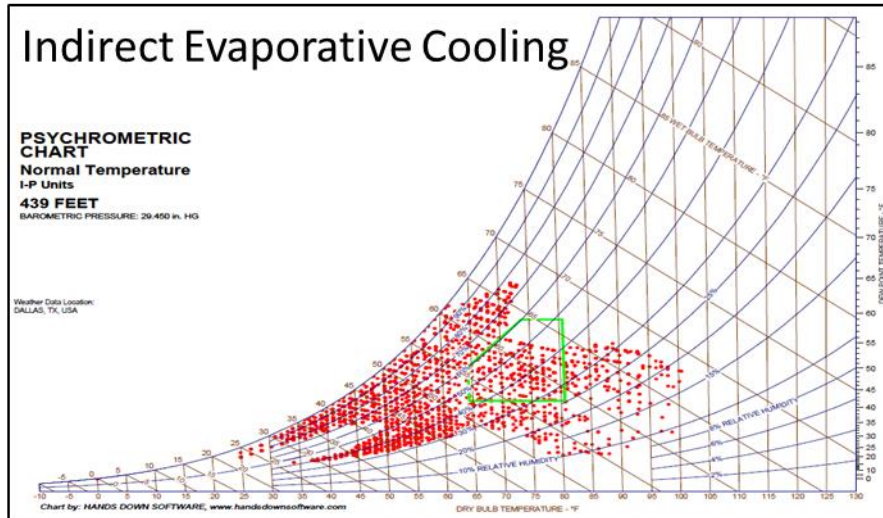


Figure 5.11 OA Bin for Recommended Region Using IEC

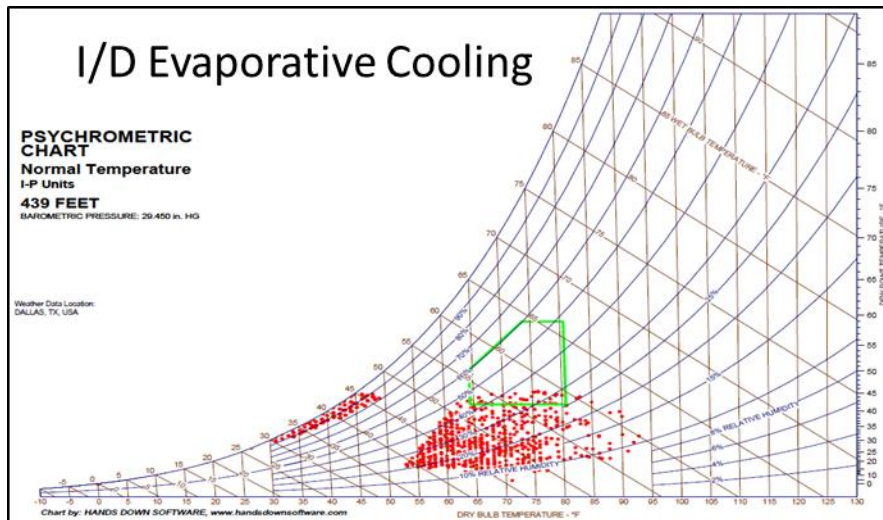


Figure 5.12 OA Bin for Recommended Region Using IDEC

Figures 5.13-5.16 represents similar results but for allowable A1 envelope. These plots represent the outside air which can be conditioned to the A1 allowable envelope using the differently modes of cooling separately. As the bounds of the allowable A1 envelope is much larger than the recommended envelope, we can see bigger weather bins with more plots in each case.

