The Time Element in Chiller Fault Detection and Diagnosis

by

Paul Lauren Riemer

A thesis submitted in partial fulfillment of

the requirements for the degree of

Master of Science

(Mechanical Engineering)

at the

UNIVERSITY OF WISCONSIN – MADISON

2001

Abstract

Vapor compression chillers are the central component of most large commercial air conditioning systems. Process supervision including fault detection and diagnosis (FDD) has been proposed as a means to identify unacceptable performance of chillers. FDD methodologies are often based on the comparison of the actual performance to some expected performance. An installation provided the case study data presented in this thesis which revealed a periodic variation in the chilled water flow rate and several other chiller monitored quantities. Time series analysis techniques were applied and shown to be capable of quantifying the time variation, resulting in increased fault detection and diagnosis capabilities.

Acknowledgements

First and foremost, thanks to my wife Brook. While she may not appreciate the content of this work, she has understood its importance to me and supported me unconditionally.

Secondly, thanks to the members of the Solar Energy Laboratory. My project advisors Professor John Mitchell and Professor Bill Beckman provided valuable overall guidance while allowing me to follow the data and make this work my own. Ian McIntosh assembled the data set used in this work and laid the groundwork for chiller fault detection and diagnosis in his Ph.D. Professor Sandy Klein suggested considering autocorrelation of the data, providing the real spark for this work. Professor Doug Reindl shared valuable insight into the real world of chiller installations.

Finally, thanks to my graduate student peers, who have provided hours upon hours of camaraderie, technical knowledge, and humor during the past 16 months.

Table of Contents

Summary of Figures

Summary of Tables

Nomenclature

a – autoregressive coefficient

ß – moving average coefficient

e – heat exchanger effectiveness

? – efficiency

AR – autoregressive

ARIMA – autoregressive integrated moving average

C_p – specific heat

Evap – evaporator

Chw – chilled water

Cw – condenser water

Cond – condenser

F – degrees Fahrenheit, units of temperature measurements

Gpm – gallons per minute, units of volumetric flow rate measurements

GPMCHW – chilled water volumetric flow rate

GPMCW – condenser water volumetric flow rate

h – enthalpy

isen – isentropic

k – lag

kW – kilowatts, units of power measurements

lbm – pounds mass

m – mass flow rate

MA – moving average

MMBtu – 10^6 British Thermal Units of Energy

Ntu – number of transfer units, a heat exchanger performance quantity

P – chiller electrical power draw

psia – pounds per square inch absolute

Q – heat transfer

s - entropy

t – time

T2 – refrigerant temperature exiting the compressor

TCHWR – chilled water return temperature

TCHWS – chilled water supply temperature

TCOND – refrigerant saturation temperature in the condenser

TCWR – condenser water return temperature

TCWS – condenser water supply temperature

TEVAP – refrigerant saturation temperature in the evaporator

UA – conductance area product

x_t – sample observation at time t

1. Introduction

1.1. Motivation

Vapor compression chillers serve as the workhorse for the air conditioning systems of most commercial applications. Air conditioning systems are one of the major components in a complex configuration of structures, equipment, control devices, and people all working together to maintain a suitable indoor environment for the occupants and their activities.

For all applications, the space conditioning systems represent a huge financial investment to the building industry simply in the first costs of system design, equipment purchase, and installation. For commercial applications, there are possible additional costs if the system fails to maintain the desired conditions and leads to reductions in productivity.

Vapor compression equipment requires energy to operate. For commercial applications, this energy is generally purchased which raises the financial stakes. While the monetary costs of energy are significant and increasing, energy use also has environmental costs. The majority of electrical energy in the world is produced from fossil fuels that must be harvested from the Earth and which release numerous by-products during combustion. These by-products have the capability to significantly alter atmospheric chemistry, but their full effects on the environment and future quality of life is not known. Due to these energy costs, even a slight improvement in the efficiency of

vapor compression equipment is a worthwhile achievement and similarly, any significant performance degradation of these systems is cause for concern.

Additionally, some of the refrigerants, or working fluids, currently used in vapor compression equipment have the potential for a more direct effect on the environment because upon release they ascend to the upper atmosphere and break down ozone. Ozone at lower levels is dangerous but at higher altitudes it is key for controlling the temperature of the earth by preventing the transmission of ultraviolet radiation.

The number of vapor compression systems in existence, the number of hours these systems typically operate, and all of their related costs draw a great deal of attention to their design, operation and maintenance. One specific ongoing response to these costs is the development of supervision techniques to detect when the behavior of the equipment deviates from some acceptable range of operating conditions.

1.2. Purpose

While the burden to design and produce a robust, efficient chiller rests on the manufacturer, the preceding section establishes the need for the means to identify when a chiller develops an operating fault. Such a fault identification system should be able to not simply detect a fault but also diagnose it or determine its nature. This system would preferably be on-line, either within the chiller control system or as an added component. The purpose of this work is to enhance fault detection and diagnosis (FDD) methodologies that have resulted from prior research efforts by considering the time variations of the data.

1.3. Scope and Organization

To examine the issue of detecting and diagnosing faults in vapor compression chillers, this thesis begins with an overview of the basic chiller components and the vapor compression refrigeration cycle. Next, existing FDD methodologies are examined along with how they might be advanced. The fundamentals of time series analysis are introduced as well as their potential for improving FDD.

Field data from a case study exhibits a periodic nature and demonstrates the need to address time variant behavior in the chiller. While the field data reveals odd behavior, optimization and fault remediation of this particular equipment is not an objective of this thesis. The benefits of including time series analysis in FDD methodologies are revealed. Finally, conclusions of this work are stated and recommendations for future work are laid forth.

2. Vapor Compression Chillers

2.1. Overview

The primary purpose of an air conditioning system is to transfer energy, in the form of heat, out of a conditioned space. To do so, an amount of air that is cooler than the current space temperature is supplied to the space while an equal amount of warm air is removed from the space. In many large air-conditioning systems, the supply air passes through an air handling unit where it is cooled using a coil heat exchanger containing cold or chilled water. This chilled water loops between the coil and the chiller. The chiller then rejects energy to a condenser water loop, which is connected to the cooling towers. The cooling towers reject energy to the ambient. To transfer the energy from the chilled water loop to the condenser water loop, a vapor compression chiller consisting of four main components and a refrigerant is employed. The refrigerant moves through the components consecutively (evaporator, compressor, condenser, and expansion device) in a closed loop.

Figure 2-1 shows a simplified schematic of the chiller, the chilled and condenser water loops, and the conditioned and ambient air streams. The dotted line box denotes the physical boundary of the chiller. The numbers (1-4) refer to the four refrigerant states, one between each of the components. For both of the water flow loops, the supply and return sides are identified by the temperature measurement locations. The chilled water supply is the water leaving the evaporator and being sent to the air handling units. The

condenser water supply is the water leaving the cooling towers and being sent to the condenser.

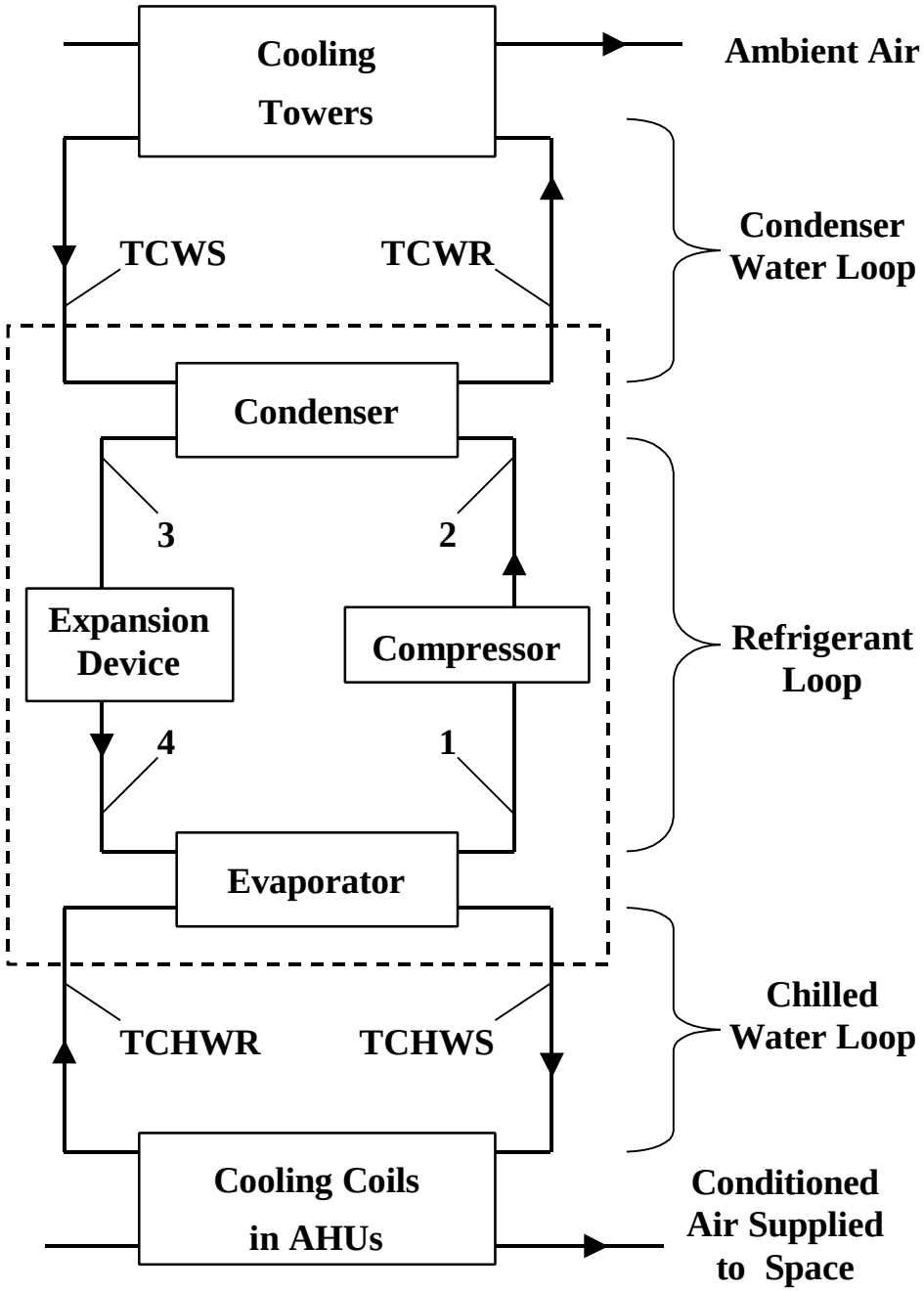

Figure 2-1: Simplified Chiller Schematic

A common method to represent the thermodynamics of the vapor compression cycle occurring in a chiller is a pressure-enthalpy (P-H) diagram, shown in Figure 2-2 including the labels of the numbered states in Figure 2-1.

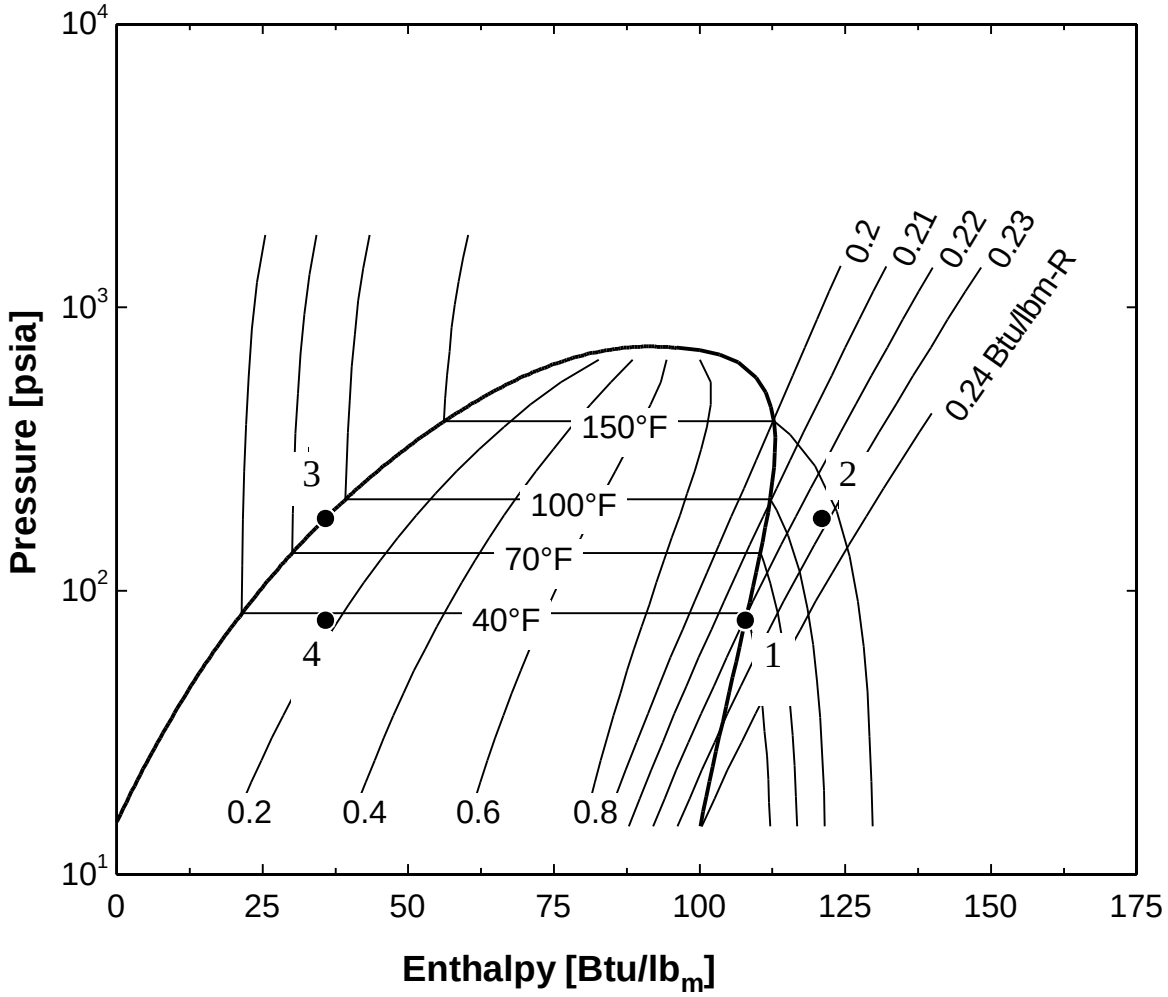

Figure 2-2: Pressure - Enthalpy Plot of R22

2.2. Physical Definitions and Relationships

The cooling provided is equal to the amount of energy transfer in the evaporator (Q_{EVAP}). For the entire chiller, a coefficient of performance (COP) is defined as the cooling provided divided by the electrical power draw (P).

$$COP = \frac{Q_{EVAP}}{P} \tag{2.1}$$

The remainder of this chapter describes each of the four chiller components and the physical relationships that exist. The descriptions are for the chilled water systems found in large facilities such as studied in this thesis.

2.2.1. Evaporator

The evaporator is a shell and tube heat exchanger with the chilled water passing through the tubes. The tubes are submerged in a pool of low-pressure refrigerant. As the water returns from the coil in the air handler, it is warm and causes the refrigerant to boil or evaporate. The refrigerant enters the evaporator at state 4 and leaves at state 1. The direction of positive heat transfer is from water to refrigerant. The amount of heat transfer can be determined from an energy balance of the waterside mass flow rate, specific heat, and temperature change, and refrigerant mass flow rate and enthalpy change. The energy balance is

$$Q_{EVAP} = m_{chw} C_{p,water} (TCHWR - TCHWS) = m_{refrigerant} (h_1 - h_4) \tag{2.2}$$

Heat exchanger performance can be quantified by defining a few parameters. The chilled water temperature change(? T_{CHW}) is the difference between the chilled water return and supply temperatures, denoted as TCHWR and TCHWS respectively. The approach of a heat exchanger is defined as the smallest temperature difference between the two fluids. In the case of a chiller, the evaporator approach ($APPR_{EVAP}$) is the chilled water supply temperature minus the temperature of the refrigerant in the evaporator

(TEVAP). The water temperature differences and the approaches for both the evaporator and condenser are shown on Figure 2-3.