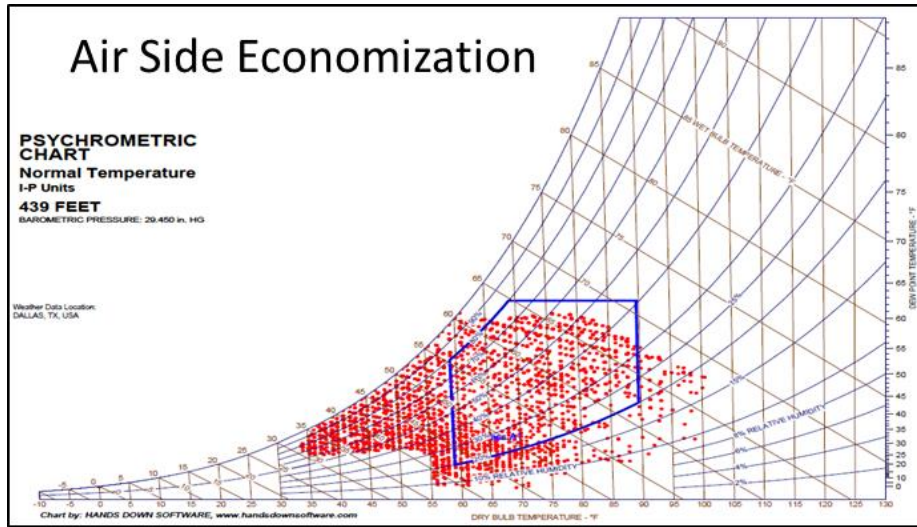


Figure 5.13 OA Bin for A1 Allowable Envelope Using ASE

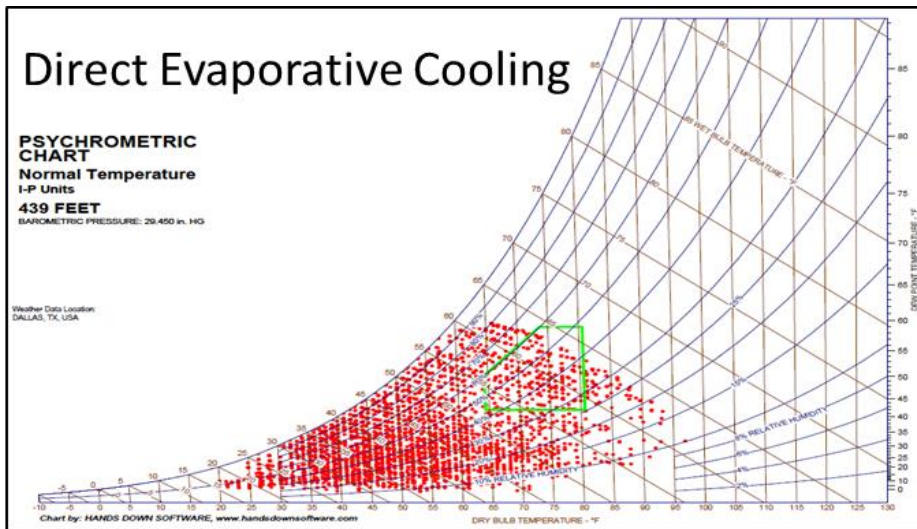


Figure 5.14 OA Bin for A1 Allowable Envelope Using DEC

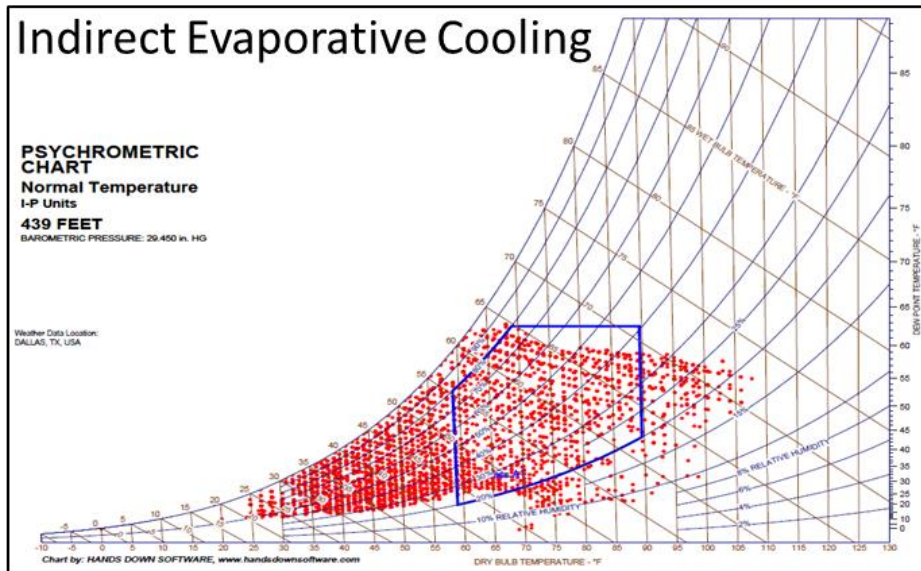


Figure 5.15 OA Bin for A1 Allowable Envelope Using IEC

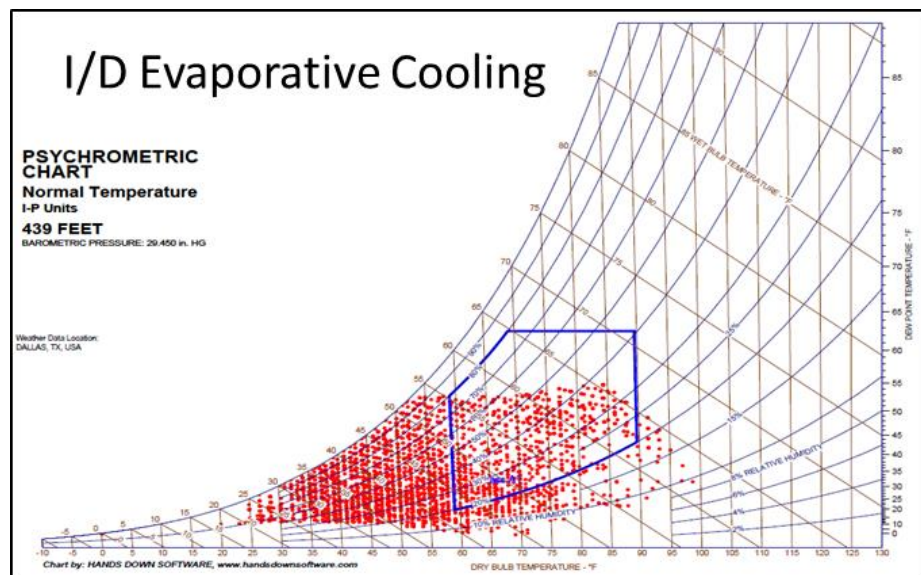


Figure 5.16 OA Bin for A1 Allowable Envelope Using IDEC

5.3 Cooling Mode - Hours of Operation

In this set of results, the data from the weather bin analysis was used to come up with a month-wise usage for each cooling mode. The data from the previous results was re-sorted to predict the hourly utilization of each type of cooling when the cold aisle is to be maintained at either the recommended or A1 allowable conditions. Figures 5.17 – 5.20 show the percentage hourly usage where the blue line represents the results for

within the recommended envelope and the orange line for the allowable A1 envelope. The percentage hourly usage is very high for winter months between November through April but falls drastically during summer months of May, June July, August and September. The predicted results for ASE ON are shown in figure 5.17. The usage is highest in April as the temperature is moderate but decreases as the climate gets warmer in summers. The minimum ASE is seen in July and is almost 0%. Similar trend is observed for all the cooling modes.