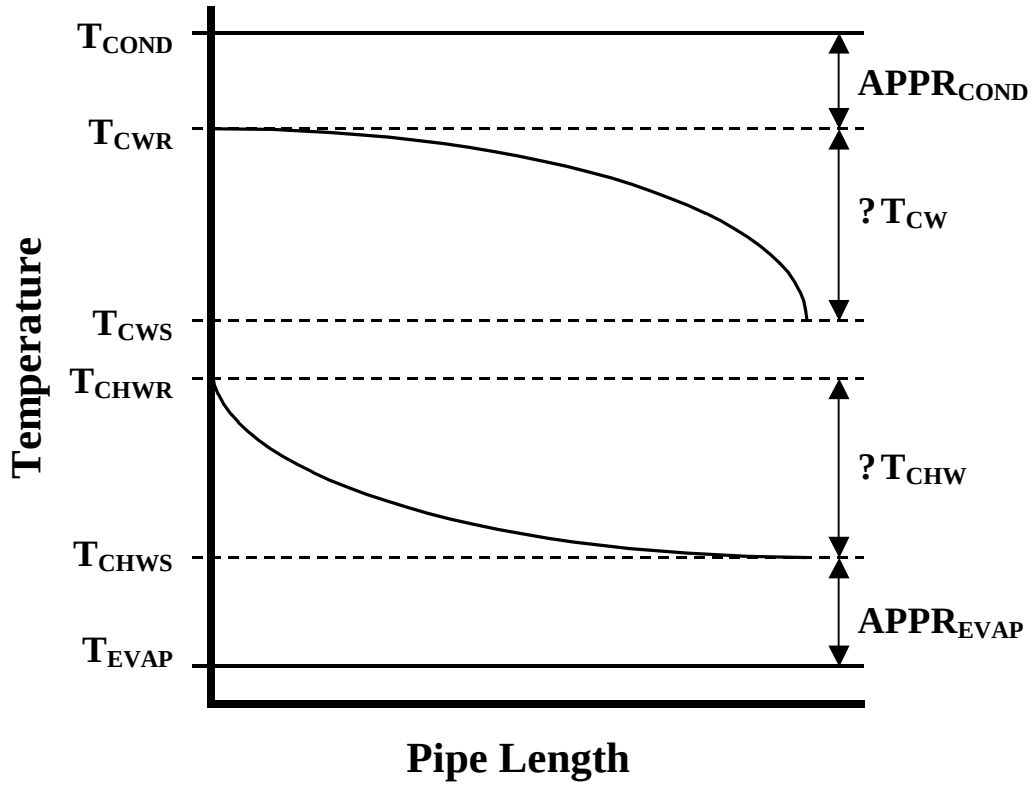

Figure 2-3: Heat Exchanger Temperature Diagram

An effectiveness of the evaporator (ε_{EVAP}) is defined as the actual energy transferred divided by the maximum possible energy transfer. The maximum possible energy transfer would occur if the water entering the evaporator was cooled to the refrigerant temperature.

$$e_{EVAP} = \frac{(TCHWR - TCHWS)}{(TCHWR - TEVAP)} \tag{2.3}$$

Another heat exchanger performance quantity, the Number of Transfer Units (Ntu), is defined as the conductance-area product (UA) divided by the water capacitance, which is the product of the mass flow rate (m) and specific heat (C_p).

$$Ntu_{EVAP} = \frac{UA_{EVAP}}{m_{chw} C_{p,water}} \tag{2.4}$$

The effectiveness and the Ntu are related depending on the type of heat exchanger. Equation 2.5 gives the equation for a shell and tube heat exchanger where one fluid undergoes a phase change and is considered to have infinite capacitance.

$$e_{EVAP} = 1 - e^{-Ntu_{EVAP}} \tag{2.5}$$

2.2.2. Compressor

The principle component of the chiller is the compressor, which is generally a centrifugal, screw, scroll, or reciprocating mechanism driven by an electric motor. The compressor lifts the refrigerant vapor leaving the evaporator, state 1, to a higher pressure entering the condenser, state 2. An energy balance relates the required work input (W) to the compressor to the product of the refrigerant mass flow rate through the compressor and the refrigerant enthalpy (h) change.

$$W = m_{refrigerant}(h_2 - h_1) \tag{2.6}$$

This real process can be compared to an ideal process by an isentropic efficiency ($?_{ISENTROPIC}$). In equation 2.7, $h_{isen,2}$ denotes the enthalpy of a fictional state that is at the entropy of the refrigerant entering the compressor (s_1), and the saturation pressure of the condenser (P_2).

$$h_{ISENTROPIC} = \frac{(h_{isen,2} - h_1)}{(h_2 - h_1)} \tag{2.7}$$

The motor turning the compressor is also a real device and cannot convert all of the electrical power into mechanical power. The efficiency of the motor (?$_{MOTOR}$) is defined as the work input to the fluid divided by the electrical power (P).

$$h_{MOTOR} = \frac{W}{P} \tag{2.8}$$

2.2.3. Condenser

The energy balance for the condenser is quite similar to that for the evaporator. The difference is that the direction of heat transfer in the condenser is from the refrigerant to the water. The energy balance is

$$Q_{COND} = m_{cw} C_{p,water} (TCWR - TCWS) = m_{refrigerant} (h_2 - h_3) \tag{2.9}$$

The condenser water temperature change (? T_{CW}) is the difference between the condenser water return and supply temperature, denoted as TCWR and TCWS respectively. The condenser approach ($APPR_{COND}$) is the difference between the refrigerant temperature and the condenser water return temperature. The water temperature difference and the approach of the condenser were represented graphically along with the evaporator values in Figure 2-3. The water enters cooler but heats up as the high-pressure refrigerant vapor coming from the compressor condenses. Like the evaporator, the effectiveness and Ntu of the condenser can be defined and related to each other.

$$e_{COND} = \frac{(TCWR - TCWS)}{(TCOND - TCWS)} \tag{2.10}$$

$$Ntu_{COND} = \frac{UA_{COND}}{m_{cw} C_{p,water}} \tag{2.11}$$

$$e_{COND} = 1 - e^{-Ntu_{COND}} \tag{2.12}$$

2.2.4. Expansion Device

The expansion device is often a valve that drops the refrigerant pressure from the condenser to the evaporator, whereas the compressor increased the pressure. High-pressure saturated liquid refrigerant leaving the condenser passes through the expansion device and becomes a low-pressure mixture of liquid and vapor, which then enters the evaporator. An energy balance for the expansion device shows that the enthalpy is constant for the process.

$$h_3 = h_4 \tag{2.13}$$

3. Fault Detection and Diagnosis Methodologies

3.1. Background

Isermann (1984) summarized some of the early concepts of the supervision of technical processes including defining a fault as " a non-permitted deviation of a characteristic property which leads to the inability to fulfill the intended purpose." Isermann also mapped a progression for supervising processes and equipment consisting of:

1. Monitoring
2. Fault Detection
3. Fault Diagnosis
4. Fault Evaluation
5. Decision
6. Action

Several other more recent works have considered all or some of these process supervision components and applied them to different HVAC systems or pieces of equipment. The remainder of this section outlines some FDD research and the components various researchers have considered. Rossi and Braun (1997) developed a detection and diagnosis methodology for vapor compression air-conditioners with a direct expansion coil. Breuker and Braun (1998b) applied that method to a rooftop air-conditioner. Breuker and Braun (1998a) considered the evaluation step by looking at the common faults in rooftop air conditioners and their impacts. Fasolo and Seborg (1995)

examined monitoring and fault detection for an HVAC controls system. Shiozaki and Miyasaka (1999) and Yoshidea and Kumar (1999) developed diagnostic tools for VAV systems. Salsbury (1999) proposed a controller with detection capabilities for VAV systems. Pan, Zheng, and Nakahara (1999) studied fault detection and diagnosis on water thermal storage HVAC systems. Li et al (1997) proposed a method to combine the detection and diagnosis stages for hot water radiator heating system in variable occupancy buildings.

Stylianou and Nikanpour (1996) presented a methodology to monitor the performance of a reciprocating chiller including fault detection and diagnosis during the time the chiller is off, in start-up, or in steady state operation. McIntosh (1999) developed a fault detection and diagnosis process for chillers and other HVAC subsystems using actual data from four identical centrifugal chillers in a single plant. Chapter 5 will consider specific aspects of McIntosh's proposed FDD methodologies and the comparative study utilizes data from the same four chillers McIntosh examined.

3.2. Typical Progression

In this section, each step in the progression proposed by Isermann is considered. This thesis considers Isermann's process supervision for a centrifugal chiller. Consequently, the key steps outlined in this section will include examples of what has been proposed or done relating to HVAC applications, specifically to chillers where possible.

3.2.1. Equipment/Processor Monitoring.

The supervision process can be applied to equipment like a chiller where faults occur and monitoring is desired. For any FDD methodology utilized in the supervision process, two types of information are relevant. First, any background knowledge about the equipment or its internal processes is beneficial. This may include general information like the underlying physics, the type of forcing inputs affecting the operating point, or installation specific information such as equipment geometry or details about controls or coupled components. Anything related to the deterministic nature of the operation of the equipment could be utilized.

Second but more importantly, all FDD methods require some amount of data the quantifies performance. These data are essential in determining the capabilities of the supervision process or conversely the desired FDD capabilities could dictate the data collection.

3.2.2. Preprocessor/Data Reduction

The raw data are not necessarily the best information to be passed on to the detection and diagnostic components. Filtering and transformations are two techniques utilized within a preprocessor. The output of a preprocessor can vary dramatically depending on the equipment involved and the detection and diagnosis schemes to be utilized. Filters and transformations are described in the next two subsections.

3.2.2.1. Filters

Filters are applied to data for several reasons but generally are employed to keep the meaningful information contained in the data and to remove or minimize the

unimportant aspects of the data. Another primary goal of employing a filter is to classify the data. The best example of data classification may be to detect and remove non-steady state data. The majority of fault detection studies to date have been based on steady state operation.

Given the dynamic nature of weather and building space utilization, HVAC systems are often cycled on and off between operating conditions. Data collected after the system has begun transitioning but before completing the transition to the new operating condition is considered non-steady-state. Rossi and Braun (1997) applied FDD techniques to all of the data for a specific period of time but "announced" the results only if the filter detected steady-state operating conditions. Stylianou and Nikanpour proposed a simple generalized steady-state detector based on temperature rates of change. The output of that detector dictated which of their three FDD modules should be used.

McIntosh filtered data to consider only steady-state operation and to remove the higher frequencies so as to preserve and accentuate the lower frequencies, which were deemed more important than the higher frequencies. After considering several filtering and smoothing approaches, McIntosh's final technique consisted of five-minute sampling combined with a 60-minute or 13 point moving average. A 13 point moving average means that the value at time t, x_t, is replaced by the average of the values x_{t-6} through x_{t+6}. In other words, the 13 points consists of the current point, the 6 previous points, and the 6 future points.

3.2.2.2. Transformations

In this preprocessor technique, the collected data may be manipulated into quantities that are more relevant and sensitive to faults. The transformed outputs of the

preprocessor are not necessarily physically significant. McIntosh wrote a data reduction module that utilized chiller thermodynamic relationships similar to those outlined in the previous chapter to convert ten measured quantities into eleven quantities that were more intuitive in characterizing the performance of the chiller and its individual components. For example, equation 2.2 shows that chiller load (i.e. energy transfer across the evaporator) can be calculated from the chilled water volumetric flow rate and entering and exiting temperatures using the density and thermal capacitance of water. Equation 2.1 defines the coefficient of performance (COP) as the cooling load divided by the power draw. COP is a common measure that combines the raw data and that has physical meaning, thus making its interpretation more intuitive. The outputs McIntosh's preprocessor were eleven characteristic quantities:

1. Q_{EVAP}
2. UA_{EVAP}
3. $APPR_{EVAP}$
4. $? T_{CHW}$
5. Q_{COND}
6. UA_{COND}
7. $APPR_{COND}$
8. $? T_{CW}$
9. $?_{ISENTROPIC}$
10. $?_{MOTOR}$
11. COP

3.2.3. Fault Detection and Fault Diagnosis

Of the process supervision steps proposed by Isermann, the detection and diagnosis stages have received the most attention and have been the focus for a large amount of research in the last decade on HVAC systems and components. Isermann and others have distinguished detection and diagnosis from each other. In separating detection and diagnosis, the objective of the detection module is to produce a binary output to denote whether or not a fault exists in the monitoring target. The objective of the diagnostic module is to determine the location and nature of the fault. Rossi and Braun (1997) propose a fault isolation strategy where fault diagnosis is achieved by performing detection for specific faults in the individual components.

3.2.3.1. Detection and Diagnostic Techniques

Regardless of whether or not detection and diagnosis occur sequentially or simultaneously, the vast majority of FDD systems are based in comparing actual values of performance characterizing quantities to corresponding predicted or expected values for the same quantities. The distinctions exist as to what quantities are compared, how the expected quantities are obtained, and what type of comparison is performed. The characteristic quantities for the actual operating condition are usually the outputs of the preprocessor. The expected quantities can be representative of correctly operating equipment or faulty operating equipment. The difference between the actual quantities and the expected quantities are referred to in the literature as innovations, residuals, or deviations. Both the expected quantities and the comparison criteria can be based either on collected data or models or a combination of the two.

The expected quantities could be from the equipment manufacturer, from previous equipment periods that were assumed fault-free, or from other similar equipment operating acceptably. This data would presumably have to be sent through the same preprocessor as the actual data. Data could also be used to construct the comparison criteria. Faults could be induced into operating equipment and recorded data would document the effects of the faults.

Models are another way to obtain the expected values of the characteristic quantities and the comparison criteria. Models can be physical or empirical and generalized or specific. Physical models for HVAC systems attempt to represent the equipment's underlying principles of mass and energy balances including heat transfer and could do so for both fault free and faulty conditions. Non-physical models would apply empirical techniques to data to quantify the chiller performance using indices that are not physically representative. For either data or model-based techniques, the binning of the comparison values is important. Binning allows the grouping of similar operating conditions together for determining the expected values.

McIntosh used data from an entire cooling season for four centrifugal chillers. He examined the data of the chillers from various portions of the season. One period of data from a chiller was accepted as being fault free. Data early in the cooling season was chosen based on the assumption that since the chillers were serviced prior to start up for the season, the performance of the chillers would only degrade as the season progressed. The acceptable data were put through a preprocessor, which filtered the data and applied transformations similar to those in the code in Appendix A. The inputs and outputs of the preprocessor became the baseline data set which were put into a lookup table. The inputs

represented the forcing inputs on the chiller and the outputs represented the corresponding expected performance characterizing values. McIntosh determined that the five input variables that were independent of chiller performance were:

1. Chilled Water Flow Rate
2. Chilled Water Supply Temperature
3. Chilled Water Return Temperature
4. Condenser Water Flow Rate
5. Condenser Water Supply Temperature.

McIntosh concluded that the chilled water supply temperature was independent of the chiller performance because it was a constant set point that the chiller would operate to meet. McIntosh further reasoned that the evaporator temperature could be used in place of both of the chilled water temperatures because the chiller must maintain the evaporator temperature in order to provide the necessary heat transfer. This reduced the first the first three variables (chilled water flow rate, supply temperature, and return temperature) to one (evaporator temperature). Further, the condenser water flow rate was also maintained constant. Thus only two variables were need to categorize the baseline data set. The lookup table was divided into 16 bins of two levels of evaporator temperature and 8 levels of condenser water supply temperature. The data from the chiller that fell into each bin were averaged to a single point. This provided the fault free data set for use in FDD.

A general regression neural network (GRNN) was used to interpolate within the lookup table to determine the expected performance at each operating condition. Actual process supervision was simulated by stepping through collected data not used in the

construction of the lookup table. For each operating point, all of the measured quantities were used to calculate the actual values of the characteristic quantities. The monitored quantities that were the significant forcing inputs (evaporator temperature and condenser water supply temperature) were input to the GRNN to estimate the expected values of the characteristic quantities. The differences between the expected and actual values of the characteristic quantities were calculated. Statistical techniques were used to show which deviations were significant. The patterns of the deviations were matched to specific faults, based on the analysis of a detailed physical model of the monitored chiller.