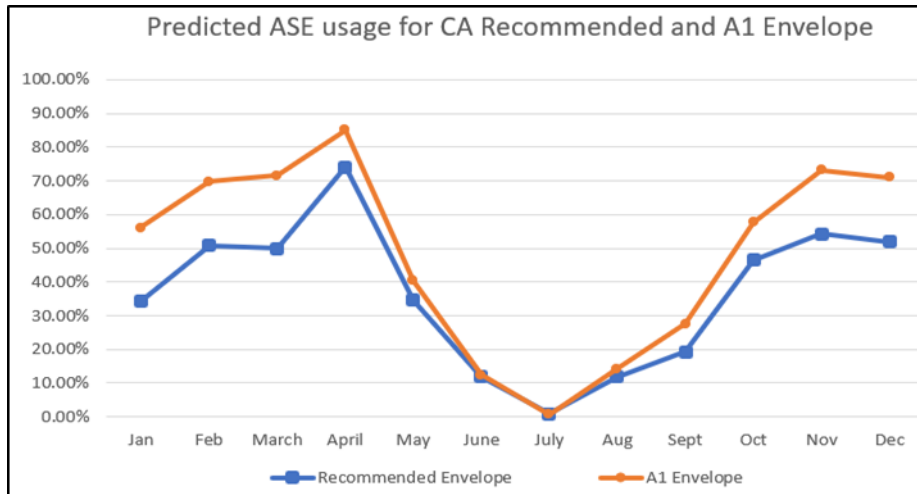


Figure 5.17 Percentage Hourly ASE usage for Recommended and A1 Envelopes

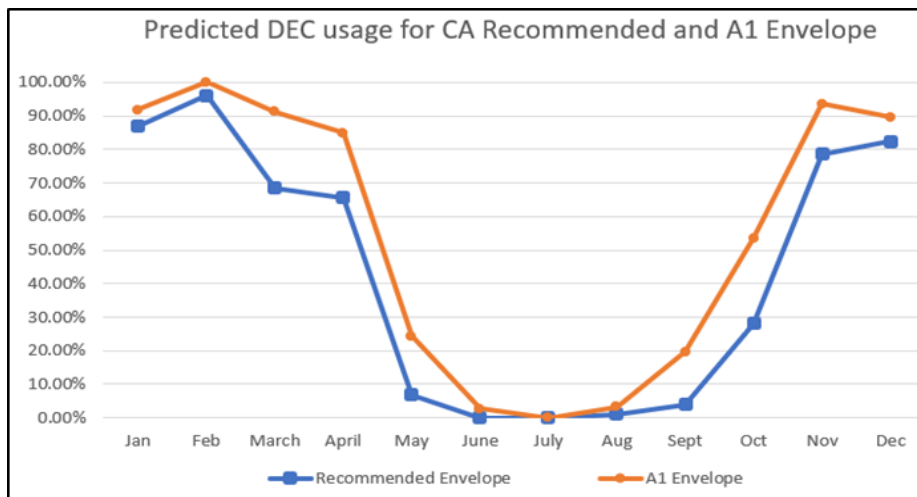


Figure 5.18 Percentage Hourly DEC usage for Recommended and A1 Envelopes

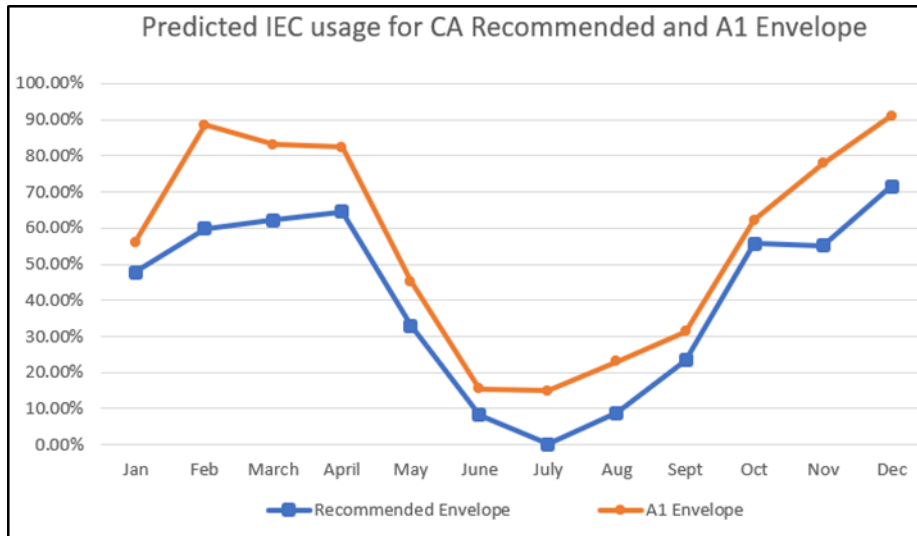


Figure 5.19 Percentage Hourly IEC usage for Recommended and A1 Envelopes

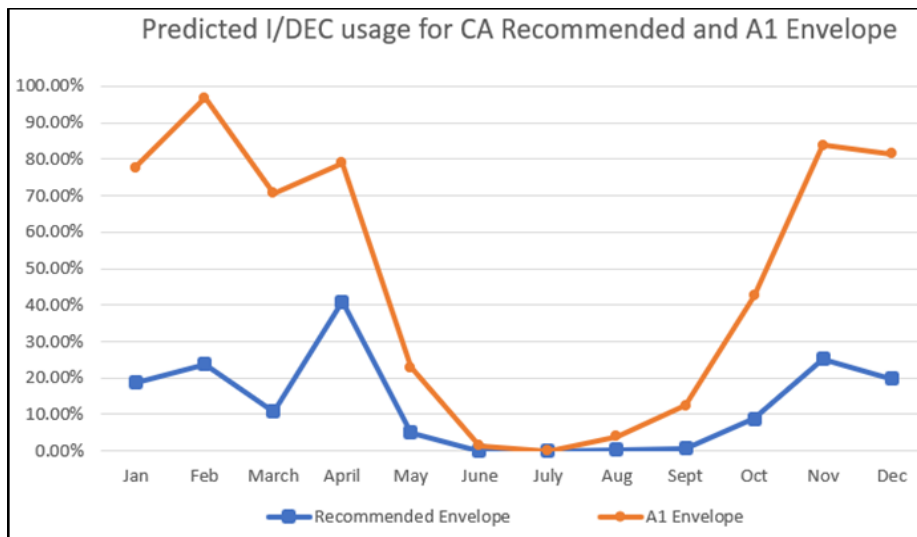


Figure 5.20 Percentage Hourly IDEC usage for Recommended and A1 Envelopes

5.4 Optimum Cooling of Data Centers

Above results give an idea of how a data center would react if only one of the cooling mode was available. These results were then filtered to find the most optimum cooling that can be achieved. The next set of results represent the percentage hourly usage for a combination of all the cooling modes. These results are then compared with the ASHRAE results.