Rossi and Braun (1997) utilized a steady state vapor compression model to predict thermodynamic states, which were compared to the measured states. A rule based classifier assembled from actual faulty data and from model simulated faults, was used to determine the most likely of five faults to have occurred based on the deviations between the states.

3.2.4. Fault Evaluation and Decision

A fault that is detected and diagnosed can also be assessed to determine its effects. Depending on those effects, a decision can be made regarding how to proceed. Breuker and Braun (1998a) considered 12 fault classifications for rooftop air-conditioners by rate of occurrence and effects on performance. Fault evaluation for other systems could be done for any HVAC system and programmed in a manner to show the current effects of the diagnosed fault to the decision-making entity (human individual or programmed component).

Depending on the current or future consequences, there are several possible reactions that range from warnings to automatic shut down sequences. For certain faults, the equipment will eventually need to be serviced to remedy the fault. Rossi and Braun (1996) defined four criteria for when service should be performed. Service is justified when it provides for net economic savings (1) on energy and future servicing. For HVAC systems service should be performed whenever occupant comfort (2) can not be maintained or personal safety (3) (e.g. of the operator) is threatened. Finally, when the fault is causing harm to the environment (4) the equipment must be serviced.

3.3. Advancing Process Supervision

FDD concepts for HVAC systems and specifically chillers are largely in a developmental stage, as evident by the nature of the previous research. Therefore, the advancements being made in FDD are to increase and test the capabilities of the methodologies, as opposed to a more developed field where progress is defined primarily by standardizing and streamlining the methodologies. To increase the capabilities of any FDD system means being able to detect and diagnosis more types of faults, detect and diagnosis faults with higher precision and accuracy, and detect and diagnosis faults sooner.

For an FDD system to identify more faults requires a larger knowledge base of faults, which comes primarily from models and experimentation. For equipment as complicated and dynamic as chillers, physical and thermodynamic modeling is difficult and quickly becomes equipment and even installation specific rather than generally applicable. Experimentation to induce faults in an operating chiller in a controlled and

monitored environment also presents challenges. Chilled water plants are too expensive to build simply for experimentation, and the importance of existing installation to the buildings they serve makes operators understandably reluctant to yield control of their equipment.

Improving the time response of FDD systems is necessary for eventual on-line implementations. The response time of FDD techniques is affected by the data acquisition because any system can only detect or diagnose faults that happened at or prior to the last measurements. Then it is an issue of computing resources, in that whatever manipulation and analysis to be completed with the data should be done as soon after the data is available as possible.

To increase the precision or accuracy of FDD systems is essentially a matter of sensitivity and statistical significance. A tempting step is to simply utilize premium and redundant instrumentation, but a cornerstone of the FDD movement has been to not require such expense. Applying FDD techniques to existing systems will be considerably easier if the current instrumentation of the equipment can be used rather than requiring retrofit. Setting aside increasing the quality of the collected data, improvements in the precision or accuracy of FDD systems will require better analysis techniques. There are several additional types of analysis that could be performed, so it remains to be determined which are appropriate by considering their benefits and also their costs. The next chapter introduces time series analysis as one possible technique that could supplement existing FDD techniques and provide expanded capabilities.

4. Time Series Analysis

4.1. Overview

Most engineers are quite familiar with correlation and regression amongst two or more distinct variables by collecting data ((a_1, b_1), (a_2, b_2), etc.), plotting them against each other and applying a curve fit to estimate how they are related. The idea of using statistical techniques to relate quantities is quite old and is utilized when the exact nature between the quantities is not known, which is true for most interesting problems. If a process or relationship can be completely described it is called deterministic, otherwise it is said to be probabilistic. To contrast the two, the precise time that sunrise or sunset occurs at a geographic location can be calculated using astronomical geometry and is therefore deterministic, whereas the actual sunlight an area receives during the course of a day can be only be estimated within certain limits because of unpredictable cloud cover and is therefore probabilistic.

A time series is a set of observation of a single quantity collected sequentially in time (x_t, x_{t+1}, x_{t+2}, etc.) and the basis of time series analysis is that the value of a variable at the present time, x_t, is related to the value of the same variable at a previous time, x_{t-t}. If a process or relationship occurs in time and has some random, probabilistic component it is said to be stochastic. A commonly studied example of a stochastic process is a stock price.

The definitive resource for time series analysis is Box, Jenkins, and Reinsel (1994). Chatfield (1989) provides a more abbreviated treatment of the material including

four basic objectives: description, explanation, prediction (or forecasting), and control. The remaining portion of this chapter will cover the basics of time series. Box, Jenkins, and Reinsel developed the essential time series analysis tools presented in this chapter. The equations in this chapter are included as functions in several common statistical software packages and can presumably be defined by the user in other packages.

4.2. Differencing

A prerequisite for time series tools is that the data series be stationary, which means that the mean and variance are not changing significantly with time. For a non-stationary series, differencing is utilized to transform the data into a new stationary series. Equation 4.1 defines a first order difference.

$$y_t = x_{t+1} - x_t \tag{4.1}$$

If a first order difference is not sufficient to define a stationary series, an additional differencing can be performed resulting in the second order equation 4.2. Differencing should be performed until a stationary series is achieved.

$$z_t = y_{t+1} - y_t = x_{t+2} - 2x_{t+1} + x_t \tag{4.2}$$

4.3. Correlation Functions

A review of the correlation of two quantities will aid in understanding time series correlation. To determine if observations of two separate quantities are correlated, consider the example of M boxes each containing some number of apples (a) and some number of bananas (b). The M representative data pairs are (a_1, b_1), (a_2, b_2)…(a_M, b_M).

Equation 4.3 defines a cross-correlation coefficient (r_{cc}) using the observation terms (a_i and b_i) and the mean values of each series ($\bar{a}$ and $\bar{b}$).

$$r_{cc} = \frac{\sum_{i=1}^{M}(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\left[\sum_{i=1}^{M}(a_i - \bar{a})^2 \sum_{i=1}^{M}(b_i - \bar{b})^2\right]}} \tag{4.3}$$

Other works may refer to this simply as a correlation coefficient but the "cross" is necessary here to distinguish it from the time series correlation coefficient to be defined later in this section. The cross correlation coefficient can be either positive or negative. The larger the magnitude of the cross correlation coefficient the more likely the two quantities are dependent on each other. For the fruit box example, a large r_{cc} means that if the number of apples, *a*, in a given box is known a good estimate of the number of bananas, *b*, in the same box can be made, possibly using some curve fit of the data. Alternatively, if the value of *b* is known, the value of *a* could be estimated. A small r_{cc} means that knowing one quantity does not help in estimating the other. The sign of the cross correlation coefficient is also significant as it corresponds to the type of relationship between the quantities. In the fruit box example, a positive r_{cc} means that observing a large number of apples, a, for a given box will suggest a large number of bananas in the box. A negative r_{cc} would mean that for a large number of apples in a box, a small number of bananas in the box would be expected.

Now considering time series correlations, for a series of N observations in time of a single quantity (x_t, x_{t+1}, x_{t+2}, etc.) a similar approach can be used by constructing N-1 data pairs, (x_1, x_2), (x_2, x_3)…(x_{N-1}, x_N). This data pairing technique is generally coupled

with the assumption that for a large N the average of all the terms in the series (x_1 to x_N) is a good approximation for both the average of the first terms in the pairings (x_1 to x_{N-1}) and the average of the second terms in the pairings (x_2 to x_N).

Equation 4.4 incorporates these ideas to define the autocorrelation coefficient (r_{ac}) for a series of N observations.

$$r_{ac} = \frac{\sum_{t=1}^{N-1}(x_t - \bar{x})(x_{t+1} - \bar{x})}{\frac{(N-1)}{N}\sum_{t=1}^{N}(x_t - \bar{x})^2} \tag{4.4}$$

In time correlation, the concept of lag is important. In the preceding definition of the autocorrelation coefficient, a series of observations were compared to the series of observations occurring just prior, a case referred to as a lag of a 1. The autocorrelation coefficient can be calculated for any lag value (k) less than the number of observations. Equation 4.4 can be rewritten to define the autocorrelation function (r_k) as shown in equation 4.5.

$$r_k = \frac{\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\frac{(N-k)}{N}\sum_{t=1}^{N}(x_t - \bar{x})^2} \tag{4.5}$$

For large numbers of observations (N) and small lags (k), the fraction (N-k)/N approaches unity and can be approximated as such. As a means of determining the appropriateness of continuing with time series analysis, the autocorrelation function for several lag values are calculated and plotted against the lag. This type of plot is called a correlogram and consistent with the previous section it is only significant for a stationary time series.

Every calculation of r_k uses all of the observations, but statistically calculations using the smaller lag values are more meaningful because they utilize the data better. For a lag of 1, N-1 pairs ((x_1, x_2), (x_2, x_3)… (x_{N-k-1}, x_{N-1}),(x_{N-k}, x_N)) were constructed but two observations, x_1 and x_N, are each only included in one pair while the other N-2 observations are each included in two pairs. Generalizing for all lags, only N-k pairs can ever be assembled and k+1 observations are each included in only one pair while the remaining N-(k+1) observations are each included in two pairs. This variable significance of r_k makes the interpreting the correlogram difficult but Chatfield ascertains that "if a time series is random, 19 out of 20 of the values of r_k can be expected to lie between ±2/vN." Chatfield's statement can be reversed to state that if more than 1 of 20 r_k values is outside of ±2/vN, the series is not random in time. The number and location of the significant r_k values can provide some insight to the experienced time series statistician as to the nature of the model that should be fit to the data.

4.4. Time Series Modeling

Standard correlation and regression techniques can lead to a model of a quantity based on other quantities. For the fruit box example, perhaps the number of bananas can be approximated as five times the number of apples. Mathematically this can be expressed as $b_i = 5a_i + e_i$ with e_i representing some random error. Similarly, a time series analysis can lead to model for an observation in time based on prior observations of the same quantity. Model construction can begin on a stationary data series, x_t, when the correlogram suggests it. The determination of model parameters is outside the scope of

this work, but Box et al, Chatfield, and several statistical software packages contain algorithms to estimate the values for the model parameters.

To begin modeling, a difference term between time points can be defined as shown in equation 4.6. This difference term is assumed to have a mean of zero and a variance that does not vary with time.

$$z_t = x_t - x_{t-1} \qquad \text{and} \qquad z_{t-1} = x_{t-1} - x_{t-2} \tag{4.6}$$

A time series model using q difference terms each multiplied by a constant coefficient is referred to as moving average, or MA(q), process and described by equation 4.7.

$$x_t = b_0 z_t + b_1 z_{t-1} + \ldots + b_q z_{t-q} \tag{4.7}$$

If a process can be described using some number (p) of previous observations, than a p^{th} order auto regressive, or AR(p), model is appropriate. The general form of the AR (p) model is shown in equation 4.8.

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_p x_{t-p} + z_t \tag{4.8}$$

Using successive substitutions, equation 4.8 becomes 4.9, which is also the representation for an infinite order moving average process.

$$x_t = z_t + a\ z_{t-1} + a^2 z_{t-2} + \ldots \tag{4.9}$$

A process may not be exclusively autoregressive or moving average, and the two relationships can be combined into a mixed model, ARMA (p,q), as shown in equation 4.10.

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_p x_{t-p} + z_t + b_1 z_{t-1} + b_2 z_{t-2} + \ldots + b_q z_{t-q} \tag{4.10}$$

For the case of applying an ARMA model to a non-stationary time series that must be differenced, the notation can be written compactly to define an autoregressive integrated moving average process, ARIMA (p,d,q).

5. Case Study

This chapter describes a chiller installation and details about the data that were collected. Simple statistical quantities were calculated for a portion of the collected data. Then existing FDD methodologies and finally time series techniques were applied to the same portion of the collected data. The results were compared and evaluated to determine the appropriateness and capabilities of time series analysis. Reiterating from the first chapter, the goal of this thesis is the improvement of fault detection and diagnosis methodologies. The data are presented as means of achieving that goal and are not presented as a means of optimizing the performance of this equipment installation.

5.1. Data Source and Composition

The data used in this case study was collected at a privately owned and operated industrial facility during the 1998 cooling season. The chilled water plant at the site has four centrifugal chillers (CC 1-4), four absorption chillers (AC 5-8), two water to water heat exchangers (HX 1-2), a ten cell cooling tower, seven primary chilled water pumps, and four condenser water pumps. A simplified schematic is provided in Figure 5-1.

The four absorption chillers are steam-fired units each rated at 1500 tons of cooling. The heat exchangers are used in an economizer cycle between the chilled water returned from the building and the condenser water returned from the cooling towers. If the chilled water leaving the heat exchanger is cool enough it is directed to the chilled water supply line, otherwise it is directed to a centrifugal chiller. When the water leaves

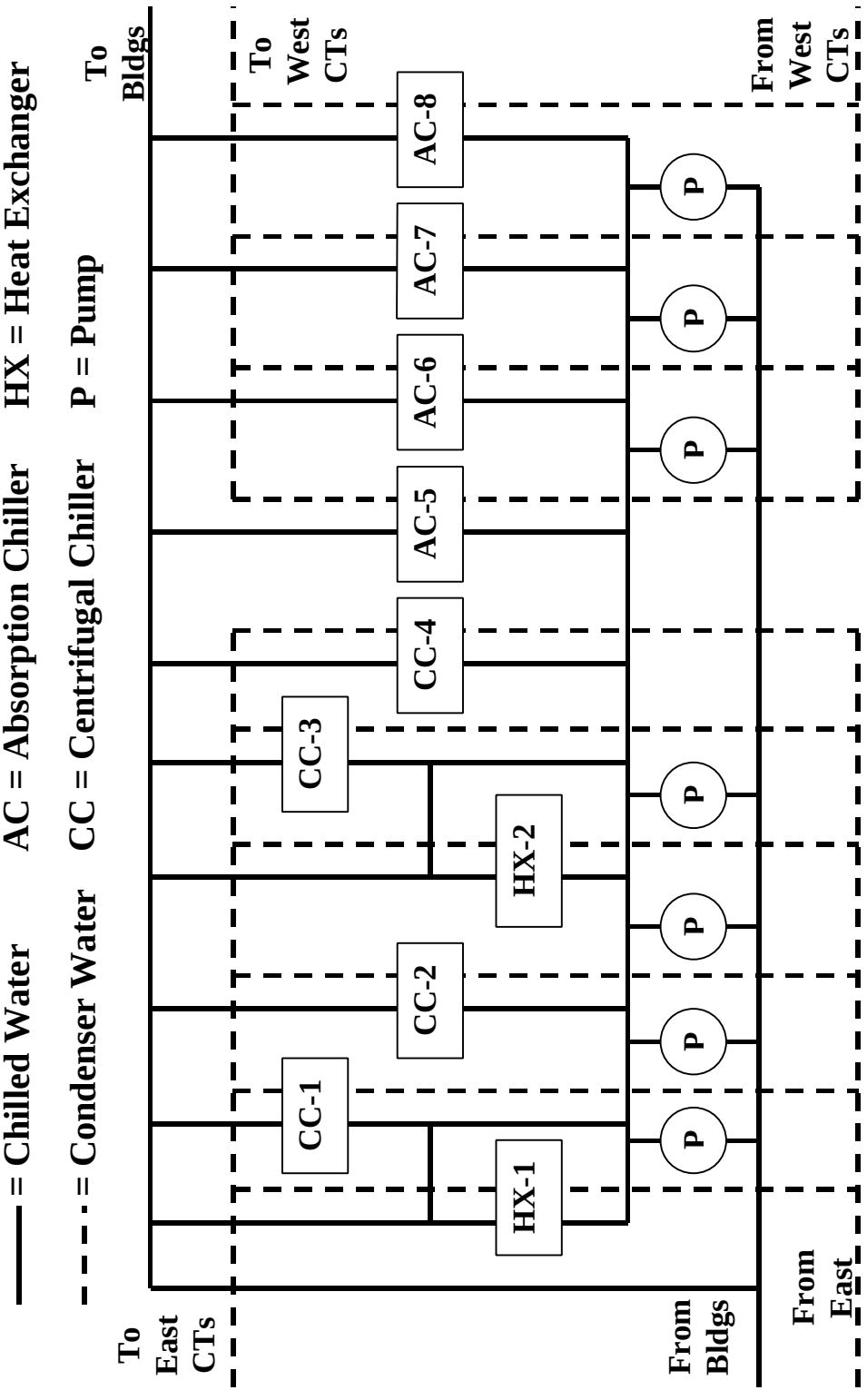

Figure 5-1: Plant Schematic

the heat exchanger and progresses to a chiller, then the heat exchanger is pre-cooling the water to reduce the load the chiller must meet. Both heat exchangers are connected to specific chillers, heat exchanger 1 to centrifugal chiller 1, and heat exchanger 2 to centrifugal chiller 3. Operating data were not collected for the absorption chillers, the heat exchangers or the pumps.

The four centrifugal chillers (referred to simply as 1-4 for the remainder of this thesis) are the same model (2000 ton R-22 units) and were installed at the same time. For each of the centrifugal chillers, the same ten variables were monitored. The following table denotes the variables (notation consistent with chapter 2), an abbreviated code, and units.

Table 5-1: Chiller Monitored Parameters

Parameter	**Code**	**Units**
Chilled Water Volumetric Flow Rate	GPMCHW	gpm
Condenser Water Volumetric Flow Rate	GPMCW	gpm
Chilled Water Supply Temperature	TCHWS	°F
Chilled Water Return Temperature	TCHWR	°F
Condenser Water Supply Temperature	TCWS	°F
Condenser Water Return Temperature	TCWR	°F
Evaporator Saturation Temperature	TEVAP	°F
Condenser Saturation Temperature	TCOND	°F
Compressor Discharge Temperature	T2	°F
Electrical Power Draw	P	kW

Some portions of the data were collected on five minute intervals in a non-simultaneous staggered manner as shown in Table 5-2 using a portion of the data from chiller four on June 1st.

Table 5-2: Staggered Data Set

Time	GPMCHW	TCHWR	TCHWS	T2	Power	TEVAP
12:00:22 AM	4944.5	49.07			709.1	
12:01:22 AM				129.2		
12:02:22 AM						42.71
12:04:22 AM			47.05			
12:05:22 AM	4921.2	49.07			710.9	
12:06:22 AM				128.7		
12:07:22 AM						42.71
12:09:22 AM			46.73			
12:10:22 AM	4874.4	49.07			710.5	

Other portions of the data were collected in a more conventional method of logging each channel every minute as shown in Table 5-3 using a portion of data from chiller three on June 28th.

Table 5-3: Normal Data Set

Time	GPMCHW	TCHWR	TCHWS	T2	Power	TEVAP
10:38:21 AM	4946.7	49.60	45.14	130.90	729.7	40.99
10:39:21 AM	3884.1	49.97	45.14	131.40	755.5	42.40
10:40:21 AM	3106.4	49.60	45.77	131.40	763.4	43.10
10:41:21 AM	3372.0	49.97	47.05	132.00	770.8	43.81
10:42:21 AM	4496.2	49.60	48.00	132.00	713.5	44.51
10:43:21 AM	5628.8	49.26	48.00	132.00	717.0	44.51

5.2. Data Segment

For comparison of existing methodologies and time series techniques, a suitable set of data was chosen. Variations in the chilled water flow rate through chiller one during June

28th to 30th were intriguing and consequently that set of operating data was chosen for analysis. The chilled water flow rate over this period is plotted in Figure 5-2.

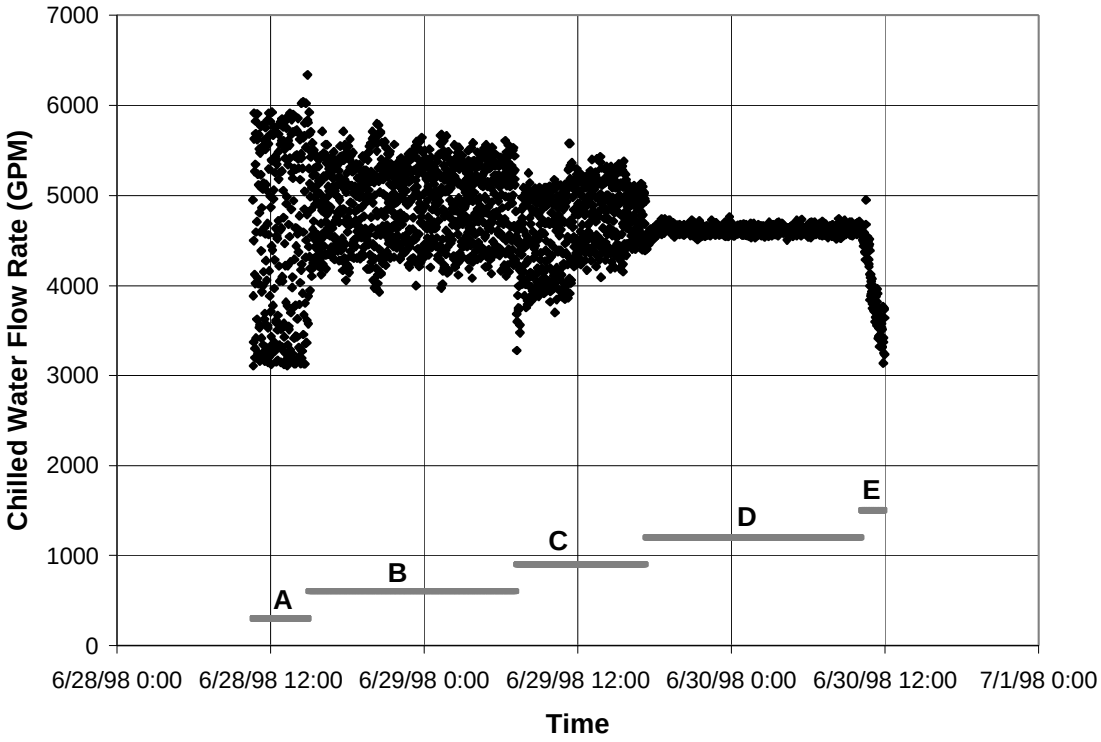

Figure 5-2: Chilled Water Flow Rate for Chiller 1 in June

The flow rate has four evident transitions that motivated dividing the data set into five segments, labeled A-E. Segment A starts when the chiller was brought on-line shortly before 12:00 on the 28th. The chilled water flow rate varied between 3000 and 6000 gpm for 5 hours after start up. The mean and variance of the flow rate changes at about 15:00 and the second section, segment B, has a higher mean and a smaller variance and extents until about 7:00 on the 29th when the flow rate dips significantly. Segment C continued until 17:30 on the 29th when the chilled water flow rate appears to settle to a more constant value. The duration of the nearly constant flow, segment D, lasted until 10:00 on the 30th, when the flow dropped. Segment E, the fifth and final section, consists

of the dropping chilled water flow rate as chiller one approached shutoff at 12:00 on the 30th.

The transition between segments B and C was attributed to the start up of chiller two and was the only change of behavior that could be explained. Chiller 2 remained on through segments C through E. The causes of the other transitions in the chilled water flow rate through chiller 1 were unknown. Chillers three and four were both off during the entire time considered and no data were collected for the absorption chillers, heat exchangers, or pumps.

All ten monitored quantities listed in Table 5-1 were collected in one minute intervals for chiller one between the 28th and 30th of June, but only the four evaporator side quantities will be examined in depth here. Specifically, these four quantities and their short hand codes are: the previously plotted chilled water flow rate (CHWGPM), the chilled water supply (TCHWS) and return (TCHWR) temperatures, and the refrigerant temperature in the evaporator (TEVAP). The three temperatures are plotted in Figure 5-3.

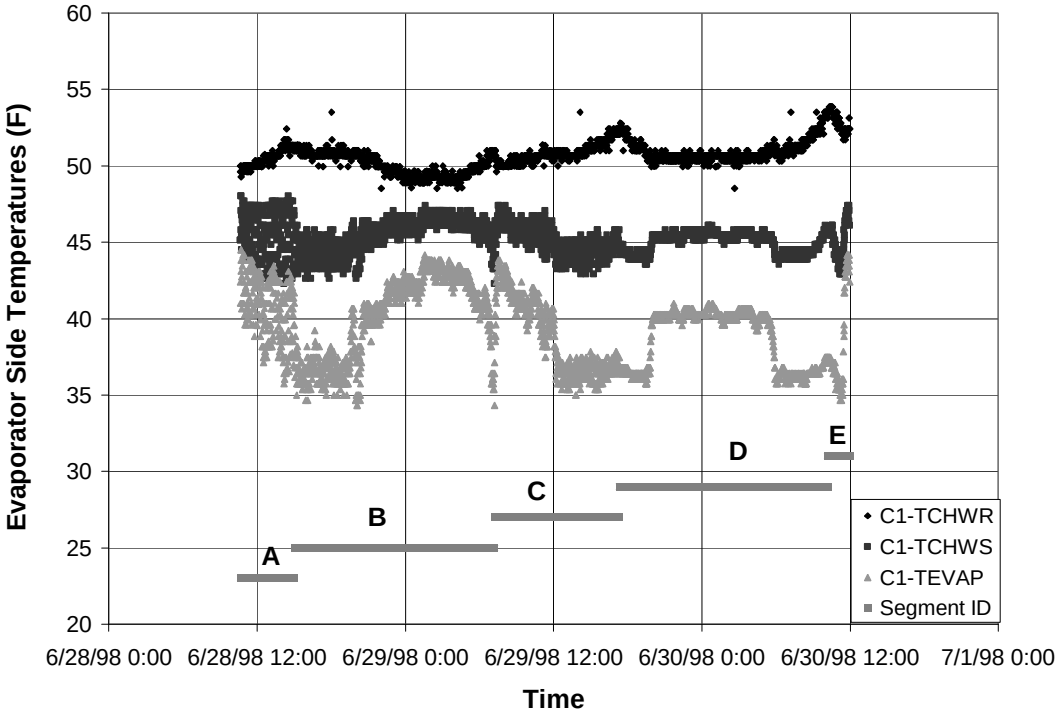

Figure 5-3: Evaporator Side Temperatures for Chiller 1 in June

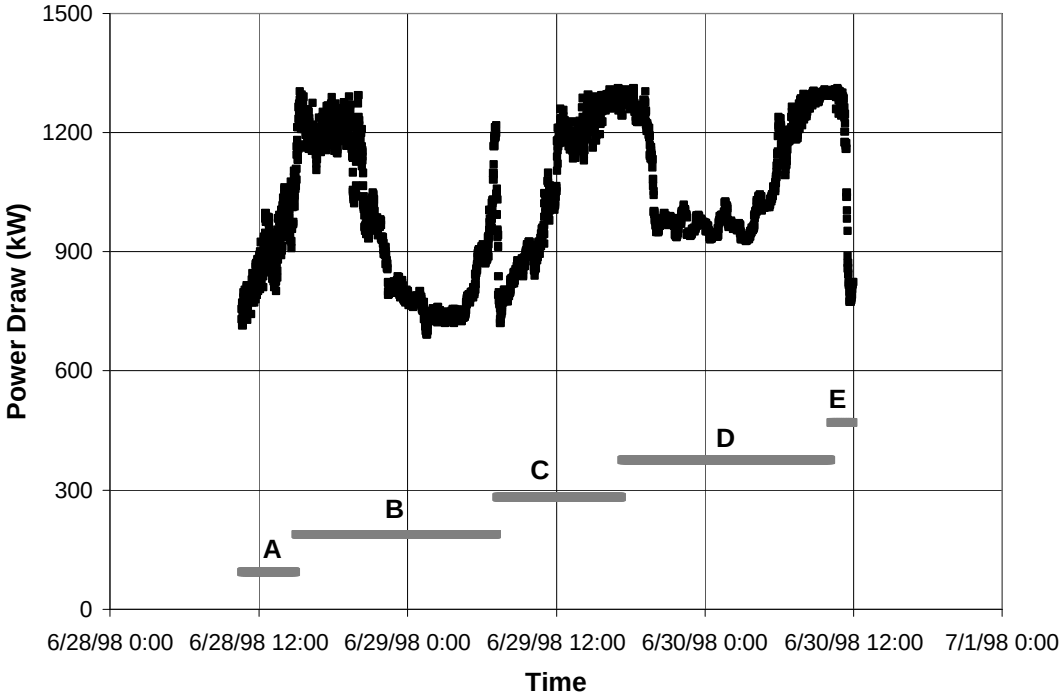

Figure 5-4: Power Draw for Chiller 1 in June

While the electrical power draw is not one of the four quantities being examined in depth, Figure 5-4 provides insight into the overall operating condition of chiller 1 during the specified period. The power draw is directly related to the cooling load met by the chiller. The daily afternoon peaks correspond to presumed peak cooling loads. There is also the expected drop in power when chiller 2 was started and began to meet a portion of the load, reducing the amount of the load that chiller 1 had to meet.

Table 5-4 contains the values of the mean (μ) and standard deviation (s) for the four monitored quantities during the five segments.

Table 5-4: Statistics of Unfiltered Data for Chiller 1 in June by Segment

	GPMCHW (GPM)		**TCHWR (F)**		**TCHWS (F)**		**TEVAP (F)**	
ID - Pts.	**μ**	**s**	**μ**	**s**	**μ**	**s**	**μ**	**s**
A – 263	4510.1	966.2	50.48	0.59	45.72	1.50	40.78	1.96
B – 972	4898.4	434.0	50.03	0.74	45.50	1.04	40.05	2.75
C – 606	4658.6	413.8	50.80	0.61	45.09	0.97	38.70	2.45
D – 1012	4613.8	43.5	50.98	0.70	44.98	0.64	38.62	1.90
E – 108	3967.6	450.9	52.84	0.61	45.07	1.21	38.17	2.98

The standard deviation of the chilled water flow rate segment A is over 950 gpm and while for segments of B, C and E, the standard deviation values all exceed 400 gpm. Segment D was essentially constant value flow rate and accordingly it has the smallest standard deviation.

5.3. Analysis Utilizing Existing Methodologies

This section will apply some of the currently prevalent ideas for analysis to the data introduced in the preceding section. McIntosh examined an earlier portion of data from chiller 2, where the data were staggered five-minute samples similar to the data shown in Table 5-2. McIntosh decided upon a filtering technique consisting of a 60-minute moving average, which utilized 13 data points. For a 61 point moving average calculation, to calculate the smoothed value at a given time requires the actual value at that time, the 30 points prior and the 30 points after the time. For the data from chiller 1 during June 28th to the 30th 1-minute samples were collected. Consequently, for this thesis a 61 point moving average filter was applied thereby preserving McIntosh's 60 minute time length averaging window. Figure 5-5 contains the plot of the filtered chilled water flow rate through chiller 1.

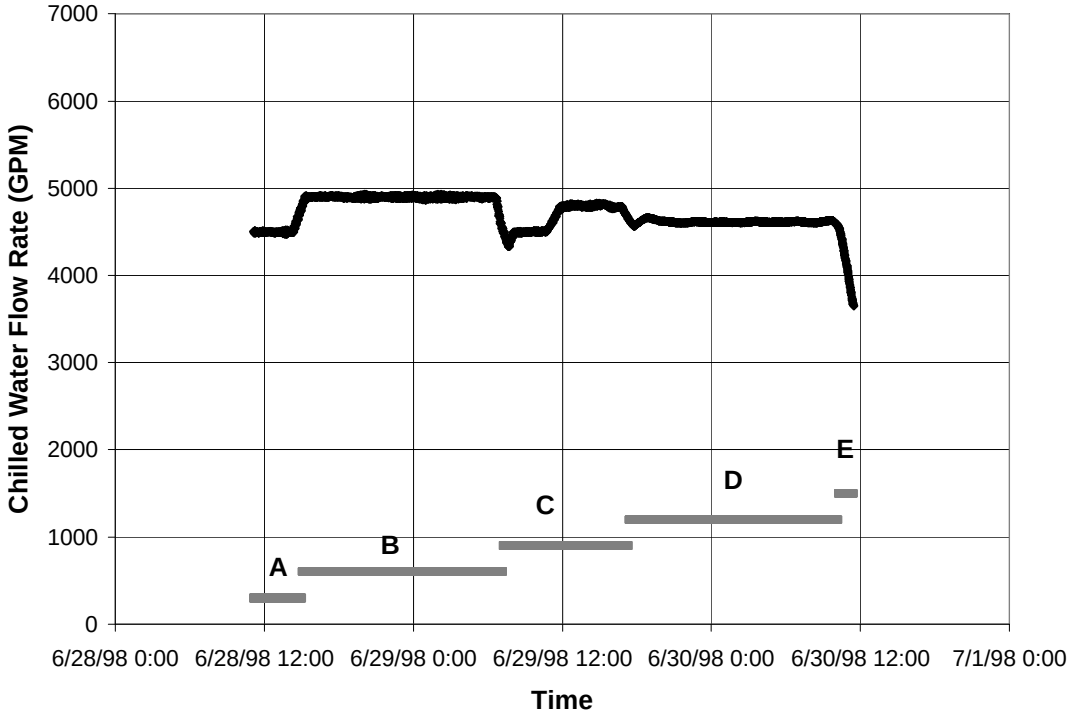

Figure 5-5: Filtered Chilled Water Flow Rate for Chiller 1 in June

The chilled water flow rate variations evident in Figure 5-2 are not evident in the filter results plotted in Figure 5-5. The transition between the segments become more pronounced and an additional, fifth, transition appears in the middle of the C segment where the flow rate rises about 300 gpm.