Figure 5.21 shows the comparison between the ASHRAE results on the top and the ANN prediction-based results at the bottom when the cold aisle needs to be maintained within the Recommended envelope. It can be observed that both these results show a similar trend with maximum usage in months of January to April and October to December and minimum in months of June, July and August. Similarly, Figure 5.22 shows the comparison between the ASHRAE results and the ANN prediction-based results when the cold aisle needs to be maintained within the Allowable A1 envelope.

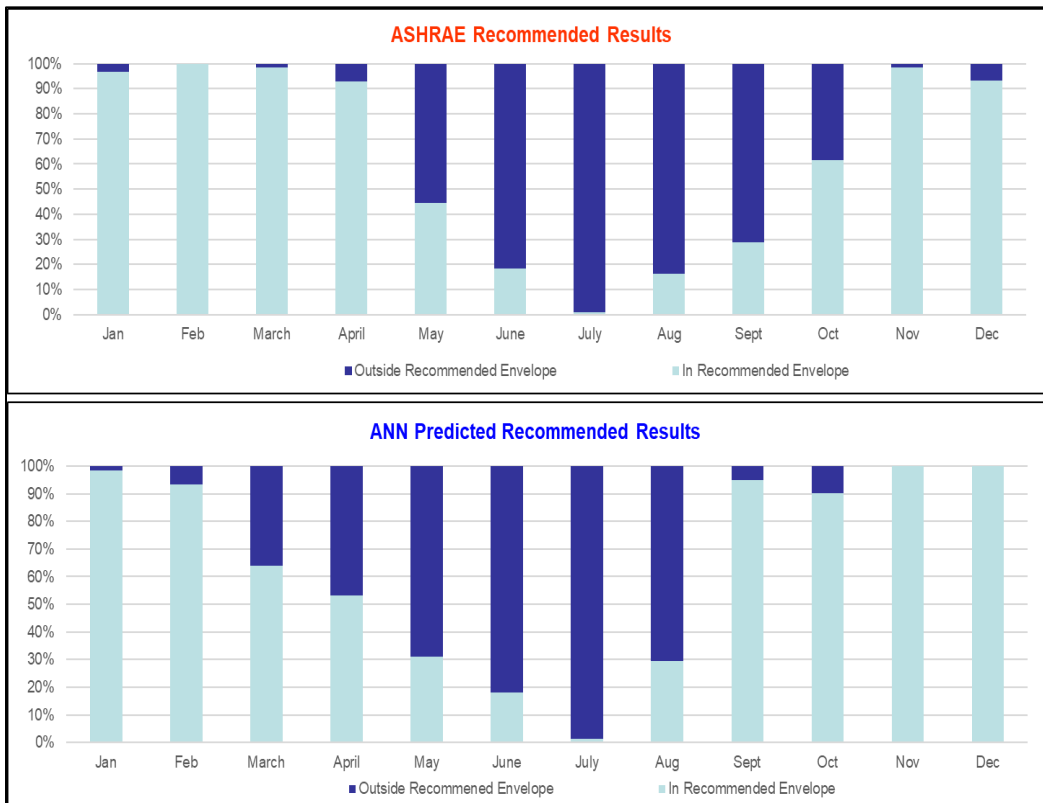


Figure 5.23 Percentage Hours of Operation in Recommended Envelope

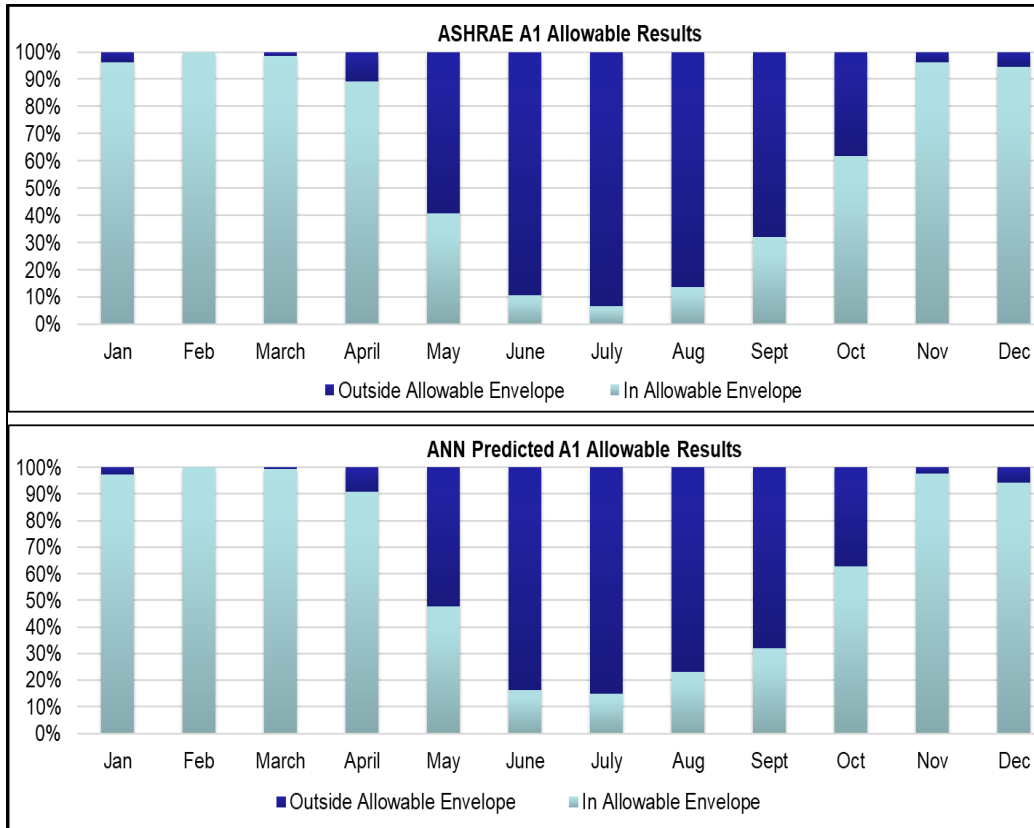


Figure 5.24 Percentage Hours of Operation in Allowable Envelope

5.4 ANN Predicted PUE Vs Actual PUE

The ANN model was also used to predict the PUE for each type of cooling as well as for the optimum working conditions. Table 3 below shows the average ANN predicted PUE values for ASE, DEC, IEC, IDEC and the combined compression less cooling. The last column gives actual PUE value for the Mestex modular data center. The lowest value for the PUE is for ASE as it is a mixing process and the only energy consumed is by the fans and the vents. PUE for DEC is bit higher as energy is utilized to run the pumps to supply water to the cooling pads. The energy consumption for IEC is further increased due to the use of cooling tower and hence shows higher PUE. Maximum energy is utilized by multistage IDEC as both DEC and IEC are used in series.

Table 3 Average PUE

ASE	DEC	IEC	I/DEC	ANN PREDICTED	ACTUAL
1.09	1.15	1.21	1.32	1.20	1.34

Further it can be observed that the ANN based predicted PUE is less than the actual PUE of the modular data center at Mestex. This shows that using the ANN predicted bins for controlling the data center operation can reduce its overall PUE resulting in a 10% energy saving over the year.