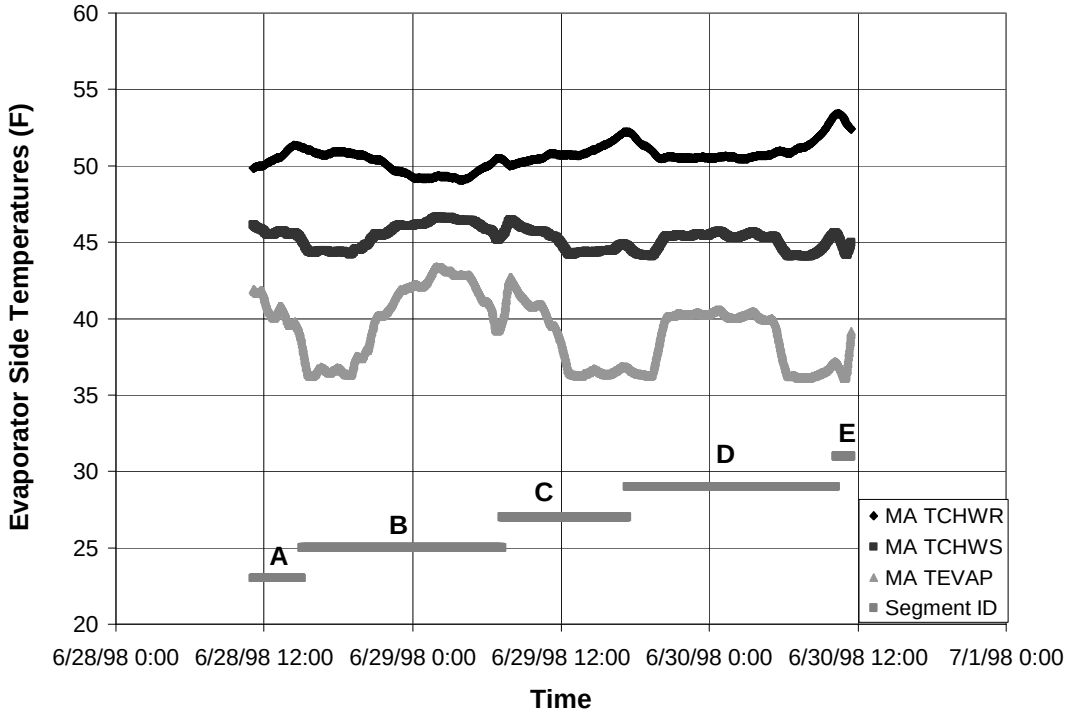

Figure 5-6: Filtered Evaporator Side Temperatures for Chiller 1 in June

The difference between temperature plots, Figures 5-6 and 5-3, is more subtle than the chilled water flow rate plots, but the filter has removed the short term variations of the temperature but preserved the larger scale transitions.

From the results of the moving average filtered data, the mean and standard deviations of each of the four quantities were calculated. Consequently filtered values are not available for the first and last 30 points in each series. Notice that for each segment or string of segments in Table 5-5, the number of points is 60 less than the corresponding value in Table 5-4 because for 30 values at both ends of each segment a filtered value cannot be calculated.

Table 5-5: Statistics of 60 Min MA Filtered Data for Chiller 1 in June by Segment

	GPMCHW (GPM)		TCHWR (F)		TCHWS (F)		TEVAP (F)	
ID - Pts.	**μ**	**s**	**μ**	**s**	**μ**	**s**	**μ**	**s**
A – 203	4500.0	14.3	50.47	0.46	45.67	0.17	40.68	0.76
B – 912	4899.0	11.2	49.98	0.68	45.55	0.85	40.22	2.58
C – 546	4676.0	144.5	50.77	0.47	45.05	0.73	38.62	2.28
D – 952	4615.4	12.4	50.87	0.50	44.98	0.58	38.73	1.82
E – 48	3945.0	195.8	52.81	0.29	44.49	0.24	37.05	0.97

For the chilled water flow rate, the mean values of the filtered data for segments A, B, and D are quite similar to the corresponding mean values of the unfiltered data. The plots of the unfiltered and filtered data for segments C and E both showed changes in the mean over the whole segment. Consequently, the points trimmed off the front and back of each segment caused the discrepancy between the mean values of the filtered and unfiltered data for segments C and E. The small values of the standard deviations of all four quantities in Table 5-5 confirms the conclusion from Figure 5-5, that the moving average filters removed the short time scale variations.

5.4. Analysis Including Time Series Methodologies

The motivation for the time series analysis performed in this thesis resulted from examining the data at an expanded time scale. There is an obvious regular time dependency of the data. Figure 5-7 shows this in the unfiltered data during the transition between segments A and B.

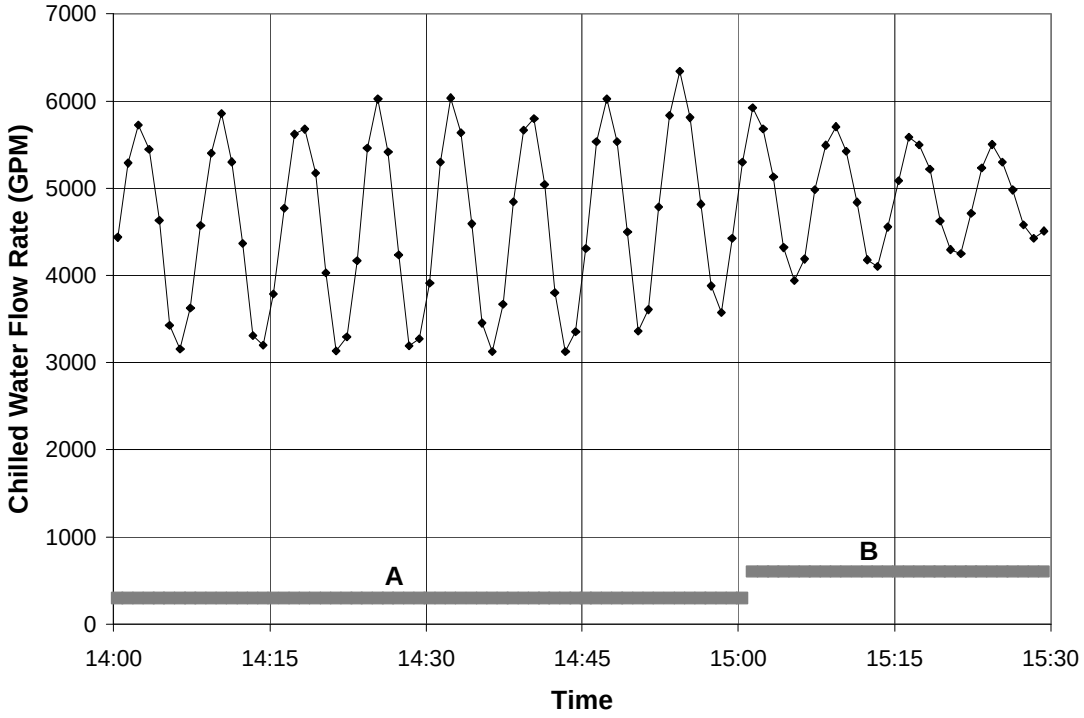

Figure 5-7: Transition Between Segments A and B on June 28th

The lines connecting the data points make the time variations evident. During the last hour of segment A, the chilled water flow rate varied approximately sinusoidally with a mean of 4500 gpm, a mean to peak amplitude of 1300 gpm, and a period of about 8 minutes. At the start of segment B, the period was similar but the mean rose and the amplitude of the fluctuation decreased. The chilled water flow rate during the transitions between segments B and C, segments C and D, and segments D and E are plotted in Figures 5-8, 5-9, and 5-10 respectively.

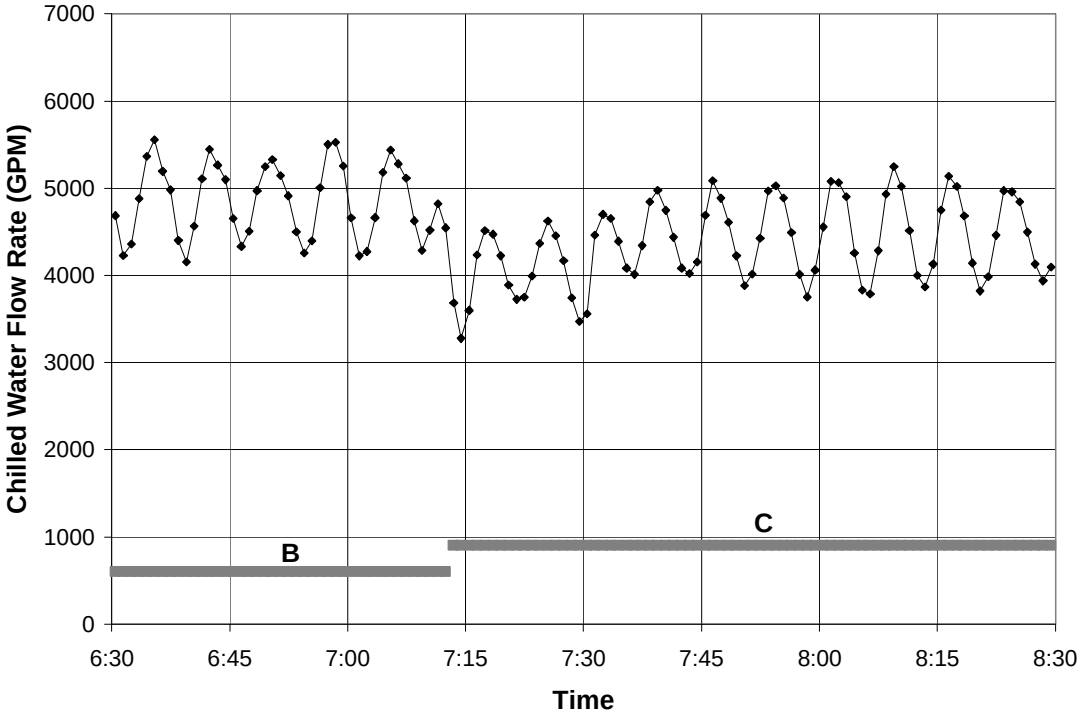

Figure 5-8: Transition Between Segments B and C on June 29th

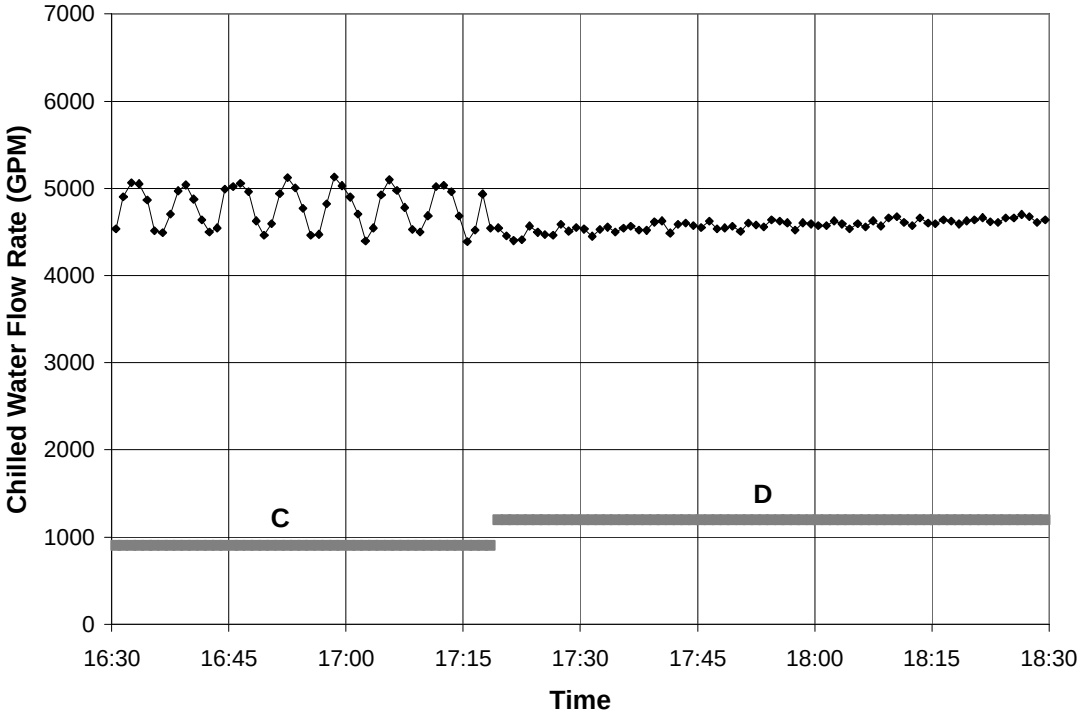

Figure 5-9: Transition Between Segments C and D on June 29th

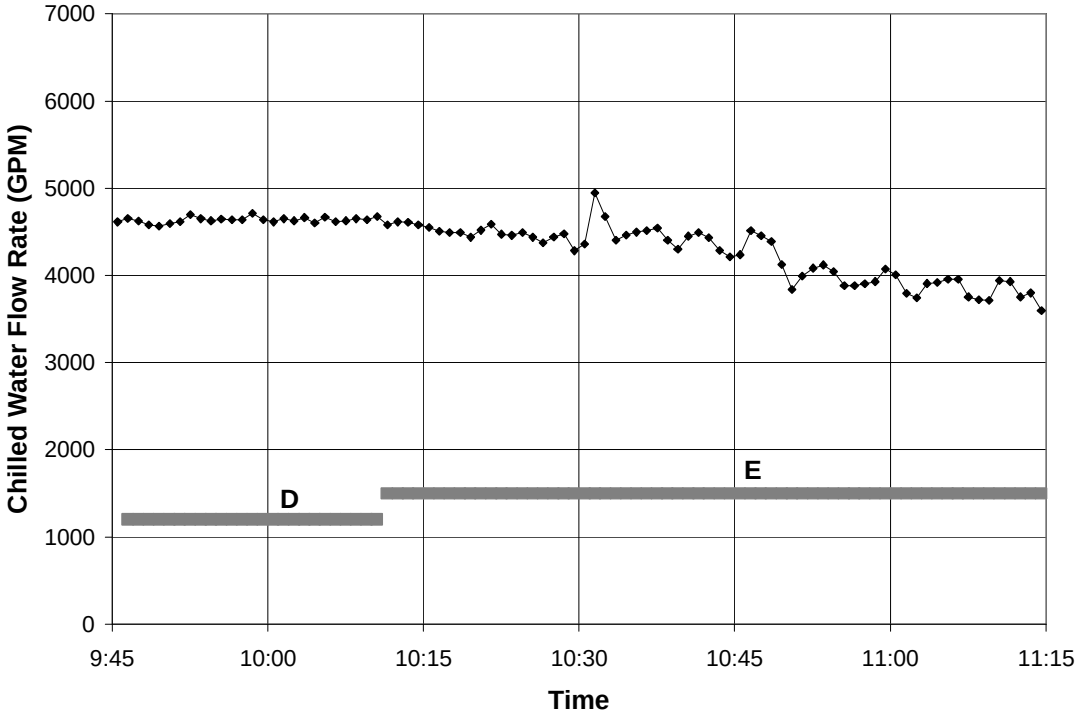

Figure 5-10: Transition Between Segments D and E on June 30th

The autocorrelation function (r_k) defined in Equation 4.5 was introduced as a means of determining the appropriateness of proceeding with time series analysis. Values of r_k for the chilled water flow rate during June 28th-30th were calculated for lags up to 100 and are shown in correlogram form in Figure 5-11.

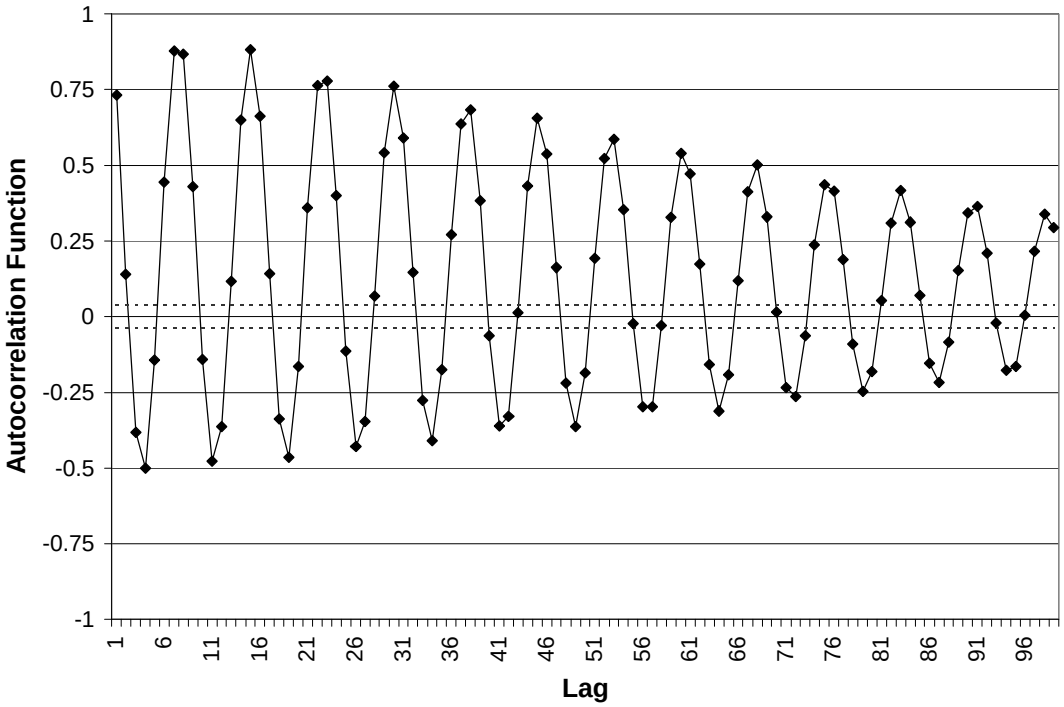

Figure 5-11: Autocorrelation Function of Chilled Water Flow Rate

The dashed horizontal lines correspond to Chatfield's estimation of confidence intervals for determining significance. It is noteworthy that while the values of r_k decrease for larger lags, the lag at every lag step up to 100 remains significant using Chatfield's criteria, meaning that even the value of the chilled water flow rate 100 minutes ago can provide a reasonably good estimate of the current value. The periodic nature of the chilled water flow rate appears in the correlogram.

The chilled water flow rate is one of the forcing inputs on a chiller. The time variant nature is also evident in other measured quantities of the chiller, especially on the evaporator side. Figure 5-12 shows the correlograms for the three evaporator side temperatures: evaporator refrigerant temperature and chilled water supply and return temperatures.

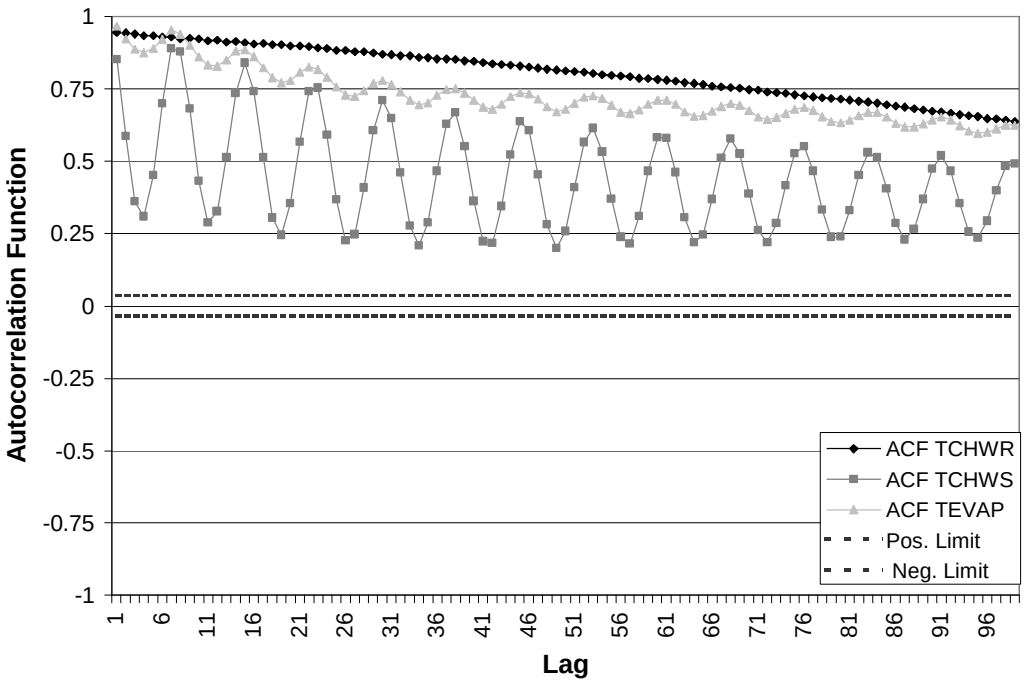

Figure 5-12: Autocorrelation Function of the 3 Evaporator Side Temperatures

The chilled water supply temperature and evaporator temperature are both strongly autocorrelated and periodic in nature, reflecting the effects of the time variant chilled water flow rate. The chilled water return temperature, which is returning from the building loop shows a strong autocorrelation but is not periodic in nature. The chilled water return temperature is not directly affected by performance of the chiller or the chilled water mass flow rate, rather it is one of the forcing inputs directly affecting the operation of the of the chiller.

The correlograms in Figures 5-11 and 5-12 motivated fitting an ARIMA model to the chilled water flow rate. The end purpose was to incorporate time series analysis techniques into on-line chiller fault detection and diagnosis. The question arose as to how to specifically apply an ARIMA model to the data. A “sliding ARIMA approach”

was developed as a macro in a commercial software package and is included as Appendix B.

The sliding window ARIMA approach fits an ARIMA model of specified form to a string of data points of a particular length and then steps through fitting a new model for overlapping windows of data. This technique requires the values of p, d, and q of the ARIMA model to be determined by the user. The values of p, d, and q, are held constant throughout the macro but the values of the autoregressive (a) and moving average (ß) coefficients are calculated for each window of data. For this work, a window length of 60 points was chosen consistent with Chatfield's suggestion of a minimum of 50 points and also to correspond to McIntosh's moving average filter length.

A walk through of one loop of the macro code helps clarify the procedure. If an ARIMA(1,1,1) was chosen for application to a string of points, the macro begins by reading in points 1 – 60 in the series, the first window. The mean of the window is calculated and subtracted from each data point. The ARIMA command in the software package takes the difference between two data points and estimates the two coefficients (one autoregressive and one moving average). These coefficients allow forecasting of the next value (point 61) and the 95% upper and lower confidence intervals of the forecast. The values of the confidence intervals are calculated based on the distribution of the random residual component. The values of the coefficients, the forecast, and the forecast limits are all output (starting with point 61 not point 1). The code then loops to read in a new window consisting of points 2-61 and the process repeats through the entire string of data.

The sliding ARIMA approach only considers one ARIMA model for one quantity series. The macro must be rerun for each combination of model and quantity to be fitted. Several ARIMA models were applied with the sliding window approach to the chilled water flow rate through chiller 1. An ARIMA (3,0,0) model was found to adequately fit the chilled water flow rate based on visual inspection and the sum of the squares of the residuals. The generalized form of an ARIMA (3,0,0) model is shown in Equation 5.1.

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + a_3 x_{t-3} + z_t \tag{5.1}$$

The actual data point, its forecasted value, and the confidence intervals from the ARIMA (3,0,0) results are plotted in Figures 5-13, 5-14, 5-15, and 5-16 for the same periods of data plotted in Figures 5-7, 5-8, 5-9, and 5-10. The coefficients calculated by the sliding window macro for the ARIMA(3,0,0) model of the chilled water flow rate are plotted in Appendix C. The macro code also sets a binary alarm value to unity if the actual value of the next point is outside of the forecast confidence intervals. This alarm is plotted on the right axis of all four plots.