5.5 ANN Based Application for Cold Aisle Predictions.

The trained ANN model was used to come up with an application that would predict the cold aisle conditions (temperature and relative humidity) and PUE based on the user inputs. The GUI for the application is shown in figure 5.25. Here, the user inputs include the OA conditions, IT load, delta T and the type of cooling while the application predicts the cold aisle conditions, PUE and bound in which the cold aisle will end up.

OA Temperature	70	CA Temperature	73.55
OA Humidity	65	CA Humidity	63.14
IT Load	15	PUE	1.1
Delta T across Racks	20	Wet Bulb Temperature	61.84
Button Group		Bound	A1 Allowable
<input checked="" type="radio"/> ASE			
<input type="radio"/> DEC			
<input type="radio"/> IEC			
<input type="radio"/> IDEC			
		Calculate	

Figure 5.25 ANN Based Cold Aisle Prediction App

Chapter 6

Summary

The main aim of this research was to decrease the cooling cost of the data centers by increasing the utilization of compression less cooling modes. This could be achieved by defining new control strategies for the cooling modes. The ASHRAE bin-based control strategies currently being implemented do not account for the non-linearities developed in the data center. So, the main objective of this study was to implement ANN to capture the non-linearities developed in the data center and then use this ANN model for predictions.

A static fitting tool ANN model was trained using the real time sensor data from a fully functional data center at Mestex Ltd. This trained model was then used to predict the PUE and the cold aisle conditions in the IT pod. A one-year TMY 3 weather data for Dallas Love Field was used as an input to the ANN model along with IT load, delta T across racks and the cooling mode. Four different set of results were obtained for each type of cooling mode on. These predicted results were then used for weather bin and CA bin analysis.