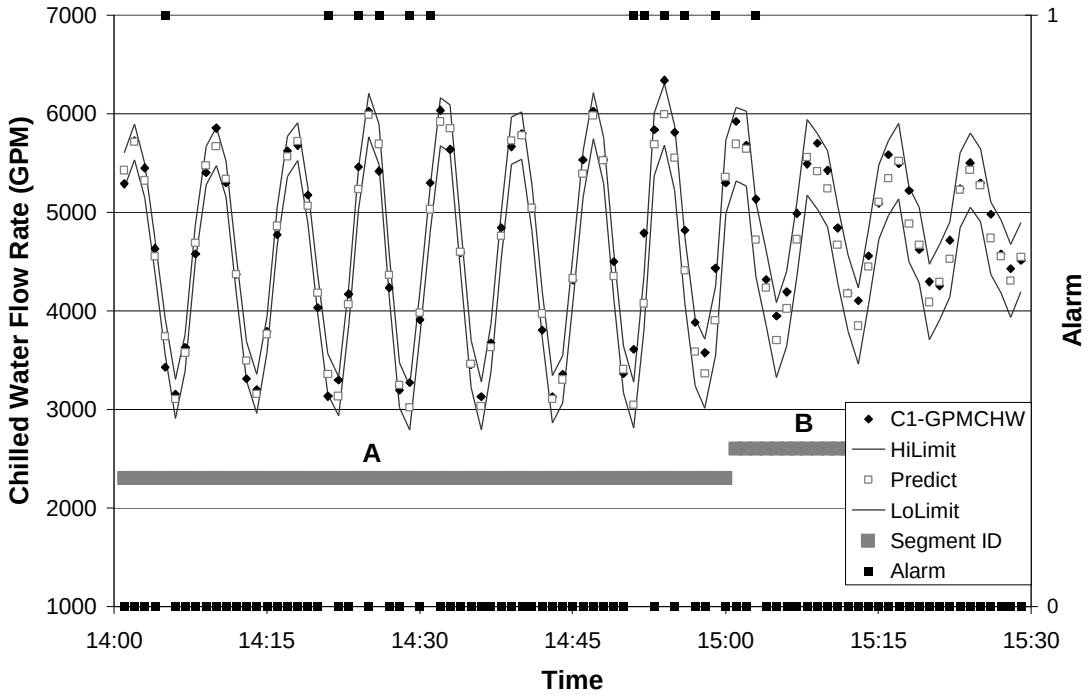

Figure 5-13: ARIMA (3,0,0) Transition Between Segments A and B on June 28th

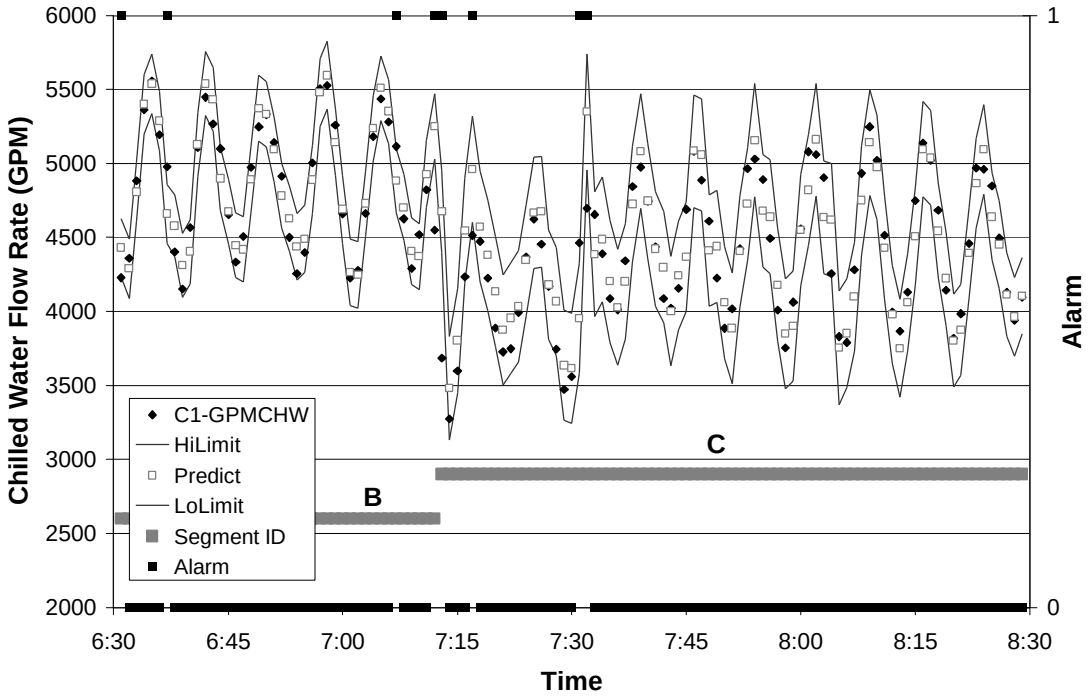

Figure 5-14: ARIMA (3,0,0) Transition Between Segments B and C on June 29th

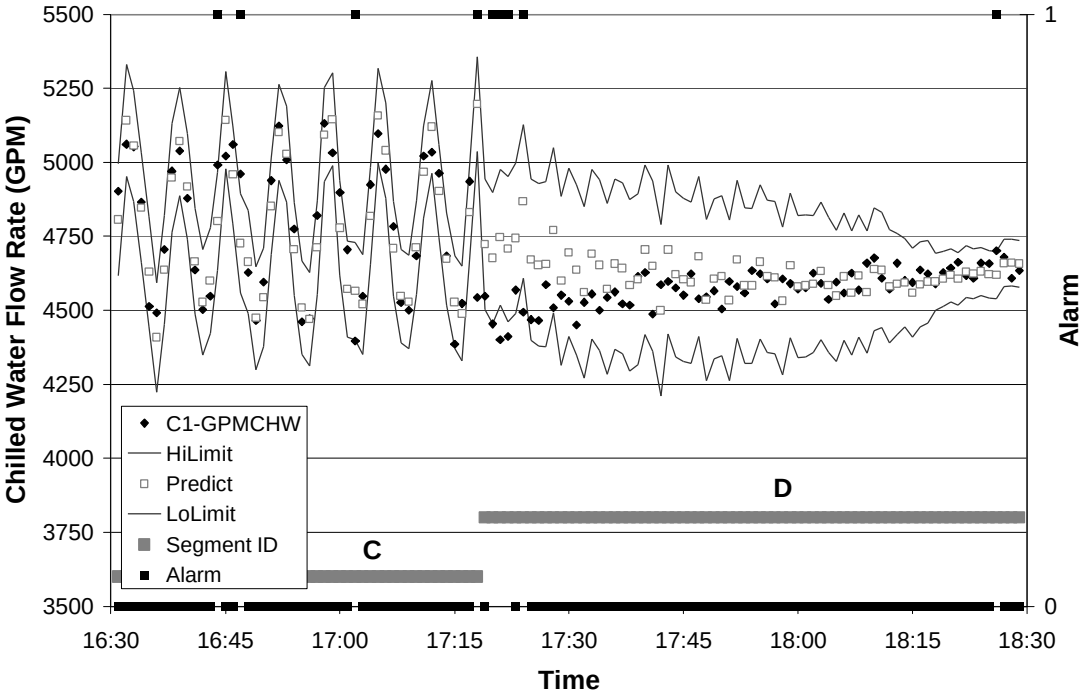

Figure 5-15: ARIMA (3,0,0) Transition Between Segments C and D on June 29th

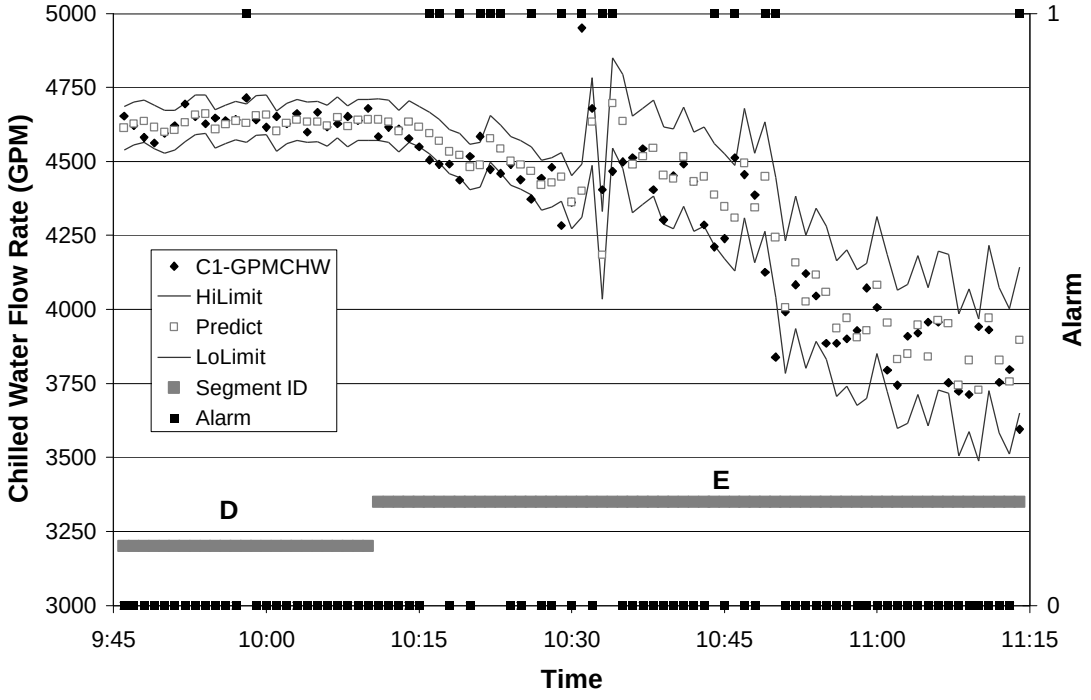

Figure 5-16: ARIMA (3,0,0) Transition Between Segments D and E on June 30th

For each of the four plots of the ARIMA(3,0,0) results, the number of alarm occurrences was high during the transition from one segment to another, indicating that the sliding window ARIMA identified the transitions as changes in the behavior of the chilled water flow rate. Considering the plots together shows the sliding ARIMA model was able to adapt to the new pattern of behavior after the transition. Figure 5-15 in particular shows the model predictions are periodic like the data during segment C, but as the window passed over the transition region the predicted values showed less and less periodic variability better matching the data of segment D.

The ARIMA (3,0,0) model form was selected from several models fitted to the chilled water flow rate and since the chilled water flow rate is strongly influencing two of the three temperatures (chilled water supply and evaporator refrigerant), intuition suggests that the ARIMA (3,0,0) model is a good first guess. Consequently, an ARIMA(3,0,0) model was fitted to all three temperatures. It is not necessarily the most suitable model for any of the three evaporator side temperatures and no other models were fitted to the temperature quantities.

As with the original data and filtered data, the statistics for the ARIMA (3,0,0) model results were calculated and are presented in Table 5-5. The standard deviations were calculated using the residuals between the actual values and the ARIMA model values.

Table 5-6: Statistics of ARIMA (3,0,0) Results for Chiller 1 in June by Segments

	GPMCHW (GPM)		TCHWR (F)		TCHWS (F)		TEVAP (F)	
ID - Pts.	**μ**	**s**	**μ**	**s**	**μ**	**s**	**μ**	**s**
A – 203	4495.8	128.8	50.63	2.19	45.65	1.43	40.52	1.29
B - 912	4899.0	106.5	49.97	3.30	45.57	1.86	40.26	0.55
C – 546	4679.2	108.7	50.83	2.58	45.01	1.51	38.37	0.51
D – 952	4616.8	39.9	50.89	2.33	45.01	1.34	38.74	0.20
E – 48	3619.9	145.8	52.29	3.90	45.17	2.39	39.89	0.61

The standard deviations of the chilled water flow rate for segments A, B, and C were all between 100 and 150 gpm, while the values in segments D and E, were about 40 and 150 respectively. The standard deviations of the chilled water return temperature for all segments exceeded 2 degrees F.

5.5. Discussion of Results

The standard deviation and mean values in Tables 5-4, 5-5, 5-6 representing the original collected data, the 61 point moving average filtered data, and the ARIMA modeled data respectively, collectively quantify the comparison of the analysis techniques. For ease of comparison, the standard deviation and mean values for the collected, filtered, and ARIMA fitted data are presented by variable in Tables 5-7 (GPMCHW), 5-8 (TCHWR), 5-9 (TCHWS), and 5-10 (TEVAP).

Table 5-7: GPMCHW Statistics for Chiller 1 in June by Technique and Segments

	Mean (GPM)			Standard Deviation (GPM)		
ID	**Collected**	**Filtered**	**ARIMA**	**Collected**	**Filtered**	**ARIMA**
A	4510.1	4500.0	4495.8	966.2	14.3	128.8
B	4898.4	4899.0	4899.0	434.0	11.2	106.5
C	4658.6	4676.0	4679.2	413.8	144.5	108.7
D	4613.8	4615.4	4616.8	43.5	12.4	39.9
E	3967.6	3945.0	3619.9	450.9	195.8	145.8

For each of the first four segments of the chilled water flow rate, the mean values that were calculated by the three different techniques are similar, as expected and as desired. In the E segment, the mean of the ARIMA fitted values is significantly lower than the other two. As discussed in chapter 4, the time series techniques are only valid for stationary data series. Because the chilled water flow rate drops throughout segment E, it cannot be considered stationary. In plotting the filtered value chilled water flow in Figure 5-5, segment C was observed to also have an overall trend and therefore also not stationary. There is no differencing in the ARIMA(3,0,0) model to account for the original data not being stationary, consequently while this ARIMA model form adequately fits the data for segments A, B, and D, it is not suitable for segments C or E.

Examining the standard deviation values of the chilled water flow rate in segments A, B, and D, the collected data had the largest standard deviation values and the filtered data had the smallest. The 61 minute moving average filter has removed all of the short time scale variations, namely the 8 minute periodic oscillation and as a result, the values of standard deviation of the filtered data are quite small. There is an actual physical randomness to the chilled water flow rate, but because the moving average filter

fails to capture the obvious time variant behavior of the data, the calculated standard deviation are not reasonable estimates of the actual randomness.

In Figures 5-13 through 5-16, an ARIMA (3,0,0) model was shown to quantify the time variant nature of the data and the resulting standard deviation values are then presumably better estimates than those from the collected or filtered data. This is confirmed by the standard deviation values from the ARIMA model being smaller than those values from the collected data. The ARIMA model (3,0,0) provides the best estimate of the randomness in the chilled water flow rate of these three techniques.

Table 5-8: TCHWR Statistics for Chiller 1 in June by Technique and Segments

	Mean (F)			Standard Deviation (F)		
ID	**Collected**	**Filtered**	**ARIMA**	**Collected**	**Filtered**	**ARIMA**
A	50.48	50.47	50.63	0.59	0.46	2.19
B	50.03	49.98	49.97	0.74	0.68	3.30
C	50.80	50.77	50.83	0.61	0.47	2.58
D	50.98	50.87	50.89	0.70	0.50	2.33
E	52.84	52.81	52.29	0.61	0.29	3.90

In addition to the chilled water flow rate, the chilled water return temperature is a forcing input to the chiller rather than a performance dependent quantity because it is the water coming back from the building entering the chiller. The correlogram of the chilled water return temperature in Figure 5-11 did not exhibit a periodic nature like the chilled water flow rate or other two evaporator side temperatures did. Since the ARIMA (3,0,0) was selected on the basis of fitting the chilled water flow rate, there is no reason to believe it will adequately fit the chilled water return temperature. While the mean values of the segments based on the results of the ARIMA (3,0,0) model are similar to those of

the collected data and the moving average filtered data, the standard deviations from the ARIMA model fit are actually larger than those of the collected data. The ARIMA (3,0,0) model is not appropriate for the chilled water return temperature data. The correlogram suggests that a different time series might yield better results but no other models were fitted as part of this thesis.

Table 5-9: TCHWS Statistics for Chiller 1 in June by Technique and Segments

	Mean (F)			Standard Deviation (F)		
ID	**Collected**	**Filtered**	**ARIMA**	**Collected**	**Filtered**	**ARIMA**
A	45.72	45.67	45.65	1.50	0.17	1.43
B	45.50	45.55	45.57	1.04	0.85	1.86
C	45.09	45.05	45.01	0.97	0.73	1.51
D	44.98	44.98	45.01	0.64	0.58	1.34
E	45.07	44.49	45.17	1.21	0.24	2.39

The chilled water supply temperature is a function of the chiller performance and therefore a function of at least some of the chiller forcing inputs including the chilled water flow rate. The correlogram of the chilled water supply temperature in Figure 5-11 showed a periodic nature similar to that of the chilled water flow rate suggesting that the ARIMA (3,0,0) model might be a suitable time series model. However, after a comparison of the standard deviation values in the Table 5-9, it can be concluded that the ARIMA model is inadequate.

Table 5-10: TEVAP Statistics for Chiller 1 in June by Technique and Segments

	Mean (F)			Standard Deviation (F)		
ID	**Collected**	**Filtered**	**ARIMA**	**Collected**	**Filtered**	**ARIMA**
A	40.78	40.68	40.52	1.96	0.76	1.29
B	40.05	40.22	40.26	2.75	2.58	0.55
C	38.70	38.62	38.37	2.45	2.28	0.51
D	38.62	38.73	38.74	1.90	1.82	0.20
E	38.17	37.05	39.89	2.98	0.97	0.61

Similar to the chilled water supply temperature, the temperature of the refrigerant in the evaporator is a function of the chilled water flow rate and had a period correlogram in Figure 5-11. The standard deviations calculated from fitting the evaporator refrigerant temperature with an ARIMA (3,0,0) model presented in Table 5-10 show that the time series model was also able to quantify the time variance and better estimate the random component. The values of the standard deviations from the moving average filtered data are actually larger than those of the ARIMA (3,0,0) model for four of the five segments.

For two of the four evaporator side quantities examined, namely the chilled water flow rate and the refrigerant temperature in the evaporator, the ARIMA (3,0,0) time series model was able to better quantify the time variance of the data and thereby accurately estimate the standard deviation. It should be noted that while this model performed well for this data, the calculated standard deviation values are still only estimates. However, the time series standard deviation estimates are better than the 61 point moving filter results which were small and appealing but actually false as the moving average filter hid the time variation.

6. Conclusions and Recommendations

6.1. Conclusions

The chilled water flow rate had a strong periodic time dependency. Since the chilled water flow rate is one of the primary driving inputs to a chiller, other dependent chiller quantities are affected as noted by the chilled water supply temperature and the refrigerant temperature in the evaporator. Both the initial analysis of simply computing statistical quantities and a second approach of hourly averaging both overlooked the periodic variation. Only close examination of the chilled water flow rate at an expanded time scale revealed the periodic nature.

Time series analysis enabled the variance to be accurately quantified which reduced the estimates of the residual random component of the measured quantities. The representative error bars on calculated characteristic quantities such as the evaporator conductance area product (UA_{EVAP}) are based on the standard deviation estimates of the measured quantities that are used in the calculation. Consequently, reducing the standard deviation estimates shrinks the error bars, which yields higher accuracy and increased the sensitivity to detect changes in the values of the characteristic quantities. The data collected for the evaporator side quantities at 14:34 on June 28th (Segment A) serves as an example. The four measured quantities each are presented with their standard deviation from both the collected data and the ARIMA(3,0,0) model used as estimates of the measurement uncertainty. The chilled water flow rate standard deviation values are from Table 5-7, but since the time series models were not optimized for the temperature

quantities, a uniform standard deviation of a half of a degree Fahrenheit was assumed for both cases and all three temperature quantities. Using the evaporator relationships in section 2.2.1 of this thesis, the value of UAEVAP and the corresponding uncertainty can be calculated.

Table 6-1: Uncertainty Analysis of UAEVAP for a Data Point in Segment A

	GPMCHW (GPM)	**TCHWR (F)**	**TCHWS (F)**	**TEVAP (F)**	**UAEVAP (MMBtu/hr-F)**
Collected Data	4590±966	51.01±0.5	42.91±0.5	37.45±0.5	2.10±0.51
ARIMA (3,0,0)	4590±128	51.01±0.5	42.91±0.5	37.45±0.5	2.10±0.27

The uncertainty of the evaporator conductance area product has been almost reduced in half by utilizing the ARIMA (3,0,0) estimates of the standard deviation of the chilled water flow rate. If detecting a change in the conductance area product was the goal, the time series results allow a 13% change to be identified as significant where only change greater than 24% could be consider significantly different using only the collected data. Fitting adequate time series models to each of the evaporator temperatures would provide still better accuracy in the evaluation of the conductance area product, which is only one of eleven typical characteristic quantities used for fault detection and diagnosis of chiller performance. The uncertainty of all of the characteristic quantities could be reduced by applying time series model to all of the measured quantities that were time correlated.

The collected data also included changes in the nature of the time behavior. This sliding window ARIMA provided two important benefits relating to those transitions.

Plotting the chilled water flow rate versus time at the correct scaling made the changes obvious to the human eye, but time series analysis was able to computationally detect those changes as evident by the binary alarm. The alarm could potentially detect more subtle changes that are not evident to the human eyes, such as the transition in the middle of segment C. These transitions may have known or unknown causes that could possibly include faults.

One alternative to the sliding window approach could be a growing window approach where all of the data after a specified point are utilized to attempt to fit a model. The window "grows" because for each subsequent point the number of points used to fit the model increases by one. A growing window ARIMA approach might be able to detect the first behaviorally transition but not subsequent transitions. The sensitivity of the detection would also be compromised after the each transition, because a single model being built around data from one or more past regions and the current region is being asked to quantify the behavior of the current region. An additional important benefit of the sliding window approach was the ability to adapt to the change in behavior. When the time behavior changed, whether a result of an intentional control action or some unknown but acceptable occurrence, the sliding window ARIMA adapted to the new behavioral nature. A time lag is required until the sliding window no longer includes data prior to the transition but after that time period the ARIMA model is again capable of detecting changes in the time behavior.

6.2. Recommendations

The techniques presented in this thesis have the ability to improve chiller fault detection and diagnosis, but there are numerous optimization questions that remain regarding the full inclusion of time series analysis. The size of the window used in the sliding window approach should be studied further. The 60 minute window was used in this thesis for comparison to McIntosh's 60 minute moving average, but a more appropriate length may exist.

Statistical techniques including time series generally work better for more data making a larger window advantageous. However, given the four transitions that occurred in chiller 1 during the examined three days in June, additional transitions are probably quite prevalent. A future algorithm could combine the sliding and growing window approaches to improve statistical significance yet remain adaptable to new behavior patterns then the sliding window.

The sliding window ARIMA approach using one ARIMA model type was applied only on the evaporator side of the chiller and the suitable ARIMA (3,0,0) model was found for the chilled water flow rates. The most suitable model should be found for the three evaporator temperatures and the other monitored quantities of the chiller or even other coupled HVAC components such as the cooling towers or air handling units. The possibility of changing model types dynamically could be considered. Such a technique might be built upon calculation of the autocorrelation function and some amount of differencing and inputting those results to an algorithm that would identify the suitable model and then calculate the necessary coefficients.

Bibliography

Box, G. E., Jenkins, G. M., and G. C. Reinsel. 1994. *Time Series Analysis, Forecasting and Control.* 3rd Ed. Prentice Hall, New Jersey.

Breuker, M.S. and J.E. Braun. 1998a. Common Faults and Their Impacts for Rooftop Air Conditioners. *International Journal of Heating, Ventilating, Air-Conditioning, and Refrigerating Research* 4(3): 303-318.