As the ANN accounts for the non-linearities in the data center, it gives more realistic predictions. The results show different and unique patterns of CA and weather bins for each cooling mode. This ANN predicted bins are very different from the ASHRAE defined bins/regions. ANN results show an increase in the number of hours of compression less cooling as compared to the ASHRAE results. Also, the average PUE calculated using ANN predicted bins is much lower than the actual average PUE of the data center over a year. The results show about 10% decrease in the cooling power consumption.

The next step in this study would be to set up new control strategies based on these ANN bins for an actual data center and compare those results with the predictions. Also, a dynamic ANN model can be trained for step ahead predictions which can be used for proactive controls of data center operations.

References

- [1] ASHRAE, ASHRAE. 2005. ASHRAE Handbook-Fundamentals. Atlanta: American Society of Heating Refrigeration and Air Conditioning Engineers, Inc., 2005.
- [2] A. Guhe, "CONTROL STRATEGIES FOR AIR-SIDE ECONOMIZATION, DIRECT AND INDIRECT EVAPORATIVE COOLING AND ARTIFICIAL NEURAL NETWORKS APPLICATIONS FOR ENERGY EFFICIENT DATA CENTERS," Arlington, Tx, 2016.
- [3] [Online]. Available: https://en.wikipedia.org/wiki/Power_usage_effectiveness.
- [4] Jimil M Shah, "Effects of mineral oil immersion cooling on IT equipment reliability and reliability enhancements to data center operations," in *ITHERM*, Las Vegas, NV, 2016.
- [5] A. Siddarth, "EXPERIMENTAL STUDY ON EFFECTS OF SEGREGATED COOLING PROVISIONING ON THERMAL PERFORMANCE OF INFORMATION TECHNOLOGY SERVERS IN AIR COOLED DATA CENTERS," Arlington, TX, 2015.
- [6] Ian Metzger, "Psychrometric Bin Analysis for Alternative," in *ASHRAE*, 2011.
- [7] J. Gao, "Machine Learning Applications for Data Center Optimization," Google.
- [8] Zhihang Song, "Airflow and temperature distribution optimization in data centers using artificial neural networks," *International Journal of Heat and Mass Transfer*, 2013.
- [9] Zhimin Du, "Fault diagnosis for temperature, flow rate and pressure sensors in VAV systems using wavelet neural network," *Applied Energy*, vol. 86, no. 9, pp. 1624-1631, 2009.
- [10] Ryohei Yokoyama, "Prediction of energy demands using neural network with model identification by global optimization," *Energy Conversion and Management*, vol. 50, no. 2, pp. 319-327, 2009.
- [11] Lizhe Wang, "Task scheduling with ANN-based temperature prediction in a data center: a simulation-based study," *Engineering with Computers*, vol. 27, no. 4, pp. 381-391, 2011.
- [12] S. Shrivastava, "Data center cooling prediction using an artificial neural network," in *IPACK*, Vancouver, Canada, 2007.
- [13] ASHRAE, ASHRAE TC 9.9. 2008. ASHRAE Environmental Guidelines for Datacom Equipment: Expanding the Recommended Environmental Envelope. Atlanta: American Society of Heating Refrigeration and Air Conditioning Engineers, Inc..
- [14] [Online]. Available: <https://www.quora.com>.

Appendix A – MATLAB Code for ANN Modeling

```
clear all;

clc;

% Solve an Input-Output Fitting problem with a Neural Network
% Script generated by Neural Fitting app
% Created 22-Feb-2018 19:22:48
%
% This script assumes these variables are defined:
%
% hourlyInput - input data.
% hourlyOutput - target data.
data=load('train data deltaT.csv');
x = data(:,1:7);
t = data(:,8:10);

% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.

% Create a Fitting Network
hiddenLayerSize = 20;
net = fitnet(hiddenLayerSize,trainFcn);

% Choose Input and Output Pre/Post-Processing Functions
% For a list of all processing functions type: help nnprocess
net.input.processFcns = {'removeconstantrows','mapminmax'};
net.output.processFcns = {'removeconstantrows','mapminmax'};
```

```

% Setup Division of Data for Training, Validation, Testing
% For a list of all data division functions type: help nndivide
net.divideFcn = 'dividerand'; % Divide data randomly
net.divideMode = 'sample'; % Divide up every sample
net.divideParam.trainRatio = 85/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 0/100;

% Choose a Performance Function
% For a list of all performance functions type: help nnperformance
net.performFcn = 'mse'; % Mean Squared Error

% Choose Plot Functions
% For a list of all plot functions type: help nnplot
net.plotFcns = {'plotperform','plottrainstate','ploterrhist', ...
    'plotregression', 'plotfit'};

% Train the Network
[net,tr] = train(net,x,t);

% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)

% Recalculate Training, Validation and Test Performance
trainTargets = t .* tr.trainMask{1};
valTargets = t .* tr.valMask{1};
testTargets = t .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,y)