Breuker, M.S. and J.E. Braun. 1998b. Evaluating the Performance of a Fault Detection and Diagnostic System for Vapor Compression Equipment. *International Journal of Heating, Ventilating, Air-Conditioning, and Refrigerating Research* 4(4): 401-425.

Chatfield, C. 1989. *The Analysis of Time Series: An Introduction.* 4th Ed. Chapman & Hall. London, England.

Fasolo, P.S. and D.E. Seborg. 1995. Monitoring and Fault Detection for an HVAC Control System. *International Journal of Heating, Ventilating, Air-Conditioning, and Refrigerating Research* 1(3): 177-193.

Isermann, R. 1984. Process Fault Detection Based on Modeling and Estimation Methods – A Survey. *Automatica* 20(4): 387-404.

Li, X., Visier, J. -C., and H. Vaexi-Nejad. 1997. A Neural Network Prototype for Fault Detection and Diagnosis of Heating Systems. *ASHRAE Transactions.* 103(1): 634-644.

McIntosh, I.B.D. 1999. A Model-Based Fault Detection and Diagnosis Methodology for HVAC Subsystems. *Ph.D. Thesis*, Madison, WI: University of Wisconsin-Madison.

Pan, S., Zheng, M., and N. Nakahara. 1999. Study on Fault Detection and Diagnosis of Water Thermal Storage Systems. *Proceedings of Sixth International Building Performance Simulation Association Conference (BS'99)* Vol. 1: 169-176. Kyoto, Japan.

Rossi, T.M. and J.E. Braun. 1996. Minimizing Operating Costs of Vapor Compression Equipment with Optimal Service Scheduling. *International Journal of Heating, Ventilating, Air-Conditioning, and Refrigerating Research* 2(1): 3-26.

Rossi, T.M. and J.E. Braun. 1997. A Statistical, Rule-Based Fault Detection and Diagnosis Method for Vapor Compression Air Conditioners. *International*

Journal of Heating, Ventilating, Air-Conditioning, and Refrigerating Research 3(1): 19-37.

Salisbury, T.I. 1999. A Controller for HVAC Systems with Fault Detection Capabilities Based on Simulation Models. *Proceedings of Sixth International Building Performance Simulation Association Conference (BS'99)* Vol. 1: 147-154. Kyoto, Japan.

Shiozaki, J. and F. Miyasaka. 1999. A Fault Diagnosis Tool for HVAC Systems Using Qualitative Reasoning Algorithm. *Proceedings of Sixth International Building Performance Simulation Association Conference (BS'99)* Vol. 1: 155-161. Kyoto, Japan.

Stylianou, M., and D. Nikanpour. 1996. Performance Monitoring, Fault Detection and Diagnosis of Reciprocating Chillers. *ASHRAE Transactions.*102 (1): 615-627.

Tsutsui, H. and K. Kamimura. 1996. Chiller Condition Monitoring Using Topological Case-Based Modeling. *ASHRAE Transactions.* 102(1): 641-648.

Yoshida, H. and S. Kumar. 1999. RARX Algorithm Based Model Development and Application to Real Time Data for On-Line Fault Detection in VAV AHU Units. *Proceedings of Sixth International Building Performance Simulation Association Conference (BS'99)* Vol. 1: 155-161. Kyoto, Japan.

Appendix A – EES Data Reduction Code

```
Refrigerant$ =          'R22'      "Refrigerant used"
DELTAT_superheat=0.01
DELTAT_subcool=0.01
T2_meas=T[2]

"!State 1 is entering the compressor"
T[1]      =    T_EVAP + DELTAT_superheat"[F]"
P[1]      =    PRESSURE(Refrigerant$,T=T_EVAP,x=1)"[psia]"
h[1]      =    ENTHALPY(Refrigerant$,T=T[1],P=P[1])"[Btu/lbm]"
s[1]      =    ENTROPY(Refrigerant$,T=T[1],P=P[1])"[Btu/lbm-R]"
P[1]      =    P_evap
v[1]      =    VOLUME(Refrigerant$,T=T[1],P=P[1])"[ft^3/lbm]"

"!State 2 is  leaving the compressor and entering the condenser"
P[2]          =    P[3]"[psia]"
s_isen[2]     =    s[1]"[Btu/lbm-R] "
h_isen[2]     =    ENTHALPY(Refrigerant$,P=P[2], s=s_isen[2])"[Btu/lbm]"
s[2]          =    ENTROPY(Refrigerant$,T=T[2],P=P[2])"[Btu/lbm-R]"
h[2]          =    ENTHALPY(Refrigerant$,P=P[2], T=T[2])"[Btu/lbm]"
P_cond        =    P[2]
v[2]          =    VOLUME(Refrigerant$,P=P[2], T=T[2])"[ft^3/lbm]"

"!State 3 is leaving the condenser "
T[3]          =    T_COND"[F]"
P[3]          =    PRESSURE(Refrigerant$,T=T_COND,x=0)"[psia]"
h[3]          =    ENTHALPY(Refrigerant$,T=T_COND,x=0)"[Btu/lbm]"
s[3]          =    ENTROPY(Refrigerant$,T=T_COND,x=0)"[Btu/lbm-R]"
v[3]          =    VOLUME(Refrigerant$,T=T_COND,x=0)"[ft^3/lbm]"

"!State 4 is entering the expansion valve"
T[4]          =    T_COND - DELTAT_subcool"[F]"
P[4]          =    PRESSURE(Refrigerant$,T=T_COND,x=0)"[psia]"
h[4]          =    ENTHALPY(Refrigerant$,T=T[4],P=P[4])"[Btu/lbm]"
s[4]          =    ENTROPY(Refrigerant$,T=T[4],P=P[4])"[Btu/lbm-R]"
v[4]          =    VOLUME(Refrigerant$,T=T[4],P=P[4])"[ft^3/lbm]"

"!State 5 is leaving the expansion valve and entering the evaporator"
T[5]          =    T_EVAP"[F]"
P[5]          =    PRESSURE(Refrigerant$,T=T[5],h=h[5])"[psia]"
h[5]          =    h[4]"[Btu/lbm]"
s[5]          =    ENTROPY(Refrigerant$,T=T[5],h=h[5])"[Btu/lbm-R]"
v[5]          =    VOLUME(Refrigerant$,T=T[5],h=h[5])"[ft^3/lbm]"

"!State 6 is leaving the evaporator"
T[6]      =    T_EVAP"[F]"
P[6]      =    PRESSURE(Refrigerant$,T=T[6],x=1)"[psia]"
h[6]      =    ENTHALPY(Refrigerant$,T=T[6],x=1)"[Btu/lbm]"
s[6]      =    ENTROPY(Refrigerant$,T=T[6],x=1)"[Btu/lbm-R]"
v[6]      =    VOLUME(Refrigerant$,T=T[6],x=1)"[ft^3/lbm]"
```

```
"!Evaporator Side Calculations"
T_bar_CHW       =    (T_CHWS + T_CHWR)/2"[F]"
P_CHW           =    1*convert(atm,psia)"[psia], pressure of chilled water"
rho_chw =DENSITY(Water,T=T_bar_chw,P=P_chw)*convert(lbm/ft^3, lbm/gal)"[lbm/gal]"
m_dot_chw=v_dot_chw*rho_chw*convert(lbm/min, lbm/hr)"[lbm/hr]"
Cp_CHW=SPECHEAT(Water,T=T_bar_CHW,P=P_CHW)"[Btu/lbm-F]"
DELTAT_CHW=T_CHWR - T_CHWS"[F]"
Q_dot_EVAP=m_dot_CHW*Cp_CHW*DELTAT_CHW*convert(Btu/hr,Tons)"[Tons]"
“calculates Q_dot_EVAP"
Q_dot_EVAP=m_dot_refrig*(h[6] - h[5])*convert(Btu/hr,Tons)"[Tons]"
“ calculates m_dot_refrig"
epsilon_EVAP=   (T_CHWR - T_CHWS)/(T_CHWR - T_EVAP)
epsilon_EVAP=   1  - exp(-Ntu_EVAP)
Ntu_EVAP=  UA_EVAP/(m_dot_chw*Cp_CHW)
APPR_EVAP=     T_CHWS - T_EVAP"[F], evaporator approach"

"!Condenser Side Calculations"
"This portion of code will calculate T_CWR_calc for comparison to the measured value"
T_bar_CW        =    (T_CWS + T_CWR_calc)/2      "[F]"
P_CW            =    1*convert(atm,psia)"[psia], pressure of condenser water"
rho_cw          =    DENSITY(Water,T=T_bar_cw,P=P_cw)*convert(lbm/ft^3,
lbm/gal)"[lbm/gal]"
m_dot_cw        =    v_dot_cw*rho_cw*convert(lbm/min, lbm/hr)          "[lbm/hr]"
Cp_CW           =    SPECHEAT(Water,T=T_bar_CW,P=P_CW)"[Btu/lbm-F]"
DELTAT_CW       =    T_CWR_calc - T_CWS"[F]"
Q_dot_COND      =    m_dot_refrig*(h[2] - h[3])*convert(Btu/hr,Tons)"[Tons], calculates
Q_dot_COND"
Q_dot_COND      =    m_dot_CW*Cp_CW*DELTAT_CW*convert(Btu/hr,Tons)"[Tons], calculates
T_CWR"
epsilon_COND    =    (T_CWR_calc - T_CWS)/(T_COND - T_CWS)

epsilon_COND    =    1  - exp(-Ntu_COND)
Ntu_COND        =    UA_COND/(m_dot_cw*Cp_CW)
APPR_COND       =    T_COND - T_CWR_calc"[F], condenser approach"

"!Efficiencies and Performance Quantities"
T_bar_refrig=(T[1] + T[3])/2"[F], average temperature of refrigerant"
P_bar_refrig=(P[1] + P[3])/2"[psia], average pressure of refrigerant"
rho_refrig=DENSITY(Refrigerant$,T=T_bar_refrig,P=P_bar_refrig)"[lbm/ft^3]”
“density of refrigerant"
v_dot_refrig=(m_dot_refrig/rho_refrig)*convert(ft^3/hr, gal/min)"[gal/min]”
“volumetric flow rate of refrigerant"
W_dot_actual=m_dot_refrig*(h[2] -h[1])*convert(Btu/hr,kW)"[kW], work input from compressor"
W_dot_tons=W_dot_actual*convert(kW, Tons)"[Tons], work input from compressor"
W_dot_isen=m_dot_refrig*(h_isen[2]-h[1])*convert(Btu/hr,kW)"[kW]”
“Isentropic work input from compressor"
DELTAQ_dot=(Q_dot_COND-Q_dot_evap)*convert(Tons,kw)"[kW], difference between evap
and cond heat transfer rates"
eta_isen=W_dot_isen/W_dot_actual"Isentropic Efficiency of Compressor"
COP =(Q_dot_EVAP*convert(Tons,kW))/Power"Coefficient of Performance of Cycle"
COP_comp= (Q_dot_EVAP*convert(Tons,kW))/W_dot_actual"COP of Compressor"
COP_ideal=(Q_dot_EVAP*convert(Tons,kW))/W_dot_isen"Ideal COP of Compressor"
eta_motor                =          W_dot_actual/Power"Motor Efficiency"
```

Appendix B – Sliding ARIMA Macro Code

```
MACRO
SLIDEARIMA DATA JUMPTO WIDTH P D Q LOLIMIT PREDICT HILIMIT ALARM co1 co2 co3 co4
co5 co6
# This macro applies an arima model to a data set in a sliding window fashion.
# It takes 6 inputs and provides 4 columns of ouput.
#
# The inputs:
# Data - the column or matrix of data on which to apply the macro
# Jumpto - an integer representing the first data point in the column to be used
# Width - an integer value of the number of points in the window
# P - an integer for the number of auto-regressive terms in the model
# D - an integer for the model differencing
# Q - an integer for the number of moving average terms in the model
#     Note that the model form stays the same but the coefficients will be
#     recalculated each iteration
#
# The outputs:
# LoLimit - the lower 95% confidence interval of the forecast
# HiLimit - the upper 95% confidence interval of the forecast
# Predict - the forecast value
# Alarm   - an on/off value to denote if data point is inside (0) or outside (1)
#           of 95% confidence intervals
#
# Defining constant variables
MCONSTANT JUMPTO WIDTH P D Q PNQ NUMBER I1 I2 I3 NITER1 NITER2 BEG NEXT
# Defining column variables
MCOLUMN  DATA SEGMENT ZSEGMENT ZPREDICT PREDICT ZLOLIMIT LOLIMIT ZHILIMIT
HILIMIT ALARM SEGMEAN RESI FITS COEF co1 co2 co3 co4 co5 co6
# Label the output columns
Name LOLIMIT = 'LoLimit' PREDICT = 'Predict' HILIMIT = 'HiLimit' ALARM = 'Alarm'
Name co1 = 'Coeff 1' co2 = 'Coeff 2' co3 = 'Coeff 3' co4 = 'Coeff 4' co5 = 'Coeff 5' co6 = 'Coeff 6'
# Sum to determine the number of coefficients in the model, maximum of 6 for full output
LET PNQ = P+Q
# Count the number of points in the data set
LET NUMBER = COUNT(DATA)
# Calculate the number of times to iterate based on the number of points,
# the width of the window and the first point used
LET NITER1 = NUMBER-WIDTH-JUMPTO+1
# Begin the overall loop
DO I1=1:NITER1
        DO I2=1:WIDTH
                # Read a windows worth of data points into a column
                LET SEGMENT(I2) = DATA(JUMPTO-2+I1+I2)
        ENDDO
        # Read the point just after the window
        LET NEXT = DATA(JUMPTO-1+WIDTH+I1)
        # Calculate the mean of the window data points
        Statistics SEGMENT;
        Mean SEGMEAN.
        # Subtract the mean from the data
```

```
        DO I2=1:WIDTH
                LET ZSEGMENT(I2) = SEGMENT(I2) - SEGMEAN
        ENDDO
        # Apply the ARIMA model of order (P,D,Q)
        ARIMA P D Q ZSEGMENT RESI FITS COEF;
          # Since we took out the mean, no need for a constant term
          NoConstant;
          # Use the Arima model to forecast the next 1 point
          # and the 95% confidence limits for it
          Forecast WIDTH 1 ZPREDICT ZLOLIMIT ZHILIMIT;
         # No plots shown
         #GSeries;
         #GACF;
         #GPACF;
         #GHistogram;
         #GNormalplot;
         #GFits;
         #GOrder;
          # Do not display any characteristics of the arima model
          Brief 0.
        # Add the mean to the predicted and limit values
        # and write them into the columns to be output
        # Place up to 6 coefficients into column for storage and output
        IF PNQ>0
                LET co1(JUMPTO-1+WIDTH+I1)=COEF(1)
        ENDIF
        IF PNQ>1
                LET co2(JUMPTO-1+WIDTH+I1)=COEF(2)
        ENDIF
        IF PNQ>2
                LET co3(JUMPTO-1+WIDTH+I1)=COEF(3)
        ENDIF
        IF PNQ>3
                LET co4(JUMPTO-1+WIDTH+I1)=COEF(4)
        ENDIF
        IF PNQ>4
                LET co5(JUMPTO-1+WIDTH+I1)=COEF(5)
        ENDIF
        IF PNQ>5
                LET co6(JUMPTO-1+WIDTH+I1)=COEF(6)
        ENDIF
        LET PREDICT(JUMPTO-1+WIDTH+I1) = ZPREDICT+SEGMEAN
        LET LOLIMIT(JUMPTO-1+WIDTH+I1) = ZLOLIMIT+SEGMEAN
        LET HILIMIT(JUMPTO-1+WIDTH+I1) = ZHILIMIT+SEGMEAN
        # Compare the actual data value to the limit values
        # and set alarm accordingly
        IF NEXT>LOLIMIT(JUMPTO-1+WIDTH+I1) AND NEXT<HILIMIT(JUMPTO-1+WIDTH+I1)
                LET ALARM(JUMPTO-1+WIDTH+I1)=0
        ELSE
                LET ALARM(JUMPTO-1+WIDTH+I1)=1
        ENDIF
ENDDO
ENDMACRO
```

Appendix C – ARIMA (3,0,0) GPMCHW Coefficients