```

```

valPerformance = perform(net,valTargets,y)
testPerformance = perform(net,testTargets,y)

% View the Network
view(net)

% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
% figure, plotfit(net,x,t)

% Deployment
% Change the (false) values to (true) to enable the following code blocks.
% See the help for each generation function for more information.
if (false)
    % Generate MATLAB function for neural network for application
    % deployment in MATLAB scripts or with MATLAB Compiler and Builder
    % tools, or simply to examine the calculations your trained neural
    % network performs.
    genFunction(net,'myNeuralNetworkFunction');
    y = myNeuralNetworkFunction(x);
end
if (false)
    % Generate a matrix-only MATLAB function for neural network code
    % generation with MATLAB Coder tools.
    genFunction(net,'myNeuralNetworkFunction','MatrixOnly','yes');
    y = myNeuralNetworkFunction(x);
end

```

```

if (false)
    % Generate a Simulink diagram for simulation or deployment with.
    % Simulink Coder tools.
    gensim(net);
end

% temp=1;
% testa=(trainTargets);
% for i=1:30
%     if testa(i)>0
%         c(temp)=testa(i);
%         d(temp)=y(i);
%         temp=temp+1;
%     end
% end
%
%
% figure
% plot(y,'r--')
% hold on;
% plot(testa,'b-')
% errorr=(y-testa)./(max(y)-min(y));
% figure
% plot(errorr)
% avgererrorval=mean(abs(c-d)./(max(testa)-min(testa)))

```

Appendix B – MATLAB Code for ANN Based Prediction App

% Button button pushed function

```
function Calculate(app)
    T=[];
    T(1,6)=app.NumericEditField.Value;
    T(1,5)=app.NumericEditField2.Value;
    T(1,4)=app.NumericEditField3.Value;
    T(1,7)=app.NumericEditField4.Value;
    WBT=(-5.806+0.672*T(1,6)-
0.006.*T(1,6).*T(1,6)+(0.061+0.004.*T(1,6)+0.000099.*T(1,6).*T(1,6)).*T(1,5)+(-0.000033-
0.000005.*T(1,6)-0.0000001.*T(1,6).*T(1,6)).*T(1,5).*T(1,5))
    if app.RadioButton.Value==true
        T(1,1:3)=[0,0,1];
    elseif app.RadioButton2.Value==true
        T(1,1:3)=[0,1,0];
    elseif app.RadioButton3.Value==true
        T(1,1:3)=[1,0,0];
    elseif app.RadioButton4.Value==true
        T(1,1:3)=[1,1,0];
    end
    inputdata=T;
    inputdata=inputdata';
    load net_deltaT;
    outputdata=net_deltaT(inputdata)'
    if outputdata(1,1)<1
        outputdata(1,1)=outputdata(1,1)+0.1;
    else
        outputdata(1,1)=outputdata(1,1);
    end
end
```

```

app.EditField.Value='Out Of Bounds';
app.NumericEditField5.Value=outputdata(1,3);
app.NumericEditField6.Value=outputdata(1,2);
app.NumericEditField7.Value=outputdata(1,1);
app.NumericEditField8.Value=WBT;
flag=true;
if app.NumericEditField5.Value<81 && app.NumericEditField5.Value>64
    if app.NumericEditField6.Value<60 && app.NumericEditField6.Value>25
        app.EditField.Value='Recommended';
        flag=false;
    end
end
if app.NumericEditField5.Value>59 && app.NumericEditField5.Value<90
    if 20<app.NumericEditField6.Value && app.NumericEditField6.Value<80
        if (WBT<62.6 && flag)
            app.EditField.Value='A1 Allowable';
        end
    end
end
end
end

```

Biographical Information

Abhishek Uday Walekar received his Bachelor's degree in Mechanical Engineering from Maharashtra Institute of Technology, University of Pune, India in June 2015. He received his Master's degree in Mechanical Engineering from University of Texas at Arlington in December 2018. His research focused on optimization of data center cooling systems using artificial neural networks. Abhishek served as a research assistant in Electronics, MEMS and Nanoelectronics Packaging Center and worked on various industry-funded projects as a graduate student researcher in NSF funded Industry University Cooperative Research Center called Center of Energy-Smart Electronic Systems (ES2).